2016-03-01

Pragmatic Quotation Use in Online Yelp Reviews and its Connection to Author Sentiment

Mary Elisabeth Wright
Brigham Young University

Pragmatic Quotation Use in Online Yelp Reviews

and Its Connection to Author Sentiment

Mary Elisabeth Wright

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

William G. Eggington, Chair
Wendy Baker-Smemoe
Jesse Egbert

Department of Linguistics and English Language

Brigham Young University

March 2016

ABSTRACT

Pragmatic Quotation Use in Online Yelp Reviews
and Its Connection to Author Sentiment

Mary Elisabeth Wright
Department of Linguistics and English Language, BYU
Master of Arts

Previous research has established that punctuation can be used to communicate nuances of meaning in online writing (McAndrew & De Jonge, 2011). Punctuation, considered a computer mediated communication (CMC) cue, expresses tone and emotion and disambiguates an author's intention (Vandergriff, 2013). Quotation marks as CMC cues can serve pragmatic functions and have been understudied. Some of these functions have been generally described (Predelli, 2003). However, no corpus study has specifically focused on the pragmatic uses of quotations in online text. Consumer reviews, a genre of online text, can directly impact business profits and influence customers' purchasing decisions (Floyd, Freling, Alhoqail, Cho & Freling, 2014). Businesses are investing in sentiment analysis to gauge their target market's opinions (Salehan & Kim, 2016). Sentiment analysis is the computerized appraisal of a text to determine whether its author is expressing a positive or negative opinion (Novak, Smailovic, Sluban & Mozetic, 2015). Sentiment analysis programs are still limited and could be improved in accuracy. Most programs rely on lexicons of words given a pre-determined polarity value (positive or negative) out of context (Novak et al., 2015). However, context is crucial to communication, and sentiment analysis programs could incorporate a better variety of contextual linguistic features to improve their accuracy. Quotations used for pragmatic communication is such a feature.

This study discovered seven pragmatic quotation uses in a 2014 Yelp review corpus: Collective Knowledge, Non-standard, Grammatical, Non-literal, Narrative, Idiolect, and Emphasis. An ANOVA and Tukey HSD test were performed, and the results were significant. Pragmatic category accounted for 15% of the variance in review star rating. The Collective Knowledge category and the Narrative and Non-literal categories were significantly different from each other. The Collective Knowledge category showed a correlation with positive sentiment, while the Narrative and Non-literal categories displayed a correlation with negative sentiment. These three categories are likely present in several types of online text, making them valuable for further sentiment analysis research. If these pragmatic patterns could be detected automatically, they could be used in sentiment algorithms to give a more accurate picture of author opinion.

Keywords: quotations, quotes, pragmatics, CMC, online reviews, Yelp, sentiment analysis, opinion mining, computer mediated communication

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

Introduction

In recent years, online consumer reviews (OCRs) have become increasingly important sources of information to shoppers purchasing a variety of products and services (Salehan & Kim, 2016). Would-be consumers can now read a review rating anything from clothes, restaurants, and vacation spots to dentists, doctors, and lawyers. Like never before, online reviews "facilitate [the] offering and sharing [of] information between service providers and consumers as well as between consumers themselves" (Park & Nicolau, 2014, p. 1). The sharing of information between consumers is not new, but online reviews have increased the scope of this sharing tremendously. According to marketing professor and consumer complaint behavior researcher, Mary C. Gilly, "The balance of power has really swung toward the consumer. . . In the past, unhappy customers . . . might tell their family and friends (about a problem). Now they're telling hundreds of thousands of people, potentially – or even millions" (Ferrell, 2013, p. 1).

Consumers know that companies will present their products and services in the best light possible on their websites. While company information on a product or service may be accurate, it may not tell the whole story. "Some studies suggest that customers show more interest toward user-generated product information on the Internet than the information vendors provided" (Salehan & Kim, 2016, p. 30). In addition, "A recent study shows that OCRs . . . are the second most-trusted source of product information after recommendations from family and friends" (Salehan & Kim, 2016, p. 30). Consumers want to know about the experiences that others have with a product or service. As the above study suggests, consumers trust their friends, family, and online reviewers above sources of information from businesses themselves. Online reviews are

1

quickly becoming a critical communication tool that has the ability to directly affect business profits.

In a 2012 online survey of 2,005 American consumers, researchers studied how reviews influenced shoppers' buying decisions. They found that "65% of potential consumers selected a brand that had not been in their original consideration set" (Floyd, Freling, Alhoqail, Cho & Freling, 2014, p. 1). This means that because these consumers read online reviews about a product, well over half chose a brand they had not even considered buying prior to reading the reviews. Further studies confirm the power of online reviews. "According to a recent report by market research firm Nielsen (2012), 70% of consumers indicate they trust online product reviews" (Floyd et al., 2014, p. 1). Consumers look to online reviews as a reliable source of information to assess the strengths and weaknesses of products and services. In addition, consumers' "utilization of online recommendations in decision-making appears to be on the rise" (Floyd et al., 2014, p. 1). Customers can easily access reviews on a company website or through other review sites. It is becoming more prevalent for customers to read these reviews before buying products. It is clear that online reviews are swaying consumer decisions.

Companies are mindful of this growing trend and have created their own platforms for customers to voice their opinions about a product or service. Now businesses "actively try to harness electronic word-of-mouth (eWOM) as a new marketing tool by inviting their consumers to post personal product evaluations on seller websites . . . " (Floyd et al., 2014, p. 1). In addition to their ability to post reviews on company or seller websites, consumers can go to third party sites, like Yelp or Angie's List, and post their opinions on a wide range of products and services. Yelp in particular is a highly popular site for online reviews and contains a wealth of opinions for potential consumers to peruse through (Ferrell, 2013). Yelp was founded in 2004 and has

more than 42 million reviews. Over 108 million consumers visit their website every month (Ferrell, 2013). Yelp is so large and popular that it competes directly with Google. Google has its own reviews on Android mobile devices, while "Yelp has a deal with rival Apple" (Ferrell, 2013, p. 1). Millions of Apple and web users have access to Yelp's collection of reviews and star ratings assessing a vast array of products and services.

Online review sites like Yelp are not only helpful for the information-seeking consumer, but they are also helpful for businesses themselves. Businesses have begun processing and analyzing online reviews to determine how customers feel about a particular product or about the company in general (Salehan & Kim, 2016). However, the sheer volume of online reviews makes it difficult for companies to quickly gain useful sentiment information from them in an efficient manner. Online review data is readily available, but what is the best way to collect the data, process it, and derive a meaningful analysis from it? To businesses "being able to recognize useful data which is already available on the Internet, as well as being able to acquire it in a format which can later be interpreted and processed, can prove extremely useful" (Chifu, Letia, Budisan & Chifu, 2015, p. 7). This is why sentiment analysis is so valuable to businesses.

"Sentiment analysis, or opinion mining, is the computational study of people's opinions, sentiments, emotions, and attitudes. It is one of the most active research areas in natural-language processing and is also extensively studied in data mining, web mining, and text mining" (Novak, Smailovic, Sluban & Mozetic, 2015, p. 2). Sentiment analysis is increasing in importance because of the expanding amount of social media text like Twitter, Facebook, blogs, reviews, etc. (Novak et al., 2015). Sentiment analysis is a complex process; however, there are different kinds of online programs that take a text and analyze its sentiment automatically. "The basis of many sentiment-analysis approaches is the sentiment lexicons, with the words and

phrases classified as conveying positive or negative sentiments. Several general-purpose lexicons of subjectivity and sentiment have been constructed" (Novak et al., 2015, p. 2). These programs rely on algorithms that incorporate online lexicons with words already assigned a sentiment value based on their meaning and the frequency of their co-occurrence with other sentiment-laden words in patterned syntactic structures (Novak et al., 2015). These programs, in addition to other sentiment calculations, are already being run on customer reviews and other types of social media (Salehan & Kim, 2016). Some sentiment analysis programs display a high accuracy when compared with human sentiment evaluations (Subhabrata, Akshat & Balamurali, 2012). Generally, sentiment analysis is a much more efficient way to gauge and interpret customer attitudes instead of human readers scouring each and every review. However, there is still much progress to be made.

While some of these sentiment programs and other calculations take into account part of speech, syntactic structure, lexicons, and even emoticons, they do not always take into account punctuation features like quotations, ellipses, parentheses, and so on. These features and others, like capitalization and unusual spelling, are used by online writers to communicate pragmatically (Subhabrata et al., 2012). Pragmatics is a "subfield of linguistics which studies how the transmission of meaning depends not only on the linguistic knowledge (e.g., grammar, lexicon, etc.) of the speaker and listener, but also on the context of the utterance, knowledge about the status of those involved, the inferred intent of the speaker etc." (Subhabrata et al., 2012, p. 2532). The use of pragmatics is found in online written communication, not just speech. A sentiment analysis study of tweets found that writers communicated pragmatically their happiness, joy, or excitement by elongating a word (e.g., gooooood). Writers also communicated the intensity of their feelings by using all caps (e.g., I HATED that) or by capitalizing a specific word (e.g., She

4

is a Loving mom.) (Subhabrata et al., 2012). Because many sentiment analysis programs rely on lexicons where words are assigned a pre-determined polarity value (positive, negative, and neutral) outside of any linguistic context, other communication strategies, like pragmatics, should be taken into account to improve the accuracy of the sentiment analysis.

A particularly understudied area, where punctuation features communicate pragmatic intent, is the use of quotations. Other studies have analyzed quotations in news reports and other genres to determine the sentiment of each quotation, but they have not delved into how the quotations were being used pragmatically (Balahur, Steinberger & van der Goot, 2009). Likewise, studies on sentiment analysis have used emoticons, capitalization, and so on to help improve the accuracy of the analysis (Desner & Herring, 2010). However, no studies have focused on quotations and their pragmatic uses in contributing to sentiment.

This can be a particular problem because quotations are used to express verbal irony in writing (Carvalho, Silva, Sarmento & Oliveira, 2009). Verbal irony is defined as "the rhetorical process of intentionally using words or expressions for uttering a meaning different (usually the opposite) from the one they have when used literally" (Carvalho et al., 2009, p. 1). The study by Carvalho, Silva, Sarmento, and Oliveira focused on sentences from a Portuguese newspaper and showed that 68.29% of the time quotations were used to express irony (2009). Sentiment analysis programs have difficulty spotting this type of pragmatic communication and can mistake negative attitudes for positive ones (Carvalho et al., 2009). In addition to expressing irony, quotations have several other pragmatic uses, like creating distance between the writer and the quoted text, drawing attention to the quoted word or phrase, expressing slang or jargon, etc. (Predelli, 2003). This study aims to conduct further research into the pragmatic uses of quotations and their relation to sentiment analysis.

The research questions this paper addresses are as follows: What pragmatic functions do quotations serve in online writing, specifically Yelp reviews? How do non-linguist, lay readers perceive the use of these quotations in Yelp reviews? Can the pragmatic functions of quotations in online reviews help predict reviewer sentiment? These questions will be addressed through a three-part study: 1) a close reading analysis of Yelp reviews containing quotations 2) a Mechanical Turk survey asking participants to categorize quotations from Yelp review excerpts 3) a one-way ANOVA comparing pragmatic categories to review star ratings.

Identifying the pragmatic uses of quotations will add to the sparse body of linguistic research in understanding how quotations are used in non-traditional ways (not just to report another's speech or writing) in online communication, specifically reviews. Surveying lay readers familiar with online communication will determine how readers interpret a reviewer's pragmatic use of quotations. Central to effective communication is a listener/reader's ability to decode pragmatic messages sent by a speaker/writer. The survey results will also help support or refute the researcher's pragmatic classifications found through her reading analysis. Lastly, the statistical analysis will either support or refute the idea that a quotation's pragmatic function can help predict its reviewer's sentiment. If there is a statistical relationship between quotation use and star rating, sentiment analysis programs can take this pattern into account while gauging the sentiment of social media text containing quotations. The following section reviews the existing literature about the potential for quotations' pragmatic functions and explains how they are an overlooked type of punctuation in studies about computer mediated communication.

Chapter 2

Review of Literature

As described in the Introduction, pragmatic quotation use could be an additional linguistic feature used to improve the accuracy of sentiment analysis programs. To do this, researchers need to understand how quotations are being used in online text to communicate pragmatically. This type of quotation use is different from an author using quotations in the conventional way of only reporting another's speech or writing. Scare quotes are one example of a common type of quotation that is not conventional but that is used pragmatically. Scare quotes are defined as "quotation marks used to foreground a particular word or phrase, esp. with the intention of disassociating the user from the expression or from some implied connotation it carries" (Scare Quotes, 2015). As seen later in this section, there are other quotation uses that can communicate an author's intention as well. Different types of quotations and their uses is an understudied area of linguistics with little previous research. As such, the following review of literature is derived from what is available on the subject and is not a comprehensive review of quotation research in general. This section begins by discussing quotations' place in the world of online text and then discusses different types of quotation uses more generally.

<u>**Quotations as CMC Cues**</u>

It has been acknowledged that computer mediated communication (CMC) eliminates "vocal and visual nonverbal cues that we depend on in face-to-face interactions to convey nuances of meaning and feelings" (McAndrew & De Jonge, 2011, p. 403). CMC includes social media text like Facebook posts, tweets, blogs, and online reviews. Previous studies have demonstrated that "expressive punctuation such as question marks and exclamation points appears to be effective as 'cyberlinguistic cues' that one can use to convey subtle information

about how a message should be interpreted" (McAndrew & De Jonge, 2011, p. 406). Other types of punctuation, aside from question marks and exclamation points, can also be expressive. Quotation marks are one such type of punctuation. From the researcher's analysis of a 2014 Yelp corpus of online reviews, discussed later in the paper, quotations can also serve as "cyberlinguistic cues" conveying pragmatic information to readers about the language in quotations.

Recent studies suggest that CMC cues are critical to online language use. They are described as "abundant, diverse, and an inseparable aspect of every CMC message" (Vandergriff, 2013, p. 2). CMC cues help communicate subtle information like emotional state (excitement, frustration, etc.) and tone (Vandergriff, 2013). In a face-to-face communication context, these subtleties might be communicated through facial expressions, body language, voice intonation, etc. However, these communicative options are of course not available in online written settings. This is why CMC cues are so important, especially in cases where language is not meant to be interpreted at face value.

These cues are especially important where non-literal language is used in online communication (Eisterhold, Attardo, & Boxer, 2006). One study in 2009 by Whalen, Pexman, and Gill researched the use of non-literal language in email communication between friends. Their aim was not to specifically study quotation marks, but one of their variables for marking non-literal language was punctuation. Quotation marks were included in their study as punctuation. They discovered that quotation marks were used to signal the following four types of non-literal language: sarcasm, jocularity, hyperbole, and rhetorical questions (Whalen, Pexman & Gill, 2009). So quotations can disambiguate an author's tone, and they can also express positive and negative emotion, or sentiment.

A study conducted in 2007 by Hancock, Landrigan, and Silver investigated strategies used to convey emotions in text messages. The authors found that "people self-reported that they used punctuation to express positive and negative emotion" (Hancock, Landrigan & Silver, 2007). They also found that the punctuation category was statistically significant in the expression of positive and negative emotion when compared to other strategy categories (Hancock et al., 2007).

These studies mentioned are all helpful. However, CMC cues, including punctuation, are usually all grouped together in studies that examine their uses. The CMC cue that has received the most individual attention is the emoticon (Vandergriff, 2013). Other forms of punctuation fall into a broad category and are all treated the same. Research explicitly examining the use of quotation marks as a CMC cue seems to be unavailable. The researcher has observed that many studies focus on punctuation like ellipses, question marks, exclamation marks, capitalization, parentheses, and so on but not on quotation marks.

Another deficiency in CMC cue studies is that most are solely quantitative (Vandergriff, 2013). Researchers may simply run the online text through the Linguistic Inquiry and Word Count program (LIWC). This is not bad, but the researchers may not actually examine the CMC cues in context. Examining CMC cues in context is a time-intensive process but is crucial in understanding how they are used for pragmatic communication. Of the studies found on CMC cues, many do not rely on corpora of naturally occurring language. Rather studies had participants write text that was analyzed for the experiment rather than rely on naturally occurring CMC (Hancock et al., 2007; McAndrew & De Jonge, 2011; Whalen et al., 2009). Fortunately, the proliferation of social media has provided a rich resource containing CMC cues that is currently available to be explored.

Pragmatic Functions of Quotations

As seen, there is limited research on quotation marks used specifically as a CMC cue, but there is some research on quotations used pragmatically in general. In their article, *Varieties of Quotation Revisited*, Cappelen and Lepore briefly discuss scare quotes and say the following, "Scare quotes are a central and pervasive device for effecting distance. It enables speakers to indicate that they distance themselves, for whatever reason or in whatever way, from certain aspects of certain linguistic practices" (2003, p. 56). It is useful to know that scare quotes function to promote distance between an author and the quoted language; however, Cappelen and Lepore's description is not precise or specific. Unfortunately, they really do not expand on it more in their article. They are vague about what the author of scare quotes wants to distance her or himself from or how using quotations effectively accomplishes this.

Cappelen and Lepore do go on to say that scare quotes are a "speech-act heuristic," meaning they are a non-semantic feature of an utterance that adds content (2003, p. 55). They discuss the distancing effect of scare quotes saying, "distancing is a component of the speech-act content of the utterance, but not its semantic content" (Cappelen & Lepore, 2003, p. 56). These authors argue that scare quotes are more relevant to the field of pragmatics than to the field of semantics. It is true that the presence of quotation marks can signal the presence of hidden meaning, adding tone or emotion to the actual language in the quotes. This supports the idea that quotations are used to communicate pragmatic information. But the authors do not go into detail about what other pragmatic functions quotations serve (Cappelen & Lepore, 2003).

Another author, Predelli, has written two interesting articles that reference quotation marks and their use. In his article, *'Subliminable' Messages, Scare Quotes, and the Use Hypothesis*, Predelli gives an example of a quotation but only discusses its intended purpose in

vague terms (2003). The example quotation is as follows: "life is 'the farce which everybody has to perform'" (Predelli, 2003, p. 153). To explain why there would be quotations around the select phrase, Predelli says, "Furthermore, the use of quotation marks flagging 'the farce which everybody has to perform' aims at calling the audience's attention to that expression, presumably in order to highlight some relevant feature of it" (2003, p. 153). Calling the audience's attention to quoted language may well be a function of quotation marks; however, what is the author's intent in calling attention to the quoted language? What does the author want to communicate? Predelli goes on to admit that "The identification of the feature in question appears to be a contextually determined matter" (2003, p. 153). This is why studying CMC cues, specifically quotation marks, in-depth and in context is so crucial to understanding them and their pragmatic patterns. Once these patterns can be identified, they can then possibly be linked to author sentiment in a meaningful way.

In another one of his articles, *Scare Quotes and their Relation to Other Semantic Issues*, Predelli gives more examples of quotations and expands on their purpose and the intent of the author who used them (2003). In this article, he offers interesting insight into some possible pragmatic functions of quotations. The different examples he gives were very helpful to the researcher of this study when trying to categorize quotations by pragmatic function, as seen later in this paper.

The first quotation example Predelli gives comes from the *Chicago Manual of Style*: "in offset printing 'proofs' of illustrations come from the darkroom, not the proof press" (2003, p. 2). In this case, "proofs" is not the correct technical term. The author knows this and puts "proofs" in quotations to highlight this. The *Chicago Manual of Style* advises that this is a proper way to use scare quotes. It states that it is better "to apply a standard technical term in a

nonstandard way than to invent a new term" (Predelli, 2003, p. 2). The style guide also acknowledges that when this is done, many times quotation marks are around the technical term (Predelli, 2003). So this is one use of quotations that is not simply conventional. Writers know when they are inexperienced in a subject and sometimes have to come up with an approximate way to describe something. They signal their awareness of this to the reader through quotations.

Predelli uses the following quote for his next example: "had it not been for Bryce, the 'copper's nark', Collins would have made his escape" (2003, p. 2). This example also comes from the *Chicago Manual of Style*. In this case, Predelli explains, "Expressions of slang or jargon, not unlike nonstandardly applied technical terms, are also suited to an apologetic use of quotation marks" (2003, p. 2). Here the author uses quotations because the phrase he or she is referencing is "unfitting for the purpose of formal, respectable writing" (Predelli, 2003, p. 2). The writer wants to openly acknowledge that the expression is "inappropriate for the setting," and so he or she uses quotation marks to accomplish this (Predelli, 2003). Predelli's depiction of this type of quote as "apologetic" is interesting. Signaling a deficiency in one's writing and openly acknowledging it does seem apologetic. In this situation, the writer is communicating this awareness with quotation marks.

Predelli gives a third quotation example that serves "the purpose of indicating that the expression they enclose is 'foreign to the normal vocabulary of the speaker' (Chicago, 173), and is borrowed from some other source's idiolect" (2003, p. 2). He cites the following example for this use of quotation marks: "national greed has disguised itself in mandates to govern 'inferior' races" (Predelli, 2003, p. 3). In this case, the author would not normally use this term to describe other races, so they put the term in quotes to indicate it is not a part of their own idiolect. This type of quotation use is similar to the one creating distance in that the author does not want to be

associated with the quoted language. In this case, association with the quoted language implies that the author thinks certain races are less valuable or in some way unequal to others (Predelli, 2003).

In his fourth and fifth examples, Predelli mentions that quotes can be used to indicate sarcasm. The fourth example states, "the 'debate' resulted in three cracked heads and two broken noses" (Predelli, 2003, p. 3). Here the author uses the term "debate" to characterize a violent interaction that seems closer to a fist fight than to an exchange of words between two parties. The author is using the term sarcastically (Predelli, 2003). The fifth example Predelli gives also demonstrates sarcasm, "this remarkable piece of 'art' consists of a large canvas covered with mud and old bus transfers" (2003, p. 3). Here the author does not really consider what they are viewing as worthy of the term "art." The author is using the quotation marks to express a sarcastic attitude (Predelli, 2003).

In his last three examples, Predelli claims that authors use quotations when their quoted expressions refer to "a general background that will be recognized by the reader" (2003, p. 3). He also says that quotation marks can be used when necessary relevant background information is specific rather than general (Predelli, 2003). His sixth example states, "myths of 'paradise lost' are common in folklore" (Predelli, 2003, p. 3). In this case, the author "merely intends to 'single out a word or phrase, not quoting it from a specific document . . . but referring it to a general background that will be recognized by the reader'" (Predelli, 2003, p. 3). So here the author is not trying to quote something specific, rather they want to point out their general reference using quotation marks. In this case, the author is assuming the reader shares a common literary knowledge or background and signals this by quoting the title of Milton's famous work.

Predelli's seventh and eighth examples demonstrate the use of quotation marks in the most conventional way. His seventh example states, "life is 'the farce which everybody has to perform'" (Predelli, 2003, p. 3). His eighth example reads, "life is 'what happens while you are making other plans'" (Predelli, 2003, p. 3). In these cases, "the author appeals to specific sources, respectively the French poet Arthur Rimbaud and the pop icon John Lennon, which she thinks her audience may (though need not) be able to identify" (Predelli, 2003, p. 4). In the last two examples, the author is quoting what another has said or written. This is the way quotation marks are commonly used. However, the author seems to not only be using quotes because they are reporting another's speech or writing. They have an additional purpose in using quotes. This purpose is similar to the example in the preceding paragraph where the author thinks the reader has a common understanding and will recognize these famous lines. Even conventional quotations, in the right context, may be used to communicate pragmatically.

These examples of quotation marks and the attached explanations for their purpose are valuable to research on pragmatic quotation use. Predelli's examples and analysis are a great foundation. However, his examples and explanations are limited, and they are mostly from one source, the *Chicago Manual of Style*. To explore what other uses for quotations exist, a more robust study with numerous samples is necessary. Once numerous samples of quotation use have been analyzed, it will be possible to see whether the uses described by Predelli are common and whether there are additional uses he did not describe. Furthermore, these patterns of quotation use can be compared to author sentiment to see what influence, if any, quotation use has on predicting an author's attitude, be it positive or negative.

From previous research, it is possible to construe quotations as CMC cues that communicate subtle but important pragmatic messages. These communicated messages can deal

14

with ironic tone, positive and negative emotion, and the marking of non-literal language. Other pragmatic uses for quotations have been identified as well: using quotes to draw attention; signaling sarcasm, creating distance, marking language outside of one's idiolect, etc. This study aims to explore an area of research that has been largely neglected: the pragmatic uses of quotations and their ability to help predict author sentiment. Building upon this foundation, this paper seeks to answer the following questions regarding quotations: 1) What are the pragmatic uses of quotations in online writing, specifically reviews? 2) How do lay readers perceive the use of these quotations? 3) Can quotation use help predict author sentiment? The means for answering these questions are found in the next section of this paper.

Chapter 3

Research Design

In early 2015, a pilot study was conducted that focused on analyzing quotations in Yelp reviews and exploring the connection of their pragmatic uses to star rating. Alterations were made as a result, and additional research was carried out in the fall of 2015.

<u>Pilot Study</u>

The pilot study was conducted to determine how quotations are used pragmatically in online reviews and to determine the relationship between quotation use and star rating. A corpus consisting of 1.6 million online Yelp reviews from 2014 was used to analyze quotation mark use. The corpus was obtained from Yelp at their website, http://www.yelp.com/dataset_challenge. The corpus contained reviews from four countries, 10 cities, and surrounding areas: Edinburgh, UK; Karlsruhe, Germany; Montreal, Canada; Waterloo, Canada; Pittsburgh, US; Charlotte, US; Urbana-Champaign, US; Phoenix, US; Las Vegas, US; and Madison, US. The Yelp dataset contained information like businesses reviewed, the review text, review star rating, etc. The star rating refers to a score (1-5) that customers give the business based on their experience. One is the lowest rating possible, while five is the highest rating possible. Each customer review has an associated star rating.

The Yelp dataset was in JSON format. A Python script was written to extract the review texts containing quotation marks and their associated star ratings from the dataset. Because of the script's limitations, 1.4 million of the 1.6 million reviews were processed. These extracted reviews were not representative of reviews with quotations in general because the script extracted only reviews with quotations right at the beginning of the text. For the current study, the script was modified to extract all reviews containing quotations regardless of their placement

within the text. The extracted reviews and star ratings were written out to a CSV file. The limited script extracted 2,132 reviews containing quotation marks.

The Yelp corpus was chosen for a few reasons: 1) It had an accessible format with all of the information needed for the study (star rating, the review text, etc.), so a Python script could extract the useful information without the researcher having to create her own corpus of reviews. 2) The corpus's large size of 1.6 million reviews ensured a sample containing a high number of reviews with quotations and quotations used a variety of ways. 3) Yelp is a widely used and popular review site that has a good representation of reviews for different types of businesses and reviews covering the entire star-rating spectrum.

Of the 2,132 reviews, one in every six was used for analysis. This was done in order to obtain a systematic sample from the reviews. However, among the one in six reviews analyzed, there were some that were omitted from the results. The omitted reviews were not analyzed for the following reasons: errors in writing the data to the CSV file, reviews written in a foreign language, missing text, and reviews where half a set of quotation marks was missing. For example, sometimes a reviewer began their review with quotations but never added the closing quotations to the text. There were 335 reviews analyzed in total.

A careful reading analysis was performed on the review texts to determine in what pragmatic ways the reviewer was using the quotations. The resulting quotation use categories from the reading analysis were not predetermined. The star ratings were not referenced when there was ambiguity in how the quotation marks were used in the text. For example, if the researcher was unsure whether a quoted word or phrase expressed sarcasm, she did not check the star rating to see if that could help inform her decision, as sarcasm might expectedly be associated with a lower star rating.

17

Each review was read individually. The quoted word or phrase was considered on its own and considered within the context of the entire review. A rubric containing the pragmatic category name, definition, and examples was created as the analysis was conducted. The analysis was conducted in Excel, where the researcher would mark a "1" under a category heading if the review text contained quotation use matching that category. More than one category could be marked for each review if the review contained multiple quotations. However, each quotation was only assigned to one category. For example, if a review contained two quotations, one use could fit under the sarcasm category, while the other use could fit under the narrative category. However, if a review contained one quotation, that quotation could not fit under both the sarcasm and narrative categories. One category would have to be chosen. The researcher would ask herself what the author's reason was for using the quotations and use her best judgment when categorizing quotations with ambiguous pragmatic intention. The frequency for each categorical use of quotation marks was calculated in the above manner.

While the reality is that a reviewer could use a quotation for more than one pragmatic purpose, assigning one quote per category made a linear regression analysis of the data possible, which was important for determining whether a relationship between category and sentiment existed. Also, only assigning one category per quotation helped the researcher clarify pragmatic patterns during the reading analysis and focus on what the author was trying to communicate with the quote.

Once the frequencies for all of the categories were determined, the category results were "dummy coded" for a linear regression analysis in SPSS. "Dummy coding" in this context refers to placing ones or zeros in Excel under the columns of categorical variables to indicate their presence or absence. The star ratings (1-5) were used as the dependent variable, while the

categories were used as the independent variable. There were some data limitations that affected the results. The distribution of the residuals was skewed. However, the data was linear and the residuals mainly homoscedastic. The regression analysis was carried out with the Entire Review in Quotes category as the reference variable.

The discourse analysis of the 335 reviews yielded nine categories describing pragmatic quotation use: Expressive/Descriptive, Proper Noun/Special Term, Narrative, Quoting Another Reviewer, Sarcasm, Quoting Outside Source, Entire Review in Quotes, Company Slogan/Claim, and Idiomatic/Known Expressions. These categories are discussed in more detail in the Findings and Discussion section. The results of the linear regression were statistically significant ($F_{(8, 334)} = 5.161$, $p < .005$) and accounted for about 11% of the variance in the star rating. The independent variables of Expressive/Descriptive, Narrative, and Sarcasm were all significant and were all associated with a decreased star rating. For every additional instance of Expressive/Descriptive use, the star rating decreased by .564 points when compared to an instance of Entire Review in Quotes use. For every additional instance of Narrative use, the star rating decreased by .358 points when compared to an instance of Entire Review in Quotes use. For every additional instance of Sarcasm use, the star rating decreased by 1.045 points when compared to an instance of Entire Review in Quotes use.

<u>Current Research</u>

The Python script previously mentioned was modified for the current study (see Appendix) to search through all 1.6 million reviews and to extract all reviews containing quotations, not just reviews with quotations at the beginning. 120,169 reviews were found containing quotations. Those reviews were narrowed down to only reviews from the following cities: Urbana-Champaign, Gilbert, Las Vegas, Madison, Mesa, Peoria, Phoenix, Pittsburgh, San

19

Diego, and Scottsdale. The initial results included reviews from over 200 cities and towns. The researcher chose reviews from cities she knew were American. This was done because the researcher was primarily interested in quotation use in American online writing. Since some of the 1.6 million reviews were from different countries, the author wrote the script to extract the city information as well as star rating, review text, etc. Non-native speakers of English and those from other countries may use quotations for different pragmatic purposes in online review writing. Because this may be true, the author did not want to add variables like country of origin or native speaker versus non-native speaker to her research. The above cities are somewhat concentrated in the Western/Southwestern US. However, the author tried to include available cities from other regions as well, like the Midwest, West coast, and the East.

Of the remaining reviews from only American cities, the researcher took one out of every six reviews for analysis. Some of the reviews were omitted because of errors. Two kinds of errors occurred: 1) only an opening set of quotation marks was found with no closing set 2) contractions were mistaken for quotation marks. In total the author read 425 reviews, and 411 were error-free and analyzed. In total 621 quotations were analyzed.

A careful reading analysis was performed on the review texts to determine in what pragmatic ways the reviewer was using the quotations. The resulting quotation use categories from the reading analysis were not predetermined. The star ratings were not referenced when there was ambiguity in how the quotation marks were used in the text. For example, if the researcher was unsure whether a quoted word or phrase was used narratively, she did not check the star rating to see if that could help inform her decision.

Pragmatic Quotation Use

During her analysis, the researcher created a rubric with category names, descriptions, and examples (see Appendix). The quotations were considered on their own and within the broader text of the entire review. Each quotation was assigned to only one category. This was done for the same reason as the pilot study, so the researcher could perform a statistical analysis comparing pragmatic category and star rating. It also helped the researcher focus on the perceived communication intention of the reviewer and organize the quotation uses into a finite and reasonable number of categories.

Sometimes pragmatic intent was difficult to determine and took careful analysis. Some quotations in form seemed straightforward, while in function were more complex. For example, in the following quotation, the word "hot" could be quoted for a number of reasons: "For bar food it was pretty good. The wings aren't too good since I had the 'hot' and it was more like mild to me (yet I love hot food)." The reviewer could be quoting the menu, in that he or she got "hot wings." The reviewer could be trying to emphasize the word "hot." The reviewer could be using quotes to communicate what he or she assumes is shared knowledge (that restaurants offer different levels of spicy chicken wings from mild to hot). In situations where a quote fit into multiple pragmatic categories, the researcher looked at similar quoted language in other reviews and relied on any helpful contextual information in the rest of the review. She continually asked herself what she thought the reviewer's main reason was for using quotations, and she used her best judgment. The categorization process was subjective, so the researcher relied on another rater to analyze many of the reviews, and she had lay readers categorize quotations as well.

After conducting the reading analysis, the researcher gave the first 50 reviews analyzed and the category rubric to a second analyst to determine inter-rater reliability. The other rater was

compensated for her analysis. The researcher and the other rater disagreed on the categorization of three quotations in three different reviews. The inter-rater reliability rate was calculated at 94% agreement.

Lay Reader Perceptions of Quotations

A pilot Qualtrics survey was created with five example reviews, a list of the pragmatic categories, and basic demographic questions (see Appendix). The participants' task was to read each of the reviews and to assign a specific quotation to a pragmatic category by choosing one of the provided categories or by creating one of their own. The categories discovered by the researcher along with simplified definitions and examples were given to the participants. The survey was given to an introductory linguistics class taken as a general requirement, where students completed the survey in exchange for extra credit. 46 participants took the survey. Based on the results and student feedback, the researcher modified some of the survey format, category names, definitions, and included more quotation examples to make the survey clearer and more user-friendly.

The researcher then created three versions of the modified survey with five different review excerpts in each survey for a total of 15 unique reviews (see Appendix). The three reviews the researcher and rater disagreed on were included in the first survey. The other reviews were chosen so there were about equal numbers of reviews representing the different pragmatic categories. Each category was represented one to two times total, with the non-literal category as the exception, represented three times. The researcher also chose reviews where the entire text could be included (not too long) or where an excerpt with a reasonable amount of context could be included. Lastly, some reviews were chosen because of the pragmatically

ambiguous nature of the quoted language. In these instances, the researcher was curious to know how a lay population would categorize these specific quotations.

Demographic questions, pragmatic categories, definitions, examples, and a question connecting category to star rating were included in these surveys as well. The surveys also gave the participants the option to choose one of the provided categories for the quotation or to create their own category with a name and description. As a part of the demographic questions, the participants were asked how often they read or wrote online reviews to gauge their familiarity with online review language. They were also asked if they were a native speaker of English and if they were primarily raised in the US. This was done to exclude the unnecessary variables mentioned earlier: native versus non-native speaker influence and non-US influence over pragmatic perceptions.

The surveys were posted on Amazon's Mechanical Turk site. Mechanical Turk was chosen because it is a surveying method used in previous research. Mechanical Turk workers are representative of a non-linguist lay population that is familiar with online language use. With limited research funding, it was important to use a survey platform where the researcher could choose how much to pay each participant. Each participant was paid a $1.00 for completing each survey. Each participant was able to complete one, two, or all the surveys. 97 participants completed version one of the survey; 87 participants completed version two; and 92 participants completed version three.

Because of the subjective nature of the questions, the only way to check response quality was through the participants' time spent on the survey. The average Mechanical Turk worker took nine minutes to complete the survey. The work of those who took under six minutes was rejected. Six minutes was too short of a time for a worker to read every instruction, description,

review, category, etc. in such an information-laden survey with material unfamiliar to them, at an unrushed pace. Rushing through the survey would compromise the quality of the answers to many of the questions. The researcher felt that three minutes below the average time accounted for those who read unusually quickly, but that accepting work from those who took a shorter time than that, would be too permissive. However, a few workers emailed the researcher saying they had taken more than one of the surveys. In this case, the worker's rejection was overturned since the worker's familiarity with the format and content of the survey would have reduced his or her time. Responses were also rejected from participants who were non-native English speakers or not primarily raised in the United States. These responses were rejected because a participant's native language or cultural upbringing could affect how they interpret quotations in online reviews. This research concerns only American perceptions of quotation use.

Statistical Analysis

The researcher reflected on the statistics used in the pilot study, and after careful consideration, she decided to use a different statistical test in this study. After the reading analysis was conducted, that results were formatted in Excel to perform a one-way ANOVA in SPSS with star rating as the dependent variable and pragmatic category as the independent variable. A one-way ANOVA was used because there was only one independent variable. A post-hoc Tukey HSD test was also performed.

An ANOVA was used instead of a linear regression because while categorical independent variables can be "dummy coded" to perform a linear regression, an ANOVA is meant to calculate the effect of a categorical variable on a continuous variable. In this study, the independent variable of pragmatic use was categorical, while the dependent variable of star rating was continuous, which fits the assumptions for an ANOVA. The descriptive statistics,

24

ANOVA results, and post-hoc Tukey HSD test are discussed in the Results and Discussion section of this paper.

Results and Discussion

Pilot Study Categories

As stated earlier, there were nine pragmatic categories from the pilot study:

Expressive/Descriptive, Proper Noun/Special Term, Narrative, Quoting Another Reviewer,

Sarcasm, Quoting Outside Source, Entire Review in Quotes, Company Slogan/Claim, and

Idiomatic/Known Expressions.

Expressive/Descriptive. The Expressive/Descriptive category is defined as the

following: Where the reviewer uses the word/words in quotations to express how they feel

about the business or to describe their experience. Some examples include the following:

1) "Meh" is kind of how I felt about this place . . .
2) "WooHoo! As good as it gets!"
3) "Love this Freemont Street downtown hotel."

Proper Noun/Special Term. The Proper Noun/Special Term category is defined as the

following: Where the reviewer puts quotes around a proper noun, a nickname, specialized terms,

jargon, or titles. Some examples include the following:

1) "Beer Cave"
2) "The Duck" as I like to call it never disappoints
3) It is not uncommon to see an executive of one of the banks in town playing the role of
    'shark' (referring to the game of pool).

The last use of quotations is similar to Predelli's second quotation example where "Expressions

of slang or jargon, not unlike nonstandardly applied technical terms, are also suited to an

apologetic use of quotation marks" (2003). However in this category, the terms need not just be

slang or jargon, they are sometimes simply proper nouns.

Narrative. The Narrative category is defined as the following: Where the reviewer

quotes another person or people involved in their experience with the business in order to

transform their experience into a narrative for the reader. This can include the reviewer quoting him or herself or thoughts that he or she had. Some examples include the following:

1) "You're in the hot seat tonight. That's where Christian Bale sat on Thursday!" said our sweet waitress to my husband;
2) "Not your typical sushi restaurant." [First impression].

This category contains instances of conventional quotation use, where someone is quoting another or him or herself. The main purpose for doing this seems to be that the reviewer wants to tell the reader a story about their experience.

Quoting Another Reviewer. The Quoting Another Reviewer category is defined as the following: Where the current reviewer quotes a previous review about the business he or she is reviewing. This can be done whether the current reviewer agrees with the previous reviewer's opinion or not. Some examples include the following:

1) "Not quite what I expected . . . not quite WHERE I expected either . . ." 4 people found this review helpful I have to agree with another reviewer . . .
2) "Expect to be treated like a sub human by the doctors." I am happy to have to amend this review.

This is similar to the Narrative category in that one person is quoting another. However, it seems that for this category, the quotes are used because the current reviewer either wants to support his or her opinion or because the current reviewer disagrees with already established opinions.

Sarcasm. The Sarcasm category was defined more broadly and encompasses concepts that some may not strictly consider sarcastic. The definition is as follows: Where the reviewer puts a specific word or words in quotations to indicate that the quoted item(s) means the opposite of what it would normally refer to. This implied "oppositeness" can also extend to an entire phrase, sentence, etc., even if the phrase or sentence is not all in quotations. This definition also includes times when the reviewer uses quotations around a word or words that have a non-literal

description or express a non-literal opinion. This is often done for comic effect. Some examples include the following:

1) The front desk clerk seemed "oh-so enthused" to be working that night . . .
2) Overall I probably wouldn't come back here though as it really didn't seem like anything special or maybe I just don't "get" it
3) They run the gamut from "I need a surgical mask and gloves to shop here" to "I think I need to put on something nice to buy toilet paper."

The first example of these is quite similar to the examples Predelli cites that include sarcasm. The last example stated here is curious. The author is definitely doing it for comic effect. However, it is not sarcastic in a strict sense perhaps, but it does fit within the broader definition that the researcher has outlined.

Quoting Outside Source. The Quoting Outside Source category is defined as follows: Where the reviewer quotes a famous saying or some other statement, relevant to the review, that enhances, supports, or conflicts with their opinion or narrative. In this case, the reviewer does not quote a person he or she interacted with or overheard in his or her experience with the business. This is why it is called an outside source. Some examples include the following:

1) "I guess nothing is ever as good as you remember it" – Marge Simpson
2) "The single best Thai restaurant in North America"? (quoting a published food critic's review).

These examples are in some way like Predelli's examples where the author quotes a well-known saying or phrase. In some of these cases, the reviewer expects the reader to know of the quoted statement, and in other cases, this is less clear.

Entire Review in Quotes. The Entire Review in Quotes category is defined as follows: Where the reviewer puts his or her entire review within quotation marks. For short reviews, the reviewer could possibly just be quoting him or herself. For longer reviews, it is unclear why the

28

reviewer is doing this. If the reviewer is in fact quoting another person, another review, or an outside source, he or she has not made it known. Some examples include the following:

1) "Farren's has the best burgers in town"
2) "Welcome to Costco I love you."

This type of quotation use may be quite specific to Yelp reviews. The reason the reviewer might do this is inconclusive.

Company Slogan/Claim. The Company Slogan/Claim category is defined as follows: Where the reviewer quotes the actual slogan, the name of the company, or something the company claims and then usually evaluates it, agreeing or disagreeing with it. Even though the company name would be a proper noun or special term, instances where the reviewer used the name to evaluate the company were put into this category. Some examples include the following:

1) "Worlds friendliest airport" they have free wifi and the food selection in the airport is good!
2) "Don't let all the cars in the parking lot fool you." Boasts the menu. No seriously. Don't.

This type of quotation use would be specific to Yelp or other kinds of business reviews. However, it may also be found on social media sites or other forums where people are casually discussing a business and not giving a formal review.

Idiomatic/Known Expressions. The Idiomatic/Known Expressions category is defined as follows: Where the reviewer puts quotes around a word or words that are a common idiom or generally known to the reader. The quotes could be around items that the audience is expected to know about or understand. Some examples include the following:

1) "Back in the day"
2) "You get what you pay for"
3) "over-cooking" seems to be the overriding theme for this evening.

This is similar to Predelli's 'paradise lost' example. A shared general background between the reviewer and reader is perhaps assumed. However, this also goes along with one of his other examples where slang terms are enclosed in quotations.

Current Study Categories

The reading analysis resulted in the discovery of seven pragmatic categories: Collective Knowledge, Narrative, Grammatical, Non-literal, Non-standard, Idiolect, and Emphasis. While analyzing the quotations from this study, the researcher was more mindful of grouping the quotations into categories that may be found in other social media writing and not just reviews. Some of the categories from the pilot study were very specific to the review genre and were too narrow. The researcher has named, defined, and provided examples for the seven pragmatic categories below.

Collective Knowledge. Collective Knowledge is defined as the following: Where the reviewer quotes something because they expect the audience knows what they are referring to based on an assumed common experience, culture, or background. The researcher identified three main patterns within this category which she termed the following: slang/stock phrases, connotation, and shared experience. The slang/stock phrase subcategory is when the reviewer puts quotes around a word or phrase that is a neologism, a slang term, an idiomatic expression, or a commonly used phrase in English. In the connotation subcategory, the reviewer quotes a word or phrase to make the audience aware of an added meaning or connotation. However, the reader is supposed to know what the connotation means or implies based on this collective knowledge. Lastly, the shared experience subcategory is when the reviewer expects the readers to have had shared a similar experience. The following are excerpts from Yelp reviews exemplifying the four subcategories of Collective Knowledge:

1) The inside is very "old school" and there seems to be a deli attached although I didn't do any ordering from that side of the business. *(slang)*
2) He is "one of a kind" employee. *(stock phrase)*
3) Ordering "hot italian sausage" from a deli guy that calls you "babe" is as amusing as Sunday morning can get. :-)   *(connotation)*
4) So if you are a frantic "gotta get it ready & it's gotta be perfect" last-minute-crafty-maven like myself THAT'S A GREAT THING. *(shared experience)*

Narrative. The Narrative category is defined as the following: Where the reviewer quotes another person, an outside text, their own thoughts, or something they said. The quote is relevant to the reviewer's experience at the business they reviewed. It seems reviewers do this for at least two reasons. The central reason is to make their review into a story for the reader. Reviewers like to narrate their experiences interacting with others at a business or interacting with their friends. Another reason is to support their opinion or to justify the rating they gave the business. Several of the quotations in the narrative category detail negative interactions with employees or others that created an unpleasant experience for the reviewer when dealing with the business they were reviewing. The following are excerpts from Yelp reviews exemplifying the Narrative category:

1) She hung up the telephone and grabbed some paper towel to start cleaning a chicken oven. I asked if we were invisable. She said.. "oh I am so sorry" and then proceeded to clean the window of the oven!!! Really? I will never again shop at this Frys or any other Frys.... you can dress it up and make the store look nice.... but you need to hire employees that understand that a customer is the number one priority
2) There's a metal claw over stuffed animals where you put your quarter in and the claw doesn't even move. It's broken if not enough to have not taken your money. And you're staring at this metal claw thinking "Ew I'm underneath a casino. What smells like swamp water?" I like Treasure Island.

Grammatical. The Grammatical category is defined as the following: The reviewer uses quotes to mark a proper noun, to refer to a specific word or letter, or to mark an aside or appositive. This is the most straightforward category. It is unclear why the reviewer uses quotes to mark proper nouns when capitalization is usually sufficient. The reviewer may do this to

ensure there is no ambiguity in what they are referring to. The marking of a specific letter or word makes sense. This same practice is seen in other kinds of writing. The marking of an aside or an appositive is unusual since the author could use a dash, parentheses, or commas to achieve the same effect. However, this may again be for clarity's sake. The following are excerpts from Yelp reviews exemplifying the Grammatical category:

1) I read the book titled "1000 Places to See or Do in the US." *(proper noun)*
2) Lulu "our yorkie" was extremely comfortable with the Doc and staff *(appositive)*
3) Don't misunderstand "divey" I mean no attitudes no frills and tons of fun. *(specific word)*

Non-literal. The Non-literal category is defined as the following: Where the reviewer uses quotes to signify that something should not be taken literally, is counterfactual, metaphorical, or possibly has another meaning. This can happen when the author intends to communicate sarcasm, irony, understatement, etc. Reviewers use quotes to draw attention to non-literal language because tone is difficult to communicate in writing, and the reviewer has limited means to signal that some parts of their writing were figurative, ironic, etc. Quotations are such a means to pragmatically communicate the reviewer's intent and tone to readers. The researcher identified four main patterns within this category that she termed the following: sarcasm, metaphor, contradiction, and understatement. The sarcasm subcategory is when the reviewer uses quotes to signal a negative ironic tone. This was the most prevalent subcategory with 76.47% of Non-literal quotations using sarcasm. The metaphor subcategory is when the reviewer puts quotes around a concrete word or phrase used in a metaphorical way. The contradiction subcategory is when the reviewer puts quotes around a word or phrase but means the opposite; although, the opposite meaning is not necessarily negative. It is unlike sarcasm in this way. The understatement subcategory is when the reviewer puts quotes around a word or

phrase whose meaning does not match the level of the reviewer's true sentiment. The following are excerpts from Yelp reviews exemplifying the Non-literal category:

1) I love Dr. Shealey she is awesome. However the "referral coordinator" Dawn and "aide" Mary are completely rude and condescending. They do NOT want to help they just want to get you off the phone. *(sarcasm)*
2) I know that not everyone will understand why this HURTS my heart but in **"2013"** for someone to tell me that "Black people" are not worth a space on a shelf speaks volumes about why we as a country can't move forward! *(metaphor)*
3) The "meat" and "seafood" dishes could really make you worry that you are eating meat. The stuffed tomatoes are delicious as well as the sweet and sour pork I also really liked the lemon grass fish. I took a friend who was a Janeist (VERY strict diet!) and he wouldn't believe that it wasn't meat lol. *(contradiction)*
4) This is one of the only spots I bring my car because they do the best job by far! I had an issue with a major "spill" let's call it (vomit) in my back seat. *(understatement)*

Non-standard. The Non-standard category is defined as the following: Where the reviewer quotes a word or phrase because they have used it in an unusual, less common, or atypical way. The reviewer knows their usage is abnormal, so they use quotes to communicate their awareness of this to the reader. Reviewers also use quotes around words or phrases for non-standard descriptions that they use as approximations to convey meaning. The following are excerpts from Yelp reviews exemplifying the Non-standard category:

1) The portions were unbelievably small (we ordered 4 sushi rolls for 2 people and were definitely still hungry - I mean not even a tad bit full). I figured that could somewhat be the case but I seriously think these were the "skinniest" rolls (not sure how to describe it!) I've ever seen. *(approximation)*
2) It's the first time I've "Yelped" anything other than a restaurant. However after spending the better part of a day enjoying the Art/Natural History Museums I'm branching out...guess I'm way overdue! *(Yelp, the name of the review site, is starting to be used as a verb, but this usage would be unusual for someone unfamiliar with the site, hence the quotes)*

Idiolect. The Idiolect category is defined as the following: Where the reviewer quotes a word or phrase they do not normally use in their idiolect, or they quote a word or phrase that is exclusive to their idiolect or group's idiolect. In the first case, when the quoted language is foreign to the reviewer, they may quote it to distance themselves from it as described in the

Review of Literature. In the second instance, the reviewer realizes the quoted language is unique to their speech and writing and foreign to the reader, so they draw attention to it to communicate this awareness to the reader. Quoted language specific to the reviewer could be, as described, something that only they say or write, or it could be something said or written specifically by a particular social group they belong to. The following are excerpts from Yelp reviews exemplifying the Idiolect category:

1) When our name was called the hostess (speaking a mile a minute) said we had to sit at a "two-fer" or outside. Mind you it was about 90 degrees. I had to ask several times what a "two-fer" was before she finally responded and said a table for two. *(foreign to their idiolect)*
2) This meant that I had to go out "en homme" the first time I had ever done so with the Yelp crowd. I figured at first it wouldn't bother me. I wasn't thrilled by it but I thought I would handle it well. But I was wrong. *(This is unique to their idiolect/group's idiolect. This term is used by the transgender and cross-dressing community to describe presenting a masculine appearance and persona in public)*

Emphasis. The Emphasis category is defined as the following: Where the reviewer wants to stress the quoted word or phrase and make it stand out. Sometimes this is done to make whatever sentiment expressed more forceful. This is a difficult category. There is the temptation to read any review and simply state that the reviewer used quotes to draw attention to the specific language quoted language. This is a hypothesis, but it is too simplistic and does not expand on the reviewer's intent behind emphasizing particular language. In the Yelp reviews analyzed by the researcher, she found that reviewers seemed to use quotes as an intensifier for the word or phrase they wanted to emphasize. The following are excerpts from Yelp reviews exemplifying the Emphasis category:

1) Fantastic dinner spot in the heart of Scottsdale! The service was impeccable and the choices were very interesting. The "famous" chopped salad was a must get.
2) She told us we couldn't do boring colors because it wasn't "VEGAS" enough. Um last time I checked WE were paying YOU for the SERVICE. Not your decision.
3) There are other items on the menu that may be better but since this is "Pizza" Picazzo I'd expect better. Rosati's is a much better option in the area.

The Grammatical category was very similar to the Proper Noun/Special Term category in the pilot study. The Narrative categories were the same. The Sarcasm category was broadened to be the Non-literal category. Quoting an Outside Source was similar to the Collective Knowledge category. The Idiomatic/Known Expressions category was broadened to Collective Knowledge (slang) subcategory. The seven categories from the current study also show similarities to some of Predelli's quotation examples. Predelli's second example about the "copper's nark" is similar to the Collective Knowledge (slang) category. Predelli's third example about "inferior" races is the same as the Idiolect category where the writer quotes something foreign to his or her vocabulary. His fourth and fifth examples about "debate" and "art" are the same as the Non-literal (sarcasm) category. His last three examples referencing known sayings, expressions, or well-known literature are similar to the Collective Knowledge category, especially the shared experience subcategory.

Category Frequency

Some pragmatic quotation uses were more common than others. Reviewers used quotations often to signal the expectation of Collective Knowledge and shared experience, with a "you know what I mean" feel or intention. They also used quotations to narrate their business experiences or misadventures frequently. The least frequent uses of quotations were Idiolect and Non-standard. Below is a graph displaying the frequency of each pragmatic category for the 621 analyzed quotations.

Figure 1. Count of Pragmatic Categories from Analyzed Reviews
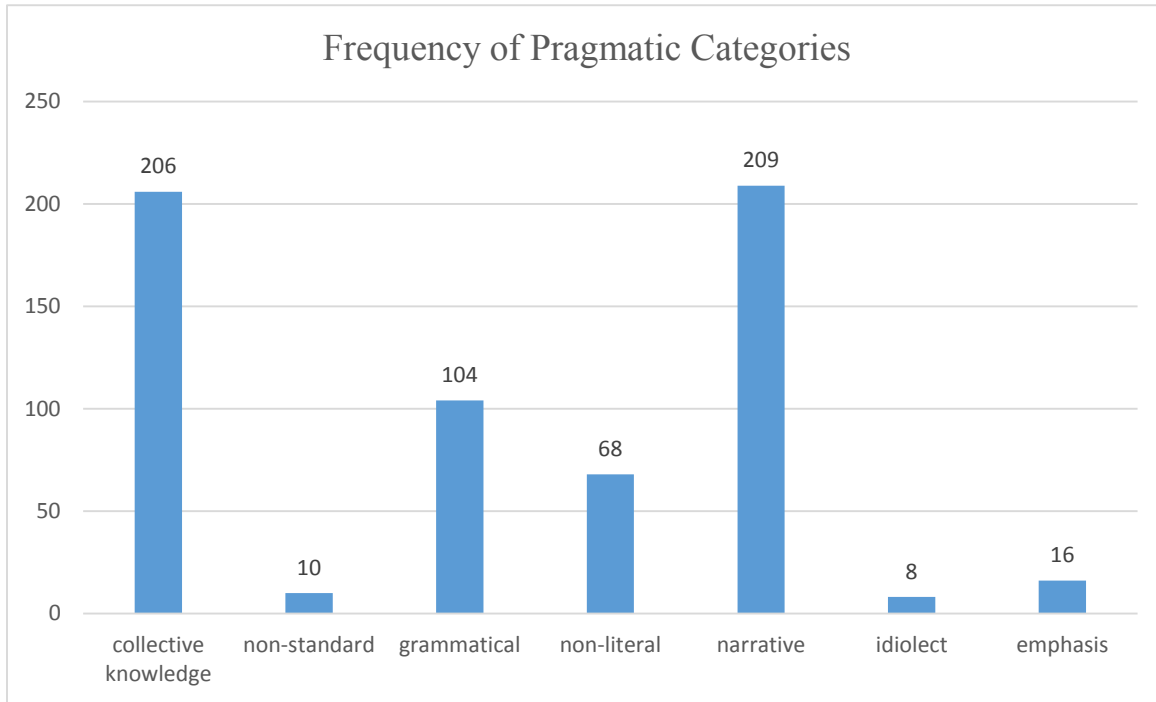
The Collective Knowledge category may have such a high frequency because it contains three distinct subcategories related by the broad concept of common knowledge and lived experience. It encompasses a range of quotations used to communicate that the reviewer believes the reader shares the same linguistic knowledge, cultural background, or similar life experiences. Many quotations fall into this category.

The Narrative category may have such a high frequency because so many of the reviewers not only assessed a business, but they told a story about their experience with it. Most of the reviews were about a reviewer going into a company, interacting with employees, purchasing a product or service, and then reflecting on it. This chain of events lends itself well to a narrative structure which includes quoting the "characters" in the story.

The Non-literal category had the fourth highest frequency. It was noted in the Introduction that quotations were a popular way to signal irony in online text. Much non-literal language

contains irony of one sort or another. In this particular category, the subcategories of sarcasm and understatement specifically are a form of irony. Non-literal language in the cases of irony is also connected to tone. Tone is difficult to express in writing. Writers have limited options, and quotations draw a reader's attention and communicate that there is an additional meaning in what is being written that is beyond face value. Also, in reviews with low star ratings, reviewers sometimes liked to use sarcasm to express their disappointment or frustration. These could all be explanations for the prevalence of quotations marking non-literal language in the reviews.

The Idiolect category may have the lowest frequency because reviewers want to be understood by their readers, and writing terms specific to their own language use, even though quoted, may confuse readers. The high frequency of the Collective Knowledge category suggests reviewers feel connected to readers through culture, background, and language. Because of this it seems, reviewers are more apt to use slang, idioms, or colloquialisms shared by the broader American society and not just unique to themselves. Also, there was only one instance of the Idiolect category where the reviewer quoted a term foreign to their language use. In many instances with such a relatively short writing sample, it is difficult to tell whether a quoted word or phrase is something a reviewer does not normally use. There were cases where reviewers used language that maybe they would not normally, but often they were quoting others they interacted with. This type of quotation fits best in the Narrative category. These reasons may explain why the Idiolect category is so low.

The Non-standard category likely has a low frequency for the same reasons as the Idiolect category. While the non-standard words or phrases used may not be unique or completely foreign to the reviewer writing them, they are not common to general American readers. Reviewers quote language they use in an atypical way as a signal to readers they are

37

aware of its unusualness. However, reviewers want to be understood and connect with readers, so they probably do not want to write strange usages and non-standard descriptions too frequently.

Another category with a low count is Emphasis. This has a few possible explanations. One explanation is that using all caps or capitalization in general is a popular way for writers to emphasize the meaning of their words. If bolding is an available formatting option, this too is a way to emphasize language. Quotations seem to be a lesser used option for expressing emphasis. Another explanation is that the researcher was wary of using this category as a default while analyzing the quotations, so she only put quotations in this category where the reviewer seemed to be using the quotes as an intensifying marker.

The prevalence of the Grammatical category is difficult to explain. Most of the instances were reviewers using quotes around proper nouns. As explained earlier, this could be done as a clarifying measure. Since capitalization can be used for emphasis as well as to designate certain proper nouns, maybe reviewers felt the need to add quotes to make it known they were referring to something specific and by name. The other instances were reviewers using quotes around a word or letter or using quotes to mark an aside or appositive. The quotes around a word or letter is pretty conventional. When a writer refers to the word "review" for instance, he or she may use quotes or italics to mark this. The same can be true for letters of the alphabet. The third instance of asides and appositives is more puzzling. Writers typically use dashes, parentheses, or commas to mark these features. This use of quotations in the Grammatical category was very rare however. It is unclear why some review writers chose quotations as a convention for marking asides and appositives.

Lay Reader Perceptions of Quotations

A total of three surveys with 15 unique Yelp reviews and review excerpts was posted on Amazon's Mechanical Turk site. These were reviews the researcher has previously analyzed that contained quotations she categorized according to the seven categories mentioned above. The Mechanical Turk workers were given the option of categorizing one quotation in each review according to the seven categories above or the option of creating an entirely new category if they found the provided categories insufficient. 97 Mechanical Turk workers completed version one of the survey; 87 completed version two; and 92 completed version three.

First Survey.

Table 1. Categories Chosen by Largest Group and Second Largest Group in Survey One

	Largest Group (%)	Second Largest Group (%)
Non-literal	46.39 Narrative	16.49 Non-literal
Narrative	28.87 Grammatical	19.59 Emphasis
Emphasis	53.61 Emphasis	13.4 Idiolect
Collective knowledge (connotation)	68.04 Non-literal	13.40 Non-standard
Collective knowledge (slang)	44.33 Idiolect	24.74 Collective Knowledge

This table shows that in the first survey, the largest percentage of Mechanical Turk participants agreed with the researcher in the classification of the third review's quotation. The second largest percentage of participants agreed with her in the classification of the first review and last review's quotations. Neither the largest nor second largest percentages agreed with the researcher in the classification of the remaining two reviews' quotations.

The five reviews in version one of the survey contained a quotation from the following categories: Non-literal (sarcasm), Narrative, Emphasis, Collective Knowledge (connotation), and Collective Knowledge (slang) (See Appendix). For the first review, 46.39% of participants chose

the Narrative category with 45 votes, while 16.49% chose the Non-literal category with 16 votes. This was one of the quotations that the researcher and rater disagreed on. In this case, the rater also put this quotation into the Narrative category. This example shows the difficulty of choosing one category for each quotation as they can be interpreted as having more than one pragmatic purpose. However, the participants' agreement with the rater that this quotation was used to help narrate the story of the review is interesting.

The researcher labeled the second review's quotation as being Narrative. The rater labeled it as Collective Knowledge (shared experience). The survey participants disagreed with both categorizations. 28.87% of workers put the quote in the Grammatical category with 28 votes. The second highest choice was the Emphasis category with 19.59% and 19 votes. The third highest was the Narrative category with 15.46% and 15 votes. The workers' assessment makes a great deal of sense because of the ambiguity of whether the phrase "Black Hair Care Products" is quoted because the reviewer sees it as a proper noun. The context is a bit more complicated because the reviewer is quoting herself as she was speaking to a store associate, which falls into the Narrative category. However, the rater's assessment also has some merit. In order to understand why the reviewer is asking for "Black Hair Care Products," one has to know that, generally speaking, African American hair and Caucasian hair is different, and some African American women prefer products made specifically for their hair type. A reader must also know that these types of products are not found in every store. This example also illustrates the complicated nature of categorizing these quotations. There are cases where arguments can be made that a reviewer uses quotations for multiple pragmatic reasons.

The researcher put the quotation from the third review in the Emphasis category. The rater put the quotation in the Narrative category. 53.61% of Mechanical Turk workers agreed

with the researcher and chose the Emphasis category with 52 votes. The second choice for participants was the Idiolect category with 13.4% and 13 votes. In this case, there seemed to be a high percentage of participants agreeing and interpreting the quotation as having a main pragmatic function.

The categorizations of the quotations in the remaining reviews were all agreed upon by the researcher and rater. The quotation in the next review was categorized as Collective Knowledge (connotation) by the researcher. 68.04% of survey participants disagreed and chose the Non-literal category with 66 votes. The second highest group chose the Non-standard category with 13.40% and 13 votes. The majority choosing the Non-literal category was a logical and fair assessment on the part of the survey participants. It should be noted that the researcher did not give the participants information about the subcategories in the survey, and the category definitions were simplified. This is an added complication that makes it difficult for the participants and the researcher to come to the same conclusion on all quotations. In this case, the quoted word, "accidentally" was used in a non-literal way. It also had a connotation that depended on a reader's background knowledge in order to be understood.

The last quotation in the version one survey was categorized as Collective knowledge (slang) by the researcher. Some survey participants disagreed, and 44.33% chose the Idiolect category with 43 votes. The second highest category, however, was Collective Knowledge with 24.74% and 24 votes. The workers' assessment of this quotation also makes sense. This particular review describes a person going with their best friend to a sushi restaurant. The reviewer says the friend has become a self-proclaimed "sushi snob." If taken literally, the self-proclaimed phrase could mean that the term "sushi snob" is part of the friend's idiolect but not part of the reviewer's idiolect. Therefore, the reviewer quoted the term because they normally do

not use it in their speech. The Collective Knowledge interpretation sees "snob" as a general slang word that can be preceded by a number of nouns to describe someone who is a connoisseur of one kind or another. For example, someone could be an "Indian food snob" or a "beer snob" and so forth. Both interpretations are valid.

Second Survey.

Table 2. Categories Chosen by Largest Group and Second Largest Group in Survey Two

	Largest Group (%)	Second Largest Group (%)
Grammatical	29.89 Grammatical	27.59 Collective Knowledge
Idiolect	35.63 Idiolect	35.63 Collective Knowledge
Narrative	59.77 Narrative	12.64 Non-literal
Non-literal (metaphor)	26.44 Non-literal	24.14 Idiolect
Non-standard	43.68 Emphasis	29.89 Non-standard

This table shows that the largest percentage of Mechanical Turk participants agreed with the researcher in the classification of the quotations in the first four reviews. The second largest percentage of participants agreed with her in the classification of the quotation in the last review.

The second version of the survey had five reviews with quotations that fall into the following categories: Grammatical, Idiolect, Narrative, Non-literal (metaphor), and Non-standard (See Appendix). For the quotation in the first review, 29.89% of survey participants agreed with the researcher with 26 categorizing the quote as Grammatical. However, the Collective Knowledge category was neck and neck with 27.59% and 24 votes. This particular quotation was about a business offering a "Ladies Night." This term could be seen as a proper noun as the name of the event at the business. The term could also be seen as referring to a culturally understood practice where a group of female friends get together and do something fun. Both interpretations have merit.

The next quotation was put into the Idiolect category. 35.63% or 31 survey takers agreed with this classification. An additional 35.63% or 31 participants believed this quote to be demonstrating Collective Knowledge. So the categories of Idiolect and Collective Knowledge were tied. The quoted phrase refers to the reviewer talking about their neighborhood growing up. The quotation reads "I grew up in the same little 'edge of Phoenix' neighborhood." The words "edge of Phoenix" is not something that means much of anything to someone unfamiliar with this area. However, substituting "Phoenix" for the name of any other city is a common expression in our language (e.g., "the edge of Dallas"). Readers know this refers to the outskirts of a particular city. However, "edge of Phoenix" could refer to a specific type of neighborhood that has a set of connotations only known by those who have grown up in a similar area. In this case, the quote would fall into the Idiolect category since a select group would use this expression to describe their specific types of neighborhoods around Phoenix. Both interpretations seem valid.

The next quotation falls in the Narrative category. The survey takers agreed, and 59.77% chose the same category with 52 votes. The second highest category was Non-literal with only 12.64% or 11 votes. This quotation involved the reviewer quoting their own thoughts about a particular business location. This quotation's use was straightforward and did not present a lot of ambiguity. The reviewer was describing the details of a location to the reader and narrated their own thoughts to make the review story-like and entertaining.

The next quotation was placed in the Non-literal (metaphor) category by the researcher. 26.44% of participants agreed with this classification with 23 votes. However, two other categories were close behind: Idiolect with 24.14% or 21 votes and Non-standard with 22.99% or 20 votes. The Mechanical Turk workers were quite divided on how to categorize this quotation.

The reviewer wrote the following: "Also their marinated steak tips are apparently marinated in 'unicorn tears' lol." The phrase in question is "unicorn tears." The researcher's classification of Non-literal makes sense since the reviewer does not actually believe the steak tips are marinated in a fictional animal's tears. The subcategory of metaphor was chosen because something concrete, like the tears of an animal, represents the concept of something amazing – the steaks have been marinated in something amazing to taste so good. Although, of course the steaks were marinated in an actual sauce of some sort and not marinated in an abstract concept, but the classification is sound.

Those who chose the Idiolect category must be thinking that the reviewer uses the phrase "unicorn tears" to represent something amazing on a somewhat regular basis in their speech, making this phrase a part of their idiolect. This could be true. However, the researcher has encountered this phrase elsewhere and used in the same humorous way. This suggests that the expression is not unique to this reviewer. Those who chose the Non-standard category must think that the expression is out of the ordinary and a bit strange. They perhaps have not come across it used in other contexts and assume the author is grasping at a sufficient description, so they took some creative license and used this phrase in a non-standard way. This makes sense. The Non-literal category and Non-standard category are fair classifications for this quotation. The Idiolect category is less convincing; however, it is logical that a participant would classify it as such without exposure to this specific phrase in other contexts.

The last quotation in the version two survey was classified as Non-standard. 43.68% of participants disagreed and chose the Emphasis category with 38 votes. However, the second chosen category was Non-standard with 29.89% or 26 votes. The review reads, "why is the place to be after work in Scottsdale a restaurant in a mall 'basement'?" The word in question is

"basement." The reviewer could be trying to emphasize this word because it is obvious, from the rest of the review, that they found the location of this restaurant strange. Not only this, but the reviewer is incredulous that a popular restaurant would be located in such a place and uses quotes to emphasize their attitude of bewilderment. This is a possibility. It is also possible that the reviewer had a difficult time knowing how to describe the location of this restaurant. The reviewer knows that readers are not accustomed to talking or reading about mall basements. Malls are structures that are thought of as having a ground floor and possibly other floors above, with each floor containing a variety of stores. Sometimes parking garages can be located below the ground floor, but these are not referred to as basements. "Basements" is a term typically used to describe the subterranean floor of a residential dwelling. If this is the interpretation, then the Non-standard category would be the most appropriate. However, both interpretations have merit.

Third Survey.

Table 3. Categories Chosen by Largest Group and Second Largest Group in Survey Three

	Largest Group (%)	Second Largest Group (%)
Emphasis	63.04 Emphasis	10.87 Collective Knowledge
Collective Knowledge (shared experience)	43.38 Narrative	18.49 Idiolect
Non-literal (sarcasm)	48.91 Non-literal	25.00 Emphasis
Non-standard	47.83 Non-standard	17.39 Emphasis
Idiolect	39.13 Idiolect	25.00 Collective Knowledge

This table shows that the largest percentage of Mechanical Turk participants agreed with the researcher in the classification of the quotations in the first, third, fourth, and fifth reviews. Neither the largest percentage nor the second largest percentage agreed with her in the classification of the quotation in the second review.

The third version of the survey had quotations from five reviews put into the following categories: Emphasis, Collective Knowledge (shared experience), Non-literal (sarcasm), Non-

standard, and Idiolect (See Appendix). The quotation in the first review was classified as Emphasis. 63.04% of survey takers agreed and chose the Emphasis category with 58 votes. The second highest category was Collective Knowledge with 10.87% or 10 votes. In this review, the reviewer uses quotes to emphasize their disgust with a bar's lack of sanitation. Not only did the reviewer put quotes around the word "ick," but they also put the word in all caps. This, along with the context of the rest of the review, helped the researcher categorize this quotation as Emphasis. Its classification was straightforward, and the majority of survey takers did not have trouble classifying it either.

The quotation in the second review was classified as Collective Knowledge (shared experience). The Mechanical Turk workers disagreed a great deal with 43.38% choosing the Narrative category with 40 votes. The second highest category was Idiolect with 18.49% or 17 votes. Then came Non-literal with 14.13% or 13 votes and then Collective Knowledge with 11.96% or 11 votes. The review reads, "I mean the brand Deisel alone makes me irritated. It so 'I hate my life! Its so hard!'." The quotation in question is "I hate my life! Its so hard!" A narrative interpretation of this quote is possible if one believes the reviewer is quoting their own thoughts and using such a quotation to make the review more story-like. However, this review is not about the reviewer's specific experience at this store. The reviewer does not tell a story. Rather it is a general rant about what the reviewer hates about the store and why. The researcher does not think that Narrative is the best suited category to describe this quotation.

Other survey takers chose the Idiolect category. This interpretation makes sense. Maybe the reviewer often uses the expressions "I hate my life!" and "Its so hard!" to describe things they do not like. This is a possibility. The Non-literal category makes sense as well because a brand of clothing is inanimate and, therefore, cannot hate its life or experience hardship.

46

However, the researcher argues that Collective Knowledge (shared experience) is the best category option for this quotation. The reviewer also uses the following quotation to describe the store in general: "It's all like 'I wanna look like I have nothing but wanna pay through the nose for it!'" This expression combined with the one above suggest the reviewer expects readers to have a common enough background that they would understand these whiny sentences to be approximate descriptions of some upscale clothing stores. The reviewer assumes that readers are familiar with the vibe these kinds of stores emit. Sometimes this vibe can communicate a pained false-sense-of-suffering existence for those who, in reality, are doing quite well but want to portray the persona of a common person with everyday difficulties. The reviewer assumes that readers know people like this who shop at the kind of store they describe. This interpretation would place the quote in the Collective Knowledge (shared experience) category.

The quotation in the next review was classified as Non-literal (sarcasm). 48.91% of survey takers agreed and chose the non-literal category with 45 votes. The second highest category was Emphasis with 25% or 23 votes. The review reads, "Then upon being seated was told by the 'server' that there were five tables ahead of me and it would be a while." This reviewer was very disappointed with the service they received at a restaurant. They put quotes around "server" to indicate that they did not consider their server to really be doing a good job, and therefore, the term "server" was not a fitting description for the person's occupation. The reviewer was using sarcasm by suggesting that their server was not really serving them. The survey takers who chose the Emphasis category were maybe thinking the reviewer was trying to emphasize the term in a pejorative way by putting quotes around it. However, the Non-literal (sarcasm) category is more fitting for this quotation.

The quotation in the next review was classified as Non-standard. The participants agreed, with 47.83% choosing this category as well with 44 votes. The second highest category chosen was Emphasis with 17.39% or 16 votes. This reviewer is writing about her positive experience at a salon and writes that she has "found a 'hair' home." Hair is not a common adjective used to describe the word "home." Because of this, the researcher and many survey takers classified this quotation as Non-standard. Those who chose Emphasis maybe thought the reviewer was drawing attention to "hair" because the reviewer had not really found a home; she had found a salon where she felt comfortable. But the reviewer maybe wanted to emphasize what kind of "home" she found, and that it was for her hair. Both of these interpretations make sense. However, it seems more plausible that the reviewer was conscious of her unusual pairing of the words "hair" and "home," so she used quotations in acknowledgement of this.

The quotation in the last review was placed into the Idiolect category. 39.13% of Mechanical Turk workers agreed with this classification, and 36 categorized the quote this way too. The second highest category chosen was Collective Knowledge with 25.00% or 23 votes. The reviewer is writing about a restaurant/bar that has a variety of beer options. The review reads, "It helped me expand my 'beer horizons.'" The researcher and many survey takers interpreted "beer horizons" to be a phrase specifically used by the reviewer in his own idiolect. However, some chose the Collective Knowledge category. These participants were perhaps thinking of the general phase "expand your horizons" or variants of it. It is true that readers would have to know about this phrase in order to make sense of the reviewer's quotation. However, it is not necessarily common for people to put a specific noun in front of "horizons" when using this saying. While other reviewers might write, "It helped me expand my horizons,"

this reviewer seems unique in writing "beer horizons" instead of just "horizons." Both interpretations have merit.

Overall Agreement

Table 4. Total Agreement of Largest Group and Second Largest Group for All Surveys

	Collective knowledge	Non-standard	Grammatical	Non-literal	Narrative	Idiolect	Emphasis
1	other	N/A	N/A	other	other	N/A	53.61%
	24.74%	N/A	N/A	16.49%	other	N/A	other
2	N/A	other	29.89%	26.44%	59.77%	35.63%	N/A
	N/A	29.89%	other	other	other	other	N/A
3	other	47.83%	N/A	48.91%	N/A	39.13%	63.04%
	other	other	N/A	other	N/A	other	other

In the above table, the numbers in the left-hand column indicate the version of the survey. The term "other" indicates that the participants chose a different category from the researcher, or from the category listed in the top row. "N/A" means that in a specific version of the survey (1, 2, or 3), there were no reviews with quotations matching the category heading in the top row. This table shows that of the 15 reviews given in the three surveys, the largest percentage of Mechanical Turk participants agreed with the researcher's quotation classification nine times, and the second largest percentage agreed with her three times.

In summary, the participants' categorization matched the researcher's to some extent, whether the highest percentage agreed with her or the second highest percentage agreed. This section certainly illustrates the challenges in interpreting quotations through a pragmatic lens. Pragmatic messages are so context dependent and are at the mercy of those interpreting them, whose life experiences inform their interpretation.

Reading Online Reviews. The participants were not only asked to categorize quotations, but they were asked demographic questions as well. They were asked how often they read online reviews and how often they read Yelp reviews. Their responses are shown in the table below.

Table 5. Frequency of Survey Participants Reading Online Reviews and Yelp Reviews

2-3/Month and More	Survey 1	Survey 2	Survey 3
Online Reviews	75.2%	79.3%	82.6%
Yelp Reviews	45.4%	46.0%	51.1%

75.2% of those who took version one of the survey read online reviews at least two to three times a month and more frequently. 45.4% of those same participants read Yelp reviews at least two to three times a month and more frequently. 79.3% of those who took version two of the survey read online reviews at least two to three times a month and more frequently. 46% of the same participants read Yelp reviews with the same frequency. 82.6% of those who took version three of the survey read online reviews at least two to three times a month and more frequently. 51.1% of those same participants read Yelp reviews at the same frequency. The participants' familiarity with online reviews would help them interpret quotations because they have likely come across reviews with quotations before.

Sentiment Connection. In all versions of the surveys, participants were also asked to predict a connection between a pragmatic category and star rating. There was not a lot of consensus on these questions, but there were some trends. 51.4% of participants across all surveys thought the Collective Knowledge category would likely be associated with a four star rating. 44.6% of participants across all surveys thought the Grammatical category would likely be associated with a four star rating. 40.6% of participants across all surveys thought the Non-standard category would likely be associated with a three star rating. These perceptions were based on the definitions of the categories, the examples provided, and the review excerpts.

The majority of participants responded that the Collective Knowledge category would have a high, positive star rating. This may be explained by the fact that reviewers seemingly use this category because they feel a common connection with the reader. They expect that readers share a common background, culture, and lived experience. This connection is positive, and so the reviewer is maybe more likely to use collective knowledge quotations in positive reviews.

A large percentage of participants responded that the Grammatical category would have a high, positive star rating. This may be because reviewers sometimes name specific products and put quotations around the product names. This happens often when reviewers are recommending something they purchased to the reader. In this case, the reviewer would have had a positive experience with the business and with the product they purchased, so the review would have a high star rating.

A sizable percentage of participants responded that the Non-standard category would have a neutral star rating. This could be because the participants perceived no pattern in how they thought non-standard quotations were used. Reviewers could describe things in a non-standard way when they had a positive or negative experience. Perhaps this is why survey takers thought it was best to link this category to a neutral star rating.

Statistical Analysis

In the sample of 411 reviews analyzed, the majority had a four or five star rating. The one star rating group was the third highest, and the two and three star groups were about even after that. The chart below displays the reviews distributed according to star rating.

51

Figure 2. Count of Five Star Ratings in Analyzed Reviews
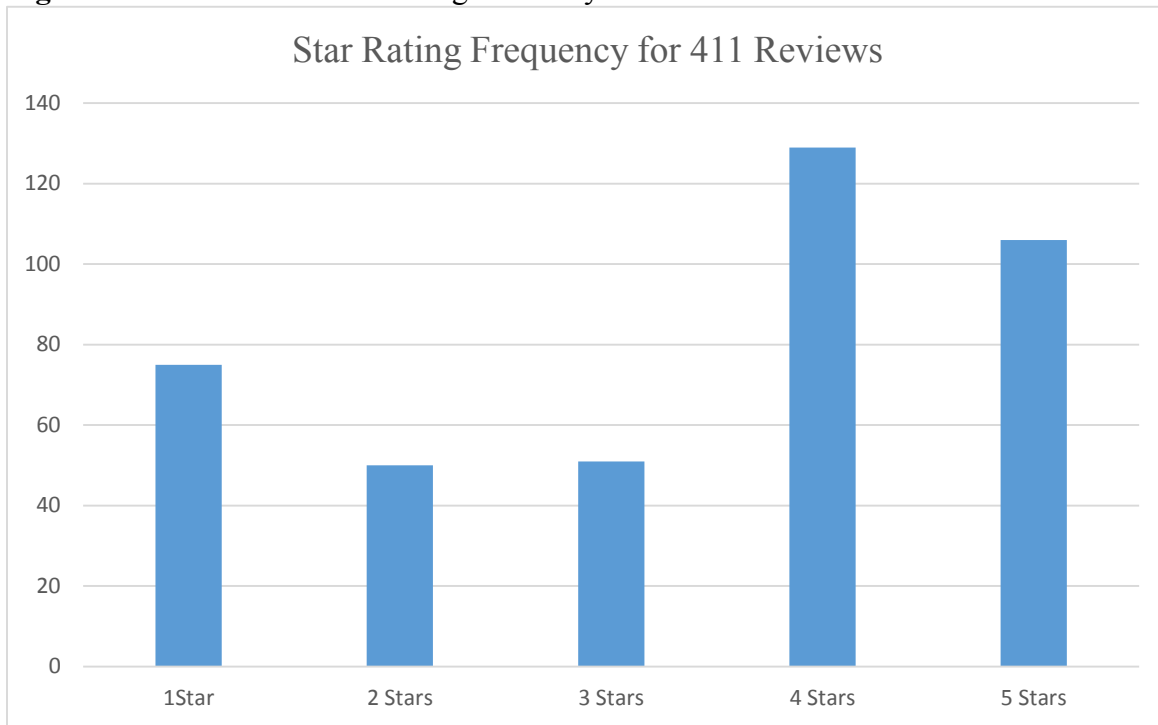
Star ratings are generally skewed towards the positive end on Yelp. So this sample follows that trend. This sample also shows that the reviews tend to be more polarized, either positive or negative.

Once the quotations were categorized, they were separated according to star rating. The median star rating was calculated for each category, and the chart below displays the results. The medians for each category represent a more accurate picture of the general star rating frequency for the pragmatic categories. It is interesting to note that two of the categories (Collective Knowledge and Grammatical) have medians that match the predictions chosen by a large percentage of survey takers. In these two cases, their predictions seemed to be quite accurate. These medians are also similar to the ANOVA results for the Narrative and Non-literal categories.

Figure 3. Median Star Rating for Seven Quotation Categories
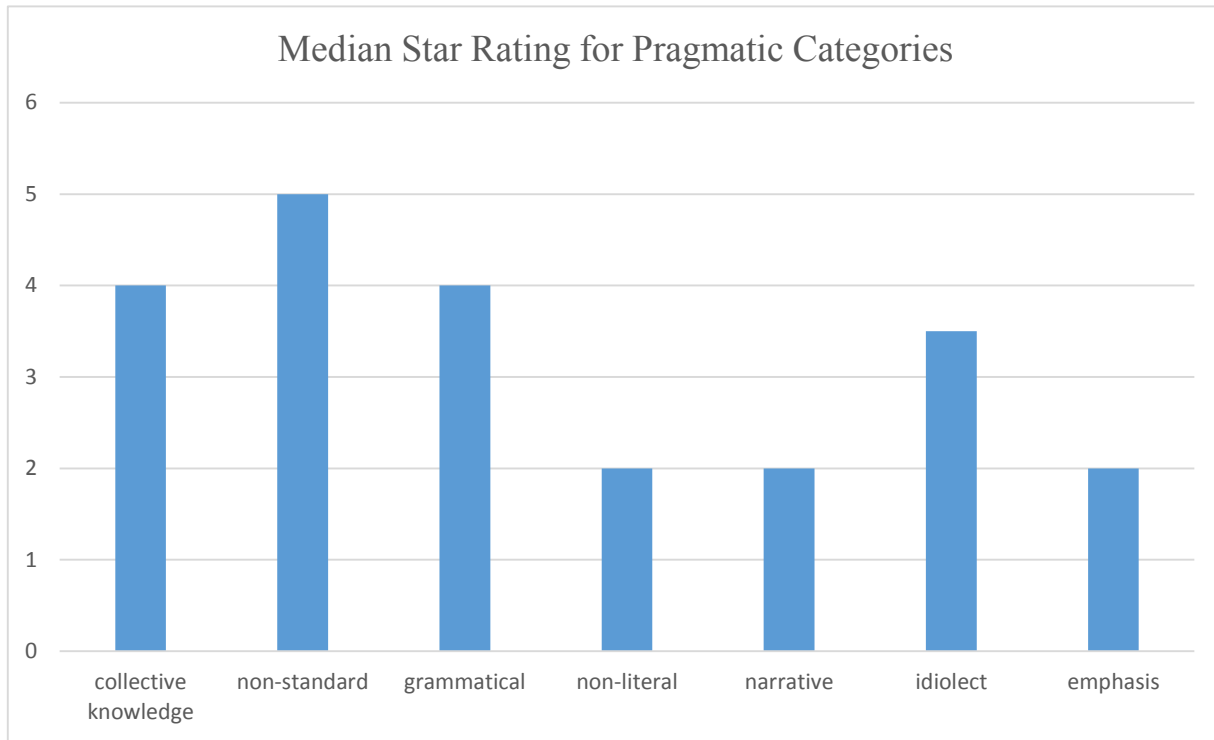
The results from the ANOVA are as follows: $F_{(6, 35)} = 17.991$, $p < .0005$. These seven categories accounted for about 15% of the variance in star rating with an R^2 of .150. A Tukey HSD test was run to determine which groups were statistically significant from each other. The Collective Knowledge category was significantly different from the Narrative and Non-literal categories. The Grammatical category was significantly different from the Narrative, Non-literal, and Emphasis categories.

As shown in the chart above, the Collective Knowledge category generally had a higher star rating, while the Non-literal and Narrative categories had lower star ratings. The Grammatical category generally also had a higher star rating, while the Non-literal, Narrative, and Emphasis categories had lower star ratings. The polarity in the star ratings for these categories likely accounts for the significance in the differences. It is surprising that the Non-standard category was not significantly different from the Non-literal, Narrative, or Emphasis

categories. However, the Non-standard category had few samples, so the median value for that category may be misleading, and this category may generally have a lower star rating than appears. It likely this category was not statistically significant because it only contained 10 quotations. Other categories with few samples were also not significant in the Tukey HSD test.

The Collective Knowledge category likely has a higher star rating because reviewers use this type of quotation to be understood and to connect with readers. The connection between reviewers and readers is positive, and so perhaps these types of quotations are associated with a higher star rating because they are present in reviews where the reviewer expresses a positive attitude. The Grammatical category may have a higher star rating for a different reason. Many times reviewers put quotes around specific products mentioned in their review. Often times they are recommending these products or a particular business to the readers. When a reviewer recommends a business or a product, it suggests they had a positive experience, which would result in a higher star rating.

The Non-literal, Narrative, and Emphasis categories all have a lower star rating. Included in the Non-literal category are the subcategories of sarcasm and understatement. The quotations in these subcategories are often used to convey a negative tone. Sarcasm especially is used to express a critical attitude. The Sarcasm subcategory was the most prevalent subcategory within the Non-literal category at 76.47%. This could account for that category's lower star rating.

The Narrative category has a lower star rating as well. While some reviews with narration are positive, many others are negative. Reviewers often quote their thoughts and interactions with others at a business because they are telling a story about an experience that was rude, outrageous, or unpleasant. It seems reviewers like to justify their low star ratings through

54

narration with specific quotes to prove that their assessment of a business is valid. Once a reviewer states their negative opinion in their review, they feel compelled to explain it.

The Emphasis category is also associated with a lower star rating. Emphasis is often used to make the meaning of the quoted language more forceful. Sometimes these quotes are negative, and the reviewer wants to draw attention to the negative language. The negative quotes indicate the reviewer was not happy with a business or product and probably gave it a low star rating. However, this category also had few samples. The above is a possible explanation for the lower star rating, but with so few samples, it is not wise to make a generalization about Emphasis always being connected to a low star rating.

Summary. After an analysis of hundreds of Yelp reviews, seven main pragmatic functions were identified in regard to quotation use: Collective Knowledge, Emphasis, Non-literal, Non-standard, Grammatical, Narrative, and Idiolect. These categories align well with the quotation functions described by Predelli in the Review of Literature section. Although some of these quotation types may be more common in online reviews, others may be more generalizable to additional online genres and social media.

The survey takers, or lay readers, sometimes had a different perception of how a quotation was being used from the researcher. However, there were several times where the perceptions of the largest percentage and the researcher coincided exactly. There were other times where the researcher's view matched that of the second largest percentage of survey takers. The participants' responses helped validate the researcher's discovery of these seven categories. The survey takers opted to create their own categories only three times. One of these times, the participant described the quote as sarcastic, which the researcher covered in the Non-literal category. In all other instances, the survey participants chose one of the categories the researcher

55

provided for them. These seven categories seem to be useful descriptions of how reviewers are pragmatically communicating with quotations in Yelp reviews.

The results of the ANOVA were statistically significant with quotation use accounting for 15% of the variance in the data. The Tukey HSD test revealed the Collective Knowledge and Grammatical categories to be different from the Non-literal, Narrative, and Emphasis categories. Some categories seem to have a strong connection to star rating, particularly the Narrative and Non-literal categories. These seem to be correlated with a lower star rating. If these pragmatic functions can be identified in reviews or other online writing, then it is likely the text has a lower star rating, and the author is expressing negative sentiment. These types of patterns could be incorporated into algorithms that calculate sentiment analysis. Quotation use combined with lexicons and other CMC cues could improve the accuracy of sentiment analysis.

Chapter 5

Conclusion

<u>**Summary of Findings**</u>

The research questions this paper sought to answer were the following: What are the pragmatic uses of quotations in online writing, specifically reviews? 2) How do lay readers perceive the use of these quotations? 3) Can quotation use help predict author sentiment? The first research question was answered through a reading analysis of 411 Yelp reviews and 621 quotations. Drawing on the researcher's pilot study, previous research about quotations, specifically Predelli's work, and observations during the current study, it was determined that there are at least seven pragmatic uses for quotations in online Yelp reviews. These seven categories include the following: Collective Knowledge, Narrative, Non-literal, Non-standard, Emphasis, Idiolect, and Grammatical.

It was determined that Narrative and Collective Knowledge were the categories with the highest frequency, while Idiolect and Non-standard had the lowest frequency. This suggests that reviewers like to make their reviews into stories and that they like to connect with readers and assume readers have a similar background to their own. This same use of quotations may be in other review texts as well. These seven categories are very similar to the quotation uses Predelli described, and his examples were not from online reviews, so perhaps these pragmatic quotation uses appear in other genres too. However, Yelp reviews are a specific sub-genre. Much more research on quotations in other social media genres would be needed to confirm these patterns of pragmatic quotation use.

The second research question was answered by collecting survey data on Mechanical Turk from non-linguist participants. These participants proved to be familiar with online reviews.

The largest percentage of participants or the second largest percentage often times agreed with the researcher's pragmatic categorization of specific quotations within Yelp reviews. Only three times did a survey participant create their own category because they felt none of the seven categories was sufficient to describe how a quotation was being used. It seems lay readers perceive the quotation use in online reviews along the same lines as the categories described by the researcher. Although, this was not always the case. A reader's interpretation of a quotation is influenced by numerous variables, including their own background, experiences, personality, etc. However, there was enough consensus generally to suggest that the categories really do describe real pragmatic communication strategies. It would be interesting to conduct additional in-depth studies about lay reader pragmatic interpretation.

The third research question was answered by calculating the median star rating for each category and by performing an ANOVA comparing the categories and star ratings against each other. The results here showed that these pragmatic categories accounted for 15% of the star rating variance. So pragmatic quotation use has some effect on predicting reviewer sentiment. It was also discovered that the Narrative and Non-literal categories were significantly different from the Collective Knowledge category. The Non-literal, Narrative, and Emphasis categories were all significantly different from the Grammatical category as well.

The Collective Knowledge category's connection with high star rating was possibly explained by the desire of reviewers to connect with readers in a positive way. This may mean reviewers are more likely to use quotations demonstrating collective knowledge when they have written a positive review, which would mean a higher star rating. The Grammatical category's association with a higher star rating may be the result of reviewers specifically naming and putting quotes around product names or business names because they are recommending them to

58

readers. In these cases, the reviewer would have given a higher star rating because of their positive experience.

The Narrative category's association with lower star ratings could be the result of reviewers choosing to tell negative stories about their interactions with a particular establishment. This would seem similar to face-to-face communication as well. When someone has a positive experience or neutral experience with a business, they may tell their friends or family a little about it. However, when someone has had a very negative experience, they seem likely to discuss it at length. It could also be true that reviewers who have given a company a low star rating feel a need to explain or justify their choice, and so they write a story about their experience and use narrative quotations to make it more entertaining and clearer for the reader.

The Non-literal category's connection with low star rating is likely a result of the presence of the sarcasm and understatement subcategories. The sarcasm subcategory was particularly prevalent. Sarcasm and understatement can be used to communicate a negative tone. This was often the case in reviews with the Non-literal (sarcasm) category. Quotations used to communicate a critical ironic intent would result in a low star rating.

The Emphasis category's connection to low star rating could be explained by reviewers placing negative words or phrases in quotations to emphasize their negative sentiment. However, this category did not have many examples, so generalizations and explanations should be made with caution. Given the connections with these categories and a high or low star rating, it seems plausible that pragmatic quotation use could be used to help predict author sentiment. These pragmatic functions seemed to be used by reviewers towards one end of the star rating spectrum or the other. Quotations are definitely an option reviewers have for making their tone and other

59

non-verbal cues clearer. They seem to be taking advantage of this option to strengthen the sentiment behind their opinions of products and services.

Limitations

Although this study yielded some interesting and thought-provoking results, there were many limitations. The reviews from the Yelp 2014 corpus were chosen because they contained quotations; however, many of the reviews were specifically about restaurant experiences. This is not bad, but it is a subcategory of reviews that makes generalizing about the language use within them difficult. Although a significant number of quotations were analyzed (621), it would still be helpful for even more to be analyzed with regards to pragmatic quotation use. The researcher of this study did have a rater to verify her analysis, but getting a broader range of perspectives on how quotations are used in reviews would strengthen the argument for the seven categories found here or refute it.

There were limitations to the Mechanical Turk survey as well. One of the central limitations was that the participants were able to take one, two, or all three of the surveys. This was not a problem for the section of the surveys asking participants to classify the quotations. However, it was a problem for the question asking participants to predict a connection between category and star rating, as this question was the exact same for all versions of the survey. So there may have been some double counting on that question.

Another difficulty with the survey is that it is hard to know whether the participants read and understood all of the material and if they made a thoughtful effort in answering the questions. This is true of all surveys, and the researcher tried to account for this by removing data collected from participants who took under six minutes to complete the survey. Despite

removing this data, there could have been participants that made snap decisions about how to categorize the quotations or participants who did not read the entire review excerpt, and so on.

There were limitations when analyzing the collected data as well. Some categories had high frequencies, while others had low frequencies. This means that the median for each category, though useful, could only give so much insight depending on the quotation count for the category. The category's median star rating was more telling for categories with high frequencies.

The data was not normally distributed. While ANOVAs are robust tests, this characteristic of the data was still not ideal for a statistical analysis. Also because each review could have multiple quotations, there was no way getting around double counting some of the star ratings. This means the star ratings were not weighted equally. Despite these difficulties, the ANOVA still yielded some important insights.

<u>Future Research</u>

Pragmatic quotation use is an area of linguistics with little research, especially quotation use in online reviews and other types of social media. Many more studies could be carried out to gain further knowledge about this subject. Additional reading analyses of quotations from various online written corpora would be useful. These pragmatic patterns may be unique to the Yelp sub-genre, or they may be widely used across online genres. More research using texts from a variety of online genres would need to be conducted to know this for sure.

Surveys targeted at lay readers with a more qualitative approach may be useful. It would be interesting to see how the survey takers would describe these quotations, left to their own devices without the aid of predetermined categories. Free responses could be compared against

the category descriptions to see if lay readers really interpreted the pragmatic messages in the same way as the researcher. In addition demographic information could be used to see if participants' interpretations varied based on gender, race, nationality, etc.

Lastly, more research would need to be done to find a way to write rules for these pragmatic patterns, so they could be detected with scripts or algorithms. This would be necessary if pragmatic quotations were to be added to existing sentiment analysis programs to improve their accuracy. This task would prove extremely difficult but not impossible, especially if the same methods or rules were applied to quotations that apply to other snippets of text. For example, if a quotation contained nouns or adjectives already in a sentiment lexicon that had a polarity value, it may be feasible to attach these words to different pragmatic uses. Once a quotation's use was identified, that information could be incorporated into the larger sentiment analysis of the entire text to provide added insight. It will be interesting to see where the future of sentiment analysis goes and what kinds of linguistic features will prove to be the most salient in improving these programs' accuracy.

References

Balahur, A., Steinberger, R., Van der Goot, E., Pouliquen, B., & Kabadjov, M. (2009). Opinion mining on newspaper quotations. In 2009 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2009: proceedings, 15-18 September, 2009 : Università degli Studi di Milano Bicocca, Milano, Italy (pp. 523-526). doi:10.1109/WI-IAT.2009.340

Cappelen, H., & Lepore, E. (2003). Varieties of quotation revisited. *Belgian Journal Of Linguistics*, *17*(1), 51-75.

Chifu, E., Letia, T., Budisan, B., & Chifu, V. R. (2015). Web harvesting and sentiment analysis of consumer feedback. *Electronics and Telecommunications*, 56 (3), 7-14.

Dresner, E., & Herring, S. C. (2010). Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication Theory* (10503293), 20(3), 249-268. doi:10.1111/j.1468-2885.2010.01362.x

Eisterhold, J., Attardo, S., & Boxer, D. (2006). Reactions to irony in discourse: Evidence for the least disruption principle. *Journal of Pragmatics, 38*(8), 1239-1256. Retrieved from http://search.proquest.com/docview/85637907?accountid=4488

Ferrell, D. (2013, November 12). Yelp effect reshapes how businesses interact with customers. Orange County Register, The (Santa Ana, CA).

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing*, 90(2), 217-232. doi:http://dx.doi.org/10.1016/j.jretai.2014.04.004

Hancock, J., Landrigan, C., & Silver, C. (2007). Expressing emotion in text-based communication. In CHI'07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: ACM New York, NY (pp. 929-932). doi:10.1145/1240624.1240764

McAndrew, F. T., & De Jonge, C. R. (2011). Electronic person perception: What do we infer about people from the style of their E-mail messages? *Social Psychological and Personality Science, 2*(4), 403-407. Retrieved from http://search.proquest.com/docview/1373478553?accountid=4488

Novak, P. K., Smailovic, J., Sluban, B., & Mozetic, I. (2015). Sentiment of emojis. *PLoS One, 10*(12) doi:http://dx.doi.org/10.1371/journal.pone.0144296

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals Of Tourism Research*, 5067-83

Paula Carvalho, Luís Sarmento, Mário J. Silva, and Eugénio de Oliveira. 2009. Clues for detecting irony in user-generated contents: oh...!! it's "so easy" ;-). In Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion (TSA '09). ACM, New York, NY, USA, 53-56. DOI=http://dx.doi.org/10.1145/1651461.1651471

Predelli, S. (2003). Scare quotes and their relation to other semantic issues. *Linguistics and Philosophy, 26*(1), 1-28. doi:http://dx.doi.org/10.1023/A:1022278209949

Predelli, S. (2003). 'Subliminable' messages, scare quotes, and the use hypothesis. *Belgian Journal Of Linguistics*, *17*(1), 153-166.

Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*, 8130-40. doi:10.1016/j.dss.2015.10.006

Scare quotes [Def. 1]. (n.d.)*. OED Online.* In Oxford English Dictionary. Retrieved April 23, 2015, from http://www.oed.com/view/Entry/172014?redirectedFrom=scare+quotes.

Vandergriff, I. (2013). Emotive communication online: A contextual analysis of computer-mediated communication (CMC) cues. *Journal of Pragmatics*, 51, 1-12. Retrieved from http://search.proquest.com/docview/1541992714?accountid=4488

Whalen, J. M., Pexman, P. M., & Gill, A. J. (2009). 'Should be fun-not!'. incidence and marking of nonliteral language in E-mail. *Journal of Language and Social Psychology, 28*(3), 263-280. doi:http://dx.doi.org/10.1177/0261927X09335253

Subhabrata, M., Akshat, M., & Balamurali, A.R. (2012). TwiSent: A multistage system for analyzing sentiment in twitter. In CIKM '12 Proceedings of the 21st ACM international conference on Information and knowledge management: ACM New York, NY (pp. 2531-2534). doi:10.1145/2396761.2398684

<c<center>**Appendices**</center>

Appendix A − Python Script

```python
import json
import re
import glob
import io

file_out = io.open('yelp_final.csv', 'w', encoding='utf8')
file_out.write("ReviewID, BusinessID, city, stars, Review Count, text\n")

count = 0
overload = 156925
ticker = 0
review_count = 0
one_star = 0
two_star = 0
three_star = 0
four_star = 0
five_star = 0

data = []
reviews = []
city = []

with open("Yelp/Yelp/yelp_academic_dataset_business.json") as file:
    for line in file:
        data.append(json.loads(line))

with open("Yelp/Yelp/yelp_academic_dataset_review.json") as file_2:
    for line in file_2:
        if ticker == overload:
            print(ticker)
            for review in reviews:
                for record in data:
                    if review["business_id"] == record["business_id"]:
                        city = record["city"]
                        #print(city)
                        # if record["city"] == ('Phoenix|Pittsburgh|Charlotte|Urbana-Champaign|Las Vegas|Madison'):
                        #     review_count += 1
                # if review["stars"]== 1:
                # one_star += 1
                if re.match('.*("|\\b|\'\.)', review["text"]):  #follow or precede by white space (send him my program)
                    text = re.sub('\n', ' ', review["text"])
                    text = re.sub(',', '', text)

                    #print(review["business_id"])
                    #print(text)
                    #print(review["review_id"])
                    file_out.write(str(review["review_id"]) + ',' + (str(review["business_id"]) + ',' + city + ',' +
str(review["stars"]) + ',' + str(review_count) + ',' + text + '\n'))
            ticker = 0
            reviews = []

        else:
            reviews.append(json.loads(line))
            ticker += 1
    print(ticker)
```

65

```
# print(review_count)
# print(one_star)
```

Appendix B – Category Rubric

Pragmatic Classifications

1) Collective Knowledge (e.g., slang, shared experience, connotation, stock phrase)
2) Emphasis (e.g., makes something more forceful, calls audience's attention to it)
3) Idiolect (whether foreign or unique to author's idiolect/usage)
4) Narrative
5) Non-literal (sarcasm, metaphor, contradiction, understatement)
6) Non-standard
7) Grammatical

Collective Knowledge: Where the author quotes something because they expect the audience knows what they are referring to based on an assumed common experience, culture, or background. Sometimes this takes the form of the author expecting the audience to have shared experience. Other times the author simply quotes slang words (e.g., neologisms, colloquial language) or idiomatic expressions (e.g., stock phrases). Sometimes the author quotes something to make the audience aware of an added connotation to the word. However, the audience is supposed to know what the connotation means or implies based on this collective knowledge.

Examples
- The inside is very "old school" and there seems to be a deli attached although I didn't do any ordering from that side of the business. *(slang)*
- I don't golf at all but like to ride my bike along the multi-use path from McDonald Drive to Pima. I watch out for the golfers and occasionally have to duck or at least honker down when I hear "fore". *(shared experience: assumes audience knows golfing terminology)*
- Ordering "hot italian sausage" from a deli guy that calls you "babe" is as amusing as Sunday morning can get.   :-)   *(connotation)*
- He is "one of a kind" employee. *(stock phrase)*

Emphasis: Where the author wants to call specific attention to the language in quotes. Sometimes this is done to make whatever emotion expressed more forceful.

Example
- There are other items on the menu that may be better but since this is "Pizza" Picazzo I'd expect better.

Idiolect: Where the author quotes something they don't normally use in their idiolect (foreign) or something that is exclusive to their idiolect (unique).

Example
- It was 10:00 PM on a week night and not much is open around Central Phoenix aside from fast food so went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's" due to its renown as a repository for the after-closing-time crowd - that didn't go home with somebody. *(unique)*

Narrative: Where the author quotes another person, an outside source, or him/herself. The quote is relevant to his/her experience at the business he/she is reviewing. The author could also be quoting his/her own thoughts. Authors do this possibly for many reasons; however, a main reason is to make their review into a story for the reader. They also often use the quotations to support their opinion or to justify the rating they gave the business.

Examples
- I got a cup to get water but then juice came out. My dad told the cashier who answered "oh we don't have any more water." wow.
- Out here for a business trip found this place around during lunch. Huge mistake! Never go there! The food was not good and to top it off my fry chicken was hairy! It had hair all over! No manager was available and the girl in charge did not care. She just said (in her Chinese accent) "Oh that's normal we get those from time to time cuz we buy our chicken it comes like that." What?! Not to mention she made me pay full price! I paid and left! Never again!

Non-literal: Where the author uses quotes to signify that something should not be taken literally, is counterfactual, metaphorical, has a connotation, or is contradictory. This can happen when the author intends sarcasm, metaphor, contradiction, understatement, and more.

Examples
- When I came to pick up my items I was extremely embarrassed because I had to tell him I could not afford it. He then told me he could "help me out " and make it $80. I tried explaining to him that the two blazers were only $25 each to begin with. *(sarcasm)*
- I know that not everyone will understand why this HURTS my heart but in **"2013"** for someone to tell me that "Black people" are not worth a space on a shelf speaks volumes about why we as a country can't move forward! *(metaphor)*
- The "meat" and "seafood" dishes could really make you worry that you are eating meat. The stuffed tomatoes are delicious as well as the sweet and sour pork I also really liked the lemon grass fish. I took a friend who was a Janeist (VERY strict diet!) and he wouldn't believe that it wasn't meat lol. *(contradiction)*
- This is one of the only spots I bring my car because they do the best job by far! I had an issue with a major "spill" let's call it (vomit) in my back seat. *(understatement)*

Non-standard: Where the author uses a term or phrase in a non-standard or a-typical way.

Examples
- It's the first time I've "Yelped" anything other than a restaurant.
- Had to go out of town and looked for many different places to "board" my dog - he's about the size of a Corgi so he didn't need a lot of space but I definitely preferred to not keep him in a cage.

Grammatical: Where the author puts quotes around a proper noun such as a title or menu item. It is not very clear why reviewers do this, but a possible reason is that they do not have italics as an option.

Examples
- I read about this place in the book titled "1000 Places to See or Do in the US" and decided to go here when I was at a conference in Pittsburgh.

- I love this place great bar great selection of beer and food great atmosphere to watch games or hang out with friends the servers and bartenders are all nice and friendly....the "Bare knuckle stout" beer is awesome!! and the "Big Unit" Hot dog is something you gotta see for yourself........and if you are an Alice Cooper fan like myself is a plus....you'll have a great time.

Appendix C – Pilot Survey

What is your age?

- ❍ 18-24 (1)
- ❍ 25-34 (2)
- ❍ 35-44 (3)
- ❍ 45-54 (4)
- ❍ 55-64 (5)
- ❍ 65-74 (6)
- ❍ 75 or older (7)

What is your gender?

- ❍ Female (1)
- ❍ Male (2)

What is your highest level of education?

- ❍ high school diploma or equivalent (1)
- ❍ some college (2)
- ❍ trade/technical/vocational training (3)
- ❍ Bachelor's degree (4)
- ❍ Master's degree (5)
- ❍ Doctorate degree (6)

What is your ethnicity?

- ❍ Caucasian (1)
- ❍ Hispanic or Latino/Latina (2)
- ❍ Black or African American (3)
- ❍ Native American or American Indian (4)
- ❍ Asian/Pacific Islander (5)
- ❍ Other (6) _____

Are you a native English speaker?

○ Yes (1)
○ No (2)

Were you primarily raised in the United States?

○ Yes (1)
○ No (2)

How often do you read or write online reviews?

○ Never (1)
○ Less than Once a Month (2)
○ Once a Month (3)
○ 2-3 Times a Month (4)
○ Once a Week (5)
○ 2-3 Times a Week (6)
○ Daily (7)

How often do you read or write Yelp reviews?

○ Never (1)
○ Less than Once a Month (2)
○ Once a Month (3)
○ 2-3 Times a Month (4)
○ Once a Week (5)
○ 2-3 Times a Week (6)
○ Daily (7)

You will now read a series of 5 reviews. In each review, there will be at least one word or phrase in quotations. Your task is to categorize one quotation in each review based on what you think the author is trying to communicate. In other words, consider why the author chose to put quotes around a particular word or phrase. Each category, along with a description and examples, is below. The categories are based on previous linguistic research and on the researcher's own analysis. If none of the

categories describes what you consider to be the author's intentions, then create your own category, complete with a description. Please read the category descriptions and examples carefully.

1) Common Experience (it's assumed that the author and reader share a common experience, culture, or background) For example: includes slang like "old school" or idioms like "one of a kind" or descriptions like "college sized" refrigerator

2) Emphasis (the author wants the quoted word or phrase to stand out) For example: Since this is "Pizza" Picazzo I'd expect better.

3) Story (the author quotes someone or something to make the review more story-like) For example: The cashier answered "oh we don't have any more water." wow.

4) Connotation (the author quotes a word or phrase that has an additional meaning) For example: I had an issue with a major "spill" let's call it (vomit) in my back seat.

5) Uncommon Use (the author quotes a word or phrase when it is used in an uncommon way) For example: I looked for many different places to "board" my dog.

6) Author Speech (the author quotes something only they would say, or the author quotes something they would not normally say) For example: I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's."

7) Proper Noun (The author puts quotes around a specific person, place, or thing) For example: I read the book titled "1000 Places to See or Do in the US."

8) Other (Write your own category description)

Please read the following review carefully and choose one category that best describes how the phrase "help me out" is being used.

I brought two pairs of pants and two (cheap) blazers there to be altered. He was very nice. However he failed to mention to me that the cost of these 4 things- one of which only needed alternation tape- would come to $96!!!!!! When I came to pick up my items I was extremely embarrassed because I had to tell him I could not afford it. He then told me he could "help me out" and make it $80. I tried explaining to him that the two blazers were only $25 each to begin with. Overall this was an extremely humiliating experience and it all could have been solved if he were more professional and would have

told me that simple alterations would cost so much. I will definitely be searching for a new tailor in the future. Way too expensive if you ask me.

- ○ Common Experience (it's assumed that the author and reader share a common experience, culture, or background) For example: includes slang like "old school" or idioms like "one of a kind or descriptions like "college sized" refrigerator" (4)
- ○ Emphasis (the author wants the quoted word or phrase to stand out) For example: Since this is "Pizza" Picazzo I'd expect better. (19)
- ○ Story (the author quotes someone or something to make the review more story-like) For example: The cashier answered "oh we don't have any more water." wow. (6)
- ○ Connotation (the author quotes a word or phrase that has an additional meaning) For example: I had an issue with a major "spill" let's call it (vomit) in my back seat. (21)
- ○ Uncommon Use (the author quotes a word or phrase when it is used in an uncommon way) For example: I looked for many different places to "board" my dog. (23)
- ○ Author Speech (the author quotes something only they would say, or the author quotes something they would not normally say) For example: I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (20)
- ○ Proper Noun (The author puts quotes around a specific person, place, or thing) For example: I read the book titled "1000 Places to See or Do in the US." (8)
- ○ Other (Write your own category description) (24) _____

Please read the following review carefully and choose one category that best describes how the word "2013" is being used.

Soooooo... I'm at "Bed Bath and Beyond" in Paradise Valley picking up items for a trip to Texas. I ask the associate who works there where I might find "Black Hair Care Products." She walks up and down five aisles and is unable to locate products for Black hair. She calls another associate who also walks up and down the same aisles and is also unable to find anything for my hair. As embarrassed as I was the expression on both of these women's faces was heartbreaking for me because they both had to acknowledge that there was nothing for my hair on the shelves. Graciously they call their manager Bryan who walks over to me and says "We used to have a whole row of product for that kind of hair but no longer carry anything for Black hair. Not enough is sold so we discontinued selling the product. " I know that not everyone will understand why this HURTS my heart but in "2013" for someone to tell me that "Black people" are not worth a space on a shelf speaks volumes about why we as a country can't move forward! I have friends and family members from every race creed and walk of life and am completely enraged at what I just had to face. I do not know anyone more culturally sensitive then I am and find

myself frustrated that the gentleman that I just spoke to could not understand why I was so upset. The moral of this story forgive but never forget that we have a long way to go.

- ○ Common Experience (It's assumed that the author and reader share a common experience, culture, or background.) For example: includes slang like "old school" or idioms like "one of a kind or descriptions like "college sized" refrigerator" (1)
- ○ Emphasis (the author wants the quoted word or phrase to stand out) For example: Since this is "Pizza" Picazzo I'd expect better. (2)
- ○ Story (the author quotes someone or something to make the review more story-like) For example: The cashier answered "oh we don't have any more water." wow. (3)
- ○ Connotation (the author quotes a word or phrase that has an additional meaning) For example: I had an issue with a major "spill" let's call it (vomit) in my back seat. (4)
- ○ Uncommon Use (the author quotes a word or phrase when it is used in an uncommon way) For example: I looked for many different places to "board" my dog. (5)
- ○ Author Speech (the author quotes something only they would say, or the author quotes something they would not normally say) For example: I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (6)
- ○ Proper Noun (The author puts quotes around a specific person, place, or thing) For example: I read the book titled "1000 Places to See or Do in the US." (7)
- ○ Other (Write your own category description) (8) _____

Please read the following review carefully and choose one category that best describes how the phrase "overly harsh" is being used.

I confess. I thought that the negative reviews could be "overly harsh" or somehow just plain wrong. And for that I apologize. Not one bird there seemed to be well-adjusted pet material. Flocks of smaller birds (read quakers conures) were in 50? gallon aquariums. Maybe smaller. None of them seemed used to being handled. The large birds that were out of their cages all showed signs of distress - only about 1/3 of the birds had any of their long tailfeathers about 1/2 showed obvious signs of feather plucking in the form of bare patches across back or chest. In the back were amazons african greys and other parrots of similar size in birds in cages I MIGHT have temporarily put budgies in. I didn't have the heart to look at

them all. Their selection of food overpriced toys and other accessories was better than the average pet store but how can I justify spending money here? I'm looking elsewhere. Ugh.

○ Common Experience (It's assumed that the author and reader share a common experience, culture, or background.) For example: includes slang like "old school" or idioms like "one of a kind or descriptions like "college sized" refrigerator" (1)

○ Emphasis (the author wants the quoted word or phrase to stand out) For example: Since this is "Pizza" Picazzo I'd expect better. (2)

○ Story (the author quotes someone or something to make the review more story-like) For example: The cashier answered "oh we don't have any more water." wow. (3)

○ Connotation (the author quotes a word or phrase that has an additional meaning) For example: I had an issue with a major "spill" let's call it (vomit) in my back seat. (4)

○ Uncommon Use (the author quotes a word or phrase when it is used in an uncommon way) For example: I looked for many different places to "board" my dog. (5)

○ Author Speech (the author quotes something only they would say, or the author quotes something they would not normally say) For example: I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (6)

○ Proper Noun (The author puts quotes around a specific person, place, or thing) For example: I read the book titled "1000 Places to See or Do in the US." (7)

○ Other (Write your own category description) (8) _____

Please read the following review carefully and choose one category that best describes how the phrase "gotta get it ready & it's gotta be perfect" is being used.

This location is THE PLACE to take your kid's team for an end-of-season party. They accommodate large parties take reservations for large parties AND allow you to decorate your tables with your team theme!   Equally important: they LEAVE YOU ALONE during the process--- so if you are a frantic "gotta get it ready & it's gotta be perfect" last-minute-crafty-maven like myself THAT'S A GREAT THING. They play fantastic-classic rock & roll as in 50's & 60's & they have really I MEAN REALLY cleaned the joint

up. New upholstery fresh paint & all done WELL. Top that with 'cooked to order' burgers that are CONSISTENTLY out of this world & you too will shout WINNER!

- ○ Common Experience (It's assumed that the author and reader share a common experience, culture, or background.) For example: includes slang like "old school" or idioms like "one of a kind or descriptions like "college sized" refrigerator" (1)
- ○ Emphasis (the author wants the quoted word or phrase to stand out) For example: Since this is "Pizza" Picazzo I'd expect better. (2)
- ○ Story (the author quotes someone or something to make the review more story-like) For example: The cashier answered "oh we don't have any more water." wow. (3)
- ○ Connotation (the author quotes a word or phrase that has an additional meaning) For example: I had an issue with a major "spill" let's call it (vomit) in my back seat. (4)
- ○ Uncommon Use (the author quotes a word or phrase when it is used in an uncommon way) For example: I looked for many different places to "board" my dog. (5)
- ○ Author Speech (the author quotes something only they would say, or the author quotes something they would not normally say) For example: I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (6)
- ○ Proper Noun (The author puts quotes around a specific person, place, or thing) For example: I read the book titled "1000 Places to See or Do in the US." (7)
- ○ Other (Write your own category description) (8) _____

Please read the following review carefully and choose one category that best describes how the phrase "Chinatown" is being used.

First time here for us. Nestled in the heart of Sin City's "Chinatown;" this place is a gem of qualify Chinese food with an authentic atmosphere that keeps you feeling welcome even after you leave.   We had the "Dinner A" which includes mushroom eggdrop soup; a small appetizer medley of egg rolls wontons and pot stickers; chicken fried rice; beef choy mein; sweet and sour pork; and cashew chicken.

We also had hot tea and Thai iced coffee with boba! Literally everything was perfectly prepared and delicious! What do we recommend? Yes. Thank you Joyful House we'll be back!

- ❍ Common Experience (It's assumed that the author and reader share a common experience, culture, or background.) For example, includes slang like "old school" or idioms like "one of a kind or descriptions like "college sized" refrigerator" (1)
- ❍ Emphasis (the author wants the quoted word or phrase to stand out) For example: Since this is "Pizza" Picazzo I'd expect better. (2)
- ❍ Story (the author quotes someone or something to make the review more story-like) For example: The cashier answered "oh we don't have any more water." wow. (3)
- ❍ Connotation (the author quotes a word or phrase that has an additional meaning) For example: I had an issue with a major "spill" let's call it (vomit) in my back seat. (4)
- ❍ Uncommon Use (the author quotes a word or phrase when it is used in an uncommon way) For example: I looked for many different places to "board" my dog. (5)
- ❍ Author Speech (the author quotes something only they would say, or the author quotes something they would not normally say) For example: I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (6)
- ❍ Proper Noun (The author puts quotes around a specific person, place, or thing) For example: I read the book titled "1000 Places to See or Do in the US." (7)
- ❍ Other (Write your own category description) (8) _____

How could this survey be improved?

Appendix D – Survey: Version 1

Some of the category names in the surveys do not match the category names in the thesis. The names were simplified to sound less academic.

Common Experience = Collective Knowledge

Emphasis = Emphasis

Story = Narrative

Added Meaning = Non-literal

Uncommon Use = Non-standard

Reviewer Speech = Idiolect

Grammatical = Grammatical

Consent to be a Research Subject

Introduction

This research study is being conducted by Mary Wright, a linguistics master's student at Brigham Young University, to determine how quotations are used in online reviews. You were invited to participate because you are a Mechanical Turk user.

Procedures

If you agree to participate in this research study, the following will occur:

You will begin the survey after reading this consent form. By beginning the survey, you are agreeing to participate in this research. The survey is online and can be completed on a computer with internet access. You will read five Yelp reviews and answer questions about the quotation use in each review. The survey will take about 20 minutes to complete. There are three versions of this survey. You may complete each version once, if you desire. Risks/Discomforts

There are minimal risks for this study. However, you may become tired or bored while answering survey questions. It may take you some time to contemplate how the quotations are being used in each review. The researcher has made the survey as clear and concise as possible. Only five reviews are included in each survey version, so you will not become fatigued while reading and answering questions. Some of the reviews included are entertaining also, so reading them should not be too boring.

Benefits

There will be no direct benefits to you. It is hoped, however, that through your participation researchers may learn about how quotation marks are used in online communication. Quotations are typically used to quote other sources or people, but they can also be used to express tone, emotion, etc. This research can help linguists understand the communication strategies of online users with regards to quotations.

Confidentiality

The research data will be kept on a password protected computer and on a password protected Qualtrics account, and only the researcher will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept on the researcher's personal password protected computer. On the survey, you will be asked some basic demographic questions, and the researcher will not be able to identify you from this information.

Compensation

You will receive $1.00 per completed survey. You have the option of completing all three versions of the survey and receiving a total of $3.00. You may only complete each version once. Compensation will not be prorated. Those who do not wish to complete this research study will not be compensated. Participation Participation in this research study is voluntary. You have the right to stop taking the survey at any time or refuse to participate entirely. Questions about the Research If you have questions regarding this study, you may contact Mary Wright at mewright10@yahoo.com for further information. Questions about Your Rights as Research Participants If you have questions

regarding your rights as a research participant contact IRB Administrator at (801) 422-1461; A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu.

Statement of Consent I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study. By proceeding with this survey, I agree to participate in this research study.

This study is about how to categorize quotations in online reviews based on the reviewer's communication intentions. The researcher has so far identified 7 quotation categories. You will now read about each of the 7 categories along with some examples from real Yelp reviews.

Category 1 Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background.

Quotation Examples 1) slang like "old school"   2) stock phrases like "one of a kind" employee   3) I expect better than "slightly better than Outback"   4) I am not a "Vegas Person"

Category 2 Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out.

Quotation Examples 1) Since this is "Pizza" Picazzo I'd expect better   2) The orchestra was either "blaring" or too quiet.

Category 3 Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like.

Quotation Examples 1) The cashier answered "oh we don't have any more water." wow   2) Booked what the site called a "full size SUV (Chevy Tahoe or similar size)" online

Category 4 Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning.

Quotation Examples 1) I had an issue with a major "spill" let's call it (vomit) in my back seat.   2) However the "referral coordinator" Dawn and "aide" Mary are completely rude and

condescending. 3) Never again! I was told right away that my car "said" it required $235 worth of work. 4) the "faux" cobblestone streets do not mix well with 4 inch heels.

Category 5 Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way.

Quotation Examples 1) I looked for many different places to "board" my dog.   2) I seriously think these were the "skinniest" rolls (not sure how to describe it!) I've ever seen.

Category 6 Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say.

Quotation Examples 1) I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (only the reviewer says)   2) I had to ask several times what a "two-fer" was (not something the reviewer usually says)

 Category 7 Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter.

Quotation Examples 1) I read the book titled "1000 Places to See or Do in the US." (proper noun) 2) Lulu "our yorkie" was extremely comfortable with the Doc and staff (extra information) 3) Don't misunderstand "divey" I mean no attitudes no frills and tons of fun. (specific word)

Now that you have read the 7 quotation categories, you will read a series of 5 reviews and decide how a specific quote in each review should be categorized. Think about what the reviewer wants to communicate. Consider why the reviewer put quotes around a specific word or phrase.   If you are sure the quote does not fit into any of the categories, you can create your own category with a name and short description. Please only do this as a last resort.

Please read the following review carefully and choose one category that best describes how the phrase "help me out" is being used.

I brought two pairs of pants and two (cheap) blazers there to be altered. He was very nice. However he failed to mention to me that the cost of these 4 things- one of which only needed alternation tape- would come to $96!!!!!! When I came to pick up my items I was extremely embarrassed because I had to tell him I could not afford it. He then told me he could "help me out " and make it $80. I tried explaining to him that the two blazers were only $25 each to begin with. Overall this was an extremely humiliating experience and it all could have been solved if he were more professional and would have

told me that simple alterations would cost so much. I will definitely be searching for a new tailor in the future. Way too expensive if you ask me.

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the phrase "Black Hair Care Products" is being used.

Soooooo... I'm at "Bed Bath and Beyond" in Paradise Valley picking up items for a trip to Texas. I ask the associate who works there where I might find "Black Hair Care Products." She walks up and down five aisles and is unable to locate products for Black hair. She calls another associate who also walks up and down the same aisles and is also unable to find anything for my hair.   As embarrassed as I was the

expression on both of these women's faces was heartbreaking for me because they both had to acknowledge that there was nothing for my hair on the shelves.

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the word "jerks" is being used.

I realize you don't want to lose money who would...but I'm sure we can come to some arrangement since you quoted me the wrong prices after I gave you all the information you asked for. He was not interested kept his negative poor service attitude and went back into the shop to work on a bike part. I was shocked. My little boys were shocked! As we drove toward Curbside Cyclery and Javelina Cycles (trustworthy shops) to get the work done my boys both said they could not believe how unreasonable

they were at South Mountain Cycles and how impressed they were with me for talking nicely and trying to get it worked out into a win/win solution when the guys at SMC were such "jerks".

- ❍ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ❍ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ❍ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ❍ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ❍ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ❍ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ❍ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ❍ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the word "accidentally" is being used.

 beautiful clean and roomy (esp the bathroom - perfect for a bunch of girls to get ready for a night out on the town). sleeps 6+ on beds (two queens + pull-out sofa bed); plus the housekeeping is nice about bringing more towels/pillows/blankets/toiletries in the case that you *ahem* "accidentally" have more

than 2 people sleeping in the room. perfect location for partying - tao beach and tao nightclub right there!

- ❍ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ❍ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ❍ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ❍ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ❍ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ❍ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ❍ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ❍ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the phrase "sushi snob" is being used.

My best friend and I stayed at Bally's. One morning slightly hungover we needed to nourish ourselves. We wanted something healthy and fresh to counter act the horrible things we did to our liver the night before. As we explored and got distracted by the free bread samples at Sbarro we saw this sushi place. The staff was so friendly! I ordered the vegetarian roll (which was sufficient not wow) and my friend

ordered the spicy tuna (which she said was SPICY! and also very good quality. She works in a sushi restaurant and has now become a self proclaimed "sushi snob".) Our bill was also super cheap!

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

In many online reviews, the reviewer gives a product or company a star rating. The reviewer can usually assign between 1 and 5 stars to the product or company (1 star being the worst and 5 stars being the best). You will now decide how the 7 quotation categories relate to star rating.

For this question, try to predict a reviewer's evaluation. What star rating might a reviewer give a company if they used the following type of quotation in their review?

	1 Star (1)	2 Stars (2)	3 Stars (3)	4 Stars (4)	5 Stars (5)
Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (1)	○	○	○	○	○
Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (2)	○	○	○	○	○
Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (3)	○	○	○	○	○
Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (4)	○	○	○	○	○
Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (5)	○	○	○	○	○
Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (6)	○	○	○	○	○

Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (7)	○	○	○	○	○

You're almost done! Now you will answer some basic demographic questions.

What is your age?

○   18-24 (1)
○   25-34 (2)
○   35-44 (3)
○   45-54 (4)
○   55-64 (5)
○   65-74 (6)
○   75 or older (7)

What is your gender?

○   Male (1)
○   Female (2)

What is your highest level of education?

○   high school diploma or equivalent (1)
○   some college (2)
○   trade/technical/vocational training (3)
○   Bachelor's degree (4)
○   Master's degree (5)
○   Doctorate degree (6)

What is your ethnicity?

○ Caucasian (1)
○ Hispanic or Latino/Latina (2)
○ Black or African American (3)
○ Native American or American Indian (4)
○ Asian/Pacific Islander (5)
○ Other (6) _____

Are you a native English speaker?

○ Yes (1)
○ No (2)

Were you primarily raised in the United States?

○ Yes (1)
○ No (2)

Please answer the questions below.

	Never (1)	Less than Once a Month (2)	Once a Month (3)	2-3 Times a Month (4)	Once a Week (5)	2-3 Times a Week (6)	Daily (7)
How often do you read online reviews? (1)	○	○	○	○	○	○	○
How often do you write online reviews? (2)	○	○	○	○	○	○	○
How often do you read Yelp reviews? (3)	○	○	○	○	○	○	○
How often do you write Yelp reviews? (4)	○	○	○	○	○	○	○

Appendix E – Survey: Version 2

Consent to be a Research Subject

Introduction

This research study is being conducted by Mary Wright, a linguistics master's student at Brigham Young University, to determine how quotations are used in online reviews. You were invited to participate because you are a Mechanical Turk user.

Procedures

If you agree to participate in this research study, the following will occur: You will begin the survey after reading this consent form. By beginning the survey, you are agreeing to participate in this research. The survey is online and can be completed on a computer with internet access. You will read five Yelp reviews and answer questions about the quotation use in each review. The survey will take about 20 minutes to complete.   There are three versions of this survey. You may complete each version once, if you desire.

Risks/Discomforts

There are minimal risks for this study. However, you may become tired or bored while answering survey questions. It may take you some time to contemplate how the quotations are being used in each review.   The researcher has made the survey as clear and concise as possible. Only five reviews are included in each survey version, so you will not become fatigued while reading and answering questions. Some of the reviews included are entertaining also, so reading them should not be too boring.

Benefits

There will be no direct benefits to you. It is hoped, however, that through your participation researchers may learn about how quotation marks are used in online communication. Quotations are typically used to quote other sources or people, but they can also be used to express tone, emotion, etc. This research can help linguists understand the communication strategies of online users with regards to quotations.

Confidentiality

The research data will be kept on a password protected computer and on a password protected Qualtrics account, and only the researcher will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept on the researcher's personal

password protected computer. On the survey, you will be asked some basic demographic questions, and the researcher will not be able to identify you from this information.

Compensation

You will receive $1.00 per completed survey. You have the option of completing all three versions of the survey and receiving a total of $3.00. You may only complete each version once. Compensation will not be prorated. Those who do not wish to complete this research study will not be compensated.

Participation

Participation in this research study is voluntary. You have the right to stop taking the survey at any time or refuse to participate entirely.   Questions about the Research If you have questions regarding this study, you may contact Mary Wright at mewright10@yahoo.com for further information.   Questions about Your Rights as Research Participants If you have questions regarding your rights as a research participant contact IRB Administrator at (801) 422-1461; A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu.

Statement of Consent

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study. By proceeding with this survey, I agree to participate in this research study.

This study is about how to categorize quotations in online reviews based on the reviewer's communication intentions. The researcher has so far identified 7 quotation categories. You will now read about each of the 7 categories along with some examples from real Yelp reviews.

Category 1 Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background.

Quotation Examples 1) slang like "old school"   2) stock phrases like "one of a kind" employee   3) I expect better than "slightly better than Outback"   4) I am not a "Vegas Person"

Category 2 Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out.

Quotation Examples 1) Since this is "Pizza" Picazzo I'd expect better   2) The orchestra was either "blaring" or too quiet.

Category 3 Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like.

Quotation Examples 1) The cashier answered "oh we don't have any more water." wow   2) Booked what the site called a "full size SUV (Chevy Tahoe or similar size)" online

Category 4 Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning.

Quotation Examples 1) I had an issue with a major "spill" let's call it (vomit) in my back seat.   2) However the "referral coordinator" Dawn and "aide" Mary are completely rude and

condescending. 3) Never again! I was told right away that my car "said" it required $235 worth of work. 4) the "faux" cobblestone streets do not mix well with 4 inch heels.

Category 5 Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way.

Quotation Examples 1) I looked for many different places to "board" my dog.   2) I seriously think these were the "skinniest" rolls (not sure how to describe it!) I've ever seen.

Category 6 Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say.

Quotation Examples 1) I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (only the reviewer says)   2) I had to ask several times what a "two-fer" was (not something the reviewer usually says)

Category 7 Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter.

Quotation Examples 1) I read the book titled "1000 Places to See or Do in the US." (proper noun) 2) Lulu "our yorkie" was extremely comfortable with the Doc and staff (extra information) 3) Don't misunderstand "divey" I mean no attitudes no frills and tons of fun. (specific word)

Now that you have read the 7 quotation categories, you will read a series of 5 reviews and decide how a specific quote in each review should be categorized. Think about what the reviewer wants to communicate. Consider why the reviewer put quotes around a specific word or phrase.   If you are sure the quote does not fit into any of the categories, you can create your own category with a name and short description. Please only do this as a last resort.

Please read the following review carefully and choose one category that best describes how the phrase "Ladies Nights" is being used.

This place is great. The staff is super friendly and helpful. We recently attended one of their "Ladies Nights" in which they gave us instruction on cleaning a firearm. The instructor was awesome. The only reason I am not giving them a 5 star is because they were SO incredibly busy the place was packed and

we did have to wait a bit. I was a little disappointed I didn't get more range time but other than that great place! I will definitely be back.

- ◯ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ◯ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ◯ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ◯ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ◯ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ◯ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ◯ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ◯ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the phrase "edge of Phoenix" is being used.

Finding good authentic Chinese food is a bit of a challenge nowaday. Wong's Restaurant has been around for years and I grew up in the same little "edge of Phoenix" neighborhood that Wong's is located

in. From the outside Wong's looks small but on the inside it's huge and "wedding reception"ish. I've been going to Wong's for a long time and have yet to find a better Sesame Chicken.

○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)

○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)

○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)

○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)

○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)

○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)

○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)

○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the phrase "Ew I'm underneath a casino. What smells like swamp water?" is being used.

After you've been to Las Vegas for an hour the glamour wears off. The arcade below the Excalibur? It makes Chuck E. Cheese look rather classy and educational. There's a metal claw over stuffed animals where you put your quarter in and the claw doesn't even move. It's broken if not enough to have not

taken your money. And you're staring at this metal claw thinking "Ew I'm underneath a casino. What smells like swamp water?" I like Treasure Island.

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the phrase "unicorn tears" is being used.

The owner Baja and staff are super friendly and very personable. This is my one of my favorite breakfast spots in town due to their specialty pancake batters. The batters come in hot cakes as well as waffle form. My absolute favorite is their pumpkin pancake with cream cheese topping....it's a foodgasm! Also

their marinated steak tips are apparently marinated in "unicorn tears" lol. The corn beef hash is homemade and savory. It's a great place to go with pals and or family.

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the word "basement" is being used.

One of the more bizarre phenomenons I have seen... why is the place to be after work in Scottsdale a restaurant in a mall "basement"? Is it the choice view of the parking garage entrance? The gentle

breeze laden with the scent of car exhaust ? I swear it's true. Flocks of shiny happy people packed like sardines onto the postage stamp sized patio.

- ⭕ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ⭕ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ⭕ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ⭕ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ⭕ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ⭕ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ⭕ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ⭕ Other (Name your own category, and write a description.) (8) _____

In many online reviews, the reviewer gives a product or company a star rating. The reviewer can usually assign between 1 and 5 stars to the product or company (1 star being the worst and 5 stars being the best). You will now decide how the 7 quotation categories relate to star rating.

For this question, try to predict a reviewer's evaluation. What star rating might a reviewer give a company if they used the following type of quotation in their review?

	1 Star (1)	2 Stars (2)	3 Stars (3)	4 Stars (4)	5 Stars (5)
Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (1)	○	○	○	○	○
Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (2)	○	○	○	○	○
Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (3)	○	○	○	○	○
Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (4)	○	○	○	○	○
Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (5)	○	○	○	○	○
Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (6)	○	○	○	○	○

Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (7)	❍	❍	❍	❍	❍

You're almost done! Now you will answer some basic demographic questions.

What is your age?

- ❍ 18-24 (1)
- ❍ 25-34 (2)
- ❍ 35-44 (3)
- ❍ 45-54 (4)
- ❍ 55-64 (5)
- ❍ 65-74 (6)
- ❍ 75 or older (7)

What is your gender?

- ❍ Male (1)
- ❍ Female (2)

What is your highest level of education?

- ❍ high school diploma or equivalent (1)
- ❍ some college (2)
- ❍ trade/technical/vocational training (3)
- ❍ Bachelor's degree (4)
- ❍ Master's degree (5)
- ❍ Doctorate degree (6)

What is your ethnicity?

- ○ Caucasian (1)
- ○ Hispanic or Latino/Latina (2)
- ○ Black or African American (3)
- ○ Native American or American Indian (4)
- ○ Asian/Pacific Islander (5)
- ○ Other (6) _____

Are you a native English speaker?

- ○ Yes (1)
- ○ No (2)

Were you primarily raised in the United States?

- ○ Yes (1)
- ○ No (2)

Please answer the questions below.

	Never (1)	Less than Once a Month (2)	Once a Month (3)	2-3 Times a Month (4)	Once a Week (5)	2-3 Times a Week (6)	Daily (7)
How often do you read online reviews? (1)	○	○	○	○	○	○	○
How often do you write online reviews? (2)	○	○	○	○	○	○	○
How often do you read Yelp reviews? (3)	○	○	○	○	○	○	○
How often do you write Yelp reviews? (4)	○	○	○	○	○	○	○

97

Appendix F – Survey: Version 3

Consent to be a Research Subject

Introduction

This research study is being conducted by Mary Wright, a linguistics master's student at Brigham Young University, to determine how quotations are used in online reviews. You were invited to participate because you are a Mechanical Turk user.

Procedures

If you agree to participate in this research study, the following will occur: You will begin the survey after reading this consent form. By beginning the survey, you are agreeing to participate in this research. The survey is online and can be completed on a computer with internet access. You will read five Yelp reviews and answer questions about the quotation use in each review. The survey will take about 20 minutes to complete.   There are three versions of this survey. You may complete each version once, if you desire.

Risks/Discomforts

There are minimal risks for this study. However, you may become tired or bored while answering survey questions. It may take you some time to contemplate how the quotations are being used in each review.   The researcher has made the survey as clear and concise as possible. Only five reviews are included in each survey version, so you will not become fatigued while reading and answering questions. Some of the reviews included are entertaining also, so reading them should not be too boring.

Benefits

There will be no direct benefits to you. It is hoped, however, that through your participation researchers may learn about how quotation marks are used in online communication. Quotations are typically used to quote other sources or people, but they can also be used to express tone, emotion, etc. This research can help linguists understand the communication strategies of online users with regards to quotations.

Confidentiality

The research data will be kept on a password protected computer and on a password protected Qualtrics account, and only the researcher will have access to the data. At the conclusion of the study, all identifying information will be removed and the data will be kept on the researcher's personal

password protected computer. On the survey, you will be asked some basic demographic questions, and the researcher will not be able to identify you from this information.

Compensation

You will receive $1.00 per completed survey. You have the option of completing all three versions of the survey and receiving a total of $3.00. You may only complete each version once. Compensation will not be prorated. Those who do not wish to complete this research study will not be compensated.

Participation

Participation in this research study is voluntary. You have the right to stop taking the survey at any time or refuse to participate entirely.   Questions about the Research If you have questions regarding this study, you may contact Mary Wright at mewright10@yahoo.com for further information.   Questions about Your Rights as Research Participants If you have questions regarding your rights as a research participant contact IRB Administrator at (801) 422-1461; A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu.

Statement of Consent

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study. By proceeding with this survey, I agree to participate in this research study.

This study is about how to categorize quotations in online reviews based on the reviewer's communication intentions. The researcher has so far identified 7 quotation categories. You will now read about each of the 7 categories along with some examples from real Yelp reviews.

Category 1 Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background.

Quotation Examples 1) slang like "old school"   2) stock phrases like "one of a kind" employee   3) I expect better than "slightly better than Outback"   4) I am not a "Vegas Person"

Category 2 Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out.

Quotation Examples 1) Since this is "Pizza" Picazzo I'd expect better   2) The orchestra was either "blaring" or too quiet.

Category 3 Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like.

Quotation Examples 1) The cashier answered "oh we don't have any more water." wow   2) Booked what the site called a "full size SUV (Chevy Tahoe or similar size)" online

Category 4 Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning.

Quotation Examples 1) I had an issue with a major "spill" let's call it (vomit) in my back seat.   2) However the "referral coordinator" Dawn and "aide" Mary are completely rude and

condescending. 3) Never again! I was told right away that my car "said" it required $235 worth of work. 4) the "faux" cobblestone streets do not mix well with 4 inch heels.

Category 5 Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way.

Quotation Examples 1) I looked for many different places to "board" my dog.   2) I seriously think these were the "skinniest" rolls (not sure how to describe it!) I've ever seen.

Category 6 Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say.

Quotation Examples 1) I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's." (only the reviewer says)   2) I had to ask several times what a "two-fer" was (not something the reviewer usually says)

 Category 7 Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter.

Quotation Examples 1) I read the book titled "1000 Places to See or Do in the US." (proper noun) 2) Lulu "our yorkie" was extremely comfortable with the Doc and staff (extra information) 3) Don't misunderstand "divey" I mean no attitudes no frills and tons of fun. (specific word)

Now that you have read the 7 quotation categories, you will read a series of 5 reviews and decide how a specific quote in each review should be categorized. Think about what the reviewer wants to communicate. Consider why the reviewer put quotes around a specific word or phrase.   If you are sure the quote does not fit into any of the categories, you can create your own category with a name and short description. Please only do this as a last resort.

Please read the following review carefully and choose one category that best describes how the word "ick" is being used.

Although there is lots of fun to be had here- the atmosphere of the show does not make up for the lack of cleanliness and overall "ICK" factor. I love me some dive bars but this one was a little too

subterranean. That and please ladies- when you are lip syncing at least know HALF of the words to the song. Just moving your mouth and jaw like an automaton is NOT fun or attractive.

- ❍ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ❍ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ❍ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ❍ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ❍ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ❍ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ❍ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ❍ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the phrase "I hate my life! Its so hard!" is being used.

I hate The Hub. It is just obnoxious. It all like "I wanna look like I have nothing but wanna pay through the nose for it!". I mean the brand Deisel alone makes me irritated. It so "I hate my life! Its so hard!".

Also the staff prertend to know the product but if you ask enough good questions you'll find they know nothing more than you could have deduced. Eww. Yuck. gag.

○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)

○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)

○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)

○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)

○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)

○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)

○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)

○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the word "server" is being used.

Waited for way to long to be acknowledged. Then upon being seated was told by the "server" that there were five tables ahead of me and it would be a while. No drink... No hope. No food. No tip. Forget this joint!

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the word "hair" is being used.

Upon entering the salon I was welcomed graciously. Shaunte was there to greet me ready to consult. She immediately understood what I desired and I scheduled an appointment for a cut and color

the next day! I'm thrilled to report that I have found a "hair" home! I absolutely love everything she did...even how she styled it!

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

Please read the following review and choose one category that best describes how the phrase "beer horizons" is being used.

 The best beer selection anywhere. This is the place I originally started learning about new beers from around the world. Before YH was even in Vegas I was frequenting them in LA. It helped me expand my "beer horizons" and now I can't even go back to bud light or anything like that.

- ○ Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (slang like "old school" or I am not a "Vegas person") (1)
- ○ Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (The orchestra was either "blaring" or too quiet.) (2)
- ○ Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (The cashier answered "oh we don't have any more water." wow) (3)
- ○ Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (I had an issue with a major "spill" let's call it (vomit) in my back seat.) (4)
- ○ Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (I looked for many different places to "board" my dog.) (5)
- ○ Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (I went with the old reliable Denny's at 7th Street & Camelback aka "the gay Denny's" or "Jenny's.") (6)
- ○ Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (I read the book titled "1000 Places to See or Do in the US.") (7)
- ○ Other (Name your own category, and write a description.) (8) _____

In many online reviews, the reviewer gives a product or company a star rating. The reviewer can usually assign between 1 and 5 stars to the product or company (1 star being the worst and 5 stars being the best). You will now decide how the 7 quotation categories relate to star rating.

For this question, try to predict a reviewer's evaluation. What star rating might a reviewer give a company if they used the following type of quotation in their review?

	1 Star (1)	2 Stars (2)	3 Stars (3)	4 Stars (4)	5 Stars (5)
Common Experience: The reviewer assumes he/she and the reader share a common experience, culture, or background. (1)	○	○	○	○	○
Emphasis: The reviewer wants to stress the quoted word or phrase and make it stand out. (2)	○	○	○	○	○
Story: The reviewer quotes him/herself, another person, or something else to make the review more story-like. (3)	○	○	○	○	○
Added Meaning: The reviewer quotes a word or phrase used in a non-literal way or that has another meaning. (4)	○	○	○	○	○
Uncommon Use: The reviewer quotes a word or phrase when it is used in an uncommon way. (5)	○	○	○	○	○
Reviewer Speech: The reviewer quotes something only he/she would say, or the reviewer quotes something he/she would not normally say. (6)	○	○	○	○	○

Grammatical: The reviewer uses quotes to mark a proper noun, extra information, or to refer to a specific word or letter. (7)	❍	❍	❍	❍	❍

You're almost done! Now you will answer some basic demographic questions.

What is your age?

❍ 18-24 (1)
❍ 25-34 (2)
❍ 35-44 (3)
❍ 45-54 (4)
❍ 55-64 (5)
❍ 65-74 (6)
❍ 75 or older (7)

What is your gender?

❍ Male (1)
❍ Female (2)

What is your highest level of education?

❍ high school diploma or equivalent (1)
❍ some college (2)
❍ trade/technical/vocational training (3)
❍ Bachelor's degree (4)
❍ Master's degree (5)
❍ Doctorate degree (6)

What is your ethnicity?

❍ Caucasian (1)
❍ Hispanic or Latino/Latina (2)
❍ Black or African American (3)
❍ Native American or American Indian (4)
❍ Asian/Pacific Islander (5)
❍ Other (6) _____

Are you a native English speaker?

❍ Yes (1)
❍ No (2)

Were you primarily raised in the United States?

❍ Yes (1)
❍ No (2)

Please answer the questions below.

	Never (1)	Less than Once a Month (2)	Once a Month (3)	2-3 Times a Month (4)	Once a Week (5)	2-3 Times a Week (6)	Daily (7)
How often do you read online reviews? (1)	❍	❍	❍	❍	❍	❍	❍
How often do you write online reviews? (2)	❍	❍	❍	❍	❍	❍	❍
How often do you read Yelp reviews? (3)	❍	❍	❍	❍	❍	❍	❍
How often do you write Yelp reviews? (4)	❍	❍	❍	❍	❍	❍	❍

108