Spatial Language Learning

Zhengling Fu
College of International Studies, Southwest University, Chongqing 400715, China

This research is financed by The Social Science Planning Project of Chongqing (2015YBYY134), The Higher Education and Teaching Reform Project of Chongqing (153030), and The Fundamental Research Funds for the Central Universities (SWU1509176)

Abstract
Spatial language constitutes part of the basic fabric of language. Although languages may have the same number of terms to cover a set of spatial relations, they do not always do so in the same way. Spatial languages differ across languages quite radically, thus providing a real semantic challenge for second language learners. The essay first examines the variables that underpin the comprehension and production of spatial prepositions in English. Then the essay reviews the functional geometric framework for spatial language and a computational model of the framework that grounds spatial language directly in visual routines. Finally, the essay considers the implications of the model for both first and second language acquisition.
Keywords: spatial language, second language acquisition, functional, geometric

1. Introduction

Spatial language, such as asking directions for places in large-scale space, or asking where a misplaced object is in small-scale space, constitutes part of the basic fabric of language. Languages carve up space in different ways. However, spatial language exhibits problems for second language learners.

2. An experimental history of spatial languages

In the early 1990s empirical evidence for the importance of location control and other extra-geometric variables started to emerge. This body of experimental work led Coventry and Garrod (2004) to develop the functional geometric framework for spatial language comprehension. In the functional geometric framework a distinction is made between three types of information that in combination underpin the comprehension and production of spatial prepositions: geometric routines, dynamic-kinematic routines and specific knowledge of how objects are likely to interact in standard situations.

2.1 Dynamic-kinematic routines

The first dynamic-kinematic route to be identified was location control which underlies the comprehension and production of *in* and *on*. For a located object to be *in* a container, this entails more than just spatial enclosure/inclusion; containment also entails the notion of location control whereby the reference object constrains the location of the located object over time. Location control is a concept which is amenable to experimental testing, and a number of studies showed just how important location control is for the comprehension of *in* and *on*.

In one series of experiments video images of fruit and balls were shown in various positions in relation to containers. For example, an apple was shown perched high on top of a pile of other apples in a bowl (the reference object). The distance between the apple and the bottom and rim of the container was varied so that the apple could be almost touching the bottom of the container in one scene and positioned on top of a large pile of apples high above the rim of the container in another scene. Critically, location control was also manipulated. In the strong location control condition, the whole scene was shown moving from side to side, with all the objects maintaining the same relative positions over time. In contrast, in the weak location control condition the apple was shown wobbling from side to side while the other objects in the scene remained stationary. So in the strong location control condition, location control is seen to hold because the bowl is clearly controlling the location of the apple over time. However, in the weak location control condition the bowl is not seen to control the location of the apple because the apple is moving independently of the other objects in the scene. A control (no movement; static scene) condition was also included. If location control is important, it should be expected that *in* would be preferred more in the strong location control condition than in the control condition, and preferred least in the weak location control condition.

A second way in which location control has been manipulated involved static scenes where there were alternative potential sources of location control. Garrod, Ferrier and Campbell (1999) presented participants with video-clips of different arrangements of a pile of Ping-Pong balls and a glass bowl. They manipulated the geometric relationship between a black Ping-Pong ball and bowl by varying the distance between the located object and the reference object. The other factor that was manipulated was the degree to which the location of the black Ping-Pong ball could be seen to be controlled by an external source. In half of the scenes the ball was attached to a thin (but visible) piece of wire suspended from above the bowl. According to the extra-geometric routine of location control, viewers' confidence in describing the black Ping-Pong ball as being *in* the bowl should relate directly to the degree to which they see the container as controlling the location of the located object (i.e., the black Ping-Pong ball). In this study half the participants made judgments regarding the appropriateness of different descriptions of the configuration of ball and bowl. The other half were asked what would happen to the black Ping-Pong ball if the bowl were moved sideways. The proportion of viewers who judged the ball as maintaining its relation to the bowl following the hypothetical movement was taken as an indicator of the degree to which the bowl was seen as controlling the location of the ball. Garrod, Ferrier and Campbell found that location control affected linguistic judgments. Additionally, comparing the two sets of judgments—location control judgments and judgments as to the appropriateness of *in* descriptions—there was a strong correlation between the two sets of judgments.

So there is considerable evidence for the application of what we have termed the dynamic-kinematic routine of location control to *in*. However, there are two points to note. First, location control in the studies thus far mentioned was important only when the containment relation was weak (i.e., when the located object was positioned above the rim of the container). Second, there are many cases where location control does not seem to apply. For example, *The marble in the circle* does not seem to involve location control. Therefore, although this routine is clearly central to the comprehension and production of *in*, and children are very sensitive to it when they are acquiring spatial terms, we need to retain a geometric routine which allows calculation of degree of containment/ enclosure independent of the extra-geometric routine of location control.

2.2 Situational knowledge

The application of geometric and dynamic-kinematic visual routines is driven by knowledge of the objects involved in the scene and how those objects typically interact in past learned interactions between those objects. Objects are associated with particular routines, both geometric and dynamic-kinematic, and prepositions have weightings for these parameters. As we have seen above, the comprehension of *over/under* is better predicted by extra-geometric relations than the comprehension of *above/ below*, while conversely the comprehension of *above/below* is better predicted by geometric routines than the comprehension of *over/under*. In the functional geometric framework it is how these constraints "mesh" together (Glenberg, 1997) that underpins the comprehension of spatial prepositions. Here we give some examples of how object and situational knowledge impact on spatial language use.

Although containers all have the function of constraining contents over time (location control), some are more appropriately designed for liquids whereas others are only suited to hold solids. Jugs are usually associated with containing liquids, bowls usually contain either liquids or solids, and sieves are designed to contain solids alone. Coventry, Carmichael and Garrod (1994) found that *in* was judged to be more appropriate for a solid in a bowl compared to a solid in exactly the same position in a jug with the same dimensions as the bowl when the solid object was on top of a pile of similar objects. Furthermore, adding liquid to the jug further decreased the use (and rating) of *in* but made no difference in the case of the bowl. Coventry et al. (1994) argued that the addition of water appears to make the object-specific function of the jug more salient, further reducing the appropriateness of the jug as a container of solids. This finding is not limited to jugs and bowls, but generalizes to a basic difference between containers that are primarily containers of liquids versus those which are primarily containers of solids.

How an object is labeled also influences its perceived function in a scene and in turn influences how one describes the location of an object with reference to that object. Coventry et al. (1994) found that the specific labels given to a reference object influenced the use of prepositions to describe the spatial relation between a located object and that reference object. When the same reference object was labeled a *dish* versus a *plate, in* was rated as more appropriate and was used more frequently in the sentence completion task. When the object was labeled *plate, on* was preferred and rated more highly. Therefore, objects become associated with particular geometric and/or dynamic-kinematic routines.

The fact that an object may be conceptualized in many different ways leads to situations in which the same object may be processed using different routines, and these may vary according to the language itself. In English a person can be *in* or *on* a bus or plane, but o*n* a car would only be appropriate if the person was standing on top of the external shell of the car. Vehicles afford location control, but can be conceptualized as containers or supporting surfaces. For English the application of *on* relates to the size and the length-to-width ratio of such vehicles; large long vehicles are usually regarded as supporting surfaces, but small objects with a low height-to-width ratio are more usually regarded exclusively as containers. In contrast vehicles are all conceptualized as containers in Polish—the equivalent of *in* is appropriate for vehicles but the equivalent of *on* is not.

It is not only objects that drive geometric and dynamic-kinematic routines— relations between objects are also important. Carlson-Radvansky, Covey and Lattanzi (1999) asked participants to place an object above a second object, and they varied the association between the two objects. For instance, when the reference object was a toothbrush (with the bristles on the left and the handle of the toothbrush extending to the right), the object to be placed was either a toothpaste tube or a tube of paint. They found that placements of the objects were between the middle (center-of-mass) of the toothbrush and bristles. However, the associated objects (toothpaste tube) were placed nearer the functional part of the reference object than the unrelated objects (the tube of paint). So knowledge of how objects typically interact affects spatial language judgments.

Finally it is important to note that how objects interact is defined in context. Coventry et al. (2001) compared objects that do not have a known protecting function with those that do. While an object such as a shield has an obvious protecting function, the function of the stool is as a supporting surface. Although Coventry et al. found that ratings for the inappropriate functional objects (e.g., *The stool is over the Viking*) were lower overall than for the sentences involving appropriate protecting objects, no interactions were found between this variable and any of the other variables examined. In other words, the effects of functionality and geometry were present for the non-stereotypically functioning objects just as they were with the stereotypically functioning objects. This is evidence that how objects are functioning in context is important, irrespective of our stereotypic knowledge about those objects.

2.3 The functional geometric framework

Within the functional geometric framework, dynamic-kinematic routines and the knowledge of the objects involved in the scene and how those objects typically interact in past learned interactions between those objects all come together to establish the situation-specific meaning of spatial language. The meaning of a spatial expression does not simply derive from the addition of the fixed meanings of the preposition together with the meanings of other elements in the sentence (e.g., nouns and verb). Rather, meaning is constructed on-line as a function of how these multiple constraints come together.

The view that meaning is constructed on-line as a result of putting together multiple constraints is consistent with recent work on embodiment, and in particular the work emanating from the labs of Glenberg, Barsalou (e.g., Barsalou, 1999) and Zwaan (e.g., Zwaan, 2004; Zwaan & Taylor, 2006). Common to these accounts is the notion that words are associated with a range of types of perceptual information. Knowing what an object is requires knowing what one does with it, and therefore its representation should reflect how one can interact with that object, a representation that can prepare you for situated action. For instance Kaschak and Glenberg (2000) have argued that the meaning of a sentence is constructed by indexing words or phrases to real objects or perceptual analog symbols for those objects, deriving affordances from the objects and symbols and then meshing the affordances under the guidance of syntax.

3. First language learning

3.1 Prelinguistic knowledge of geometric and dynamic-kinematic routines

According to the model developed by Coventry and Garrod (2004) and the implementation of it, spatial language involves a combination of grounding symbols directly in perceptual representations as well as learning how symbols co-occur in the language being learned. The goal of the language learner is therefore to bind linguistic and perceptual information together in order to map language onto meaningful events.

The starting point for learning spatial language is the pre-linguistic perceptual knowledge of the world that is present from the early stages of life. There is good evidence that infants have knowledge of both geometric and dynamic-kinematic routines. In relation to geometric routines, preferential looking studies have established that babies only a few months old can distinguish between a range of relations including *left* and *right* and *above, below* and *between*. Preferential looking relies on the much replicated finding that infants have a tendency to look longer at novel than at familiar stimuli. For example, Quinn (1994) habituated 3-and 4-month-old infants to a single diamond presented in different positions above a (horizontal) bar. Infants then saw two diamonds in novel positions; one above the bar and one below the bar. Infants consistently showed a visual preference for the diamond presented in the novel position below the bar. This strongly suggests that they had formed a category for *above*.

There is also evidence that infants as young as 2.5 months have expectations about containment events and location control. Hespos and Baillargeon (2001) showed infants an object lowered inside a container with either a wide opening or no opening in its top surface. Infants looked longer at the closed-container event rather than the open-container event indicating the infants have knowledge about containers. In another experiment infants saw an object lowered either inside or behind a container.

The container was then moved forward and to the side showing the object visible behind it. Infants looked longer when the object was revealed in the condition where the object was initially placed inside the container than in the condition where it was initially placed behind the container. This was probably as the infants realized that the object could not pass through the wall of the container, and hence should have moved with the container to the new location—an early appreciation of location control.

Infants also have early knowledge of gravity. Kim and Spelke (1992) showed infants videotaped events in which a ball rolled downwards or upwards while speeding up or slowing down. At 7 months infants looked longer at test events with inappropriate acceleration. Similarly, when they were shown a stationary object released on an incline they looked longer when the object moved upwards. In another study Needham and Baillargeon (1993) have shown that 4.5-month infants have expectations that an object will fall towards the ground when it has its supporting surface removed from beneath it. Infants were shown a hand placing a box either on a platform and leaving it there or placing the box beyond the platform and leaving it seemingly floating in mid-air. The impossible event attracted the infants' attention for longer than the possible event.

In summary, before language is learned, infants have developed quite sophisticated knowledge of the spatial world in terms of geometric and dynamic-kinematic routines. But what happens to these conceptual building blocks when language is learned?

3.2 Spatial language acquisition and linguistic relativity

From the pre-linguistic knowledge foundations that the child brings to language learning the child has to acquire the symbol-to-symbol relations in the language and visuo-symbol-to-symbol relations. Given that languages differ in how they carve up the spatial world, one can ask what happens to this pre-linguistic knowledge during language learning.

Comparing English and Korean speakers, Choi and Bowerman (1991) presented evidence to suggest that the first language children are exposed to affects the way in which space is conceptualized and categorized. English expresses path notions (movement into, out of, etc.) in a constituent which is a "satellite" to the main verb (e.g., a particle or preposition), while Korean expresses path in the verb itself. A video cassette is put *in* a video case in English, a lid is put *on* a kettle, a pear is put *in* a bowl, and a glass is put *on* a table. In Korean the verb *kkita* is used for tight-fit path events whereas *nehta* is used for loose-fit containment relationsand *nohta* is used for loosefit support relations. Therefore, Korean carves up the spatial world according to degrees of location control as much as it does according to (geometric) containment and support relations, while English carves up containment and support relations primarily in terms of geometric routines rather than location control. Furthermore, Choi and Bowerman (1991) found evidence that children learning English and Korean respectively extend meanings of terms according to the semantic structure of their input language. English children produced *in* for paths into both tight- and loose-fit containers and extended their use of *in* accordingly, whereas Korean children produced *kkita* for putting objects into tight places, *nehta* for putting objects into loose containers, and extended their use accordingly. Choi, McDonough, Bowerman and Mandler (1999) used a preferential-looking task to assess the generalizations made by children learning either English or Korean. By the age of one-and-ahalf to two years the children in both cases spent more time looking at the language-appropriate aspects of spatial relations. English children looked more at containment scenes than non-containment scenes on hearing *in*, whereas Korean children looked more at tight-fit scenes than loose-fit scenes when hearing *kkita*. However, testing even younger children, McDonough, Choi and Mandler (2003) provide evidence that infants have conceptual readiness for learning location control in either language at an earlier age. Testing 9- to 14-month-old Korean and English infants, they found that both groups categorized both tight and loose containment and tight and loose support.

Overall, the data indicate that, prior to the acquisition of L1, prelinguistic infants have information available regarding both geometric and extra-geometric properties of spatial relations, and when language learning gets under way the language learned actually structures how semantic categories are formed.

4. Second language acquisition

4.1 Predictions from the functional geometric framework

Languages differ in low they carve up the spatial world. Given that one may have acquired a particular language that carves up the world in a particular way, what are the consequences for the acquisition of a second language? Previous analysis allows us to enumerate possibilities regarding how English may be acquired as L2 as follows:

1) A single dimension is used for L2 learning (either geometric or extra-geometric dimensions), but weightings for this dimension from L1 are not used.
2) A single dimension is used for L2 learning and weightings from L1 are used.
3) Both geometric and extra-geometric dimensions are used to acquire L2, but weightings from L1 are not used.
4) Both dimensions are used to acquire L2, and weightings from L1 are used.

With regard to motion events in English, there is some evidence that transfer does occur in L2. Talmy has noted that languages package motion events in different ways. Some languages, including English and most Indo-European languages, have been termed "satellite-framed" languages (Talmy, 1985). These languages express path motions using a constituent that is a satellite to the main verb, such as a preposition (in the case of English). Other languages, such as Spanish and Korean (as discussed above), have been termed "verb-framed" languages in that they express path in the verb itself (and in some of these languages, such as Korean, they lack spatial prepositions completely). So the way that languages carve up motion involves a difference in the extent to which particular constituents express motion. This difference allowed a number of authors to examine whether acquiring a verb-framed or a satellite-framed language is affected by whether L1 has the same or a different typology.

In relation to other spatial expressions there are a number of studies providing evidence for L1 to L2 transfer across a range of languages. Although these studies vary methodologically in ways that affect their direct comparison, as well as the reliability and validity of the conclusions drawn in some cases, they all suggest to some degree that language transfer occurs.

Overall, there is some preliminary evidence that transfer does occur in the case of spatial language, but what about the more fine-grained predictions made above? One issue in the first and second language learning literature is the extent to which learners use single cues versus multiple cues when learning a language. There is evidence that language learners are selective in the parameters they focus on, at least in the early stages of L2 acquisition. For example, in artificial language learning, it has been shown learners often focus on one cue at a time. Consistent with this, in first language acquisition, Richards, Coventry and Clibbens (2004) found that children modified their spatial descriptions more in response to location control changes than to changes in geometry when describing scenes showing containment and support relations in L1 English. Of course, it does not follow that second language learners focus on the same cues when acquiring L2. As we reviewed earlier in the chapter, a natural history of research on spatial language shows that a focus on the importance of extra-geometric relations underpinning the comprehension and production of spatial terms has been a relatively recent development. It is possible that native L1 speakers also focus on geometric relations as a main cue to learning distinctions between spatial terms in L2.

To recap, from existing data in the SLA literature we might expect that learners acquiring prepositions in L2 Spanish or English would be likely to focus on minimal cues, either geometric or extra-geometric, and to become sensitive to the subtle interplay between elements of the functional geometric framework only later in acquisition.

4.2 Further implication and conclusions

The model we have outlined takes us to some more important methodological issues regarding second language acquisition. According to the Coventry el al. (2005) model, language learning involves acquiring symbol–symbol relations and symbol–visuo-symbol relations. Yet much testing of second language ability involves using language alone rather than testing how spatial language co-varies with the spatial world. Following the preferential looking work we have briefly reviewed, how language drives attention is a critical part of being a competent speaker of a language in the spatial domain. The use of *in* and *on* in English involves more than knowing which nouns prepositions can co-occur with—these terms should drive how spatial relations are conceptualized in the scene being described. A consequence of this is that when presented with a picture, one might expect that English speakers might misremember the shape of the reference object in line with the spatial preposition paired to conceptualize the spatial scene. So if the sentence presented with scene (a) is *The bird is in the dish*, one might expect that (b) would be more likely to be false-alarmed to than (c) on an old-versus-new recognition task. In contrast, we might expect that (c) would be more likely to be false-alarmed to than (b) if *The bird is on the dish* was presented with picture (a). Performance on such tasks may be revealing regarding the extent to which the L2 learner has truly grasped language.

So we have suggested that second language learning research would do well to separate out knowledge of symbol–symbol relations and knowledge of symbol–visuo-symbol relations when considering competence for spatial language. Furthermore, using existing computational models as a means of testing and developing theories of Second Language Acquisition offers much potential for fruitful development in the SLA field in the coming years. Already computational models for spatial language have been shown to operate using the same architecture across a range of languages. For example, Regier's (1996) constrained connectionist model for spatial language has been trained on a range of languages from different language families, including Japanese, Mixtec and Russian as well as English. Using the end states of these training sets to then train on a second language would be an exciting avenue to explore.

A final point to note concerns the current disparity between first and second language learning. First language acquisition clearly involves learning symbol–symbol and symbol–visuo-symbol relations in tandem. Furthermore, there is evidence that learning language with associated images/imagery enhances children's reading comprehension. However, second language acquisition rarely presents images with language as this knowledge is assumed to be given in first language acquisition and is presented in the form of language. It would be worthwhile to examine whether grounding language during second language learning affects the success with which one acquires that second language. Moreover, distinctions between prepositions in language learning guides frequently map spatial prepositions in L2 onto the geometric equivalents in L1. Sensitivity to both geometric and extrageometric constraints in language teaching may help direct the attention of the second language learner in ways that ground spatial language more directly in line with narrative goals and events.

5. Conclusions.

The use of experimental techniques varying the spatial world and measuring linguistic responses to changes in spatial array has produced a comprehensive picture of how spatial prepositions are comprehended and produced in English (and in some other languages). The acquisition of spatial prepositions has been shown to be underpinned by pre-linguistic knowledge of geometric and dynamic-kinematic routines consistent with the functional geometric framework and the acquisition of spatial language across languages is sensitive to components of the framework. Such fine-grained analyses afford fine-grained predictions regarding second language acquisition. We hope that forthcoming experimental work will allow us to establish whether particular constraints dominate the acquisition of spatial prepositions in a second language. The offered model represents the various and extended semantic features associated with the spatial prepositions in both English and Spanish. We believe that this cognitive model gives a more systematic account of the semantics of English prepositions than other more traditional accounts, therefore reducing the necessity for rote item learning on Second Language Acquisition. Moreover, we believe that such an analysis has real implications for the classroom context of the different semantic interpretations of each of the spatial prepositions in English.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22(4)*, 577–660.

Carlson-Radvansky, L. A., Covey, E. S., & Lattanzi, K. M. (1999). “What” effects on “where”: Functional influences on spatial relations. *Psychological Science*, *10*, 516–521.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition*, *41*, 83–121.

Choi, S., McDonough, L., Bowerman, M., & Mandler, J. (1999). Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development*, *14*, 241–268.

Coventry, K. R., & Garrod, S. C. (2004). *Saying, Seeing and Acting. The Psychological Semantics of Spatial Prepositions*. Essays in Cognitive Psychology Series. Hove and New York: Psychology Press.

Coventry, K. R., & Prat-Sala, M. (2001). Object-specific function, geometry and the comprehension of “in” and “on.” *European Journal of Cognitive Psychology*, *13(4)*, 509–528.

Coventry, K. R., Carmichael, R., & Garrod, S. C. (1994). Spatial prepositions, object-specific function and task requirements. *Journal of Semantics*, *11*, 289–309.

Garrod, S., Ferrier, G., & Campbell, S. (1999). *In* and *on*: investigating the functional geometry of spatial prepositions. *Cognition*, *72*, 167–189.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20(1)*, 1–55.

Hespos, S. J., & Baillargeon, R. (2001). Reasoning about containment events in very young infants. *Cognition*, *78*, 207–245.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, *43*, 508–529.

Kim, I. K., & Spelke, E. S. (1992). Infants’ sensitivity to effects of gravity on visible object motion. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 385–393.

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-monthold infants. *Cognition*, *47*, 121–148.

Quinn, P. C. (1994). The categorization of above and below spatial relations by young infants. *Child Development*, *65*, 58–69.

Regier, T. (1996). *The human semantic potential: spatial language and constrained connectionism*. Cambridge, MA: MIT Press.

Richards, L. V., Coventry, K. R., & Clibbens, J. (2004). Where’s the orange? Geometric and extra-geometric factors in English children’s talk of spatial locations. *Journal of Child Language*, *31*, 153–175.

Talmy, L. (1985). Lexicalisation patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description*, Volume 3, pp. 36–149. Cambridge: Cambridge University Press.

Zwaan, R. A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation*, Vol. 44 (pp. 35–62). New York: Academic Press.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting and understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1–11.