Mapping Space: A Comparative Study

Michele I. Feist (feist@louisiana.edu)
Department of English, University of Louisiana at Lafayette, P.O. Box 43719
Lafayette, LA 70504, USA
Yuan Zhang (mzrwhw@gmail.com)
Independent scholar, Ottawa, Ontario K2W 1H2, Canada

Abstract

The semantics of spatial terms has attracted substantial attention in the cognitive sciences, revealing both compelling similarities and striking differences across languages. However, much of the evidence regarding cross-linguistic variation pertains to fine-grained comparisons between individual lexical items, while cross-linguistic similarities are found in more coarse-grained studies of the conceptual space underlying semantic systems. We seek to bridge this gap, moving beyond the semantics of individual terms to ask what the comparison of spatial semantic systems may reveal about the conceptualization of locations in English and Mandarin Chinese and about the nature of potential universals in this domain. We subjected descriptions of 116 spatial scenes to multidimensional scaling analyses in order to reveal the structures of the underlying conceptual spaces in each language. In addition to revealing overlaps and divergences in the conceptualization of space in English and Mandarin, our results suggest a difference in complexity, whereby Mandarin terms are accommodated by a lower-dimensional similarity space than are English terms.

Keywords: spatial semantics; universals; cross-linguistic variation

Introduction

All peoples, in all languages, have occasion to talk about the locations of objects in their environments – environments which are fundamentally similar. Despite this, the vocabularies of space differ strikingly across languages, fueling interest in spatial semantics across the cognitive sciences. Most notably, scholars have observed that both the number and the nature of the contrasts that are encoded vary markedly from language to language (Bowerman 1996; Bowerman & Choi 2001; Feist 2000; Gentner & Bowerman 2009; Landau & Jackendoff 1993), with the result that "translation equivalents" for spatial terms can be quite different in meaning (Feist 2013; Trujillo 1995).

As a case in point, the range of spatial configurations that can be described using the English preposition *on* is divided amongst three prepositions – *op, aan*, and *om* – in Dutch (Gentner & Bowerman 2009); Dutch thus routinely encodes distinctions that are optional in English. More strikingly, even the dimensions of contrast encoded in spatial semantic systems may vary across languages: whereas English encodes a distinction between containment and support, Korean encodes a distinction between tight and loose fit (Bowerman & Choi 2001) that neutralizes the containment/support contrast.

Tempering these findings of variation is an overall structuring of the semantic domain of topological relations which appears to be shared cross-linguistically. For example, despite finding evidence of a "fractionated picture of overlapping contrasts" (Levinson & Wilkins 2006, p. 520) which echoes the variation briefly reviewed above, Levinson and Wilkins argue that the extensional ranges of the adpositions in the dozen languages they studied suggest a common underlying conceptual space. This is consistent with earlier findings suggesting that topological notions may be organized in a coherent conceptual space characterized by a small set of "attractors" – groups of situations that are likely to be lexicalized in similar ways across languages (Levinson & Meira 2003), including "ATTACHMENT", "IN", and "ON-TOP". Thus, while the semantics of individual spatial terms in different languages may differ from one another, the underlying conceptual components that make them up are argued to be drawn from a common set. This conclusion is supported by the work of Feist (2008), who found that the extensional ranges of spatial terms across a sample of 24 languages could be accommodated by a two-dimensional similarity space, with one dimension encoding the degree to which the reference object constrains the location of the located object, while the second dimension encodes the relative vertical positions of the two objects. Taken together, these studies in semantic typology suggest that the cross-linguistic variation that has often been noted is overlaid upon a common conceptual core.

The stark contrast between the word-level evidence of cross-linguistic variation and the system-level evidence of a common conceptual core raises many questions regarding the conceptualization of space. Is cross-linguistic variation limited to fine-grained details of lexical encoding, leaving a substantial universal conceptual basis intact? This would suggest that, while languages vary in the contrasts they mark, each structures its semantic system around fundamentally the same topological concepts. Or is the fine-grained cross-linguistic variation evidence of deeper differences in the nature of the topological concepts underlying the meanings of spatial terms? This would suggest that the system-level similarities that have been observed are in fact quite abstract, with variation arising within the set of topological concepts upon which the meanings of lexical items are based.

Zhang, Segalowitz, and Gatbonton (2011) began to address questions such as these, asking whether Mandarin Chinese and English differ with respect to the conceptual specification of containment and support rather than merely in the mapping of these two concepts onto spatial lexemes. They had speakers describe a set of 116 line drawings depicting a range of topological relations in order to examine the lexicalization of containment and support in the two languages. Each of the elicited spatial terms was classified as encoding containment, support, or "other concepts", and the extensions of containment-encoding and support-encoding lexemes in the two languages were compared. They found that approximately half of the pictures were categorized similarly at this broad level of detail (either as examples of containment or as examples of support) by speakers of the two languages, representing a large overlap in how the concepts of containment and support may be represented in English and in Mandarin. However, Zhang and her colleagues also noted differences in the uses of containment-encoding and support-encoding adpositions: Mandarin speakers described a larger proportion of the pictures using support-encoding lexical items than using containment-encoding lexical items, while English speakers evidenced the opposite pattern. Much of this difference could be accounted for via a difference in the encoding of partial inclusion and of part-whole relations: these relations tended to be described using support-encoding adpositions in Mandarin, but containment-encoding adpositions in English. This pattern of results suggests cross-linguistic differences in the boundaries separating the two conceptual categories, despite overlap in their cores, thus situating variation as a lexicalization phenomenon, rather than as evidence that the conceptual systems – and the topological concepts themselves – differ.

In a similar vein, Johannes and her colleagues asked whether the cores of lexicalized containment and support concepts were similar across languages (Johannes et al. 2015; Landau et al. 2017). They asked speakers to describe scenes predefined as representing subtypes of containment or support, then examined the rate of use of the Basic Locative Construction[1] for each subtype. For both concepts, they found that the rate of use of this construction was highest for a similar range of subtypes across the languages sampled, suggesting that these subtypes may constitute universal conceptual cores for containment and support.

Before we can conclude that the conceptual cores are indeed universal, however, we need to take a closer look at the extensions of containment- and support-encoding spatial terms as a source of evidence for the underlying structures of the concepts, without prejudging the status of either the adpositions or the scenes as exemplars of containment or support. In their study, Zhang et al (2011) classified each adposition *a priori* as encoding support, containment, or "other concepts"; they then used this classification to explore the kinds of situations that will be encoded as either containment or support in Mandarin and in English. In so doing, they neutralized fine-grained contrasts marked by the lexical items in the two languages, in essence positing that coherent, unified concepts of support and containment are encoded in English and Mandarin. In a parallel fashion, Johannes and her colleagues (2015; Landau et al. 2017) classified the scenes used in their studies as exemplars of either containment or support. In addition, they limited the scope of their study to variation in the use of BE *in/on* (and its translation equivalents), leaving fine-grained semantic contrasts unexplored. This methodology likewise assumes the existence of coherent, unified concepts of support and containment. Such unified concepts, however, cannot be assumed. As a case in point, the coherence of support as a universally salient concept has been contradicted by cross-linguistic evidence, with support relations clustering with two different groups of scenes in Levinson and Meira's (2003) analysis. In abstracting away from the semantic richness of the spatial adpositions and the complexity of the scenes, these studies may have inadvertently introduced a universal structure to the systems rather than objectively testing for its presence. In this paper, we reintroduce the semantic richness of the spatial terms while removing the *a priori* categorization of the scenes and adpositions in order to better assess the degree of similarity between the Mandarin and the English spatial semantic systems.

Spatial semantics in paradigmatic perspective

In order to better understand the comparison between the Mandarin spatial descriptors and the English ones, we shift the focus of our attention from the spatial terms as exemplars of abstract concepts to the spatial terms as indicators of linguistically-relevant degrees of similarity amongst spatial scenes (cf., Croft 2010; Feist 2008; Levinson & Meira 2003). Because words name categories, when speakers use a single word to describe two scenes, they are relying on a perceived similarity between the scenes that enables the sameness of description. Conversely, when two scenes are described using different words, speakers are highlighting differences between the scenes.

We can examine the patterns of similarity that underlie a language's semantic system via statistical techniques such as multidimensional scaling (MDS). MDS uses the co-occurrence of lexical items and pictures to construct a similarity space in which the placement of each picture is a function of the extent to which the lexical items used to describe it overlap with the lexical items used to describe each of the other pictures in the set. For example, consider the two pictures in Figure 1. If one speaker described the apple as *in* the bowl and the boat as *in* the water, this would provide evidence that the two scenes are similar, and should be placed close together in the similarity space. However, if another speaker instead described the boat as *on* the water, this would temper that judgment of similarity, and result in some distance between the two pictures. By adding in evidence from multiple speakers, a fuller picture may emerge of the extent to which each pair of pictures is treated as similar by speakers of a language.

[1] Although Levinson and Meira (2003, p. 486) define this construction as "answers to *where* questions", Johannes and her colleagues limited their investigation to BE *in/on* (and its equivalent in the other languages studied).

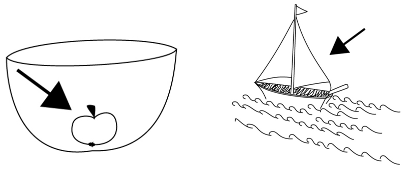

Figure 1: Two spatial scenes (from Bowerman & Pederson 1992b)

The output of multidimensional scaling is a similarity space reflecting this evidence, in which the placement of points (each representing a picture in the set) is a function of the overlap between the set of spatial terms used to describe that picture and the sets of spatial terms used to describe all the other pictures under study. The more shared spatial terms in the elicited descriptions for a pair of pictures, the closer the pictures will be in the final similarity space (see Croft 2010; Croft & Poole 2008; Feist 2008; Levinson & Meira 2003). Hence, the similarity space presents a visual representation of the conceptual space underlying the uses of the spatial terms.

To assess the adequacy of the account of the patterns in the data provided by the similarity space, we can examine both the percent correct classification, indicating the proportion of the pictures in the solution that are placed correctly relative to the elicited naming patterns, and the aggregate proportional reduction of error (APRE), indicating the extent to which the resulting solution improves upon a solution which places all of the pictures in a single category (see Croft & Poole 2008; Poole 2000, 2005 for further discussion). The APRE is measured on a scale from 0-1, with higher values indicating fewer errors in the model.

With a cross-linguistic data set, MDS returns a representation of a space upon which all the languages in the set may be overlaid such that the distinctions marked in each language isolate contiguous sets of points (Croft & Poole 2008; Feist 2008); the fewer that are miscategorized, the better the solution. As such, MDS provides a means by which we may identify potential universals underlying the semantic systems of a varied set of languages (Feist 2008; Levinson & Meira 2003). In the current study, we use MDS to construct separate similarity spaces for a set of simple spatial scenes as described by speakers of Mandarin and by speakers of English. With single-language data sets such as these, the conceptual space that MDS returns is one that only respects the distinctions marked in that language, thus providing a representation of the similarities amongst the pictures in the set as encoded in the naming patterns of the language under study. Comparison of the conceptual spaces resulting from separate MDS analyses, thus, gives a novel view into the fine-grained differences in the contrasts marked within the semantic systems of the examined languages, thus enabling a richer comparison than has been possible in previous work.

Method

The Corpus We used the 5800 picture descriptions (2900 descriptions from each language) collected by Zhang et al. (2011). The descriptions were elicited using 116 simple line drawings: 65 pictures from Bowerman and Pederson's (1992b) Topological Relations Picture Series (pictures 18, 20, 24, 33, 47, and 59 were excluded; see Zhang 2013) and an additional 51 developed by Zhang (2013). Each drawing depicts two objects – one highlighted in yellow, and one in black and white – in a simple spatial relation, with the names of the objects printed below the picture (in English or in Mandarin, as appropriate). The pictures depicted a range of topological relations; example pictures are shown in Figure 1. The pictures were printed two to a page, vertically aligned.

The set of pictures was described by 25 native speakers of English living in Montreal, Canada, and 25 native speakers of Mandarin living in Harbin, China. All speakers reported themselves either to be monolingual, or to have only limited knowledge of a second language. The pictures were presented in random order, and participants were asked to describe for each the location of the yellow object with respect to the black and white one.

Analysis In order to be able to compare the structuring of space in the two languages, the English and Mandarin descriptions were analyzed separately. We used Poole's Optimal Classification nonparametric unfolding algorithm (Poole 2000, 2005; see also Croft 2010; Feist 2008) to perform MDS analyses of the two sets of descriptions.

Our procedure was as follows. First, we identified the spatial terms used in each of the elicited descriptions. Because our aim was to analyze spatial term usage at a fine level of detail, we considered each adpositional expression to be a separate spatial term, hence *in* was separate from *inside; on,* from *on top.* This resulted in identification of 36 spatial terms in Mandarin and 38 in English. Next, we constructed two matrices – one for each language – with the 116 pictures defining the rows and with the elicited spatial terms defining the columns. Within each matrix, we then filled in each cell to indicate whether the spatial term heading the column had been used by any participant to describe the picture heading the row (cf., Feist 2008). These matrices were then input into the Optimal Classification algorithm as implemented within the R programming environment.

Results

We look first at the results for each language separately, beginning with English. Next, we turn to the comparison between the English solution spaces and the Mandarin ones.

English The lowest dimensional fit that provided a high rate of correct classification and a substantial improvement over a null model (i.e., one in which all the pictures are in a single category) was the two-dimensional solution, with 97.7% correct classification and an APRE of .765. The conceptual space associated with this solution is presented in Figure 2.

A close examination of Figure 2 reveals that the dimensions in the solution space readily admit of semantic interpretation. The x-axis, anchored by pictures of a ball underneath an upside-down bowl (and, hence, located at its interior) [picture number 110] and of a muscle in a leg [83] on the left end, and by pictures of a city at the shore of an ocean [109] and of a rope wound around a tree stump [43] on the right, corresponds to a continuum between interior location and surface contact. The y-axis, on the other hand, is anchored by pictures of a dog resting beside a dog house [6] and of a garden on the bank of a river [108] on the upper end, and by pictures of a gate in a fence [136] and of a muscle in a leg [83] at the lower. This axis thus corresponds to variation in the amount of control that the ground exerts over the figure. Along the y-axis we also see variation in the alienability of the objects (i.e., the extent to which the relation between them is inherent to their nature [Strazny 2005]), with more alienable connections (including tree/house [49] and garden/river [108]) anchoring the upper end of the dimension, and more inalienable connections (including muscle/leg [83] and gate/fence [136]) anchoring the lower end.

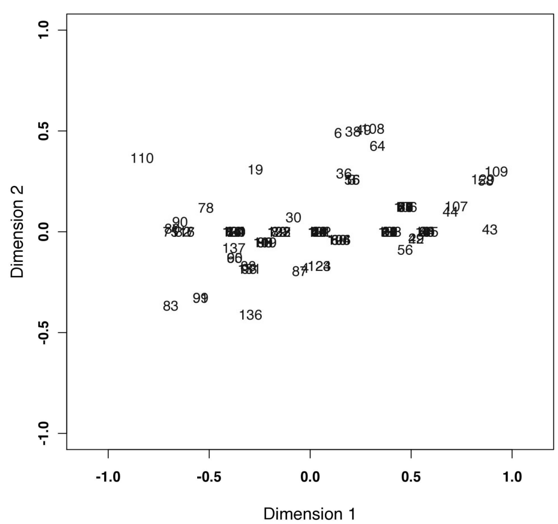

Figure 2: 2-dimensional solution for English. Each number represents one picture in the set.

The addition of more dimensions improved the fit and, even more so, the APRE. The analysis in four dimensions provided the best fit, with 99.5% correct classification, and an APRE of .947 (see Tables 1 and 2).

Mandarin The lowest dimensional fit that provided a high rate of correct classification and a substantial improvement over a null model was again the two-dimensional solution, with 98.8% correct classification and an APRE of .889. The conceptual space associated with this solution is presented in Figure 3.

A close examination of the solution space reveals that the dimension located along the x-axis encodes a continuum between interior location and surface contact. This dimension is anchored at the left end by pictures of a ball underneath an upside-down bowl [110], of a circle surrounded by a rectangle [91], and of a house surrounded by a fence [60], and at the right end by pictures of a garden on the bank of a river [108], of a city on the shore of an ocean [109], of a crease in a pair of pants [86], and of a tree at the top of a hill [65]. The y-axis, on the other hand, encodes variation in the alienability of the two objects. This axis is anchored at the upper end by alienable pairs such as a ball underneath an upside-down bowl [110], a ball under a chair [16], and a garden on the bank of a river [108]; the axis is anchored at the lower end by inalienable pairs such as a curve in a road [88], a tree growing at the top of a hill [65], and a bump in a road [123].

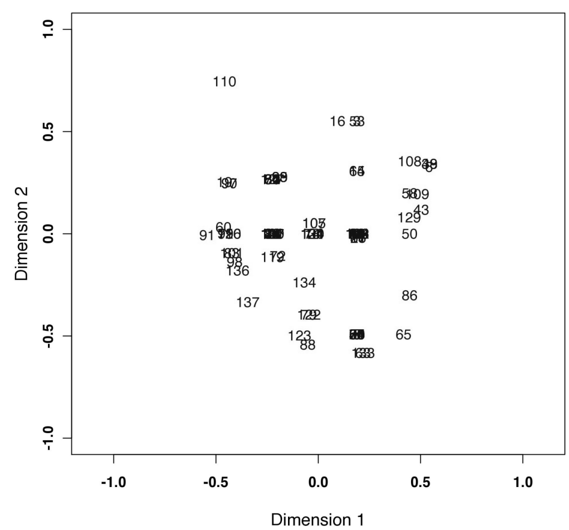

Figure 3: 2-dimensional solution for Mandarin. Each number represents one picture in the set.

The analysis in three dimensions provided the best fit, with 99.4% correct classification, and an APRE of .947, while the gains associated with a higher dimensional fit were more modest (see Tables 1 and 2).

Comparing the solution spaces

There are a number of ways in which we can compare the structurings of the spatial semantic domain in English and Mandarin. At a broad level, we can ask whether the two languages differ in the complexity of the structure of the domain, whereas at a more fine-grained level we can ask whether – and how – conceptual distinctions differ across the two languages.

Turning first to the question of complexity, we compared the two languages with respect to the adequacy of the similarity spaces produced by the MDS analyses. Despite the fact that the two languages yielded comparable numbers of spatial terms (36 in Mandarin; 38 in English), we observed that the Mandarin data was better accommodated at all dimensionalities than was the English data, for both measures of fitness. Table 1 shows the correct classification rates at one, two, three, and four dimensions for both languages. Although the adequacy of the solutions for the two languages was comparable, we note that the Mandarin solution correctly classified a slightly higher proportion of the pictures than did the English solution at each of the four levels of dimensionality considered. This pattern is replicated for the other fitness statistic, the APRE, for which the differences in adequacy between the Mandarin solutions and the English ones are more pronounced.

Table 1: Correct classification rates for four MDS solutions in both languages

	Mandarin	English
1 dimension	96.4%	95.5%
2 dimensions	98.8%	97.7%
3 dimensions	99.4%	98.6%
4 dimensions	99.7%	99.5%

Table 2: APRE for four MDS solutions in both languages

	Mandarin	English
1 dimension	.659	.551
2 dimensions	.889	.765
3 dimensions	.947	.864
4 dimensions	.976	.947

Pushing this observation farther, we found that the optimal solution in Mandarin was achieved with fewer dimensions than in English, underscoring differences between the two languages in the degree of complexity encoded in topological spatial terms and hinting at differences between the two languages in the semantic structuring of this domain. Notably, the fitness statistics for the three-dimensional Mandarin solution and the four-dimensional English solution were almost identical (99.4% and 99.5% correct classification, respectively, and APREs of .947).

Looking more closely at the placements of the individual pictures, we can ask whether the semantic structurings associated with the two languages differ at a finer-grained conceptual level. To do this, we compared the one-. two-, and three-dimensional semantic spaces across the two languages, asking in each case whether the placements of the pictures along each dimension correlated across the two languages.

We looked first at the one-dimensional solutions, which correctly classified a substantial proportion of the pictures for each language, but presented a relatively modest improvement over a null model. Our analysis revealed a substantial overlap in the placement of pictures along the one-dimensional solution (r = .68, p < .0001), suggesting significant similarity in the ways in which English and Mandarin group situation types in the spatial domain.

In both languages, we observed that the lowest dimensional solution that provided both a high rate of correct classification and a substantial improvement over a null model was the two-dimensional solution, so a comparison of the two-dimensional solutions will be especially important to our understanding of cross-linguistic variation in this domain. At first blush, the English and Mandarin two-dimensional solutions share many similarities: both include one dimension that encodes a continuum between inclusion and surface contact and one dimension that encodes the alienability of the figure-ground relation. However, a closer look reveals that these similarities are but part of the story, co-existing with important differences in the details of the solution spaces.

We consider first the details of the continuum between interior and surface contact. While the English and Mandarin continua overlap, reflected in high correlation between the coordinates along this dimension across the two languages (r = .75, p < .0001), we noted important differences in the placements of many of the pictures in our set. First, we observed that some pictures, such as the crease in pants [86] and the light bulb in a socket [133], are located toward the surface contact end of Mandarin's Dimension 1 but more centrally in the English solution space. In addition, many examples of three-dimensional full inclusion (e.g., an apple in a bowl [2] and a fish in a fishbowl [32]) can be found towards the center of the expanse of Mandarin's Dimension 1, but farther towards the inclusion end of Dimension 1 in English. Looking more closely, we observed that the sets of pictures anchoring the inclusion end of Dimension 1 differed between the two languages: in English, this dimension is anchored by examples of three-dimensional inclusion such as a ball underneath an upside-down bowl [110] and of a muscle in a leg [83], whereas in Mandarin this dimension is anchored by examples of two-dimensional inclusion such as a circle surrounded by a rectangle [91] and a house surrounded by a fence [60]. Whereas all these scenes could be classified as "containment" (cf., Johannes et al. 2015; Landau et al. 2017), these differences suggest that even though both Mandarin and English draw upon a contrast between inclusion and surface contact, the Mandarin system privileges two-dimensional over three-dimensional inclusion, whereas the English system privileges three-dimensional over two-dimensional inclusion. In addition, this dimension is far more spread out in English than in Mandarin, suggesting not only differences in the nature of the inclusion concept, but also differences in the linguistically-relevant degree of similarity amongst the pictures along this dimension.

Turning to the second dimension, we noted less overlap in the semantic interpretation (above), reflected in weaker correlation between the coordinates along this dimension across the two languages (r = .36, p < .0001). Furthermore, whereas this dimension encodes alienability in both solution spaces, this factor is connected to the amount of control exerted by the ground in English, but not in Mandarin. This suggests that control may play a larger role in the semantics of English spatial terms than in the semantics of the Mandarin terms. In addition, this data suggests that the closeness of the relation between two objects may be more likely to be independently assessed and taken into account for speakers of Mandarin than for speakers of English.

Finally, we compared the three-dimensional solutions, which improved substantially beyond the two-dimensional solution in English, but less so in Mandarin. Just as the gains in moving from two dimensions to three differed in the two languages, we also observed that the placements of the pictures were less congruent in the three-dimensional solutions (dimension 1: $r = .60$, $p < .0001$; dimension 2: $r = .37$, $p < .0001$; dimension 3: $r = .28$, $p = .0023$) than what had been observed for the two-dimensional solutions. This decrement in congruence suggests greater differences between the two semantic systems become evident at more fine-grained levels of analysis.

Conclusions

Our study extends the evidence regarding cross-linguistic variation in the semantics of spatial terms beyond comparisons of the meanings of individual terms by comparing and contrasting the conceptual spaces underlying the semantics of spatial terms in English and in Mandarin. Whereas our findings reinforce the conclusion from past work (i.e., Feist 2008; Levinson & Meira 2003; Levinson & Wilkins 2006) that the conceptual spaces underlying spatial relational inventories are subject to similar factors across languages, this similarity is tempered by evidence that both the factors and the conceptual spaces themselves differ in subtle ways across languages.

We consider first the findings regarding the spatial semantic systems in the two languages. Our Mandarin data was accommodated with fewer dimensions than was our English data, suggesting that the semantic complexity of spatial relational systems varies cross-linguistically. Furthermore, whereas the optimal scaling solution in Mandarin drew upon two dimensions, echoing Feist's (2008) optimal solution for a cross-linguistic dataset, the optimal solution for our English dataset required additional dimensions. This may indicate that each language elaborates on and, hence, may add complexity beyond a universal conceptual core. Thus, while this conceptual core may provide a skeletal structure for how humans think and talk about spatial location (cf., Feist 2008; Levinson & Meira 2003), it markedly underspecifies what we need to encode in order to effectively function in a spatial world.

The strongest correlation between the solution spaces for the two languages involved Dimension 1 of the two-dimensional solutions, which encoded a continuum between interior location and surface contact. In addition to reflecting the importance of the distinction between containment and support (cf., Zhang 2013), this dimension echoes continua that have emerged from other cross-linguistic studies of the semantics of spatial terms, including Bowerman and Pederson's (1992a; see also Bowerman & Choi 2001) similarity gradient and the dimension corresponding to location control in Feist's (2008) MDS analysis. However, whereas Feist's (2008) MDS solution conflated location control and the interior-surface continuum, suggesting that the two may often be inseparable, the English solution reported here separates the two as individual dimensions. This suggests that, whereas both factors are important cross-linguistically and are related, individual languages will make use of different options regarding the extent to which factors are separated in their semantic systems.

Our findings further suggest that cross-linguistic variation may extend beyond fine-grained details of lexical encoding into the nature of the topological concepts themselves. As in the work of Johannes and her colleagues (2015; Landau et al. 2017), our findings underscore the importance of containment and support concepts in the semantics of topological terms. However, a close examination of the continuum between interior and surface contact in the solution spaces for English and Mandarin revealed that the distribution of inclusion scenes differs considerably across the two languages, suggesting differences in the underlying containment concepts. In English, we observed that scenes in which the ground surrounded the figure in three dimensions were placed farther toward the interior end of the continuum than were scenes in which the ground surrounded the figure in two dimensions, suggesting that three-dimensional inclusion is more prototypical than is two-dimensional inclusion in this language. In contrast, in Mandarin we observed the opposite pattern, suggesting that two-dimensional inclusion constitutes a better example of the concept than does three-dimensional inclusion in this language. While inclusion played an important role in structuring spatial semantics in each case, the conceptual cores around which the inclusion concepts were structured differed. Thus, cross-linguistic differences lie not only in the ways terms are distributed relative to a conceptual distinction, but also in the kinds of scenes considered to be best examples of the anchoring conceptual categories.

Whereas MDS allows visualization of the conceptual space underlying semantic systems, it does not afford a picture of the meanings of the lexical items themselves, nor does it afford a close look at the encoding possibilities for individual spatial scenes. The lower correlations observed between the dimensions of the three-dimensional solutions suggest that important cross-linguistic differences may only become evident when viewed at this fine-grained level of analysis. In future work, we will complement the current analysis with analyses focused on these two aspects of spatial semantics. To better understand the detailed ways in which the lexical items relate to one another, we will examine frequencies of use of each term for each picture – i.e., the behavioral profiles (Gries 2010) of each of the elicited terms. To better understand variation in the codability of the scenes, we will assess the breadth of descriptions elicited by each scene. Taken in combination with the current study, these analyses will afford a better understanding of the ways in which individual lexemes fit together to create a semantic system.

Taken together, the current results present a rich picture of the interplay between universals and variation in the semantics of spatial terms. While languages may draw upon a common set of concepts to structure meanings in this domain, these concepts may in fact be quite underspecified. As a result, the variation in meaning that has been observed across languages may be indicative of variation in the ways in which the universal conceptual core has been developed in each language in order to produce a useful set of concepts for communication.

References

Bowerman, M. (1996). The origins of children's spatial semantic categories: Cognitive versus linguistic determinants. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity*. Cambridge, MA: MIT Press.

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press.

Bowerman, M., & Pederson, E. (1992a). Cross-linguistic perspectives on topological spatial relationships. Paper presented at the 91st Annual Meeting of the American Anthrolopogical Association, San Francisco, CA.

Bowerman, M., & Pederson, E. (1992b). Topological relations picture series. In S. C. Levinson (Ed.), Space stimuli kit 1.2. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.883589.

Croft, W. (2010). Relativity, linguistic variation and language universals. *CogniTextes, 4*. Retrieved from http://journals.openedition.org/cognitextes/303.

Croft, W., & Poole, K. T. (2008). Inferring universals from grammatical variation: Multidimensional scaling for typological analysis. *Theoretical Linguistics, 34*: 1-37.

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

Feist, M. I. (2008). Space between languages. *Cognitive Science, 32*(7), 1177-1199.

Feist, M. I. (2013). Experimental lexical semantics at the crossroads between languages. In A. Rojo & I. Ibarretxe-Antuñano (Eds.). *Cognitive linguistics and translation: Advances in some theoretical models and applications*. Berlin: Mouton de Gruyter.

Gentner, D., & Bowerman, M. (2009). Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. In: J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura and Ş. Özçalişkan (eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*. New York: Psychology Press.

Gries, S. Th. (2010). Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon, 5* (3), 323--346.

Johannes, K., Wang, J., Papafragou, A., & Landau, B. (2015). Similarity and variation in the distribution of spatial expressions across three languages. D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp.997 - 1002). Austin, TX: Cognitive Science Society.

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

Landau, B., Johannes, K., Skordos, D., & Papafragou, A. (2017). Containment and support: Core and complexity in spatial language learning. *Cognitive Science, 41* (4), 748-779.

Levinson, S. C., Meira, S., & The Language and Cognition Group. (2003). "Natural concepts" in the spatial topological domain - adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language, 79* (3), 485-516.

Levinson, S. C., & Wilkins, D. P. (Eds.) (2006). *Grammars of space: Explorations in cognitive diversity.* Cambridge: Cambridge University Press.

Poole, K. T. (2000). Non-parametric unfolding of binary choice data. *Political Analysis, 8*, 211-237.

Poole, K. T. (2005). *Spatial models of parliamentary voting*. Cambridge: Cambridge University Press.

Strazny, P. (2005). *Encyclopedia of linguistics*. New York: Fitzroy Dearborn.

Trujillo, A. (1995). Towards a cross-linguistically valid classification of spatial prepositions. *Machine Translation, 10*, 93-141.

Zhang, Y. (2013). *Spatial representation of topological concepts IN and ON: a comparative study of English and Mandarin Chinese.* Doctoral dissertation), Individualized Program, Concordia University.

Zhang, Y., Segalowitz, N., & Gatbonton, E. (2011). Topological spatial representation across and within languages: IN and ON in Mandarin Chinese and English. *Mental Lexicon, 6*(3), 414-445.