

# How do restaurant reviews differ across price segments?\*

**Analysing written review text with natural language processing**

Lucca Zachmann

June 21, 2019

## **Abstract**

Restaurant visitors have a critical tongue. The degree of criticism, however, depends on the expectations that visitors have for a gastronomic experience. This means, expectations are an increasing function of the restaurant's price category. This paper argues that restaurant reviewers tend to "show off" their culinary capital in written restaurant reviews for the most prestigious establishments, i.e. fine dining, but take a moderate approach when reviewing commonplace restaurants. Review texts are more extensive for fine dining establishments as compared to mid range restaurants or cheap eats. Fine dining reviewers are more negative in their written review text at both ends of the numerical review scale, i.e. for terrible and excellent experiences. Word clustering increases and linguistic diversity decreases with the restaurant's price category. Central elements in excellent reviews of fine dining establishments are the interplay between food and wine, the food itself in mid range restaurants and quality/price considerations in cheap eats.

---

\*. This research paper acts as the final assessment of the course *Psychology of Eating* taught by Prof. Peter Todd in May 2019 at the University of Gastronomic Sciences in Pollenzo, Italy.

# 1 Introduction

Restaurant visitors are a critical crowd. Expectations for gastronomic services are high. The food is expected to be tasty and attractively presented while service is anticipated to work efficiently, prompt, friendly and professionally. However, the degree to which restaurant visitors expect their eating experience to be depends on the type or style of the establishment and may even be proportional to the price range in which it operates.<sup>1</sup> For instance, an experience that would be considered legitimate in a fast food restaurant might be deemed unacceptable in a fine dining environment.

In the last years, user-generated online reviews have tremendously gained in popularity. Restaurant reviews in the world wide web are a powerful tool to produce and disseminate information in a peer-to-peer communication network. Restaurant visitors are therefore no longer just consumers of gastronomic services but also producers of online content. They are perceived as *prosumers*.

Several studies have been undertaken to better understand the patterns behind restaurant reviews. For instance, a study in the US has shown that restaurant reviews depend greatly on the locality where the restaurants are located. Restaurants in predominantly White neighbourhoods were considered “authentic” and “cozy,” while they were criticised for a lack of dining options and an “atmosphere of dirt and danger” in Black neighbourhoods.<sup>2</sup> Another study compared restaurant reviews from two different geographic contexts, that are webpages from the US and Hong Kong. The findings suggest that there exist similarities in the formal properties and differences in the review content among the two webpages.<sup>3</sup>

This paper is inspired by Vásquez and Chik (2015) who “demonstrate that online restaurant reviews provide a means through which individuals can display their culinary capital to an audience who is likely to share similar interests”.<sup>4</sup> Having something to say about restaurants has become “part of a modern, sophisticated, urban identity.”<sup>5</sup> Therefore, the hypothesis in this paper is that restaurant reviewers tend to “show off” their culinary capital in written restaurant reviews for the most prestigious establishments, i.e. fine dining, but take a moderate approach when reviewing commonplace restaurants. In doing so, their verbal rating is more critical for pricy high-expectation diners than it is for everyday restaurants.

---

1. Michele L. Bailey, “Meeting customer expectations in the foodservice industry: A “Mystery Diner” program” (Thesis at the Rochester Institute of Technology, 1994).

2. Sharon Zukin, Scarlett Lindeman, and Laurie Hurson, “The omnivore’s neighborhood? Online restaurant reviews, race, and gentrification,” *Journal of Consumer Culture* 17, no. 3 (2015): 459–479.

3. Alice Chik and Camilla Vásquez, “A comparative multimodal analysis of restaurant reviews from two geographical contexts,” *Visual Communication* 16, no. 1 (2017): 3–26.

4. Camilla Vásquez and Alice Chik, ““I Am Not a Foodie...”: Culinary Capital in Online Reviews of Michelin Restaurants,” *Food and Foodways* 23, no. 4 (2015): 231–250.

5. Mitchell Davis, *A taste for New York: Restaurant reviews, food discourse, and the field of gastronomy in America* (New York University, 2009).

The hypothesis is tested by analysing restaurant reviews from fifteen best-in-class restaurants which operate in three different price categories. The paper is therefore structured as follows. Section 2 describes the data whereas section 3 explains the methodologies with which the data is analysed. Sections 4 and 5 present and discuss the results respectively. Finally, section 6 concludes.

## 2 Data

The data for this study comes from <https://www.tripadvisor.com>, the largest social travel website in the world. I collect 5’080 online restaurant reviews that are written in English from Copenhagen (DK) by means of web scrapping.<sup>6</sup> The five best ranking restaurants in their respective price category, i.e. cheap eats, mid range and fine dining, are considered (best-in-class). Cheap eats are predominantly fast food restaurants, mid range restaurants offer international cuisine while fine dining focuses on Scandinavian cuisine. Table 2 provides additional summary statistics.

	Restaurant type	# of reviews	Avg. rating	Avg. word count per review
1	Cheap eats	1’113	4.85	24.70
2	Mid range	2’027	4.82	27.94
3	Fine dining	1’940	4.77	43.78

Table 1: Summary statistics

At first glance becomes apparent that cheap eats are with 1’113 reviews the least frequently reviewed restaurant type, followed by fine dining (1’940 reviews) and mid range (2’027 reviews). The average rating is a decreasing function of the restaurant’s price range, i.e. cheap eats are on average higher rated than mid range or fine dining restaurants.<sup>7</sup> However, customers write more extensive reviews in terms of words for fine dining restaurants (44 words per review) as compared to mid range (28) or cheap eats (25).

## 3 Methodology

Tripadvisor provides the reviewer with two different rating schemes. The first, and more obvious, is the numeric rating in which the reviewer can rate the restaurant visit on a scale from 1 (terrible) to 5 (excellent) in one-unit increments. The second rating occurs through a verbal rating in which the reviewer writes a short review text. A crucial difference in the two

---

6. The choice of the location is random.  
7. The average ratings are comparably high due to the web scraping of the five best-in-class restaurants in their respective category.

ratings is that the reviewer can only express culinary capital in the written review. The words used in this text leave the reader of the review with a sentiment rating, i.e. does the sentence or clause classify it as expressing a positive, negative or neutral opinion?<sup>8</sup> Therefore, I apply in a first step a sentiment analysis to numerically rate the sentiment expressed in the written review text. Each review text is continuously classified from -1 to +1 depending on the sentiment that is expressed. I call this rating verbal review score.

In a second step, I use only excellent restaurant reviews, i.e. those with a numeric rating of 5, and analyse the wording in these written reviews for each restaurant category. This word clustering allows to determine the used language and gives insight into what reviewers considered to be important for an excellent review. I therefore split the dataset into word bigrams. A word bigram is a sequence of each two adjacent words in a string variable. I remove stop words such as *the*, *a*, *an*, *in* and many more from the dataset and keep only bigrams that appear more than 25 times. The bigrams are then discussed according to six categories based on theoretical constructs used in the restaurant customer evaluation literature.<sup>9</sup> The categories are: (1) Quality of food and beverage (2) Relative importance of food versus beverage (3) Labour (4) Ambience and atmosphere (5) Price and value and (6) Other amenities.

## 4 Results

### 4.1 Sentiment analysis

Figure 1 depicts the numeric and verbal review rating scores for the three restaurant categories. Four observations can be taken from the figure. First, verbal review ratings are an increasing function of numeric review ratings. This is intuitive in the sense that the average reviewer who ranks a restaurant numerically low (high) also writes a negative (positive) verbal review. Second, mid range restaurants are consistently better rated in terms of verbal reviews compared to fine dining restaurants. Verbal review ratings of cheap eats perform worst for poor (2) or average (3) experiences. Henceforth, the third observation from the figure is that, at the extremes, i.e. for terrible (1) and excellent (5) restaurant experiences, verbal review scores for fine dining are lower compared to mid range restaurants and cheap eats. Fourth, terrible experiences are verbally rated particularly critical for fine dining restaurants since the average verbal review score is negative.

---

8. Bing Liu et al., "Sentiment Analysis and Subjectivity.," *Handbook of natural language processing* 2, no. 2010 (2010): 627–666.

9. Karl Titz, Je'Anna Lanza-Abbott, and Glenn Cordúa Cruz, "The Anatomy of Restaurant Reviews," *International Journal of Hospitality and Tourism Administration* 5, no. 1 (2004): 49–65.

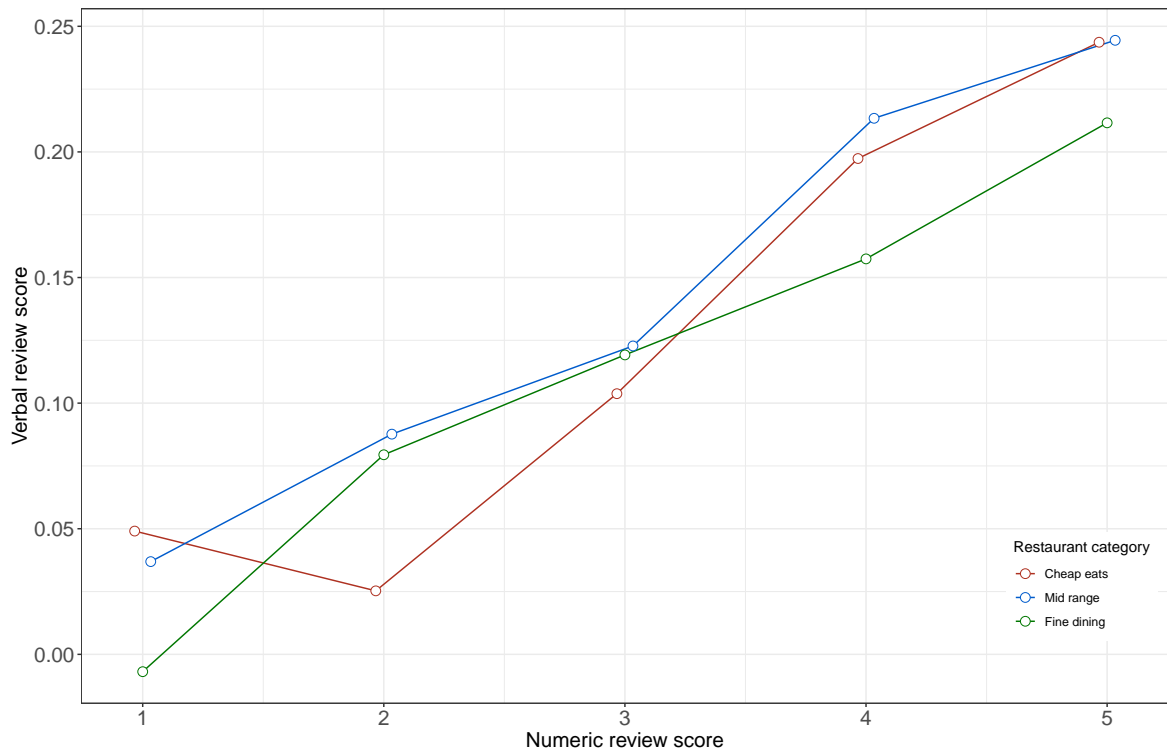


Figure 1: Numeric vs. verbal review ratings per restaurant category

## 4.2 Word clustering

Figures 2, 3 and 4 show bigram networks of the most frequently used wordings of excellent eating experiences for cheap eats, mid range and fine dining respectively. The figures are read by following the direction of the arrows (for instance: late  $\bullet \rightarrow$  breakfast is read as late breakfast). The darker the edges are, the more frequently the bigrams are mentioned.

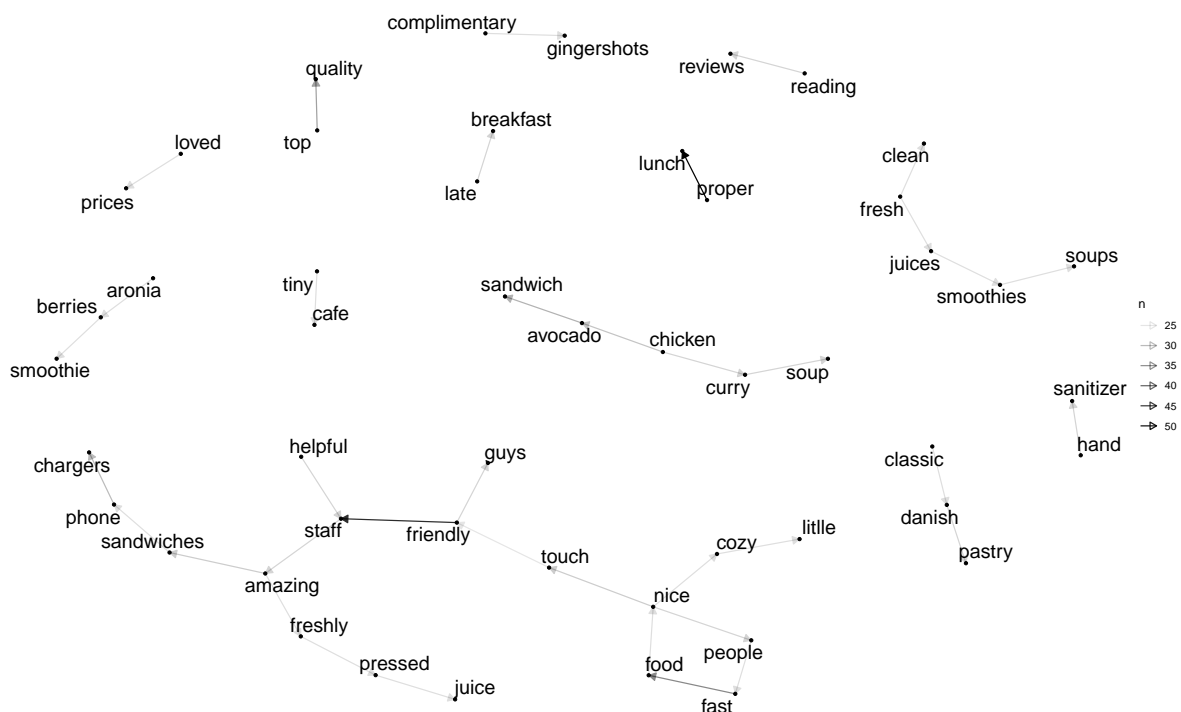


Figure 2: Bigram network of cheap eats

Figure 2 shows the bigram network for cheap eats. The five most frequently mentioned bigrams are: proper lunch (52 appearances), friendly staff (47), fast food (36), top quality (33) and avocado sandwich (31).

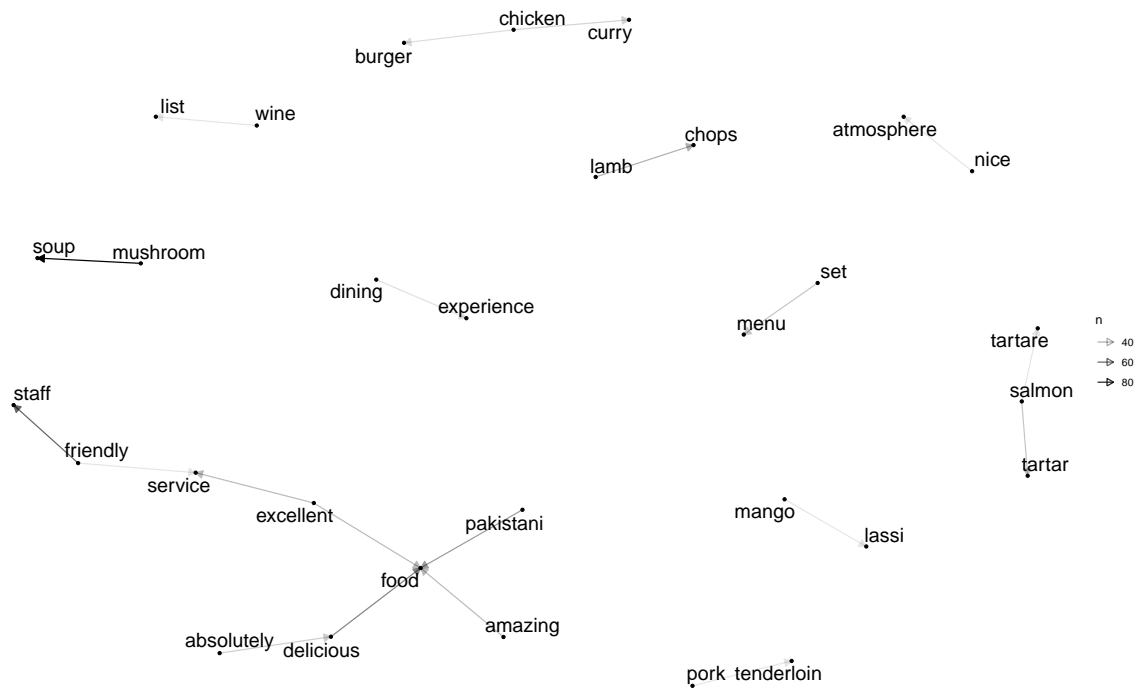


Figure 3: Bigram network of mid range restaurants

Figure 3 displays the bigram network for mid range restaurants. Top five bigrams are mushroom soup (89), friendly staff (60), delicious food (52), pakistani food (45) and lamb chops (39).

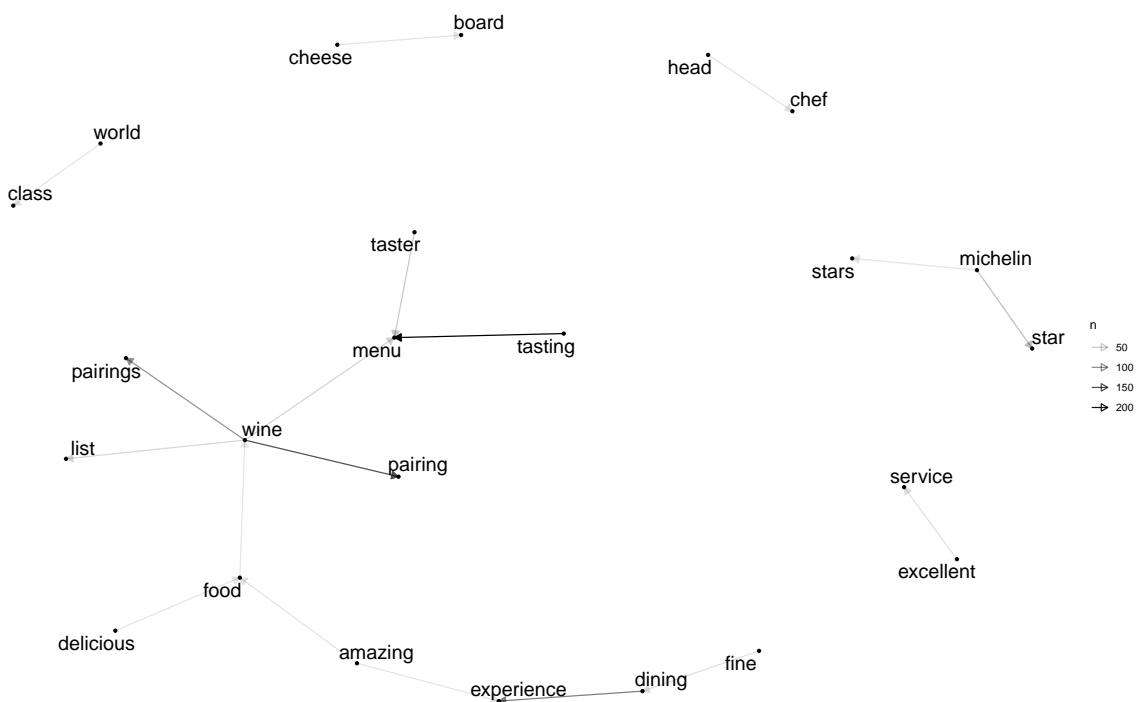


Figure 4: Bigram network of fine dining restaurants

Figure 4 presents the bigram network for fine dining restaurants. The dominant bigrams

are tasting menu (215), wine pairing (150), dining experience (110), wine pairings (94) and michelin star (58).

## 5 Discussion

The figures in the previous subsection 4.2 make clear that the diversity of bigrams, i.e. the number of different nodes, decreases and the clustering, i.e. the number of occurrences of each node, increases with the restaurant’s price range. That is, cheap eats show 48 different nodes that appear more than 25 times, mid range 31 nodes and fine dining 24 different nodes. However, the sum of the appearances<sup>10</sup> of the top five mentioned bigrams is 199 for cheap eats, 285 for mid range and 627 for fine dining. Therefore, when reviewing fine dining restaurants, a specific vocabulary seems to exist that reviewers make use of which leads to greater word clustering.

Category	Restaurant type		
	Cheap eats	Mid range	Fine dining
(1) Quality of food and beverage	Top quality	-	-
(2) Relative importance of ... ... food	Amazing sandwiches	Excellent, absolutely delicious, amazing food	Delicious, amazing food
	Classic danish pastry	Mushroom soup	Tasting menu
	Chicken curry	Lamb chops	Cheese board
	Curry soup	Chicken burger	
	Nice food	Chicken curry	
	Fast food	Salmon tartar	
... versus beverages		Pork tenderloin	
	Berries smoothie	Wine list	Wine pairing
	Fresh juices	Mango lassi	Wine list
	Complimentary gingershots		Wine menu
(3) Labour	Helpful and friendly staff	Friendly staff	Excellent service
	Friendly guys	Excellent and friendly service	Head chef
	Nice people		
(4) Ambience and atmosphere	Tiny cafe	Nice atmosphere	Fine dining experience
		Dining experience	
(5) Price and value	Loved prices	-	-
(6) Other amenities	Hand sanitizer	-	Michelin star(s)
	Phone chargers		World class

Table 2: Word classifications

Table 2 classifies the bigrams into the six categories mentioned in section 3. Quality and price are only mentioned for cheap eats. Additionally apparent for cheap eats are the occurrences of food unrelated amenities such as hand sanitiser or phone chargers. Decisive for an excellent review of cheap eats are therefore price/quality considerations and additional amenities. However, Table 2 suggests that food is the central element in excellent mid range restaurant reviews. Mango lassi, pork tenderloins, salmon tartar, chicken curry and chicken burger are additionally dominant bigrams that were not already mentioned in subsection 4.2. Also, food is labeled as absolutely delicious, amazing and excellent. For fine dining on the other hand, the decisive element for an excellent review is the interplay between food and

10. This is the summation of the bigram appearances in parentheses from the previous subsection 4.2. The legends of the graphs (n) depict how many times the bigrams appear.

wine. Wine is as equally important as food (if not to say that it is even more important). Food is summarised in the tasting menu and there are no specific dishes mentioned as it is the case for cheap eats and mid range restaurants.

Apparent from Table 2 is also the evolution of the wording used for labour. While labour is termed as helpful and friendly staff, friendly guys or nice people in cheap eats, it is labelled as friendly staff or excellent and friendly service in mid range restaurants and it is exclusively characterised as excellent service or becomes even personal by mentioning the head chef in fine dining restaurants.

## 6 Conclusion

This short research paper has looked into restaurant reviews from different price segments. Review format and content differ across restaurant price categories. Amongst the best-in-class restaurants in Copenhagen, cheap eats receive the highest numerical rating (4.85), followed by mid range (4.82) and fine dining (4.77) restaurants. However, when verbally reviewing the restaurants with a review text, restaurant visitors in fine dining establishments describe their experience with significantly more words (44) than they do in mid range restaurants (28) or cheap eats (25). At the extremes, that is for terrible and excellent restaurant experiences, the reviews for fine dining restaurants are more negative compared to mid range restaurants or cheap eats. This suggests that reviewers "show off" their culinary capital by providing extensive and more critical reviews for the most prestigious establishment to similarly interested peers.

Central element in excellent fine dining reviews is the interplay between food and wine. Wine is as important as food, if not to say that it is even more important. The dominant element in excellent mid range restaurant experiences is food. Quality/price considerations characterise together with additional amenities excellent reviews for cheap eats.

It needs to be noted that this study considered due to time limitations only fifteen best-in-class restaurants in Copenhagen. Future research should therefore increase the sample size by including more restaurants in their respective price category and geographically extend the analysis to gastronomic strongholds from different cultural settings.