

[Cute title]: The Pragmatics of Spatial Language

Author 1 (email@place.edu)

Department of Psychology, University of Gondor

Author 2 (email@place.edu)

Department of Cognitive Science, University of Mordor

Author 3 (email@people.edu)

Department of Computer Science, University of Shire

Abstract

People use spatial language

Keywords: Pragmatics, implicature, spatial language

Introduction

The spatial world is rich and complex, but our spatial vocabulary is often coarse and limited. For example, English—like many other languages—has a restricted and closed class of spatial prepositions (Talmy, 1983, 2000; Landau & Jackendoff, 1993) such as “in,” “on,” and “near” that describe a potentially infinite set of spatial relations. This raises a natural question: How do we make accurate inference in spatial situations when our spatial vocabulary is somewhat impoverished?

A plausible solution to this question is via pragmatics (Grice, 1975; Horn, 1984). On this view, listener would make inferences that go beyond the literal meaning of speaker’s utterance by taking speaker’s perspective in a given context. After all, inference is cheap and articulation is expensive (Levinson, 2000), and therefore we expect pragmatics to play a key role in enriching the core spatial vocabulary that can be limited under many different situations.

To illustrate this idea, we describe a simple scenario that provides the basis of our study. Figure 1 shows the map of a small city with two quarters (i.e. represented by the red and blue rectangles respectively) and a plaza (i.e. represented by the dashed circle), where the plaza is located inside the red quarter. Suppose you were told that “a gold lily grew in the red quarter.” Where would you think the flower had grown on the map? There are two possibilities. In the first case, a listener who takes the literal meaning of “in” would infer the lily to have grown anywhere within the red square. This follows from the fact that the core meaning of “in” would simply refer to locations within the enclosure of the red rectangle as bounded by its sides. However, a more sophisticated listener might instead reason as follows for a more precise spatial inference: “If the lily had grown in the plaza, the speaker would have said that the lily grew in the plaza. However, given that she didn’t say so, the lily must have grown within the red quarter but in areas excluding that plaza.” Thus, this pragmatic listener would be able to make a more specific prediction than the literal listener about where the lily might have grown, even though the speaker’s utterance is identical in both cases. In this respect, pragmatic reasoning helps the listener to locate things beyond the precision of what the meaning of “in” conventionally renders by combining perspective

taking (i.e. playing the role of speaker) with one’s knowledge of the spatial situation (i.e. the fact that the plaza is located inside the red square).



Figure 1: Illustration of a simple spatial situation.

In this study, we explore the relation between pragmatics and spatial language using probabilistic programs (.). MAYBE NOAH CAN SAY SOMETHING HERE? Previous work has discussed the role of pragmatics in spatial locative expressions (Herskovits, 1985, 1987). However, to our knowledge, formal work in this domain is scarce. Moreover, recent computational studies have emphasized speaker’s role in the pragmatic use of spatial language (Carstensen, Kon, & Regier, 2014; Golland, Liang, & Klein, 2010), but not how listener makes pragmatic inference based on implicatures. We seek to bridge this gap with a formal model that makes quantitative predictions about listener’s behavior as guided by spatial language and pragmatics.

To preview our results ... MAYBE TOMER CAN FILL IN THE EXCITING RESULTS PREVIEW HERE?

Modeling spatial implicature

We model spatial implicature using the rational speaker-listener model applied previously to domains such as scalar implicature and social reasoning (Goodman & Stuhlmüller, 2013; Stuhlmüller & Goodman, 2013). We consider a speaker and listener as in the sketch shown in Figure (**). Both speaker and listener have access to a shared lexicon that contains a mapping from possible utterances to predicates that return Boolean values over states of the world. For example, in our specific lexicon the utterance “in Red Quarter” maps to a predicate that takes in a location and returns the value *True* if that location falls within the boundaries of the Red Quarter rectangle, and *False* otherwise. We expand on the lexicon details further down.

The listener attempts to infer the true state of where an event occurred in the world (s), by considering the utterance they heard from the speaker (u), and the listener's knowledge of the world and the speaker. We consider two possible listeners. The first is a *literal* listener (L_0) who understands an utterance by strictly referring to its meaning in a lexicon, modulated by the prior probability of a state occurring:

$$P_{L_0}(s|u) \propto P(s|u(s) = \text{True})P(u), \quad (1)$$

A speaker faced with such a literal listener (S_0) will choose an utterance that will maximize the probability of the listener inferring the correct state s . In continuous spatial distributions an exact equality between the observed state s and the listener's inferred state s' cannot be expected, and it is only required that an approximate equality condition holds $s \approx s'$. In this paper we use $|s - s'| < \epsilon$, though other approximate equality conditions are possible. The speaker is then:

(*** Not sure about the exact equation to put down here ***)

$$P_{S_0}(u|s) \propto *** \quad (2)$$

The second possible listener is a *pragmatic* listener (L_1), who takes into account the speaker's reasoning when choosing an utterance:

$$P_{L_1}(s|u) \propto P_{S_0}(u|s)P(s) \quad (3)$$

Such recursive, probabilistic reasoning can be captured as a probabilistic program. We implemented the model in the Church language (cite, website?) along the following lines:

```
(define (speaker state)
  (rejection-query
    (define utterance (sample utterance-prior))
    ;; Return utterance...
    utterance
    ;; Conditioned on...
    (about-equal state (listener utterance 0)))))

(define (listener utterance depth)
  (rejection-query
    (define state (sample state-prior))
    ;; Return state...
    state
    ;; Conditioned on...
    (if (= depth 0)
        (utterance state)
        (equal? utterance (speaker state))))
```

Listeners of depth 0 and 1 are *literal* listener L_0 and *pragmatic* listener L_1 as defined above. While higher combinations of listener and speaker depths are useful in games of social coordination (CITE), in many tasks the *pragmatic* listener model is sufficient, and here we are interesting is seeing how well it models people's behavior in a spatial task above that of a *literal* listener.

Utterance lexicon The utterances are implemented as boolean functions that take in a state of the world (the x,y pair where an event occurred) and return *True* or *False*. We considered three types of spatial utterances: "In" utterances (e.g. "In the Red Quarter") return *True* if the event falls within the boundaries of a region and *False* otherwise. "Near-edge" utterances (e.g. "Near the edge of the Red Quarter") return *True* if the event is within a distance d of the boundaries of a region and *False* otherwise. "Near" utterances (e.g. "Near the Red Quarter") return *True* if the event is near the edge of a region but not in it, as defined by the "In" and "Near-edge" utterances. People no doubt have access to many more spatial utterances and their combinations, but we restrict ourselves to this set for simplicity.

Parameter choice The model as stated leaves open the question of what distance counts as "Near" or "Near the edge" in the shared lexicon (what d to use), as well as what counts as "approximately equal" between the true state and the inferred state (what ϵ to use). We set both d and ϵ at 10 units (pixels). A fuller treatment of "Near" will need to use a more region-size-dependent choice of d ('near your coffee mug' is not the same as 'near your coffee house'), but is not considered here.

*** NEED TO REWRITE THIS TO REFLECT THE NEW MODEL THAT INTEGRATES OVER NEAR DISTS. ALSO MENTION 'SHARPENING PARAMETER'.

Experiment

We examined people's spatial inferences and the predictions of our model regarding the spatial terms 'in' and 'near', by putting participants in the role of a listener and asking them to guess where an event happened on a map in response to an utterance of a speaker who had access to the location of the event.

Participants, materials and methods

Participants ($N = 49$, 13 female, median age 29) were recruited through Amazon's Mechanical Turk service.

We constructed 4 simple city maps (see Fig. XX), each containing 2 "Quarters" of different size and color, and a circle marked as "Victory Plaza". The location of Victory Plaza varied among the 4 cities, while the location of the Quarters remained the same. To broadly control for effects of color and position, we created another set of 4 maps by flipping the original maps along the vertical and horizontal axis, and changing the color of the quarters, making 8 maps in total. Participants were randomly assigned to one of the two map groups.

Participants were shown an example map including a legend (as in Figure ***), and were informed that throughout the experiment they would see similar maps. Participants were told that anywhere in the city they're shown, a special flower called a 'Gold Lily' can grow, and that their task is to find the Gold Lilies.

Participants were further told that in their task a person will

IN... Red/Blue/Plaza/City



Figure 2: Results across ‘In’ utterances. Dots are guesses of where a Gold Lily grew, color coded by region ‘In’ refers to.

NEAR... Red/ Blue/ Plaza



Figure 3: Results across ‘Near’ utterances.

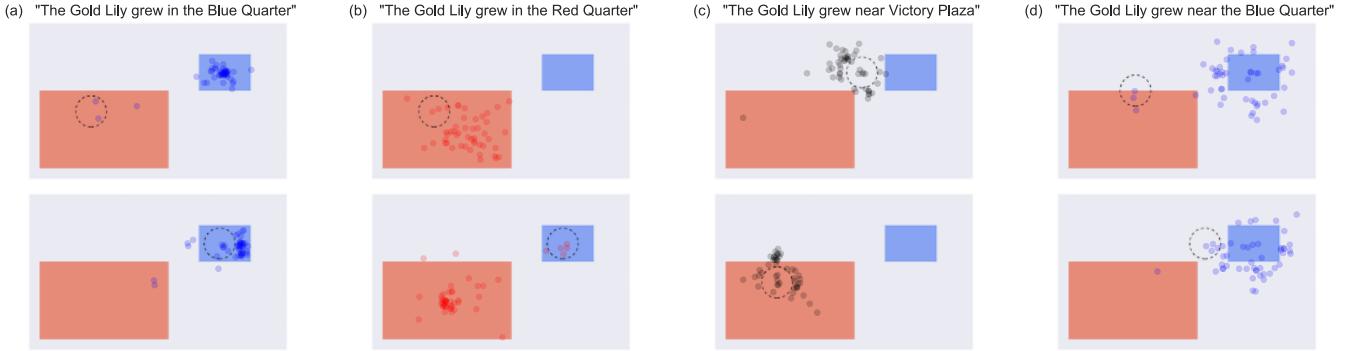


Figure 4: zoom in on results

tell them where a Gold Lily grew, and that this person can say the Gold Lily grew *in* a location or *near* a location. This person was also said to be reasonable, and honest. As an example, a participant might read the sentence ‘A person tells you: A Gold Lily grew in the Red Quarter’. Participants made their guess by clicking directly on the maps they were shown.

For each of the 4 maps in their group, participants were prompted with a sentence made of *word* \times *location* combinations, where *word* \in [In, Near] and *location* \in [Red Quarter, Blue Quarter, Victory Plaza, City]. We chose not to include the combination “Near the City”, as this area was not in the scope of the image and likely to create confusion. In total, each participant was prompted with 28 different sentences, shown in random order. Each map had a legend clearly displayed to its right.

Qualitative Results

Figures 2 and 3 show participants’ guesses on the maps by spatial word, location and city, collapsing across the rotated-and-inverted-color cities and color-coded by the reference of the word. For example, in the top-left 2, the blue dots correspond to all participant guesses for where the Gold Lily grew in City 1 when prompted with ‘In the Blue Quarter’.

Any combination of word, location and city out of the 28 possibilities is informative, but space does not allow out. Instead, we focus on several general findings, making reference to particular examples in Figure 4, and the reader is encouraged to consider other combinations for themselves.

‘In X’ implies ‘In X except Y’ As shown for example in Figure 4(a): When people hear ‘In Blue’ in City 2 (bottom) they infer ‘In Blue, but not in the Plaza’. Both the top and bottom of (a) show participant guesses for the lily location when told it grew in the Blue Quarter. In the top figure (Blue Quarter without Plaza) people place most of their guesses in the center of the Quarter (a tic-tac-toe-like division of the Quarter shows that the grid center, accounting for 11% of the area, captures 59% of the guesses). In the bottom figure (Blue Quarter with Plaza), people avoid the same center (the grid center captures 8% of the guesses) while shifting to the right of the Plaza. Such a pattern of results holds for the other re-

gions as well as shown in Figure 2.

Edge avoidance When there is no direct intersection between regions (except with the City, that all regions intersect with), people are not placing a uniform probability on the region, but rather on its center. For example, in Figure 4(b): When people hear ‘In Red’ in City 2 that has no Plaza in it(bottom), they place most of their guesses in the center of the region. A grid-analysis shows the red-grid center accounts for 65% of the guesses, as opposed to the 11% expected by chance. Using 10,000 bootstrapped simulations of 49 subjects drawn from a uniform distribution on the Red Quarter shows that in *none* of them does the center account for more than 30% of the responses.

‘Near’ is non uniform on edges As shown in Figure 4(c) and (d), when told ‘Near Plaza’ or ‘Near Blue’, people do not place a uniform probability on edges of regions. In (c), people place most of their guesses to the top-and-left of the Plaza (top) or in the top-and-right of the Plaza. In (d), people place most of their guesses to the left-and-bottom of the Blue Quarter (top) or the bottom of the Blue Quarter (bottom).

While showing some examples of spatial implicature, this pattern of results is explicitly qualitative. In the next section, we turn to a more quantitative analysis, comparing people’s results with the different listener models.

Comparison to Model(s)

In order to quantitatively compare people to the literal listener and pragmatic listener models (L_0 and L_1), we converted their responses into two-dimensional distributions. We used non-parametric, multivariate kernel density estimation to infer these distributions. The same estimation was applied to samples drawn from the probabilistic programs that represent the L_0 and L_1 .

Example comparisons are shown in Figure 5. When hearing ‘In Red’ in City 1 (left-most column) The literal listener L_0 places a uniform posterior distribution on the Red Quarter, while L_1 and people avoid Victory Plaza and lean to the right of the Red Quarter. This correspondence can be measured in terms of the Kullback-Leibler (KL) distance between the dis-

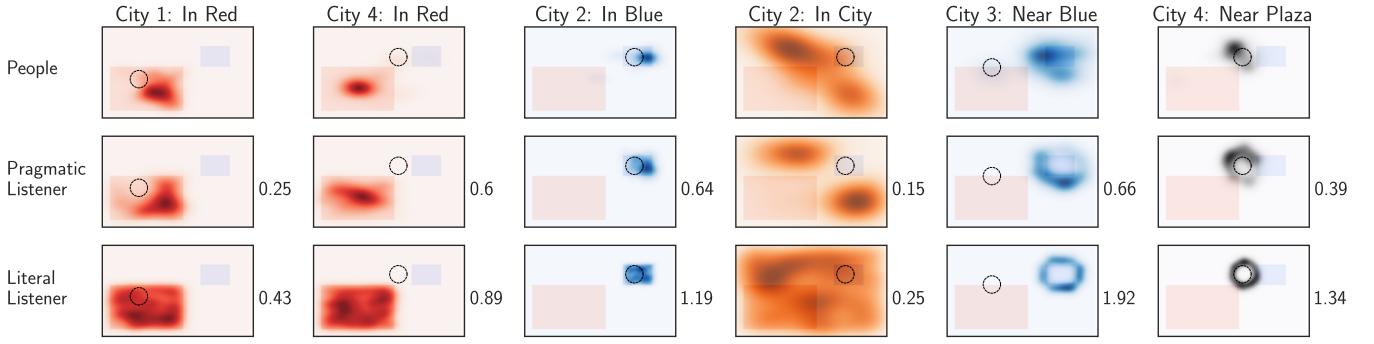


Figure 5: Example comparisons between people and the two listener models. Colored patches indicate probability distributions inferred from people’s responses and model samples. Numbers to the right of each listener subplot indicate the KL distance between the distributions of people and the model samples, lower is better.

tributions. In this particular example, $KL(\text{people}, L_0) = 0.43$, while $KL(\text{people}, L_1) = 0.25$. Other columns in Figure 5 show examples of correspondence between people’s implicature inferences and L_1 . The L_1 listener also shows the same general qualitative patterns as people, discussed in the previous section. A full comparison figure is beyond the scope of this paper, but can be found here: (*LINK*).

In terms of KL distance, the pragmatic listener L_1 is closer to people than the literal listener for 26 of the 28 (93%) possible comparisons. Figure 6 shows the distribution of the distances for both listener models. These results suggest that the pragmatic listener model is able to account for the quantitative pattern of pragmatic spatial inferences of people, within our domain.

Discussion

References

- Carstensen, A., Kon, E., & Regier, T. (2014). Testing a rational account of pragmatic reasoning: The case of spatial language. , 2009–2013.
- Golland, D., Liang, P., & Klein, D. (2010). A game-theoretic approach to generating spatial descriptions. In *Proceedings of the 2010 conference on empirical methods in natural language processing* (pp. 410–419).
- Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition.
- Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics, vol.3* (pp. 41–58). Academic Press.
- Herskovits, A. (1985). Semantics and pragmatics of locative expressions*. *Cognitive Science*, 9(3), 341–378.
- Herskovits, A. (1987). *Language and spatial cognition*. Cambridge University Press.
- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and r-based implicature. *Meaning, form, and use in context: Linguistic applications*, 11–42.

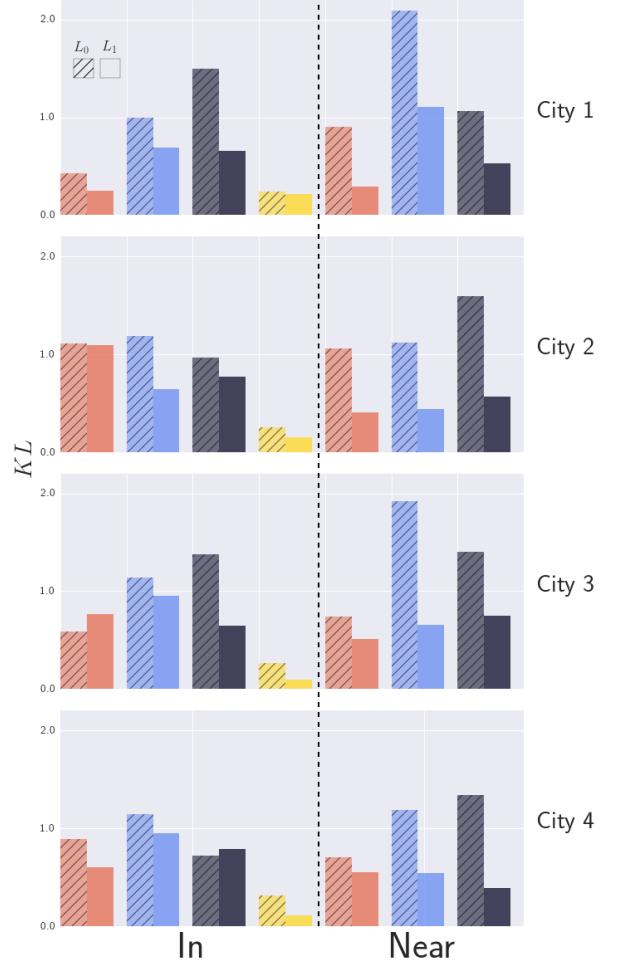


Figure 6: KL distances between distributions inferred from people and the two listener models, organized by question and city. Hatched bars are the literal listener L_0 and plain bars are the pragmatic listener L_1 . The color of the bar codes the question as before.

- Landau, B., & Jackendoff, R. (1993). Whence and whither in spatial language and spatial cognition? *Behavioral and Brain Sciences*, 16(02), 255–265.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT Press.
- Stuhlmüller, A., & Goodman, N. D. (2013). Reasoning about reasoning by nested conditioning: Modeling theory of mind with probabilistic programs.
doi: <http://dx.doi.org/10.1016/j.cogsys.2013.07.003>
- Talmy, L. (1983). *How language structures space*. Springer.
- Talmy, L. (2000). *Towards a cognitive semantics, Volume 1: Concept Structuring Systems*. MIT Press.