Tweet it like you mean it: Exploring expressive lengthening in Tweets

Introduction

In spoken language, the prosody of a given utterance is a rich source of information between interlocutors (Schnoebelen, 2012; Rao, 2013, Bolinger; 1989). Pitch, duration and emphasis can encode the nuances of sarcasm, affection, or emotionality, and lead to inferences about speaker intention or psychological state far beyond what is encoded in lexical content alone. By contrast, written communication might be considered a relatively flat affective domain given the absence of spoken prosodic indicators and more stringent conventions around grammatical well-formedness. How could one possibly convey the boundless enthusiasm of having enjoyed the third installment of the Hunger Games trilogy? Surely not through lexical content alone:

"I saw Mockingjay and thought it was extremely good."

As it is for most important linguistic phenomena, @KimKardashian presents us with a path out of the dark:

@KimKardashian: "mocking jay was soooooooo good!!!!!!"

Computer mediated communication appears to be at the forefront of change in language convention, especially in written language (Eisenstein et. al, 2014). Increasingly, speakers incorporate written cues that seem to function as prosodic indicators to encode extra-linguistic information around emotionality, sentiment and subjectivity (Jen Doll, 2013; Schnoebelen, 2012). Communication through social media text such as tweets, and texting, appear to be especially promising domains (and a wild west of sorts) for studying these evolving conventions. In part, this is driven by the amount and accessibility of linguist data available from these sources. Additionally, the informal nature of these mediums and the rapid evolution of devices and terms observed therein, warrant continued investigation (Brody & Diakopoulos, 2011).

One example of a device that at the very least conveys emotional stance and also appears to convey information about the speaker, is the emoticon. In "Do you smile with your nose? Stylistic variation in Twitter Emoticons," Tyler Schnoebelen examined the 28 most common emoticons used in American English tweets. Schnoebelen points out that the emoticon was originally proposed, "in order to guide affective interpretations" and found the presence of at least one emoticon in 9.7% of the over 30 million American English tweets in his corpus. He claims that emoticon usage is an example of an attempt to "preserve" some of the information conveyed in spoken language. Like prosody, interlocutors make use of visual cues such as facial expression and gesticulation to convey an intended message. Emoticons, Schnoebelen offers, characterize "stylized representations of what gets lost" when communication is confined to text. While the focus of our analysis will be on the affective information communicated through expressive lengthening of sentence final exclamation marks, Schnoebelen's study showed that emoticons not only encode affective information around emotional stance, but also represent stylistic choices that vary by author. In this way, nose or non-nose choice for emoticons was a viable predictor of other lexical variations including frequency of emoticon use, misspelling, and expressive lengthening.

Expressive lengthening in texting and communication via social media is a well-known and often parodied phenomenon (Lynton, 2014; Doll, 2013). While some critics consider such linguistic variation the product of "laziness" or declining "proper language skills" (Merritt, 2013) others have argued expressive lengthening is an intentional and expressive device for encoding affective information in a prosody-less medium (Schnoebelen, 2012; Brody & Diakopoulos 2011). In "Using Word Lengthening to Detect Sentiment in Microblogs," Brody and Diakopoulos show in three experiments that the prevalence of expressive lengthening makes it a viable metric for assessing subjectivity in microblog and social media text, that expressive lengthening is itself associated with subjectivity and sentiment, and they present a method utilizing expressive lengthening as a sentiment measure. For the purposes of this paper we are most interested in the association between expressive lengthening and sentiment as observed in Experiment II of their study. Brody & Diakopoulos assess the hypothesis that "lengthening represents a textual substitute for prosodic indicators in speech." It follows from this that expressive lengthening represents an intentional and informative *choice* used to strengthen the sentiment or emotion intended in an utterance. In particular, Brody & Diakopoulos demonstrated the association between lengthening and subjectivity using a preexisting sentiment lexicon (Wilson et al., 2005) such that words that are more likely to be expressively lengthened are more likely to be rated as "subjective."

Motivations for the current study

The current study builds on the ideas presented by Schnoebelen (2012) and Brody & Diakopoulos (2011). In particular, we focus on a fairly specific, but prevalent device in tweets – sentence final lengthening (SFL) of exclamation marks (!). Building off the finding in Experiment II of Brody & Diakopoulos (2011) - that expressive lengthening is not arbitrary, but in fact associated with sentiment - we can make a to-date untested and fairly general, but intuitive hypothesis: that increasing the *degree* of lengthening of a lengthened device such as exclamation marks, denotes some kind of extra-linguistic meaning given the competition of less lengthened alternatives. We set out to examine the effects of increasing degree of sentence final (!) on clauses immediately preceding the device.

Data / Methodology

We scraped 1,000 tweets from the top 100 most followed twitter accounts using the command line tool t (https://github.com/sferik/t). After filtering for non-English accounts we parsed the data into clauses with at least one sentence final (!). In total, we were left with 26,624 tweet clauses with sentence final (!) with lengthening from a single (!) to six (!!!!!!). These 26,624 sentence final (!) tweet clauses came from approximately 86 unique accounts from @BillGates to @KimKardashian. Original tweets can be found here:
https://github.com/benpeloquin7/Winter_2015/tree/master/Linguist230a/Final_Project/data-hold/account-store

Analysis

The primary independent variable of interest throughout our analysis is the degree of expressive lengthening of sentence final (!) on a range of dependent measures, from clause character count, token count, average token emotionality rating, average co-occurrence of lengthened tokens in the preceding clause, and average length of a lengthened token. Among clauses that end with a sentence final (!) the vast majority are single (!) marks with only .55% (n=148) of our clauses containing an expressive lengthening of degree six (see Figure 1). Given the nature of this analysis, in which there is a high degree of stylistic variability between accounts, our main form of statistical analysis will make use of mixed-effects models to accommodate these account specific stylistic differences as random effects.

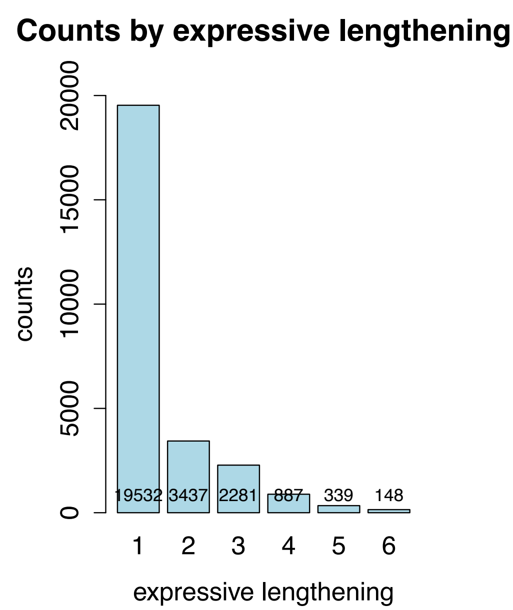

Figure 1 Clause counts grouped by degree of lengthening

Character counts, token counts and average token length

We examined expressive lengthening as a predictor of preceding clause character counts (see Figure 2) and token counts by constructing separate mixed effects models to test both of these dependent variables with number of target (!) as a fixed effect and random slopes and intercepts for account and account by number of target interactions. We used ANOVA to assess goodness of fit model differences with and without our fixed effects. Degree of expressive lengthening significantly improved model predictions of character counts, chi-square(1, 24.672) ($p < 0.001$) and token counts, chi-square(1, 55.8) ($p<0.001$). We observed that both mean character counts and token counts decrease as a function of SFL.

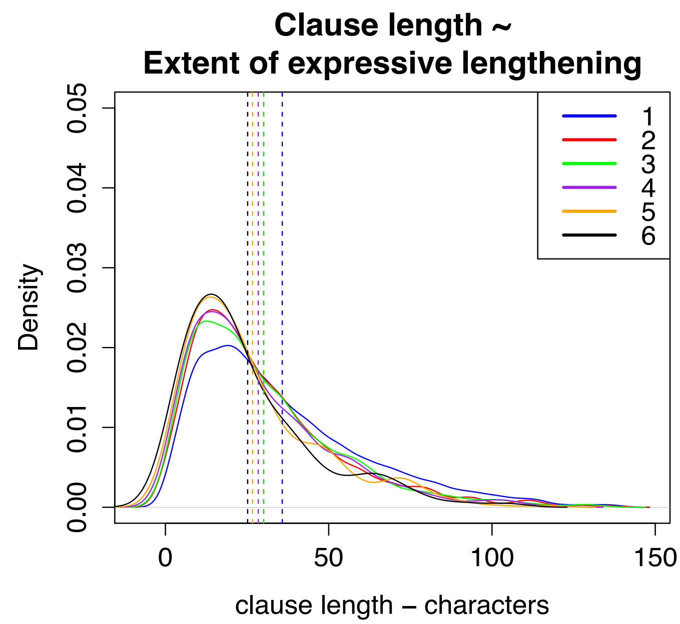

Figure 2 Character count by degree of expressive lengthening

It's obvious that character and token counts should correlate positively, however these measures tell us nothing about the nature of the lexical items we are counting. If character counts decline more rapidly than token count as a function of SFL we should observe longer tokens on average in clauses with higher degrees of SFL. We construct mixed-effects models as before, this time including average token length as our dependent measure. Degree of expressive lengthening does not appear to improve model predictions: chi-square(1, 3.1) (p=.079) and average token length does not appear to differ qualitatively.

Thus far our dependent measures have been fairly blunt. We have only considered counts and average token length to differentiate between clauses with varying degrees of SFL. Ultimately, we are most interested in variations that may exist in the *lexical content* of those preceding clause tokens. One approach we might have taken would incorporate information theoretic notions around information content (surprisal) to assess the lexical variations in clauses preceding SFL devices. This kind of measure would require constructing conditional word probabilities and a corresponding language model from our corpus. Given scope limitations we turn to other measures that might provide insight into the nature of the content in the clauses preceding our SFL (!). We incorporate an additional data set into our analysis to better assess whether there are qualitative differences in the linguistic content of our clause groups. Given observations in the work cited earlier, a natural route would be to assess our clauses along emotionality/sentiment dimensions.

BRM – Norms of valence, arousal, and dominance

Warriner, Kuperman and Brysbaert's "Norms of valence, arousal and dominance for 13,915 English lemmas" (2013) extends previous work by researchers trying to categorize and measure affective word meaning. For each of nearly 14,000 word lemmas, measurements of three components of emotion - *valence* (pleasantness of a stimulus), *arousal* (intensity of emotion provoked) and *dominance* (the degree of control exerted by a stimulus) were collected. We incorporate this data set into our analysis, constructing average emotionality ratings for each clause. 6,304 clauses were ineligible for this analysis because of data sparsity. That left us with 20,320 clauses that had at least one token match in the emotional ratings database.

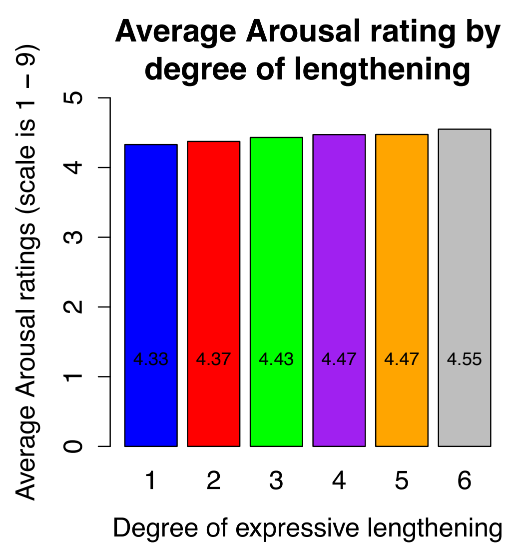

Figure 3 Average Arousal ratings

Each clause was assigned an average emotionality rating for Arousal, Valence and Dominance. These averages account for each token match in the norms database by simply taking the mean across all matches if multiple matches exist in a given clause. While the effect appears slight, we see Arousal ratings increase monotonically as a function of SFL (Figure 3). We conducted a Student's t-test for average Arousal ratings for the non-lengthened clauses vs clauses with any degree of lengthening (2+). There was a significant difference in mean emotionality ratings $t(8529)=6.2773$ ($p<.001$) between these groups. Additionally, we tested the performance of mixed-effects models with degree of lengthening as our fixed effect with random slope and intercepts for account and account by degree of lengthening interactions. Model predictions for Arousal ratings are improved with the addition of degree of lengthening, chi-square(1, 6.6) ($p=.01$), suggesting that there may be a correlation between the use of more emotional words in clauses with greater degrees of expressive lengthening.

While expressive lengthening does not seem to have an affect on average token word length, it does seem to correlate with shortened clauses that encode more emotional content.

Co-occurrence of expressive lengthening

Perhaps the most obvious channel to explore is whether SFL (!) influences expressive lengthening of the tokens that precede it. We parse our clause data for tokens that contain 3+ repetitions of the same character. Each clause is assigned a count of lengthened tokens so we can examine averages across our grouping variable. Brody & Diakopoulos (2011) found a correlation between the subjectivity of a token and its tendency to be lengthened. They inferred that the more lengthened forms a word has, the more likely it (the base lemma) is to be subjective. We estimate the likelihood of observing a lengthened token preceding a SFL (!) by dividing the number of lenghthened tokens in a clause by the total number of tokens. We then average across all of our clauses in which there is at least one lengthened token, by degree of SFL. We see a dramatic increase in the likelihood of observing a lengthened token as a function of SFL (see figure 4). Clauses with a six degree lengthening are nearly ten times more likely to contain a lengthened token (8.8%) than clauses with only a single (!) (0.95%). Of course, it could be the case that this effect is primarily driven by @accounts that are more likely to lengthen *everything*. We ran a mixed effects model to predict likelihood of encountering a lengthened token in the preceding clause given degree of SFL with account and account by degree of lengthening random effects. Model predictions are significantly improved when we incorporate a fixed effect for degree of sentence final lengthening, chi-square (1, 21.44) ($p < .001$).

While it appears that an increase in the degree of sentence final (!) lengthening influences the occurrence of other lengthened tokens, does it also influence the *degree* of lengthening in these other lengthened tokens? Put differently, a potential hypothesis might posit that the clause "I looooooved it!!!!!!" is preferred over "I loooved it!!!!!!" because speakers want to balance the degree of sentiment across the linguistic signal. In order to assess this we reduced our corpus to clauses that contained at least one lengthened item. This reduced our total number of clauses to n=879. We generated an average for length of lengthened tokens (in characters) within individual clauses and grouped clauses by degree of SFL. We assessed the effect on lengthening of preceding-clause-lengthened-tokens by degree of SFL (!) (see Figure 5). Degree of lengthening in SFL (!) is correlated with increased lengthening of lengthened tokens in the preceding clause. Mixed-effect model predictions of preceding token length are significantly improved with the inclusion of the degree sentence final (!) as a fixed effect

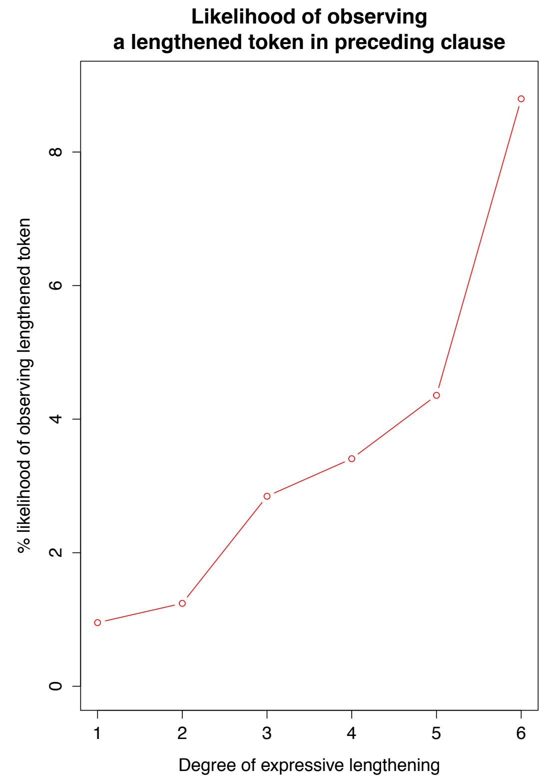

Figure 4 percent likelihood of observing lengthened token in preceding clause

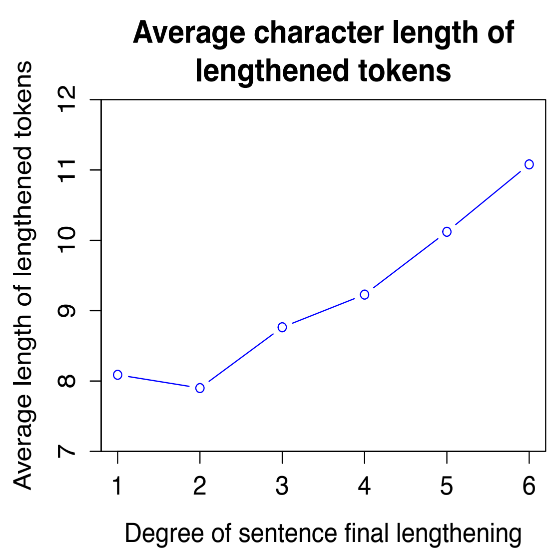

Figure 5 Degree of lengthening for expressively lengthened tokens in preceding clause

chi-square(1, 13.41)($p < .001$).

Brief Summary

In summary, we've observed some interesting associations between degree of sentence final (!) lengthening and content in the preceding clause. We enumerate these observations below.

As one increases the degree of SFL (!), the preceding clause:

1) character and token counts *decrease* as a function of SFL
2) average emotional content of tokens *increase* as a function of SFL
3) occurrences of other expressively lengthened words *increase* as a function of SFL
4) average length of lengthened tokens *increase* as a function of SFL

Theoretical implications

Thus far we have presented evidence that *something* is going on in clauses that have been expressively lengthened with sentence final (!), but have not placed these observations in theoretical context. We now introduce two theoretical frameworks to make testable predictions for work moving forward.

Uniform Emotion Density (UED)

One potential explanation borrows the general idea that optimal speakers strive to maintain uniformity of information across a linguistic signal (Levy & Jaeger 2007 - Uniform Information Density). An adaptation of this approach might posit that speakers strive to maintain consistency in the degree of emotional content in a similar fashion. This kind of hypothesis, which we will call Uniform Emotion Density, predicts that an utterance a)"I loooooved it!!!!!!" would be favored over b) "I loooooved it!" because the single sentence final (!) is inconsistent (given its lengthened alternatives) with the degree of sentiment expressed by "loooooved." Similarly, this hypothesis predicts that the utterance c) "I loved it!!!" should be favored over d) "I liked it!!!" because "loved" is a more arousing term than its alternative "like." These predictions align with our findings that increases in the degree of SFL are correlated with increased emotional content, increased occurrences of lengthened tokens, and increased lengthening of lengthened tokens.

However, UED makes no predictions as to why we saw reductions in character and token counts. These length reductions seem to fit with the intuition that higher degrees of emotionality reduce clause length and are produced in shorter durations. For example, a tweet clause such as "It's beeeen foreveeerrrr, caaan't waait to seeee youuuu!!!!!!" begins to feel exhausting – almost a kind of emotional informativity overload. UED only predicts that speakers should try to maintain consistency over the linguistic signal. A modified UED approach might posit an optimal cumulative emotional informativity amount for a given utterance such that speakers try to maintain uniformity *and* orient around this optimal cumulative emotional information amount. This hypothesis could be assessed in additional studies using similar corpora, however another approach could utilize human judgments around the naturalness / awkwardness of utterances that deviate from hypothetical uniform emotion rates.

Prosodic Imitation Hypothesis (PIH)

Another account might frame these observations in terms of natural speech production; we'll call this the "Prosodic Imitation Hypothesis." This seems to be a fairly widespread intuition in the literature and deserves to be expressed more formally. In "Robust Emotion Recognition using Spectral and Prosodic Features," author K.S. Rao describes the affect of emotions on "basic prosodic parameters such as pitch energy and duration." (p.47) In particular, pitch and energy values are greater for active emotions such as anger or excitement, but have shorter duration. This observation aligns with a) the nature of our tweets, which tended toward more active emotions (because the SFL (!)) and b) the overall shortening of clauses as a function of increases in SFL. Expressive lengthening may be a very literal transcription of vocal prosodic contours and follow the same apparent prosodic conventions as spoken language. Simply put, Tweeters may be translating prosodic emphasis as best they can to convey the degree and nature of sentiment intended. For instance, if we consider the fictional tweet, given earlier: "It's beeeen foreveeerrrr, caaan't waait to seeee youuuu!!!!!" we can imagine physical constraints (lung capacity) disfavoring this kind of utterance simply from a speech production perspective (let alone the intensity of the emotion). If Tweeters are literally transcribing an auditory signal, this utterance should be disfavored on biological grounds.

PIH aligns with our observations of shorter preceding clause length and indirectly with the increases in occurrence and lengthening of lengthened tokens as a function of SFL. This hypothesis would be fairly hard to address quantitatively, however, as its central tenant is that people tweet as they speak. We might dig a little deeper into naturalistic studies of speech production and try to back into the corresponding text-based equivalents for prosodic indicators. Alternatively, future experimental studies could test the productivity of this hypothesis by actually having subjects listen and transcribe audio recordings to text. Manipulations could include the context of the transcription: transcribe for a legal proceeding, for a tweet, for a TV show, etc., to assess the strategies subjects might use to directly transcribe the prosodic contours of the audio recording.

Closing remarks

In addition to investigating the hypotheses just presented, future work should try to replicate these findings using larger data sets and data from less-public figures. The fact that our corpus was made up of celebrity tweets may have biased our findings in ways currently unknown. Additionally, future studies may want to investigate other common SFL devices such as (?) and (.). Lastly, to the extent that we can utilize existing sentiment analysis tools, it would be interesting to extend these ideas to emotions other than the more active emotions, which appear to dominate this data set (excitement, anger, enthusiasm).

In this study we have provided evidence that expressive lengthening is a flexible device and *degree* of lengthening is not arbitrary. Rather, degree of lengthening represents an intentional choice by speakers to modulate the intensity of sentiment associated with a message and has clear repercussions for linguistic

content elsewhere in an utterance. I've included the full R code for my statistical analyses, my command line scripts for web-scraping preliminary data processing, and a PDF of this report on Github:
https://github.com/benpeloquin7/Winter_2015/tree/master/Linguist230a/Final_Project

References

Ball, A.E. (2011, July 1). Talking (Exclamation) Points. *The New York Times*. Retrieved from http://www.nytimes.com/2011/07/03/fashion/exclamation-points-and-e-mails-cultural-studies.html?_r=0&pagewanted=print

Bolinger, Dwight. 1989. *Intonation and Its Uses: Melody in Grammar and Discourse*. Stanford: Stanford University Press.

Brody, Samuel and Diakopoulos, Nicholas. 2011. Cooooooooooooooollllllllllllll!!!!!!!!!!!!!! Using word lengthening to detect sentiment in microblogs. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 562–570. Edinburgh, Scotland, UK.: Association for Computational Linguistics.

Doll, J. (2013, February 20). Why Drag it out? *Atlantic*. Retrieved from http://www.theatlantic.com/magazine/archive/2013/03/draggin-it-out/309220/

Doll, J. (2013, February 21). Why Twitter makes us want to Add Extra Letterssss. *The Wire*. Retrieved from http://www.thewire.com/entertainment/2013/02/why-twitter-makes-us-want-add-extra-letterssss/62348/

Eisenstein J, O'Connor B, Smith NA, Xing EP. 2014. *Diffusion of Lexical Change in Social Media*. PLoS ONE 9(11): e113114. doi:10.1371/journal.pone.0113114

K.S Rao and S.G Koolagudi, *Robus Emotion Recognition using Spectral and Prosodic Features,* SrpingerBriefs in Speech Technology, DOI: 10.1007/978-1-4614-6360-3_3. 2013

Levy, R. & Jaeger, T.F. 2007. *Speakers optimize information density through syntactic reduction*. Retrieved from http://www.bcs.rochester.edu/people/fjaeger/papers/E_LevyJaeger07_full.pdf

Lynton, E. (2014, April 22) What Your Text Punctuation Really Means. *The Harvard Crimson*. Retrived from http://www.thecrimson.com/article/2014/4/22/a-guide-to-texting/

Merritt, A (2013, April 3). Text-speak: Language evolution or just laziness?. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/education/educationopinion/9966117/Text-speak-language-evolution-or-just-laziness.html

Schnoebelen, Tyler. 2012. *Do You Smile with Your Nose? Stylistic Variation in Twitter Emoticons*. University of Pennsylvania Working Papers in Linguistics, vol 18.

Warriner, A.B., Kuperman V., Brysbaert M. 2013. *Norms of valence, arousal and dominance for 13,915 English lemmas*. Behavior Research Methods, 45, 1191-1207

Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*. ACL, Stroudsburg, PA, USA, HLT '05, pages 347-354