Do You E-Care? Analyzing the Impact of Conversational Agreement in Online Customer Service

Francisco Ordenes, University of Massachusetts - Amherst, USA
Dhruv Grewal, Babson College, USA

To assess online consumers’ service evaluations, many firms use automated sentiment analysis, yet this method does not capture dialogues dynamics between consumers and employees. Drawing on relational communication and text mining, we assess the non-linear effect of the dialogue agreement, content positivity, feedback type, and their interaction, on conversational sentiment.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/1024264/volumes/v45/NA-45

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
Impacts of Language on Consumer Behavior

Chairs: Ernest Baskin, Saint Joseph’s University, USA
Peggy Liu, University of Pittsburgh, USA

Paper #1: The Unexpected Implications of Product Descriptors on Product Perceptions
Ernest Baskin, St. Joseph’s University, USA
Peggy Liu, University of Pittsburgh, USA

Paper #2: Make Your Tweety Bird Tweet: Use of Textual Paralanguage in Brand and Spokescharacter Online Communications
Andrea Webb Luangrath, University of Iowa, USA
Joann Peck, University of Wisconsin – Madison, USA
Victor A. Barger, University of Wisconsin – Whitewater, USA

Francisco Villarroel Ordenes, University of Massachusetts-Amherst, USA
Dhruv Grewal, Babson, USA

Paper #4: Asymmetry in Emotion Language Is Consequential For Evaluative Judgments
Chiara Longoni, Boston University, USA
Geeta Menon, New York University, USA

SESSION OVERVIEW

Since the days of the Whorfian hypothesis (Whorf 1956), much research has been done on how language shapes both cognition and emotion. The role of language in affecting consumer behavior is less well-understood, however. Yet language is highly important in consumers’ lives, whether in the form of descriptors on product packages or even in terms of consumers’ own linguistic engagement with brands. In this session, we have assembled four papers that collectively examine how different facets of language shape consumer behavior. As a group, these papers raise two main questions: how can language be used to influence consumer cognitions and emotions, and how can marketers take advantage of these influences and structure the language content of their promotions to take advantage of consumer biases.

First, Baskin and Liu examine the effects of unknown language descriptors on product perceptions, finding that consumers make inferences about products, both in terms of price and taste, despite not understanding the descriptor’s meaning and having no a priori associations with the descriptor. They present nine studies showing how price perceptions are increased while taste perceptions are decreased when consumers are uncertain about the meanings of product descriptors.

Second, Luangrath, Peck, and Barger examine language use on Twitter, in particular, textual paralanguage and explore how its usage affects perceptions of brand competence and warmth. They present four studies showing that introducing non-word items into a brand’s messaging can decrease competence perceptions of the brand.

Third, Ordenes and Grewal examine language through automated text mining on a sample of online dialogues (N=2,084) between customers and employees of various companies including Amazon and Tesco as a way of understanding the impact of interactional control on the service resolution. They use this data to show that, beyond sentiment, interactional control within a dialogue has a non-linear effect (diminishing returns) on the service resolution. In addition, they show that the non-linear effect of interactional control is stronger for dialogues in which the trend of employee positivity across messages is increasing.

Finally, Longoni and Menon examine emotional language and its effects on evaluative judgments. Across six studies, they show that the use of positive emotion words lowers emotion intensity and product evaluation.

Overall, all of these papers are at advanced stages, and a range of methods—from lab experiments to text mining—are employed. By examining the multi-faceted impact of aspects of language on consumer perceptions, we hope that this session helps to spur on a new wave of research that delves deeper into aspects of language and how they affect consumer behavior. We expect this session to appeal to researchers interested in the topics of language, word-of-mouth, and online marketing.

The Unexpected Implications of Product Descriptors on Product Perceptions

EXTENDED ABSTRACT

Imagine going to a restaurant and seeing meaningless descriptors on the menu to describe the food items. For example, envision the items, “fiducated cornbread” or “faldered hamburger steak” as you peruse the menu. Even though you are unaware of the meaning of “fiducated” or “faldered” (in fact, they have no meaning), how might they change your perception of the item they are associated with? In particular, would the descriptors affect your price or taste perception of the menu items?

In this research, we examine the effects of meaningless descriptors on perceptions of products, both in a restaurant menu context, as well as more broadly. In particular, we examine how adding descriptors, even meaningless ones with no inherent meaning, affect product price perceptions and in the case of food, taste perceptions. We show that meaningless descriptors increase price perceptions through increasing consumer uncertainty about the product. Thus, consumers may assume the product is more exceptional. Regarding taste, we draw from the notion that consumers are risk averse both generally and especially in the food domain to predict that the uncertainty about product characteristics may lead to increased risk perceptions and thus a decrease in predicted taste.

We present a series of nine studies to test these hypotheses. First, we pretest meaningless descriptors in order to ensure that they are, in fact, unknown to the majority of our participants. In study 1, we establish our main price effect in the food domain. In addition, we show that meaningless descriptors actually hurt predicted taste and enjoyment and have negative downstream consequences for food purchase intentions. Next, in studies 2a and 2b, we replicate our main price effect and show that it holds both inside and outside the food domain. We also find evidence against processing fluency as an alternative explanation by showing that descriptor complexity, one measure of fluency, does not affect price perceptions in 2a. Study 2b further rules out fluency by using a different manipulation of fluency, hard-to-read font, and showing that meaningless descriptors affect price and taste perceptions over and above the effect of fluency. Next, study 3a, shows our serial mediation mechanism for price perceptions in that meaningless descriptors create uncertainty about the general characteristics of a product which then leads people to predict that the product will be more exceptional thus leading to
increased price perceptions. Study 3b shows our serial mediation mechanism for taste perceptions in that meaningless descriptors decrease taste perceptions by increasing uncertainty and increasing potential product risk therefore leading to decreased taste perceptions. Additionally, we look at two potential boundary conditions for the price effect. Study 4a provides additional support for the role of uncertainty by showing that, when product uncertainty is decreased by showing participants a photograph of the product in question, the effect is attenuated. Study 4b then provides additional support for the role of product exceptionality by showing that special occasion products, which are already exceptional, also act as a moderator for our effect. Finally, study 5 shows that the price effect obtains even when price is given to the consumer.

In summary, this research contributes to a number of literatures. First, it contributes to the area of linguistic psychology and the price perception literature writ large. The literature in this area has not been able to disentangle the effects of descriptors as all studies that we are aware of confound the existence of descriptors with the meaning that they give to the products that they describe. While we do not believe that the meaning gained from descriptors has no effect on perceptions [many studies have shown that it does in areas that we find a negative effect in, such as taste (Wansink, Painter, and Van Ittersum 2001)], we do believe that a large portion of the pricing perceptions effect may arise from the existence of the descriptor rather than anything related to the content of the descriptor itself and its relation to the product being described. In addition, we contribute to the linguistics literature by showing an area where a lack of linguistic meaning is taken by consumers as a cue to a change in their visualization about the product being described. We grant further evidence that language is a factor in forming cognitions and biases in everyday life. Finally, we also contribute to the nascent field of textual paralanguage by showing a new instantiation of meaning that supplements the literal meaning of the written language inherent in the descriptor.

Make Your Tweety Bird Tweet: Use of Textual Paralanguage in Brand and Spokescharacter Online Communications

EXTENDED ABSTRACT

For years, Kellogg’s has been utilizing textual paralanguage in its advertising for Frosted Flakes cereal, printing “THEY’RE GR-R-REAL!” next to the image of Tony the Tiger. Through repetition of the letter ‘r’, Kellogg’s communicates that the word “great” is to be spoken in a drawn out, tiger-like growl. Recent research defines textual paralanguage (TPL) as written manifestations of nonverbal audible, tactile, and visual communication (Luangrath, Peck, and Barger 2017). In this research, we examine the effects of TPL use by brands and brands’ spokescharacters on social media.

There is growing interest in online communications and the insights to be gained from text-rich data. Much of the existing work on language in consumer research has focused on word usage, such as explanatory phrases (Moore 2015), assertive advertisements (Zemack-Rugar, Moore, and Fitzsimons, forthcoming), refusal words (Patrick and Hagtvedt 2012), vowel sounds (Lowrey and Shrum 2007), figurative language (Kronrod and Danziger 2013), closeness-implying pronouns (Sela, Wheeler, Sarial-Abi 2012), and meaningless words (Liu and Baskin, working paper). Recent work has been instrumental in advancing sentiment analysis beyond dichotomous classifications of positive/negative to a more nuanced consideration of sentiment strength, implicit meanings, and patterns of sentiment (Villarroel-Ordenes et al. forthcoming).

We approach the study of language from an alternative perspective, shifting focus from the meaning and interpretation of specific words to the ways in which the words are communicated. While word-based language assessment is a critical component of linguistic analysis, we believe that aspects of messages conveying properties of speech and communication context are important as well. TPL in text is analogous to nonverbs in face-to-face communication.

Study 1a examines how prominent U.S. brands use TPL in text-based marketing communications. Two sets of brands were selected for this study: Forbes’ list of “The World’s Most Valuable Brands” (Badenhausen 2013), including brands such as Nike, Google, and Disney, and Time’s list of “The 13 Sassiest Brands on Twitter” (Grossman 2014), including brands such as Hamburger Helper, Orbitz, and Old Spice. A Python program was written to collect brand tweets and TAMS Analyzer was used to code the tweets for TPL. Of 5,214 brand tweets, 1,011 (19.4%) contained TPL. Use of TPL was markedly higher by the sassier brands (32.1%) than the Forbes’ brands (12.7%). Using these lists as proxies for brand warmth and competence, we hypothesize that sassier brands are likely to be perceived as warmer while Forbes brands, which are more traditional and conservative, are likely to be considered more competent.

Study 1b explores the relationship between TPL and consumer perceptions of warmth and competence. Participants on Amazon’s MTurk (N=1,847) were randomly presented with one brand name from the list of Forbes’ and Time’s brands (study 1a) and asked for their perception of the brand’s warmth (Aaker et al. 2010) and competence (Aaker 1997). Results indicate that the Forbes’ brands are perceived to be more competent ($M_{ForbesComp} = 3.78, M_{TimeComp} = 3.29, F(1,1846) = 165.28, p < .001$), but less warm than Time’s sassier brands ($M_{ForbesWarm} = 3.02, M_{TimeWarm} = 3.27, F(1,1846) = 25.70, p < .001$). Pearson correlation coefficients reveal that TPL use correlates positively with perceptions of brand warmth (.16, p < .01) and negatively with perceptions of brand competence (-.05, p < .05).

To test this relationship, Study 2 investigates whether TPL affects perceptions of brand warmth and competence. Ninety-eight undergraduate students participated in an online study and were asked to imagine that a brand tweets the following: “Big things come in small packages [BIG, Awww, ☺].” Tweets either contained TPL or did not. Results reveal that a tweet containing TPL as compared to a tweet without TPL does not significantly affect perceptions of brand warmth ($M_{NoTPL} = 3.52, M_{TPL} = 4.06; F(1, 97) = 2.89; p = .09$) but negatively affects perceptions of brand competence ($M_{NoTPL} = 4.73, M_{TPL} = 4.09; F(1, 97) = 4.58; p = .04$).

In Study 3, we investigate whether a brand’s spokescharacter can mitigate the negative effects of TPL on perceptions of brand competence. Spokescharacters are often created to imbue brands with personality (Fournier 1998), which may give them leeway to be more informal, humorous, and relatable. A 2 (TPL/no TPL) x 2 (brand/brand mascot) study was conducted with 476 participants on Amazon’s MTurk. Participants were asked to imagine that a [brand/ brand mascot] tweets the following: “A penny saved is a penny earned [!!!!!!, Yippee!, *thumbs up*].” The avatar displayed next to the tweet was either an image of a brand mascot (i.e., a frog) or a neutral brand image (i.e., a green circle). Results indicate that perceptions of brand warmth are not significantly affected by TPL ($M_{NoTPL} = 4.89, M_{TPL} = 4.69, F(1,475) = 2.48, p > .10$). Similar to study 2, we find that TPL negatively affects brand competence ($M_{NoTPL} = 4.96; M_{TPL} = 4.62; F(1, 475) = 6.61; p = .01$). A significant interaction between the spokescharacter and the TPL condition ($F(1, 475) = 3.79; p = .05$) suggests that when a brand tweets using TPL, perceptions of brand competence decrease significantly ($M_{BrandNoTPL} = 5.13, M_{BrandTPL} = 4.52; F(1, 472) = 9.72; p = .002$). However, when
a spokescharacter tweets using TPL, there is no significant positive or negative effect on perceptions of brand competence ($M_{\text{Frog/NoTPL}} = 4.80; M_{\text{Frog/TPL}} = 4.71; F(1, 472) = .21; p = .65$). In addition, we rule out the possibility that the presence of a smiling frog led to greater positive affect ($M_{\text{Frog}} = 5.54, M_{\text{Brand}} = 5.67, F(1,475) = 1.49, p > .10$).

This research suggests that, when using TPL in online communications, a brand can avoid negative effects on brand competence by tweeting from a spokescharacter account rather than the brand’s main Twitter account. We reason that consumers expect an informal character to use more informal language, but when a brand, a more formal source, uses informal language, it makes the brand seem less formal and thus less competent. As social media platforms continue to add to the range of available text-based symbols, images, and stickers, Frosted Flakes should leave the growling to Tony.

Do you e-care? Analyzing the Impact of Dialogue Dynamics in Online Service Resolution

EXTENDED ABSTRACT

Consumers are increasingly providing feedback and interacting with brands in digital communicational channels. This has given rise to the growing concept e-care, that is, customer service through social online channels such as Twitter and Facebook. For example, the volume of tweets targeted to brands has doubled, and the percentage of users visiting Twitter for customer services increased by 70% during 2013–2014 (McKinsey 2015). Many companies use sentiment analysis—an automated process for deciphering the sentiment contained in text—to assess the sentiment of these conversations, yet this method is limited by its inability to capture dialogue dynamics resulting from consumer-employee interactions. Limited consumer research explicates how language patterns at a dialogue level, can inform the use of sentiment analysis for consumer research (Ordenes et al. 2017).

Relational Communication Theory (Soldow and Thomas 1984) posits that every message comprises content (i.e., substance of a message) and form (i.e., structure of the content). Whereas content relates to the informational aspect, form serves the interpretation of the message in terms of the interactional control within a dialogue. Through the form of the message, an interactor positions him- or herself toward dominance, deference, or equality. The interactors’ reciprocal propositions entail various combinations of control positions which have an influence on the effectiveness of the interaction.

By theorizing on relational communication (Soldow and Thomas 1984) we posit that the interactional control has non-linear effect (diminishing returns) on the service resolution (hypothesis 1). Then, in line with research on affective meaning (Warriner and Kuperman 2014), we suggest that an increasing trend in positivity across employee messages, has a positive effect on the service resolution (hypotheses 2). Finally, in line with research on the interplay between content and style (Ludwig et al. 2013) we posit that the positive effect of interactional control is stronger when positivity across employee messages increases (hypothesis 3).

A dataset of Twitter and Facebook, consumer initiated dialogues with an employee (all complaints), between November 2014 and May 2015, was scraped and pre-processed (e.g., duplicates deleted and data was structured at a dialogue level) by using Knime Analytics 3.2. We identified 2084 dialogues of the retail firms Amazon, Tesco, and Walmart, as well as the car manufacturer Ford.

The dependent variable, service resolution, was coded manually for each dialogue. Two coders coded the service resolution in terms of whether the customer ended up the dialogue satisfied (1) or dissatisfied (0). Coder inter-reliability was assessed using Krippendorff’s alpha, resulting in a value of 0.82.

Interactional control was operationalized using the Relational Communication Framework (Soldow and Thomas 1984). Two coders manually coded each sentence for every message of a dialogue. Each sentence was assigned to a Digit 1 (message author), Digit 2 (grammatical form), Digit 3 (response mode). Then a control position relative to the previous sentence from the customer and the employee was determined (i.e., 1 = deference, 2 = equality, 3 = dominance). Finally, to determine the overall interactional control, an average was calculated from the sentence control positions (SCRL) across the dialogue (1 being the lesser control and 3 being the greater control). Coder inter-reliability was examined by using the Krippendorff’s alpha resulting in .99, .89 and .84 for digits “1” “2” and “3” respectively.

To assess the trend of employee positivity across dialogue messages, LIWC (Linguistic Inquiry and Word Count) text analysis software was used (Tausczik and Pennebaker 2010). In line with Ordenes et al (2017), a single positivity trend per dialogue was obtained by computing the slope of positivity across the employee messages within a dialogue. Several control variables such as the number of messages exchanged, picture presence, average employee response time (minutes), social media type, industry type, third participant, and message arousal (Ordenes et al. 2017; Schweidel and Moe 2014; de Vries, Ginsler and Leeflang 2012) were included.

In line with our hypotheses, we specified a hierarchical logistic regression approach to estimate 1) the non-linear effect interactional control, 2) employee positivity trend, and 3) their interaction effect. Model 1 represents the baseline of the hierarchical regressions in which we only evaluated the effect of basic sentiment analysis on the service resolution. Model 2 included the nonlinear effect of the interactional control on the service resolution. Our findings supported H1 by finding a diminishing returns relationship of interactional control. The results of Model 3 supported the proposition of H2. A higher level of employee positivity across sentences increases the likelihood of a satisfactory service resolution. Finally, Model 4 incorporated the interaction effect between dialogue interactional control and employee positivity trend. Our results partially supported H3 by confirming that dialogues in which employees increase their positivity across sentences, strengthen the relationship between interactional control and the service resolution.

By using the lenses of relational communication theory, the present study informs research on online consumer-employee dialogues. Theoretical contributions in the area of consumer language and managerial implications concerning the development of conversational analytics’ methods are offered.

Asymmetry in Emotion Language Is Consequential For Evaluative Judgments

EXTENDED ABSTRACT

Imagine two football fans Tweeting about their team winning the Super Bowl; one fan uses 2 positive emotion words whereas the other fan uses 6 positive emotion words. Which fan will rate the victory more emotionally intense? Which fan will enjoy the victory more? Content-focused models of judgments (Higgins, 1996) predict that the fan who used 6 positive emotion words experiences the victory more intensely, and therefore enjoys the win more. Contrary to this prediction, we show that for positive (but not for negative) emotion words, the effect is reversed: the greater the number of positive emotion words a person associates with an event, the lower the emotion intensity and evaluation of an event.
We ground our predictions in diverse literatures spanning emotions and how emotions are conceptualized (Barrett, 2007; Lindquist et al., 2006). We also rely on several lines of research that have examined production and development of emotion language (i.e., the vocabulary we use to refer to emotions) in both children and adults. These lines of research are suggestive — although not testing directly — that positive emotion words are less readily available than negative emotion words. For example, in 4-month olds, negative emotions such as anger and fear grab attention more than positive emotions such as happiness, as measured via gaze orientation, recall, and recognition (Montague and Walker-Andrews, 2001). Young children spontaneously and more frequently talk with their parents about the causes of unpleasant emotions than the sources of positive feelings (Dunn et al., 1987; Dunn et al., 1991), and are more prone to reminisce about negative past events than positive ones (Miller and Sperry, 1988). The qualities of parent-child narrative discourse and interactions also suggest that positive emotions are less salient than negative ones. For example, parental questioning, the act of parents enquiring, “what is the matter?”, “what is bothering you?”, is more likely to prompt children to elaborate and think about negative than positive emotions (Lagattuta and Wellman 2001).

Together, these lines of enquiry indirectly suggest that negative emotions are more salient than positive emotions. We test this hypothesis directly and examine its implications for evaluative judgments in a variety of domains. Specifically and across 6 studies, (a) we show that emotion language affects evaluative judgments, (b) we document an asymmetry in positive (vs. negative) active emotion vocabulary, (c) we show that this asymmetry has hedonic consequences across various contexts, and (d) we show that the effect of emotion language on judgments is mediated by the subjective experience of effort associated with emotion generation.

First, we directly test the hypothesis that there is lower chronic accessibility of positive (vs. negative) emotions words. In study 1, we asked participants to generate as many positive or negative emotion words as they could in a set time frame. Participants were able to generate fewer positive emotions than negative emotions, suggesting that negative emotion words dominate the working emotion vocabulary of the average individual (p < .001). This difference reflects the English vocabulary (406 positive emotions vs. 499 negative emotions).

Second, we show that this differential accessibility of positive (vs. negative) emotion words is consequential for evaluative judgments in various domains, including recalls of personal experiences, evaluation of photos and blog posts, and product reviews. In studies 2 and 3 we measured the relationship between the number of positive (negative) emotion words used when recalling a personal event, and how intensely positively (negatively) one felt about that personal event. In study 2, participants wrote a post about a personal event that was meaningful to them, had taken place in the recent past, and in which they were happy or unhappy (order counterbalanced). In study 3, participants wrote about current positive and pleasant (or negative and unpleasant) events. In both studies, participants who reminisced about happy past life experiences or recalled pleasant current events subsequently felt less happy the greater the number of positive words they had included in the recall (for past events, p = .02; for current events, p = .01), whereas the number of negative emotion words they used did not affect mood ratings.

The same effect emerged for evaluation of photos and blog posts in studies 4 and 5. Participants who described positive stimuli (a photo of a smiling child, a photo of the zoo) using 6 positive emotion words rated these stimuli less favorably compared to participants who used 2 positive emotion words (p = .001 and p < .001 respectively), with the difference in evaluations fully mediated by the effort associated with emotion generation (indirect effect bootLLCI: 0.044 bootULCI: 0.745; bootLLCI: 0.012 bootULCI: 0.518 respectively). However, the number of negative emotion words used to describe negative stimuli did not predict evaluation, nor did effort in emotion generation mediate evaluation.

Finally, in study 6 we tested the proposed phenomenon in the context of product reviews. Participants who used 8 positive emotion words as hashtags after reviewing a cafe subsequently gave the establishment lower star ratings compared to participants who used 3 positive emotion words as hashtags (p = .001), unless participants assigned the star rating before reviewing the cafe.

Overall, across several studies, measurements, and contexts, we present a novel framework that advances our understanding of emotions and emotion language, and furthers our knowledge of the consequences of emotions on judgments. Importantly and from a practical standpoint, this research shows how encouraging consumers to express positive affective states, like it is often customary online by use of hashtags, tweets, and short posts, may have aversive consequences on evaluations.

REFERENCES

Paper #1

Paper #2


Paper #3


Paper #4


Miller and Sperry (1988), The socialization of anger and aggression, Merrill-Palmer Quarterly, 33, 1-31
