A framework for work on frames of reference

Anna Shusterman and Peggy Li

Humans, like all navigating animals, need to think about where they are in the world. What mental capacities underlie our spatial knowledge and behavior? In this chapter, we review a central organizing concept in theoretical and empirical work on spatial cognition, that of *frames of reference*. Frames of reference refer to coordinate axes that can describe a space and the spatial relationships within it. Different researchers have proposed different ways of thinking about frames of reference, leading to different typologies drawing on distinct bases of empirical work. We suggest that studies of children's cognition and language acquisition can provide insight into these questions, in two ways: first, by shedding light on the developmental starting points of frame-of-reference language concepts, and second, by revealing points of developmental change and divergence between cultures and languages. Drawing on this work, we propose that the development of spatial language and concepts, on the one hand, and the coordination of linguistic communities, on the other, present unique constraints. We argue that these distinct pressures lead to cross-cutting distinctions in categorizing spatial frames of reference.

COORDINATE AXES

In essence, frame-of-reference concepts are complex, abstract mental structures. A frame of reference (FoR) is a coordinate framework that organizes a set of spatial relations. This coordinate framework can be derived from any entity or set of entities in the world onto which axes may be imposed. One common framework divides frames of reference into egocentric and allocentric, or those that are from one's own perspective versus that of another entity. However, this is quite a broad cut: the term *allocentric* often refers to non-egocentric viewpoints (e.g., port side of a ship), viewpoint-independent reference frames (e.g., bird's-eye view), cardinal directions (e.g., north), and sometimes to another *person's* perspective. Similarly, a common distinction between viewer-based and place-based representations (Levinson, 1996) conflates across different kinds of viewers – self and other – but it is not clear that these frames are represented the same way in the mind (Newcombe & Huttenlocher, 2000; Schober, 2009; Duran, Dale, & Kreuz, 2010). Clearly, then, an adequate typology of reference frames is going to require finer distinctions.

We find it useful to make the first cut by distinguishing two kinds of entities that can define the axes of the coordinate framework. In an *object-based frame of reference*, these entities can be things that move relative to the earth, such as a person, or a fronted object. In an *environment-based reference frame*, the coordinate axes stem from stationary entities like the earth or smaller sub-environments.

The *egocentric* FoR is really just one case of an object-centered reference frame, because the axes are defined over an object – namely, one's own body. The egocentric FoR may hold a privileged status in cognition because it is the perspective from which one takes in information about the world and plans one's own movements through space.

In *non-egocentric* reference frames, the expression changes its meaning depending on the orientation of an object other than the speaker. As the name implies, these reference frames are neither egocentric nor environment-based. They are one kind of allocentric frame.

Finally, *environment-based FoRs*, such as rooms, buildings, and local terrain, are also allocentric. However, they are defined over stable areas fixed to the earth rather than moveable objects, and they provide axes such as *window-side/wall-side*, *front/back*, and *uphill/downhill*. In sum, we conceptualize three distinct categories of frames of reference: egocentric object-based, non-egocentric object-based, and environment-based. In the rest of this chapter, we assess the psychological status of frames of reference. How are these reference frames represented? What spatial computations do they support? How do they map on to language?

Table 1. Ways of distinguishing frame of reference. The egocentric (vs. allocentric) distinction captures the perspective by which we take in information about the world and program our motor movements. The object-based vs. environment-based distinction captures whether the frame of reference is anchored to earth, the space in which we live.

EGOCENTRIC	**ALLOCENTRIC (NON-EGOCENTRIC)**	
OBJECT-BASED		**ENVIRONMENT-BASED**
Front-back *at my back* *in front of me**	Front-back *the back of the doll* *in front of the chair**	Principle axis *uphill-downhill* *sunrise-sunset*
Left-right *my left hand* *the cup to my right*	Left-right *Richard's right foot* *stage left*	Derived axis *Crosshill*

SCOPE AND ORIGIN

In addition to the entities that define the coordinate axes, spatial relations require two other considerations: *scope* and *coordinate origin*. First, FoRs vary in their scope – the scale of space to which they apply. In principle, one can refer to locations inside or outside of the entities defining the frame, covering regions of space varying dramatically in size. For instance, terms like *left* and *right* can be used to refer to the left and right sides of a person's body, limiting the scope to the body. They can also be used to refer to other objects, not only immediately off to the left and right sides of the person (e.g., the box to the person's left), but also of locations miles away from the person; for example, for a person standing facing north in New York, California is to the person's left (see Figure 1a). In practice, the scope is often restricted for various reasons. For instance, it is difficult to think about and extend a frame of reference defined by a person in New York to talk about locations and relations of entities in California; it is also typically not useful or necessary to think about such relations.

The scope covered by FoRs in the world's languages is often additionally influenced by the conventions of one's linguistic community. For example, Brown and Levinson (1993) describe Tseltal Mayans who refer to spatial locations with the terms meaning *uphill* and *downhill*, derived from the terrain of the local environment. The *uphill-downhill* terms, derived from the important hill on which they reside, are also used to refer to entities off and beyond the hill. In contrast, terms like *uptown* and *downtown* for Manhattan are restricted in scope to the island itself even though in principle the terms could be extended to talk about directions and locations off and beyond the island (Levinson, Kita, Haun, & Rasch, 2002).

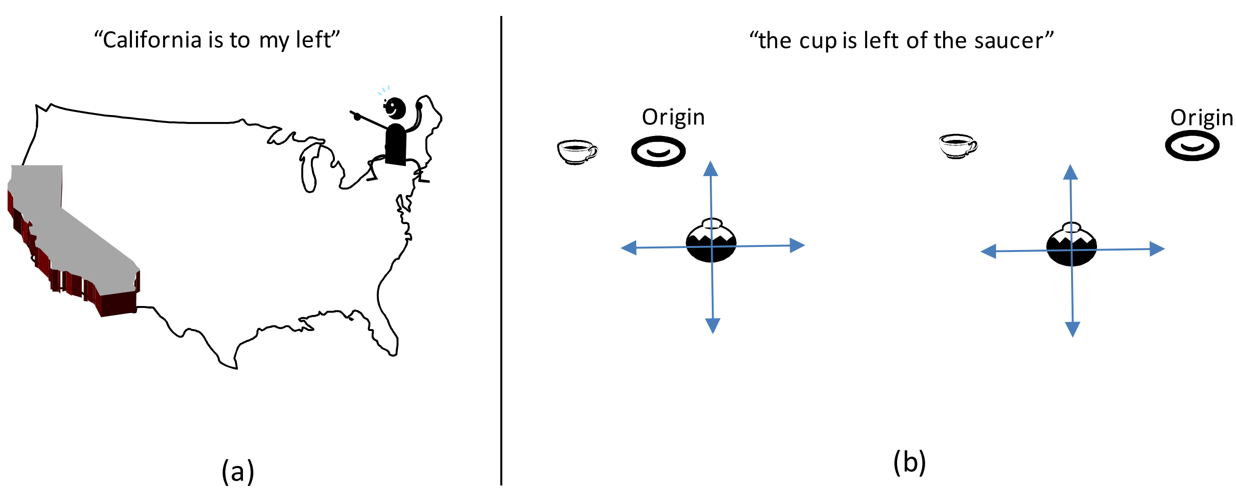

Figure 1. Scope and Origin considerations. In principle, all frames of reference can be extended to talk about locations of objects in far-reaching places, as in (a); e.g., the depicted person in (a) is referencing the position of California while in NYC. Origins can also be

flexible relocated within the coordinate framework, as in (b); in one case, the saucer is to the left of the person and the other case, the saucer is to the right.

The second consideration is coordinate origin. One can imagine the frame of reference as a grid of coordinate axes, defined by some entity (say, a person) and laid over some area of space (say, a room). For any given spatial relation, then, the origin could be located anywhere within that scope. If we said, "the cup is to the left of the saucer," the saucer could be anywhere in the room, left *or* right of the person (see Figure 1b). The coordinate framework, the scope, and the origin would jointly allow us to interpret the possible range of locations for the cup.

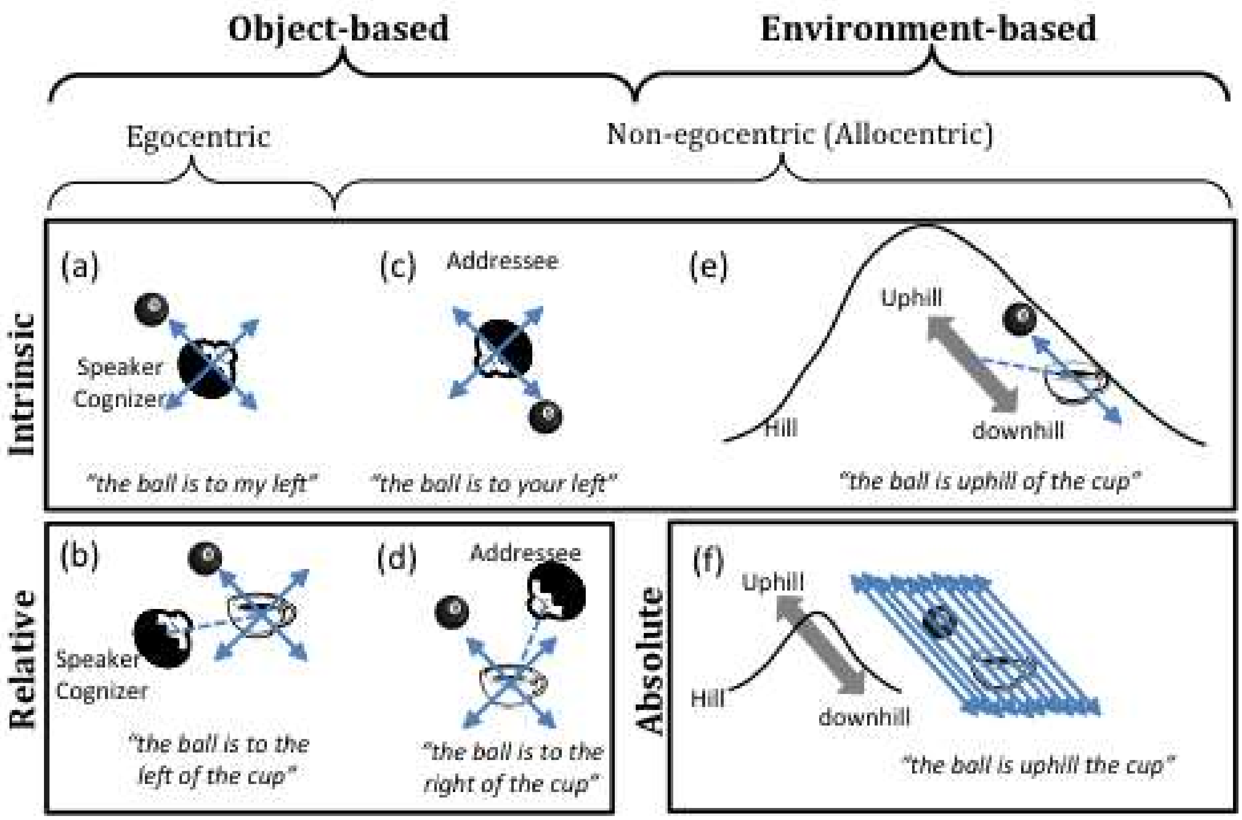

Figure 2. The figure depicts how Levinson (1996)'s tripartite taxonomy, relative, intrinsic, and absolute, cross-cuts the egocentric/allocentric and object-base/environment-based distinctions. In the examples, the ball serves as the figure object and the ground object is either the cup or the viewer (i.e., the speaker or addressee), depending on the frame in question.

LEVINSON'S TYPOLOGY: VARIATION IN FRAME-OF-REFERENCE TERMS ACROSS LANGUAGES

A different, well-known taxonomy of spatial frames of reference terms comes from a survey of over twenty languages by Levinson and colleagues from the Max Planck Institute at Nijmegen (Levinson, 1996; 2003). They observed there is considerable cultural variation in the conventions that speakers choose to adopt for talking about spatial relations (Levinson, 1996; 2003). Some cultures may adopt different features of the environment (e.g., the slope of the mountain, the direction of river drainage; "the cow is uphill of house") or different features of objects (e.g., nose, front; "the cow is at the nose of the house") as conventions. Cultures vary in the number of different conventions they use at all, which ones they use most frequently, and which ones are preferred for communication.

They suggest that three kinds of FoRs (*intrinsic*, *relative*, and *absolute*) can classify the world's languages. In this taxonomy, the three kinds of frames differ in how they talk about a *figure* object (i.e., the entity to be located) in relation to a *ground* object (i.e., the reference entity) via a *viewpoint* (i.e., the entity that defines the coordinate system). Levinson defines *intrinsic* frames as binary relations, involving a figure and a ground/viewpoint. The ground and the viewpoint are one and the same. An example would be "the ball is to the left of me," where the figure, *ball*, is related to the ground/viewpoint, *me*. *Relative* frames differ from *intrinsic* frames because they describe ternary relations, where the relation between the figure and ground is specified by a third party's viewpoint. An example would be "the ball is to the left of the cup," where the figure, *ball*, is related to the ground, *cup,* via some third party's viewpoint, such as the *speaker* or the *addressee's.* Thus, *relative* frames expand the instances covered by the *intrinsic* frames by allowing the ground to be another entity besides the one that defines the coordinate system. Finally, *absolute* frames are reserved for relating the figure to the ground via the viewpoint of an environment-based frame. However, in order to qualify as an absolute frame, the region of space covered by the frame must be infinite. Therefore, environment-based FoRs with scope restrictions (like Manhattan's uptown-downtown example) are classified by Levinson as intrinsic, not absolute.

The three frames of reference that arise from this world-language analysis -- *relative*, *intrinsic*, and *absolute* – are often equated with egocentric, non-egocentric object-centered, and geocentric frames, respectively (Majid, Bowerman, Kita, Haun, & Levinson, 2004). However, the taxonomy is actually not the same (see Figure 2; Newcombe and Huttenlocher, 2000; Watson, Pickering, & Branigan, 2006). Importantly, the distinctions made in Levinson's typology are neither more nor less specific than in ours; they are cross-cutting. First, Levinson's taxonomy focuses not just on the choice of *perspective* (coordinate system), but also on the *figure* (i.e., the entity to be located) and *ground* (i.e., the reference entity) mentioned in the linguistic expression. For instance, Levinson's typology distinguishes relative expressions, in which all three are different (the BALL is to left of the CUP from MY perspective), from intrinsic expressions, in which the ground and the perspective are collapsed (the BALL is to the left of ME from MY perspective). Note that both of these would be 'egocentric object-centered' in our framing, albeit with a different coordinate origin/figure for the spatial relation (CUP or myself).

Second, what we would call environment-based reference frames might be categorized by Levinson's framework as intrinsic, such as a room or even the island of

Manhattan, or as absolute, as in earth-based cardinal directions, if the spatial relations were binary (e.g., north side of the island, window side of the room).

Third, Levinson collapses across egocentric and non-egocentric object-based frames. Take the example of a mug to the left of a computer from my viewpoint. This would qualify as a relative expression for Levinson, because there are three separate entities for the perspective, the figure, and the ground, and egocentric left right in our conception. For Levinson, this relative expression would have the same structure as a mug to the left of a computer from your viewpoint, while for us, the shift in perspective reveals that a different frame of reference is involved (non-egocentric left-right). This is not an exhaustive list of differences, but meant to highlight that the two typologies use different defining principles.

Levinson and his colleagues' framework brings to order cross-linguistic work, articulating a clear difference between 'relative' and 'absolute' languages. Thus, one startling finding from this linguistic anthropological foray was the discovery that some languages appear not to use words like left and right at all, but only words like "sunrise" and "sunset" or "uphill" and "downhill," functionally equivalent to the cardinal directions in English (north/south/east/west). This analysis stands in contrast to previous Western theorists like Kant and Piaget, who held the view that egocentric spatial relations are more 'natural' and that absolute spatial terms are more 'abstract' and thus constructed over the course of conceptual development.

> "…our geographical knowledge, and even our commonest knowledge of the position of places, would be of no aid to us if we could not, by reference to the sides of our body, assign to regions the things so ordered and the whole system of mutually relative positions."
>
> *- Kant, Ak II, p. 379f*

In light of the Max Planck analysis, Kant and Piaget's position appears poignantly Eurocentric—if some languages of the world rely exclusively on absolute terms, and exclude relative terms like left and right, humans must be capable of entertaining a variety of spatial concepts.

The second startling finding is that this typology yields some clear ordering principles in how languages organize their spatial communication. For example, Levinson and colleagues report that languages that have absolute frames of reference always have intrinsic frames of reference, but sometimes lack relative frames of reference. Languages that have relative frames of reference always have intrinsic frames of reference. Thus, there is clearly a value to this framework, as it brings order to and reveals clear patterns in cross-linguistic data.

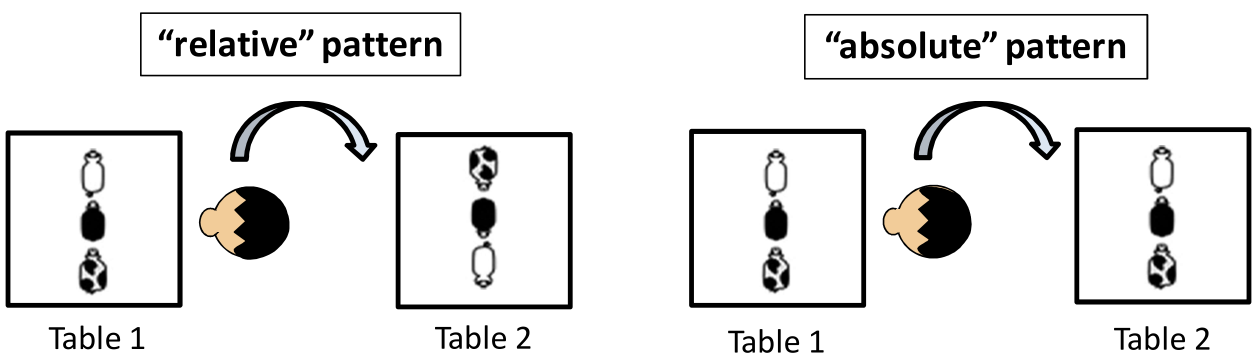

Figure 3. Animals-in-a-row task in which participants are asked to "make it the same" at Table 2 as Table 1. Most relative language speakers prefer to align animals using their own (egocentric) body axes while most absolute language speakers prefer to align animals using environment-based axes.

In addition to a valuable cross-linguistic analysis, much research has highlighted a correlation between speaking one of these two types of languages and choosing certain kinds of responses on open-ended tasks. For example, in the Animals-in-a-Row task, participants view an array of stimuli (e.g., rabbit-dog-fish), and are asked to recreate it after a 180-degree turn. Typically, speakers of absolute languages recreate the array following an absolute pattern (i.e., maintaining north-south relationships) while speakers of relative languages recreate the array following a relative pattern (i.e., maintaining left-to-right relationships; see Figure 3).

Many researchers have taken this paradigm as suggesting a correlation between language and concepts, though this position has met with considerable criticism as well (e.g., Pederson et al. 1998; Majid et al., 2004). In earlier writings, Levinson has used such findings to argue that language drives behavior and actually sets up frames of reference. As Brown and Levinson (2000) put it:

> [There are] reasons to think that some spatial concepts are intrinsically linguistic: for example, south [or left] is not given by any individual's independent act of cognition,… These kinds of spatial notions, then, are the sort of concepts that may be learned first through language. (p. 175)

This hypothesis is intuitively appealing: north/south concepts seem unnatural to many speakers of western languages, and left-right concepts are difficult for many children and even adults in 'relative-language' cultures. Therefore, language seems like a reasonable candidate for setting up such thorny concepts, and the reported correlation between language and thought is consistent with this view.

Nevertheless, there is a different way to think about the situation. Mappings between word and world may reflect cognitive architecture (e.g., see Landau & Jackendoff 1993; Barner, Li & Snedeker, 2010; Papafragou & Gleitman, 2005). In this light, the different ways that languages construct reference frames may be revealing about how our minds structure space. To test this possibility, we turned to an exploration of children's conceptual repertoire for thinking about frames of reference.

CONCEPTUAL ROOTS OF FRAME-OF-REFERENCE LANGUAGE

Two pieces of data, one experimental and one anecdotal, led us to take the possibility seriously that absolute concepts and other spatial reference frames are readily accessible even for speakers of relative, not absolute, languages. In one study, Li & Gleitman (2002) showed that English speakers, who typically respond in a 'relative' pattern on Animals-in-a-Row, could quickly be induced to shift to an 'absolute' pattern, simply by raising the blinds in the testing room to provide more visual access to the outside world. This finding suggests that, contrary to previous interpretations, both reference frames were quite readily available to these participants. In another study (Shusterman & Spelke, 2005), a child who was being trained on the words "left" and "right" had his hand held by an experimenter, who said "this is your right hand." The child then turned one half-circle, raised his left arm, and said "and now this is my right hand!" then turned another half-circle, raised his right arm, and said "and now this is my right hand!" He turned one more half-circle, raised his right arm and said "and now this is my left hand!" This child had interpreted the terms "right" and "left" as though they were environment-centric, not body-centric. This anecdote, like the blinds-up study, suggests that environment-based reference frames are surprisingly easy for people to think about, even for children.

To test this intuition, we assay children's interpretations new spatial frame-of-reference terms using a word-learning paradigm (Brown, 1957), following on the example of previous studies using this method to articulate spatial concepts. We had two notions in setting up the study. First, how would children approach a word-learning situation for novel spatial words, especially if, as others had suggested, language set up these concepts in the first place? Would they be confused about our new words, or systematic in their assumptions about what they meant? Second, if children did show systematic patterns of responses in word-learning, what could we conclude about their conceptual representations of reference frames?

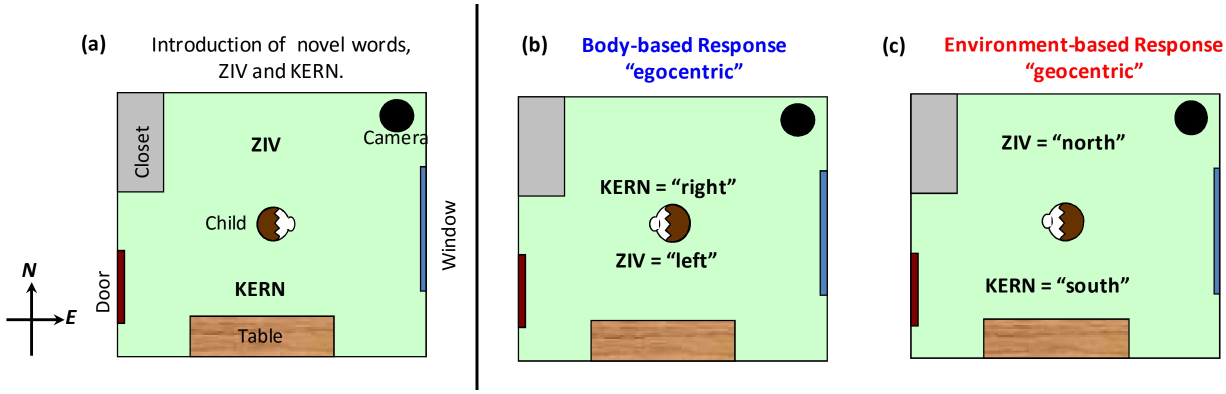

Figure 4. Children were introduced to novel spatial words, ZIV and KERN, while facing one direction. Then they were turned 180 degrees and asked to indicate the directions and locations of objects with the novel words. (b) and (c) depicts two types of interpretations children might have regarding the novel words.

In a series of experiments, 4-year-old English speaking were given instructions about fake directions, along the lines of: "This way is ZIV. This way is KERN. This is your ZIV arm. This your KERN arm. This toy is on your ZIV side. This toy is on your KERN side." The specific wording varied in different studies, but in all of them, children were asked as series of questions (e.g., Can you point to the ZIV?) to ensure they had encoded the presented words. Then, the children were turned 180 degrees and asked to point ZIV, to raise their KERN arm, to find the toy on their ZIV side, and so on. This allowed us to test whether they interpreted ZIV and KERN as referring to their bodies, in which case the words would rotate with them, or as referring to external locations or directions, in which case the words would stay fixed despite the child's rotation.

What we found was that it did not matter whether we started by talking about the children's ZIV and KERN sides, about toys next to them, or about the ZIV and KERN sides of the room; subtle changes in the syntax of the instructions also did not matter (see Figure 4). Children overwhelmingly and systematically interpreted the terms as having environment-based meanings (Shusterman & Li, under revision; i.e., Figure 4c rather than Figure 4b). After they turned 180 degrees, the very arm we had just labeled ZIV while shaking that hand, became KERN to them, just as it had for the little boy who inspired the study in the first place.

We then pushed to see how robust children's representations were. In one task, we taught children ZIV and KERN, giving feedback consistent with environment-based interpretations (i.e., each time the child rotated, we asked them to respond to ZIV/KERN requests and corrected them according to an environment-based interpretation). We then took them out of the room and up the hallway, asking them to point ZIV and KERN from various positions. Children's overall performance is well illustrated by one participant in the original study, who left the study room and went to meet her mother in the waiting area. She said, "I learned ZIV and KERN!" and proceeded to point—correctly—in the corresponding cardinal directions.

In another test of children's generalization, we assessed whether children could flexibly use the environment-based reference frame with a new coordinate origin. We placed an array of three objects on one side or the other of the child: one reference object, like a stuffed animal, with small identical toys on either side. Thus the whole array could be to the ZIV of the child, and we could ask for the toy to the KERN or to the ZIV of the stuffed animal. If children answered only relative to themselves as the coordinate origin, they should have chosen the toys at chance; but if they could answer relative to the new coordinate origin, they should have chosen the correct toy. Again, four-year-old *English-*

speaking children were excellent at this game: they could use ZIV and KERN flexibly relative to a new reference point. Curiously, under Levinson's framework, absolute frames of reference are not supposed to support the computation of spatial relations like this; but here, children made these computations flexibly and spontaneously.

In a final test of the robustness of environment-based reference frames in children, we ran the same experiment but in a white, windowless room devoid of any landmarks at all. The only way to track one's position was to maintain an internal sense at all times of which way had been labeled ZIV and KERN, and to update those labels after each 180-degree turn. Despite the complete lack of external cues to facilitate a representation of the environment, some children nevertheless interpreted the novel terms using an environment-based framework and maintained it through multiple rotations.

Thus, environment-based reference frames, sometimes called geocentric reference frames, appear to be quite trivial for children to represent and use in word learning. Data using quite different methods has revealed similar patterns not just in English-speaking children but also in diverse language communities (Haun et al., 2006).[1] Furthermore, like children, several species of great apes found it easier to learn the pattern of a hiding game when the pattern was geocentric, not egocentric (Haun et al, 2006).

RELATIVE FRAMES OF REFERENCE

What about the so-called "relative" concepts? Are they deceptively simple, like the environment-based reference frames seem to be? Or are they constructed through

[1] Importantly, we found similar patterns of performance for the few children who passed a test of left-right knowledge at above chance levels as for children who did not know left from right.

language and culture, as some have argued (Majid et al, 2004). And more mysteriously, why don't all languages have them?

One possibility for explaining why some languages never develop relative frames of reference is that reasoning about relative frames of reference is computationally difficult. Another possibility is that adding relative frame of reference expressions to a language introduces ambiguities. For instance, in a language that exclusively uses intrinsic frames of reference and lacks relative frames of reference, the expression "the ball is to the right of the girl" is unambiguous: The ball is on the side of the girl determined by the right side of her body. However, in a language with relative frame of reference, such an expression could be ambiguous between the intrinsic meaning and the relative one (girl's right or someone else's right). Since the speaker must craft expressions such that the listener can infer the speaker's intended meaning, perhaps it is simpler if languages do not make use of relative frame of reference expressions, especially when intrinsic or geocentric frames of reference are available, sufficient, and unambiguous for indicating directions and locating objects.

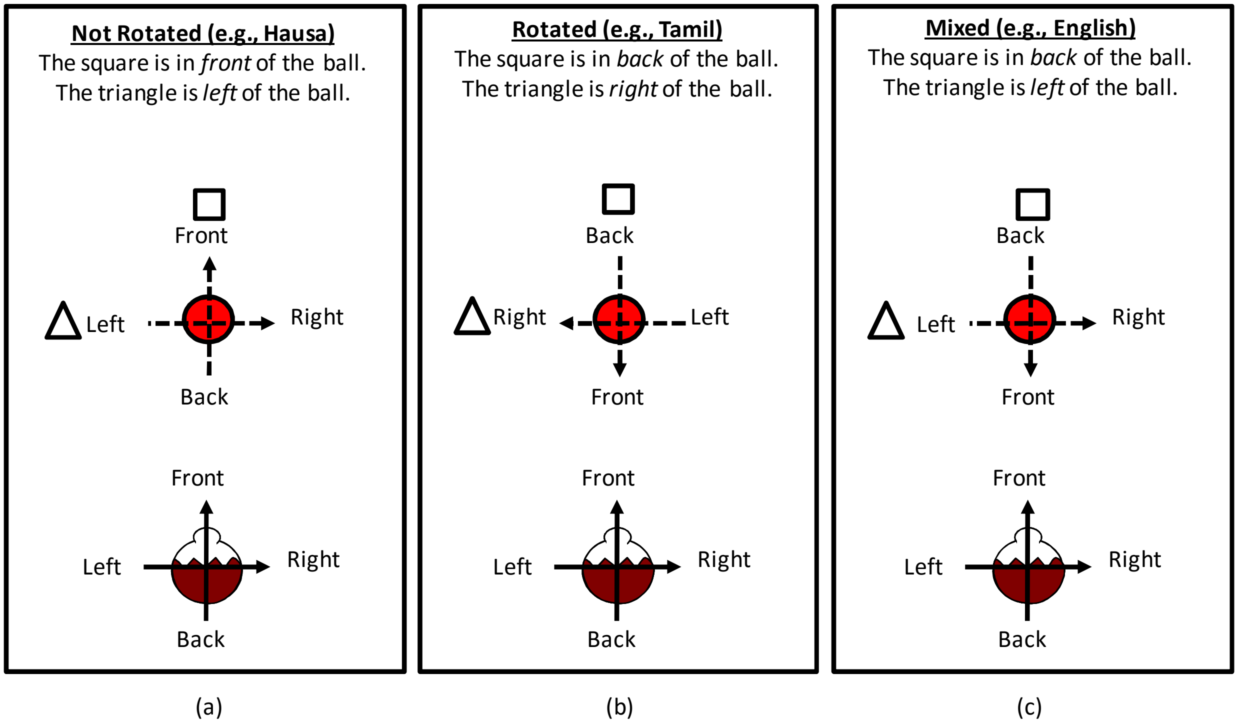

Figure 5. The figure depicts the conventions that three relative languages adopt to assign axes to non-fronted objects, such as the ball. In (a), speakers' axes are translated onto the ball. In (b), speakers' axes are rotated and translated onto the ball. In (c), the front-back axis is rotated and then translated while the left-right axis is simply translated.

Given these ambiguities, communities tend to develop conventions to reduce ambiguity when using relative frames of reference. Idiosyncracies in these conventions are likely to contribute to cross-linguistic variation. For example, some languages use the speakers' perspective in relating the figure to the ground, while others take the perspective of an imaginary listener who is facing the speaker (Levinson, 2003). Sometimes the developed conventions differ depending on the particular coordinate axes (front-back vs. left-right). These various instantiations can be seen respectively in languages such as Hausa, Tamil, and English, when these languages locate a figure relative to a non-fronted object (i.e., an object that has no fronts, such as a ball; see Figure 4 for example; Levinson, 2003). In Hausa, one can think of the coordinate system of the speaker being projected onto the ball to determine front-back and left-right. In Tamil, one can think of the coordinate system as being rotated onto the ball so that the front-back and left-right relation is from the perspective of someone imaginary facing the speaker. In English, determining left-right is a direct projection just like in Hausa, but determining front-back is a rotation like Tamil.

How can we differentiate these possibilities – whether relative expressions are avoided in some languages because they are computationally difficult, or because they create ambiguities in communication? To address this, we asked which frames of reference were learnable by children. If it is the case that relative frames are challenging, but intrinsic ones are straightforward, then children should readily learn left and right on their own bodies, and on other objects, as long as they don't have to engage in relational reasoning about them. If, on the other hand, the cognitive distinction is not between intrinsic and relative but rather between egocentric and non-egocentric object-based frames, then one might expect children to readily handle left-right relations from their own perspectives but struggle with the left-right of another entity. Children have long been reported to struggle with left and right, even their own. However, the input conditions are typically messy and ambiguous between whose left or right is being discussed. We wondered, therefore, if our structured introduction of terms and deliberate feedback would help children learn, and if so, which meanings would stick. Drawing on children's success in learning environment-based spatial terms, we applied the same training method to left and right.

BEYOND THE BODY: THE PROBLEMS OF EXTENSION, TRANSLATION, AND ROTATION

We started by considering what about left and right could make it so challenging, which we summarize as three problems a child must solve in order to flexibly use these words in the way that adults conventionally do in English and other "relative" languages.

First, a child has to realize that the meanings of these terms are not limited to the borders of one's own body. "Left" does not just mean a certain side of one's body; rather, "left" also means in the direction of the coordinate axes set up by one's body (i.e., a leftward direction). This is the problem of *extension* because the meaning of left and right is extended beyond the body.

Furthermore, the egocentric axes set up by one's body can be applied onto another object through *translation*. In the example "the cup is to the left of the saucer," the speaker's left-right axes can be imposed onto the saucer, which serves as the ground in the spatial relation. We call this the problem of *translation* because the left-right relation is defined by transferring the center of the body's coordinate axes a certain distance, without rotating them, to the new ground object.

Lastly, "left" can refer not just to coordinates set up by one's own body, but any entity to which a front-back axis can be ascribed. These left-right axes may be rotated with respect to the speaker's own left and right, such that the intended meaning of the speaker's "right" is actually on a speaker-facing listener's left. This is the problem of *rotation*.

We found that children readily mapped novel spatial terms onto their bodies with simple instructions drawing children's attention to their bodies, for example by placing a bracelet on one hand for a few trials or by explaining "your body has two sides, a ZIV side and a KERN side." They could rapidly transfer this mapping to objects next to them, suggesting they had solved the problem of extension. They could also readily apply them to a new coordinate origin and compute novel spatial relations like "KERN (right) of the stuffed animal." Like in the environment-based task, this might have been a confusing request if she used herself as the coordinate origin for the spatial relation, since both toys would be to the KERN of her on half the trials and neither toy would be to the KERN of her on the other half. Nevertheless, children performed remarkably well on these trials, indicating that they could easily solve the problem of translation to a new coordinate origin. Thus, somewhat surprisingly, left and right is not a big challenge for children, as long as the instructions are clear, the input unambiguous, and the issue is their own left and right.

However, rotation, or trying to represent a non-egocentric left-right frame, proved to be an intractable problem. Across many manipulations, children could not learn the ZIV and KERN sides of a doll. Instead, they systematically mapped the terms to themselves or to the environment. Attempting to show them how their own left and right sides corresponded to the doll's did not help, nor did putting a bracelet on the doll for a few practice trials, a manipulation that was so beneficial when learning their own left and right. To test whether children could even notice or remember left and right on another entity, they were asked to remember which pocket the coin was hidden on a doll, one sewn onto its right hand or the one on the left. The child would watch the coin be hidden, close their eyes while the doll was rotated, and open their eyes to retrieve the coin. Performance was at chance. It is noteworthy that, in contrast to left-right, front-back relations are represented easily and early, by two years of age (Levine & Carey, 1982).

Other evidence from a variety of spatial behavior tasks is concordant with the conclusion that object-centered, non-egocentric left-right reference frames are uniquely challenging, for children and possibly for adults too. In spatial memory tasks, in which participants had to memorize the relative positions of objects in a spatial array, both children and adults perform better when they walk around an array than when the array rotates on a table (Nardini, Burgess, Breckenridge, & Atkinson, 2006; Simons & Wang, 1998). This may be because it is hard to represent objects rotating relative to the table-based frame of reference. Similarly, Lourenco and Huttenlocher (2006) found different performance on a search task when the search arena was rotated than when the child was rotated; again, this requires representing the object-based frame of reference and tracking it through rotation.

In summary, then, it does not seem to be the case that children are “good at egocentric but bad at allocentric” frames of reference, or that they are “good at intrinsic but bad at relative.” Rather, our studies suggest that children are uniquely limited in reasoning about non-egocentric, object-centered, left-right relations. They can contemplate other kinds of non-egocentric spatial relations (such as environment-based ones), other kinds of object-centered spatial relations (such as ones involving front and back), and other kinds of left-right relations (such as egocentric ones involving their own bodies).

OUTSTANDING QUESTIONS

We are left with a number of questions. First, how should we align these developmental data with the cross-linguistic analyses? The developmental data does not align cleanly with Levinson’s typology, yet the cross-linguistic typology clearly has explanatory force. There are some points of alignment: Environment-based frames of reference map at least somewhat onto *absolute* (Fig. 2f)[2]. Within the object-based frame of reference, where the

[2] Whereas Levinson would identify the difference between 2e and 2f as a major cut across the intrinsic and absolute reference frames, we would identify both within the environment-based reference frame but differing in scope. A bounded environment, like

alignment is messier, there is at least an implicit agreement that we should distinguish front-back from left-right (not pictured in figure). In the cross-linguistic typology, front-back is not much discussed, likely because such terms are prevalent across most languages and thus don't constitute an important dimension of difference, while left-right relations are emphasized. The developmental data make clear that primary axes like front-back, are easier for children to learn than secondary axes like left-right (which can only be computed once the primary axis is established). This is especially the case in non-egocentric relations, where others' front-back is available to children by the age of 2 while others' left-right is challenging until late childhood and even into adulthood. Thus, the cross-linguistic and developmental data both highlight the relative ease of primary axes.

But how should we make sense of Levinson's 'relative' reference frames? One way of looking at it is that the natural cut-points that matter for language groups are different than those that matter for development. As we point out, Levinson's *relative* frame includes both egocentric left-right (Fig. 2b) *and* non-egocentric left-right (Fig. 2d) frames; the major cut point between intrinsic and relative is having or not having flexible left-right language at all (Fig. 2 a/c vs. Fig 2 b/d). In the cross-linguistic analysis, collapsing these makes sense: languages that use left-right language for egocentric spatial relations also use left-right language for non-egocentric ones. In the developmental data, however, it is clear that one is

---

the island of Manhattan, is analyzed by Levinson as intrinsic (i.e., 2e), not absolute, because the uptown-downtown axes do not extend beyond the boundaries. He argues that this contrasts with aboriginal languages whose uphill/downhill or sunrise/sunset distinctions that extend beyond the hill or anywhere on earth (i.e., 2f). We would argue that this is just a difference in conventions about scope: New Yorkers *could* extend those terms off of the island, for example by connecting the terms to east/west/sunrise or other supports for thinking about fixed bearings. Our developmental data on children's learning of environment-based reference frames suggest that their reasoning is flexible and abstract in the ways that Levinson has said are characteristic of true absolute frames (Levinson, 2002; Fig. 2f).

much harder for children than the other; perhaps languages ease the burden by adopting the language from the simpler frame to talk about the more challenging one.

Conversely, there are distinctions that don't seem to matter developmentally, but have been argued to matter in categorizing language groups. For instance, children seem indifferent to the distinction between a left-right frame without translation (to my left; Fig. 2a) and with translation (to the left of some other ground object; Fig. 2b); these are both easy for children in the egocentric case (Fig. 2 a and b) and both hard in the non-egocentric case (Fig. 2 c and d). However, this distinction is central to Levinson's cross-linguistic analyses: translation is the essential difference between his intrinsic (Fig. 2 a and c) and relative (Fig. 2 b and d) reference frames.

There is no easy answer, then, about how to align the developmental and cross-linguistic typologies; the processes of learning new language and concepts, on the one hand, and coordinating a linguistic community, on the other, are subject to different constraints.

In particular, children's word-learning skills are likely to be impacted by cognitive development, and the full development of spatial cognition takes many years. We speculate that children's developing conceptual representations limit the kinds of meanings they can entertain, especially non-egocentric left-right relations. This leads to the second question: What is it about non-egocentric left-right that makes it a uniquely hard frame of reference? Which conceptual representations need further development before this reference frame is tractable for children? Is it fundamentally a problem of perspective taking, mental rotation, symmetry, or something else entirely? Is it a modality-specific problem only in the visual system – might this challenging frame of reference be more easily represented with other senses, such as the haptic modality, or is it just hard to think about non-egocentric left-right relations?

Finally, what gives rise to human adults' eventual ability to think with an object-based frame of reference, sufficiently to dream up buildings, visualize the fit of an enzyme and a protein, or imagine another person's point of view? Relative languages have emerged all over the world, and people do talk about left and right of external objects all the time. Our understanding of the development of these capacities, however, remains elusive.

REFERENCES

Brown, P., & Levinson, S. C. (1993). 'Uphill' and 'downhill' in Tzeltal. *Journal of Linguistic Anthropology*, 3, 46-74.

Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology* 55, 1-5.

Duran, N., Dale, R., Kreuz, R. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. Cognition, 121(1), 22-40.

Haun D.B., Rapold, C.J., Call, J., Janzen G., & Levinson S.C. (2006). Cognitive cladistics and cultural override in hominid spatial cognition. *Proceedings of the National Academy of Sciences, 103*(46), 17568-17573.

Kant, I. (1902ff). Von dem ersten Grunde des Unterschiedes der Gegenden im Raume. In Kants Werke. Akademie Ausgabe II (Ak II, pp. 375–384). Berlin, Georg Reimer/de Gruyter (Original work published 1768). (J. Handyside, Trans. 1991). On the first ground of the distinction of regions in space. In J. Van Cleve & R. Frederick (Ed.), The philosophy of left and right: Incongruent counterparts and the nature of space (pp. 27–33). Dordrecht & Boston: Kluwer.

Levine, S. C. & Carey, S. (1982). Up Front: The acquisition of a concept and a word. *Journal of Child Language*, *9*, 645-647.

Levinson, S. C. (1996) Frames of reference and Molyneaux's question. In Bloom P, Peterson MA, Nadel L, Garrett MF (eds) *Language and Space*, pp 109-170. MIT Press, Cambridge MA.

Levinson, S. C. (2003) *Space in Language and Cognition: Explorations in Cognitive Diversity*, Cambridge University Press, Language, Culture and Cognition 5.

Levinson, S. C., Kita, S., Haun, D., & Rasch, B. (2002). Returning the tables: Language affects spatial reasoning. *Cognition*, *84*, 155-88.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition*, *83*, 265-94.

Lourenco, S. F., & Huttenlocher, J. (2006).How do young children determine location? Evidence from disorientation tasks. *Cognition*, 100, 511-529.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004) *Trends in Cognitive Science, 8*, 108–114.

Nardini, M., Burgess, N., Breckenridge, K., & Atkinson, J. (2006). Differential developmental trajectories for egocentric, environmental and intrinsic frames of reference in spatial memory. *Cognition*, 101, 153–172

Newcombe, N., & Huttenlocher, J. (2000). *Making space: The development of spatial representation and reasoning*. Cambridge: MIT Press.

Pederson, E., Danziger, E., Levinson, S., Kita, S., Senft, G., & Wilkins, D. (1998). Semantic typology and spatial conceptualization. *Language*, *74*, 557-589.

Piaget, J. (1928) *Judgement and Reasoning in the Child*, Routledge.

Piaget, J. & Inhelder, B. (1971). Mental imagery in the child. New York: Routledge & Kegan Paul.

Schober, M. F. (2009). Spatial dialogue between partners with mismatched abilities. In K. R. Coventry, T. Tenbrink, & J. A. Bateman (Eds.), *Spatial language and dialogue* (pp. 23-39). Oxford: Oxford University Press.

Shusterman, A. & Spelke, E. (2005). Language and the development of spatial reasoning. In P. Carruthers, S. Laurence and S. Stich (eds.), *The Structure of the Innate Mind*. Oxford University Press.

Shusterman, A. & Li, P. (under review). Conceptual precursors and fast mapping in spatial language acquisition.

Simons, D.J. and Wang, R.F. (1998) Perceiving real-world viewpoint changes. *Psychological Science*, 9, 315–320

Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: Insights from animals. *Trends in cognitive sciences*, *6*(9), 376-382.

Watson, M. E., Pickering, M. J., & Branigan, H. P. (2006). An empirical investigation into spatial reference frame taxonomy using dialogue. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.