

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Breadth of the Stars: Exploring Adjectival Breadth in Online Reviews

Permalink

<https://escholarship.org/uc/item/6cp1k0sq>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Lin, Lin

Dale, Rick

Stroessner, Steve

Publication Date

2024

Peer reviewed

Breadth of the Stars: Exploring Adjectival Breadth in Online Reviews

Lin L. Lin (llin001@g.ucla.edu)

Rick Dale (rdale@ucla.edu)

Steven J. Stroessner (ss233@g.ucla.edu)

Department of Communication, University of California, Los Angeles, CA 90095, USA

Abstract

Language is fundamental in human cognition and communication, helping us to encode the world around us. Adjectives represent a linguistic form used extensively, particularly in the social domain. Adjectives vary in both their valence and their breadth (e.g., “punctual” is narrower than “dependable”). Variations in adjectival breadth have not been studied extensively but may have significant consequences across various domains. The present study explores how subtle distinctions in adjective use may relate to descriptions of experiences that people share. To assess linguistic breadth in online communication, we examine whether online reviews with different star ratings are associated with differences in adjectival breadth. Through an analysis of over 200,000 reviews from Amazon digital music (Study 1) and Yelp restaurants (Study 2), we find evidence that linguistic desirability and breadth of adjectives in reviews positively correlate with their ratings. Specifically, higher-rated reviews tend to use broader and more desirable adjectives. However, this relationship varies between product categories, with high-rated music reviews showing increased linguistic breadth and desirability, while top-rated restaurant reviews demonstrate a decrease in breadth. This paper contributes to understanding linguistic breadth in social media contexts, highlighting how subtle language variations in evaluations can reflect different cognitive and communicative processes.

Keywords: social cognition; linguistics; language and thought; language understanding; big data

Introduction

Imagine a lab member showed up on time for a meeting. How would you describe this action in one word? “Punctual” or “Dependable”? Now imagine an intern who missed a deadline for an important project. Would you describe the intern as “Tardy” or “Irresponsible”? Moreover, what if those individuals are people you know well and like versus someone you have just met or dislike?

Subtle linguistic factors can play crucial roles in communicating about various aspects of the external world. They can convey information about communicators and reveal opinions, attitudes, and even prejudicial biases. People regularly exhibit subtle biases in their language, often without conscious awareness or intent. The linguistic category model (LCM; Semin & Fiedler, 1991) was developed to account for variations in linguistic descriptions that produce different understandings of events and people. The LCM contends that an identical event or behavior can be described variably. Observed a person kicking another individual might be described using narrow concrete descriptive action verbs (e.g., kick), broader interpretive action verbs (e.g., attack), or

very broad state verbs (e.g., hate). In the LCM, adjectives (e.g., aggressive) represent the broadest linguistic category.

Research using the LCM and other theoretical advancements has shown that linguistic variations in verb abstractness represent a nuanced but crucial aspect of language, affecting consequential judgments and actions. For example, Wigboldus et al.’s (2000) research on linguistic expectancy bias (LEB) showed that stereotypic expectancies can impact linguistic descriptions of others. Stereotype-consistent events tend to be described at a higher level of abstraction than stereotype-inconsistent actions. For instance, descriptions of a woman who appears dominant in a manner inconsistent with the female stereotype tend to be characterized using relatively concrete language (e.g., “she raises her voice and points fingers”). In contrast, people tend to use more abstract language when describing a man demonstrating the same traits (e.g., “he is aggressive”).

Similarly, research on stereotypic explanatory bias (SEB; Sekaquaptewa et al., 2003) shows that people generally create more elaborate explanations when describing behaviors that are inconsistent rather than consistent with stereotypes. Encountering stereotype incongruent actions appears to trigger explanatory reasoning, manifesting in efforts to rationalize and comprehend incongruity. Moreover, the greater the number of external, situational explanations White participants offered for the stereotype-incongruent events regarding Black individuals (e.g., attributing employment at a prestigious firm to personal connections), the more negative behavior they exhibited toward a Black partner in an interracial interaction.

Furthermore, the differences in the level of abstractness allow message recipients to draw different inferences about a protagonist (Maass, 1999; Maass & Arcuri, 1992; 1996; Maass et al., 1989). As a statement’s level of abstraction increases, more information is believed to be provided about the individual, producing stronger expectations regarding temporal and cross-situational stability. In addition, abstract descriptions produce difficulties verifying the veridicality of statements and imagining disconfirming examples.

Evaluating levels of breadth in adjectives

Identifying cues in linguistic patterns has long been an agenda in applied natural processing in domains ranging from psycholinguistics to education (for reviews, see Graesser et al., 2004, 2011). In the social cognitive realm, evidence supportive of the LCM has accumulated over the decades since its inception. However, this research has focused almost exclusively on differences in verb abstractness (Dragojevic et

al., 2017; Franco & Maass, 1996b; Karpinski & Von Hippel, 1996; Von Hippel et al., 1997). However, implications of variations in the breadth of adjectives have been relatively neglected. All languages have adjectival word classes (Dixon & Aikhenvald, 2004; Evans & Levinson, 2009), and they play a crucial attributional role in English and related languages. Adjectives are central means for communicating detail (e.g., big, old, red) and, when describing social groups, constituting the semantic content of stereotypes (e.g., lazy, rude, aggressive).

Adjectives are the most abstract category according to the LCM, but substantial differences exist in their breadth (Hampson et al., 1987). The causes and consequences of variation in adjectival breadth have yet to be extensively studied. Moreover, adjectival variation may produce critical inferential consequences for both message recipients and senders. Using the example provided earlier, the words "punctual" and "tardy" are narrow descriptors limited to the specific acts performed by the protagonist. In contrast, the descriptors "dependable" and "irresponsible" are broad and imply much information about the protagonist beyond one context, such as their personalities and work ethic. Examining different levels of breadth within adjectives will not only allow a more fine-grained semantic category for research into social-cognitive implications of language usage but also pose new questions about the behavioral and psychological impacts of adjectives on language and communication.

Linguistic abstractness in online review content

Much work on linguistic abstractness in social-cognitive domains has focused on interpersonal and intergroup settings. This research uses linguistic abstractness as a tool to examine the transmission and maintenance of socially shared beliefs at different levels of abstractness (Dragojevic et al., 2017; Franco & Maass, 1996b; Geschke et al., 2010; Karpinski & Von Hippel, 1996; Von Hippel et al., 1997). For instance, Geschke et al. (2010) found that media coverage about minorities is linguistically biased (either narrowly positive or broadly negative) and produces higher levels of prejudice.

While prejudice and stereotypes are critical topics in communication, another intriguing question is whether similar linguistic biases emerge in descriptions of nonhuman entities. Indeed, this domain may serve as a particularly useful arena to test the role of adjectival breadth in descriptive language. Will people vary in linguistic abstractness regarding their characterizations of liked versus disliked nonhuman targets such as experiences? To address this question, we turned to online reviews suitable for large-scale semantic analysis. Reviews are particularly helpful for studying variations in linguistic breadth because reviewers' valenced judgments (i.e., attitudes) are explicit and clear, typically provided through numerical or star ratings. In addition, unlike in the social domain, there are not generally strong normative pressures to view experiences and products equally positively. Consumers are often eager to report their

views of products and experiences that vary in the pleasure they provide. Therefore, online reviews are particularly useful for examining adjectival breadth at different valence levels.

There is currently limited research on the abstractness of descriptions in online reviews (Aerts et al., 2017a; Aerts et al., 2017b; Capocasa, 2014; Krishnamoorthy, 2015; Shin et al., 2019). Only two studies have attempted to investigate the relationship between valence and abstractness of verb usage with mixed results (Aerts et al., 2017b; Capocasa, 2014). By randomly sampling 50 reviews from eight websites (Yelp, Epinions, Qype, Rate It All, Amazon, TripAdvisor, Lens, and Kieskeurig), Aerts et al. (2017b) found that when reviews are written more abstractly, those reviews tend to be more positive. Other relevant studies have investigated factors that influence linguistic abstractness, such as platform design, previous reviews, and customers' prior attitudes toward the brand (Aerts et al., 2017a). Additionally, Capocasa (2014), Krishnamoorthy (2015), and Shin et al. (2019) have explored the relationship between the helpfulness and linguistic abstractness of reviews. In contrast, Capocasa (2014) demonstrated no relationship between the valence of text descriptions and language abstractness by analyzing Amazon ratings of books and appliances.

It is important to emphasize that those studies analyzed linguistic abstractness in terms of the categories provided by LCM. In other words, they all focused on verbs and not adjectival breadth.

Current study

We analyze the relationship between linguistic abstractness and valence (positive and negative) in online reviews by explicitly focusing on the breadth of adjectives. Online reviews constitute an indispensable part of modern lives that regularly guide consumer behavior. We focused on descriptions of experiences, allowing the identification of both conceptually similar effects to those found in the social domain and potentially unique aspects of communication about nonsocial entities.

Two categories of reviews were selected for this research: digital music (Amazon, Study 1) and restaurants (Yelp, Study 2). To quantitatively analyze the breadth of adjectives, this study used a recently developed database (Lin, Dale, & Stroessner, in preparation) containing participants' subjective ratings for 1,214 adjectives on dimensions of familiarity, breadth, and desirability. The list of trait terms was created by combining words used in previous research focused on human traits (Anderson, 1968; Hampson et al., 1987) and other common adjectives gleaned through online searches. During the study, each participant was randomly presented with 100 words, and they rated each word on the dimensions of familiarity, breadth, and desirability. Approximately 100 individual ratings were collected for each word on a 1 to 9 scale. Thus, stable current estimates of these dimensions were derived (Table 1). For example, the words "good" and "normal" have the highest score on breadth, while "unpunctual" and "untalkative" have the lowest score on

breadth. Table 2 illustrates how each word could be classified in terms of both breadth and desirability.

Table 1: Rating of words on breadth (1= narrow, 9= broad) and desirability (1= undesirable, 9 = desirable)

BREADTH		DESIRABILITY	
<i>5 broadest terms</i>	<i>5 narrowest terms</i>	<i>5 most desirable terms</i>	<i>5 least desirable terms</i>
Good 7.33	Unpunctual 2.33	Kind 8.66	Fraudulent 1.30
Normal 7.10	Untalkative 2.37	Healthy 8.61	Malicious 1.36
Fun 7.10	Loquacious 2.62	Happy 8.58	Stupid 1.36
Open 7.04	Punctual 2.65	Successful 8.56	Unlovable 1.40
Creative 7.02	Concise 2.88	Talented 8.56	Narcissistic 1.41

Table 2: Categorization of traits by desirability and breadth

	NARROW	BROAD
DESIRABLE	Punctual	Good
	Well-spoken	Fun
	Honest	Creative
	Articulate	Successful
	Studious	Able
UNDESIRABLE	Fraudulent	Bad
	Abusive	Horrible
	Misogynistic	Evil
	Horny	Nasty
	Hypochondriac	Sick

Study 1

Method

Data Collection A publicly available Amazon review dataset including ratings (on a scale of 1 to 5) and comments from May 1996 to October 2018 (Ni et al., 2019) was used for this study. We randomly sampled 100,000 Amazon product reviews on Digital Music.

Analysis and Results We analyzed the relationship between the rating of music and linguistic features (desirability and breadth) in the review content using regression models. Analyses were conducted by first identifying adjectives used in reviews and assigning desirability and breadth scores from the database (Lin et al., in preparation). Aggregate scores were then computed for each review on dimensions of desirability and breadth.

The median length of those reviews was 39 words with a maximum count of 4477 words, and a minimum count of 1 word. Of all the selected music reviews, each review, on average, contained 2.24 adjectives, with a maximum of 126 adjectives and a minimum count of at least one adjective used in one review. Linguistic desirability ranged from 1.30 to 8.66, with an average score of 6.95 and a standard deviation of 1.68 on a scale of 1 to 9. Linguistic breadth ranged from 2.61 to 7.33, with an average score of 6.09 and a standard deviation of 0.85. To illustrate these results, we provide

examples of reviews that have breadth approximately at the observed average breadth for their rating (Table 3).

A multiple linear regression analysis was conducted to examine the effects of linguistic desirability and star ratings on the linguistic breadth of review content. The model also included an interaction term to investigate whether the impact of linguistic desirability on linguistic breadth varied with star ratings. The regression analysis revealed that linguistic desirability was significantly associated with the linguistic breadth of review content ($\beta = 0.15$, SE = 0.008, $t = 17.18$, $p < .001$, Fig. 1). Star ratings were also a significant predictor ($\beta = 0.32$, SE = 0.009, $t = 35.94$, $p < .001$, Fig. 2). Interestingly, the interaction between product desirability and star ratings was significant ($\beta = 0.017$, SE = 0.002, $t = 9.14$, $p < .001$), suggesting that the impact of linguistic desirability on linguistic breadth varied across valence. Specifically, increases in linguistic desirability for highly-rated music were associated with greater linguistic breadth (Figs. 3a and 3b).

In addition, we employed part-of-speech tagging to verify that the adjectives identified within the reviews were functioning syntactically as adjectives. We resampled 1,000 reviews several times, applied spacyr speech tagging in R, and calculated the percentage of adjectives from our list that were indeed averages in this sample. Across several rounds of this resampling, we found that approximately 85% of those adjectives were indeed used as adjectives.

Overall, analyses revealed some meaningful semantic differences in reviews varying in star ratings. First, the analysis provided important evidence that adjectives are used regularly in communication, especially when describing a product or an experience. Secondly, the positive correlation between star ratings and linguistic desirability validates the robustness of our normed adjectives (Lin et al., in preparation). Specifically, and as expected, highly-rated music was described with more desirable adjectives in the written review, compared to poorly-rated music. Thirdly, the significant interaction term showed that reviews written more positively tended to be more abstract, except for the reviews with five-star ratings. The overall patterns align well with the findings reported by Aerts et al. (2017b).

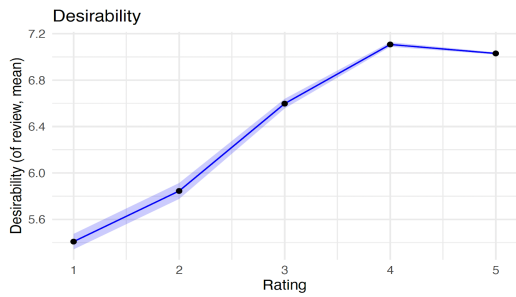


Figure 1: Music ratings by linguistic desirability.

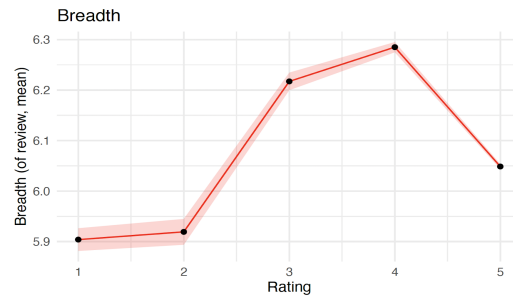


Figure 2: Music rating by linguistic breadth.

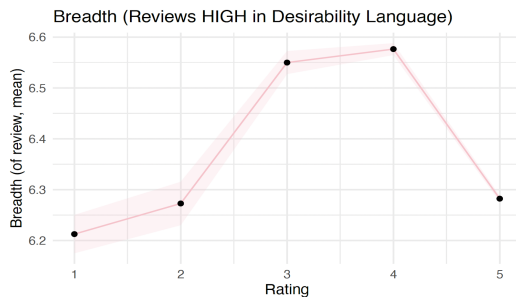


Figure 3a: Rating by breadth for high desirability reviews.



Figure 3b: Rating by breadth for low desirability reviews.

Table 3: Example of Amazon music ratings at the average level of linguistic breadth (6.09).

RATING	REVIEW (ADJECTIVES IN BOLD)
1	" Worst album he has ever made. honestly, I am a Redman fan and I have all of his cds. I hate to do this to Mr Reggie but this cd is a joke. Only two songs are good and the rest are just horrible . Stay away from this album. or buy the two songs that are actually okay . Redman, you owe me for this turd."
3	"I'm into hypnosis and meditation and have listened to several over the years. These sessions 1 and 2 are average . Many parts of the script make me feel like it's not for someone still working at exercising and losing weight but maintaining. The voice and the music is soothing and overall okay which is why I give it 3 stars."
5	"It feels funny finally buying songs like this one, but I needed music from my childhood to make me feel more secure in this " modern " day and age."

Study 2

Method

Data Collection In Study 2, we aimed to assess the relationship between evaluations and linguistic features in a related but distinct experiential category: restaurants. Assessing these relationships in a different domain provides opportunities to replicate and extend the effects we report in a different domain of experience.

We used a segment of restaurant review data from the Yelp Dataset Challenge (<https://www.yelp.com/dataset>) that includes restaurant ratings (on a scale of 1 to 5) and rater comments (Vinson et al., 2019). As in Study 1, 100,000 reviews were randomly sampled for use in this study.

Analysis and Results The median length of those reviews was 131 words with a maximum count of 998 words, and a minimum count of 1 word. Among all the selected reviews,

each restaurant review contained 4.24 adjectives on average, with a maximum count of 38 adjectives and a minimum count of at least one adjective used in one review. The linguistic desirability of each restaurant review ranged from 1.36 to 8.66, with an average score of 6.70 and a standard deviation of 1.54 on a scale of 1 to 9. The linguistic breadth of each restaurant review ranged from 2.65 to 7.33, with an average score of 6.02 and a standard deviation of 0.64. Table 4 provides examples of reviews varied by the ratings, which were all at the average level of breadth (Table 4). Again, many of the words were used as adjectives. Part-of-speech tagging suggested that approximately 87% of those adjectives were indeed used as adjectives.

Using the same analytic approach as in Study 1, we observed a similar linguistic pattern with restaurant reviews. Again, linguistic desirability ($\beta = 0.19, SE = 0.004, t = 53.15, p < .001$; Fig. 4) and restaurant ratings ($\beta = 0.46, SE = 0.004, t = 114.6, p < .001$; Fig. 5) both were significantly associated

with the linguistic breadth of review content. We also found a significant but negative interaction between linguistic desirability and restaurant ratings ($\beta = -0.007$, $SE = 0.001$, $t = -6.93$, $p < .001$), suggesting that the impact of linguistic desirability on linguistic breadth varied depending on the

restaurant rating. In particular, the negative interaction coefficient indicates that the positive effect of linguistic desirability on linguistic breadth decreased as the restaurant rating increased (Figs. 6a and 6b).

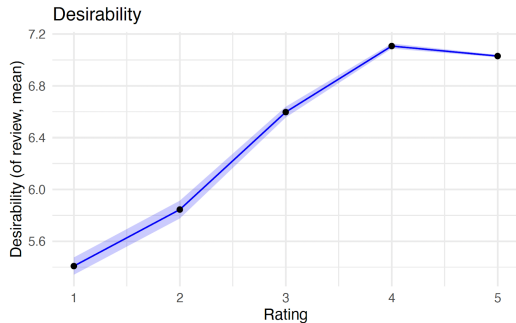


Figure 4: Restaurant rating by linguistic desirability.

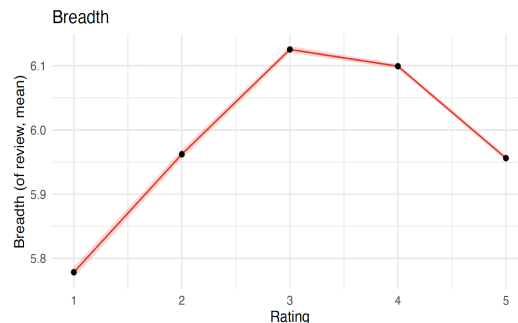


Figure 5: Restaurant rating by linguistic breadth.

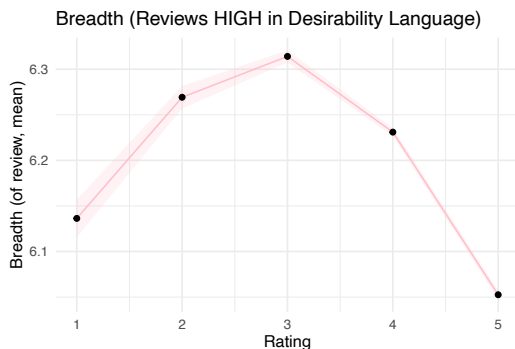


Figure 6a: Rating by breadth for high desirability reviews.

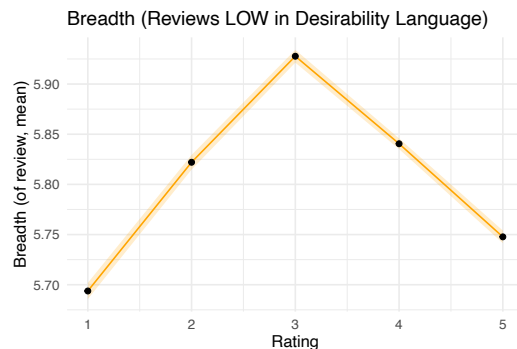


Figure 6b: Rating by breadth for low desirability reviews.

Table 4: Example of Yelp ratings at the average level of linguistic breadth (6.02).

RATING	REVIEW (ADJECTIVES IN BOLD)
1	"The sound is awful , the food is awful , and anything edible or potable is grossly overpriced even by venue standards. And the tickets... Dear lord, the tickets. If you've going to pack us like sardines into a 100+ degree venue, without even the courtesy of breathing room, lay off the ridiculous prices and fees..."
3	"Not 5 star legendary service. Not one 100% customer service award awarded here. Wasn't greeted at the door. There was one person in line. Two including me. Anywhere. The lady was nice . The barista was hot. But slow . Just my kind of man! The coffee was good . So I guess perfection takes time. I mean look at the 16th chapel. Rome wasn't built in a day. But a espresso machine doesn't take ten minutes to extract a shot."
5	"I have a truly deep love of this place. The service is great , the food is wonderful , and the beer...is amazing . If you need recommendations, I'd recommend trying the 8th st chicken strips as well as (for beers) the hefeweizen and sunbru."

Discussion

In this paper, we explored the linguistic context of online reviews, focusing on the breadth of adjectives. Specifically, we investigated the relationship between evaluative ratings and the semantic features of written reviews of Amazon digital music and restaurants on Yelp. Across the two studies, we found that (1) adjectives are prevalent in the context of online reviews, (2) ratings are significantly and positively

associated with the linguistic desirability of the review, (3) there is a meaningful relationship between the breadth and the desirability of the adjectives used in reviews. As linguistic desirability increases, so does linguistic breadth (except for five-star reviews), and finally (4) evaluative ratings interact with the positive relationship between linguistic desirability and breadth differently with music and restaurant reviews.

This work illustrated a combination of novel methodological approaches to test critical social-cognitive perspectives in linguistic cues in communication and

cognition. Previous studies looking at language abstraction using the LCM predominantly focused on the different levels of abstraction within verbs only, with little attention paid to adjectives (Dragojevic et al., 2017; Franco & Maass, 1996b; Geschke et al., 2010; Karpinski & Von Hippel, 1996; Von Hippel et al., 1997). With an explicit focus on adjectives, this work recognizes and establishes the prominent role of linguistic breadth in communication. Moreover, existing findings on language abstraction primarily involve interpersonal settings, with a heavy interest in stereotyping and prejudice. The current work shows that linguistic breadth is implicated in broader forms of communication.

This work also advances our understanding of the relationship between linguistic abstraction and ratings in online reviews, which is a limited area of investigation (Aerts et al., 2017a; Capocasa, 2014). It is encouraging that the results of Study 1 are consistent with Aerts et al. (2017), who found that linguistic desirability positively correlates with linguistic breadth. However, their approach only used 400 reviews, and we randomly sampled 100,000 reviews in each study. In addition, while our results are inconsistent with Capocasa (2014), in which the researcher found a null relationship, different categories of online reviews were analyzed. Capocasa (2014) analyzed appliances and books, a mix of product goods and experience goods. Here, we focused on two different experience goods: music and restaurants.

Intriguingly, high ratings enhance the positive relationship between linguistic desirability and breadth for music reviews yet attenuate the same relationship for restaurant reviews. In other words, when people write a review for a restaurant, on average, they use broader and more positive adjectives as they give higher ratings but use narrow and positive adjectives when they give 5-star ratings. Several potential factors could lead to this different pattern of results. First, although both are experience-good, music and restaurants are distinct categories, and dining out is often more costly in terms of finance, physical health (e.g., food poisoning), and mental health (e.g., encountering a rude server) than purchasing a piece of music. Therefore, the goal in communication might differ (i.e., describing vs. advising vs. persuading). According to Douglas and Sutton (2003), explicit communication goals have substantial effects on language abstraction independent of the impact of senders' beliefs.

Furthermore, Wakslak and Joshi (2020) demonstrated that increased persuasion is associated with lower linguistic abstraction. Thus, it is possible that due to the high cost of dining in restaurants (versus purchasing music), reviewers have different communication goals in writing restaurant reviews, which decreases the linguistic abstraction of those top-rated restaurant reviews (i.e., persuading others that this is a fantastic restaurant). Furthermore, evidence suggests that high abstractness correlates with deception (Gröndahl & Asokan, 2019), and reviews written in concrete language are considered more reliable and helpful than those written in abstract language (Capocasa, 2014). It could be argued that

due to the higher cost of dining out, reviewers with extremely unpleasant and pleasant experiences wanted their reviews to be perceived as trustworthy, thus writing with more concrete adjectives.

Regardless of the explanation for these effects, it is important to note that the coefficient associated with the interaction term is notably low, suggesting that interpretative caution is warranted.

Limitations and future directions

Our results build on the previous research in linguistic abstractness while extending the scope beyond human referents. However, several limitations are addressed here. First, the review lengths varied between Studies 1 and 2, possibly contributing to these findings. Second, the adjectives used in the review did not always target the referent. Third, future researchers could utilize an advanced computational approach beyond part-of-speech tagging to filter out the instances where those words were not used as adjectives (example of the word “kind” highlighted in red in Table 4).

Additionally, this work is limited to only two categories. Future studies should investigate other categories to provide a more comprehensive picture of the relationship between linguistic desirability and breadth. Another promising future direction is incorporating additional semantic features, such as concreteness (Brysbaert et al., 2014) and semantic contextual diversity (Hoffman et al., 2013). In addition, semantic divergence analyses could shed light on the (dis)similarity of the distribution of different adjectives under each rating category. Future investigation of the psychological status of breadth may help us understand its distinctive qualities, especially in the attributional power of adjectives. By linking breadth to other semantic measures, it offers a unique lens into how speakers (and listeners) are affected by these subtle aspects of language.

References

- Aerts, G., Smits, T., & Verlegh, P. W. (2017a). How online consumer reviews are influenced by the language and valence of prior reviews: A construal level perspective. *Computers in Human Behavior*, *75*, 855-864.
- Aerts, G., Smits, T., & Verlegh, P. W. (2017b). The platform shapes the message: How website design affects abstraction and valence of online consumer reviews. *Decision Support Systems*, *104*, 104-112.
- Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Social Psychology*, *9*, 272-279.
- Beukeboom, C. J. (2014). Mechanisms of linguistic bias: How words reflect and maintain stereotypic expectancies. In *Social cognition and communication* (pp. 313-330). Psychology Press.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, *46*, 904-911.
- Capocasa, L. (2014). Can you tell why some reviews result more helpful than others? An empirical analysis of utility votes in online consumers reviews adopting the abstractness-concreteness framework of the linguistic category model.
- Dixon, R. M. W., & Aikhenvald, A. Y. (eds) (2004). *Adjective classes: A cross-linguistic typology*. Oxford: Oxford University Press,
- Douglas, K. M., & Sutton, R. M. (2003). Effects of communication goals and expectancies on language abstraction. *Journal of Personality and Social Psychology*, *84*(4), 682.
- Dragojevic, M., Sink, A., & Mastro, D. (2017). Evidence of Linguistic Intergroup Bias in U.S. Print News Coverage of Immigration. *Journal of Language and Social Psychology*, *36*, 462-472. <https://doi.org/10.1177/0261927X16666884>
- Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and brain sciences*, *32*(5), 429-448.
- Franco, F. M., & Maass, A. (1996). Implicit versus explicit strategies of out-group discrimination: The role of intentional control in biased language use and reward allocation. *Journal of Language and Social Psychology*, *15*, 335-359.
- Geschke, D., Sassenberg, K., Ruhrmann, G., & Sommer, D. (2010). Effects of linguistic abstractness in the mass media: How newspaper articles shape readers' attitudes toward migrants. *Journal of Media Psychology: Theories, Methods, and Applications*, *22*, 99.
- Gröndahl, T., & Asokan, N. (2019). Text analysis in adversarial settings: Does deception leave a stylistic trace?. *ACM Computing Surveys (CSUR)*, *52*(3), 1-36.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior research methods, instruments, & computers*, *36*(2), 193-202.
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational researcher*, *40*(5), 223-234.
- Hampson, S. E., Goldberg, L. R., & John, O. P. (1987). Category-breadth and social-desirability values for 573 personality terms. *European Journal of Personality*, *1*, 241-258.
- Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior research methods*, *45*, 718-730.
- Karpinski, A., & Von Hippel, W. (1996). The role of the linguistic intergroup bias in expectancy maintenance. *Social Cognition*, *14*, 141.
- Krishnamoorthy, S. (2015). Linguistic features for review helpfulness prediction. *Expert Systems with Applications*, *42*(7), 3751-3759.
- Lin, L., Dale, R., & Stroessner, S. (In preparation). An extensive dataset evaluating breadth and desirability for 1,214 adjectives.
- Maass, A. (1999). Linguistic intergroup bias: Stereotype perpetuation through language. In Zanna, M. P. (Ed.), *Advances in Experimental Social Psychology* (Vol. 31, pp. 79-121). San Diego, CA: Academic Press.
- Maass, A., & Arcuri, L. (1992). The role of language in the persistence of stereotypes. In G. R. Semin & K. Fiedler (Eds.), *Language, Interaction and Social Cognition* (pp. 129-143). Newbury Park, CA: Sage.
- Maass, A., & Arcuri, L. (1996). Language and stereotyping. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and Stereotyping* (pp. 193-226). New York: Guilford Press.
- Maass, A., Salvi, D., Arcuri, L., Semin, G. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*, 981-993.
- Ni, J., Li, J., & McAuley, J. (2019, November). Justifying recommendations using distantly-labeled reviews and fine-grained aspects. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 188-197).
- Sekaquaptewa, D., Espinoza, P., Thompson, M. Vargas, P., & Von Hippel, W. (2003). Stereotypic explanatory bias: implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology*, *39*, 75-82.
- Semin, G. R., & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. *European review of social psychology*, *2*, 1-30.
- Shin, S., Chung, N., Xiang, Z., & Koo, C. (2019). Assessing the impact of textual content concreteness on helpfulness in online travel reviews. *Journal of Travel Research*, *58*(4), 579-593.
- Vinson, D. W., Dale, R., & Jones, M. N. (2019). Decision contamination in the wild: Sequential dependencies in online review ratings. *Behavior research methods*, *51*, 1477-1484.

- Von Hippel, C., Mangum, S. L., Greenberger, D. B., Heneman, R. L., & Skoglund, J. D. (1997). Temporary employment: Can organizations and employees both win? *Academy of Management Perspectives, 11*, 93-104.
- Wakslak, C., & Joshi, P. (2020). Expansive and contractive communication scope: A construal level perspective on the relationship between interpersonal distance and communicative abstraction. *Social and Personality Psychology Compass, 14*(5), 271-284.
- Wigboldus, D. H. J., Semin, G. R., & Spears, R. (2000). How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology, 78*, 5-18.