9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

Ethnic succession in US restaurant review framing

Keywords: NLP; text analysis; cultural analysis; framing; food

Extended Abstract

Introduction. Implicit attitudes toward ethnic groups are nuanced and challenging to measure as people may censor their discourse. The way people evaluate food, however, can be revealing, since people may be less self-conscious when sharing opinions of different cultures' cuisines. Work in sociology has posited a theory of "ethnic succession of taste," with attitudes toward ethnic groups and cuisines being linked to migration patterns: as European immigration to the US slowed in the 20th century and immigrants assimilated up the socioeconomic ladder, so too did the prestige of European food; conversely, Asian and Hispanic food remain lower status, since ongoing Asian and Hispanic immigration continues to populate low-prestige jobs [6, 7].

We computationally evaluate the theory of ethnic succession of taste with a large-scale study of framing differences between US and European vs. Asian and Latin American cuisine in Yelp reviews. Our work extends theoretical work on the sociology of taste [1, 7] and earlier work on framing in review language [3, 9]. We extract dominant per-region frames from 5.2M reviews and show that Asian and Latin American cuisine is framed more negatively and stereotypically while controlling for customer satisfaction and restaurant price level. For example, Asian restaurants tend to be evaluated for hygiene, and Latin American waiters for their language skills. Our results provide the first empirical corroboration of the ethnic succession of taste and reveal linguistic processes by which its attendant attitudes are reified.

Methods. We use the latest release of Yelp's open dataset (`www.yelp.com/dataset`, retrieved January 2023) containing 5.2M reviews of 64K US-based businesses (after subsetting to food-related establishments). We use businesses' self-declared category tags to obtain cuisine labels for each business (e.g., *Italian*, *Mexican*). We then map country-level cuisine labels to broader geographic regions (e.g., *Europe*, *Latin America*). Since country distribution within regions is non-uniform, we replicate analyses with the most frequent cuisine category per region removed.

To study framing along different restaurant attributes, we analyze the adjectives that modify food tokens (e.g., *chicken, noodles*), waitstaff (e.g., *waiter, server*), and general restaurant-related tokens (e.g., *place, spot*). By targeting adjectives anchored to noun sets of interest, we obtain more controlled and interpretable results compared to aggregate measures of sentiment over the entire text of reviews. We compile anchor sets with WordNet [4] and augment our set of food anchors with a dataset of menu items [2]. We parse all review texts with Stanza and leverage dependency parse relations to retrieve adjectives modifying each anchor set. We then identify dominant frames by extracting the adjectives most strongly associated with each cuisine region using the Fightin' Words method [5]. In addition, we obtain an overall measure of framing valence by computing the proportion of adjectives for each cuisine region belonging to the positive emotion lexicon from Linguistic Inquiry and Word Count (LIWC) [8].

Results. Controlling for rating (number of stars given by the reviewer) and price level (number of dollar signs labeled by Yelp), we find that framing in reviews differs strongly for European and American cuisines versus Asian and Latin American cuisines (Table 1). We observe the starkest differences in the descriptions of restaurant waitstaff: waiters at American restaurants are framed as outgoing (*fun, personable*), whereas waiters at Asian restaurants are framed as shy (*courteous, quiet*). American and European waitstaff also tend to be described with more generic positive terms (*fantastic, wonderful*), while non-Western waitstaff are praised for their efficiency (*quick*) or language proficiency (*bilingual, fluent*).

Similarly, whereas American and European food tends to be praised through the use of strong positive adjectives like *amazing, excellent* and *charming*, Asian and Latin American food of the same star and price tier are evaluated more in terms of literal tastes and smells (*flavorful, fresh, bland*). We confirm this result with a lexicon of taste terms (`www.chefsbest.com/taste-terminology/`, retrieved February 2023), finding small but significant effect sizes. Finally, while US and European restaurants tend to be described in more general positive terms (*great, perfect, lovely*), Asian and Latin American restaurants are evaluated along more specific dimensions, such as hygiene (*clean*) or authenticity (*authentic, americanized*). These patterns hold even with the most frequent cuisines per region removed (Table 2).

Crucially, the rubric for US and European cuisines tends to be one of unmarked and generic "goodness," independent of any particular frame, whereas Asian and Latin American cuisines are more restricted in how they are perceived and appraised. The greater limitation on what Asian and Latin American food, restaurants, and waitstaff "get to be" in consumers' imaginations is reminiscent of social stereotyping more broadly: default groups are described with unmarked language, whereas non-default groups are associated with specific stereotypes. Perhaps particularly revealing is the framing of Asian and Latin American cuisines as *typical*, a word which implies the existence of an archetypal X and evokes a singular image of what a *typical* X may be. Finally, we find that US and European restaurants are framed with more positive valence (see Figures 1 & 2) across food, waitstaff, and general restaurant quality, followed by Latin American restaurants and finally, Asian restaurants.

Conclusion. Overall, we found that US and European restaurants are framed as having more cultural capital than Asian and Latin American restaurants, which are framed more negatively and in marked terms related to stereotypical traits. Our results corroborate the theory of ethnic succession of taste, showing how negative, stereotyping attitudes toward less assimilated immigrant groups are reflected in attitudes toward those groups' cuisines.

References

[1] P. Bourdieu. *Distinction: A social critique of the judgement of taste*. Harvard University Press, 1987.

[2] D. Jurafsky, V. Chahuneau, B. Routledge, and N. Smith. Linguistic markers of status in food culture: Bourdieu's distinction in a menu corpus. *Journal of Cultural Analytics*, 1(1), 2016.

[3] D. Jurafsky, V. Chahuneau, B. R. Routledge, and N. A. Smith. Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*, 2014.

[4] G. A. Miller. Wordnet: a lexical database for English. *Communications of the ACM*, 38(11):39–41, 1995.

[5] B. L. Monroe, M. P. Colaresi, and K. M. Quinn. Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4):372–403, 2008.

[6] K. Ray. Ethnic succession and the new American restaurant cuisine. *The restaurants book: ethnographies of where we eat*, 97:113, 2007.

[7] K. Ray. Bringing the immigrant back into the sociology of taste. *Appetite*, 119:41–47, 2017.

[8] Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of lang. & social psych.*, 29(1):24–54, 2010.

[9] S. Zukin, S. Lindeman, and L. Hurson. The omnivore's neighborhood? Online restaurant reviews, race, and gentrification. *Journal of Consumer Culture*, 17(3):459–479, 2017.

US	LATIN AMERICA	EUROPE	ASIA
		Food modifiers	
great* (19.4)	authentic* (24.8)	homemade* (10.4)	spicy* (42.6)
local* (13.9)	complimentary* (10.1)	wonderful* (10.3)	sour* (38.5)
daily* (10.3)	skinny* (7.3)	romantic* (8.2)	authentic* (26.4)
good* (9.7)	strong* (6.2)	excellent* (7.9)	fresh* (18.3)
amazing* (9.5)	bland* (5.8)	personal* (7.6)	flavorful* (15.5)
juicy* (9.3)	complementary* (5.1)	warm* (7.3)	vegetarian* (14.6)
cold* (9.2)	free* (4.8)	large* (6.9)	americanized* (13.7)
creative* (9.1)	mild* (4.8)	delicious* (6.4)	sweet* (12.3)
seasonal* (8.9)	chunky* (4.6)	authentic* (6.1)	traditional* (10.0)
crispy* (8.6)	flavorful* (4.6)	crusty* (6.0)	hot* (9.6)
cheesy* (8.5)	tropical* (4.2)	outstanding* (6.0)	oily* (8.9)
eclectic* (8.2)	warm* (3.7)	creamy* (5.7)	crunchy* (8.4)
		Waitstaff modifiers	
good* (9.3)	bilingual* (3.4)	knowledgeable* (3.9)	polite* (6.6)
great* (8.6)	authentic* (2.1)	attentive* (3.5)	friendly* (6.1)
amazing* (8.5)	tender (1.7)	professional* (2.5)	attentive* (5.1)
awesome* (6.5)	awesome (1.5)	charming* (2.0)	kind* (4.9)
fantastic* (5.3)	millenial (1.4)	personable (1.9)	nice* (4.9)
attentive* (4.5)	stingy (1.3)	warm (1.9)	courteous* (3.3)
excellent* (4.4)	slow (1.3)	obliging (1.8)	quiet* (2.7)
fun* (3.4)	silly (1.2)	oblivious (1.8)	quick* (2.4)
personable* (3.3)	super-attentive (1.2)	wonderful (1.8)	fluent* (2.4)
		Establishment modifiers	
great* (24.3)	authentic* (9.7)	cozy* (7.6)	clean* (15.4)
cool* (15.2)	colorful* (5.7)	romantic* (7.0)	authentic* (12.5)
fun* (12.2)	pure* (5.2)	quaint* (5.7)	ethnic* (6.1)
nice* (11.0)	typical* (4.3)	little* (4.5)	typical* (5.9)
beautiful* (10.0)	historic* (3.7)	charming* (4.4)	favorite* (5.1)
perfect* (8.6)	old* (3.4)	small* (4.3)	modern* (4.9)
trendy* (7.7)	favorite* (3.4)	wonderful* (4.0)	americanized* (4.8)
busy* (6.4)	festive (3.0)	loud* (3.8)	traditional* (4.5)
upscale* (6.4)	traditional (3.0)	lovely* (3.5)	vegetarian* (4.0)

Table 1: Adjectives most associated with the food, waitstaff, and establishments of different cuisine regions among 4-star reviews of restaurants with a $$ price level, with trivially associated ethnic and geographic modifiers (e.g., *italian, hispanic*) removed using a lexical filter compiled from WordNet. Strength of association is computed as a z-score of the log odds ratio between adjectives of a given cuisine region vs. adjectives of all other cuisine regions. Significant values (z-score > 1.96) are marked with an asterisk.

US	LATIN AMERICA	EUROPE	ASIA
		Food modifiers	
amazing* (11.4)	authentic* (6.2)	warm* (7.8)	spicy* (38.2)
creative* (11.0)	national* (3.3)	authentic* (5.4)	fresh* (20.5)
local* (10.5)	fiery* (2.4)	wonderful* (4.8)	authentic* (19.6)
great* (10.0)	flavorful* (2.2)	savory* (4.2)	sour* (15.9)
seasonal* (8.5)	tender* (2.0)	romantic* (3.9)	vegetarian* (14.7)
eclectic* (8.5)	tropical (1.9)	traditional* (3.6)	flavorful* (13.8)
interesting* (8.3)	original (1.8)	tender* (3.4)	specialty* (9.9)
crispy* (7.7)	tough (1.7)	lovely* (3.3)	crunchy* (9.4)
fun* (7.5)	famous (1.6)	artful* (3.1)	special* (8.1)
unique* (7.5)	cilantro-infused (1.6)	crusty* (3.0)	non-sushi* (8.0)
		Waitstaff modifiers	
good* (9.4)	bilingual* (2.3)	knowledgeable* (3.2)	polite* (6.0)
amazing* (8.9)	super-casual (1.1)	inept (1.6)	attentive* (5.1)
great* (7.0)	exasperated (1.1)	appalling (1.5)	friendly* (4.9)
awesome* (6.1)	gaunt (1.1)	able (1.4)	kind* (4.7)
attentive* (6.0)	slimey (1.1)	snide (1.4)	nice* (4.2)
fantastic* (4.9)	uninformative (1.1)	charming (1.4)	courteous* (3.3)
cool* (3.8)	unconvincing (1.1)	pleasant (1.4)	quiet* (3.0)
excellent* (3.8)	wooden (1.1)	attentive (1.3)	quick* (2.4)
knowledgeable* (3.5)	overdressed (1.1)	gracious (1.3)	fluent* (2.3)
phenomenal* (3.2)	morose (1.1)	lovely (1.2)	timid* (2.1)
		Establishment modifiers	
great* (14.2)	historic* (4.7)	historic* (5.0)	clean* (12.2)
cool* (11.1)	old* (4.1)	beautiful* (4.1)	authentic* (9.2)
beautiful* (9.7)	beautiful* (2.5)	old* (3.7)	ethnic* (6.7)
trendy* (8.3)	vegan* (2.3)	cozy* (3.5)	small* (4.6)
nice* (8.2)	authentic (1.9)	romantic* (3.2)	favorite* (4.4)
fun* (8.0)	huge (1.9)	lovely* (2.8)	traditional* (4.3)
touristy* (7.2)	large (1.4)	charming* (2.5)	typical* (4.1)
upscale* (7.0)	elusive (1.3)	wonderful* (2.2)	modern* (3.6)
amazing* (6.9)	iconic (1.3)	gorgeous* (2.2)	americanized* (3.6)
busy* (6.4)	themed (1.2)	quaint* (2.0)	oriental* (3.5)

Table 2: Adjectives most associated with the food, waitstaff, and establishments of different cuisine regions with the most frequent cuisine removed ("American (traditional)"; "Mexican", "Italian", "Chinese"). Adjectives are among 4-star reviews of restaurants with a $$ price level, with trivially associated ethnic and geographic modifiers (e.g., *italian, hispanic*) removed using a lexical filter compiled from WordNet. Strength of association is computed as a z-score of the log odds ratio between adjectives of a given cuisine region vs. adjectives of all other cuisine regions. Significant values (z-score > 1.96) are marked with an asterisk.

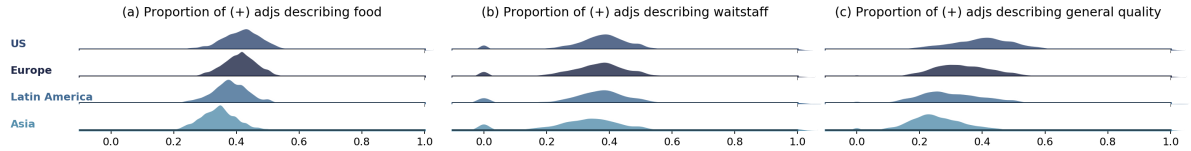

Figure 1: Proportion of LIWC positive emotion adjectives across different cuisine regions, showing that US and European restaurants tend to be described more positively than their Latin American and Asian counterparts across (**a**) the food served, (**b**) employees servicing, and (**c**) overall restaurant quality. US restaurants are framed the most positively, followed closely by European restaurants (β_{food}=-4.2e-3, p=0.01, $\beta_{\text{waitstaff}}$=-0.01, p=0.001, β_{general}=-0.05, p < 1e-68), then Latin American restaurants (β_{food}=-0.04, p < 1e-93, $\beta_{\text{waitstaff}}$=-0.02, p < 1e-4, β_{general}=-0.09, p < 1e-140), then Asian restaurants (β_{food}=-0.07, p=0.0, $\beta_{\text{waitstaff}}$=-0.04, p < 1e-25, β_{general}=-0.13, p=0.0). Coefficients and significance were obtained by fitting a linear regression model to predict framing valence from cuisine region while controlling for price level and stars awarded.

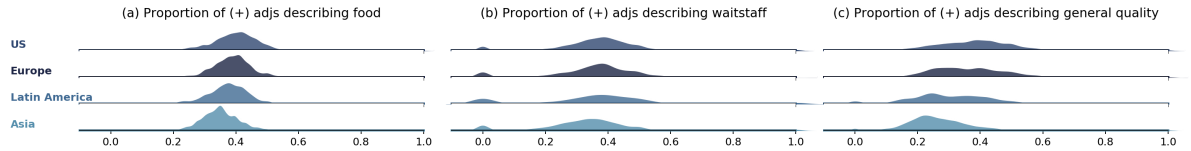

Figure 2: Proportion of LIWC positive emotion adjectives across different cuisine regions, with the most frequent cuisine per region (“American (traditional)”; “Mexican”, “Italian”, “Chinese”) removed for robustness. US and European restaurants tend to be described more positively than their Latin American and Asian counterparts across (**a**) the food served, (**b**) employees servicing, and (**c**) overall restaurant quality. US and European restaurants are framed equally positively (β_{food}=-4.5e-3, p=0.04, $\beta_{\text{waitstaff}}$=-0.01, p=0.06, β_{general}=-0.01, p=0.07); Latin American restaurants are framed less positively along food and overall quality but not significantly worse for waitstaff (β_{food}=-0.02, p < 1e-6, $\beta_{\text{waitstaff}}$=-0.01, p=0.14, β_{general}=-0.04, p < 1e-15), and Asian restaurants are framed significantly worse along all three attributes (β_{food}=-0.05, p < 1e-129, $\beta_{\text{waitstaff}}$=-0.03, p < 1e-9, β_{general}=-0.09, p < 1e-147). Coefficients and significance were obtained by fitting a linear regression model to predict framing valence from cuisine region while controlling for price level and stars awarded.