Extending Beyond Space

Brooke O. Breaux (brookebreaux@louisiana.edu)
Institute of Cognitive Science, University of Louisiana at Lafayette, Lafayette, LA 70504 USA

Michele I. Feist (feist@louisiana.edu)
Institute of Cognitive Science, University of Louisiana at Lafayette, Lafayette, LA 70504 USA

Abstract

Investigations into the semantics of the spatial and non-spatial uses of *in* and *on* have tended to assume that a type-level similarity exists between these two prepositions. However, their syntactic distributions, while overlapping, are not equal in scope (Navarro, 1998). In this paper, we ask whether these distributional differences might be related to semantic differences between the two terms. The preliminary evidence collected here suggests that *in* and *on* have slightly different levels of interpretability, even in their prepositional uses. Thus, both semantically and syntactically, the assumption of type-level similarity may need to be qualified.

Keywords: Semantics; prepositions; metaphor; language

Introduction

Investigations into the semantics of prepositions such as English *in* and *on* have tended to treat these lexical items as though they are different tokens of the same semantic and syntactic type. Such treatment seems to follow from the generative grammar tradition in which lexical category – rather than meaning – determines syntactic behavior. For example, Cook and Newson (2007) suggest “that arguments are interpreted in a particular way due to the structural positions they occupy” (p. 263). This assumption is also reflected in introductory linguistics and psycholinguistics text books, which state that words belonging to the same lexical category, or word class, are typically interchangeable syntactically (cf., Carroll, 2004; O’Grady, Archibald, Aronoff, & Rees-Miller, 2005). Together these suggest that different lexical items drawn from the same word class may interact with the rest of language in very similar ways.

Even in more cognitive views of language, we find evidence that prepositions are treated as a lexical category without indication that the individual differences between the prepositions will have important repercussions for the functions of the individual lexical items within the linguistic system. As a case in point, type-level equivalence has been assumed in examinations of the semantics of prepositions (e.g., Coventry & Garrod, 2004; Feist, 2000, 2008, in press; Feist & Gentner, 2003; Tyler & Evans, 2003; Vandeloise, in press). Much of this work focuses on the criteria that distinguish the meaning of one preposition from that of another, without discussion of the possibility that prepositions may differ in additional ways beyond their meanings. For example, while Tyler and Evans (2003) do acknowledge the importance of context in establishing the meaning of a lexical item and the fact that different prepositions will occur in different contexts, such contextual factors do not lead to different proposals regarding the nature of the meanings of individual prepositions.

However, evidence from corpus-based studies of prepositions challenges this assumption of distributional equivalence. For example, in his investigation into the semantic structure of English topological prepositions, Navarro (1998) found a differentiation between *in* and *on* based not only on their meanings but also on their syntactic distributions. While *on* tends to occur primarily in prepositional constructions, *in* is also quite prevalent within “a wide range of morphosyntactic usages that make it controversial to categorise it on behalf of a single syntactic construction” (Navarro, 1998, p. 273), including use as a full adverb, as an adverbial particle of a phrasal verb, and as a prefix for nouns, adjectives, and adverbs. This difference in syntactic distribution suggests that, despite their similarity as topological prepositions, *in* and *on* may behave quite differently within the language system as a whole.

Following up on these observations, we searched for uses of *in* and *on* in the more than 400 million word Corpus of Contemporary American English (COCA; www.americancorpus.org). Our first observation was that the frequencies of occurrence of *in* and *on* are highly unequal overall, with *in* (7,333,378 instances) appearing more than 2½ times more frequently than *on* (2,723,768 instances). Secondly, and more importantly, we examined the combinatorial possibilities for both *in* and *on* across a set of naturally occurring uses within a limited syntactic context (i.e., prepositional phrases containing the preposition immediately followed by a noun). Within the hundred most frequent collocations for each preposition, we observed an inequality in the distribution of uses, $\chi^2(1, N = 200) = 21.34, p = .0003$ (see Table 1).

Table 1: Noun types collocating with *in* and *on*

<table>
<thead>
<tr>
<th></th>
<th>Proper Nouns</th>
<th>Noun Phrases</th>
<th>Idioms</th>
<th>Concrete Nouns</th>
<th>Abstract Nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>In</em></td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>33</td>
<td>54</td>
</tr>
<tr>
<td><em>On</em></td>
<td>2</td>
<td>23</td>
<td>2</td>
<td>45</td>
<td>28</td>
</tr>
</tbody>
</table>
Taken together, these results suggest an imbalance between *in* and *on* that has yet to be thoroughly investigated.

Clearly, *in* and *on* have different meanings, which will result in the two prepositions collocating with different sets of nouns. However, these differences have not thus far led to a challenge to the assumption of type-level similarity based on their shared lexical category. As such, the differences in distribution and in combinatorial possibility that have been observed in corpus-based studies of *in* and *on* remain unexplained by the current state of thinking regarding their meanings.

There are two possible explanations for the observed differences between *in* and *on*. First, it may be that the differences are an artifact of the searches that yielded them, and that these differences would disappear given a large enough sample drawn from the corpus. In this case, the assumption of type-level similarity would remain intact, with the differences, which would be attributable to differences in meaning, limited to differences in the sets of nouns that collocate with each, but not to differences in the sizes of the sets or in the ranges of meaning types within the sets.

The second possibility is that *in* and *on* differ not only in meaning, but in meaning potential, with *in* able to collocate with a wider range of nouns than *on*. In this case, the particular semantics of *in* and *on* will have a direct influence on their potential to combine with other lexical items, rather than that potential being determined by their belonging to the lexical class of prepositions, and the assumption of type-level similarity inherited from generative grammar will need to be abandoned.

In order to discriminate between these two explanations, we will seek evidence regarding the reality of the noted imbalance using a separate methodology. If the evidence gathered from an experimental investigation of the combinatorial possibilities of *in* and *on* fails to replicate the corpus evidence, this would support a type-level similarity-based explanation wherein the noted imbalance is an artifact of the corpus searches performed. If, on the other hand, the experimental data replicates the imbalance noted in the corpus, this would support the explanation that the range of combinatorial possibilities of a lexical item is not determined by its lexical class. Rather than having their influence limited to the specific referential situations within which prepositions are deemed appropriate, meaning differences may significantly determine prepositions’ ranges of combinatorial possibilities.

In order to experimentally examine the combinatorial possibilities displayed by the prepositions *in* and *on*, we asked English speakers to interpret prepositional uses of *in* and *on* presented in the same novel syntactic and semantic contexts (i.e., the same novel sentence frames). If there is an imbalance in the combinatorial possibilities of these prepositions, then we should see different levels of interpretability for the two prepositions. To be clear, while we anticipate their different meanings to result in different interpretations of the sentences, if there are indeed differences in interpretability these should be evident in the rates at which participants attempt to provide interpretations for the novel sentences. Such an imbalance in interpretability between these lexical items when presented in identical sentence frames would suggest that the assumption of type-level similarity within lexical classes is unwarranted.

### Experiment 1

Experiment 1 tested whether novel non-spatial uses of the preposition *in* would be more easily interpretable than matched non-spatial uses of the preposition *on*. If so, participants should make more attempts to interpret sentences containing *in* than sentences containing *on*.

#### Method

**Participants** A total of 82 UL Lafayette students participated in this experiment in exchange for course credit. One student, a native speaker of Vietnamese, was subsequently removed from further analysis; a second participant was removed for not following the task instructions. The 80 remaining participants were all native speakers of English. Of these, 39 took part in the *in* condition and 41 took part in the *on* condition.

**Materials** The stimuli consisted of forty sentences constructed from twenty sentence frames. Sentence frames were in the form *These Xs are Y*; each *Y* was a non-spatial prepositional phrase (i.e., *in* or *on* followed by an abstract noun), and each *X*, a concrete noun. Each sentence frame had both an *in* variant and an *on* variant (see Table 2).

In order to provide the prepositions with a neutral playing field, the sentence frames needed to constitute unfamiliar contexts for both prepositions under consideration. At the same time, we wanted the interpretability of each sentence as a whole to hinge on the interpretability of its prepositional phrase. Thus, in constructing our sentences, we (1) selected abstract nouns that would be considered unfamiliar prepositional objects for both *in* and *on* and (2) chose as the sentential subjects nouns which would be as stable in their meanings as possible.

To accomplish these goals, we searched for twenty abstract nouns that do not frequently occur as objects of either *in* or *on*. Francis & Kučera’s (1982) rank list of lemmas was used to formulate a list of highly frequent nouns from which we could extract 100 that could potentially serve as abstract prepositional objects. Beginning with the most frequent lemma, one of us (B.B.) categorized each noun as either concrete or abstract. Nouns were categorized as concrete if they could refer to a concrete object, a concrete set of objects, a part of a concrete object, or the location of a concrete object; otherwise, they were labeled as abstract and set aside for potential use as a non-spatial prepositional object. Two-hundred and twenty-eight nouns had to be categorized in order to find 100 that fit the abstract criterion.
Table 2: The twenty sentence frames used to construct the experimental stimuli.

<table>
<thead>
<tr>
<th>Sentence frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 These houses are in/on system.</td>
</tr>
<tr>
<td>2 These rooms are in/on reason.</td>
</tr>
<tr>
<td>3 These cars are in/on idea.</td>
</tr>
<tr>
<td>4 These streets are in/on result.</td>
</tr>
<tr>
<td>5 These lights are in/on month.</td>
</tr>
<tr>
<td>6 These books are in/on hour.</td>
</tr>
<tr>
<td>7 These roads are in/on sense.</td>
</tr>
<tr>
<td>8 These tables are in/on moment.</td>
</tr>
<tr>
<td>9 These pictures are in/on voice.</td>
</tr>
<tr>
<td>10 These walls are in/on century.</td>
</tr>
<tr>
<td>11 These buildings are in/on situation.</td>
</tr>
<tr>
<td>12 These plants are in/on term.</td>
</tr>
<tr>
<td>13 These windows are in/on difference.</td>
</tr>
<tr>
<td>14 These floors are in/on statement.</td>
</tr>
<tr>
<td>15 These radios are in/on feeling.</td>
</tr>
<tr>
<td>16 These boats are in/on organization.</td>
</tr>
<tr>
<td>17 These parks are in/on basis.</td>
</tr>
<tr>
<td>18 These mountains are in/on event.</td>
</tr>
<tr>
<td>19 These blocks are in/on opportunity.</td>
</tr>
<tr>
<td>20 These apartments are in/on association.</td>
</tr>
</tbody>
</table>

We then compared our concreteness categorization with concreteness judgments collected from the MRC Psycholinguistic Database (Wilson, 1988; http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm) for each of the 228 categorized nouns. Of the 201 queries that resulted in concreteness ratings (concrete, n = 117; abstract, n = 84), the mean concreteness rating for the nouns we labeled as concrete (M = 507.99) was significantly higher than for the nouns we labeled as abstract (M = 357.51; F (1,199) = 175.52, p < .0001).1

Next, we searched COCA for combinations of *in* and *on* with each abstract noun. The 20 abstract nouns chosen for the experiment were those for which (1) combinations with both *in* and *on* produced frequency totals lower than 100 and (2) the absolute differences between the frequencies of combinations with *in* and *on* was at a minimum. A post-hoc one-way ANOVA revealed that the average frequency of *in* combinations (M = 32.70) was not significantly different from the average frequency of *on* combinations (M = 23.75), F (1, 38) = 2.77, p = .1040.

Because we wanted the interpretability of our sentences to hinge on the prepositional phrase and, hence, the compatibility of the preposition and the abstract noun, we needed the other content words to be more stable in their meanings. Previous research has suggested that object terms may be more stable in their meanings than other terms (Feist & Cifuentes Pérez, 2007; Gentner & Asmuth, 2008; Gentner & France, 1988). Thus, only count nouns that were considered by the experimenters to normally refer to inanimate concrete objects – especially when considered as a group of objects (i.e., when the noun is preceded by the adjective *these*) – were selected for use as sentential subjects.

Finally, to ensure that the *in* variants and the *on* variants of our resultant sentences were equally novel, a frequency search was conducted in COCA for each of the subject noun-prepositional phrase combinations (e.g., *house together with in system*). This search revealed that none of the final combinations appeared in the corpus.

**Procedure** Participants were randomly assigned to interpret either the *in* variants or the *on* variants. They were presented with all twenty sentences in their assigned condition in a randomized order on a computer screen. For each sentence, they were asked to either explain its meaning in the text box provided or, if they were unable to formulate a meaningful interpretation, to simply type *uninterpretable* in the text box instead of an interpretation.

**Design** We used a 2 (Preposition: *in* or *on*) x 20 (Sentence Frame) design with preposition as a between-subjects factor and sentence frame as a within-subjects factor.

**Analysis and Results**

**Stimuli Check** Before turning to our results, we ask whether the sentential subjects were less likely to shift in meaning within the context of the sentences than were the objects of the prepositions, as required by the design. To test this, we calculated for each sentence (e.g., *These houses are in/on system*) the proportion of times the nouns used as sentential subjects (e.g., *houses*) and those used as prepositional objects (e.g., *system*) in the stimulus sentences were reproduced in the participants’ interpretations. A one-way ANOVA revealed that sentential subjects were reproduced in interpretations significantly more often (M = .74, SD = .12) than their prepositional object counterparts (M = .29, SD = .17; F (1, 38) = 94.25, p < .0001), suggesting that any differences in the interpretability of the sentences would have more to do with interpretation of the prepositional phrases than with interpretation of the subjects within the wider context of the sentence, as was required by our experimental design.

**Interpretability** A repeated measures ANOVA on sentence interpretability revealed a significant main effect of sentence frame, F (19, 60) = 10.89, p < .0001, indicating that participants found some sentence frames to be more interpretable than others. Because the interpretability of the sentences was dependant on the interpretability of the prepositional phrases, this result suggests that the abstract nouns were not equally interpretable as objects of the prepositions.

Of greater relevance to the question of differences in combinatorial possibilities between the two prepositions, we observed a marginally significant sentence frame by

---

1 Of the 27 noun queries that did not result in concreteness ratings, 11 were labeled as concrete and 16 as abstract.
preposition interaction, \( F(19, 60) = 1.72, p = .0574 \). Post-hoc \( t \)-tests comparing the proportion of participants willing to provide interpretations between conditions for each sentence frame, individually, revealed four significant differences in which interpretability was higher for participants in the \textit{in} condition than for participants in the \textit{on} condition and no significant differences in the opposite direction.

Although the ANOVA did not reveal a significant main effect of preposition, we did observe a trend in the predicted direction whereby participants who interpreted \textit{in} sentences were more likely to provide interpretations (\( M = .58, SD = .19 \)) than participants who interpreted \textit{on} sentences (\( M = .51, SD = .28 \)). Furthermore, we note that the lack of a significant difference between the two conditions may have been driven, in part, by the large variances in interpretability of the two groups. Therefore, we were interested in any broader patterns in the data that might be hidden within or beneath this high variability.

We turn first to the variances of interpretability for the two groups of participants. While a significant difference between the variances of interpretability for the \textit{in} condition and the \textit{on} condition would not be the original effect we were looking for, it would suggest an imbalance, or difference, between how the different groups responded to the prepositions in question. The data show that the interpretability of the \textit{in} variant sentences resulted in lower standard deviations (\( SD = .19 \)) than the \textit{on} variant sentences (\( SD = .28 \)). When interpretability was averaged across sentence frame, Levene’s test for homogeneity of variances revealed that the mean interpretability of the \textit{in} condition was significantly less variable than the mean interpretability of the \textit{on} condition (\( F(1, 78) = 11.79, p = .0010 \)). This difference in interpretation variability, while subtle, is suggestive of a difference between the two prepositions.

To see whether any broader patterns were underlying this high variability, we next categorized each of the participants as either high-percentage interpreters or low-percentage interpreters. Since overall interpretations were provided for 54.06\% of the sentences, participants who provided interpretations for ten or fewer of the twenty sentences were considered low-percentage interpreters and participants who provided interpretations for more than ten sentences were considered high-percentage interpreters. In the \textit{in} condition, 29 participants were categorized as high interpreters and 10 as low interpreters; in the \textit{on} condition, 20 participants were categorized as high interpreters and 21 as low interpreters. This difference between conditions was significant, \( \chi^2 (1, N = 80) = 5.60, p = .0179 \). Taken together, these results hint at an effect of preposition on interpretability.

**Discussion**

Although the data hint at an imbalance between the potential interpretability of the prepositions \textit{in} and \textit{on}, we did not find the main effect of preposition that we had originally predicted. The lack of a result is particularly curious because, in a separate attempt to create novel non-spatial uses of \textit{in} and \textit{on} that would be considered by participants to be nonsensical, we had the subjective experience that nonsense \textit{on} metaphors were easier to construct than nonsense \textit{in} metaphors. While this phenomenological experience was reflected in the trends from Experiment 1, the lack of a significant main effect of preposition suggests one of two possibilities. One possibility is that our phenomenological experience may simply be different in kind from the phenomenological experience of our participants. In fact, Sandra and Rice (1995) warn researchers against relying exclusively on their own linguistic intuitions since these might differ dramatically from the intuitions of the general population.

Another possibility is that our subjective experience was driven by the task at hand. It may be that attempting to gauge the interpretability of both prepositions within the same semantic and syntactic contexts is what highlights their differences in interpretability. This difference – between the task leading to our subjective experience and the experimental task performed by our participants – is not unlike the difference between a within-subjects experimental design and a between-subjects experimental design. Birnbaum (1999) argues that participants are exposed to different contexts depending on whether they are taking part in a within-subjects experiment or a between-subjects experiment, and it is this difference in context that could result in widely divergent results from the two kinds of experiments. For example, in the between-subjects design of Experiment 1, the context for each sentence was a set of sentences involving novel prepositional phrases built upon a single preposition. In contrast, the context of our subjective experience was the creation of novel prepositional phrases built upon both \textit{in} and \textit{on}, facilitating a comparison between them. This comparison is more like the everyday experience of using language, in which novel sentences are encountered in the context of similar structures built around a variety of lexical items. Similarly, a within-subjects design in which participants would be exposed to both \textit{in} sentences and \textit{on} sentences would allow for an implicit comparison of the two prepositions, akin to the range of contexts which speakers are exposed to in everyday language use. As a result of these differences in context, the lack of a between-subjects effect for preposition might reflect more about variation in the interpretability of the novel sentence frames than about similarity in the interpretability of novel \textit{in} and \textit{on} prepositional phrases.

Thus, Experiment 2 was designed to test whether the lack of a strong result in Experiment 1 was due to differences between the linguist and the language user or to differences between a task involving consideration of multiple prepositions and one involving consideration of a single preposition.

**Experiment 2**

Using a completely within-subjects design, Experiment 2 tested whether novel non-spatial uses of the preposition \textit{in} would show higher interpretability than matched non-spatial...
uses of the preposition *on*. If so, participants should make more attempts to interpret sentences containing *in* than sentences containing *on*.

**Method**

**Participants** A total of 20 University of Louisiana at Lafayette students participated in this experiment in exchange for course credit. Two students were removed from further analysis because they identified themselves as native speakers of Igbo and Arabic, respectively. The 18 remaining participants were native speakers of English.

**Materials** The materials were the same as those used in the first experiment.

**Procedure** The procedure was the same as in the first experiment, except that participants saw all 40 of the stimulus sentences.

**Design** We used a 2 (Preposition: *in* or *on*) x 20 (Sentence Frame) design. Both were treated as within-subjects factors.

**Analysis and Results**

**Interpretability** Unlike in the between-subjects design of Experiment 1, a two-way repeated measures ANOVA on the results of Experiment 2 revealed a main effect of preposition (*F* (1, 17) = 11.87, *p* = .0031), whereby participants were significantly more likely to attempt interpretations of *in* sentences (*M* = .64, *SD* = .48) than interpretations of *on* sentences (*M* = .54, *SD* = .50), as predicted.

In addition, as in Experiment 1, we observed a significant main effect of sentence frame (*F* (19, 323) = 3.45, *p* < .0001), confirming that the sentence frames were not all equally interpretable. Furthermore, as in Experiment 1, we observed a significant preposition by sentence frame interaction, *F* (19, 323) = 1.64, *p* = .0453 (see Figure 1). In support of our prediction, post-hoc contrasts revealed that for six sentence frames the *in* variant sentence was more interpretable than the *on* variant sentence, while for no sentence frame did participants find the *on* variant sentence to be more interpretable than the *in* variant sentence.

**Discussion**

In contrast to our own subjective experiences considering the interpretability of novel prepositional phrases headed by *in* and *on*, in Experiment 1 we failed to find a significant difference in the interpretability of sentences utilizing the preposition *in* and sentences utilizing *on*. The question we wanted to address in Experiment 2 was whether the difference between our experiences and the results of Experiment 1 were due to differences between the analyst and the language user (cf., Sandra & Rice, 1995) or due to differences between considering the interpretability of two prepositions and considering the interpretability of just one. In contrast to Experiment 1, in Experiment 2 we observed a difference in interpretability between *in* sentences and *on* sentences when participants were asked to interpret both kinds of sentence, suggesting that it was the task itself that masked the differences in interpretability in Experiment 1.

In line with the observed distributional differences from the corpus-based work (see Introduction), Experiment 2 revealed that *in* can more easily appear in novel combinations with other lexical items than *on* can. This difference in interpretability between *in* and *on* suggests that the two prepositions may be operating at slightly different semantic levels.

![Figure 1: Proportion of participants providing interpretations for each sentence frame paired with each preposition. Each sentence frame is represented in the graph by its prepositional object.](image.png)

**General Discussion**

Corpus-based studies of *in* and *on* have yielded observations of differences in morphosyntactic distribution (Navarro, 1998), overall frequency, and the range of non-spatial uses of the prepositions, calling into question the validity of the type-level similarity suggested by traditional treatments of prepositions in linguistics. In this study, we asked whether these differences correspond to differences in interpretability between the two prepositions, suggesting that the noted imbalance is in fact real and supporting the interpretation that the observed differences are due to a difference in meaning potential between *in* and *on*.

Across two studies, we found that *in* and *on* did evidence semantic differences in their combinatorial potentials. When participants were asked to interpret both novel *in* prepositional phrases and novel *on* prepositional phrases, we found that they were more likely to reject as uninterpretable sentences involving *on* phrases than sentences involving *in* ones, echoing the trend in interpretability found when participants were asked to interpret just one kind of sentence. In addition, we found that all sentences for which there was a significant difference in interpretability were more often interpreted in the *in* variant than in the *on* variant. In no case did we find the *on* variant to be more interpretable than the *in* variant in our novel contexts.
Taken together, this pattern of results suggests that, in addition to having different meanings, the prepositions in and on have different semantic combinatorial possibilities.

While this result is suggestive, further investigation is necessary to understand the strength and scope of the differences between in and on. For example, in balancing the frequency of co-occurrence of the abstract nouns and the two prepositions, we considered only the frequency of the collocations between the prepositions and the abstract nouns with no intervening lexical items, leaving aside co-occurrences at greater distances (e.g., in a sense, which is very high in frequency). However, our participants could potentially have used these phrases, if familiar, to interpret the novel sentences (e.g., These roads are in a sense.). Alternatively, participants may simply have been more likely to attempt an interpretation because of the high frequency of co-occurrence between the preposition and the noun at two-step (e.g., in a sense) and three-step positions (e.g., in the traditional sense). In order to gain a clearer understanding of the differences between in and on, we are planning a follow-up experiment in which these frequencies will also be balanced.

**Conclusions**

Taken together, the differences in distribution, frequency, and semantic combinatorial possibility argue against the assumption of a type-level similarity between in and on. In addition, the fact that all three types of data point toward in having a wider range of applicability than on suggests that these three phenomena may be linked.

Our results suggest that the overall combinatorial possibilities for in may be higher than those for on. In particular, this might result in a wider range of metaphorical extensions for in than for on. As a result, investigations into the semantics of non-spatial uses of these prepositions would benefit from taking into account the differences in meaning potential between these prepositions and the possibility that the structure of the extensions and their relations to spatial uses may similarly differ.

**Acknowledgements**

We would like to thank Marlene Burke for help with data collection, Carmen Comeaux for help recruiting participants, and the Language and Cognition Lab at UL Lafayette for comments and helpful discussion.

**References**


