Spatial language: Meaning, use, and lexical choice

Kristen Johannes (kjohann@WestEd.org)
Science Technology Engineering and Mathematics Program, WestEd
300 Lakeside Dr., Oakland, CA 94612 USA

Barbara Landau (landau@cogsci.jhu.edu)
Department of Cognitive Science, Johns Hopkins University,
3400 N Charles St., Baltimore, MD 21218 USA

Abstract

Accounts of spatial language aim to address both the meaning of a spatial term and its usage patterns across diverse cases, but do not always clearly distinguish these from one another. Focusing on the case of English prepositions *in* and *on*, we set out to disentangle spatial language meaning from spatial language use by comparing judgments on a series of linguistic tasks designed to tap each aspect of spatial language. We demonstrate that judgments of truth-conditional meaning and patterns of naturalistic use show different distributional signatures, with judgments of meaning giving rise to a more uniform distribution than use patterns. We explore a third aspect of spatial language: lexical choice, and propose that choice is a key factor in shaping the distribution of spatial expression use. Our analyses reveal that the distribution of lexical choice judgments is highly correlated with the distribution of expression use in spatial descriptions for the same spatial scenes, supporting a model of spatial language that differs from traditional accounts of meaning and categorization.

Keywords: Spatial cognition; spatial language; semantics; language use

Introduction

Spatial terms in languages of the world tend to constitute a small closed class set (Landau & Jackendoff, 1993; Talmy, 1985). In English, for example, this set is typically limited to the spatial prepositions, including *in, on, over, above, etc*. To linguistically encode spatial relations with this limited inventory, a speaker must systematically abstract over fine-grained properties of objects and configurations and attend to coarse-grained spatial and/or mechanical properties of their relations. Modeling the nature of this abstraction remains a long-standing problem in the cognitive sciences. The systematic ways in which speakers encode relations (i.e., generate descriptions) is often confounded with the ways in which they decode spatial descriptions (i.e., understand the meaning of descriptions).

This problem has been exacerbated by a lack of separation between definitional questions about the *meaning* of a spatial term like *in* or *on* and categorization questions about the *use* of a term by a population of speakers – questions that may ultimately have different answers. Meaning and use represent distinct and separable aspects of many semantic domains (Cruse, 2011). In keeping with this observation, we suggest that the task of formally defining spatial terms such as *in* and *on* is separate from, albeit related to, the task of specifying the conditions under which speakers will use a spatial term to describe a location or configuration. For example, formal accounts of spatial meaning come under fire when proposed meanings cannot accommodate peripheral uses (see e.g., Bennett, 1975 for examples and e.g., Feist, 2000, and Herskovits, 1986 for commentary), while accounts of spatial categorization based on language usage patterns often propose all-or-none category boundaries that mimic binary truth conditional judgments (Regier, Khetarpahl, & Majid, 2013). In this paper, we aim to disentangle spatial language meaning from spatial language use by comparing judgments on a series of linguistic tasks across the same sets of spatial stimuli, including a task designed to directly assess speakers' lexical choices, which we propose are key in accounting for spatial term use but are not necessarily active in spatial term meaning.

Below, we review a selection of research on spatial language categorization, focusing mainly on the prepositions *in* and *on*. We organize the review into work that explicates the formal meaning of spatial terms and work that targets speakers' use of spatial terms for categorization. We then introduce recent work that suggests that speakers' *choice* of spatial term from among candidates is a critical variable in reconciling categories of spatial term meaning with patterns of speakers' spatial term use. The current study addresses these relationships – between meaning, choice, and use – directly for the English prepositions *in* and *on*, evaluating two complementary hypotheses, outlined below.

Defining spatial terms

Past and present, accounts of spatial *meanings* have also had to shoulder the burden of accounting for detailed patterns of spatial expression *use* (and, in some cases abstract uses of spatial expressions, see e.g., Jamrozik & Gentner, 2015). Traditional simplified accounts of prepositional meaning such as Bennett (1975)[1] attempt to define spatial prepositions as a function of geometric

[1] As just one example, Bennett (1975, p. 71) defines *in* and *on* using the notions of location at the interior of an object (for *in*), and location at the surface of an object (for *on*).

[2] We examined the 7 items for which the rate of expression use

properties of configurations as a means of abstracting away from specific objects. These definitional theories have been consistently criticized for being, on the one hand, too vague and allowing unlikely cases into the definition (e.g., an apple under an upside-down bowl fits Bennett's denotation of "in the bowl") and for failing, on the other hand, to predict the range of peripheral cases for which *in* and *on* can apply (e.g., an apple on top of other fruit contained by a bowl, cf. Feist, 2000). These accounts have been replaced by proposals that incorporate large sets of features in order to narrow and specify the meaning of *in* and *on* based on usage patterns (see e.g., Feist, 2000; Vandeloise, 2010; Xu & Kemp, 2012), and by proposals that prioritize world knowledge and pragmatic inference (Herskovits, 1986), so as to preserve narrow denotations for *in* and *on* while accounting for peripheral cases that depend on additional processes such as chaining (Malt, Sloman, Gennari, Shi, & Wang, 1999). The current study examines whether accounting for frequent or infrequent uses of a spatial term is a necessary goal for accounts of spatial meaning.

Spatial categories inferred from language use

Studies of spatial language categorization typically measure speakers' usage of spatial terms for different spatial scenes and, based on these data, one can infer possible category boundaries for single terms and/or semantic structure across multiple terms. Many of these accounts do not start from any initial hypotheses about the semantic content or meaning of particular spatial terms, and instead use spatial descriptions to infer systematic groupings of scenes under spatial terms. One prevailing assumption, however, is that a given spatial scene will fall "all-or-none" into only one spatial term category (e.g., the same scene cannot be categorized as both *in* and *on*).

For example, Levinson and colleagues (Levinson et al, 2003) and Regier and colleagues (Regier et al., 2013) examined spatial descriptions for a diverse set of spatial scenes from the Topological Relations Picture Series (Bowerman and Pederson, 1993). Across a large sample of languages, both groups analyzed the spatial term(s) used by the majority of speakers in a language group to encode a given scene – a point we will return to shortly.

Levinson et al. used multidimensional scaling on these data and proposed underlying spatial categories that are shaped by a handful of "attractors" – salient spatial scenes that are encoded in similar ways across languages. Similarly, Regier et al. employed an inferential (semantic map) analysis to come to a similar solution. Both studies are agnostic to the lexical content of particular spatial terms, but the researchers' analytical choices reflect a critical assumption about how spatial language use relates to underlying spatial categories. Specifically, researchers in both studies identified the modal term used by the majority speakers of each language for each scene, treating language-internal variation as noise. The result of this modal assumption is binary, all-or-none categorization of a scene by spatial terms in a language, partitioning spatial scenes into language-based equivalence classes, reminiscent of binary truth-conditional meaning.

This all-or-none semantic category structure limits the inferences that can be made about the relationship between the meaning of a spatial term and its use in encoding different spatial scenes. In particular, it ignores the possibility that spatial terms might overlap in the spatial scenes they apply to, leading to probabilistic use of multiple spatial terms, and, in a similar vein, precludes the idea that spatial terms can compete with one another to encode the same spatial scene.

Recent work from Johannes and colleagues (Johannes, Wilson & Landau, 2016; Johannes 2015; Landau, Johannes, Skordos, & Papafragou, 2016) demonstrates that multiple spatial terms are used by English speakers to encode the same spatial scenes. Moreover, they find that tracking the fine-grained use of a single spatial term across a diverse set of spatial relation scenes reveals a graded, non-uniform distribution of expression use across scenes, suggesting that some terms are a "better fit" to a spatial scene than others. In this paper, we extend the observations of Johannes and colleagues, proposing that speakers' *choice of spatial term*, among many candidates, to describe a configuration is a critical variable in accounting for the non-uniform distribution found in spatial expression usage patterns.

The Current Study

In the current study, we pursue two related hypotheses aimed at exploring how speakers evaluate the meanings of spatial terms and how this process differs from their decisions to use specific terms in spatial descriptions. We propose implicit lexical competition – speakers' choice of a particular spatial term among viable candidates – as a way of accounting for differences in speakers' judgments of spatial expression meaning and patterns of spatial expression use.

We test these hypotheses using the spatial terms *in* and *on* as a case study and compare data from three different linguistic tasks, outlined in Table 1, conducted using the same diverse sets of containment and support scenes (originally from Johannes, 2015, and Johannes, Wilson, Landau, 2016; see Figures 1 and 2). A truth-value judgment task is used to assess speakers' binary truth conditions for different expressions by simply asking whether a given expression applies to a given spatial scene. A spatial description task is used to observe speakers' self-generated spatial descriptions for each spatial scene. Finally, a forced-choice judgment task is used to measure speakers' judgments about which of two (true) spatial expressions is a better fit to a given spatial scene.

Hypothesis 1: Speakers' judgments of the truth-conditional meaning of spatial expressions are subject to different criteria than their decisions to use these expressions in spatial descriptions. We predict that tasks that separate these two types of judgments (see Table 1) will show different distributional signatures across the same set of diverse

scenes, with usage patterns yielding a more articulated, non-uniform distribution. Moreover, speaker's truth value judgments for a given spatial expression are not necessarily predicted to correlate with their use of the same spatial expression for the same spatial scenes, as tasks are hypothesized to engage different linguistic processes.

Hypothesis 2: Speakers' truth-value judgments and patterns of spatial descriptions differ due to implicit competition among felicitous candidates, which at play in spatial description tasks but not in truth-value judgment tasks. We predict that the distribution of judgments from a lexical choice task (Table 1), wherein speakers must choose between two felicitous spatial terms for diverse spatial scenes, will align with speakers' usage pattern in a spatial description task for those same scenes, but are not predicted to correlate with truth-value judgments.

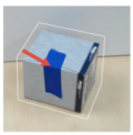
Experiment

Methods

Design. The experiment was structured as a between-subjects design with five separate groups of adult participants. Each group completed a different pairing of a linguistic task with a spatial stimulus set (see Table 1).

Linguistic tasks. Table 1 provides an example of each of the linguistic tasks, along with the range of possible responses.

Table 1. Linguistic tasks used with each stimulus set, including example prompts and possible responses.

Task	Example Prompt	Responses
Truth value judgment (Stimulus sets 1 & 2)	Is the following sentence true of the scene? "The sandwich is on the plate." 	Binary judgment: {Yes or No}
Spatial description (Stimulus sets 1 & 2)	Where is object A in relation to object B in the scene?  [A: strawberries; B: bag]	Natural language description: "The strawberries are in the bag."
Forced-choice judgment (Stimulus set 1 only)	Which of these two sentences is a better description of the scene?  A: "The tape is on the box" B: "The tape is stuck to the box"	Binary judgment: {A or B}

Participants. A total of 175 adults (mean age = 19.6 years) participated in the experiment through a series of self-paced online interfaces in return for course credit. Table 2 shows the number of participants that provided data for each linguistic task.

Table 2. Participant breakdown across tasks and stimuli sets.

Stimulus set 1 Tasks	**N**
Truth value judgment	50
Spatial Description	50
Stimulus set 2 Tasks	**N**
Truth value judgment	25
Spatial Description	25
Forced-choice judgment	25

Materials. We used two sets of stimuli to elicit linguistic judgments and descriptions. Stimulus set 1 was developed by Johannes (2015) and consisted of 64 containment scenes and 64 support scenes, for a total of 128 items (Figure 1). Stimulus set 2 came from Johannes, Wilson, and Landau (2016; adapted from Landau et al., 2016) and consisted of 18 containment scenes and 15 support scenes, for a total of 33 items (Figure 2).

Figure 1. Example containment (left) and support (right) scenes from Stimulus set 1.

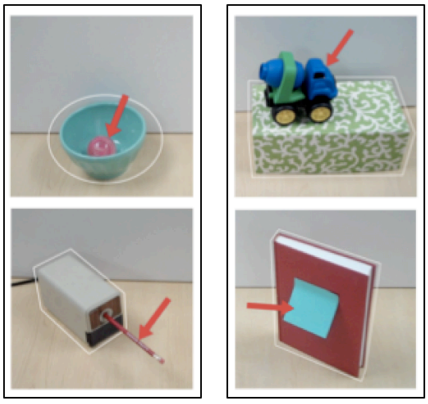


Figure 2. Example containment (left) and support (right) scenes from Stimulus set 2.

Procedure. Participants completed each task using a self-paced online interface. Tasks included critical trials, which probed linguistic judgments for the prepositions *in* and *on* for containment and support scenes (see Figures 1 and 2), respectively, as well as filler trials, which elicited judgments for other prepositions or descriptions for scenes depicting other types of spatial relationships (e.g., proximity). The number of critical trials in each task varied depending on the stimulus set: tasks employing Stimulus set 1 had 128 critical trials and 62 filler trials, while tasks that employed Stimulus set 2 had 33 critical trials and 11 filler trials.

Results

We first examined the relationship between truth-conditional meaning and expression use by comparing the patterns of participants' truth-value judgments to their expression usage patterns from the spatial description task, both carried out using Stimulus set 1. We then explored the relationship between participants' truth-value judgments, spatial descriptions and lexical choice patterns for tasks carried out with Stimulus set 2. We compared patterns of spatial expression use, from the spatial description task, across spatial scenes to patterns of truth-value judgments, from the truth value judgment task, and patterns of lexical choice, from the forced-choice task. Although participants judged, or described, both containment scenes and support scenes, we present, analyze, and discuss these spatial categories separately.

Comparing distributions of truth value judgments to spatial descriptions: Stimulus set 1. Figure 3 presents a subset of containment and support items side by side and respectively shows participants' average rates of use of *in* or *on* (i.e., proportion of descriptions using *in* or *on*) in the spatial description task (top black bars) and average truth-value acceptance rates (i.e., proportion of "True" judgments) on the truth-value judgment task (bottom white bars). Truth-value acceptance rates were greater than or equal to rates of expression use for all but 5 containment items and all but 2 support items[2].

We tested whether measures of spatial expression meaning and spatial expression use show similar distributional signatures across the same spatial scenes. Our reasoning was as follows: if participants are using the same knowledge in similar ways to make judgments about spatial expression meaning and decisions about expression use, then the resulting pattern of truth-value judgments for *in* and *on* should systematically relate to the pattern of *in* and *on* use in descriptions of the same scenes. That is, scenes that are frequently described with *in* or *on* should also show higher rates of acceptance on the truth value judgment task, and scenes for which *in* and *on* are used infrequently should show low rates of truth value acceptance. We tested this prediction using Pearson correlations, computed separately for containment and support items, between rates of *in* and *on* use in the spatial description task and rates of *in* and *on* acceptance in the truth value judgment task. The pattern of spatial descriptions and the pattern of truth-value judgments for containment items showed a weak, negative, but reliable correlation (r = -.383, n = 64, p<.01), while support items showed no reliable correlations between usage and acceptance judgment patterns (r=.059, n=64, *ns*). The weak relationship between participants' truth-value judgments and spatial descriptions aligns with the picture in Figure 3, wherein truth-value judgments show a uniform distribution across scenes, while spatial expression use in descriptions shows a more articulated usage profile.

Our analysis supports a disconnect between participants' judgments about the meaning and felicity of *in* and *on*, on the one hand, and their decision to use the expressions to describe containment and support scenes, on the other. While it is clear that meaning and use must be linked in some way (that is, speakers must have implicit knowledge of the meaning of a spatial expression in order to successfully use it to communicate), we suggest that this link is not direct and explore lexical choice – a speaker's decision about which of multiple expressions apply to a given situation – as an intervening process between meaning and use.

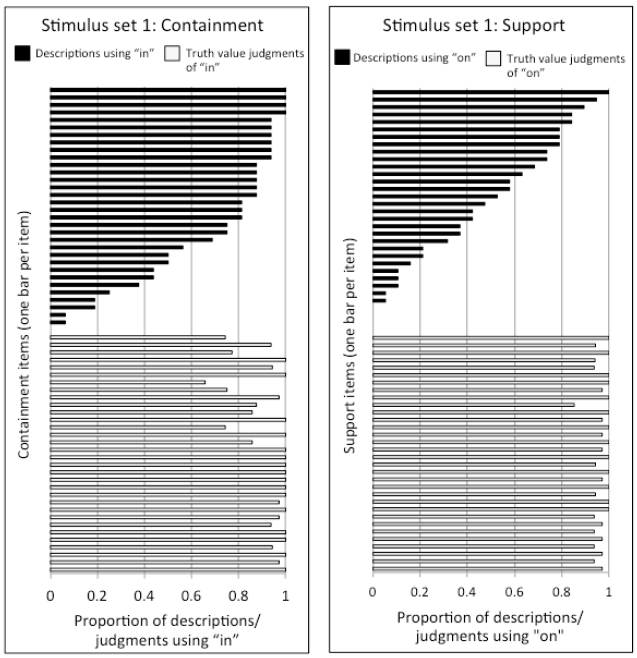


Figure 3. Patterns of expression use and truth-value judgments (bottom) of *in* (left panel) and *on* (right panel) across a subset of containment and support items.

Exploring lexical choice as an intervening variable between truth value judgments and spatial descriptions: Stimulus set 2. We collected descriptions and judgments for items in Stimulus set 2 (33 items total). Participants' responses are displayed separately for containment (Figure

[2] We examined the 7 items for which the rate of expression use exceeded the truth-value acceptance rate and found that, for all but one item, the absolute difference between use and acceptance was less than 0.3.

4) and support (Figure 5) items, which present patterns of responses from the spatial description task, forced-choice judgment task, and truth-value judgment task. As before, participants produced descriptions with *in* and *on* non-uniformly across containment and support items (top panels of Figures 4 and 5) and showed near-uniform truth value judgments across the same items (bottom panels of Figures 4 and 5). Participants' patterns of responses on the forced-choice judgment task, like their patterns of spatial expression use, showed a non-uniform distribution across items (middle panels of Figures 4 and 5).

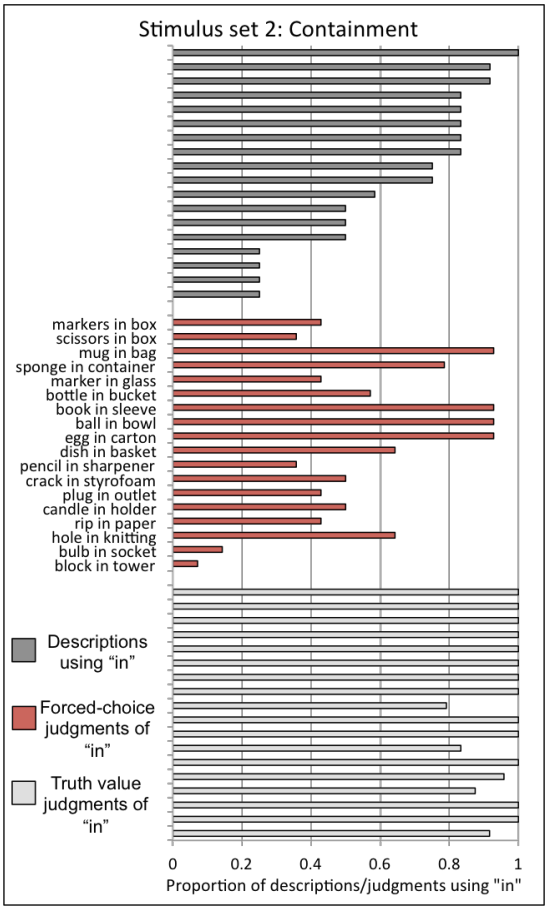


Figure 4. Patterns of use (top), forced choice judgments (middle), and truth value judgments (bottom) for *in* across containment items. Items on the y-axis are presented in the same order in all three plots.

Following our previous analysis, we first measured the relationship between participants' average rates of *in* and *on* use in their spatial descriptions and acceptance rates in their truth-value judgments. Pearson correlations between language use and truth value judgments were non-significant for both containment items (r=.306, n=18, *ns*) and support items (r=.159, n=15, *ns*).

Next, we explored the hypothesis that lexical choice, operationalized here as forced-choice judgments, serves as an implicit process in the generation of spatial descriptions but not truth-value judgments. Forced-choice judgments were not reliably correlated with truth value judgments for either containment or support. However, forced-choice judgments were strongly related to patterns of *in* and *on* use for both containment (r=.538, n=18, p<.01) and support items (r=.736, n=15, p<.01), suggesting similar variation in speakers' constrained (forced-choice) decisions about which of two expressions best applies to a spatial scene and their unconstrained decisions about how to describe the same spatial scene.

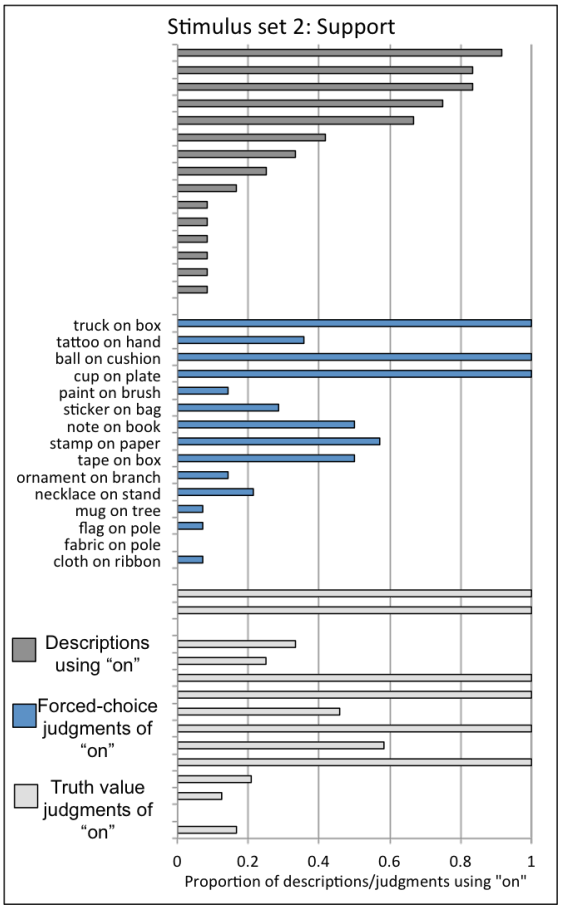


Figure 5. Patterns of use (top), forced choice judgments (middle), and truth value judgments (bottom) for *on* across support items. Items on the y-axis are presented in the same order in all three plots.

Discussion

In this paper, we compared behavior across three commonly used linguistic tasks in order to examine and elucidate the relationship between judgments of meaning, lexical choice and language use as they apply to spatial terms like *in* and *on*. We found that truth-value judgments of the meaning of *in* and *on* are nearly uniform across diverse containment and support scenes, demonstrating that these terms are true of the scenes. Speakers' *use* of these terms, however, is not uniform: some scenes are described more frequently by *in* and *on* than others.

Consistent with this distributional difference, we found no strong reliable statistical relationship between judgments of meaning vs. use for the same spatial scenes. However, when we measured judgments of lexical choice (between *in* and *on* and other truth-conditionally feasible alternatives), we discovered a non-uniform distribution of choices, similar to the distribution evidenced for spatial expression use. Our analyses confirmed a strong statistical relationship between participants' responses on these tasks across the same set of containment and support scenes.These results support a view of spatial expression meaning as partially distinct from spatial expression use.

Consequences for the possible meanings of spatial expressions. Early accounts of the meaning of terms like *in* and *on* (e.g., Bennett, 1975) came under fire (and were subsequently replaced) owing to the underspecified nature of their proposed denotation. The reasoning behind the critical reception of these theories was that a useful definition of a term like *in* should apply to exactly those cases that we most often use the term for and should rule out cases for which the term is rarely used. However, including a layer of lexical choice in the spatial encoding system, as we suggest here, allows for underspecified meanings that may over-extend to cases where the term is rarely used precisely because other better-suited terms are used in its place. For example, Johannes (2014, 2015, 2016) suggests an underspecified account of meaning for spatial terms like *in* and *on*, whereby speakers' use of these preposition is blocked by the presence of more informative lexical verbs (e.g., *hang, attach*).

Consequences for the study of spatial categorization through language use. The majority of studies on spatial categorization start by identifying a single form class – for example, prepositions in English – that serves as the primary vehicle for spatial meaning. In contrast to this, the results of the current study suggest that fine-grained spatial categorization is a function of speakers' choices between multiple felicitous expressions and not only dependent on the truth-conditional meaning of a single expression. Thus, future work on spatial categorization should expand the spatial language inventory (beyond e.g., prepositions, see Johannes, Wilson, & Landau, 2016) and focus on how categories carved out by individual spatial terms may overlap to give rise to a complex graded semantic space for this domain.

Conclusions

We have demonstrated that, for English, speakers' judgments of the truth-conditional meanings of a spatial term are not necessarily aligned with their use of that term to describe the same spatial scene. We propose that the process of choosing a spatial term among a set of felicitous competitors gives rise to speakers' non-uniform distribution of spatial expression use.

References

Cruse, A., (2011). *Meaning in Language: An Introduction to Semantics and Pragmatics*. Oxford University Press.

Feist, M.I. (2000). On *in* and *on*: An investigation into the linguistic encoding of spatial scenes. Doctoral dissertation, Northwestern University.

Herskovits, A. (1986). Language and spatial cognition: an interdisciplinary study of the prepositions in English. Cambridge, England: Cambridge University Press.

Jamrozik, A., & Gentner, D. (2015). Well-hidden regularities: Abstract Uses of *in* and *on* retain an aspect of their spatial meaning. *Cognitive Science*, *39*, 1881-1911.

Johannes, K. (2015). *Geometric and functional knowledge in the acquisition of spatial language.* Doctoral dissertation, Department of Cognitive Science, Johns Hopkins University.

Johannes, K., Wang, J., Papafragou, A., & Landau, B. (2015). Systematicity and variation in the distribution of spatial expressions in three distinct languages. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 997-1002.

Johannes, K., Wilson, C., & Landau, B. (2016). Systematic feature variation underlies adults and children's use of in and on. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 2429-2434.

Johannes, K., Wilson, C., & Landau, B. (2016). The importance of lexical verbs in spatial language acquisition: The case of *in* and *on*. *Cognition*, *157*, 174-189.

Landau, B. & Jackendoff, R. (1993). 'What' and 'where' in spatial language and spatial cognition. *Behavioural and Brain Sciences, 16(2),* 217-265.

Landau, B., Johannes, K., Skordos, D., & Papafragou, A. (2016). Containment and Support: Core and complexity in spatial language learning. *Cognitive Science, 40,* 1-32.

Levinson, S.C., Meira, S, et al. (2003). 'Natural concepts' in the spatial topological domain – adpositional meanings in cross-linguistic perspective: An exercise in semantic typolgogy. *Language, 79*, 485-516.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language, 40*, 230-262.

Regier, T., Khetarpahl, N., & Majid, A. (2013). Inferring semantic maps. Linguistic Typology, 17, 89-105.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language typology and syntactic description* (pp. 57- 149). Cambridge: Cambridge University Press.

Vandeloise, C. (2010). Genesis of spatial terms. In V. Evans & P. Chilton (Eds.), Language, cognition, and space: The state of the art and new directions (pp. 171-192). London: Equinox.

Xu, Y. & Kemp, C. (2010). Constructing spatial concepts from universal primitives. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*.