



Think 53: Food Talks

Dan Jurafsky & Yoshiko Matsumoto

Sex, Drugs, and Sushi Rolls

Tuesday, May 23, 2017



What can the language of food reviews tell us

About the author

- Gender
- Attitudes toward food, guilt
- Psychological state

About our culture more generally

About our language

- How language changes over time

Hidden in the words of Yelp reviews

Dan Jurafsky, Victor Chahuneau, Bryan R. Routledge, and Noah A. Smith. 2014. Narrative framing of consumer sentiment in online restaurant reviews. First Monday 19:4

- Our idea: look at the language of reviews online
- Correlate review words with facts about restaurants or reviewers



Let's talk methodology!

Where do I find data?

How do I choose a sample?

How do I measure that sample?

How do I draw some qualitative conclusions from the quantitative measurements?

Two Yelp datasets

1) Yelp Challenge Dataset

https://www.yelp.com/dataset_challenge

The Challenge Dataset:

- **2.2M** reviews and **591K** tips by **552K** users for **77K** businesses
- **566K** business attributes, e.g., hours, parking availability, ambience.
- Social network of **552K** users for a total of **3.5M** social edges.
- Aggregated check-ins over time for each of the **77K** businesses
- **200,000** pictures from the included businesses

[Get the Data](#)

Cities:

- U.K.: Edinburgh
- Germany: Karlsruhe
- Canada: Montreal and Waterloo
- U.S.: Pittsburgh, Charlotte, Urbana-Champaign, Phoenix, Las Vegas, Madison

2) All reviews from 7 cities (NY, LA, Chicago, DC, Boston, SF, Phil)

What one of the 900,000 reviews looks like

```
{"votes": {"funny": 0, "useful": 5, "cool": 2}, "user_id":  
"rLtl8ZkDX5vH5nAx9C3q5Q", "review_id": "fWKvX83p0-  
ka4JS3dc6E5A", "stars": 5, "date": "2011-01-26", "text": "My wife took  
me here on my birthday for breakfast and it was excellent. The  
weather was perfect which made sitting outside overlooking their  
grounds an absolute pleasure. Our waitress was excellent and our  
food arrived quickly on the semi-busy Saturday morning. It looked like  
the place fills up pretty quickly so the earlier you get here the  
better.\n\nDo yourself a favor and get their Bloody Mary. It was  
phenomenal and simply the best I've ever had. I'm pretty sure they  
only use ingredients from their garden and blend them fresh when  
you order it. It was amazing.\n\nWhile EVERYTHING on the menu  
looks excellent, I had the white truffle scrambled eggs vegetable skillet  
and it was tasty and delicious. It came with 2 pieces of their griddled  
bread with was amazing and it absolutely made the meal complete. It  
was the best \"toast\" I've ever had.\n\nAnyway, I can't wait to go  
back!", "type": "review", "business_id": "9yKzy9PApeiPPOUJEtnvkg"}
```

How to study word usage

Ask “does a **set of words W** occur more often in texts of **type A** or texts of **type B**?”

What does **type A versus B** mean?

- Expensive restaurants versus cheap restaurants
- Female reviewers versus male reviewers
- High-rated restaurants versus low-rated restaurants

Sampling

Expensive restaurants versus cheap restaurants

- Choose everything
 - "All the \$\$\$\$ restaurants versus all the \$ restaurants"
 - "All the \$\$\$ and \$\$\$\$ restaurants versus all the \$ and \$\$ restaurants"
- Choose a sample
 - Randomly select 50 \$\$\$\$ restaurants and 50 \$ restaurants
 - 50 times, choose a random number and pick the nth restaurant in a list without replacement

How to study word usage

Ask “does a **set of words W** occur more often in texts of **type A** or texts of **type B**?”

What does **set of words W** mean?

- We could ask about all words
- We could ask about words in a particular list, chosen because of a hypothesis
 - “words about health”

Word usage in text A versus B

Count Ratio: does “luxury” occur more in text A or text B?

Ratio of the counts in the two texts

$$\frac{\textit{countA}(\textit{“luxury”})}{\textit{countB}(\textit{“luxury”})}$$

Count Ratio = 3

Insufficient! Why?

Maybe Rev 2+4 are longer!

	“luxury” in A = expensive	“luxury” in B = cheap
Rev 1		3
Rev 2	9	
Rev 3		2
Rev 4	6	

Word usage in text A versus B

Likelihood Ratio: Ratio of the probabilities (percentages, normalized counts) in the two texts

$$\frac{\left(\frac{\text{count}_A(\text{"luxury"})}{\sum_{\text{word in } A} \text{count}(\text{word})} \right)}{\left(\frac{\text{count}_B(\text{"luxury"})}{\sum_{\text{word in } B} \text{count}(\text{word})} \right)}$$

Word usage in text A versus B

Likelihood Ratio: Ratio of the probabilities (percentages, normalized counts) in the two texts

	"luxury" in A = expensive	All words in A = expensive	"luxury" in B = cheap	All words in B = cheap
Rev 1			3	100
Rev 2	9	200		
Rev 3			2	200
Rev 4	6	400		
Totals	15	600	5	300

Word usage in text A versus B

Likelihood Ratio: Ratio of the probabilities (percentages, normalized counts) in the two texts

$$\frac{\left(\frac{\text{count}A(\text{"luxury"})}{\sum_{\text{word in } A} \text{count}(\text{word})} \right)}{\left(\frac{\text{count}B(\text{"luxury"})}{\sum_{\text{word in } B} \text{count}(\text{word})} \right)} = \frac{15/600}{5/300} = 1.5$$

Count ratio was 3

Likelihood ratio is 1.5, more accurate

In practice we use *log(likelihood ratio)*



What words are associated
with 1-star reviews?

Let's proceed to count!

A sample 1-star review (shortened with ... by me)

The bartender... absolutely horrible... we waited 10 min before we even got her attention... and then we had to wait 45 - FORTY FIVE! - minutes for our entrees... stalk the waitress to get the cheque... she didn't make eye contact or even break her stride to wait for a response ...

What words are most associated with * (versus *****) reviews

worst -47.498

was -44.069

not -42.365

bad -39.836

terrible -39.538

rude -39.493

minutes -39.287

horrible -38.584

asked -38.109

she -37.379

manager -34.547

no -34.391

awful -32.796

told -32.495

bland -30.655

we -29.995

he -29.901

said -28.123

waitress -27.902

didn't -27.822

were -27.745

us -27.624

disgusting -27.291

our -26.905

mediocre -26.484

ok -25.980

overpriced -25.187

to -24.850

worse -24.479

her -24.426

...

...

Top 50 words associated with one-* reviews

Linguistic Class	Words in Class
Negative sentiment	worst, rude, terrible, horrible, bad, awful, disgusting, bland, tasteless, gross, mediocre, overpriced, worse, poor
Linguistic negation	no, not
1 pl pronouns	we, us, our
3rd pronouns	she, he, her, him
Past tense verbs	was, were, asked, told, said, did, charged, waited, left, took
Narrative sequencers	after, then
Common nouns	manager, waitress, waiter, customer, customers, attitude, waste, poisoning, money, bill, minutes
Irrealis modals	would, should
Complementizers	to, that

Language of bad reviews?

Negative sentiment language

horrible awful terrible bad disgusting

Past tense narratives about people

waited, didn't, was

he, she, his, her,

manager, customer, waitress, waiter

Frequent mentions of **we** and **us**

... **we** were ignored until **we** flagged down one waiter to go get **our** waitress ...

Other narratives with this language

A genre using:

Past tense, we/us, negative, people narratives

Texts written by **people suffering trauma**

- Chat group discussions after Princess Diana's death
 - Stone, L.D. & Pennebaker, J.W. (2002). Trauma in real time: Talking and avoiding online conversations about the death of Princess Diana. *Basic and Applied Social Psychology*, 24, 172-182
- Blog posts after September 11, 2001
 - Cohn, M.A., Mehl, M.R., & Pennebaker, J.W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, 15, 687-693
- Student newspaper reports after a campus tragedy
 - Gortner, E.-M., & Pennebaker, J.W. (2003). The archival anatomy of a disaster: Media coverage and community-wide health effects of the Texas A&M Bonfire Tragedy. *Journal of Social and Clinical Psychology*, 22, 580-603

Why? Pennebaker's *social stage model of coping*

- people feel a need to tell stories expressing their negative emotion,
- “we/us/our” suggests we are seeking comfort in community
- Past tense used to distance ourselves from the traumatic event

Implications

1-star reviews are not descriptions of bad food.

They are trauma narratives!

The lesson of bad reviews:

We are very sensitive to personal interaction



What about positive reviews?

Wine reviewing changes beginning in 1980s

Linguist Adrienne Lehrer analysis:

- fleshy, muscular, sinewy, big-boned, or broad-shouldered.

Robert Parker reviews:

- sexy, sensual, supple and seductive, liquid Viagra,
- “offering voluptuously textured, hedonistic drinking”

Sean Shesgreen

- talk of wines as “pretty and caressing,” “ravishing,” “pillowy,” and “overendowed” means “gastronomy has eclipsed sex.”

Positive (*****) reviews

*the apple tart ice cream pastry caramely
thing was just **orgasmic***

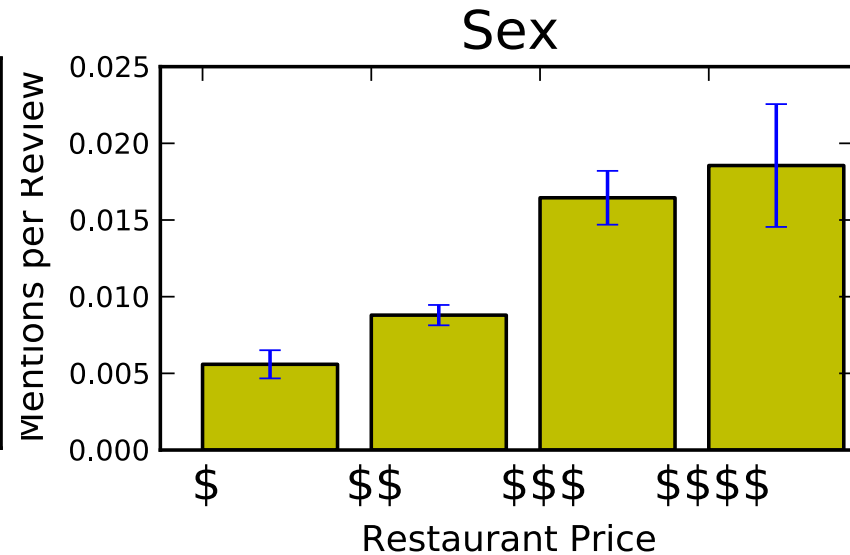
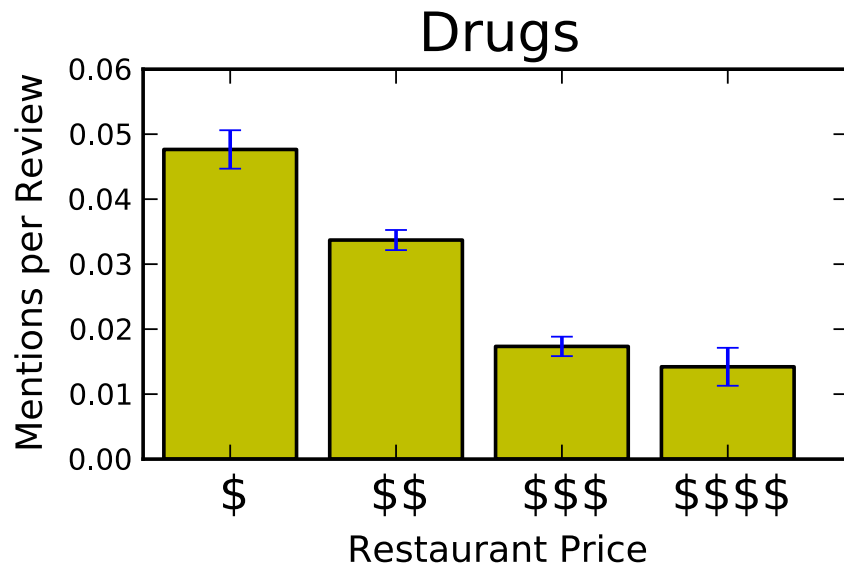
*sumptuous flavors, jaw-droppingly good,
sexy food*

*succulent pork belly paired with
seductively seared foie gras*

Positive (*****) Reviews?

addicted to wings
the fries are like crack
....crave... cupcakes

orgasmic pastry
seductively seared...
very naughty pork belly



Drug metaphors

*garlic noodles . . . are now my **drug of choice***

*these cupcakes are like **crack***

*be warned the wings are **addicting***

*. . . every time **I need a fix**. That fried chicken is so damn good!*

*I swear the fries have **crack** or some sort of **addicting drugs** in them*

Why the addiction narrative?

Table 4: Foods most likely to be described using drug metaphors.				
Meaty, fatty foods	Starchy comfort food	Sweet food	Small ethnic dishes	Descriptors
burgers	pizza	sweets	sushi	comfort
barbecue	mac and cheese	pancakes, breakfast	dim sum	fried, greasy
chicken wings	pasta/noodles	sugar	tacos, burritos	unhealthy
french fries	soups	chocolate	spam musubi	hearty, satisfying
	sandwiches	beignets	dumplings	junk
			falafel	authentic
			tapas	cheap

Craved foods aren't vegetables, or main courses like meatloaf or fish or even side dishes like mashed potatoes

They are **junk foods** or at least non-normative foods

- Assuage the guilt
- It's not my fault, I had no control, the cupcake made me eat it

Gender and the addiction metaphor?

Are women more likely to use the metaphor of drugs than men?

- women are more pressured to conform to healthy eating
- women more likely to mention food cravings
 - Rozin, *et al.* (1991): females are significantly more likely to express cravings for chocolate than males.
 - Zellner, *et al.* (1999), Weingarten and Elston (1990), Osman and Sobal (2006): female undergraduates more likely than males to report food cravings.
- How to test:
 - Figure out the gender of Yelp users, see if women use the drug metaphor more.

How to figure out the gender of Yelp users

Yelp dataset gives us the first names of the reviewer

```
{"votes": {"funny": 0, "useful": 7, "cool": 0}, "user_id":  
"CR2y7yEm4X035ZMzrTtN9Q", "name": "Jim", "average_stars": 5.0,  
"review_count": 6, "type": "user"}
```

```
{"votes": {"funny": 0, "useful": 1, "cool": 0}, "user_id":  
"_9GxoHhdx30ujPaQwh6Ew", "name": "Kelle", "average_stars": 1.0,  
"review_count": 2, "type": "user"}
```

```
{"votes": {"funny": 0, "useful": 1, "cool": 0}, "user_id": "8mM-  
nqxjg6pT04kwcjMbsw", "name": "Stephanie", "average_stars": 5.0,  
"review_count": 2, "type": "user"}
```

```
{"votes": {"funny": 0, "useful": 2, "cool": 0}, "user_id": "Ch6CdTR2IVaVANr-RglMOg",  
"name": "T", "average_stars": 5.0, "review_count": 2, "type": "user"}
```

```
{"votes": {"funny": 0, "useful": 0, "cool": 0}, "user_id": "NZrLmHRyiHmyT1JrfzkCOA",  
"name": "Beth", "average_stars": 1.0, "review_count": 1, "type": "user"}
```

```
{"votes": {"funny": 30, "useful": 45, "cool": 36}, "user_id": "mWx5Sxt_dx-  
sYBZg6RgJHQ", "name": "Amy", "average_stars": 3.79, "review_count": 19, "type":  
"user"}
```

Use Social Security Baby Name Database

<https://www.ssa.gov/OACT/babynames/limits.html>

Lisa,F,46081
Mary,F,43493
Susan,F,35750
Karen,F,35184
Linda,F,31462
Patricia,F,26538
Donna,F,25727
Cynthia,F,24021
Deborah,F,22893
Sandra,F,22415
Lori,F,22388
Brenda,F,21275
Pamela,F,21274
Barbara,F,20648
Debra,F,20600
Kimberly,F,20365
Sharon,F,19630
Teresa,F,18512
Nancy,F,18030

Michael,M,85049
David,M,81332
John,M,78450
James,M,72571
Robert,M,70212
Mark,M,53520
William,M,44862
Richard,M,39420
Thomas,M,36531
Jeffrey,M,33525
Scott,M,30708
Steven,M,30704
Joseph,M,30087
Kevin,M,28356
Charles,M,27594
Timothy,M,27089
Daniel,M,25554
Kenneth,M,25244
Paul,M,23807
Brian,M,23245
...

Susan,M,86

$P(\text{Susan born 1962 is listed female}) = \frac{35750}{35750+86} = .9976$

Some Caveats:

- Doesn't account for transgender or non-binary gender
- We don't know exactly when Yelp people are born
 - Names shift slowly from male to female
 - 1920-1950 Stacy was male name
- So we pick some confidence threshold (95% or etc.) and only use names above the threshold

Are women more likely to use the metaphor of drugs than men?

- Label a Yelp reviewer as female if name probability of people born after 1951 was female with probability $> .80$
- That assigns a most-probable-gender to 90% of Yelp users
 - Could choose a higher threshold than .80 to get more accurate conclusions
 - Still has errors due to assuming binary model of gender
- We then used linear regression to predict the number of mentions of addiction from the gender of the speaker.
- Women significantly more likely than men to talk about food as a drug ($p=0.000832$).

Are women more likely to use the metaphor of drugs than men?

- Women significantly more likely than men to talk about food as a drug ($p=0.000832$).
- But we don't know the cause:
 - women might be more likely than men to have these cravings
 - women might be more comfortable than men in admitting to these cravings
 - women might have identical desires but be more likely than men to use this particular linguistic metaphor

What about sex in reviews?

sex on the beach roll

foreplay roll

sweet temptation roll

orgasmic spicy tuna roll

sexy mama roll

sexy lady roll

hot sexy shrimp roll

sexy lizzy roll

Sushi



Other use of sex

*molten chocolate cake . . . honestly an **orgasm on a plate***

*I still **lust** for the silky panna cotta and **tantalizing sorbet***

*marshmallows . . . so . . . sticky and sweet, they're nearly **pornographic***

*warm chestnut mochi chocolate cake . . . **seductively gooey on the inside***

Descriptive words associated with **dessert** in restaurant reviews

Sex:

orgasmic, sinful, tempting, tantalizing

Mouthfeel:

rich, moist, warm, sweet, dense, hot, creamy, flaky, light, fluffy, sticky, dry, gooey, smooth, crisp, sugary, buttery, oozing, satin, soft, velvety, thick, melty, silky, oozing, tangy, thin, crunchy, silky, spongy, nutty



Reminder from Strauss paper

When we talk about desserts, we talk about their feel in the mouth, not their appearance, smell, taste, or sound.

Americans usually describe desserts as soft or dripping wet.

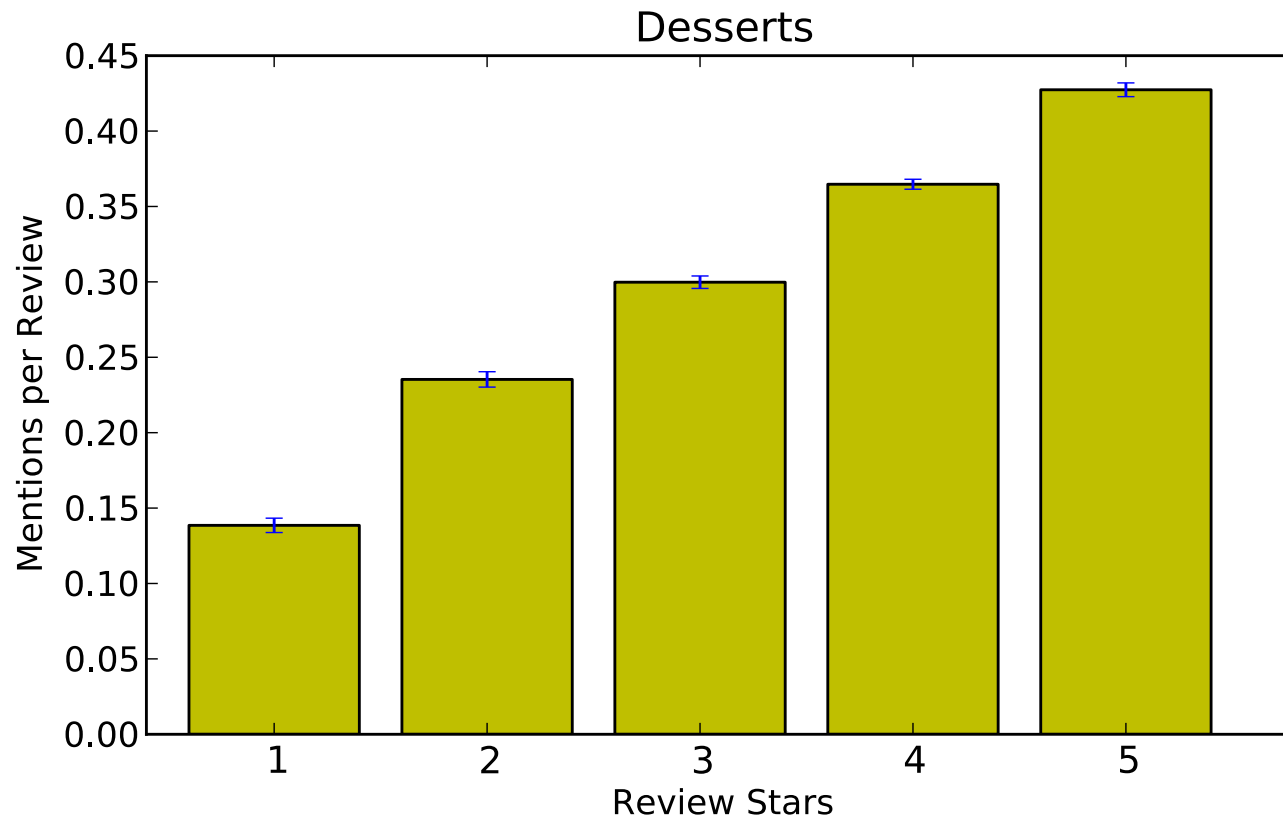
US commercials emphasize tender, gooey, rich, creamy food, and associate softness and dripping sweetness with sensual hedonism and pleasure.



Top sentiment words people use about dessert

delicious amazing yummy decadent
divine yum good OK wow fabulous
scrumptious delectable wonderful
delish refreshing awesome perfect
incredible fantastic heavenly

The more reviewers mention dessert...





Furthermore

Women (on Yelp) are more likely than men to mention desserts in their reviews.

The WOMEN AS DESSERT metaphor

Caitlin Hines

“From MTV tartlet to arthouse pastry du jour. No wonder the adoring critics who droolingly dub her 'luminous' are so eager to gobble Liv Tyler up. You can almost sniff a just-from-the-oven freshness about her as she strides into the room.”

- USA Today, 1996, cited by Hines

The WOMAN AS DESSERT Metaphor

Caitlin Hines. 1999. [Rebaking the Pie: The WOMAN AS DESSERT Metaphor](#). In *Reinventing Identities: The Gendered Self in Discourse*, Edited by Bucholtz, M. and Liang, AC and Sutton, L.A. Oxford University Press

The claim: there is an implicit metaphor in English in which women are represented as dessert

Hines data for WOMEN ARE DESSERT metaphor

- (piece of) cake
- Cheesecake
- Cookie
- Crumpet
- Cupcake
- (a tasty bit of) pastry
- cherry pie
- poundcake
- Pumpkin pie, pumpkin tart, punkin
- tart

The WOMEN AS DESSERT metaphor

Methodology: The words Hines allowed
as evidence

- Must have non-metaphorical sense (rules out *sweetiepie* and *honeybun*)
- Must appear in multiple sources

Collected in late 1990s by hand (before Google)

- Dictionary definitions
- A few newspaper examples

Her claim about what's in common in this metaphor

- Firm on the outside,
- Soft or juicy in the middle
- Able to be cut into more than one piece (cheesecake, cherry pie, poundcake) or conceptualized as one serving of a batch (cookie, crumpet, cupcake)
 - “...a piece of that”
- You don't see:
 - *custard
 - *ice cream cone
 - *mousse



Problems

Not much data by modern standards

Most of the usages seem archaic

The methodology seems a bit circular

Modern usages ("milkshake", etc) may not match her model

One follow-up study

Two Stanford frosh Debra Pacio and Linda Yu

Looked at dessert metaphors in pop music

- *Billboard Top 10 Singles Charts* from 1950 to present

They found

- The word “sweet” associated with pop and soul, declines from 1950 to present
- Mentions of candy and other desserts as a metaphor for sex in hiphop increasing in recent years
 - Kelis’s “Milkshake”
 - Li’l Wayne’s “Lollipop”



Still: Katy Perry's Bon Appetit

Sweet potato pie

Sweet tooth

Whipped cream

Melt in your mouth

Use some sugar

Fresh out of the oven

Hit that sweet tooth

World's best cherry pie

Katy
Perry
on Bon
Appetit

KATY PERRY

WORLD'S BEST
CHERRY PIE






In Summary

The Woman as Dessert metaphor

An intriguing hypothesis to test, modify or examine!



Another natural direction

Sex, Chocolate, and Gender



Office hours switch for today

My office hours today: 5-6:30pm

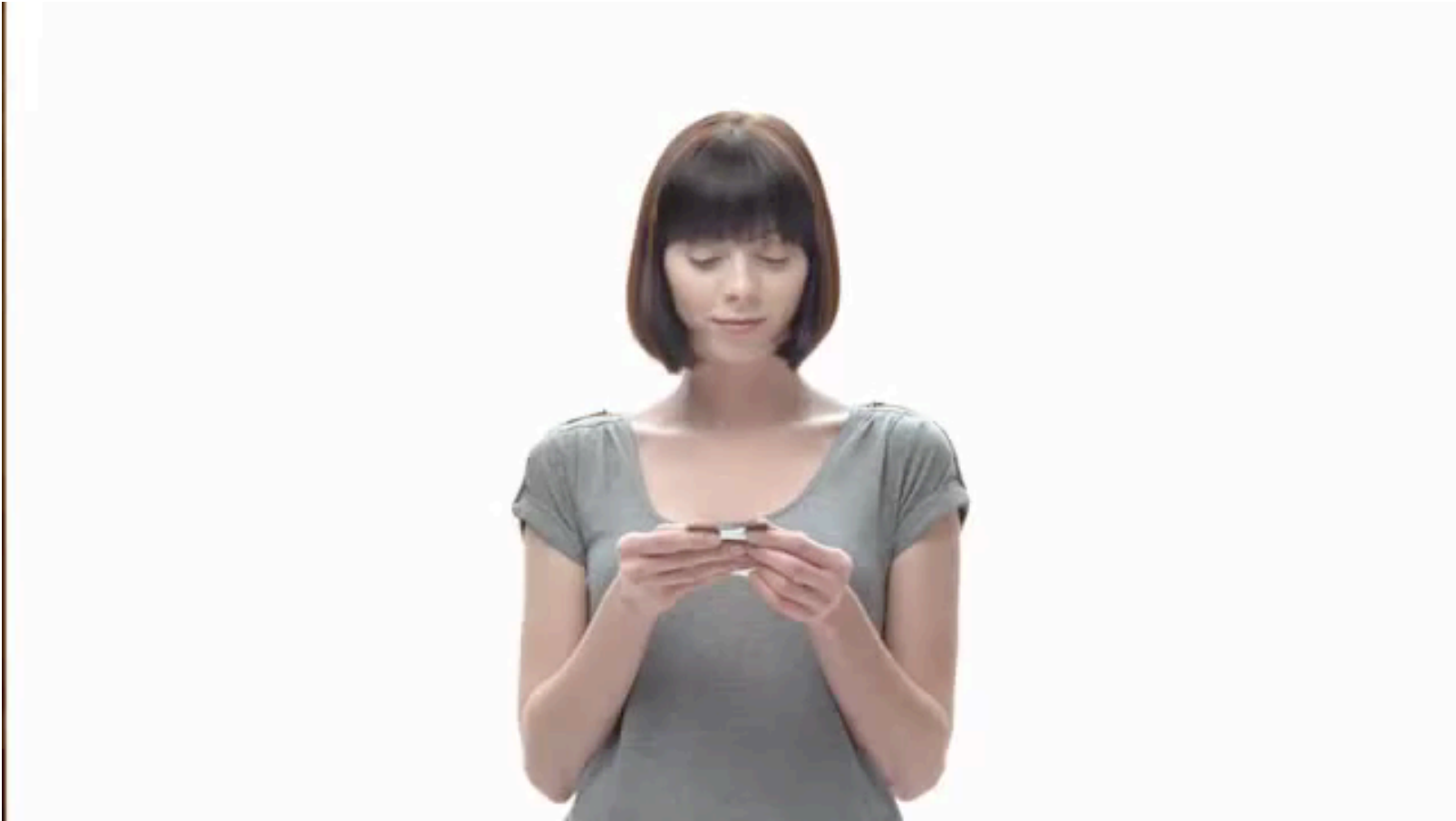
Ghirardelli chocolate ad

<https://www.youtube.com/watch?v=OnH9xOIIYwkQ>



YORK chocolate commerial

<https://www.youtube.com/watch?v=CzcxM6u7ey0>



M&M's commercial

<https://www.youtube.com/watch?v=JBvlvvKgVqs>



Conclusion

Review words tell us about human psychology

- Negative differentiation
- Trauma caused by bad personal interaction
- Attitudes toward eating and guilt

Language and its link with sex and food

- WOMEN ARE DESSERT metaphor
- Possible gendered function of craving
- Increase in sex metaphors for wine starting 1980

Methodology

- Word counting
- Metaphor finding