OFF TO A GOOD START? GRAMMAR AND SYNTAX IN THE OPENING PREDICT REVIEW HELPFULNESS

José-Domingo Mora[1]
Charlton College of Business
University of Massachusetts Dartmouth
280 Old Westport Rd., Dartmouth, MA 02747, United States
jmora@umassd.edu

Anoosha Izadi
Charlton College of Business
University of Massachusetts Dartmouth
280 Old Westport Rd., Dartmouth, MA 02747, United States
aizadi@umassd.edu

ABSTRACT

We investigate two novel predictors of review helpfulness: grammatical classes of words, such as verbs and articles, and the arrays they form, or syntax. Patterns of grammar and syntax are known to co-occur with speech varieties that affect communication outcomes. However, measuring text syntax can be methodologically challenging. To address this, we analyze ordered word positions within reviews rather than full texts and look for text segments where the diversity of unique words and classes of words peaks. We find that the five-word segment at the beginning of a review, which we call the opening, exhibits these characteristics. Through statistical modeling and content analyses, we show that grammar and syntax classes in the opening predict review helpfulness and co-occur with specific clusters of words in the full text. Additionally, experimental studies provide evidence that the opening does not work in isolation, supporting the assumption that consumers read reviews in full. These findings may help to simplify online review analyses and inform future research agendas on consumer reviews.

Keywords: Online reviews; Helpfulness; Linguistics; Consumer-generated content; Marketing

1. Introduction

Remarkably, 94% of online consumers are reluctant to purchase from companies with negative reviews, and 53% expect businesses to respond to such reviews within a week (ReviewTrackers, 2022). These attitudes and expectations raise the critical question of what review characteristics businesses should assess to prioritize responses. Scholars have addressed this issue by examining the antecedents of proxy measures of review persuasiveness, most notably review helpfulness or perceived helpfulness of reviews (Biswas et al., 2022) which we call PHR. Beyond its connection to persuasiveness, PHR is an interesting construct because measures of PHR pervade both virtual retail and high-traffic websites dedicated to product reviews. Over the years, several factors have been studied to predict PHR. These include consumer factors like involvement and susceptibility to social influence (De Pelsmacker et al., 2018), reviewer factors like experience and gender (Ravula et al., 2023), product factors like intangibility and variety (Choi & Leon, 2020), and context factors such as consistency with the valence and lexicon of other reviews (Namvar & Chua, 2023; Purnawirawan et al., 2015). Review factors, such as the number of words, customer ratings of the review (Hong et al., 2017), numerical clues in the text (Li et al., 2023), and content-analytic variables that allow the evaluation of constructs like emotions (Xu et al., 2023) and psychological distance of the reviewer from the reviewed product (Chaterjee, 2023) have also been extensively examined. These studies have contributed significantly to the understanding of PHR and have been useful in predicting its success. In addition to text analysis, researchers have begun to examine multimodal reviews and their influencing factors (Park et al., 2023; Ceylan et al., 2023; Jeong & Yeu, 2023; Li et al., 2023). However, none of these interesting contributions examine the simpler structural

[1] Corresponding author.
Cite: Mora, J. D., & Izadi, A., (2024, Aug) Off to a Good Start? Grammar and Syntax in the Opening Predict Review Helpfulness, *Journal of Electronic Commerce Research, 25(3).*

characteristics of texts, specifically grammar and syntax, which linguistics scholars have shown to reflect semantic and rhetorical text characteristics. Thus, the objective of the present study is to investigate grammar and syntax patterns in online reviews as potential predictors of PHR. Since grammar and syntax are straightforward observables known to most speakers of any language, the present investigation offers the prospect of simplifying and expediting analyses in business and academia.

Linguists employ grammatical classes of words (GCW), such as verbs, prepositions, adjectives, and articles, along with the patterns of connection between them, to define a level of analysis known as register. Registers are "'language patterns' associated with the goal and context of communication" (Biber & Conrad, 2019) and co-occur with rhetorical strategies of varying effectiveness. Crucially, in our study, registers can be detected through the analysis of shorter text segments within full and longer texts. For instance, scholarly articles and conversational English have been shown to differ in terms of patterns of three to four unique words (Biber, 2009). In scholarly articles, the pattern "the ____ of the" is very common, with the blank slot filled by nouns such as "subject" or "reality." In spoken language, this pattern is uncommon, whereas the fixed three-word patterns "I don't know" and "A lot of" are abundant. These simple patterns differentiate journal articles from colloquial conversations, regardless of the fundamental differences between the classes of speech. Readers may weigh the relevance and usefulness of a text belonging to one of these classes based either on simpler three- or four-word structures or the more fundamental differences associated with them, like vocabulary, sentence length, or perhaps both. Nevertheless, the analysis of shorter text sequences should help predict the relevance and usefulness of the text.

This study uses a mixed-methods approach, from computational to correlational to experimental, to answer two fundamental questions: Are there specific language registers within online reviews? If so, can these language registers be used to predict review helpfulness? We begin with insights from other fields concerned with complex codes: molecular biology and genetics. In the late 1980s, geneticists discovered a relatively short but highly variable segment of the human genome. Individual differences in this short segment have led to routine analyses to determine identity (Saad, 2005). Our intuition is that there may be a similar segment or segments in online reviews, which are texts (codes) written for a specific type of reader with similar goals and a likely narrow set of text structures. Our investigation reveals that the five-word segment at the start of the review, which we call opening, shows the highest variability in terms of unique words and grammatical classes of words compared with all other segments. Furthermore, we show that the GCW in the opening and the syntax patterns they form are important predictors of PHR.

2. Background

2.1. Importance of PHR

A large proportion of the literature on online reviews reports counts of "useful" or "helpful" ratings, which users assign to reviews and are subsequently reported by retail and review websites. Given the considerable overlap between the two terms and the subjective nature of helpfulness ratings, we treat them as a single measure earlier defined as PHR. This section highlights the importance and predictors of PHR and identifies the knowledge gap addressed in this study.

The capacity of various forms of electronic word-of-mouth to predict choice and consumption has been well-established in the literature (Godes & Mayzlin, 2004; Chen & Xie, 2008; Zhu et al., 2010). Anonymous online reviews influence purchase intentions more than the recommendations of friends and acquaintances (Erkan & Evans, 2018). Some product review metrics, such as the number of ratings per product and average user ratings, significantly improve the accuracy of demand models (Dellarocas, et al., 2007). Text-analysis constructs also predict marketing outcomes. Timeliness and accuracy of information predict the choice of travel destination as well as the use of review information in travel plans (Chong et al., 2018). The sales rank of products sold on Amazon.com correlates with the subjectivity and informativeness of reviews (Ghose & Ipeirotis, 2011). Importantly for our research, the influence of these metrics and constructs on consumption is modified by PHR. For instance, attitudes and purchase intentions are influenced by information in online reviews only when those reviews are perceived as useful (De Pelsmacker et al., 2018; Walther et al., 2012).

In terms of its direct effects on marketing outcomes, PHR tends to negatively affect the probability of purchasing from online retailers (Kim et al., 2018), but context factors influence this effect. For instance, PHR negatively affects purchase intention only when the queue of reviews on the web page opens with a negative entry (Kolomiiets et al., 2016). Price also attenuates the negative effects of PHR on sales (Duan & Mao, 2022). Finally, individual characteristics are important. Useful negative reviews of environmentally unfriendly products are more effective in reducing the intention to buy among individuals with higher, compared to lower, environmental moral norms (Filieri et al., 2021).

In summary, the literature provides substantial evidence that PHR plays an important role in forming attitudes and intentions and in predicting consumption, even as several other factors modify their effects. This justifies the interest of researchers in identifying the antecedents of PHR and continuing the search for newer, actionable ones. The following section identifies the known factors and points to neglected ones.

2.2. Predicting PHR

Earlier contributions on predictors of PHR focused on review extremity (as too high or low star ratings), review depth (as word count), and product type, which scholars have shown to moderate the influence of both extremity and depth on PHR (Mudambi & Schuff, 2010). A comprehensive meta-analysis (Hong et al., 2017) found that depth and age (time since posting) positively influenced PHR, whereas readability (as counts of words and characters) and customer ratings (1 to 5 stars, linear and quadratic) did not significantly affect PHR. Other formal review metrics affecting PHR include volume (as number of posted reviews; Hong & Pittman, 2020). Beyond formal metrics, researchers have explored content-analytical predictors of PHR, such as the quality of arguments (Filieri, 2015), subjectivity (Ghose & Ipeirotis, 2011) and the use of profanity (Hair & Ozcan, 2018). More recently, emotions expressed in consumer reviews have been shown to predict PHR. Specifically, negative emotions such as anger and anxiety (Xu et al., 2023), as well as various other discrete emotions (Ravula et al., 2023; Xu et al., 2022; Liao et al., 2022) have been shown to play a role. Additionally, the novel dimension of emotional arousal (Chou, 2023) and a related construct, sentiment have been implicated (Shah et al., 2023; Li et al., 2023; Biswas et al., 2022). These findings are consistent with earlier contributions on valenced emotions (Ahmad & Laroche, 2015; Felbermayr & Nanopoulos, 2016) and the effects of review slants (Maslowska et al., 2017). The predominant trend in the literature is that positive emotions decrease while negative emotions increase PHR, a pattern seemingly unaffected by cultural differences (Biswas et al., 2022). This consistency could be explained by the association of positive and negative information with two deeply ingrained heuristics in prospect choices: positive reviews might be interpreted as dishonest (e.g., as paid postings) leading to avoidance, whereas negative reviews capitalize on loss aversion (Casaló et al., 2015).

Linguistic predictors of PHR have only been identified in a limited number of studies. PHR has been shown to be affected by the match between the language styles of reviewers and readers (Liu et al., 2018) as well as by rhetorical and argumentative factors (Moradi et al., 2023; Amos et al., 2022; Ahmad, 2017; Srivastava & Kalro, 2019). These higher-level constructs may not be well-suited for everyday business decision-making, as they require interpretation and analysis. In contrast, several studies have used deep learning to train algorithms to measure ad hoc constructs (Singh et al., 2017) or find ad hoc word patterns at different levels of aggregation to predict PHR (Mitra & Jenamani, 2021). Unfortunately, this text-mining approach is not informed by the theory of human language and requires advanced computational methods; thus, its findings may be difficult to generalize.

To address these shortcomings—the lack of simple but still generalizable linguistic categories in online reviews—we discuss the empirically-informed framework known as varieties of text and derive research questions to guide our inquiry.

3. Theoretical Framework

Linguists recognize three varieties of text: genre, register, and style (Biber & Conrad, 2019). The broadest level is genre, which represents sets of texts sharing major structural commonalities, such as TV newscasts, advertisements, poetry, academic papers, and meeting minutes. The genre we are interested in is online reviews. The most specific text variety is style and it is highly influenced by the personal inclinations of individual speakers. The style reflects how an individual user of language creates text, including their personally preferred syntactical forms and vocabulary. For instance, in literature, the styles of individual writers within the genre of short fiction tend to be very different: readers can distinguish one author from another. Thus, style has limited generalizability and is not the focus of our research. Instead, we focus on register, the middle-range variety between genre and style. A register is characterized by the linguistic commonalities in texts written by language users within in a shared context, be it social or otherwise, while ignoring individual preferences. To illustrate how genre, register and style are integrated within the varieties framework, we now examine the genre football broadcasting. This linguistic genre is predominantly narrative; it is anchored in specific jargon that describes rules, roles, and actions, and reflects the customs and values of the culture in which it is embedded. The use of interjections (for example, Gosh! Incredible!), adjectives (e.g., great, impressive), and action verbs (e.g., runs, tackles) are very frequent. However, while one group of broadcasters may want to add excitement by emphasizing interjections or even introducing their own neologisms as rhetorical devices, another group may prefer a more sober, objective narrative, thus sparing interjections and grandiose adjectives. These two groups share two distinct registers within the genre, which we could label as “exciting” and “objective.” Finally, the style of each individual broadcaster is likely to differ in terms of unique words (for example, Gosh! vs. Incredible!), personal phrases, and linguistic structures (e.g., shorter vs. longer sentences).

Crucially for the present research, and as stated earlier, a text’s linguistic registers can be identified through shorter text segments (Biber, 2009; Biber & Conrad, 2019). However, it is safe to assume that not all text segments carry the same predictive power and those with greater variability across reviews are more likely to hold predictive power. We refer to these highly variable and potentially predictive sequences as diagnostic segments. We draw an analogy from genetics and molecular biology, where researchers identified relatively short DNA segments showing significant variation among human individuals. Although not all DNA can vary substantially without affecting bodily functions and survival, variations in this small segment contribute to individuality (Saad, 2005). Language, as a code, strikes a similar balance; adherence to basic norms is necessary for effective communication, but variability is also allowed. We empirically investigate whether online reviews exhibit greater variability in the beginning, middle, or possibly other sections. In contrast to DNA matching, we are not interested in individual differences but in identifying commonalities within subsets of individual reviews. For this reason, we will look at a dozen GCWs and strive to consolidate their various syntactic patterns. We do not look for unique words, which are numerous, or the larger number of arrays they form, thus departing from linguistic approaches that rely on unique words and distinctive sequences to characterize text genres or registers (Biber, 2009). The GCW and syntax patterns in diagnostic segments, as defined earlier, are expected to predict PHR as they co-vary with similar language patterns in the full text. Thus, the first question is:

RQ1: Are there diagnostic text segments in online reviews?

Linguistics scholarship typically involves the description and comparison of large corpora (bodies) of language, such as academic versus conversational, as demonstrated in the earlier example. To the best of our knowledge, there has been little research connecting lexical or syntactic patterns to any measurable outcome, such as persuasion or changes in mood. The present research aims to bridge the gap between marketing and linguistics scholarship by investigating the relationship between syntax patterns and a specific outcome, namely, PHR. Therefore, this study addresses the following questions:

RQ2: Do the proportions of words by grammatical class in the diagnostic text segments predict the perceived helpfulness of online reviews?

RQ3: Do syntax patterns (sequences of GCW) in diagnostic text segments predict the perceived helpfulness of online reviews?

4. Overview of Empirical Studies

We conducted four concatenated studies, as summarized in Table 1. In Study 1, we addressed RQ1 by measuring the linguistic variability of long text segments across reviews. We used computational methods to perform a large set of simple operations. First, we took position number one (the first word in all reviews) and found both the sum total of unique words and the sum total of grammatical classes in that position overall reviews in the sample. Each sum is an indicator of variability. This provided the first pair of data points. Next, we repeated the same process with word positions ranging from two to 100. Finally, we created profiles of linguistic variability using two alternative measures. This exploratory approach allows for the detection of text positions with more variability and, hence, better potential as diagnostic segments.

Study 2 has two parts. The first part was also exploratory and focused on calculating the proportions of GCW (RQ2) and identifying the syntax patterns of GCW (RQ3) in the diagnostic segment detected in Study 1. Syntax patterns were identified using multivariate techniques. In the second part of the study, the proportions and patterns of GCW were tested as predictors of PHR. The estimation of the statistical model in Part Two of Study 2 enables us to answer RQ2 and RQ3 and, hence: a) establish if syntax, word composition, or both predict PHR and, as a result, b) decide what type of stimuli to use in Experimental Study 3.

Study 3 corroborates the findings of Study 2 using experimental methods with an online sample. Study 4 also uses an experimental approach to test a hypothesis emerging from Study 3. Specifically, as the diagnostic segment detected in the present research is formed by the first five words of the reviews, we elucidate whether readers use this short segment as a heuristic or if they read the full text or a substantial part of it. We present each of these studies in detail below.

5. Study 1: Finding Diagnostic Segments

5.1. Sample and Data

For Studies 1 and 2, we relied on a convenience sample of customer reviews of restaurants in the metropolitan area of Phoenix-Scottsdale, Arizona, collected through *Yelp!* between March 2005 and January 2013 and made publicly available online (https://data.world/datasets/consumer). The database contained 229,907

Table 1: Research Design

	Study 1	Study 2	Study 3	Study 4
Input	Consumer reviews	Diagnostic segment/s by review	Syntax patterns predicting PHR	Syntax patterns predicting PHR
Process	• Text-analytical software LIWC 22 and SAS • Chosen measures of variability[a]: ○ Shannon index for unique words ○ Simple proportions for GCW • Both measures are computed for word positions 1 to 100	• Part 1: Theoretically guided clustering and classification trees to find syntax patterns[b]. • Part 2: GLMM regression. Proportions of GCW and syntax patterns enter the model as predictors of PHR	• Experimental study on an online panel. • Participants are given a series of short reviews with syntax patterns known to the experimenter (personal vs. impersonal) and asked to holistically rate them in terms of PHR	• Experimental study on an online panel. • Participants are given several short reviews, either with the original text intact or with an altered diagnostic segment but the same remainder of the text, and asked to holistically rate them in terms of PHR
Output	Diagnostic segment/s with higher word variability → RQ1 answered	Syntax patterns predicting PHR → RQ2 and RQ3 answered	Confirmation of syntax pattern's direction of effects in Study 2 → Emerging H1: Opening reflects full text	Altering the diagnostic segment does not change perceived PHR. → H1 is supported

[a] A variance was also considered but not used for reasons explained in 5.2.

[b] The number of possible syntax patterns is severely constrained by the fact that most of these patterns make no linguistic sense, e.g., when the sentence "It is a beautiful day" is re-arranged as "Beautiful it day is." This issue is advantageous for taxonomical purposes and allows the present study to focus on a two-layer structure (see Section 6.1).

consumer review entries for several product categories. We focused on restaurants as they predominate in the dataset (61.2%), and *Yelp!* reports a larger set of product characteristics for restaurants than for any other product, which are unique; for example, the type of food served and the presence of a bar on the premises. Of the 140,683 restaurant reviews in the restaurant dataset, we retained those with 20 or more words, for a total of 130,848 words. This judgmental word limit seeks to exclude reviews that are too short to have a well-defined narrative or analysis while preserving the sample size. Furthermore, we needed to retain only the reviews for which the chosen text-analytic software (LIWC 22, see Methods) was able to classify at least 90% of the words into one of the GCW, which is a methodological necessity. The reviews discarded in this last round contained extremely rare words, typos, neologisms, and custom words (e.g., Yesssss). The process yielded 95,917 reviews with a median of 122 words and 653 characters. Table 2 presents the statistics for this sample.

5.2. Methods

Text is operationalized as a sequence of words in a unique order, with the first word occupying the first text position and the last word occupying the last text position. Text segments are operationalized as shorter-ordered subsets within a text, starting at any position and with a small number of elements. Reviews vary greatly in length; therefore, we focus our analyses on positions 1 to 100 (41.2% of the sample). Reviews with 20-100 words comprised 40.2%, and the remaining 59.8% had at least 101 words. Only 5.9% of reviews have 250 words or more.

Positions 1 to 100, across 95,917 reviews, were characterized by the variability of entries in terms of two classification frameworks: a) unique words, which encompass over a million categories in the English language, or b) GCW, with a number of categories orders of magnitude smaller, namely nouns, adjectives, verbs, adverbs, prepositions, conjunctions, articles, auxiliary verbs, or one of the five types of pronouns. To detect unique words, we extracted individual words from each review and compiled them to calculate word frequency. To detect the GCW, we used a more elaborate two-pronged approach.

We first analyzed the review segments from word positions 1 to 100, one position at a time, for a total of 100 analyses, using the psycholinguistic software LIWC-22 (Boyd et al., 2022). LIWC compares words in a text sample

Table 2: Descriptive Statistics of the Estimation Dataset

		Mean	Std Dev	Minimum	Maximum
DV	Review is "useful" (counts)	1.45	2.19	0	82
Review factors	Length (words)	153.96	119.25	20	5,000
	Stars awarded	3.69	1.19	1	5
	"Cool" (votes)	0.88	1.85	0	77
	"Funny" (votes)	0.68	1.75	0	70
	Age (days)	733.80	507.71	0	2,796
Reviewer factors	Stars awarded	3.73	0.56	1	5
	"Useful" (votes)	344.23	1036.20	0	24,293
	"Cool" (votes)	242.02	862.35	0	22,410
	"Funny" (votes)	194.86	689.01	0	24,519
	Number of reviews	121.44	194.94	1	2,810
Product factors	Number of reviews	134.34	137.63	3	803

Product factors	Type of food[a]	American new	15.1%	Steaks/BBQ	4.2%
		Mexican	13.8%	Japanese/sushi	4.1%
		American traditional	11.2%	Asian	3.0%
		Pizza	9.3%	Thai	2.9%
		Breakfast	8.2%	Delicatessen	2.6%
		Sandwiches	7.6%	Seafood	2.6%
		Burgers	6.1%	Barbeque	2.6%
		Chinese	5.1%	Vegan/vegetarian	2.5%
		Italian	4.8%	Tex-Mex	1.6%
	Bar on premises [a]		9.2%		

[a] Proportion of the sample. The types of food and bars on the premises are non-mutually exclusive.

with dictionaries of words by linguistic category. Importantly, LIWC considers grammatical context, as words may perform more than one function, as in "I *like* it" (verb) and "It's *like* new" (preposition). LIWC-22 identifies 13 GCWs: adjectives, adverbs, verbs, conjunctions, prepositions, articles, and five types of pronouns (1st person singular and plural, 2nd person, 3rd person singular and plural, and impersonal). LIWC-22 does not identify nouns or interjections. To this end, we used text analytical functions in the SAS package to develop codes to extract, clean, and detect individual words, which were compared with a custom dictionary of 4,008 nouns and 477 interjections collected from online repositories. Each word was identified as a noun or no-noun and as an interjection or no interjection with SAS, but these labels are retained only if LIWC does not place the word in another GCW. Hence, we indirectly used LIWC's contextualization capabilities to detect these two classes through SAS.

We are not interested in unique words or GCW frequencies *per se* but in their variability over positions 1–100. Variance is an obvious candidate for a measure of variability; however, for discrete variables such as unique words and GCW, it would be necessary to know the shape of their distribution to choose proper mathematical expressions. A better choice for this type of data is the Shannon diversity index (Shannon, 1948), which captures two components of diversity: the number of classes in the set and how uniformly distributed individual observations are over those classes. For unique words in a 100-position online review segment, the Shannon index is specified as $H_i = -\sum_{w=1}^{W} p_{iw} \log(p_{iw})$, where i denotes position in text, from i=1 to I=100, w denotes unique word (W≈1,000,000) and p_{iw} is the ratio of the frequency of any given w in position i to the total count of reviews having position i. For instance, if the unique word "The" appears in the first position in 10,000 out of the 95,917 reviews in the dataset, then the proportion of w="The" in the set of reviews that have position i=1 (at least one word, all of them) would be $p_{1,\text{"The"}}$=10,000/95,917 =0.104. This ratio is calculated as many times for position i=1 as unique words are found in position 1. The resulting set of p_{iw} was inserted into the formula to obtain H_1. We repeated the entire process 100 times for as many positions as possible.

Shannon's measure loses power as the number of categories decreases, which poses a problem for the GCW, which had fewer than 10 categories in Study 1. For instance, consider that at position 3, there are five different GCW, including nouns. If all nouns were suddenly substituted for pronouns in all these reviews, the Shannon index would remain unchanged at the GCW level, even as considerable changes would take place at the unique word level. Thus, for GCW, we calculate the percentage of each GCW, subscripted as g in each review text position i, which we call q_{ig}. For instance, if articles were present in 19,183 of the 95,917 online reviews in position 3, then $q_{3Articles}$=19,183/95,917=20%. We calculated q_{ig} for positions 1 to 100 for verbs and nouns (consolidated as their profiles mirror each other), prepositions and conjunctions (consolidated for the same reasons), adjectives, adverbs, pronouns 1st person, pronouns 2nd person, and interjections (dropped from Study 2 in favor of content analysis variables). Pronouns in the 3rd person were rare and hence excluded.

5.3. Results

Figure 1 presents the results of the analyses of H_i and q_{ig}. Panel a) shows that H_i grows rapidly from positions 1 to 5 and flattens out around position 11, and steadily declines afterwards. This tendency is not explained by the loss of short reviews as review size increases, as indicated by the second curve in 1a), the ratio of diversity of word position to the number of unique words per position (H_i/N_{iw}), which follows the same pattern. The decay in H_i after position 11 may indicate that cumulative cognitive effort causes reviewers to settle on widely used language patterns, which hurts unique word diversity. An alternative, non-exclusive explanation is that reviewers put extra effort into making the opening attractive, and that effort wanes with text length. Figure 1, panel b) reveals that changes in q_{ig} are important from positions 1 to 5, stabilizing, coincidentally, at position 6. Thus, syntax patterns are less predictable in the first five positions compared to the remainder of the text, which aligns with the explanations provided for changes in unique word diversity. We conclude that positions 1 to 5 form the only well-differentiated area across texts and, hence, the sole possible diagnostic segment, which provides an answer to RQ1. We call this segment *review opening*.

6. Study 2: Predicting PHR From GCW and Syntax

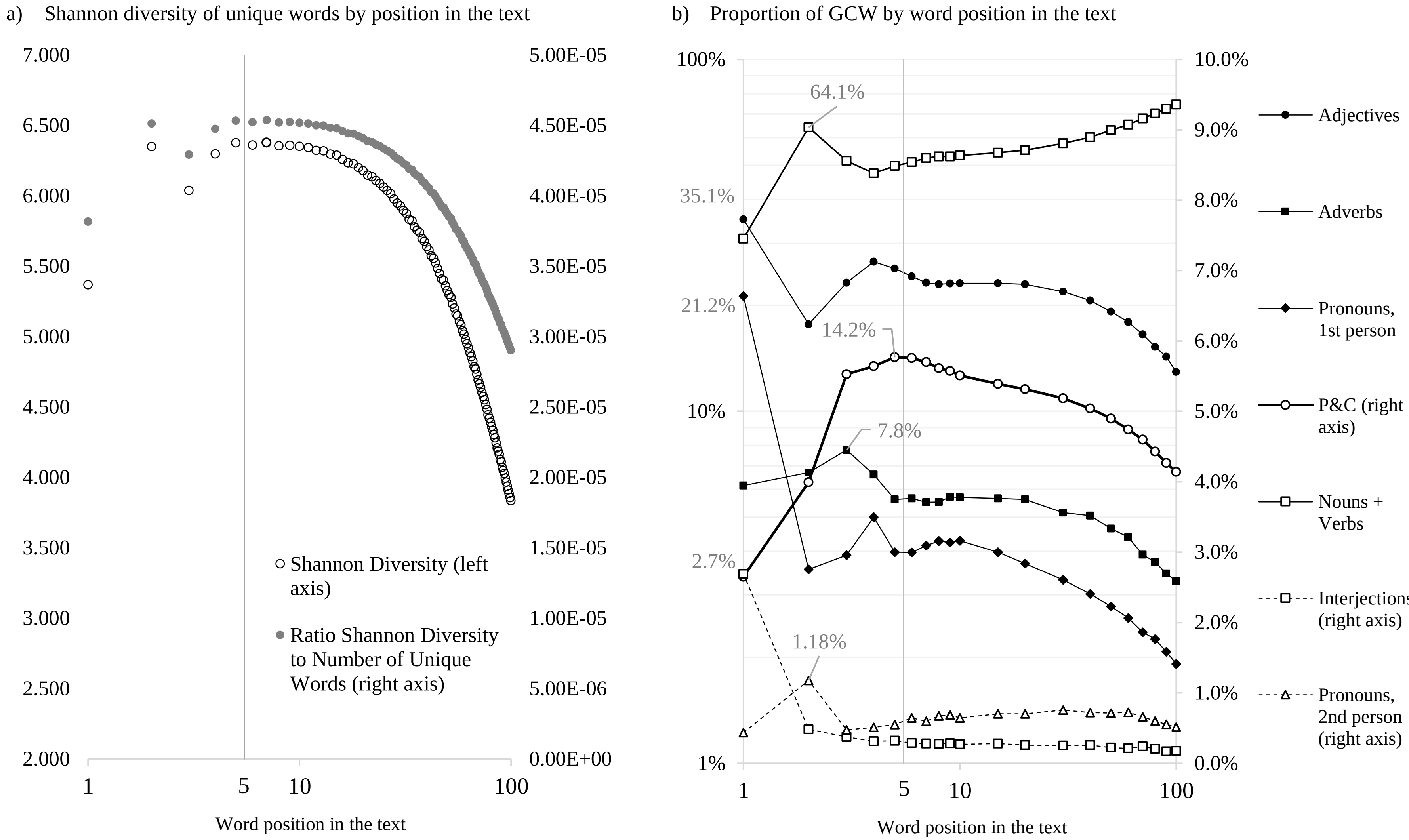
This section is divided into two parts: First, we use the GCW detected in Study 1 to empirically determine syntax patterns, that is, sequences of GCW; second, we estimate a statistical model of PHR that incorporates GCW and syntax patterns to answer RQ2 and RQ3.

6.1. Syntax Patterns

Linguists have shown that specific short-word sequences characterize genres and registers. For instance, the most frequent short-word sequences in conversational English and academic papers are very different (Biber, 2009). Registers within online genre reviews may be challenging to detect; however, as Study 1 shows, there is a potential diagnostic segment, the opening, which could provide a novel and interesting way for classifying them into meaningful categories. The potential number of syntax patterns in a set of five GCW, like the opening, chosen from a set of 13 GCW when order matters, equals the permutations of 13 objects taken as sets of five, ${}_{13}P_5$= 154,440. This total includes syntax patterns that are not semantically viable (e.g., “This place is good,” viable, vs. “Place is good this,” non-viable) or are rare (“Good is this place”). By assuming that reviewers ensure that they do not use non-sensical or rare syntactic forms, we can disregard order words. This reduces the potential number of syntax patterns to combinations of 13 objects taken in sets of five, ${}_{13}C_5$= 1,287. These 1,287 patterns need to be further reduced into a number of syntax classes small enough to be accommodated in a regression equation, by grouping similar patterns into broader groups of patterns. Such grouping must also make theoretical sense. Thus, we grouped the GCW into two widely used categories of words in linguistics: *content* and *function* (Pennebaker, 2013). Content GCW name and describe objects and people, and the actions they perform or are performed upon them. They are nouns, verbs, adjectives, and adverbs. Function GCWs connect content GCWs in a manner consistent with their tense, number, and gender; that is, pronouns, articles, conjunctions, prepositions, and auxiliary verbs. Content GCW conveys facts and assessments, an informational component different from the “meta-data” conveyed by functional words, i.e., Who’s talking? How many are there?

The potential number of variations in a set of five positions containing either content or function words is ten, some of which may not make sense, for instance, five consecutive function words (five pronouns, prepositions, or conjunctions) or five consecutive content words (nouns, verbs, adjectives, or adverbs). Thus, our starting point was reduced to a maximum of eight combinations of content and function words. Considering all combinations of the four content and seven function words, we arrived at a total of 560, many of which would be of limited or no use, while others would have disproportionate representation. Some may even be very similar to others. This makes it necessary to empirically assess syntax patterns, but at the same time, the order of magnitude of the number of syntax classes indicates that this theoretically guided classification task does not require major computational effort.
Therefore, we use classic multivariate methods.

We proceeded in three stages. First, a round of clustering analyses where content GCW and function GCW are entered separately as candidate classification criteria. This yielded the first layer of taxonomic units, which we call first-layer clusters. Second, several rounds of within-cluster clustering yield a second layer of clusters nested within the first layer. Third, consolidation of second-level clusters using classification trees substantially reduces the



Note: Vertical lines between 1 and 10 indicate word position 5. Data point labels in 1b) indicate maxima or local maxima.

Figure 1. The First 100 Words

number of syntax covariates. Trees allow cross-checking that the DV varies over consolidated clusters (Appendix A provides a detailed explanation of the entire process). Thus, syntax patterns are captured as two-layer structures in which each layer is a distinct set of GCW.

In the first round of clustering, we found that the two most abundant pronouns, 1st person singular (present in 42.9% of openings) and impersonal (25.6%), or their absence (31.5%), defined the highest-quality clusters, which we call, respectively, C_A, C_B, and C_C. This is both a welcome simplification emerging from the data and a confirmation of the conjecture formulated earlier in this section about the importance of pronouns as conveyors of critical "meta-data." This validates the expectations set in the previous paragraph.

In the second round of clustering, we find that the sets of GCW defining the new second-level clusters differ across first-level clusters. Cluster C_A of personal pronouns yielded second-tier quality clusters only with content GCW as the classification criterion for a total of 11 clusters. Conversely, in the first-level cluster, C_B is a function word, specifically an impersonal pronoun, that yields quality clusters totaling three. Cluster C_C does not yield quality clusters using content or function CGW; hence, it will remain the base case for the dummies capturing the second-level clusters in the statistical model. Finally, we used classification trees with PHR_r as the DV to aggregate these 14 second-level clusters into fewer nodes. We use the term second-level groups to designate these nodes. The number of second-level clusters within C_A is now three, and for C_B is two. In total, five empirical syntax classes.

Table 3 summarizes the characteristics of the empirical syntax classes. Each class is formed by GCW nodes that define the second-level groups and pronouns that define the first-level clusters. For instance, Verbs-C_A is the first group in the chart and represents a syntax class in which verbs and first-person singular pronouns are always present.

Table 3: Empirical Classification of Review Openings by Syntax Class

Syntax Class	1st Level: Clusters 2nd Level: Nodes	C_A - 1st Person Singular Pronouns 1PP-Verbs	 1PP-Nouns	 1PP-Mix
N		5,154	24,144	11,812
Mean PUR_r		1.53	1.28	1.39
Share of sample		5.4%	25.2%	12.3%
Content GCW forming nodes	Nouns	0%	100%	41%
	Adverbs	0%	29%	58%
	Verbs	100%	86%	55%
	Adjectives	7%	15%	51%

Syntax Class	1st Level: Clusters 2nd Level: Nodes	C_B - Impersonal Pronouns IMP-Aux	 IMP-Prep	C_C - Other N.A.	Total Sample (C_A+C_B+C_C)
N		16,302	8,278	30,227	95,917
Share of sample		17.0%	8.6%	31.5%	100.0%
Mean PUR_r		1.35	1.29	1.34	1.34
*Function** GCW forming nodes	Conjunctions	12%	11%		
	Prepositions	0%	100%		
	Articles	30%	18%		
	Auxiliary verbs	68%	49%		

* Excluding pronouns

6.2. A Model of PHR

We estimate regression models where dummies for GCW and syntax patterns in the opening enter as focal predictors of PHR of review *r* (PHR_r), operationalized as the number of "helpful" votes reported on *Yelp!* (Hong et al., 2017). Count data such as these tend to follow a Poisson or negative-binomial distribution, which can be empirically elucidated. Generalized linear mixed models (GLMM) (Lee & Nelder, 1996) simultaneously deal with count data and the hierarchical relationships embedded in the two-layered syntax classes emerging from Study 1. With GLMM, there is no need to transform the dependent variable (counts) or assume that they follow a symmetric distribution. As with generalized linear models (GLM), the DV is linked to a set of predictors through a link function, which helps capture nonlinear relationships.

We first present the model in its hierarchical form and then discuss a simpler, non-hierarchical specification that is necessary for benchmarking: models not accounting for grammar or syntax, grammar only, and grammar and syntax in a non-hierarchical form. The link function for both the Poisson and negative binomial distributions is the natural logarithm. Hence,

$$E[PHR_r] = \log(z_r) \qquad (1).$$

The linear predictor z_r in (1) is structured as

$$z_r = \beta_0 + \boldsymbol{\beta_1}'\boldsymbol{SYNTAX_{1r}} + \boldsymbol{\beta_2}'\boldsymbol{GRAMMAR_r} + \boldsymbol{\beta_3}'\boldsymbol{REVIEW_r} + \boldsymbol{\beta_4}'\boldsymbol{REVIEWER_r} + \boldsymbol{\beta_5}'\boldsymbol{PRODUCT_r} + \varepsilon_r \qquad (2).$$

Furthermore, (2) can be extended, as said earlier, by specifying:

$$\boldsymbol{\beta_1} = \boldsymbol{\beta_{21}}'\boldsymbol{SYNTAX_{2r}} + \upsilon_r \qquad (3),$$

where ε_r , $\upsilon_r \sim$ NB(r,p)[2]. The test vector $\boldsymbol{GRAMMAR_r}$ contains dummies for GCW in the opening. Test vectors $\boldsymbol{SYNTAX_{1r}}$ and $\boldsymbol{SYNTAX_{2r}}$ contain dummies for pronoun-defined clusters and second-level syntax patterns, respectively. The vectors $\boldsymbol{REVIEW_r}$, $\boldsymbol{REVIEWER_r}$ and $\boldsymbol{PRODUCT_r}$ contain other reviews, reviewers, and product characteristics, respectively. Table 4 summarizes the operationalization of the vector elements.

Table 4: Model Predictors by Model Specification

Vectors	Elements / Operationalization			Scale/Metrics	Source/ Software	Sample references	Model specifications
SYNTAX·r	Syntax patterns			Dummies	LIWC-22, SAS, SPSS	*N.A.*	3 & 4
GRAMMARr	Grammatical classes of words			Dummies			2, 3 & 4
REVIEWr	Psycho-metrics	Emotion	Positive Negative	% words in text capturing each construct	LIWC-22	A&L, X&A	1, 2, 3 & 4
		Cognitive processes				LU&A	
	Readability	Informal language		% informal words		G&I, P&N	
		Length in words		Ln(Counts+1)	As reported in the dataset	G&I, L&C	
	User evaluation	Stars (linear, squared)		Rating, 1 to 5		A&L, LA&A, M&A	
		Review is fun Review is cool		Counts			
	Time posted	Days posted		Ln(Counts+1)		G&I	
REVIEWERr	User evaluation	No. of stars		Rating, 1 to 5		LU&A	
		Reviewer cool Reviewer funny		Ln(Counts+1)		*N.A.*	
	Experience	No. of reviews posted		Ln(Counts+1)		S&K	
PRODUCTr	User ratings	No. of stars		Ordinal 1-5		H&al.	
	Volume	No. of reviews		Ln(Counts+1)			
	Features	Bar in the premises		2 categories		C&L, LU&A	
		Type of food		18 categories			

Note: The dependent variable, PHRr ≥0 is the number of "agree" answers with the sentence "This review is useful" reported on the *Yelp!* web site.

Legend: *N.A.* is "not available" to the best of our knowledge. A&L is Ahmad & Laroche 2017, C&L is Choi & Leon 2020, G&I is Ghose & Ipeirotis 2011, H&al. is Hong et a. 2018, L&C is Lee & Choeh 2016, LA&A is Liao et al. 2022, LU&A is Liu et al. 2018, M&A is Maslowska et al. 2017, P&N is Park, S., & Nicolau 2015, S&K is Srivastava & Kalro 2019, X&A Is Xu et al. 2022.

In simpler benchmark models, we omit equation (3), drop the vector $\boldsymbol{SYNTAX_{1r}}$ from equation (1), and consider a unique error term, ε_r, yielding non-hierarchical GLMs (Wedderburn, 1974). Both hierarchical and non-hierarchical forms are estimated via maximum likelihood with the Laplace approximation in SAS, which produces identical estimates as the general estimating equations. The proposed models deviated from those in the literature (Liu et al., 2018; Yi & Oh, 2022), who estimated zero-inflated models as an extension of the GLM, as they observed disproportionally large numbers of reviews with zero votes. In our data, the DV frequency peaked at zero but decreased smoothly.

6.3. Results and Discussion

Table 5 presents parameter estimates for the four nested specifications. Model 1 was the benchmark and included the

[2] In the dataset, the ratio DV variance to mean is large (3.45) which favors negative binomial over Poisson distribution.

most relevant predictors found in the literature. Model 2 introduces the vector $\boldsymbol{GRAMMAR_r}$ whose elements are GCW. Model 3 introduces $\boldsymbol{SYNTAX_{2r}}$, which contains only dummies for the second-level nodes, as shown in Table 4. As explained earlier, Models 1, 2, and 3 are nonhierarchical GLMs. Model 4 is hierarchical because it tests syntax as structures where content and function GCW influence how the two types of pronouns affect PHR. Model 4 introduces $\boldsymbol{SYNTAX_{1r}}$ containing dummies for the 1st person singular and impersonal pronouns hierarchically related to the function and content GCW in $\boldsymbol{SYNTAX_{2r}}$. As fit measures allow model comparisons, we report Akaike's Information Criterion (AIC), which gets smaller as fit improves, with differences of 2 or more considered as highly significant; McFadden's pseudo-r-squared (ρ^2)[3], with values larger than 0.2 considered "excellent" (McFadden, 1977).

In Table 5, the fit improves with the model complexity from left to right. The AIC values decreased significantly in that direction, with large differences between adjacent specifications. Furthermore, even when it changed modestly, ρ^2 decreased consistently. Model 2 shows that GCW in the opening has substantial effects on PHR, which remain unchanged as the syntax enters Models 3 and 4.

These results provide a positive response to RQ2. The significant GCW estimates fall at the lower end of all significant estimates but are still in the same order of text readability, cognitive processes, and product characteristics. The GCW estimates appear to be independent of the remaining estimates, which change little in Model 2 compared with Model 1. It is worth noting that, among the three groups of GCW estimates, pronouns showed the largest degree of variability compared to other functional and content GCW. This means that the effect of these functional GCWs is less certain than that of the other GCWs. The largest positive estimate is for 1st person singular (I, me, my, etc.), and the most negative one is for the second person (you, your, etc.). Impersonal pronouns show a large positive estimate, whereas estimates for 1st person plural (we, our, etc.) and 3rd person, regardless of number, are non-significant.
Model 3 introduces $\boldsymbol{SYNTAX_r}$, a vector of five dummies that capture a unique combination of GCW, either content (e.g., verbs and nouns) or function (e.g., article and preposition), but all have a pronoun as they emerge from the clustering process. Model 3 improved the fit over Model 2, thus providing a positive answer to RQ3. As explained earlier, the hierarchical model 4 splits $\boldsymbol{SYNTAX_r}$ into $\boldsymbol{SYNTAX_{1r}}$ and $\boldsymbol{SYNTAX_{2r}}$. The first-level vector $\boldsymbol{SYNTAX_{1r}}$ contains three dummies for pronouns: 1st person singular pronouns, impersonal pronouns, and other pronouns plus no pronouns. The second-level vector $\boldsymbol{SYNTAX_{2r}}$ contains syntax classes in $\boldsymbol{SYNTAX_r}$ minus pronouns, thus becoming strictly content or functional word classes.

Model 4 further improves the fit compared to Model 3, thus corroborating that the answer to RQ3 is positive. The syntax class estimates for Model 4 capture the cascade of effects: *GCW (level 1)* $\rightarrow$ *Pronouns (level 2)* $\rightarrow$ *PHR*, as opposed to *GCW* $\rightarrow$ *PHR* in Model 3. For instance, the estimate for the combined effect of the syntax class Content GCW (noun, verb, adverb, and adjective) and any 1st person singular pronoun was 0.045 ($p<.0001$). The estimate for Content GCW combined with any impersonal pronoun was 0.140 ($p<.0001$)[4]. Importantly, an alternative model specification reversing the causation sequence to pronouns $\rightarrow$ syntax class minus pronouns $\rightarrow$ PHR considerably reduced the fit (AIC=269,938; only better than Model 1).

These results support the view of text in general, and reviews in particular, as multilevel structures in which one pivotal GCW, pronouns in this case, anchors the narrative, while other grammatical classes further specify their effects.

To help match the opening and text characteristics, we now profile the full text of reviews belonging to each of the five syntax classes captured as dummies in Model 4 using the text-analytic software LIWC. For simplicity, we refer to the three syntax classes within the 1st person singular pronoun cluster as *personal* and the syntax classes in the impersonal pronoun cluster as *impersonal*. First, we determined their linguistic profiles as the proportion of total words belonging to each of the nine non-pronoun GCW, using paired t-tests for unequal samples and variances. We found no significant differences between the syntax classes. The coefficients of variation were equally modest, ranging from 0.80% to 4.01%.

For the second analysis of the full texts, we use 65 psycho-social variables in LIWC, each falling within two broad areas and nine groups within those two areas. The areas, which we refer to as psychological traits consist

---

[3] $\rho^2 = 1 - [\ln(L_M) / \ln(L_0)]$, where L_M and L_0 are the values of the likelihood functions for the test model and the intercept-only model (Dobson & Barnett 2008)

[4] Some hierarchical estimates in model 4 are non-significant and hence omitted for brevity, i.e. syntax patterns on other pronouns (3rd person and 1st person plural) or no-pronouns, and syntax patterns with content GCW on impersonal pronouns.

Table 5: Estimates by Model Specification

				Model 1	Model 2	Model 3	Model 4
N= 95,917			AIC	270,705	269,919	269,880	269,776
			McFadden's r^2	0.2544	0.2566	0.2568	0.2570
	Level	*2 - GCW*	*1 - Pronoun*				
Opening	Syntax patterns (hierarchical)	Content GCW	1st Person Singular				0.045***
		Nouns-Verbs	1st Person Singular				0.080***
		Verbs	1st Person Singular				0.114***
		Content GCW	Impersonal				0.140***
		Nouns-Verbs	Impersonal				0.103***
		Verbs	Impersonal				-0.183***
		No-Preposition	Impersonal				-0.095***
		Preposition	Impersonal				-0.006
	Syntax patterns (independent)	Content GCWs and 1stPS				0.066***	
		Nouns-Verbs and 1stPS				0.087***	
		Verbs and 1stPS				0.042**	
		No-Preposition & Impersonal				-0.096***	
		Preposition and Impersonal				-0.020	
	GCW	Function	1st Pers. Sing.		0.074**		
			1st Pers. Plur.		0.008	0.009	0.009
			2nd Person		-0.305***	-0.303***	-0.303***
			3rd Person		-0.038	-0.040	-0.040
			Impersonal		-0.068***		
			Conjunctions		-0.035***	-0.037***	-0.037***
			Prepositions		-0.021**	-0.039***	-0.039***
			Articles		-0.001	-0.001	-0.001
		Content	Nouns		0.017**	0.017**	0.017**
			Verbs		0.018	0.019*	0.019
			Aux. verbs		-0.022**	-0.023**	-0.023**
			Adverbs		-0.021**	-0.023**	-0.023**
			Adjectives		0.029***	0.031***	0.031***
Review	Psychometrics	Emotion	Positive	-0.139***	-0.130***	-0.129***	-0.129***
			Negative	0.008	0.011**	0.011**	0.011**
		Cognitive processes		-0.027***	-0.026***	-0.026***	-0.026***
	Readability	Informal language		0.038***	0.039***	0.039***	0.039***
		Word count		0.021***	0.024***	0.025***	0.025***
	Age			0.031***	0.017***	0.016***	0.016***
	User evaluation	Number of stars (linear)		-1.141***	-1.235***	-1.239***	-1.239***
		Number of stars (quadratic)		1.201***	1.305***	1.309***	1.309***
		Review is "fun"		0.183***	0.189***	0.190***	0.190***
		Review is "cool"		-0.138***	-0.162***	-0.163***	-0.163***
Reviewer	Expertise			0.134***	0.132***	0.132***	0.132***
	User evaluation	Number of stars		0.023***	0.025***	0.026***	0.026***
		Reviewer is "fun"		-0.145***	-0.143***	-0.143***	-0.143***
		Reviewer is "cool"		0.142***	0.140***	0.139***	0.139***
		Reviewer is "useful"		3.227***	3.556***	3.565***	3.565***
Product	Volume of reviews			0.000	-0.001	-0.001	-0.001
	Evaluations	Number of Stars		0.070***	0.068***	0.068***	0.068***
	Bar on premises			0.048***	0.050***	0.049***	0.049***

Table 5. (Contd.)

Product (Contd.)	Type of food	Seafood	0.085	***	0.086	***	0.085	***	0.085	***
		Barbeque	0.077	***	0.076	***	0.076	***	0.076	***
		Burgers	0.022		0.024		0.023		0.023	
		Sandwiches	0.023		0.023		0.023	*	0.023	
		Asian	0.024		0.021		0.021		0.021	
		American new	0.023	**	0.018		0.017		0.017	
		Tex-Mex	-0.025		-0.023		-0.024		-0.024	
		Steaks	-0.023		-0.030		-0.029	*	-0.029	
		Pizza	-0.036	**	-0.034	**	-0.035	**	-0.035	**
		Breakfast	-0.041	**	-0.036	**	-0.036	**	-0.036	**
		Mexican	-0.039	***	-0.040	***	-0.040	***	-0.040	***
		Chinese	-0.040	**	-0.040	**	-0.041	**	-0.041	**
		Delicatessen	-0.048		-0.044		-0.043	*	-0.043	
		Italian	-0.056	**	-0.055	**	-0.055	**	-0.055	**
		Japanese	-0.081	***	-0.078	***	-0.079	***	-0.079	***
		Thai	-0.113	***	-0.110	***	-0.110	***	-0.110	***
		American traditional	-0.120	***	-0.117	***	-0.116	***	-0.116	***
		Vegetarian	-0.171	***	-0.167	***	-0.166	***	-0.166	***
Intercept			0.884	***	0.957	***	0.964	***	0.964	***

Legend: ***, p≤.001; **, p≤.01; *, p≤.05.

of six groups: drive, cognition, affect, emotion, time orientation, and perception. The area we call relational and social environment comprises three groups of variables: social, cultural, and lifestyle. The 65 variables in these groups represent the proportion of words within each category in the LIWC dictionaries. In pairwise comparisons between syntax classes, we found that the proportion of these 65 variables that differ with p<.01 across pairs of syntax classes is 37% for those belonging to different pronoun clusters (personal vs. impersonal) and 42% when the comparison is between the syntax classes in C_A and C_B and the rest of the sample, C_C. The proportion fell to 17% among syntax classes within the personal cluster C_A and 7% for syntax classes within the impersonal cluster C_B. Despite the small differences in syntax classes in terms of GCW, the significant contrasts in psychosocial variables provided strong support for the diagnostic capabilities of the opening.

The psychosocial variables that showed the largest variation between clusters C_A and C_B are seven, reflecting social relations, such as friend, mother (+28% towards personal C_A), and communication (+19%), as well as focus in the present (+24%), likely to indicate the narrative's tense. Conversely, two positive valence-related variables had lower values in personal cluster C_A: positive tone (–18%) and positive emotion (–20%). Within clusters, the drive variable affiliation, the cognition variable certitude (e.g., really, actually), and the all-or-nothing variable (e.g., never, always) increased the absolute values of the estimates, either positive or negative, by +20 %, +15 %, and +10 %, respectively.

We now discuss the direction of the estimates for known predictors of PHR in relation to the literature in the first assessment of model validity. Table 5 shows that these benchmarking estimates are very consistent over models; thus, the term "estimates" in this section refers to those in models 1 to 4. The estimates for the constructs (measures in parentheses), review age (days since posted), and review readability (informal language and word count) are positive, consistent with meta-analytical research (Hong et al., 2017) showing that positive and stronger effects of these factors on PHR hold for "external" review platforms such as *Yelp!*, as opposed to "internal" ones such as an e-retailer's site. The quadratic and linear terms for review ratings (number of stars) yield a parabola with a negative slope for negative values, which becomes flat for positive values up to 1.0 SD but increases afterward, consistent with previous findings for external platforms, where extreme reviews tend to be viewed as more helpful (Hong et al., 2017). The same study validates the positive estimate for reviewer expertise (the number of reviews posted), which does not depend on the source of the review, whether external or internal.

We now compare the other estimates with those of reports from non-analytical studies. Consistent with the literature (Hong & Pittman, 2020; Jeong & Koo, 2015), positive valence (as a positive emotion) has negative estimates, whereas negative valence (as a negative emotion) has positive estimates with a smaller absolute value. This may reflect the interaction between negative valence and review objectivity; the subjective reviews in the study sample are likely to offset the effects of negative valence. We found no effect of volume (number of reviews posted for a restaurant), contrary to the findings (Hong & Pittman, 2020). These experiments were conducted on an internal review platform with low-involvement goods as opposed to an external platform and a higher-involvement service in the present study, which may explain the differences.

Positive answers to RQ2 or RQ3 hold practical significance. Given that the opening of the review contains essential information on the registers in the full text, it can be intuitively utilized by consumers. In other words, do consumers who rely solely on the opening segment, longer segments, or even the full-text impact PHR? When choosing which review to read, readers likely employ heuristics, such as reading the first or perhaps the second one. However, it remains untested whether they also use heuristics when reading individual reviews, for example, reading only the first few words. An implicit assumption that pervades the literature is that consumers read entire online review texts. This assumption is revealed by the consistent characterization of the full text (not parts of the text) in analytical and correlational studies and by the use of monolithic reviews as stimuli (not reviews where specific parts change). We refer to this assumption as the full-text assumption. To the best of our knowledge, this assumption has not been investigated previously. However, a sizable body of research conducted to date supports it. The studies cited earlier testing text-interpretive constructs suggest that to fully comprehend the complex subjective processes underlying those constructs, readers should read more than just the opening of the reviews. Some of these constructs have been listed earlier and include quality of arguments (Filieri et al., 2021), subjectivity (Ghose Ipeirotis, 2011) and, to a lesser extent, emotions (Xu et al., 2022) and review valence (Hong & Pittman, 2020). Consistent with this inference, linguistics scholarship has shown that smaller text segments tend to contain syntax patterns that are representative of those in the full text. Therefore, we propose the following hypothesis:

H1: The effect of linguistic characteristics on a diagnostic segment (opening) of a text (online reviews) can be explained by the linguistic characteristics of the full text.

7. Study 3: Confirmatory Experiments

The results of studies 1 and 2 suggest a relationship between the register prevailing in the review and PUR. Study 3 aimed to corroborate this relationship using experimental methods. Hence, we expect outcomes consistent with those in Table 5, specifically that PHR will decrease when readers are shown reviews of opening syntax patterns from impersonal pronoun clusters (C_B) and would increase when readers are shown reviews of personal pronoun clusters (C_B).

7.1. Participants and Procedure

Two hundred and three participants from Prolific (48.3% female; M_{age} = 37.10, SD_{age} =12.84) completed the study in exchange for monetary compensation. This study employed a review content (personal vs. impersonal) between-subjects design. All participants were asked to imagine a scenario in which they planned to go to a restaurant with which they were unfamiliar and decided to check the reviews before heading out to the restaurant. Participants were provided with eight reviews before being asked to make an overall judgment. Reviews of both conditions were selected from the traditional American food category. The eight reviews in the first condition were drawn from high-frequency openings in the subsample of 29,298 first-person personal pronouns, verbs, and nouns, with estimates of 0.114 ($p<.0001$) and 0.080 ($p<.0001$). As explained in Study 2, the register in these reviews is a first-person, balanced narrative of the experience. Thus, we keep *personal* as the label for this condition. Reviews in the second condition were drawn from high-frequency openings in a subsample of 16,302 reviews with impersonal pronouns and no prepositions, estimating -0.095 ($p<.0001$). Study 2 revealed that the register in these reviews is an omniscient narrative with a positive tone or emotional leaning. Thus, consistent with Study 2, we refer to this condition as *impersonal*. (Please refer to Online Appendix B for a review of the specimens used in Study 3.)

After going over the reviews, participants responded to a five-item review usefulness scale on a seven-point scale (1 = "strongly disagree"; 7 = "strongly agree") (adapted from Bailey & Pearson, 1983; Cheung et al., 2008; and Wu & Shaffer, 1987): 1) The reviews are valuable 2) The reviews are informative 3) The reviews are useful and 4) People who left reviews are trustworthy 5) People who left reviews are reliable. Scale reliability was sufficiently high (Cronbach α = .93), and hence, we averaged participants' scores across the five items to a "review usefulness" index.

7.2. Results

We conducted a one-way ANOVA on the review usefulness. Consistent with the sign of the regression estimates, the results revealed that participants in the personal review condition perceived the reviews as more useful than those in the impersonal review condition ($M_{personal}$ = 5.88, $SD_{personal}$ = 0.88 vs. $M_{impersonal}$ = 5.57, $SD_{impersonal}$ = 1.15, $F(1, 201) = 4.89$, $p = .028$). These results validate the findings of Correlational Study 2. Specifically, participants exposed to personal reviews (versus impersonal reviews) rated them as more useful.

8. Study 4: Testing the Full Text Assumption

The syntax patterns of reviews in Study 2 were created using the review opening as a single criterion, assuming that openings would consistently reflect the same linguistic register or type of full text. As explained in the background section, this simplification is not only useful but also supported by linguistic scholarship. However, as discussed earlier, there are at least two alternative explanations for the differential effects of syntax patterns on PUR: readers could use the openings of the texts as heuristics, not reading the full text, or they could read the full text or a significant part thereof. In H1, we argue for the latter. Although Study 3 corroborates the findings of Study 2, it does not test H1. Study 4 accomplished this goal.

8.1. Participants and Procedure

Two hundred and eleven participants from Prolific (47.9% female; M_{age} = 33.91, SD_{age} = 12.27) completed the study in exchange for monetary compensation. This study employed a (opening: personal vs. impersonal) between-subjects design. Similar to the previous study, all participants were asked to imagine a scenario in which they planned to go to a new restaurant (this time Asian) that they were unfamiliar with and decided to check out the reviews before heading out to the restaurant. They were provided with nine reviews before judging the entire review sample. We selected reviews from an actual sample used in Studies 1 and 2. The difference between this study and Study 3 was that, unlike Study 3, in Study 4, we used the same reviews under both conditions, except for the openings. More specifically, by controlling for content, the same nine reviews included in the personal condition were also used in the impersonal condition, except for the opening (first five words) of each review (please refer to online Appendix C for the review specimens used in Study 4). The openings used in the personal condition were all written in the first-person singular, and those used in the impersonal condition praised the business or some features of the business.

After reviewing the reviews, the participants responded to the same five-item review usefulness scale used in Study 3. Scale reliability was again sufficiently high (Cronbach α = .93), so we averaged participants' scores across the five items to create a "review usefulness" index.

8.2. Results

We conducted a one-way ANOVA on the review usefulness. The result revealed that participants in the personal opening condition did not perceive reviews as more useful than those in the impersonal opening condition ($M_{personal}$ = 4.96, $SD_{personal}$ = 1.34 vs. $M_{impersonal}$ = 5.07, $SD_{impersonal}$ = 1.00, $F(1, 209) = .44$, $p = .51$). Therefore, Study 4 demonstrates that review opening alone does not influence the perception of review usefulness. Specifically, with the remainder of the full text held constant, participants who were exposed to personal openings (versus impersonal openings) did not rate them as more useful. These results and those of the previous study suggest that it is the full text of the review that impacts individuals' judgement regarding the PUR, not the opening by itself, thus supporting H1.

9. General Discussion

Table 6 summarizes the findings of these four studies. They have consistently pointed to two linguistic varieties within the online product review genre. These two classes referred to as clusters C_A and C_B, are characterized by a specific class of pronouns. Cluster C_A is defined by first-person singular pronouns (I, me, my, mine), whereas cluster C_B is associated with impersonal pronouns (this, these, that, those) or no pronouns. Pronouns play a crucial role in defining the standpoint of the narrative and, for that reason, are usually found in the opening of reviews. This raises the question of whether, instead of the remarkable variability of the opening in terms of GCW, it is a single pronoun that gives the first text segment its diagnostic power. We believe that the answer is no. These two specific types of pronouns represent narrative types that differ in terms of their focus on interpersonal relationships, tense of the narrative, and valence of assessment. Pronouns are integral components of these more complex language structures; they do not work in isolation, even when they co-occur with other GCW to form the varieties we label personal (C_A) and impersonal (C_B).

Two characteristics indicate that personal and impersonal reviews represent distinct language registers (Biber & Conrad, 2019). First, they differ in the type of GCW they present and how they are connected, especially earlier in the review. Second, personal and impersonal reviews suggest contrasting communication goals. Personal reviews aim to offer a balanced assessment from the visitor's perspective with the implicit goal of providing balanced information. By contrast, impersonal reviews tend to praise food and venues, hinting at a persuasion goal while avoiding the inclusion of the narrator within the venue context, which can potentially diminish credibility. This explanation was empirically supported. For instance, a meta-analysis of 17 studies on health communication showed that first-person narratives are twice as likely as third-person narratives to affect attitudes and behaviors (Winterbottom et al., 2008). In the Methods section, we conjecture that the role of pronouns in openings may be related to credibility. Visiting a restaurant, ordering, and consuming a meal entails time, opportunity, and even psychological costs beyond the obvious monetary costs. Therefore, the primary motivation behind reading restaurant reviews is likely to be a reduction in uncertainty, an objective that credible personal reviews are more likely to achieve than less credible impersonal reviews. This is consistent with the three estimates for first-level personal pronouns in Table 5 being positive and with two out of four significant estimates for first-level impersonal pronouns being negative. The sign of impersonal estimates is reversed to positive only when there are words in the opening that provide information, that is, content GCWs, such as nouns and adjectives.

Beyond its composition in terms of GCW, the variability of the opening segment remains very important. As depicted in Figure 1b), as the reviews grow in length, they tend to converge toward similar proportions of GCW. This can be attributed to the reviewer's tendency to minimize cognitive effort by using more industry-specific jargon and common expressions, which are characteristics of the genre, as well as the increased use of commonplace words, as

Table 6: Summary of Results by Study

Study 1	Study 2	Study 3
Searching for diagnostic segments. - The **opening**, defined as the first five words in the review, presents two major characteristics of diagnostic segments: • *Word variability peaks*: Measured as the diversity of unique words. • *Sentence structure changes rapidly*: Major changes in composition by GCW, which stabilize after word position six	Part 1: Classifying opening word arrays into syntax patterns (SP). - Types of SP: • *Cluster C_A*: 3 SP with 1st person singular pronouns (I, me, my, mine) • *Cluster C_B*: 5 SP with impersonal (this, these, that, those) or no pronouns. - Syntax is successfully operationalized as a set of 8 simple patterns. Part 2: A statistical model of PHR with SP as predictors - *Syntax matters*: SP outperforms simple proportions of GCW. - *Pronouns matter*: A model with a hierarchical array of pronouns + other GCW in C_A and C_B outperforms a simpler model with non-hierarchical syntax patterns in C_A and C_B - *First person pronouns are better than impersonal pronouns*: Estimates are generally positive for "I, me, mine" patterns in C_A (mean 0.080) but negative for "this, that, those" patterns in C_B (mean -0.041)	Confirming that syntax patterns have measurable effects in practice and the direction of effects in the statistical model. - An experimental study shows that participants reading reviews in cluster C_A (e.g., pronouns I, me, mine) perceived the reviews significantly as more useful than participants reading reviews in cluster C_B (e.g., pronouns this, that, those) - This confirms the direction of effects found in the statistical model Study 4 Confirming the expectation that the effects of the opening occur because its syntax patterns reflect characteristics of the full review. - An experimental study finds no statistical difference in PHR between participants reading reviews as posted online vs. participants reading reviews with opening segments altered to conform to different clusters. - Collaterally, the "full text assumption" in online reviews literature is supported.

evidenced by the decline in unique word diversity in Figure 1a). Furthermore, building on a previous conjecture, review writers are not only likely to be cognitively fresher at the beginning of their writing task but also to consciously put more effort into crafting an appealing opening. Secondly, the population of reviews between 20 and 100 words is very heterogeneous in terms of length. This presents a challenge since we compute diversity and proportions of GCW across different positions in the text. Reviews, akin to short stories, exhibit diverse structures that may vary in length or reviewer preferences. Known dramatic (story) structures may vary from the triad *beginning-middle-end* in Aristotle's *Poetics* (Amos et al., 2022) to Campbell's *The Hero's Journey* (Campbell & Dudley, 2020). For instance, in Figure 1b), the words in the full texts of all 20-word reviews are lumped together with the first 20 words of reviews with 60 or more words. This results in a mix of text sections that play different roles. Averaging such varied sentences should yield more homogeneous proportions of GCW, as in word positions six and higher in Figure 4b). This mixing of text sections provides another explanation for the diagnostic significance of the opening segment, which is not affected by this process. The first five words are most likely part of the initial section of any review in the study sample, resembling the Aristotelian beginning, particularly in reviews with a minimum length of 20 words.

10. Limitations and Future Research

In addition to the point raised earlier, the mixing of text sections due to varying text lengths may also mask interesting sections of online reviews other than the opening. This is a measurement issue that has a plausible solution. For instance, in future studies, reviews can be split into sections regardless of word length and the chosen dramatic structure framework. Thus, the proportions of GCW can be averaged by section rather than by word position. This would help researchers further refine diagnostic review segments or parts, departing from the fixed-length segment established in the present research. These new diagnostic tools may help uncover even more varieties of speech in online reviews and assess their contribution to PHR or any other construct of interest. On a first read, the results of Study 4 may appear to offer no grounds for such endeavors, as it seems logical to continue characterizing full-review texts if readers indeed tend to read them in full. However, this is true only to a certain extent. Diagnostic segments can considerably simplify the analysis of large datasets in both academia and industry while also offering straightforward decision guidelines for small businesses. For instance, the present study shows that first-person pronouns with pros and cons in restaurant reviews merit special attention. It is also possible to expand research on diagnostic segments to encompass consumer reviews in different product categories or other forms of consumer-generated content, such as influencer speech. Beyond the obvious roles of personality, charisma, subject matter expertise, and brand consistency, one might wonder whether speech variety plays a role in influencer persuasion. Can we identify a few diagnostic phrases or grammatical constructions that signal a higher chance of influencer success? The range of possibilities broadens as we consider rhetoric, semiotics, and the interface between language and psychology (Pennebaker, 2013).

Acknowledgment

The costs of recruitment of participants in studies 3 and 4 were covered by the Department of Management and Marketing, University of Massachusetts Dartmouth.

REFERENCES

Ahmad, S. (2017). Uncovering the paths to helpful reviews using fuzzy-set qualitative comparative analysis. *Journal of Marketing Analytics*, *5*(2), 47–56.

Ahmad, S. N., & Laroche, M. (2015). How do expressed emotions affect the helpfulness of a product review? Evidence from reviews using latent semantic analysis. *International Journal of Electronic Commerce*, *20*(1), 76–111.

Amos, C., Zhang, L., King, S., & Allred, A. (2022). Aristotle's modes of persuasion and valence effects on online review trustworthiness and usefulness. *Journal of Marketing Communications*, *28*(4), 360–391.

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. In *International Journal of Corpus Linguistics*, *14*(3), 275–311. John Benjamins.

Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press. https://doi.org/10.1017/9781108686136

Biswas, B., Sengupta, P., & Ganguly, B. (2022). Your reviews or mine? Exploring the determinants of "perceived helpfulness" of online reviews: A cross-cultural study. *Electronic Markets*, *32*(3), 1083–1102. https://doi.org/10.1007/s12525-020-00452-1

Campbell, J. & Kudler, D. (2020). *The hero with a thousand faces: vol. 30th anniversary special edition*. Joseph Campbell Foundation; eBook Collection (EBSCOhost). https://www.libproxy.umassd.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2680592&site=ehost-live&scope=site

Casaló, L. V., Flavián, C., Guinalíu, M., & Ekinci, Y. (2015). Avoiding the dark side of positive online consumer reviews: Enhancing reviews' usefulness for high risk-averse travelers. *Journal of Business Research*, *68*(9), 1829–1835. https://doi.org/10.1016/j.jbusres.2015.01.010

Ceylan, G., Diehl, K., & Proserpio, D. (2024). Words meet photos: When and why photos increase review helpfulness. *Journal of Marketing Research*, *61*(1), 5–26. https://doi.org/10.1177/00222437231169711

Chatterjee, S. (2023). Effect of construal level on the drivers of online-review-helpfulness. *Electronic Commerce Research*, 1–29. https://doi.org/10.1007/s10660-023-09716-2

Chen, M.-J., & Farn, C.-K. (2020). Examining the influence of emotional expressions in online consumer reviews on perceived helpfulness. *Information Processing & Management*, *57*(6), 102266, https://doi.org/10.1016/j.ipm.2020.102266

Cheung, C. M., Lee, M. K., & Rabjohn, N. (2008). The impact of electronic WOM: The adoption of online opinions in online customer communities. *Internet Research*, *18*(3), 229–247.

Choi, H. S., & Leon, S. (2020). An empirical investigation of online review helpfulness: A big data perspective. *Decision Support Systems*, *139*, 113403, https://doi.org/10.1016/j.dss.2020.113403

Chou, Y.-C. (2023). How Much Is Too Much? The nonlinear link between emotional arousal and review helpfulness. *Decision Support Systems*, *175*, https://doi.org/10.1016/j.dss.2023.114035

Chong, A. Y. L., Khong, K. W., Ma, T., McCabe, S., & Wang, Y. (2018). Analyzing key influences of tourists' acceptance of online reviews in travel decisions. *Internet Research*, *28*(3), 564–586.

De Pelsmacker, P., Dens, N., & Kolomiiets, A. (2018). The impact of text valence, star rating and rated usefulness in online reviews. *International Journal of Advertising: The Review of Marketing Communications*, *37*(3), 340–359. https://doi.org/10.1080/02650487.2018.1424792

Dellarocas, C., Xiaoquan Zhang, & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45.

Dobson, A. J., & Barnett, A. G. (2008). *An introduction to generalized linear models* (3rd ed.). CRC Press.

Dong, X., Liao, J., Chang, Y. & Zhou, R. (2022). Consumer responses to online advertising and online reviews: The mediating role of cognitive responses and affective responses. *Journal of Electronic Commerce Research*, *23*(4), 207–222.

Duan, Y., Liu, T., & Mao, Z. (2022). How online reviews and coupons affect sales and pricing: An empirical study based on e-commerce platform. *Journal of Retailing & Consumer Services*, *65*, 102846, https://doi.org/10.1016/j.jretconser.2021.102846

Erkan, I., & Evans, C. (2018). Social media or shopping websites? The influence of eWOM on consumers' online purchase intentions. *Journal of Marketing Communications*, *24*(6), 617–632. https://doi.org/10.1080/13527266.2016.1184706

Felbermayr, A., & Nanopoulos, A. (2016). The role of emotions for the perceived usefulness in online customer reviews. *Journal of Interactive Marketing*, *36*, 60–76. https://doi.org/10.1016/j.intmar.2016.05.004

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, *68*(6), 1261–1270. https://doi.org/10.1016/j.jbusres.2014.11.006

Filieri, R., Javornik, A., Hang, H., & Niceta, A. (2021). Environmentally framed eWOM messages of different valence: The role of environmental concerns, moral norms, and product environmental impact. *Psychology & Marketing*, *38*(3), 431–454. https://doi.org/10.1002/mar.21440

Ghose, A., & Ipeirotis, P. G. (2011). Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. *IEEE Transactions on Knowledge & Data Engineering*, *23*(10), 1498–1512.

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, *23*(4), 545–560. https://doi.org/10.1287/mksc.1040.0071

Hair, M., & Ozcan, T. (2018). How reviewers' use of profanity affects perceived usefulness of online reviews. *Marketing Letters: A Journal of Research in Marketing*, *29*(2), 151–163. https://doi.org/10.1007/s11002-018-9459-4

Hong, H., Xu, D., Wang, G. A., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems*, *102*, 1–11.

Hong, S., & Pittman, M. (2020). eWOM anatomy of online product reviews: Interaction effects of review number, valence, and star ratings on perceived credibility. *International Journal of Advertising*, *39*(7), 892–920.

Jeong, H.-J., & Koo, D.-M. (2015). Combined effects of valence and attributes of e-WOM on consumer judgment for message and product. *Internet Research*, *25*(1), 2–29.

Jeong, J. E. & Yeu, M. (2023). Can mental imagery enhance review helpfulness and product evaluation? The role of contextual background and mental simulation. *Journal of Research in Interactive Marketing*, *17*(6), 959-974. https://doi.org/10.1108/JRIM-08-2022-0237

Kim, S. J., Maslowska, E., & Malthouse, E. C. (2018). Understanding the effects of different review features on purchase probability. *International Journal of Advertising*, *37*(1), 29–53.

Kolomiiets, A., Dens, N., & De Pelsmacker, P. (2016). The wrap effect in online review sets revisited: How perceived usefulness mediates the effect on intention formation. *Journal of Electronic Commerce Research*, *17*(4), 280–288.

Kupor, D., & Tormala, Z. (2018). When moderation fosters persuasion: The persuasive power of deviatory reviews. *Journal of Consumer Research*, *45*(3), 490–510. https://doi.org/10.1093/jcr/ucy021

Lee, S., & Choeh, J. Y. (2016). The determinants of helpfulness of online reviews. *Behaviour & Information Technology*, *35*(10), 853–863. https://doi.org/10.1080/0144929X.2016.1173099

Lee, Y. & Nelder, J. A. (1996). Hierarchical generalized linear models. *Journal of the Royal Statistical Society*, *58*(4), 619–678.

Li, C., Kwok, L., Xie, K. L., Liu, J., & Ye, Q. (2023). Let photos speak: The effect of user-generated visual content on hotel review helpfulness. *Journal of Hospitality & Tourism Research, 47*(4), 665–690. https://doi.org/10.1177/10963480211019113

Li, H., Wang, X., Wang, S., Zhou, W., & Yang, Z. (2023). The power of numbers: An examination of the relationship between numerical cues in online review comments and perceived review helpfulness. *Journal of Research in Interactive Marketing*, *17*(1), 126–139. https://doi.org/10.1108/JRIM-09-2021-0239

Li, J., Zhang, Y., Li, J., & Du, J. (2023). The role of sentiment tendency in affecting review helpfulness for durable products: Nonlinearity and complementarity. *Information Systems Frontiers*, *25*(4), 1459–1477. https://doi.org/10.1007/s10796-022-10292-3

Liang, T. P., Cheng, J. C., Saini, V., & Shih-Chieh Hsu, J. (2021). Is being helpful good enough for online reviews? Exploring the role of information credibility and data source through meta-analysis. *Journal of Electronic Commerce Research*, *22*(4), 336–362.

Liao, J., Li, C., & Filieri, R. (2022). The role of humor in management response to positive consumer reviews. *Journal of Interactive Marketing*, *57*(2), 323–342.

Liu, A. X., Xie, Y., & Zhang, J. (2019). It's not just what you say, but how you say it: The effect of language style matching on perceived quality of consumer reviews. *Journal of Interactive Marketing*, *46*, 70–86. https://doi.org/10.1016/j.intmar.2018.11.001

Liu, X., Wu, X., Shi, W., Tong, W., & Ye, Q. (2022). The impacts of electronic word-of-mouth on high-involvement product sales: Moderating effects of price, brand origin, and number of customers. *Journal of Electronic Commerce Research*, *23*(3), 177-189.

Lopes, A. I., Dens, N., & De Pelsmacker, P. (2022). Valence and attribute repetition in negative sets of online reviews: (When) Can positive reviews overcome negative ones? *Journal of Electronic Commerce Research*, *23*(1), 1–12.

Maslowska, E., Malthouse, E. C., & Bernritter, S. F. (2017). Too good to be true: The role of online reviews' features in probability to buy. *International Journal of Advertising*, *36*(1), 142–163.

McFadden, D. (1977). *Quantitative methods for analyzing travel behaviour of individuals: Some recent developments*. Retrieved from: https://eml.berkeley.edu/reprints/mcfadden/quant_methods.pdf .

Mitra, S., & Jenamani, M. (2021). Helpfulness of online consumer reviews: A multi-perspective approach. *Information Processing & Management*, *58*(3), 102538, https://doi.org/10.1016/j.ipm.2021.102538

Moradi, M., Dass, M., & Kumar, P. (2023). Differential effects of analytical versus emotional rhetorical style on review helpfulness. *Journal of Business Research*, *154*, 113361. https://doi.org/10.1016/j.jbusres.2022.113361

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185–200.

Namvar, M., & Chua, A. Y. K. (2023). The impact of context clues on online review helpfulness. *Internet Research*, *33*(3), 1015–1030. https://doi.org/10.1108/INTR-02-2021-0093

Park, K., Lee, S., Doosti, S., & Tan, Y. (2023). Provision of helpful review videos: Effects of video characteristics on perceived helpfulness. *Production & Operations Management*, *32*(7), 2031–2048. https://doi.org/10.1111/poms.13969

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, *50*, 67–83. https://doi.org/10.1016/j.annals.2014.10.007

Pennebaker, J. W. (2013). *The secret life of pronouns: What our words say about us*. Bloomsbury Publishing.

Pumawirawan, N., De Pelsmacker, P., & Dens, N. (2012). Balance and sequence in online reviews: How perceived usefulness affects attitudes and intentions. *Journal of Interactive Marketing*, *26*(4), 244–255.

Purnawirawan, N., Eisend, M., De Pelsmacker, P., & Dens, N. (2015). A meta-analytic investigation of the role of valence in online reviews. *Journal of Interactive Marketing*, *31*, 17–27. https://doi.org/10.1016/j.intmar.2015.05.001

Ravula, P., Bhatnagar, A., & Gauri, D. K. (2023). Role of gender in the creation and persuasiveness of online reviews. *Journal of Business Research*, *154*, 113386. https://doi.org/10.1016/j.jbusres.2022.113386

Saad, R. (2005). Discovery, development, and current applications of DNA identity testing. *Proceedings (Baylor University. Medical Center)*, *18*(2), 130–133. https://doi.org/10.1080/08998280.2005.11928051

Shah, A., Gohary, A., & Lie, D. (2023). Looking for negativity! How review sentiment and review subjectivity form review helpfulness across cultures. *Journal of Consumer Behaviour, 23(3):*1193-1207, https://doi.org/10.1002/cb.2277

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Singh, J. P., Irani, S., Rana, N. P., Dwivedi, Y. K., Saumya, S., & Kumar Roy, P. (2017). Predicting the "helpfulness" of online consumer reviews. *Journal of Business Research*, 70, 346–355. https://doi.org/10.1016/j.jbusres.2016.08.008

Wang, Y., Wang, J., & Yao, T. (2019). What makes a helpful online review? A meta-analysis of review characteristics. *Electronic Commerce Research*, *19*(2), 257–284.

Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear-models, and Gauss-Newton method. *Biometrika*, *61*(3), 439–447.

Winterbottom, A., Bekker, H. L., Conner, M., & Mooney, A. (2008). Does narrative information bias individual's decision-making? A systematic review. *Social Science & Medicine*, *67*(12), 2079–2088. https://doi.org/10.1016/j.socscimed.2008.09.037

Wu, Y., & Wu, J. (2016). The impact of user review volume on consumers' willingness-to-pay: A consumer uncertainty perspective. *Journal of Interactive Marketing*, *33*, 43–56. https://doi.org/10.1016/j.intmar.2015.11.001

Xu, D., Ye, Q., Hong, H., & Sun, F. (2022). Emotions for attention in online consumer reviews: The moderated mediating role of review helpfulness. *Industrial Management & Data Systems*, *122*(3), 729–751.

Xu, C., Zheng, X., & Yang, F. (2023). Examining the effects of negative emotions on review helpfulness: The moderating role of product price. *Computers in Human Behavior, 139*, 107501. https://doi.org/10.1016/j.chb.2022.107501

Yi, J., & Oh, Y. K. (2022). The informational value of multi-attribute online consumer reviews: A text mining approach. *Journal of Retailing & Consumer Services*, *65*, 102519, https://doi.org/10.1016/j.jretconser.2021.102519

Zhu, F., & Zhang, X. (Michael). (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133–148.

Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research*, *15*(4), 267-280.