

Quantifying Perceived Distance between Spatial Entities in Literary Text

Rosie Larson; Sandeep Soni
Quantitative Theory and Methods, Emory University
{rosie.larson, sandeep.soni}@emory.edu

Space as a Literary Device

Authors make stylistic choices to describe distance that modulate reader perception. As a result, how do readers perceive distance between spatial entities in narrative text? We annotated passages with masked locations for distance to assess perceived distance.

Example Annotation Passage

[...] were calmer and more composed than they had ever been since my journey to the sea of ice . A few days before we left <LOC 1> on our way to <LOC 2> , I received the following letter from Elizabeth : My dear Friend , It gave me the greatest pleasure to receive a letter [...]

Masked GPEs (as identified by BookNLP)
Less than 10 tokens between GPEs

25 tokens before and after GPEs

Annotation Task

We annotated 98 randomly selected passages that contained one of eight selected location pairs. Both locations were masked and annotators were asked to estimate distance on the following scale:

Rating	Distance (in miles)	Example
1	0-1	NA
2	1-50	SF and San Jose (40 mi.)
3	50-500	DC and NYC (200 mi.)
4	500-2,000	Miami and NYC (1100 mi.)
5	2,000-4,000	SF and Miami (2600 mi.)
6	4,000+	L.A. and London (5,400 mi.)

He felt he would go mad if he had to spend another night in **London**. Mildred recovered her good temper when she saw the streets of **Brighton** crowded with people making holiday, and they were both in high spirits as they drove out to Kemp Town.

This passage received a rating of 2, possibly due to the assumption that the characters drove between the locations.

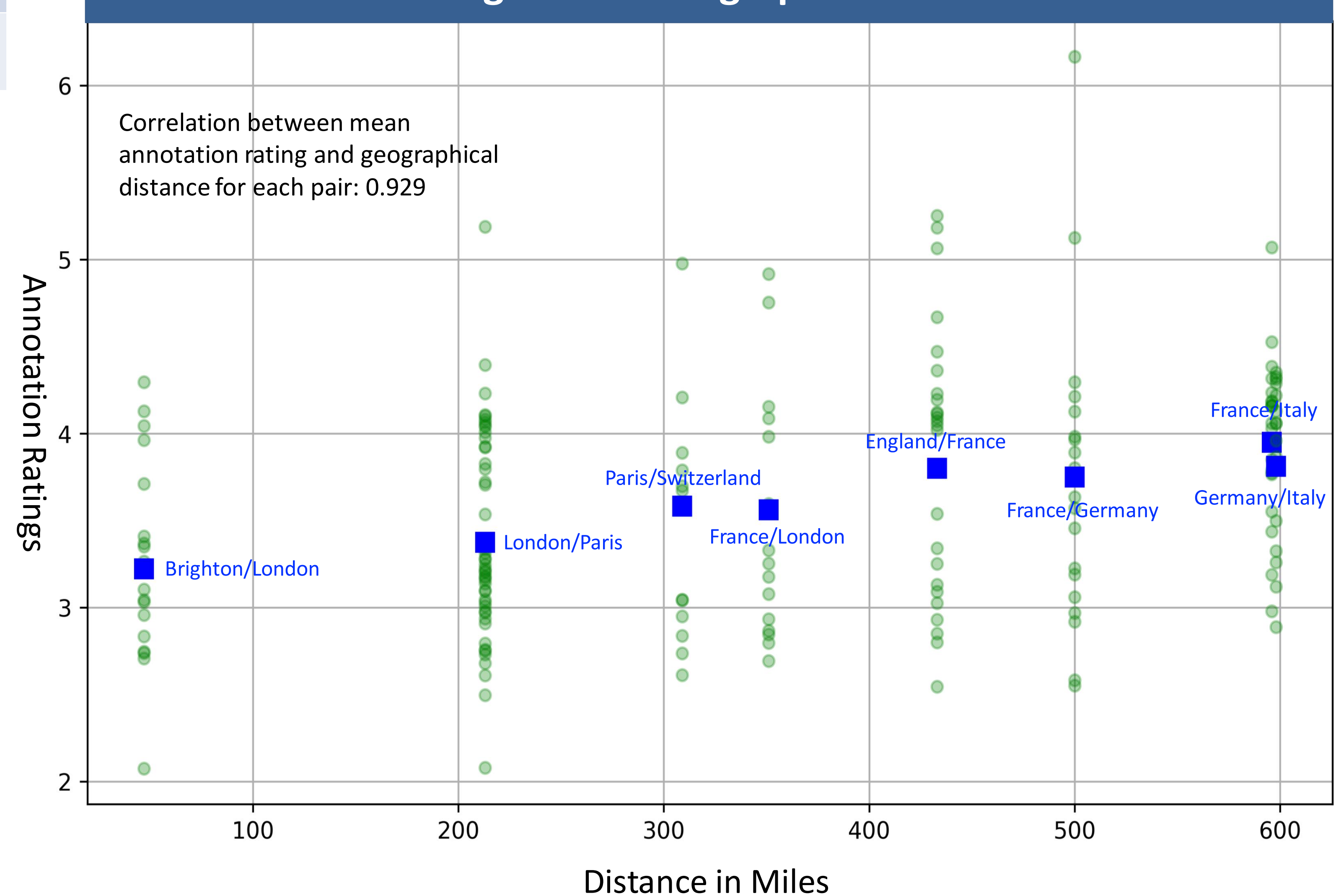
Whilst our friend George and his young wife were enjoying the first blushing days of the honeymoon at **Brighton**, honest William was left as George's plenipotentiary in **London**, to transact all the business part of the marriage.

This passage received a rating of 4, perhaps because one location is a honeymoon destination, which is usually far from home.

Initial Findings

Location Pair	# Annotations	Mean Rating	Std
Paris/London	22	3.37	0.570
England/France	12	3.80	0.707
France/London	11	3.56	0.629
France/Germany	10	3.75	0.786
France/Italy	10	3.95	0.394
Germany/Italy	9	3.81	0.403
Brighton/London	9	3.22	0.548
Switzerland/Paris	8	3.58	0.669

Annotation Rating versus Geographical Distance for 8 Pairs



Limitations & Next Steps

- Our focus on GPEs was limiting because they represent only a subset of locations mentioned in literature.
- Annotations were not equally distributed across the scale, partially because we selected only frequently occurring pairs. Reevaluation of the scale and frequency threshold is needed.