

Highways and words:

Migration patterns and linguistic diversity

Anjali Bhatt & Molly Lewis

June 30th, 2017



Summer Institute in
Computational Social Science

1. Motivation

2. Methods

3. Findings

4. Next Steps

Cultural diversity matters

1. Culture is a defining attribute of human behavior in communities
2. Diversity enables innovation, human understanding, and other forms of social and economic progress

Language is a window into culture

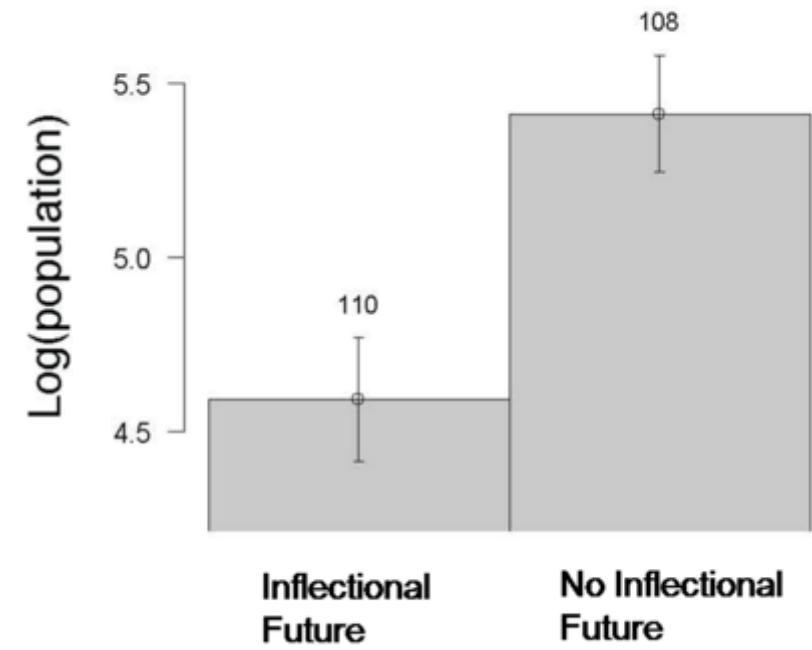
- Provides insight into how people construe meaning from their experiences
- Is a vehicle for how communities negotiate, codify, and uphold norms, values, and beliefs

What are the sources of linguistic variability?

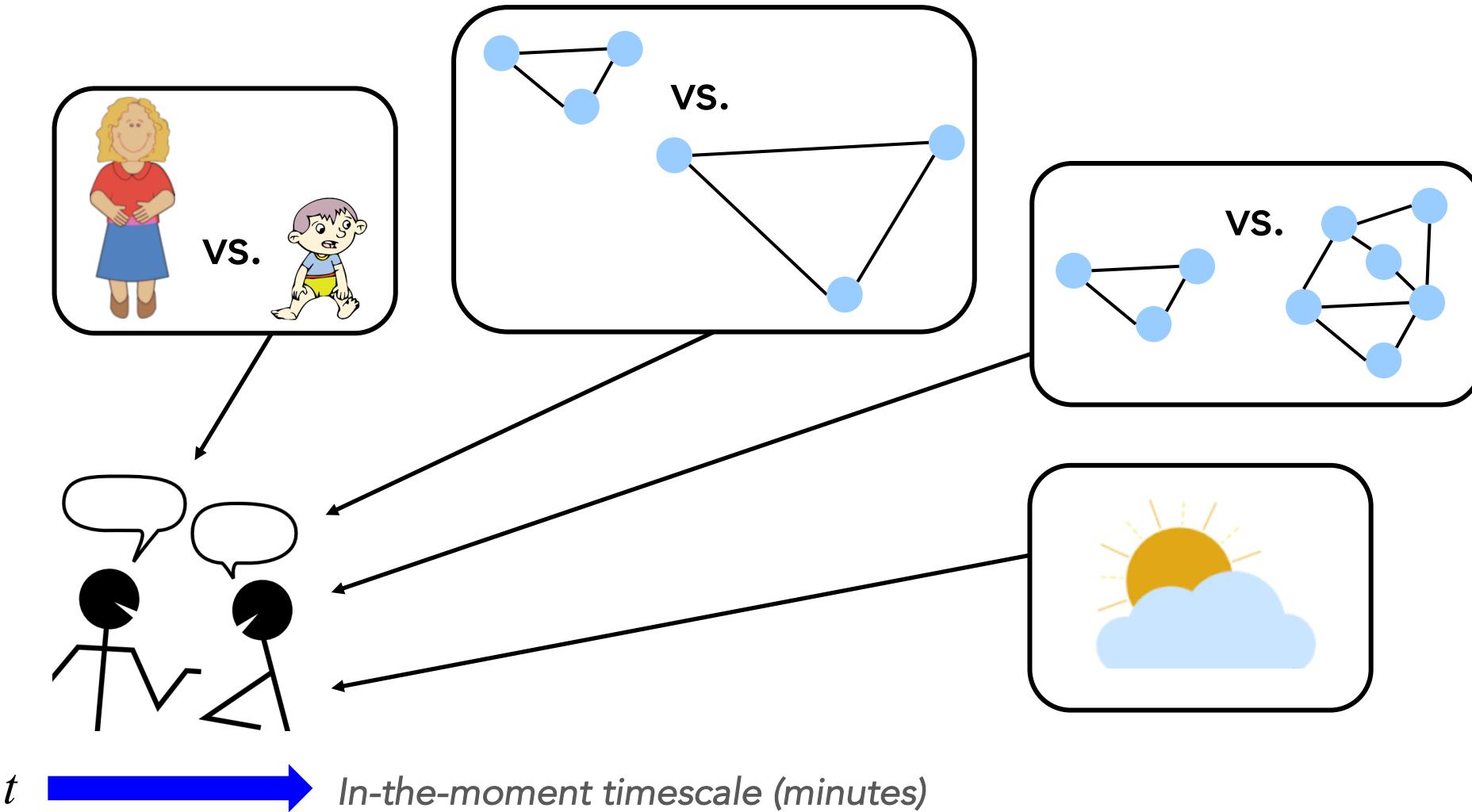
Linguistic Niche Hypothesis (Lupyan & Dale, 2010): Language systems adapt to their environments

Learnability pressures as one factor influencing the morphological complexity of a language

E.g., languages spoken by many second-language learners should be structurally simpler



Linguistic Niche Hypothesis



(Nettle, 2012)

What explains within-language variation?

- Linguistic patterns differ based on cultural customs
 - E.g., accents, formality, inside jokes, word innovation
 - Both for spoken and written communication
- “Code-switching” in more culturally diverse communities



Migration patterns are a potential mechanism

- Movement of individuals across communities can generate **cultural diffusion**
- In this study we indirectly examine mobility patterns through **proximity to highways**



Hypotheses

1. *Towns further away from highways are more linguistically homogenous*
2. *Towns further away from highways are more linguistically distinctive from each other*

Can highways be cultural connectors?

1. Motivation

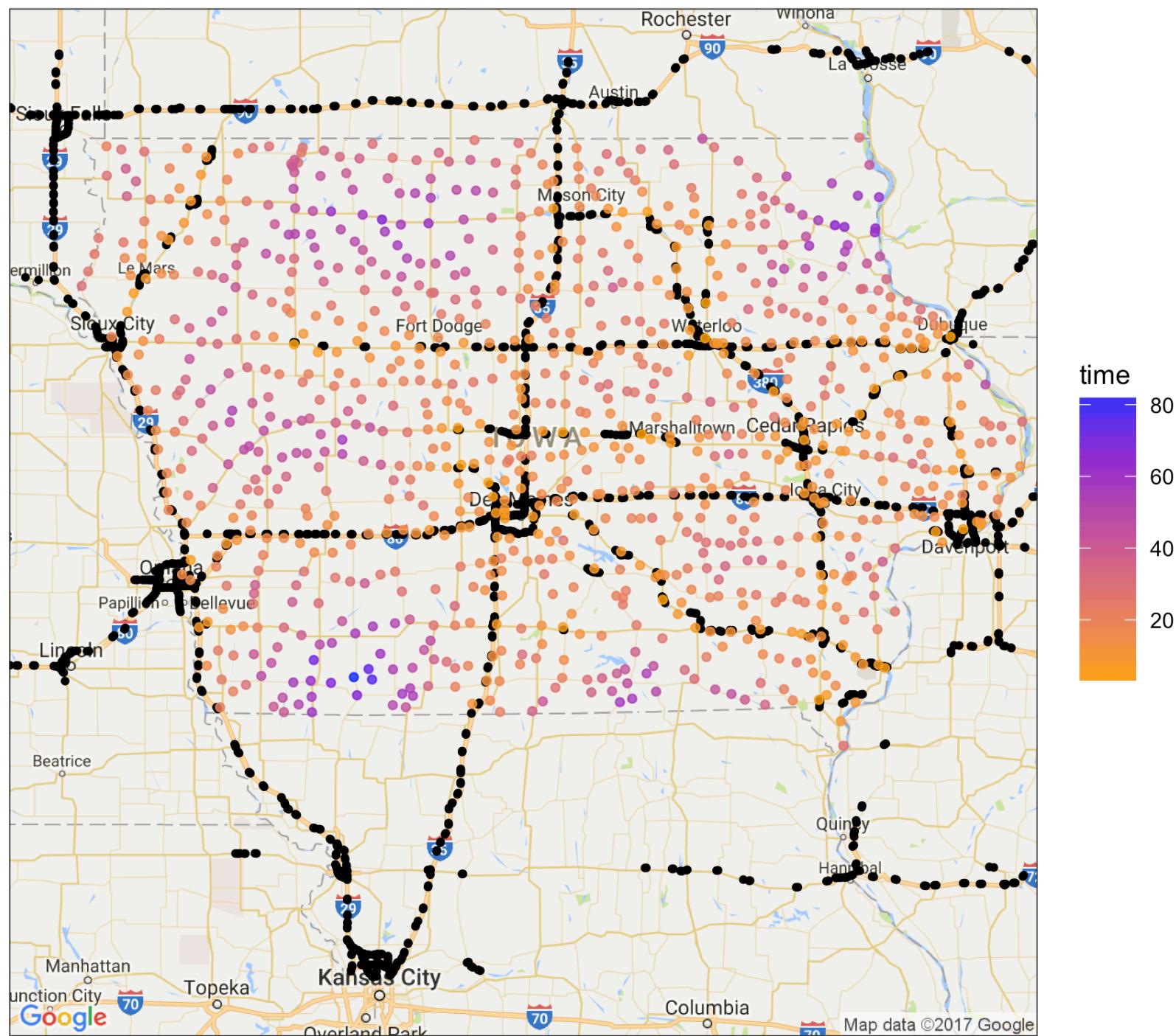
2. Methods

3. Findings

4. Next Steps

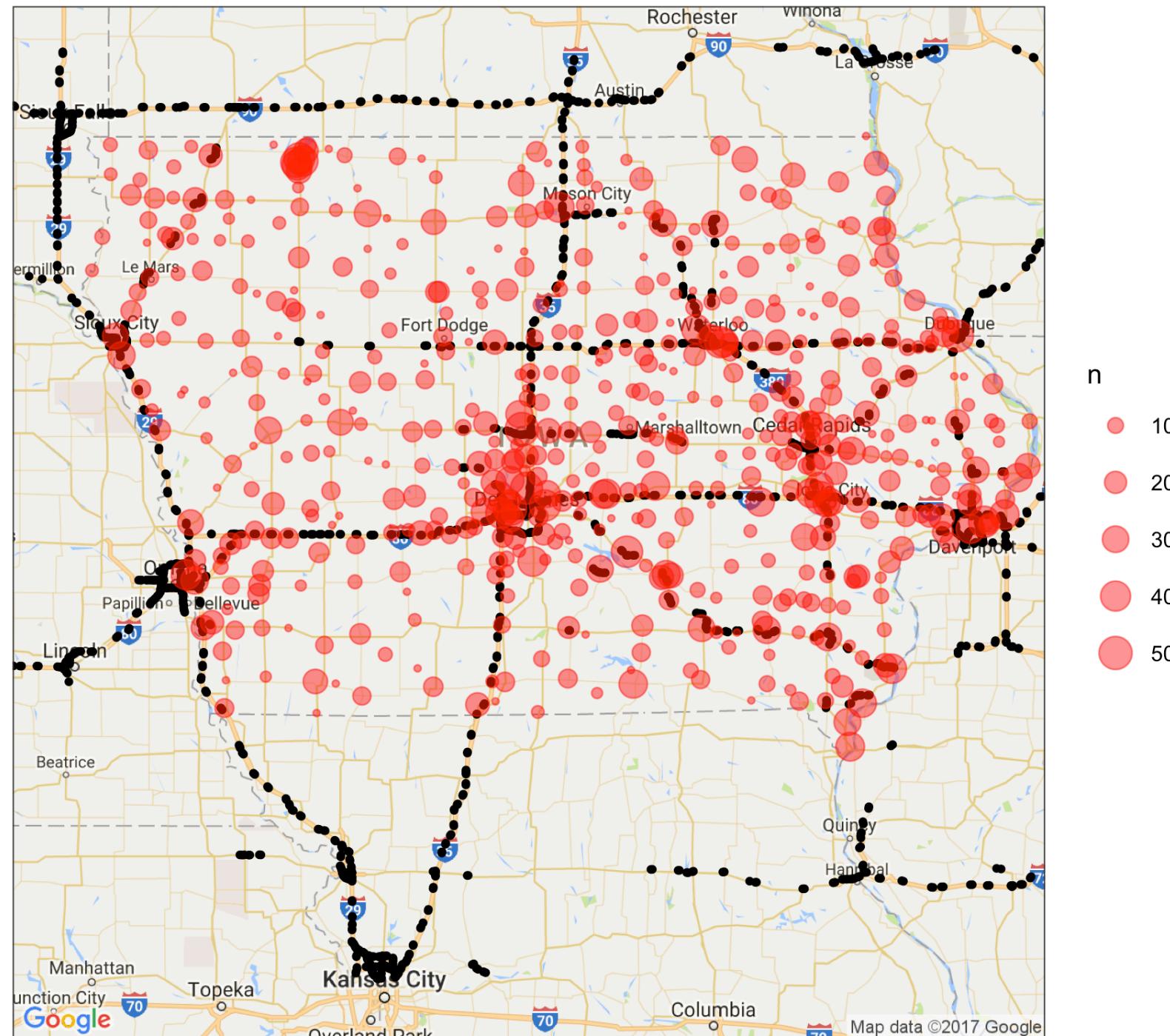
Iowa as a pilot

- Highway exits from OpenStreetMap (includes nearby exits outside of Iowa)
 - 946 incorporated cities in Iowa, excluding Des Moines
 - Driving time to highway ranges from 2 to 83 min (median of 20 min)



Text from Google place reviews

- Fetch all reviews for establishments within a 5km radius of city centers
- 24,469 total reviews
- Retain 546 cities with at least 5 reviews



Text from Google place reviews

SUBWAY® Restaurants

1561 1st Ave SE, Cedar Rapids, IA

 Write a review

4.3 ★★★★☆ 29 reviews

Sort by: Most helpful ▾



D Moore

5 months ago - 

★★★★☆ I visited this Subway on a Saturday early evening. Food was great but the staff was lazy and I as the customer had to ask was there anyone else working...then, Another Subway employee came from the back to assist the first worker. This young man proceed to complete the cold portion process. What really bugged me was that as I was requesting my items the male worker decided to stop adding my chosen topping and tell me a story of how his mother swells after she consumes a item that I choose.....WHO GIVES A F##k. I didn't stop by Subway to make a friend, I came to build a sandwich. I stated to the young man could u please stop talking about your mother and continue to build my sandwich... I'm just saying

SUBWAY® Restaurants

77 W Green St, Winterset, IA

 Write a review

Sort by: Most helpful ▾



Clarissa Thomas

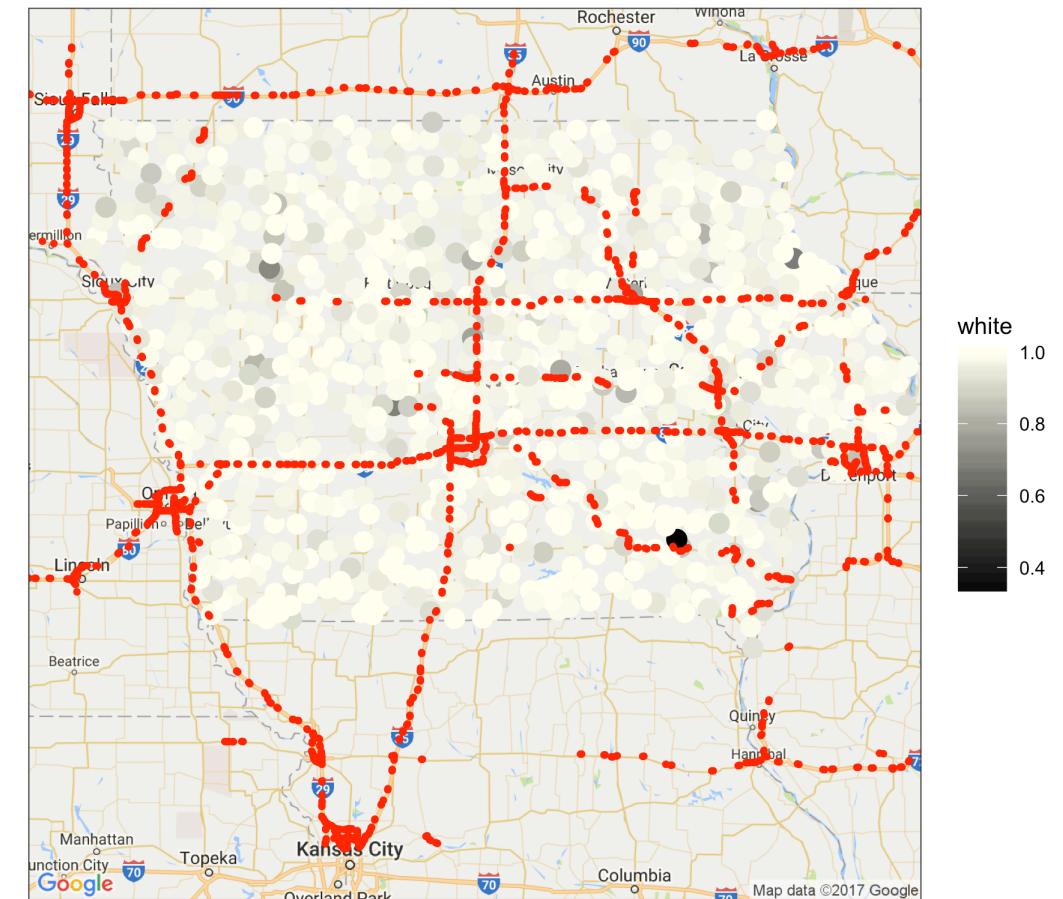
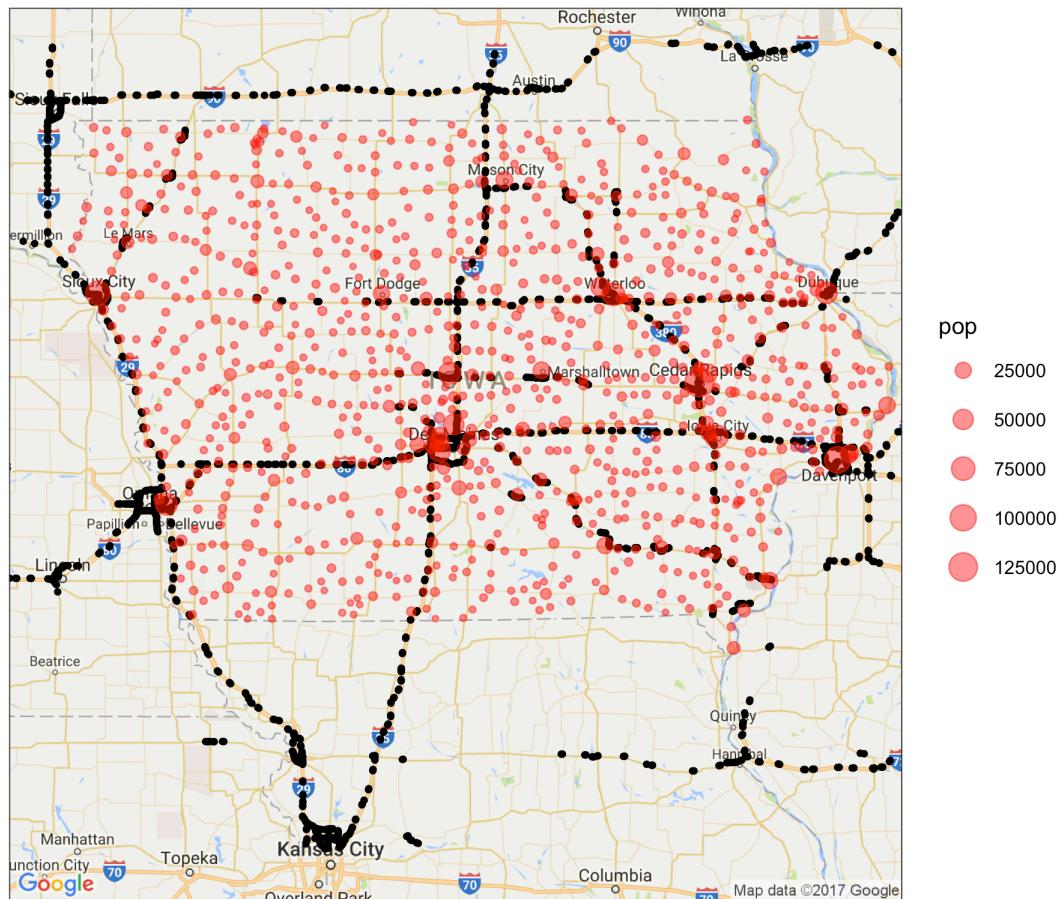
3 months ago

★★★★☆ Only reason it is getting a 4 star rating is because of the sweet girl behind the counter. Today on 3/4/17 at 1:20 pm their machine went down and are now only accepting cash atm. My sandwich was made right but the two men I encountered making it could have been a lot more pleasant, no hi how are you and no eye contact. The girl behind the counter there 10 minutes ago was so sweet even in all the chaos with it being busy and their machine being down. Thank you so much for your can do attitude and great winning smile, it's because of you that I will visit here again! :) PS ...Subway promote this girl, she is the exact definition of customer service, the other two need a lot of work.

 Helpful?

Demographic data

- Still need to incorporate data from census on population size, race/ethnicity, income, age, etc.



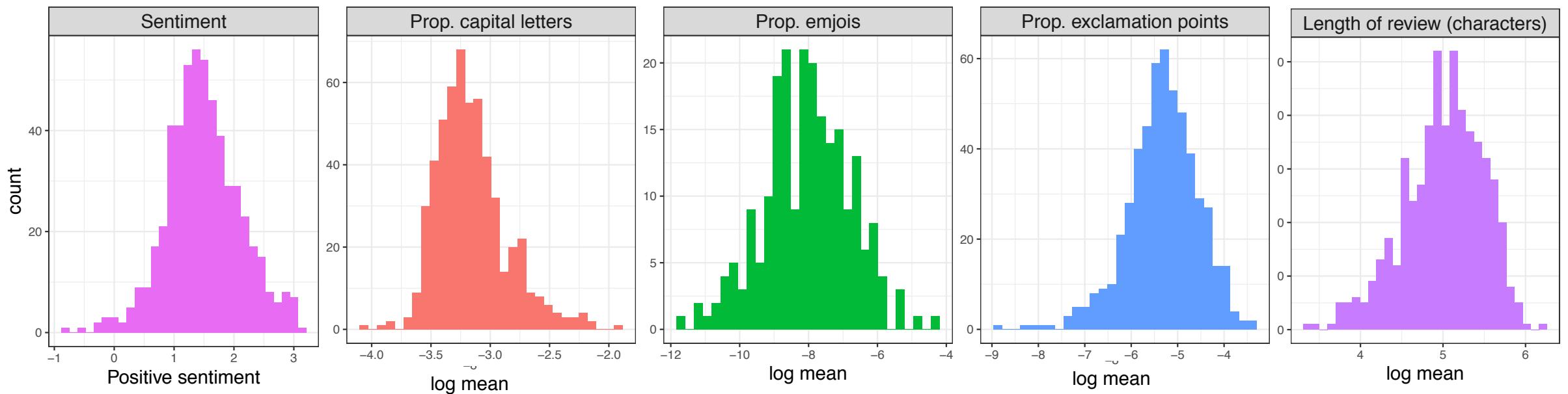
1. Motivation

2. Methods

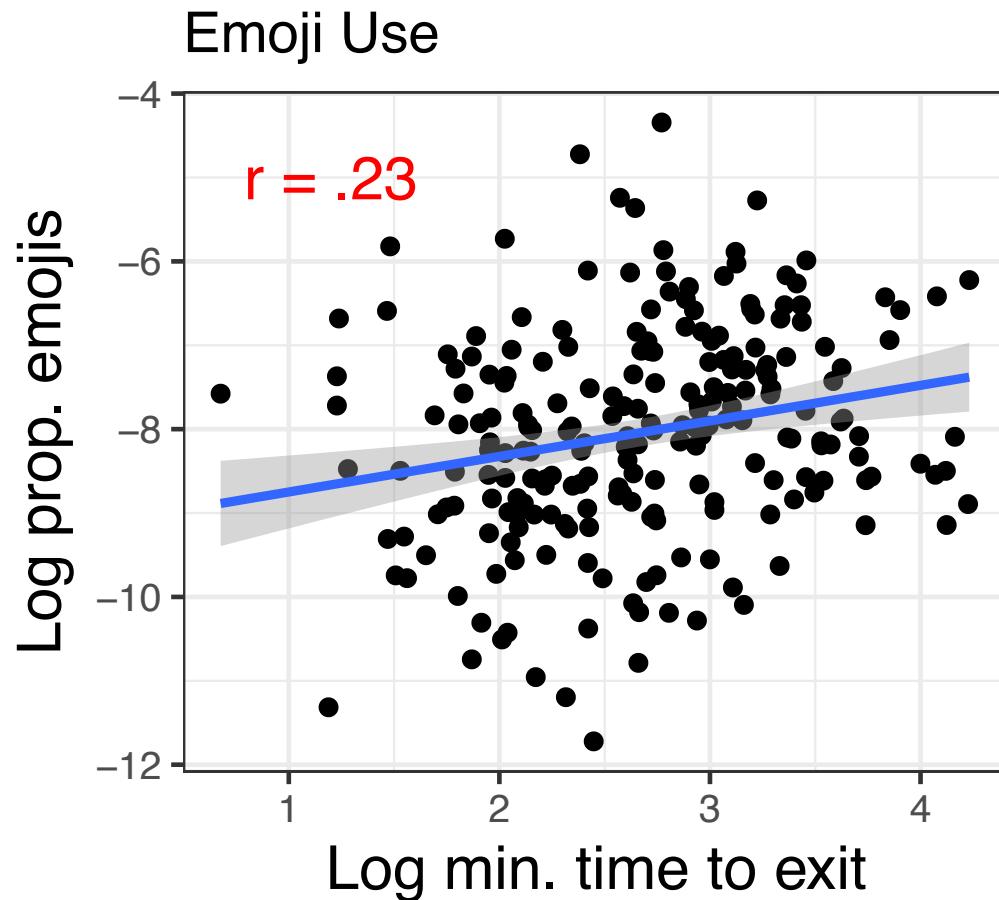
3. Findings

4. Next Steps

Linguistic features



Emotion and distance



Positive Sentiment: $r = .13$
Log. Prop. Exclamations : $r = .1$
Log Prop. Capitalization: $r = .15$

Within-town linguistic diversity

- Is town x more varied linguistically than town y?
- Multiple levels of variation
 - Vocabulary
 - Topics
 - Usage patterns
- How to measure this?
- Type-Token Ratio: Number of distinct words in each “corpus” (= town)
 - But, not measure of within-town variability
 - Highly correlated with length

Linguistic Distinctiveness

- We haven't yet explored this yet
- Will examine whether isolated towns use most distinctive vocabularies, syntax, etc.

1. Motivation

2. Methods

3. Findings

4. Next Steps

Future Directions

1. Account for heterogeneity:

- Demographics
- Content variability (e.g., only examine Subway ® or pharmacies)
- Variability in # reviews

2. Different measures of linguistic diversity/distinctiveness

3. Different measures of migration

- Inflow and outflow in cities (e.g., Craigslist ads)
- Employee movement between organizations
- Gentrification patterns

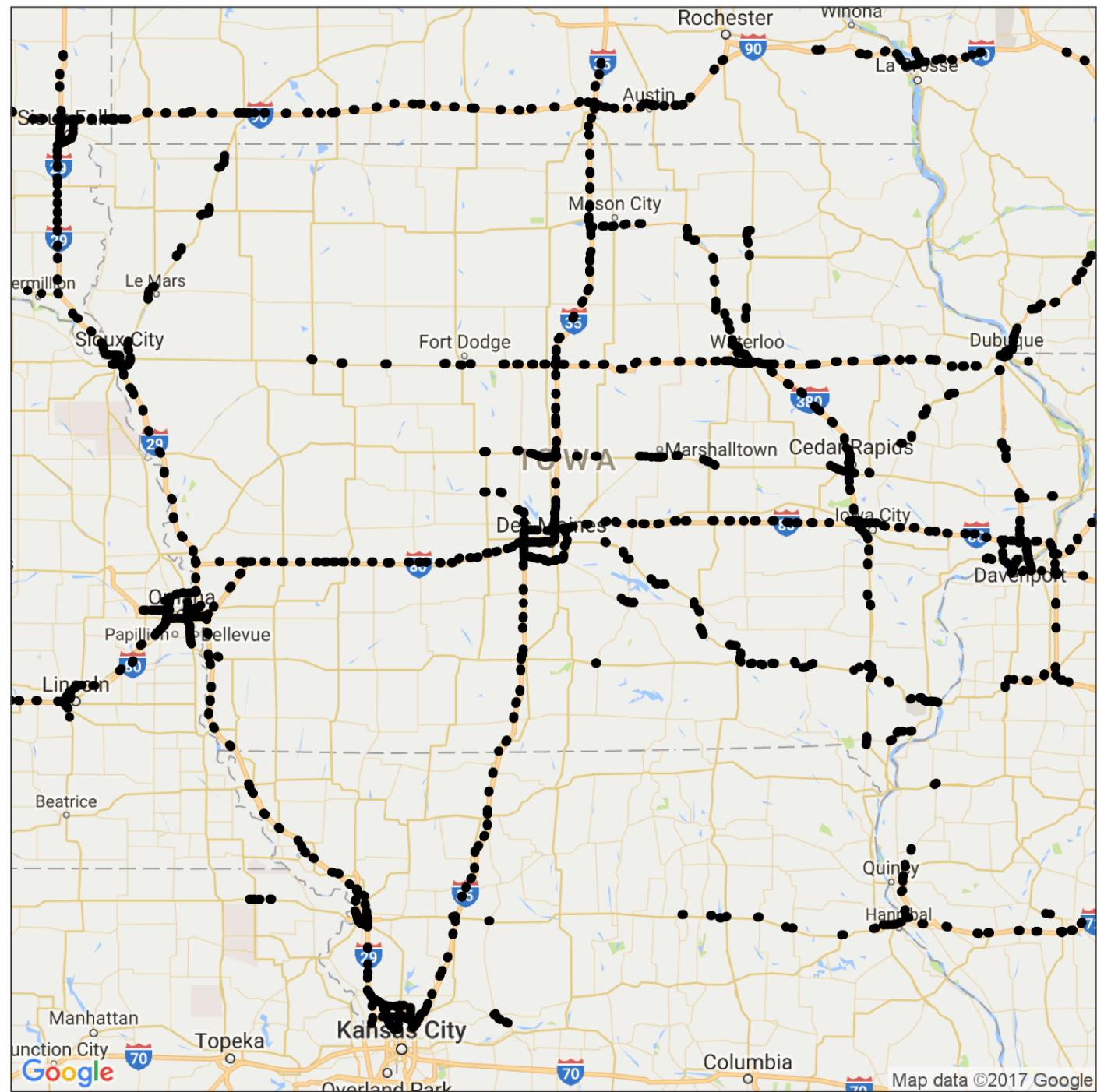
Thank you!

ambhatt@stanford.edu

mollylewis@uchicago.edu

Appendix

Highway Exits



Towns in Iowa

