

Coventry, K. R. (2013). On the mapping between spatial language and the vision and action systems. In Coello, Y., & A. Bartolo, A. (Eds.), *Language and Action in Cognitive Neuroscience (Chapter 11, pp. 209-225)*. Psychology Press.

## **On the Mapping between Spatial Language and the Vision and Action Systems**

Kenny R. Coventry

Cognition and Communication Research Centre

School of Life Sciences,

Northumbria University

### **1. Introduction**

Language learning does not take place in a bubble. Exposure to language co-occurs with other (non-linguistic) input, such as exploration of the world visual, in interaction with a caregiver who plays an active role in the manipulation of joint attention and joint action (Tomasello, 2003). The consequences of this for language learning are that the child is likely to acquire a number of associative links during learning. First the child learns that words co-occur with other words – that is, that the symbols in a language co-occur with other symbols in systematic ways. It has been shown that such relations can be powerful predictors of understanding under a range of circumstances (see for example Landauer & Dumais, 1997). Second, the child learns that words co-occur with non-linguistic information. That is, words become associated with perceptual information, such that *hamster* is associated with a certain shape, a furry feeling, and so on. Taken together, it seems likely that word-to-word (symbol-to-symbol) relations and word to percept (what I will call in shorthand, symbol-to-visuosymbol relations) might provide a more powerful means of accounting for understanding than either set of relations alone (see Andrews, Vigliocco & Vinson, 2009, for modelling and discussion). However, there is a third set of learned relations. The child learns that certain types of perceptual information occur together. For example, bottles and glasses are frequently found together, are often associated with a pouring event, and such a pouring

event affords relieving oneself from a state of thirst. It can be argued that these three sets of interconnected relations are necessary to account for meaning/understanding, and as I show below, these interconnections are important for an understanding of how spatial language is comprehended across different situations and tasks. My goal here is not to elaborate on how these interconnections are learned, nor to discuss how they constitute understanding (see Coventry & Taylor, in preparation), but rather to show how these sets of relations are important for understanding two types of spatial language: spatial prepositions and spatial demonstratives.

## **2. Spatial Language in Context**

Spatial language would appear to be an ideal candidate to start to examine how symbol-to-symbol and symbol-visuosymbol relations constrain meaning. Talking about the spatial world usually involves talking about concrete arrangements of objects, such as furniture in a room, objects in a picture, crosses and circles in a diagram, and so forth. And such language often occurs in situations where infants are manipulating objects whilst interacting with caregivers. This is exactly the type of situation where one might expect close relationships between words and non-linguistic representations of the world to be formed. At the same time spatial language exhibits considerable cross-linguistic variability, making it challenging to learn, particularly in a second language (see for example Coventry, Guijarro-Fuentes, & Valdés, 2011). So learning symbol-to-symbol relations is also important to capture regularities in how spatial terms and nouns, for example, go together in a language.

Two classes of spatial words that are among the first terms to be acquired across languages are spatial prepositions – words such as *in*, *on*, *under*, and so on – and spatial demonstratives – words such as *this* and *that*. These terms are high frequency within a language, and notably also occur in a range of other (non-spatial) guises. For example, prepositions and

demonstratives can be used temporally (*This decade and that decade, See you in a minute, The movie starts at 8pm*), metaphorically (*I'm over the moon, I'm under the weather, Do it like that*), and occur in more grammatical guises also (*That's why I like music, For the last time*). As such, they constitute important classes of terms to be understood. We consider each of them in turn.

## 2.1 Spatial Prepositions

Spatial prepositions can be classified into three basic types: the so-called topological prepositions (*in* and *on* will be our focus here), projective prepositions (e.g., *over, below, in front of*, etc.) and proximity terms (e.g. *near, beside*, etc.). Approaches to spatial language in linguistics typically employ linguistic glosses and/or pictures of spatial relations (e.g. image schemata), and these are taken to represent words without elucidating the vision and action processes underlying them. Coventry and Garrod (2004) argue that the starting point for understanding spatial language is to examine how language and the spatial world covary. Systematically manipulating the visual world and then examining how language maps on to the world is revealing regarding the vision and action parameters that are important for that mapping.

According to the 'functional geometric framework' proposed by Coventry and Garrod (2004), the comprehension and production of these terms involves three sets of interlocking parameters. These are geometric routines, dynamic-kinematic routines, and object/situational knowledge. We take each of these in turn, focusing on *in, on, over, under, above, and below*. First, spatial prepositions are all associated with geometric routines. Where an object to be located (the located object, LO) is positioned in 2 or 3 dimensional space with respect to a reference object (RO) affects which preposition is appropriate to describe where that object is. This comes as no surprise – spatial language clearly maps onto space – but perhaps what is

more of a surprise is how it does so. Two examples of geometric routines will make this clear.

First, Coventry and Garrod (2004) appeal to concepts from Cohn's region connection calculus (RCC; Cohn, 1996; Cohn, Bennett, Gooday & Gotts, 1997; Cohn, Randell & Cui, 1995) as underlying the geometric relations associated with *in* and *on*. RCC is a qualitative geometry that characterises containment and enclosure in terms of two primitive relations: connection and convexity. It is beyond the scope of this chapter to discuss how these relations are computed (see Coventry & Garrod, 2004 for discussion). The important point here is that these primitive relations afford a considerable degree of flexibility in their application, providing an elegant means to account for different degrees of enclosure – from containable insides (as in *coffee in a cup*), to graded containment relations (as in *an apple in a bowl* when it is positioned above the rim of a bowl on top of other fruit in a bowl) to scattered insides (such as *an island in an archipelago*). Moreover, although Cohn and colleagues make no claims regarding the neural mechanisms involved in their computation, Coventry and Garrod (2004) suggest that the visual routines of the type proposed by Ullman (1996) afford flexibility of computation allowing a small number of basic processes to account for what are often thought of as distinct geometric relations (and distinct senses of *in* and *on* in linguistic accounts; see for example Herskovits, 1986).

Second, *above* is associated with the weighted attention directly from a reference object to a located object. Originally Logan and Sadler (1996) and Hayward and Tarr (1995) noted that *above* seems to be most appropriate when the located object is directly above the reference object in the vertical plane, aligned with the reference object, and a spatial template was proposed to map onto spatial prepositions. Since then, the Attention Vector Sum (AVS) model (Regier & Carlson, 2001) shows how this basic idea can be grounded more appropriately in a computational mechanism that takes its inspiration from population vector

encoding in neural subsystems (Georgopolous, Schwartz, & Kettner, 1986). Like RCC for spatial relations, AVS is sensitive to the dimensions of the reference and located objects, and affords a grounded on-line mechanism for computing goodness of fit of *above* to spatial scenes.

Geometric relations on their own do not account for the comprehension and production of spatial prepositions adequately. In addition to geometric routines, dynamic-kinematic routines are important. Again we can consider *in* and *on* and *above* in turn. *In* and *on* seem to require in many cases that the reference object controls the location of the located object. This location control information goes beyond geometry – it requires knowledge of dynamics and kinematics operating within the gravitational world. Some empirical demonstrations show this to be the case. Garrod, Ferrier and Campbell (1999) presented participants with images of bowls and plates, and in some of these images objects were arranged in a pile on top and in contact with these reference objects. They varied where the located object (always a ball) was in relation to the reference object (e.g. on top of piles of objects of varying heights). However, crossed with this geometric manipulation was the manipulation of location control. In the ‘weak location control condition’ a string suspended from above (connected to a wooden frame) was attached to the located object. The motivation was that the string was the controller of the ball, not the reference object, so that when the bowl or plate was moved, the ball would not move with it. In contrast, when there was no string connected to the ball, movement of the reference object would result in the located object moving also (‘strong location control condition’). Garrod et al. asked one group of participants to rate the likelihood that the ball and bowl or plate would remain in the same relative positions should the bowl or plate be moved (an independent measure of the degree of location control) while a second (different) group of participants rated sentences of the form *The ball is in/on the bowl/plate* to describe the same images. Garrod et al. found

striking correlations between judgements of the degree of location control and ratings for *in* and *on* to describe the same scenes.

Similar results have been found with more naturalistic spatial language production in both adults and children, showing videos of fruit in various positions in relation to containers and supporting surfaces (Coventry, 1998; Richards, Coventry & Clibbens, 2004). For instance descriptions of where an apple was when it was shown positioned on top of a pile of fruit high above the rim of a bowl in a static image was compared to descriptions for the same apple either when the whole scene (apple plus other fruit plus bowl) was shown moving together from side to side in a video (the strong location control condition) or when the apple was shown moving from side to side on its own, but still in contact with the fruit below (the weak location control condition). *In* and *on* were produced more as primary descriptions of the location of the apple in relation to the bowl in the strong location control condition than in the other conditions and used least in the weak location control condition.

Analogous examples are found for *over*, *under*, *above* and *below*. In one experiment Coventry et al. (2010) presented participants with static images of a person holding an object with a protecting function, such as an umbrella, with falling objects shown a distance away from the object (e.g. rain). In an earlier study, Coventry, Prat-Sala and Richards (2001) found that whether the falling objects were shown hitting or missing the protecting object (and hitting the person holding the object instead) affected judgements of the extent to which *The umbrella is over the man* was regarded as a good description. Similar results were reported for containers, such as bottles and jugs, shown pouring various objects either missing or entering a second (recipient) container. Coventry et al. (2010) show that when the end destination of falling objects is not actually shown, judgements are nevertheless affected by the percentage of falling object that are judged to make contact with the person holding the protecting object. Moreover, this was supporting in an eye tracking study. When static images

of objects were shown with falling objects beginning to fall from the mouth of the object (e.g. a bottle beginning to pour liquid with a second container below and to the left of the bottle), eye movement data showed that participants looked at the end point of the falling objects before they returned their language judgements. One way of thinking about this is that people mentally ‘animate’ the visual scene, mapping what happens during such animation onto past experiences of pouring events where one would be more likely to describe the bottle as over the glass when the liquid ends up in the glass.

The third set of relations required for spatial language is object/situational knowledge. The shape of a reference object, for example, is not a perfect predictor of the extent to which it is regarded as a supporting surface or container. This information comes from how symbols co-occur within the language. Plates are regarded as supporting surfaces, although they do have convex hulls that afford a certain degree of containment. The same object can be labelled a *plate* or a *dish*; in the first case a support relation is appropriate but in the second case a containment relation is appropriate (Coventry, Carmichael & Garrod, 1993; Coventry & Prat-Sala, 2001). Additionally, building of situational knowledge is important. This goes beyond just symbol-to-symbol and symbol-to-visuosymbol relations – visuosymbol-to-visuosymbol relations are also needed. For example, we know that solid objects afford protection from the elements, although they may never have been seen in that context. A suitcase can protect someone from getting wet, but a sieve is not so good. Coventry et al. (2001) substituted solid objects that do not have a protecting function, such as a suitcase, for the umbrella and other protecting objects described earlier. When a suitcase is shown stopping rain hitting a person, acceptability ratings for *over*, *under*, *above* and *below* were higher than when the suitcase was shown not protecting the person. Presumably the rich information about how objects with certain types of properties can be used in various ways must be learned through experience with those objects.

I have necessarily been selective in the presentation of empirical evidence supporting the three components of the functional geometric framework. There are many other examples both within English (see for example Carlson-Radvansky, Covey & Lattanzi, 1999; Carlson-Radvansky & Radvansky, 1996; Carlson-Radvansky & Tang, 2000; Miller & Carlson, this volume, for other examples) and across languages (Feist, 2008) – and see Coventry & Garrod, 2004 for a comprehensive review. The point is that the evidence for the three components is quite robust. However, it is important to note that the three components do not seem to apply all of the time. For example, a cross above a circle doesn't seem to involve any obvious dynamic kinematic information, while location control seems to be critical for three dimensional objects in the world as it is through gravity that containers and supporting surfaces possess control properties. Rather than argue for the primacy of one set of relations compared to another (see Coventry & Garrod for discussion; see also Miller & Carlson, this volume) - the way in which symbol-to-symbol relations, symbol-to-visuosymbol relations and visuosymbol –to-visuosymbol relations become (temporally) bound together during learning may account for these differences. Prepositions become bound together with the visuosymbolic information that co-occurs with them and with the nouns that co-occur with those prepositions. Bottles and glasses become associated in learning with a pouring routine which co-occurs most frequently with prepositions such as *over* and *above*. With crosses and circles, the absence of such learning co-occurrence information means that geometric routines are used alone as positional information is the information that differentiates relative locations of crosses and circles. However, quite readily one can imagine that learning that the cross and circle co-occur in only certain ways may induce a more functional reading (for example, that the cross must be aligned with the circle, such that the circle will control where it is).

One can ask what sort of model might be able to account for how these mutual constraints operate on spatial language comprehension and production, with the flexibility required to do so. There have been a number of recent attempts to integrate both geometric and extra-geometric information within a single model (see for example Carlson, Regier, Lopez & Corrigan, 2006; Coventry, Cangelosi, Rajapakse, Bacon, Newstead, Joyce, & Richards, 2005; Lockwood, Forbus, Halstead, & Usher, 2006). The model of Coventry et al. (2005) was an attempt to model the three components of the functional geometric framework directly, employing cognitive-functional constraints through the extension of Ullman-type visual routines acting on dynamic visual input. The videos input to the model comprised films of containers pouring liquids into other containers of the sort we have considered above. Using a “what + where” code (see Joyce et al., 2003) the separate objects in each visual scene are identified (e.g., reference object, located object and liquid), and comprise an array of activations recording the visual stimuli in each area of the visual field. This output is then fed into a predictive, time-delay connectionist network akin to an Elman simple recurrent network (Elman, 1990). In essence, the network learns how liquids pour, etc., allowing the network to then predict the path of the pouring liquid when a static visual scene is presented. The model provides a mechanism for implementing perceptual symbols (see Joyce et al., 2003), and can “replay” the properties of the visual episode that was learned. So inputting a still image with implied motion following training with dynamic images, the model is able to establish where falling objects end up. The outputs of this predictive network feed further into a dual-route (vision and language) feed forward neural network to produce a judgement regarding the appropriate spatial terms describing the visual scene. Thus when seeing a static image with implied motion, the model is able to run a simulation of interaction between objects, based on the mapping between the objects shown in the static scene and past observed interactions with those objects. These simulation results feed forward to language

judgements, mirroring the results presented here and elsewhere. In addition, the model learns symbol-to-visuosymbol relations, and symbol-to-symbol relations through combined language and visual input.

While the three components of the functional geometric framework and associated models contain information regarding how objects interact with each other, what has not been investigated thus far is how and whether interaction with objects affects spatial language comprehension and production. It is possible that a dynamic-kinematic simulation of pouring involves a motor component where the hand causes the pouring. This would be consistent with work showing motor activations for verbs (see Taylor & Zwaan, this volume). However, I now turn to consider a second class of spatial language – spatial demonstratives – where the importance of interaction with objects has been established.

## 2.2 Spatial demonstratives

Spatial demonstratives are particularly important in language as they provide one of the most obvious direction connections to joint attention, and to the action system. It has been noted that the use of *this* and *that* in early development usually occur with deictic pointing, and this observation is reinforced with the knowledge that the use of such terms in some languages cannot occur without pointing (Senft, 2004). Diesel (2005), in a landmark survey of the demonstrative systems of over 200 languages, notes that the most common distinction demonstrative systems make across language (present in about half the languages sampled) is a binary proximal/distal contrast (as in English). Less common, but nevertheless prevalent across languages is a three term demonstrative system (just over one third of languages sampled) which is either distance based (e.g. Spanish) or person oriented (e.g. Japanese). Beyond these distinctions, some languages make a variety of additional distinctions, such as

contrasts based on whether an object is visible or not (e.g. Tiriyo) and whether an object is owned or not (Supyire).

In spite of the intuition that spatial demonstratives appear to be about distance, their use seems at first sight to confound any straightforward mapping between these terms and a basic (non-linguistic) distance-based distinction. *This* and *that* can be used to denote distances far removed from the speaker, such as *this universe* and *that universe*, and *this* and *that* can both be used in peripersonal space, as in *this finger* and *that finger*. Moreover, the fact that languages lexicalise distinctions other than a distance based near versus far distinction is also suggestive that a distance-based distinction may not be basic (Kemmerer, 1999, 2006).

Coventry, Valdés, Castillo and Guijarro-Fuentes (2008) set out to test whether there is a relationship between demonstrative use and a basic distal distinction between near and far space. There is robust experimental/neuropsychological evidence to support a distinction between two separate brain systems that represent near/peripersonal and far/extrapersonal perceptual space (see Berti & Rizzolatti, 2002; Làdavas, 2002; Legrand, Brozoli, Rossetti, & Farnè, 2007 for reviews). This dissociation comes from studies on nonhuman primates (e.g. Iriki, Tanaka, & Iwamura, 1996), neuropsychological studies with patients who exhibit visual neglect in near space but not in far space (e.g., Brain, 1941; Cowey, Small, & Ellis, 1994; Halligan & Marshall, 1991), and from experimental studies with healthy participants (e.g. Bjoertomt, Cowey, & Walsh, 2002; Gamberini, Seraglia, & Priftis, 2008). Moreover, work with neglect patients has shown that peripersonal space is extendable if one uses a tool. For example, Berti and Frassinetti (2000) tested a patient who showed a dissociation between near and far space in the manifestation of neglect. Using a laser pointer, the patient performed poorly on a line bisection task when the line was in near space but performed much better when the line was in far space. In contrast, when the patient performed the same task using a stick rather than a laser pointer performance on the task in far space became much worse,

resembling performance in near space (see also Longo & Lourenco, 2006; Pegna et al., 2001). It would appear that near space can be extended with the use of a tool, and that peripersonal space representation involves multisensory inputs where the body is represented in action at a functional level (Coello & Delevoye-Turrell, 2007; Farnè, Iriki, & Làdavas, 2005; Holmes & Spence, 2004; Làdavas, 2002; Legrand, et al., 2007).

Coventry et al. devised a new memory game methodology with which to elicit naturalistic use of spatial demonstratives while also manipulating distance of an object from a speaker, tool use, and who places the object. Participants played a 'memory game' where the goal of the game was to remember the positions of objects (coloured shapes) placed on coloured dots positioned on the midline of a large conference table. Participants were informed that they were taking part in an experiment on the effects of language on memory for object location and that they were in the 'language condition.' They were first told that cards would be read out with a placement instruction (e.g., "You place red square on black dot"). Following placement, participants were instructed that they had to point to each object naming it using a combination of just three words: a demonstrative, a colour and a shape (e.g., *this/that red square*), so that everyone in the 'language condition' experienced the same level of language coding. Coventry et al. manipulated the distance between the participant and placed object, whether participants used their hand or a 70cm stick when pointing at the objects placed, and who placed the object.

The results showed a mapping between the peripersonal-extrapersonal distinction and spatial demonstrative choice. As distance from the body increased, *this* was used less frequently and *that* was used more frequently, with a marked decline in the use of *this* at the boundary between near and far space. Moreover, when participants pointed at the object using a stick, there was a corresponding extension of the use of *this* to describe object location mirroring the results with neglect patients discussed above. Similarly, contacting the

object in near space by placing it also led to an increase in the use of *this* to describe its location compared to the condition where the object was placed by the experimenter. This pattern of results points to a direct mapping between demonstrative choice and perceptual space.

While it has been shown that the peripersonal-extrapersonal distinction is important for demonstrative choice, it is also the case that even within peripersonal space the labelling of an object with (the Italian equivalents) of *this* or *that* affects how participants reach towards those objects. Bonfiglioli, Finocchiaro, Gesierich, Rositani and Vescovi (2009) investigated reach to grasp actions for objects placed either nearer or further from the participant in peripersonal space. Participants were instructed to reach and contact an object referred to in a command (of the form *this/that* object name), but only if the demonstrative and noun exhibited the correct gender agreement (so it was a go/no go task). Reaction times to objects in a near position were faster when (the Italian equivalent of) *this* was used and slower when *that* was used, and conversely reaction times for objects in the far location were faster with *that* and slower with *this*. These data are consistent with other studies showing that congruence of labelling of objects affects reach to grasp actions (Gentilucci, Benuzzi, Bertolani, Daprati, & Gangitano, 2000; Gentilucci and Gangitano, 1998; Glover & Dixon, 2002; Glover, Rosenbaum, Graham & Dixon, 2004).

A number of differences between the studies of Coventry et al. and Bonfiglioli et al. make it difficult to interpret why Coventry et al. found clear differences in the selection of *this* and *that* to describe objects placed in peripersonal versus extrapersonal space while Bonfiglioli et al. (2009) found congruence effects of demonstratives with near/far distances within peripersonal space. One possibility is that the use of two distances alone repetitively across trials in the Bonfiglioli et al. study meant that participants set up a contrastive space akin to the contrastive use of these terms. (It is the case that spatial demonstratives can be

used contrastively within peripersonal or extrapersonal space, as Kemmerer has noted.) Changing the task dimensions may well play a role in the extent to which spatial versus contrastive use of demonstratives occurs.

Empirical work on spatial demonstratives is in its infancy, but clearly action plays a key role in the production and comprehension of this class of spatial language. I next consider possible generalisations across the two spatial language semantic categories.

### **3. Spatial language – what research on demonstratives can learn from research on prepositions and vice versa**

I have considered two classes of spatial language – spatial prepositions and demonstratives – and have shown that both classes of spatial language involve more than just geometric relations underpinning their use. In this section, I ask if the parameters important for spatial demonstratives are also important for spatial prepositions, and vice versa. Underpinning this issue is the possibility that a common set of vision and action components may underpin spatial language generally.

Taking peripersonal versus extrapersonal space as a starting point, one can speculate as to whether such a distinction may also be important for spatial prepositions. One possibility relates to reference frame use with spatial prepositions. The so-called ‘projective prepositions’ - such as *above*, *in front of*, *to the left of* – require a reference frame to be established before they can be used. For example, *the coin is to the left of the monkey* could mean that the coin is located in the region to the left of the monkey from the viewer’s perspective (egocentric left) or on the monkey’s left side which may not necessarily be the same as the viewer’s left. The first example is often referred to as relative reference frame, and the second, intrinsic (Levinson, 1996). Coventry, Andonova and Tenbrink (in preparation) have tested whether choice of reference frames as expressed through choice of

spatial preposition is affected by whether the visual scene to be described was placed in peripersonal or extrapersonal space. Participants described where one object on a card was positioned in relation to a second object (e.g. *Where is the marble with respect to the ladybug?*). It was predicted that the relative frame would be used more (and the intrinsic frame less) when the card was placed in the peripersonal space of the participant than when it was placed in the extrapersonal space of the participant. This was thought to be the case as the relative frame uses a person's point of view, whose peripersonal space is relevant; this is not the case with the intrinsic frame. The results supported this, providing the first evidence that near versus far space affects reference frame choice and consequent spatial preposition selection. As Borghi & Cimatti (2010) have argued, being able to act upon objects directly may have consequences for a range of types of language.

One might expect also that spatial demonstrative comprehension and production might be influenced by components of the functional geometric framework originally applied to prepositions. For example, just as the functional relations between two objects are important determinants of spatial prepositions comprehension and production, the functional relation between the speaker and object may similarly be important. For example, one might expect that *this* would be used more frequently to describe the location of a glass in peripersonal space when the handle of the cup is pointing to the speaker (thus affording interaction) versus when the handle is pointing away from the speaker (making the object more difficult to lift canonically). Just as tool use may extend peripersonal space, so might peripersonal space contract when an objects does not afford interaction, or perhaps when it is occluded, and so on.

There are also likely to be other variables as yet uninvestigated that impact on the comprehension and production of both classes of words. For example, the demonstrative systems of some languages involve contrasting objects owned versus not owned by the

speaker. One can imagine *this* being used more in English also when an object is owned by the speaker as opposed to by their interlocutor. With prepositions too, building on the results of Coventry, Andonova and Tenbrink (under review), one can imagine that the relative frame may also be more likely to be used when the card on which objects locations are to be described is owned by the participant rather than by an interlocutor. Of course these possibilities have not as yet been tested. Future studies would do well to do so.

#### **4. Conclusion**

In summary, I have examined variables that are important for the comprehension and production of two classes of spatial language – spatial prepositions and demonstratives. Far from assuming that these terms are abstract symbols unconnected from vision and action, it is clear that both classes of term exhibit systematic relations with how we perceive and interact with the world. This does not deny the importance of what I have termed symbol-to-symbol relations, but rather underlines the fact that spatial language bears a close correspondence with sensorimotor representations given the very nature of that language.

#### **References**

- Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, *116*(3), 463–498.
- Berti, A., & Frassinetti, F. (2000). When far becomes near: Remapping of space by tool use. *Journal of Cognitive Neuroscience*, *12*, 415-420.
- Berti, A., & Rizzolatti, G. (2002). Coding near and far space. In H. –O. Karnath, A. D. Milner, & G. Vallar (Eds.), *The cognitive and neural bases of spatial neglect* (pp. 119-129). New York: Oxford University Press.
- Bjoertomt, O., Cowey, A., & Walsh, V. (2002). Spatial neglect in near and far space investigated by repetitive transcranial magnetic stimulation. *Brain*, *125*(9), 2012-2022 .
- Bonfiglioli, C., Finocchiaro, C., Gesierich, B., Rositano, F., & Vescovi, M. (2009). A kinematic approach to the conceptual representations of this and that. *Cognition*, *111*, 270-274.
- Borghetti, A., & Cimatti, F. (2010). Embodied cognition and beyond: Acting and sensing the body. *Neuropsychologia*, *43*, 763-773.

- Brain, W. R. (1941). Visual disorientation with special reference to lesions of the right cerebral hemisphere. *Brain*, *64*, 244-272.
- Carlson, L. A, Regier, T., Lopez, B., & Corrigan, B. (2006). Attention unites form and function in spatial language. *Spatial Cognition and Computation*, *6*, 295-308.
- 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.
- Carlson-Radvansky, L. A., & Radvansky, G. A. (1996). The influence of functional relations on spatial term selection. *Psychological Science*, *7*, 56-60.
- Carlson-Radvansky, L. A., & Tang, Z. (2000). Functional influences on orienting a reference frame. *Memory and Cognition*, *28*(5), 812-820.
- Coello Y, & Delevoeye-Turrell Y (2007). Embodiment, spatial categorization and action. *Consciousness and Cognition*, *16*, 667–683.
- Cohn, A. G. (1996). *Calculi for qualitative spatial reasoning*. Paper presented at the Proceedings of AISMC-3, LNCS.
- Cohn, A. G., Bennett, B., Gooday, J., & Gotts, N. M. (1997). Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica*, *1*(3), 1-42.
- Cohn, A. G., Randell, D. A., & Cui, Z. (1995). Taxonomies of logically defined qualitative spatial relations. *International Journal of Human Computer Studies*, *43*, 831-846.
- Coventry, K. R. (1998). Spatial prepositions, functional relations and lexical specification. In P. Olivier, & K. Gapp (Eds.), *The Representation and Processing of Spatial Expressions* (pp. 247-262). Lawrence Erlbaum Associates.
- Coventry, K. R., Andonova, E., & Tenbrink, T. (in preparation). Primed by what we see. Spatial reference frame use in language and visual context.
- Coventry, K. R., Cangelosi, A., Rajapakse, R., Bacon, A., Newstead, S. N., Joyce, D., & Richards, L. (2005). Spatial prepositions and vague quantifiers: Implementing the functional geometric framework. In C. Freksa, B. Knauff, B. Krieg-Bruckner & B. Nebel (Eds.), *Spatial Cognition IV. Reasoning, action and interaction*, pp.98-110. Lecture Notes in Computer Science: Springer-Verlag.
- Coventry, K. R., Carmichael, R., & Garrod, S. C. (1994). Spatial prepositions, object-specific function and task requirements. *Journal of Semantics*, *11*, 289-309.
- Coventry, K. R., & Garrod, S. C. (2004). *Seeing, Saying and Acting. The Psychological Semantics of Spatial Prepositions*. Taylor Francis, Psychology Press: Hove and New York.

- Coventry, K. R., Guijarro-Fuentes, P., & Valdés, B. (2011). Spatial language and second language acquisition. In V. Cook & B. Bassetti (Eds.), *Language and Bilingual Cognition* (pp. 262-286). Taylor Francis: Psychology Press.
- Coventry, K. R., Lynott, D., Cangelosi, A., Monrouxe, L., Joyce, D., & Richardson, D. C. (2010). Spatial language, visual attention, and perceptual simulation. *Brain & Language, 112*(3), 202-213.
- 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., Prat-Sala, M., & Richards, L. (2001). The interplay between geometry and function in the comprehension of 'over', 'under', 'above' and 'below'. *Journal of Memory and Language, 44*, 376-398.
- Coventry, K. R., & Taylor, L. (in preparation). Words, percepts, situations and learning: the Sym-Sym-Vis-Vis framework. Manuscript in preparation.
- Coventry, K. R., Valdés, B., Castillo, A., & Guijarro-Fuentes, P. (2008). Language within your reach. Near-far perceptual space and spatial demonstratives. *Cognition, 108*, 889-895.
- Cowey, A., Small, M., & Ellis, S. (1994). Visuospatial neglect can be worse in far than in near space. *Neuropsychologia, 32*, 1059-1066.
- Diessel, H. (2005). Distance contrasts in demonstratives. In Haspelmath, Dryer, Gil, & Comrie (Eds.), *World atlas of language structures* (pp. 170-173). Oxford: Oxford University Press.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.
- Farnè, A., Iriki, A., & Làdavas, E. (2005). Shaping multisensory action-space with tools: Evidence from patients with cross-modal extinction. *Neuropsychologia, 43*, 238-248.
- Feist, M. I. (2008). Space between languages. *Cognitive Science, 32*(7), 1177-1199.
- Gamberini, L., Seraglia, B., & Priftis, K. (2008). Processing of peripersonal and extrapersonal space using tools: Evidence from visual line bisection in real and virtual environments. *Neuropsychologia, 46*(5), 1298-1304.
- Garrod, S., Ferrier, G., & Campbell, S. (1999). *In and on*: Investigating the functional geometry of spatial prepositions. *Cognition, 72*, 167-189.
- Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research, 133*, 468-490.
- Gentilucci, M., & Gangitano, M. (1998). Influence of automatic word reading on motor control. *European Journal of Neuroscience, 10*, 752-756.

- Georgopolous, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, 223, 1416-1419.
- Glover, S., & Dixon, P. (2002). Semantics affect the planning but not control of grasping. *Experimental Brain Research*, 146, 383–387.
- Glover, S., Rosenbaum, D. A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research*, 154, 103–108.
- Halligan, P. W., & Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature*, 350, 498 - 500
- Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, 55, 39-84.
- Herskovits, A. (1986). *Language and Spatial Cognition. An interdisciplinary study of the prepositions in English*. Cambridge University Press.
- Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation(s) of peripersonal space. *Cognitive Processing*, 5, 94–105.
- Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, 7, 2325–2330.
- Joyce, D. W., Richards, L. V., Cangelosi, A., & Coventry, K. R. (2003). On the foundations of perceptual symbol systems: Specifying embodied representations via connectionism. *The Fifth International Conference on Cognitive Modelling*. Bamberg, Germany.
- Kemmerer, D. (1999) “Near” and “far” in language and perception. *Cognition*, 73, 35-63.
- Kemmerer, D. (2006). The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia*, 44, 1607-1621.
- Làdavas, E. (2002). Functional and dynamic properties of visual peripersonal space. *Trends in Cognitive Science*, 6, 17–22.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato’s problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211-240.
- Legrand, D., Brozzoli, C., Rossetti, Y., & Farnè, A. (2007). Close to me: Multisensory space representations for action and pre-reflexive consciousness of oneself-in-the-world. *Consciousness and Cognition*, 16, 687–699.
- Levinson, S. C. (1996). Frames of reference and Molyneux’s question. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 109-169). Cambridge, MA: MIT Press.
- Lockwood, K., Forbus, K., Halstead, D. & Usher, J. (2006). Automatic Categorization of Spatial Prepositions. *Proceedings of the 28th Annual Conference of the Cognitive Science Society. Vancouver, Canada*.

- Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and Space* (pp. 493-530). Cambridge, Mass.: MIT Press.
- Longo, M. R., & Lourenco, S. F. (2006). On the nature of near space: Effects of tool use and the transition to far space. *Neuropsychologia*, *44*, 977-981.
- Pegna, A. J., Petit, L., Caldara-Schnetzer, A. –S., Khateb, A., Annoni, J. –M., Sztajzel, R. et al. (2001). So near yet so far: Neglect in far or near space depends on tool use. *Annals of Neurology*, *50*, 820-822.
- Regier, T., & Carlson, L.A. (2001) Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General*, *130*(2), 273-298.
- Richards, L. V., Coventry, K. R., & Clibbens (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.
- Senft, G. (2004). *Deixis and Demonstratives in Oceanic Languages*. Canberra: Pacific Linguistics.
- Tomasello, M. (2003). *Constructing a language: A Usage-based Theory of Language Acquisition*. Harvard University Press.
- Ullman, S. (1996). *High-level vision: Object recognition and visual cognition*. Cambridge MA: MIT Press.