fi®st
m¤ñd@¥
PEER-REVIEWED JOURNAL ON THE INTERNET

HOME   ABOUT   LOGIN   REGISTER   SEARCH   CURRENT
ARCHIVES   ANNOUNCEMENTS   SUBMISSIONS

fi®st m¤ñd@¥
PEER-REVIEWED JOURNAL ON THE INTERNET

Narrative framing of consumer sentiment in online restaurant reviews
by Dan Jurafsky, Victor Chahuneau, Bryan R. Routledge, and Noah A. Smith

Abstract

The vast increase in online expressions of consumer sentiment offers a powerful new tool for studying consumer attitudes. To explore the narratives that consumers use to frame positive and negative sentiment online, we computationally investigate linguistic structure in 900,000 online restaurant reviews. Negative reviews, especially in expensive restaurants, were more likely to use features previously associated with narratives of *trauma*: negative emotional vocabulary, a focus on the past actions of third person actors such as waiters, and increased use of references to "we" and "us", suggesting that negative reviews function as a means of coping with service–related trauma. Positive reviews also employed framings contextualized by expense: inexpensive restaurant reviews use the language of *addiction* to frame the reviewer as craving fatty or starchy foods. Positive reviews of expensive restaurants were long narratives using long words emphasizing the reviewer's linguistic capital and also focusing on *sensory pleasure*. Our results demonstrate that portraying the self, whether as well–educated, as a victim, or even as addicted to chocolate, is a key function of reviews and suggests the important role of online reviews in exploring social psychological variables.

Contents

---

1. Introduction

Consumer opinions permeate the Web. A wide variety of methods from natural language processing have been employed to process these opinions and learn which products consumers like or don't like, and also discover the particular aspects of each product that people care about (Archak, *et al.*, 2011; Blair–Goldensohn, *et al.*, 2008; Brody and Elhadad, 2010; Popescu and Etzioni, 2005; Hu and Liu, 2004; Snyder and Barzilay, 2007; Titov and McDonald, 2008a, 2008b; Jo and Oh, 2011; McAuley, *et al.*, 2012; Reschke, *et al.*, 2013).

We propose to extend this work by exploring the linguistic expression of sentiment in richer detail. Previous work has focused on tasks like predicting sentiment ranking (a value from 1–5) or extracting aspects (like learning which words focus on taste versus color versus texture in a beer review, or service versus food in a restaurant review). By contrast our goal is to use computational tools to understand the narrative structures and framings that reviewers use. How do reviewers express fine–grained
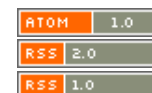
differences in sentiment beyond just positive or negative? What narratives are used to express different levels of sentiment? What are the psychological functions of these narratives?

Following the long line of previous research, we examine consumer reviews, and in particular reviews of restaurants. Restaurant reviews offer rich metadata in addition to extensive text, including the numeric rating the customer assigns, economic variables like the price level of the restaurant being reviewed, and control variables like the type of food and location. The combination of these three aspects — language, rating sentiment, and product price — allows us to address a number of open questions. What are the particular narratives used in negative versus positive reviews, or in particularly strong reviews of either valence? How do reviewers use reviews to express a particular aspect of their own psychological or social characteristics?

To answer these questions we draw on the long literature on the computational extraction of social meaning from text. Many studies in both the computational and the social psychological literature have explored extraction from texts of different kinds of social meaning, including the studies of sentiment mentioned above, as well as such factors as scores on big–five personality instruments (Pennebaker and King, 1999; Mairesse, *et al.*, 2007), perceptions of friendliness and flirtation (Ranganath, *et al.*, 2013), romantic interest (McFarland, *et al.*, 2013), deception (Larcker and Zakolyukina, 2012), and ideological positioning (Sim, *et al.*, 2013). These studies have relied on a number of computational techniques, most commonly use of the many lexicons used to model the linguistic expression of sentiment and opinions (Riloff and Wiebe, 2003; Hu and Liu, 2004; Wilson, *et al.*, 2005; Baccianella, *et al.*, 2010; Stone, *et al.*, 1966; Pennebaker, *et al.*, 1997; Reschke, *et al.*, 2013). Our plan is to draw on these lexicons and computational models of other linguistic features to understand the questions posed above.

Our hypothesis is that the function of reviews is not just to evaluate a restaurant by assigning a raw rating or even to summarize aspects of a restaurant. Instead we propose that reviews are fundamentally a kind of social discourse, in which reviewers employ narratives to portray their own social or psychological characteristics, role or stance. As such we expect that the different kinds of narratives we see in different kinds of reviews will result from the different kinds of social goals and relations that reviewers have when writing good versus bad reviews, and about expensive versus inexpensive restaurants. We test this hypothesis by using automatic methods to extract information about these social psychological functions from large corpora of online reviews.

━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

2. Data and methods

We chose a dataset of reviews large enough to investigate the full range of restaurant markets, from fast food to luxury restaurants, allowing the investigation of linguistic structures in the reviews while controlling for confounding factors like geographical location, type of food served at the restaurant, length of the review, and so on. We built our dataset by extending the datasets used by Chahuneau, *et al.* (2011), including reviews from the Web site yelp.com (http://www.yelp.com/) from 2006–2011 for a set of restaurants in seven cities: Boston, Chicago, Los Angeles, New York, Philadelphia, San Francisco, and Washington D.C. They randomly divided the restaurants into 80 percent for training and 20 percent reserved for evaluation and testing. All the analyses in this paper are performed on their training dataset. From this data, we used only restaurants that were characterized on Yelp as restaurants and bars; thus all delis, groceries, and caterers were removed from the dataset. In addition to reviews, we used two variables from their Yelp dataset: the city, and the price range, a variable on a four–level scale from $ to $$$$. We also coded two more control factors that might interact with our hypotheses: restaurant category and whether the restaurant was a chain.

The restaurant category consisted of a label from a set of 32 types of restaurants. These were constructed by hand–clustering the complete set of restaurant categories from the Yelp restaurant category into the following 32 categories, based on choosing restaurants with similar cuisines and similar price ranges:

> Pizza, Chinese, Italian, Steakhouses, American (new), Japanese, Mexican, French, American (traditional), Sandwiches, Cafes, Fast food, Thai, Indian, Asian, Diners, Seafood, Middle Eastern, Latin American, Bars, Bakeries, Spanish, Korean, Mediterranean, Barbeque, Other European, Vegetarian, Ethiopian, Soul food, Southern and Cajun, Greek, Asian fusion

Each restaurant was assigned to exactly one of these categories; restaurants listed in Yelp with multiple classes were assigned to whichever of those classes occurred most frequently in the entire dataset.

The resulting dataset consists of 887,658 reviews from 6,548 restaurants.

To detect linguistic strategies corresponding to particular hypotheses, we used standard methods from computational linguistics and sentiment analysis that measure characteristics of words and sentences. These include shallow properties like review length, as well as the number of times words appear from specialized *lexicons*, lists of words and phrases that were designed to operationalize each

Victor Chahuneau
Carnegie Mellon University
United States

Graduate Student, Language Technologies Institute

Bryan R. Routledge
Carnegie Mellon University
United States

Associate Professor, Tepper School of Business,

Noah A. Smith
Carnegie Mellon University
United States

Associate Professor, Language Technologies Institute

strategy. Lexicons were mainly drawn from the previous literature and also from an initial investigation of the menus.

The initial investigation employed the "log odds ratio informative Dirichlet prior" method of Monroe, *et al.* (2008), to find words that are statistically overrepresented in a particular category of review compared to another (such as those with one star versus five stars, or reviewing cheap versus expensive restaurants). The method estimates the difference between the frequency of word *w* in two corpora *i* and *j* via the log–odds–ratio for *w*, $\delta_w^{(i-j)}$ which is estimated as:

$$\delta_w^{(i-j)} = log\left(\frac{y_w^i + \alpha_w}{n^i + \alpha_0 - (y_w^i + \alpha_w)}\right) - log\left(\frac{}{n}\right.$$

(where n^i is the size of corpus *i*, n^j is the size of corpus *j*, y_w^i is the count of word *w* in corpus *i*, y_w^j is the count of word *w* in corpus *j*, α_0 is the size of the background corpus, and α_w is the count of word *w* in the background corpus.) In addition, Monroe, *et al.* (2008) make use of an estimate for the variance of the log–odds–ratio:

$$\sigma^2\left(\delta_w^{(i-j)}\right) \approx \frac{1}{y_w^i + \alpha_w} + \frac{1}{y_w^j + \alpha_w}$$

The final statistic for a word is then the z–score of its log–odds–ratio:

$$\frac{\delta_w^{(i-j)}}{\sqrt{\sigma^2\left(\delta_w^{(i-j)}\right)}}$$

The Monroe, *et al.* (2008) method thus modifies the commonly used log–odds ratio in two ways: it uses the z–scores of the log–odds–ratio, which controls for the amount of variance in a word's frequency, and it uses counts from a background corpus to provide a prior count for words, essentially shrinking the counts toward to the prior frequency in a large background corpus. These features enable differences even in very frequent words to be detected; previous linguistic methods used to discover word associations (mutual information (Church and Hanks, 1990), log likelihood ratio (Dunning, 1993), t–test (Manning and Schütze, 1999), and chi–square (Yang and Pederson, 1997)) have all had problems with frequent words. Because function words like pronouns and auxiliary verbs are both extremely frequent and have been shown to be important cues to social and narrative meaning, this is a major limitation of these methods, and one of the reasons we chose the Monroe, *et al.* (2008) method.

Our second method is the use of ordered logistic regression, predicting a review's rating score (a ranked category from one to five) or the restaurant's price (a ranked category of $, $$, $$$, $$$$). The regression allows us to test for the association of linguistic variables with ratings or price after controlling for factors like the type of food at the restaurant or geographical location. To operationalize linguistic hypotheses, we employed lexicons, which are groups of words or phrases that express particular hypotheses. Mainly these lexicons are drawn from the large variety of sentiment and social lexicons, including LIWC (Pennebaker, *et al.*, 1997; Pennebaker, *et al.*, 2007), the General Inquirer (Stone, *et al.*, 1966), and others developed for specific hypotheses (Ranganath, *et al.*, 2013; McFarland, *et al.*, 2013), or specifically for restaurants (Reschke, *et al.*, 2013).

■ ————————————————————————————————

3. Overall sentiment skew

Reviews generally skew toward the positive. Figure 1 shows the distribution of reviewer star values over the reviews; the mean and median review in our data is four rather than the three that would be expected if all the star values were equally likely.

Restaurant Ratings

Figure 1: Frequency of reviews at each star level showing positive sk

This strong positive skew in the star ratings is consistent with previous work analyzing reviews of movies, hotels, restaurants, and consumer products (Potts, 2011).

To check whether this rating skew is matched by a bias toward positive vocabulary, we investigated two lexicons that provide sets of psychological and social categories and lists of words with meanings in the categories, including categories representing positive and negative sentiment. (We chose these two from a larger set of sentiment dictionaries: Riloff and Wiebe, 2003; Hu and Liu, 2004; Wilson, *et al.*, 2005; Baccianella, *et al.*, 2010).

The General Inquirer (Stone, *et al.*, 1966) includes lists of 1,915 positive words and 2,291 negative words. The 70 categories in LIWC (Pennebaker, *et al.*, 1997) include 276 positive emotional word stems (love, nice, sweet) and 499 negative emotional word stems (bad, weird, hate, problem). These stems correspond to a much larger number of words, since stems like *lucki* in the positive emotional dictionary correspond to words like *luckily, luckiness, luckier*, and *luckiest*. For each of these two sentiment lexicons we computed the frequency of each word in the reviews, and then examined the total token frequency of the top 500 most frequent words in each category.

	Table 1: Ratios of total frequencies of the most frequent 500 positive to most frequent 500 negative vocabulary words in two sentiment lexicons: the General Inquirer (Stone, *et al.*, 1966), and LIWC (Pennebaker, *et al.*, 1997).	
Lexicon	**Positive–negative ratio in restaurant reviews**	**Positive–negative ratio in Google Books**
General Inquirer	1.8	1.5
LIWC	2.7	1.8

Table 1 shows a positive–to–negative ratio of between 1.8 and 2.7 to 1 in the two general–purposes lexicons.

Positive skew is consistent with a long tradition of results showing a bias toward positivity in language (the "linguistic positivity bias") in English and other languages, and in non–linguistic cognitive processes (the "Pollyanna hypothesis" of Boucher and Osgood, 1969). Positive words are more frequent in the vocabulary (Dodds and Danforth, 2009; Rozin, *et al.*, 2010; Augustine, *et al.*, 2011; Dodds, *et al.*, 2011; Zajonc, 1968), and are linguistically unmarked [1].

We hypothesized that reviews would have an even stronger bias, however, than general English. The final column in Table 1 shows the positive skew in the Google Books corpus (Michel, *et al.*, 2011), again computed by counting the frequency of each lexicon word in Google Books, examining the total token frequency of the top 500 most frequent words in each category, and computing the ratio. The bias towards positive sentiment in restaurant reviews is thus exaggerated compared to the positive biases seen in perhaps less purely opinionated genres like the Google books corpus. In the following sections we explore positive reviews and negative reviews separately in order to understand what narratives and framings accompany this exaggerated evaluative language.

■ ─────────────────────────────

4. Narrative framing in strongly negative reviews: The role of trauma

What narratives and framing accompany particularly negative reviews? One hypothesis might be that there is no characteristic framing, that negative reviews merely consist of descriptions of food with negative evaluative vocabulary. To determine if this is true we began by exploring vocabulary that was strongly associated with the lowest reviews, those that assigned only a single star.

For each word in the review corpus we compute its frequency in one–star reviews, in five–star reviews, and in the entire review corpus. We then use the log–odds–ratio informative Dirichlet prior method (Monroe, *et al.*, 2008) described earlier to find words that are more strongly associated with one–star than five–star reviews, using the distribution over frequency in the entire review corpus as the Dirichlet prior that the method requires. We then sorted all words in the corpus by their log–odds association score, and selected the 50 words most associated with one–star reviews. Table 2 shows that the top 50 words associated with a one–star review fall into eight classes. The classes are shown in the table, ordered by the average association score of the class.

Table 2: Top 50 words associated with one–star reviews by the Monroe, *et al.* (2008) method.	
Linguistic class	**Words in class**
Negative sentiment	worst, rude, terrible, horrible, bad, awful, disgusting, bland, tasteless, gross, mediocre, overpriced, worse, poor
Linguistic negation	no, not
First person plural pronouns	we, us, our
Third person pronouns	she, he, her, him
Past tense verbs	was, were, asked, told, said, did, charged, waited, left, took
Narrative sequencers	after, then
Common nouns	manager, waitress, waiter, customer, customers, attitude, waste, poisoning, money, bill, minutes
Irrealis modals	would, should
Infinitives and complementizers	to, that

The largest class in Table 2, and the one most associated with one–star reviews, consists of negative evaluative descriptors (*worst, bad, terrible*). Together with the use of linguistic negation (*no, not*), negative evaluative descriptors are characteristic of all negative sentiment genres and hence certainly to be expected from these negative reviews [2].

We also see a group of features related to narrative discourse. Biber (1988) and Biber (1995) used factor–analytic methods to analyze the linguistic features of different linguistic genres. They found a stable set of dimensions that occurred in a variety of studies in many languages, including dimensions indicative of narrative genres, informative discourse, persuasive language, and others. They extracted a number of linguistic features and assigned factor loadings to each. The most significant features associated with narrative text are past tense verbs (.90), third person pronouns (.73), perfect aspect verbs (.48), and speech act verbs (verbs of speaking like *say* or *tell*) (.43). As is clear from Table 2, each of Biber's linguistic features is also disproportionately represented in one–star reviews in our data, as are related narrative features like the narrative sequencers "then" and "after".

The combination of these two categories suggests that one–star reviews are narratives of negative emotion, stories about something bad that happened involving what other people said and did.

Who are these people in the narrative being referred to by the third–person pronouns? The following list gives the common nouns with the highest log–odds–ratio association with one–star reviews; all are references to service personnel and service failings:

> manager, customer, minutes, money, waitress, waiter, bill, attitude, management, business, apology, mistake, table, charge, order, hostess, tip

Finally, as shown in Table 2, one–star reviews also have a marked increase in the use of first person plural (we, us, our), in sentences like the following:

> ... **we** were ignored until **we** flagged down one waiter to go get **our** waitress ...

> ... **we** were both so furious **we** refused to finish the food on principle ...

This exact constellation of features (negatively emotional past tense narratives about other people, with an associated increase in the first person plural) has been associated in a number of previous studies with a particular genre: people writing after experiencing trauma. According to the standard social stage model of coping (Pennebaker and Harber, 1993), shortly after a disaster or tragedy people experience emotional upheavals and obsessive thoughts and feelings. In this phase they share these thoughts and feelings with others, including strangers, and the phase is marked by expressions of collectively shared grief, in which people seem to emphasize their belonging to groups, using the words we or us with high frequency, as a sign of solidarity and other–comforting and a way of achieving "collective closure". Pennebaker and his colleagues have tested the model by showing these linguistic tendencies in a number of domains. Stone and Pennebaker (2002) found that fans writing about the death of Princess Diana on Internet chat rooms wrote narratives using negative emotional words and more past tense and were "more collective in their orientation", using more first person plural pronouns (we, us, our) and fewer first person singular pronouns. Gortner and Pennebaker's (2003) study of articles in the student newspaper after a campus tragedy found a similar increase in negative emotion and in collective focus as represented by more first person plural pronouns.

The similarity of one–star reviews to the linguistic characteristics of these trauma narratives suggests a hypothesis that negative restaurant reviews are not simply reviews describing bad food, but rather are trauma narratives, a coping mechanism (Pennebaker and Harber, 1993) for dealing with the minor trauma people experience at the restaurants.

To confirm that these results hold more generally, we extracted measure of these linguistic tendencies from the entire set of reviews, and used an ordered logistic regression to test whether these linguistic features of trauma are indeed associated with negative reviews more generally, after controlling for potential confounds like length and price.

We extracted linguistic variables to measure negative emotion, narrativity, and first person plural, as follows:

> **Negative emotion**: We used the list of negative emotional words tagged "negemo" from the LIWC lexicon (Pennebaker, *et al.*, 2007). The original list had 500 words and word–stems. Stems were expanded (so for example the stem "fail*" expands to include "fails", "failed", "failing", "failure", and "failures") and examples that occurred only once and were likely to be errorful were eliminated. The resulting dictionary contained 2,387 word types, from very frequent examples like "bad" (100,000 occurrences out of 100 million total words) or "disappointed" (36,000 occurrences) to rare examples like "heartbreakingly" (nine occurrences) or "antagonized" (two occurrences). We then entered as a feature in the regression for each review the log of the number of negative emotional words that occurred in the review.

> **First person plural**: We counted all occurrences of the words "we", "us", "our", "ours", "ourselves".

The three narrative features with the highest factor weights in Biber's (1988) model were extracted to model Biber's narrative dimension:

> **Past tense and perfect verbs**: We ran the Stanford part–of–speech tagging software (Toutanova, *et al.*, 2003) on the text of each review

to mark all instances of past tense (preterites) and past participles. We coded the past tense variable as the number of preterites in the review. The perfect variable was the number of instances of the perfect tense, extracted following Biber (1988) as any form of the verb "have" followed by a past participle, including those with adverbs between.

Third person pronouns: We counted all occurrences of the words "he", "she", "him", "his", "her", and "hers".

We then summed the past tense, perfect, and third person pronoun variables to construct a "narrative" variable. Figure 2 shows the words per review for the three variables with confidence intervals.
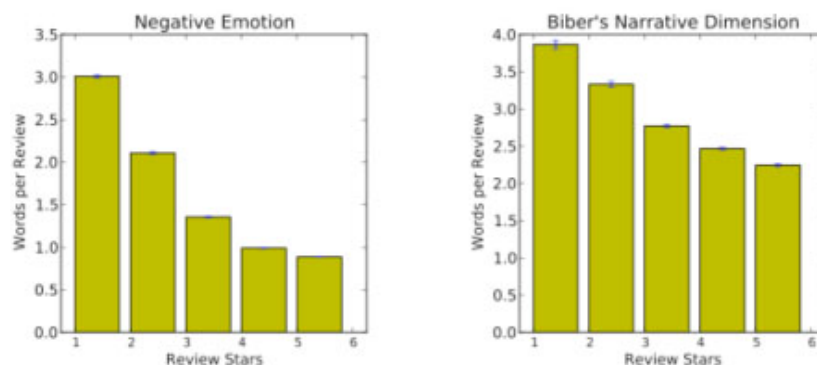
Figure 2: Count of words per reviews for the three features associated with trauma

We used the polr package in R to run an ordered logistic regression predicting the number of stars (one to five), with the length of each review in words and the price of the restaurant as control variables, and these three variables as independent variables. Each set of counts was first log–transformed. Because word counts are generally all quite collinear (the longer the review, the more words that can occur for each of the word categories) we converted each log count to a residual by linear regressing the log review length against the log count and entering the resulting residuals as variables.

Table 3 shows the regression coefficients.

Table 3: Coefficients from the ordered logistic regression predicting restaurant ranking (one–five).					
	Estimate	**Standard error**	**t value**	**Pr(>|t|)**	
categoryamerican_(traditional)	-0.118352230	0.008514940	-13.8993608	<2e-16	*
categoryasian_fusion	-0.383909166	0.019314570	-19.8766609	<2e-16	*
categorybakeries	0.152791277	0.022188440	6.8860756	5.735250e-12	*
categorybarbeque	-0.332836958	0.024549337	-13.5578797	<2e-16	*
categorybars	-0.420993085	0.011393686	-36.9496814	<2e-16	*
categorychinese	-0.205331658	0.010102779	-20.3242751	<2e-16	*
categorycoffee	0.298540967	0.017507862	17.0518227	<2e-16	*
categorydiners	-0.453196049	0.023040437	-19.6695942	<2e-16	*
categoryethiopian	0.099255106	0.031617176	3.1392780	1.693647e-03	*
categoryfast_food	-0.013952121	0.013117052	-1.0636628	2.874815e-01	
categoryfrench	0.158557299	0.010608788	14.9458448	<2e-16	*
categorygreek	0.557878356	0.062775866	8.8868286	<2e-16	*
categoryindian	0.003219127	0.013577776	0.2370879	8.125886e-01	
categoryitalian	0.060221409	0.008858718	6.7979825	1.060944e-11	*
categoryjapanese	-0.118879860	0.008754026	-13.5800208	<2e-16	*

categorykorean	-0.113884208	0.018707318	-6.0876824	1.145569e-09	*
categorylatin_american	0.195435763	0.013636119	14.3322129	<2e-16	*
categorymediterranean	-0.136183956	0.024457939	-5.5680881	2.575497e-08	*
categorymexican	-0.254312580	0.009866050	-25.7765349	<2e-16	*
categorymiddle_eastern	0.317829004	0.015938524	19.9409307	<2e-16	*
categoryother	0.055613105	0.022895170	2.4290322	1.513919e-02	
categoryotherasian	0.094814182	0.011865451	7.9907776	1.340902e-15	*
categoryothereuropean	0.345084176	0.023729731	14.5422707	<2e-16	*
categorypizza	-0.005113619	0.009866164	-0.5182986	6.042499e-01	
categorysandwiches	0.180697773	0.013160024	13.7308086	<2e-16	*
categoryseafood	-0.124489960	0.015569266	-7.9958788	1.286529e-15	*
categorysoul_food	0.367521216	0.035655747	10.3074890	<2e-16	*
categorysouthern	-0.012340794	0.037757708	-0.3268417	7.437876e-01	
categoryspanish	0.017856816	0.017378568	1.0275194	3.041760e-01	
categorysteakhouses	0.105009040	0.016068250	6.5351882	6.352958e-11	*
categorythai	-0.083068955	0.011528411	-7.2055860	5.779475e-13	*
categoryvegetarian	0.144773248	0.021379028	6.7717413	1.272416e-11	*
citychicago	0.326285955	0.010705069	30.4795751	<2e-16	*
cityla	0.122540208	0.011769003	10.4121149	<2e-16	*
citynyc	-0.064688543	0.008119377	-7.9671811	1.623347e-15	*
cityphiladelphia	0.100037878	0.011704558	8.5469161	<2e-16	*
citysf	0.068031992	0.007902410	8.6090185	<2e-16	*
citywashington	-0.225251878	0.010404948	-21.6485353	<2e-16	*
logreviewlen	-0.220581853	0.002314758	-95.2936850	<2e-16	*
price.L	0.640794311	0.009253715	69.2472525	<2e-16	*
price.Q	0.325718985	0.006752567	48.2363178	<2e-16	*
price.C	-0.001066022	0.004416229	-0.2413874	8.092549e-01	
negative_emotion	-0.950972003	0.003974972	-239.2399520	<2e-16	*
narrative	-0.242580818	0.002258191	-107.4226166	<2e-16	*
1st_person_plural	-0.119710870	0.004701538	-25.4620666	<2e-16	*
service_staff	-0.390894143	0.004942540	-79.0877094	<2e-16	*
1|2	-4.175936003	0.015116176	-276.2561105	<2e-16	*
2|3	-2.980166634	0.014612152	-203.9512514	<2e-16	*
3|4	-1.749688221	0.014357006	-121.8699904	<2e-16	*
4|5	0.082333331	0.014249947	5.7777992	7.568406e-09	*

Many of the control variables in Table 3 are significant; restaurants get higher review scores when they are higher priced, and the quadratic term for price suggests additionally that both extra-low and extra–high priced restaurants get higher rankings. Chicago restaurants get higher scores, while New York and Washington restaurants get particularly low ones, suggesting regional norms in star assignment. Bakeries, cafes, sandwich shops, vegetarian and some European restaurants (French, Italian, Greek) all tend to have higher scores, while Asian food (Chinese, Japanese, Korean, Thai) and some subsets of American food (diners, barbecue, and American (traditional)) all have lower scores. After controlling for these variables, there is a significant effect of trauma narratives: the use of Narrative ($p < 2 \times 10^{-16}$), Negative emotion ($p < 2 \times 10^{-16}$), and First person plural ($p < 2 \times 10^{-16}$) are all associated with lower rankings.

We examined a random selection of one–star reviews. While there were definitely complaints about food ("watery" chowder, "tasteless dry overfried" fish, "no flavor at all") and price ("overpriced", "outrageously expensive"), the overriding complaint was indeed about traumatic interpersonal relations: the host made the customer wait before seating or sat other people first or chose a bad table, the waiter or waitress

was rude, unavailable, or didn't apologize for mistakes, the manager didn't help, and so on.

Here are examples of two reviews (modified slightly to preserve anonymity):

> "The bartender was either new or just absolutely horrible ... we waited 10 min before we even got her attention to order ... and then we had to wait 45 — FORTY FIVE! — minutes for our entrees ... Dessert was another 45 min. wait, followed by us having to stalk the waitress to get the check ... she didn't make eye contact or even break her stride to wait for a response ... the chocolate souffle was disappointing ... I will not return."

> "So rude! We walked into [name] with a group of 6 and there were at least 3 or 4 tables that I could see were empty. I inquired about seating and before I could finish my sentence the host/waiter who was speaking to us abruptly announced 'we are sold out for the night you have to make a reservation' and pretty much chased us out the door. Will not make the mistake trying to give them business again just to be shoved out the door."

In summary, one–star reviews were overwhelmingly focused on narrating experiences of trauma rather than discussing food, both portraying the author as a victim and using first person plural to express solace in community.

---

5. Narrative framing in positive reviews: The addiction narrative

A very common narrative, appearing in both the popular and scientific literature, frames food as an addictive substance and the eater as an addict subject to cravings or desire. The use of drugs as a metaphor is common for all sorts of pleasurable experiences, including food as well as love [3]. For food, the metaphor has been most prevalent in discussing addiction (Rozin and Stoess, 1993; Rozin, *et al.*, 1991). In this section we investigate whether consumer reviews make use of this framing.

5.1. Methods

As in the previous study, we used a lexicon to operationalize the addiction framing, and then use ordered logistic regression to predict review rating from this variable after accounting for control variables. The lexicon of words and phrases designed to operationalize the metaphor includes the following words and phrases and their inflected forms: *addiction, crave/craving, chocoholic, jonesing, binge/binging*. It also includes phrases in which drugs are described as a metaphor (*drug of choice, like a drug, new drug, favorite drug*, etc.) and phrases describing food as the drug crack (including *made of crack, food crack, edible crack*, etc.).

As with the previous regressions, the counts were first log–transformed, and we again converted each log count to a residual by linear regressing the log review length against the log count and entering the resulting residuals as variables. In order to check whether this framing is used differentially by price we use a ordered logistic regression predicting restaurant price level from the addiction variable.

5.2. Results

The regression against rating shows that after controlling for restaurant category, city, and review length, the use of the language of addiction is associated with higher ratings ($p<2\times10^{-16}$). Figure 3a shows the number of mentions of addiction per review for the different rating categories, with .95 confidence intervals.
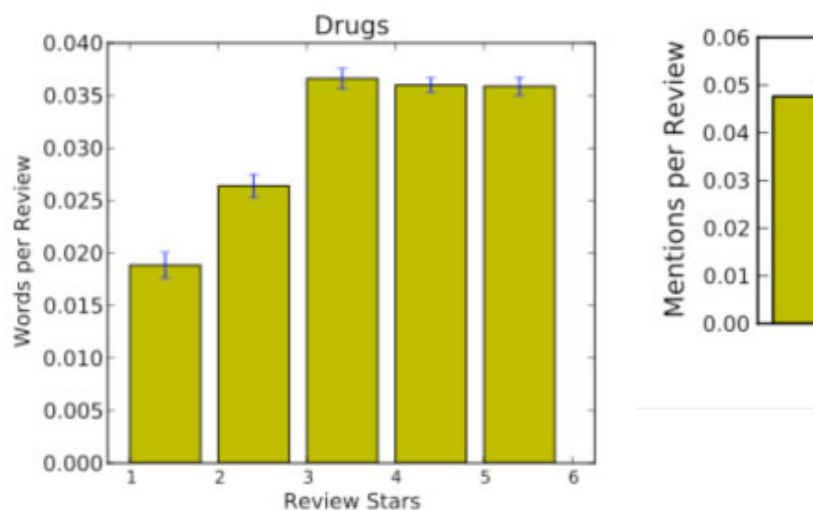
Figure 3: (a, left) Relation between the use of words or phrases related to drug/addiction and intervals; (b, right) The cheaper the restaurant, the more use of the language of drugs and ad

The regression against restaurant price level shows that after controlling for restaurant category, city, and review length, the use of the language of addiction is associated with cheaper restaurants ($p < 2 \times 10^{-16}$). Figure 3b shows the mentions per review with confidence intervals. Representative examples include:

> the ... garlic noodles should be outlawed! They are now my **drug of choice**

> these cup cakes are like **crack**

> be warned the wings are **addicting**

> ... every time **I need a fix**. That fried chicken is so damn good!

Table 4 shows the foods most likely to be described as addicting or craved.

Table 4: Foods most likely to be described using drug metaphors.				
Meaty, fatty foods	**Starchy comfort food**	**Sweet food**	**Small ethnic dishes**	**Descriptors**
burgers	pizza	sweets	sushi	comfort
barbecue	mac and cheese	pancakes, breakfast	dim sum	fried, greasy
chicken wings	pasta/noodles	sugar	tacos, burritos	unhealthy
french fries	soups	chocolate	spam musubi	hearty, satisfying
	sandwiches	beignets	dumplings	junk
			falafel	authentic
			tapas	cheap

Foods that are described as being "addicting" or "craved" are described as "comfort food", using adjectives like "fried", "unhealthy", "authentic", or "cheap". They consist of fried, starchy, or sweet foods. These are generally not normative, "sit–down–dinner" entrees, but rather take–out food, fast food, or snacks. The ethnic foods most likely to be craved are dishes that are small and perceived as non–normative, snack–like dishes: sushi, dim sum, falafel, tacos.

Finally, we explored the role of gender, testing whether women or men are more likely to frame themselves as addicted.

While the gender of Yelp reviewers is not made available, their first name is generally available on their reviewer sites. Previous research has shown that in many (although not all) cases first names can be used to estimate the gender of the writer

(Herdağdelen and Baroni, 2011; Vogel and Jurafsky, 2011; Smith, *et al.*, 2013). For a small subset of our data (4,929 reviews) we retrieved the first name of the reviewer by finding the name on the review Web page. We then assigned gender to names by using the name database of the U.S. Social Security Administration (http://www.ssa.gov/oact/babynames/names.zip), selecting names for children born after 1951. Each reviewer name was assigned a gender only if the Social Security database was sufficiently strongly biased toward one gender (constituting at least 80 percent of the births). We then used linear regression to predict the number of mentions of addiction from the gender of the speaker. We found that women were significantly more likely than men to talk about food as a drug (p=0.000832).

We confirmed the gender result by examining a second dataset released by Yelp for the Phoenix metropolitan area (http://www.yelp.com/dataset_challenge/) which has reviewer first name information for a much larger set of reviews. We looked at the 161,897 reviews for restaurants in this database, and used the Social Security name database with the same 80 percent threshold to assign gender. Our algorithm was able to assign a gender to 90 percent of the names. We then ran a linear regression on this database to predict the number of mentions of addiction from the gender of the speaker. Once again, women were significantly more likely than men to talk about food as a drug ($p < 2 \times 10^{-16}$).

5.3. Discussion

Whether there is in fact a biochemical link between junk food cravings and drug addiction is an open question in the literature [4]. Nonetheless, our results suggest that the folk model of this belief is productive and widespread in consumer reviews. Hormes and Rozin (2010) found that participants rated the words "craving" and "addiction" in various languages as being most appropriately applied to drugs, alcohol, or food. Our study extends these results to show that the metaphor of food as an addiction or craving tends to apply to a particular subset of foods. The foods that are "craved" are foods that are in some way non–normative: they are meaty, sugary, starchy foods, generally fast food and street food, or small snack–like inexpensive ethnic foods. Craved foods aren't vegetables, or main courses like meatloaf or fish or even side dishes like mashed potatoes. The folk model of what we crave or are addicted to encompasses foods that are somehow considered inappropriate for a meal, bad for you (unhealthily full of fats and sugars), inexpensive, comfort food that we feel guilty for having but eat anyhow.

The result that women are more likely to use this metaphor in our data is also consistent with previous results. Rozin, *et al.* (1991) found that females are significantly more likely to express cravings for chocolate than males. Zellner, *et al.* (1999), Weingarten and Elston (1990), and Osman and Sobal (2006) found that female undergraduates were more likely than males to report food cravings. Our results do not distinguish among the possible causes of the greater number of these expressions by female reviewers: women might be more likely than men to have these cravings or feelings, women might be more comfortable than men to admitting to these cravings, or women might simply be more likely than men to use this particular linguistic metaphor to describe their otherwise identical desires. Choosing among these or other possible causal scenarios remains for future work.

In summary, our use of automatic processing of online reviews to detect the expression of these cravings is a significant methodological extension of earlier work on food cravings, enabling a much larger–scale investigation with more details about the nature of the foods that are framed this way and who is doing the framing.

---

6. Narrative framing in reviews of expensive restaurants

In our final study we investigated two sets of frames associated with reviews of very expensive restaurants, to understand how expense is characterized in reviews. We first examined review features linked with educational capital. Education is strongly associated with differences in socioeconomic status, and in fact is one of the main ways that class status is defined in social scientific studies, along with work and income. Previous work on food advertising found that advertising of more expensive products employs longer, more complex words and longer sentences (Freedman and Jurafsky, 2011), presumably because complex words or sentences signal the writers' higher educational capital, and hence project higher social status. We therefore tested whether this use of more complex language to project "linguistic capital" was similarly associated with price in reviews, predicting that reviews more expensive restaurants would be longer and use longer words.

The second feature we investigate frames food as a sensual or even sexual pleasure. This tendency is widespread in expensive wine reviews, which make extensive use of phrases like *sexy, sensual, seductive, voluptuously textured, ravishing*, and *hedonistic* (Lehrer, 2009; McCoy, 2005; Shesgreen, 2003). Television food commercials in the United States also emphasize "sensual hedonism" with words like *luscious, indulgent, irresistible*, and *decadent* (Strauss, 2005). We therefore expected reviews of expensive restaurants to use words related to sex or sensuality.

6.1. Methods

Linguistic capital: To test the hypotheses of linguistic capital we coded two variables that mark language complexity.

1. The **total number of words** in the review. We used the log of this value.

2. The **average word length in letters** of all words in the review. We again used the log of this value.

Sensual language: The lexicon of words and phrases designed to operationalize the metaphor was drawn from the previous literature. The first lexicon models sex and sensuality, with the following words and stems (some drawn from the LIWC (Pennebaker, *et al.*, 2007) lexicon category "Sex", others from inspection of menus): *erotic, food porn, lust, lusted, lusting, naughty, orgasm*, pornographic, seductive*, sensual*, sex*, sinful, sultry, tempt, temptation, tempting, voluptuous, wine porn*. The notation* means all words beginning with this prefix (so *sex** includes *sexy, sexual, sexier*, and *orgasm** includes *orgasmic* and *orgasmically*). As with the previous regressions, the counts were first log–transformed, and we again converted each log count to a residual by linear regressing the log review length against the log count and entering the resulting residuals as variables.

We added one control factor, the restaurant category, which consisted of a label from a set of 32 types of restaurants described above. Ordered logistic regression was then used to predict the restaurant price (an ordered class ranging over $, $$, $$$, $$$$) from the variables of interest and the control variable, via the polr package in R. We also used a separate ordered logistic regression to predict the restaurant rating (an ordered class ranging over one–five stars) from the variables of interest and the control variable, again via the polr package in R.

6.2. Results

After controlling for restaurant type, expensive restaurants were significantly more likely to make use of longer words ($p < 2 \times 10^{-16}$) and longer reviews ($p < 2 \times 10^{-16}$); Figure 4 shows the values and .95 confidence intervals.

Figure 4: Longer words and longer reviews are positively associated with res

The ordered regression on price found that, by contrast to the addiction narratives in the previous section, the metaphor of sex and sensual pleasure is more likely to be used when reviewers are describing expensive restaurants ($p = 3.22 \times 10^{-5}$). Some examples:

> the apple tarty ice cream pastry caramely thing was just **orgasmic**

> sumptuous flavors, jaw–droppingly good **sexy** food

> succulent pork belly paired with **seductively** seared foie gras

Figure 5 shows the number of mentions per review by price level, comparing it with the values for the drug/addiction framing shown earlier.

Figure 5: The more expensive the restaurant, the more metaphors of sex; the cheaper the
language of drugs and addiction.

The regression on star rating showed that mentions of sex are associated with higher
ratings ($p<2\times10^{-16}$). Figure 6 shows the mentions per review with confidence
intervals.
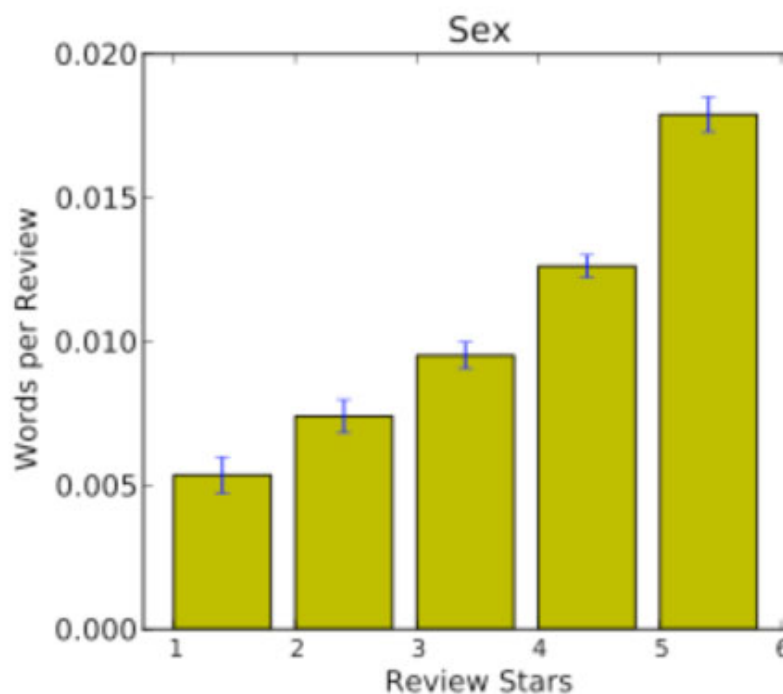
Figure 6: Number of mentions per review of words or phrases related to sex showing that
framing is associated with higher ratings.

To further explore the framing of expensive food or restaurants as sex, we extracted
the words most likely to appear near these sexual words and phrases, which we
defined as those words with the highest log likelihood ratio between the counts near
sexual words and their counts elsewhere in the reviews, using the weighted log–
odds–ratio, informative Dirichlet prior method of Monroe, *et al.* (2008) described
above. The words most associated with sex and sensuality fall into two classes:
dessert (words like *chocolate, cake, dessert, truffle, pastry, pistachio, cheesecake*),
and romantic ambiance (words like *dark, romantic, lighting, vibe, ambiance, décor*).

This relationship of sex with dessert is a common cultural meme (Rozin, 1987). To
explore other functions of dessert, we looked at the association between mentions of
dessert in a review and the rating of the review. We developed a list of 500 words and
phrases for desserts and used it to automatically code the number of mentions of
dessert in each review. Figure 7 shows the mentions per review, by review rating. We
entered the number of mentions of desserts into the ordered regression predicting
rating and found that (after controlling for restaurant category and review length)
mentioning dessert is a significant predictor of higher rating ($p<2\times10^{-16}$).
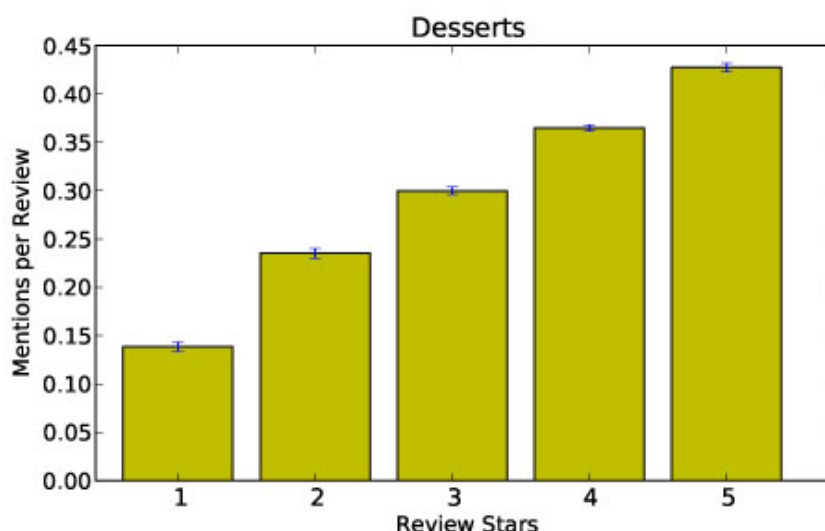
Figure 7: Number of times dessert is mentioned per review by review rating.

We again checked the gender of the reviewers by adding a gender variable to the ordered regression predicting rating. After controlling for restaurant category women are significantly more likely than men to talk about dessert (p=0.000138). However, we found no difference between women and men in the use of sexual framing of dessert or other food.

6.3. Discussion

The fact that reviewers use more complex words and write longer reviews for more expensive restaurants suggests that reviewers are adopting the stance of the high socio–economic class associated with expensive restaurants. The use of this higher level of educational capital is thus another way that the review offers a chance for self–depiction, in this case a way for the reviewer to portray themselves as well–educated. By using the metaphor of sexuality and sensuality in these long reviews the reviewer further portrays themselves as a food lover attuned to the sensual and hedonic element of cuisine.

An additional implication from this section is the important role of dessert as a psychological and social marker in food reviews. Reviews are more positive when they mention dessert, desserts are more likely to be discussed by women, and desserts are associated in the language of both men and women with sex and sensuality.

7. General discussion

Diverse narratives and framings were found across different kinds of reviews. One–star reviews are trauma narratives that help cope with face threats by portraying the author as a victim and seeking solace in community. Positive reviews appeal, presumably light–heartedly, to the author as an addict suffering from cravings for junk foods, non–normative meals, and other guilty pleasures. Reviews of expensive reviews use more complex words and wordy reviews to portray the reviewer as educated and possessed of higher linguistic capital, and use the language of sensuality to emphasize the reviewer's credentials as a sensualist. Across multiple variables, online review narratives reveal the reviewers' concern with face and the presentation of the self. Even the fact that reviews show a stronger positivity bias than general text suggests that reviews reveal a tendency toward positive self–presentation. Previous work has shown that online consumer reviews are an important source of insights into consumer sentiment about specific products. Our work shows that online reviews are also valuable as a source of insight into social psychological processes via their link with narrative framings.

These findings also offer a new methodology for using online text and automatic gender computation to confirm and extend prior work on both food cravings and gender and food. Our results suggest that the objects of food cravings are non–normative foods, snacks, unhealthy comfort foods or small foods that are generally seen as some sort of violation of cuisine norms. Our work also is consistent with previous work suggesting that women are more likely to use the metaphors of addiction to describe food desires, and women are more likely to discuss dessert. Previous research has suggested that these findings are likely quite culture–specific,

and the subject clearly calls for further cross–cultural study. The use of online reviews offers a natural way to investigate these questions across cultures by acquiring parallel data from consumer reviews across different cultures and languages. The results of this study may also have implications for the restaurant industry; the fact that negative reviews describe service–related traumas may offer an avenue for identifying problems with customer satisfaction.

Our study has a number of limitations. The reviews we consider are all in English, and limited to the United States. Considering reviews from different languages and regions, and of products other than restaurants, as well as over different time periods than our 2006–2011 window, could lead to significantly broader conclusions.

Despite these limitations, the results of our initial investigation are promising, and suggest that online reviews, with their rich affective content, offer an important new direction of inquiry in using Web data to inform and advance the behavioral sciences.
FM

About the authors

Dan Jurafsky is Professor of Linguistics and Computer Science at Stanford University.
Direct comments to: jurafsky [at] stanford [dot] edu

Victor Chahuneau is Graduate Research Assistant in the Language Technologies Institute at Carnegie Mellon University.
E–mail: vchahune [at] cs [dot] cmu [dot] edu

Bryan R. Routledge is Associate Professor of Finance in the Tepper School of Business at Carnegie Mellon University.
E–mail: routledge [at] cmu [dot] edu

Noah A. Smith is Finmeccanica Associate Professor of Language Technologies and Machine Learning in the School of Computer Science at Carnegie Mellon University.
E–mail: nasmith [at] cs [dot] cmu[dot] edu

Acknowledgments

Notes

1. Zimmer, 1964; Clark and Clark, 1977, p. 538.

2. Hu and Liu, 2004; Pang and Lee, 2008; Potts, 2011, *inter alia*.

3. Sunderland and Denny, 2003, p. 196.

4. Avena, *et al.*, 2009; Johnson and Kenny, 2010; Ziauddeen, *et al.*, 2012, *inter alia*.

References

N. Archak, A. Ghose, and P.G. Ipeirotis. 2011. "Deriving the pricing power of product features by mining consumer reviews," *Management Science*, volume 57, number 8, pp. 1,485–1,509.
doi: http://dx.doi.org/10.1287/mnsc.1110.1370, accessed 21 March 2014.

N.M. Avena, P. Rada, and B.G. Hoebel. 2009. "Sugar and fat bingeing have notable differences in addictive–like behavior," *Journal of Nutrition*, volume 139, number 3, pp. 623–628.
doi: http://dx.doi.org/10.3945/jn.108.097584, accessed 21 March 2014.

A.A. Augustine, M.R. Mehl, and R.J. Larsen. 2011. "A positivity bias in written and spoken English and its moderation by personality and gender," *Social Psychological & Personality Science*, volume 2, number 5, pp. 508–515.
doi: http://dx.doi.org/10.1177/1948550611399154, accessed 21 March 2014.

S. Baccianella, A. Esuli, and F. Sebatiani. 2010 "SENTIWORDNET 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," *LREC '10: Proceedings of the Seventh International Conference on Language Resources and Evaluation*, at http://www.lrec-conf.org/proceedings/lrec2010/summaries/769.html, accessed 21 March 2014.

D. Biber, 1995. *Dimensions of register variation: A cross–linguistic comparison*. Cambridge: Cambridge University Press.

D. Biber, 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

S. Blair–Goldensohn, T. Neylon, K. Hannan, G. Reis, R. McDonald, and J. Reynar, 2008. "Building a sentiment summarizer for local service reviews," *NLP in the Information Explosion Era*, at http://www.ryanmcd.com/papers/local_service_summ.pdf, accessed 21 March 2014.

J. Boucher and C.E. Osgood, 1969. "The Pollyanna hypothesis," *Journal of Verbal Learning and Behavior*, volume 8, number 1, pp. 1–8.
doi: http://dx.doi.org/10.1016/S0022-5371(69)80002-2, accessed 21 March 2014.

S. Brody and N. Elhadad. 2010. "An unsupervised aspect–sentiment model for online reviews," *HLT '10 Human Language Technologies: Proceedings of the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 804–812.

K.W. Church and P. Hanks. 1990. "Word association norms, mutual information, and lexicography," *Computational Linguistics*, volume 16, number 1, pp. 22–29.

P.S. Dodds and C.M. Danforth, 2009. "Measuring the happiness of large–scale written expression: Songs, blogs, and presidents," *Journal of Happiness Studies*, volume 11, number 4, pp. 441–456.
doi: http://dx.doi.org/10.1007/s10902-009-9150-9, accessed 21 March 2014.

P.S. Dodds, K.D. Harris, I.M. Kloumann, C.A. Bliss, and C.M. Danforth, 2011. "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter," *PLoS ONE*, volume 6, number 12, e26752.
doi: http://dx.doi.org/10.1371/journal.pone.0026752, accessed 21 March 2014.

T. Dunning, 1993. "Accurate methods for the statistics of surprise and coincidence," *Computational Linguistics*, volume 19, number 1, pp. 61–74.

A. Herdağdelen and M. Baroni. 2011. "Stereotypical gender actions can be extracted from Web text," *Journal of the American Society for Information Science and Technology*, volume 62, number 9, pp. 1,741–1,749.
doi: http://dx.doi.org/10.1002/asi.21579, accessed 21 March 2014.

J.M. Hormes and P. Rozin. 2010. "Does 'craving' carve nature at the joints? Absence of a synonym for craving in many languages," *Addictive Behaviors* volume 35, number 5, pp. 459–463.
doi: http://dx.doi.org/10.1016/j.addbeh.2009.12.031, accessed 21 March 2014.

M. Hu and B. Liu, 2004. "Mining and summarizing customer reviews," *KDD '04: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168–177.
doi: http://dx.doi.org/10.1145/1014052.1014073, accessed 21 March 2014.

Y. Jo and A.H. Oh, 2011. "Aspect and sentiment unification model for online review analysis," *WSDM '11: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, pp. 815–824.
doi: http://dx.doi.org/10.1145/1935826.1935932, accessed 21 March 2014.

P.M. Johnson and P.J. Kenny. 2010. "Dopamine D2 receptors in addiction–like reward dysfunction and compulsive eating in obese rats," *Nature Neuroscience*, volume 13, number 5, pp. 635–641.
doi: http://dx.doi.org/10.1038/nn.2519, accessed 21 March 2014.

D.F. Larcker and A.A. Zakolyukina, 2012. "Detecting deceptive discussions in conference calls," *Journal of Accounting Research*, volume 50, number 2, pp. 495–540.
doi: http://dx.doi.org/10.1111/j.1475-679X.2012.00450.x, accessed 21 March 2014.

C.D. Manning and H. Schütze, 1999. *Foundations of statistical natural language processing*. Cambridge, Mass.: MIT Press.

J. McAuley, J. Leskovec, and D. Jurafsky, 2012. "Learning attitudes and attributes from multi–aspect reviews," *ICDM '12: Proceedings of the 2012 IEEE 12th International Conference on Data Mining*, pp. 1,020–1,025.
doi: http://dx.doi.org/10.1109/ICDM.2012.110, accessed 21 March 2014.

E. McCoy, 2005. *The emperor of wine: The rise of Robert M. Parker, Jr. and the reign of American taste*. New York: ECCO.

J.–B. Michel, Y.K. Shen, A.P. Aiden, A. Veres, M.K. Gray, The Google Books Team, J.P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, J. Orwant, S. Pinker, M.A. Nowak, and E.L. Aiden, 2011. "Quantitative analysis of culture using millions of digitized books," *Science*, volume 331, number 6014, pp. 176–182.
doi: http://dx.doi.org/10.1126/science.1199644, accessed 21 March 2014.

B. Monroe, M. Colaresi, and K. Quinn, 2008. "Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict," *Political Analysis*, volume 16, number 4, pp. 372–403.
doi: http://dx.doi.org/10.1093/pan/mpn018, accessed 21 March 2014.

J.L. Osman and J. Sobal, 2006. "Chocolate cravings in American and Spanish individuals: Biological and cultural influences," *Appetite*, volume 47, number 3, pp. 290–301.
doi: http://dx.doi.org/10.1016/j.appet.2006.04.008, accessed 21 March 2014.

B. Pang and L. Lee, 2008. "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, volume 2, numbers 1–2, pp. 1–135.

doi: http://dx.doi.org/10.1561/1500000011, accessed 21 March 2014.

J.W. Pennebaker and L.A. King, 1999. "Linguistic styles: Language use as an individual difference," *Journal of Personality and Social Psychology*, volume 77, number 6, pp. 1,296–1,312.
doi: http://dx.doi.org/10.1037/0022-3514.77.6.1296, accessed 21 March 2014.

J.W. Pennebaker and K.D. Harber, 1993. "A social stage model of collective coping: The Loma Prieta Earthquake and the Persian Gulf War," *Journal of Social Issues*, volume 49, number 4, pp. 125–145.
doi: http://dx.doi.org/10.1111/j.1540-4560.1993.tb01184.x, accessed 21 March 2014.

J.W. Pennebaker, R.J. Booth, and M.E. Francis, 2007. "Linguistic inquiry and word count: LIWC 2007," at http://homepage.psy.utexas.edu/HomePage/Faculty/Pennebaker/Reprints/LIWC2007_Operato accessed 21 March 2014.

J.W. Pennebaker, T.J. Mayne, and M.E. Francis, 1997. "Linguistic predictors of adaptive bereavement," *Journal of Personality and Social Psychology*, volume 72, number 4, pp. 863–871.
doi: http://dx.doi.org/10.1037/0022-3514.72.4.863, accessed 21 March 2014.

A.–M. Popescu and O. Etzioni. 2005. "Extracting product features and opinions from reviews," *HLT '05: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 339–346.
doi: http://dx.doi.org/10.3115/1220575.1220618, accessed 21 March 2014.

C. Potts, 2011. "On the negativity of negation," *Proceedings of Semantics and Linguistic Theory*, volume 20, pp. 636–659; version at http://elanguage.net/journals/salt/article/viewFile/20.636/1414, accessed 21 March 2014.

K. Reschke, A. Vogel, and D. Jurafsky, 2013. "Generating recommendation dialogs by extracting information from user reviews," *51st Annual Meeting of the Association for Computational Linguistics — Short Papers*; version at http://www.stanford.edu/~jurafsky/pubs/yelp-acl2013.pdf, accessed 21 March 2014.

E. Riloff and J. Wiebe, 2003. "Learning extraction patterns for subjective expressions," *EMNLP '03: Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, pp. 105–112.
doi: http://dx.doi.org/10.3115/1119355.1119369, accessed 21 March 2014.

P. Rozin, 1987. "Sweetness, sensuality, sin, safety, and socialization: Some speculations," In: J. Dobbing (editor). *Sweetness*. New York: Springer–Verlag, pp. 99–111.
doi: http://dx.doi.org/10.1007/978-1-4471-1429-1_7, accessed 21 March 2014.

P. Rozin and C. Stoess, 1993. "Is there a general tendency to become addicted?" *Addictive Behaviors*, volume 18, number 1, pp. 81–87.
doi: http://dx.doi.org/10.1016/0306-4603(93)90011-W, accessed 21 March 2014.

P. Rozin, L. Berman, and E. Royzman, 2010. "Biases in use of positive and negative words across twenty languages," *Cognition & Emotion*, volume 24, number 3, pp. 536–548.
doi: http://dx.doi.org/10.1080/02699930902793462, accessed 21 March 2014.

P. Rozin, E. Levine, and C. Stoess, 1991. "Chocolate craving and liking," *Appetite*, volume 17, number 3, pp. 199–212.
doi: http://dx.doi.org/10.1016/0195-6663(91)90022-K, accessed 21 March 2014.

S. Shesgreen, 2003. "Wet dogs and gushing oranges: Winespeak for a new millennium," *Chronicle of Higher Education* (7 March), at http://chronicle.com/article/Wet-DogsGushing-Oranges-/20985, accessed 21 March 2014.

Y. Sim, B.D.L. Acree, J.H. Gross, and N.A. Smith. 2013. "Measuring ideological proportions in political speeches," *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 91–101, and at http://aclweb.org/anthology//D/D13/D13-1010.pdf, accessed 24 March 2014.

B.N. Smith, M. Singh, and V.I. Torvik, 2013. "A search engine approach to estimating temporal changes in gender orientation of first names," *JCDL '13: Proceedings of the 13th ACM/IEEE–CS Joint Conference on Digital Libraries*, pp. 199–208.
doi: http://dx.doi.org/10.1145/2467696.2467720, accessed 21 March 2014.

B. Snyder and R. Barzilay. 2007. "Multiple aspect ranking using the good grief algorithm," *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, pp. 300–307.

L.D. Stone and J.W. Pennebaker, 2002. "Trauma in real time: Talking and avoiding online conversations about the death of Princess Diana," *Basic and Applied Social Psychology*, volume 24, number 3, pp. 173–183.
doi: http://dx.doi.org/10.1207/S15324834BASP2403_1, accessed 21 March 2014.

P.J. Stone, D.C. Dunphy, M.S. Smith, and D.M. Ogilvie, 1966. *The general inquirer: A computer approach to content analysis*. Cambridge, Mass.: MIT Press.

P.L. Sunderland and R.M. Denny. 2003. "Psychology vs anthropology: Where is culture in marketplace ethnography?" In: T. deWaal Malefyt and B. Moeran (editors).

Advertising cultures. London: Berg, pp. 187–202.

I. Titov and R. McDonald. 2008a. "A joint model of text and aspect ratings for sentiment summarization," *Proceedings of ACL–08*, pp. 308–316, and at http://aclweb.org/anthology/P/P08/P08-1036.pdf, accessed 21 March 2014.

I. Titov and R. McDonald. 2008b. "Modeling online reviews with multi–grain topic models," *WWW '08: Proceedings of the 17th International Conference on World Wide Web*, pp. 111–120.
doi: http://dx.doi.org/10.1145/1367497.1367513, accessed 21 March 2014.

K. Toutanova, D. Klein, C. Manning, and Y. Singer, 2003. "Feature–rich part–of–speech tagging with a cyclic dependency network," *NAACL '03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*, volume 1, pp. 173–180.
doi: http://dx.doi.org/10.3115/1073445.1073478, accessed 21 March 2014.

H.P. Weingarten and D. Elston, 1990. "The phenomenology of food cravings," *Appetite*, volume 15, number 3, pp. 231–246.
doi: http://dx.doi.org/10.1016/0195-6663(90)90023-2, accessed 21 March 2014.

T. Wilson, J. Wiebe, and P. Hoffmann, 2005. "Recognizing contextual polarity in phrase–level sentiment analysis," *HLT '05: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 347–354.
doi: http://dx.doi.org/10.3115/1220575.1220619, accessed 21 March 2014.

Y. Yang and J.O. Pedersen, 1997. "A comparative study on feature selection in text categorization," *ICML '97: Proceedings of the Fourteenth International Conference on Machine Learning*, pp. 412–420.

R.B. Zajonc, 1968. "Attitudinal effects of mere exposure," *Journal of Personality and Social Psychology*, volume 9, number 2, part 2, pp. 1–27.
doi: http://dx.doi.org/10.1037/h0025848, accessed 21 March 2014.

D.A. Zellner, A. Garriga–Trillo, E. Rohm, S. Centeno, and S. Parker, 1999. "Food liking and craving: A cross–cultural approach," *Appetite*, volume 33, number 1, pp. 61–70.
doi: http://dx.doi.org/10.1006/appe.1999.0234, accessed 21 March 2014.

H. Ziauddeen, I. Sadaf Farooqi, and P.C. Fletcher, 2012. "Obesity and the brain: How convincing is the addiction model?" *Nature Reviews Neuroscience*, volume 13, number 4, pp. 279–286.
doi: http://dx.doi.org/10.1038/nrn3212, accessed 21 March 2014.

K.E. Zimmer, 1964. "Affixal negation in English and other languages: An investigation of restricted productivity," *Language*, volume 42, number 1, pp. 134–142.

Editorial history

SHARE