Deixis, gesture, and cognition in spatial Frame of Reference typology*

Eve Danziger
University of Virginia

The three Frames of Reference recognized in the current inventory of spatial-language types are differentiated by their placement of the Anchor from which the vector of search space from Ground to Figure is calculated (Levinson 1996). In certain well-recognized examples, Anchor merges with Ground. The existing analysis treats this merged component as analytically Ground rather than Anchor; its location in or out of the speech situation is therefore taken to be independent of the Frame of Reference typology. Instead, I treat this component as analytically Anchor, making its speech-situation status criterial to the typology. Four, not three, Frames of Reference now appear. The fourth, "Direct" frame, distinguishes binary locutions with a speech participant as Ground/Anchor (e.g. 'in front of you') from "Object-Centered" binary locutions in which Ground/Anchor is not a speech participant (e.g. 'in front of the kettle'). This four-frame analysis corresponds better than does the three-frame one to the logic of rotation sensitivity which has been used to show Whorfian parallels between language and conceptualization across cultures. I close by discussing the application of Frame of Reference typology to pointing gestures, and show how recognition of the fourth frame of reference allows us to bring discussion of these, and of the linguistic demonstratives and locatives with which they so frequently co-occur, fully within the Frame of Reference typology.

Introduction

The domain of space is not similarly partitioned across all languages and cultures (Pederson et al. 1998), and the notion of "Frame of Reference" (Rock 1990) has proved useful in constructing an inventory of spatial-language types across languages (Levelt 1984, Danziger 1996, Levinson 1996, 2003). One stated conclusion of this line of typological research has been that "Languages use ... just three frames of reference: absolute, intrinsic and relative". (Levinson 1996: 148).

usefully ∧

In this paper, I present arguments in favor of the view that four, not three frames of spatial reference can be distinguished ~~usefully~~ in the usage of speakers around the world. The fourth frame, here called the "Direct" frame, is the one in which canonical deictic reference, with Speaker as both Anchor and Ground, finds its place in the cross-linguistic typology of spatial reference. It is argued that full recognition of this frame is vital for the continued consideration of parallels between language and thought in the domain of space, and that it can be used to bring linguistic demonstratives and locatives, and thereby also speech-accompanying gesture, into the existing spatial language typology.[1]

Frame of Reference

The notion of Frame of Reference takes account of the fact that a speaker can often take several possible perspectives on the phenomenon that s/he is considering. These changes of perspective are particularly visible in the language of spatial reference. In a particular utterance, for example, is it the train in which the speaker sits that is moving? Or is it the landscape that is "rushing by" outside the window? Again, does the patch on your bicycle tire "change location" when you cycle to work? Or is it still "in the same place" (cf. Levinson 1996)? The different possible answers to these questions represent different Frames of Reference a speaker might adopt in describing and in conceptualizing the scenes in question.

Which of the different possible Frames of Reference a speaker uses with respect to a given scene will depend upon the selection of a conceptual Anchor for the scene. Alongside Talmy's (1983) Figure (the entity located) and Ground (the entity with respect to which the Figure is located), this notion of Anchor is a crucial element in the vocabulary needed to discuss alternative perspectives among linguistic and conceptual expressions of spatial location. The Anchor is the zero point from which the vector is calculated that narrows the search space from Ground to Figure. The Anchor therefore, is that part of the scene which the speaker treats as immovable, fixed, in relation to the others (Levinson 1996).

Absolute, relative, and intrinsic

The three Frames of Reference recognized in the current typology are distinguished from one another by their placement of the Anchor from which the relationship of Figure to Ground is calculated.[2] Consider for example, Figure 1, which shows a single spatial configuration, in which the Figure is a carton of milk and the Ground is a kettle of water.
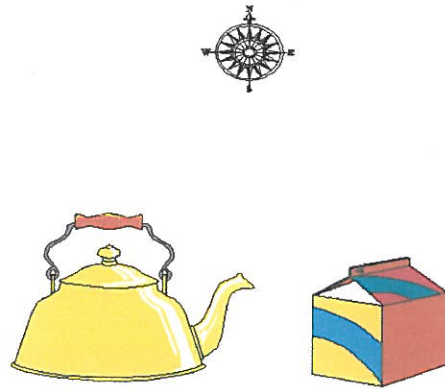
Figure 1. Where's the Milk?

To illustrate what is meant by Absolute, Relative and Intrinsic Frames of Reference, let us consider three possible answers to the question "Where's the milk?"

(1) ABSOLUTE: The milk is to the east of the kettle.

(2) RELATIVE: The milk is to the right of the kettle. (From the speaker's perspective)

(3) INTRINSIC: The milk is at the spout of the kettle.

In sentence (1), which illustrates the Absolute Frame of Reference, the Anchor is located in the landscape or the cosmology surrounding the Figure-Ground scene. In sentence (2), which illustrates the Relative frame of reference, the Anchor is embedded in one of the speech-situation participants.[3] In sentence (3), which illustrates the Intrinsic Frame of Reference, the Anchor is neither in the surrounding landscape, nor in a speech-situation participant, but in the Ground object itself. In each case, the felicity of reference is sensitive to the rotation of the Anchor that defines the Frame of Reference, and the Anchor for a given locution can therefore be defined as that component which, if rotated, renders the locution false. Rotation of the kettle, for example, renders false the locution in (3), while the other statements are unaffected. Similarly, rotation of the Speaker renders (2) false, while having no effect on the other two statements. If we graphically chart the contrasts which define the different types of Anchor, we obtain a layout like that in Table 1.

The labels for the rows and columns in Table 1 have been chosen rather freely from different literature that have separately discussed the two kinds of contrasts, but which have rarely considered them together.[4] Other terms could also have been chosen.[5]

Of prime importance in the vertical contrast (here labeled Ternary/Binary) is the question of whether the Anchor for a particular Frame of Reference is to be

Table 1. Frames of Reference — Three Types of Anchor

	Allocentric Anchor is not a speech-situation participant	Egocentric Anchor is a speech-situation participant
Ternary Anchor is not Ground	**Absolute** *The milk is to the east of the kettle.*	**Relative** *The milk is to the right of the kettle.* (From the speaker's perspective)
Binary Anchor is (part of) Ground	**Intrinsic** *The milk is at the spout of the kettle.*	

found in the Ground object or not. If it is, the Frame of Reference is considered "Binary", since only two distinct entities — Figure and Ground/Anchor play a role in it. If on the other hand the Anchor is distinct from the Ground, the Frame of Reference is considered "Ternary", since three separate entities, Figure, Ground and Anchor, can be distinguished (cf. Levinson 1996, Majid et al. 2004).

Of prime importance in the horizontal contrast (here labeled Allocentric/Egocentric) to date has been the question of whether this third term, the Anchor that is not Ground, is or is not to be found in one of the speech-situation participants (speaker or addressee). In what follows, I propose a generalization of this distinction so that the speech-situatedness of Anchor becomes criterial with regard to Frame of Reference even when Anchor is part of Ground (i.e. for Binary as well as for Ternary Frames of Reference). This will rationalize the template of rotation sensitivities among the Frames of Reference that has been the basis for empirical arguments in favor of Whorf-type parallels between language and cognition in this domain, and will pave the way for the incorporation of linguistic demonstratives and locatives, and the pointing gestures that so often accompany them, into the spatial Frame of Reference typology.

Anchor and Ground

As we have seen, the critical issue in distinguishing among Frames of Reference in this typology is to identify the nature of the Anchor, not the nature of the Ground object. A key clarifying insight emerges from this fact: Utterances within Absolute and Relative frame-types can take speech-situation participants as Ground objects equally well (Levinson 1996, Haviland 1998). The Frame of Reference classification is therefore independent of the issue of linguistic deixis, to this extent. For example:

ABSOLUTE and non-DEICTIC (Allocentric Anchor, Allocentric Ground):
(4)   The milk is to the east of the kettle.

ABSOLUTE and DEICTIC (Allocentric Anchor, Egocentric Ground):
(5)   The milk is to the east of you.

If we apply the logic of rotation sensitivity to identifying the Anchor in these ex-
amples, we see that Anchor is indeed distinct from Ground in both cases, and in
neither case is Anchor to be found in a speech participant. If the milk is "to the
east" of the kettle (example (4)), it remains so whichever way the kettle is facing,
and the same is true if we rotate the speech participant Ground ("you") in example
(5). Rotation of the Ground object does not require a change in the spatial descrip-
tor in order to maintain accuracy, whether or not the Ground object is located in
the speech situation. This observation also applies to parallel examples within the
Relative Frame of Reference:

RELATIVE and non-DEICTIC (Egocentric Anchor, Allocentric Ground):
(6)   The milk is to the right of the kettle. (From the speaker's perspective)

RELATIVE and DEICTIC (Egocentric Anchor, Egocentric Ground):
(7)   The milk is to the right of you. (From the speaker's perspective)[6]

Again, the use of the original spatial descriptor (here "to the right of") remains
accurate in both of these examples under any rotation of the Ground object, even
though in (7) the Ground object is a participant in the speech situation and in
(6) it is not. In both (6) and (7), we can see that rotation of the speaker (Anchor)
would indeed require a change of locution ("right" would no longer be accurate),
whereas rotation of the Ground ("kettle", "you") would not. Insensitivity of the
accuracy of locutions within the Ternary (Absolute and Relative) Frames of Refer-
ence to the rotation of their Ground object derives directly from the fact that it is
the orientation of the Anchor, and not that of the Ground, that dictates the neces-
sary phrasing.

The impetus for a Frame of Reference typology that recognizes only three
types comes from the application of the same logic to utterances within the Binary
Frame of Reference type. By analogy with Absolute and Relative, as exemplified
above, it has been declared (Levinson 1996, Pederson et al. 1998: 589, but see Dan-
ziger 1998) that all utterances with Ground as Anchor should fall within the same
(Intrinsic) frame of reference whether or not their Ground object is drawn from a
speech-situation participant:

INTRINSIC and non-DEICTIC (Allocentric Ground/Anchor):
(8)   The milk is at the spout of the kettle.

INTRINSIC and DEICTIC (Egocentric Ground/Anchor):

(9)   The milk is in front of me. (With reference to speaker's own front. i.e. 'at my front')[7]

But the analogy leads to a logical choice, since the nature of the Ground object is independent of the nature of the Anchor only for Ternary (Relative, Absolute) and not for Binary Frames of Reference. In Binary Frame of Reference space, if the Ground object is found in a speech-situation participant, then the same is true for Anchor. This can be confirmed by noting that in (8) and (9), in contrast to what was observed in examples (4) through (7), rotation of the Ground object ("kettle", "me") indeed results in the falsification of the locution. Precisely because Ground and Anchor are identical in these examples, rotation of the Ground object actually does create an inaccuracy of the locution — the milk is no longer "at the spout" of the kettle if the kettle has been rotated, and it is no longer "at my front" if I have turned around. But if, as in (9), Anchor is found in a speech-situation participant, then surely we are dealing with a strictly Egocentric Frame of Reference. Whereas if, as in (8), the Ground object in a Binary Frame of Reference is found outside a speech-situation participant, the Anchor is also by definition found outside any speech-situation participant. And in such a case, we must be dealing with a strictly Allocentric Frame of Reference. While it is therefore true that in the case of Ternary Frames of Reference, as exemplified in (4) through (7), the question of whether Ground is or is not to be found in the speech situation is not criterial to determining the Frame of Reference classification of the locution, the same is not true for Binary Frames, as exemplified in (8) and (9). The original analogy which suggested that it was, was a false one.

Table 2 summarizes and introduces the new terms "Object-Centered" Frame of Reference, for the intersection of Allocentric and Binary values of the matrix, and "Direct" Frame of Reference, for the intersection of Egocentric and Binary values.[8]

Table 2. Frames of Reference — Four Types of Anchor

	Allocentric Anchor is not a speech-situation participant	Egocentric Anchor is a speech-situation participant
Ternary Anchor is not Ground	**Absolute** *The milk is to the east of the kettle.* (Or: *The milk is to the east of you.*)	**Relative** *The milk is to the right* of the kettle.* (Or: *The milk is to the right* of you.*) *from the speaker's perspective
Binary Anchor is (part of) Ground	**Object-Centered** *The milk is at the spout of the kettle.*	**Direct** *The milk is in front* of me.* *with reference to speaker's own front

Both the existing three-frame and the proposed four-frame analyses define Binary Frames of Reference as those in which Anchor merges with Ground. The new four-frame analysis differs from the existing three-frame one in choosing to treat this merged component analytically as Anchor, so that its location in or out of the speech situation is criterial to the Frame of Reference typology. In the three-frame typology, this component was treated as Ground, so that its location in or out of the speech situation was independent of the assignment of the locution containing it to a position in the Frame of Reference typology.[9]

Strictly from the point of view of the logic involved, there is perhaps not much reason to prefer the four-frame analysis to the three-frame one. And the three-frame version already has considerable historical precedent as well as a certain amount of literature to its credit. Nevertheless, the four-frame version conforms better than the three-frame version to the empirical facts of rotation sensitivity that have been the key to the finding that Frames of Reference used in spatial language often have parallels in non-linguistic spatial problem-solving and/or in speech-accompanying gesture. The analogy between the sentence pairs (4)/(5) or (6)/(7) and (8)/(9) is, in short, not complete, and it is incomplete in a way that has a direct impact on the fact that rotation sensitivities within the spatial Frame of Reference typology have been used to demonstrate Whorfian parallels between language and conceptualization. Finally, the four-frame analysis is more useful than its three-frame counterpart in typological discussion not only of lexical spatial relators but of deictic demonstratives and locatives, and their associated pointing gestures, as well.

Language and thought

Not all languages make use of all Frames of Reference in every spatial-reference context. In particular, while Binary Frames of Reference are apparently universal across languages and cultures (Danziger 1999, 2001), the adoption of one or both of the Ternary frames is a matter of cultural option (Pederson et al. 1998). This latter fact, with the differential sensitivity of Anchor to rotation under the different Frames of Reference, has been the key to exploring whether and to what extent conceptual organization relevant to linguistic expression also extends to non-linguistic realms.

The interest of the present Frame of Reference typology in classifying varieties of spatial language around the world has been proved by the correlation of a linguistic preference for a Relative or Absolute frame of reference with analogous non-linguistic problem-solving preferences (Pederson et al. 1998, Majid et al. 2004), and with gestural habits (Haviland 1993, Levinson 2003) among speakers.

For example, speakers of languages such as Tzeltal (Mayan) or Arrernte (Pama-Nyungan) deploy an Absolute but not a Relative linguistic Frame of Reference to describe a configuration like that of the milk and the kettle. Speakers of these languages will, if rotated 180 degrees and asked to rebuild the scene, place the milk to the east of the kettle — thus violating the original relation of the array to their own right and left sides. Speakers of languages such as Dutch or Japanese, in which a Relative but not an Absolute spatial frame of reference is in everyday linguistic use to describe this kind of scene, will, under the same rotation circumstances, preserve the original right-left relations in their rebuilding of the scene, and violate the original north-south-east-west relations in order to do so.

These cross-linguistic findings make it clear that the differences among Frames of Reference can be understood as a psychological as well as a purely linguistic matter. They have raised the possibility for there to exist a true Whorfian connection between linguistic expression and psychological conceptualization in the domain of space (cf. Pederson et al. 1998). I show in what follows, however, that the current three-frame division of the typology is not fully suited to carry the debate forward into these non-linguistic realms. Instead, a four-frame typology is preferable.

Frame of Reference and rotation sensitivity

We have seen that the linguistic description of a scene can be identified within the Frame of Reference typology by specifying the nature of the Anchor from which the vector is calculated to narrow the search space from Ground to Figure. Since this is so, the Anchor, in turn, can be defined as that element of the scene which, if rotated, falsifies the description. We can therefore consult rotation sensitivities of any part of a spatial scene in order to identify the Anchor, and hence the Frame of Reference, for any particular description of that scene. The different types of possible rotation sensitivities have been charted (after Levinson 1996: 149, with examples from our milk and kettle array) in Table 3:

We now see why the non-linguistic tasks that were so useful in distinguishing Absolute from Relative problem-solvers relied on rotation of the individuals involved. Absolute, but not Relative linguistic descriptions need no adjustment if the speech participant is rotated, since in an Absolute frame, the speech participant is never the Anchor (although sometimes the Ground). In Relative descriptions however, the speech participant is by definition the Anchor, and such descriptions are therefore vulnerable to the rotation of the participant. Non-linguistic problem-solvers in the rotation tasks are asked to build a new Figure-Ground array such that it continues to fit their mental representation of the array they first saw. The

Table 3. Rotation Sensitivities: Three Frames of Reference

Description for diagnosis	A. Description still felicitous under rotation of speech participant?	B. Description still felicitous under rotation of Ground?	C. Description still felicitous under rotation of Figure-Ground array?	Frame of Reference diagnosis
1. milk east of kettle	Yes	Yes	No	Absolute
2. milk to right* of kettle	No	Yes	No	Relative
3. milk at spout of kettle	Yes	No	Yes	Intrinsic

* from the speaker's perspective

differences in rebuilding strategies that become visible when the participants are rotated demonstrates that these mental representations have Frame of Reference properties (rotation sensitivities) that differ across cultures in ways that run parallel to those which characterize the linguistic descriptions that the individuals involved would also be inclined to use.[10]

Rotation sensitivities and the Direct Frame of Reference

Notwithstanding its usefulness in designing and interpreting non-linguistic tasks that require distinction between Absolute and Relative frames of reference in language and in non-linguistic problem-solving, it is clear that the chart of rotation sensitivities is incomplete. Only three of the four logical possibilities in which the felicity of a description is sensitive to some but not to all kinds of rotation are included in the existing table. Missing is the case in which a description remains felicitous only under the rotation of the entire Figure-Ground array, but not under the rotation of speech participant, or of Ground. One example of a description that fulfills these rotation-sensitivity criteria might be "The milk is in front of me." (with reference to speaker's own front) in which the merged Ground/Anchor of the locution is found in one of the speech participants. Such a description is indeed rendered infelicitous through the rotation of the relevant speech participant, requiring a No instead of a Yes in column A. (This would happen for example if the speaker turned his or her back to the carton of milk).[11] We can re-draw the entire table to include this case as in Table 4:

Protestations of inclusion to the contrary, it is now clear that a description like (9) "the milk is in front of me" (with reference to the speaker's own front), does not share the rotation sensitivities of the Intrinsic Frame of Reference as characterized in Table 3. Where rotation sensitivities are concerned, four, not three Frames of Reference possibilities are clearly in play. And since rotation sensitivity has been

Table 4. Rotation sensitivities: Four Frames

Description for diagnosis	A. Description still felicitous under rotation of speech participant?	B. Description still felicitous under rotation of Ground?	C. Description still felicitous under rotation of Figure-Ground array?	Frame of Reference diagnosis
1. milk east of kettle	Yes	Yes	No	Absolute
2. milk to right[a] of kettle	No	Yes	No	Relative
3. milk at spout of kettle	Yes	No	Yes	Object-Centered
4. milk in front[b] of me.	No	No	Yes	Direct

[a]from the speaker's perspective
[b]with reference to speaker's own front (i.e. "at my front")

the key to psychological task designs for the investigation of Whorfian effects in the domain of spatial reference, the addition of the fourth set of rotation sensitivities and the corresponding Direct Frame of Reference is crucial to filling out the typology. Its absence, for example, could explain the apparent finding (Li et al. 2005) that Tzeltal speakers, lacking a Relative Frame of linguistic reference, nevertheless sometimes solve rotation puzzles in an "Egocentric" fashion.[12]

While the three-frame analysis which followed the decision to treat the merged Anchor/Ground component of Binary descriptions as analytically Ground (and therefore independent of Frame of Reference classification) was logically defensible, it resulted in collapsing two distinct cases of rotation sensitivity into a single typological category. In carrying the discussion into the realm of non-linguistic tasks that depend upon differences in rotation sensitivity to distinguish one type of mental representation from another, the three-frame typology is therefore less useful than a four-frame typology which keeps these two cases apart. I will now continue to a more recent realm of interest to spatial Frame of Reference typology — that of pointing gestures — and show that here also, recognition of the Direct Frame of Reference makes an essential contribution.

Gesture, demonstration, and Frame of Reference

The interest in pointing gestures was a new departure for the spatial Frame of Reference paradigm, since, in the original research that gave rise to the three-frame typology, deictic locatives and demonstratives were considered undesirable data — precisely because it was felt that they were likely to rely on pointing gestures to establish the spatial vectors involved. Methods were elaborated specifically so as to

make it unlikely that such speech would be elicited, and speakers were deliberately positioned so that they could not see one another (Baayen and Danziger 1994). But after the demonstration of principled alignments between Frames of Reference used in speech and in non-linguistic problem-solving, attention turned to the pointing gestures that accompany speech (Haviland 1993, cf. Levinson 2003). The question was again asked whether Frames of Reference used in habitual speech showed parallels to Frames of Reference observed outside language — this time in gesture.

It had been observed relatively early on (Haviland 1993) that speakers of languages in which an Absolute Frame of Reference was preferred to a Relative one in speech also showed distinctive tendencies toward preserving cardinal direction orientation in their spontaneous gestures. In later, more controlled, investigations of gesture within the Frame of Reference paradigm (Levinson 2003), participants watched a simple film of an activity that occurred along their own left-right axis on a video screen (for example, a boy kicking a ball from left to right across the screen). The activity of course also occurred on some cardinal direction axis (say, north to south). Participants were then asked to rotate 180 degrees before recounting the action of the film. As expected, speakers frequently accompanied their accounts with gestures, and it was noted that speakers of languages in which the Relative Frame of Reference was common in speech maintained gestural reference to the boy, for example, on their own left. While speakers of languages in which the Absolute Frame of Reference was habitual also maintained gestural reference under the rotation, the difference was that they continued to gesture northward (rather than leftward) when discussing the boy. It was concluded that the former group of participants could be said to be gesturing within a Relative Frame of Reference, and the latter within an Absolute one (Levinson 2003).

It is noteworthy that, just as in the rebuilding tasks, the distinction between Absolute and Relative Frame of Reference in the gesture study was achieved via the rotation of the participant. As in the earlier tasks, the rotation allowed for differences in the Frame of Reference properties of the two types of behavior, which might have been otherwise indistinguishable, to become visible. It is equally noteworthy however, that neither the type of gesture classified as Relative nor that classified as Absolute in the study involved geographically accurate gestures toward the spot where the screen actually stood during the viewing (i.e. back over the gesturer's shoulders). This is evidence that the gestures involved were actually transposed (Bühler 1934, Haviland 1993) from some simpler deictic form in which pointing gestures would indicate the actual direction in which the speaker (Ground) must search in order to find the Figure. That simpler form is precisely the type of gesture that the original Frame of Reference elicitation efforts went to some lengths to avoid. It is the type of gesture, for instance, that might be expected

to accompany the following very plausible answer to the original question about the location of the milk:

(10)   The milk is right over there!

As before, it is clear that the milk is the Figure. As long as the locution has not undergone transposition, the Ground object with respect to which this Figure is to be located is the speaker him or herself (cf. Bühler 1934). But which part of this locution supplies the Anchor, the critical information as to the particular direction from Ground in which the addressee must search in order to locate the Figure? Within the speech itself, this information is of the vaguest sort. The English word 'there' at best encodes a direction "away from the speaker". The speech alone does not provide vector information that would allow the hearer to narrow the search space from Ground to Figure.[13] The sentence, in short, requires a pointing gesture in order to provide such vector information.[14] Many theorists (Bühler 1934: 108, see also Enfield et al. 2007, de Ruiter and Wilkins 1998) have indeed taken it as axiomatic that pointing gestures are obligatory (or nearly so, see Clark et al. 1983) for any spoken speech-situated deictic form.[15]

locative

With a pointing gesture toward the milk, and assuming a non-transposed utterance, the distribution of information over the parts of the proposition is as follows:

(11)   The milk is {in the direction of my gesture} from me.
      FIGURE   ANCHOR   GROUND

The gesture vector that constitutes Anchor for the locution is conveyed by extension of a part, or facet, of the speaker's body. Anchor is thus located as part of the Ground, and we see that the rotation of the Ground will indeed disrupt the felicity of the expression (try turning the pointing speaker with the pointing gesture frozen in place, but without moving the milk). But this Ground/Anchor is also the speaker, so that the locution is also vulnerable to the rotation of the speech participant. Finally, if the pointing speaker is rotated along with the milk (rotation of Figure-Ground array), no disruption of felicity results.

In short, non-transposed pointing gestures such as this show rotation sensi- *utterances* tivities that fit the profile of the Direct Frame of Reference. This is so because the speaker/gesturer in such cases functions as the Ground of his or her own deictic locative utterance (by definition of the semantics of such forms), and also as the Anchor of his or her own pointing gesture (by definition of the physiology of such gestures, and of our understanding of the notion of "pointing" as an indicator of the search-space vector to be followed in order to locate Figure). Such gestures must therefore be Binary in the typology (Anchor = Ground), and indeed Direct (Anchor/Ground = speech participant).

In deictic transposition, however, (Bühler 1934) an entity other than the speaker/gesturer is assigned to the role of Ground. Such transposition may separate Ground from Anchor and bring us into the Ternary portion of the Frame of Reference Typology.[16] If such cases are examined for the nature of their Anchor, through observation of their rotation sensitivities, they can, as we have seen (Haviland 1993, Levinson 2003) be assigned Relative, Absolute, or other positions within the Frame of Reference typology.[17]

In sum, although without the gesture, speech involving deictic demonstration may not contain vector information and therefore display no rotation sensitivities, the multi-modal gesture-speech pairing does supply such information. The examination of rotation sensitivities in such pairings allows us to discover the identity of the Anchor in them, and therefore the Frame of Reference category to which they belong. The placement of pointing gestures within the categories of the Frame of Reference typology has already begun (Levinson 2003). But the identification of the Direct Frame of Reference as distinct from other Binary Frame of Reference cases is essential to the full application of the typology to pointing gestures. Without it, it is impossible to assign a place within the Frame of Reference typology to the simplest kind of non-transposed locational gestures, and the existing understanding of Relative and Absolute pointing gestures remains incomplete.

Concluding summary

Frame of Reference analyses are explicitly intended to provide a template for the consideration not only of linguistic but also of conceptual construals of speakers. Evidence for the particular Frame of Reference that individual speakers entertain can be drawn from their speech, but also — and even in the absence of speech — from their actions in solving spatial puzzles, and from their gestures. Researchers have extracted such evidence by designing task situations for both gesture and problem-solving in which the diagnostic distinctions in rotation sensitivities under the different Frames of Reference can be observed. But we have now seen that the erstwhile Intrinsic frame of reference as defined by rotation sensitivity did not, as claimed, subsume cases in which the speaker was Ground/Anchor. For that, we needed to specify an additional set of rotation sensitivities, filling out the chart in Table 4 to its logical capacity. A fourth Frame of Reference, dubbed the Direct Frame, was recognized. This Direct Frame of Reference in turn furnished the basis for building a more complete understanding of the pointing gestures already being discussed in the Frame of Reference literature. Overall, unless the discussion of Relative and Absolute non-linguistic problem-solving strategies and gesture types is based on an understanding of their relation to Direct expressions

in the same modalities, it will remain incomplete and vulnerable to confusion and misinterpretation.

Wider implications

> *I repeat, there is no phonetic deictic sign that could do without the gesture or a sensory guide equivalent to the gesture or finally an orientation convention that takes their place.*
> Karl Bühler 1934, p. 108

Pointing gestures with deictic demonstratives execute the same semantic function (providing vector information about the search space from Ground to Figure) as do the lexical spatial relators (translation equivalents of 'left', 'beside', 'north' etc.) that have been traditionally studied in Frame of Reference typologies. If paired with the use of a pointing gesture, speech-situated locatives and demonstratives can be considered to have a vector-specifying Anchor, and therefore to be properly considered within the Frame of Reference typology. Bühler (1934) proposes a theory of language in which commonalities between pointing gestures and lexical spatial relators serve as the pivotal point for the idea that spatial (and other) forms of language take their origins from pointing gestures.

Arguments from language acquisition, cross-linguistic distribution, and gesture transposition all suggest that the Direct Frame of Reference may indeed have a degree of psychological primacy over the other Frames. The fact that special methods were employed in the early linguistic research into Frame of Reference to suppress deictic forms in hopes of eliciting lexicalized spatial relators suggests that these latter were indeed, as Bühler suggests, understood to be functional substitutes — or even dispreferred replacements — for linguistic and gestural demonstration. In studies of child language, use of what is here called the Direct Frame of Reference has long been recognized as a form of spatial language acquired especially early in child development, with even Object-Centered locutions appearing only later in acquisition (Tanz 1980). Ternary Frames of Reference — if present at all — are acquired latest of all, often requiring formal instruction (Johnston and Slobin 1979, de León 1994). Pointing gestures within the Direct Frame appear to constitute the template from which gestures in the other Frames are transposed. Finally, we can note that, like the Object-Centered Frame of Reference, the Direct Frame is apparently universal across languages and cultures (Danziger 1999). It would surely be perverse to exclude this basic component and proposed point of origin of the Frame of Reference architecture from our discussions of the types of spatial language, of language and conceptualisation in the domain of space, and of the application of the notion of Frame of Reference to pointing gestures. In the longest run, the present clarification of the Frame of Reference typology will

provide an opening to re-formulate more explicitly the relationship of deixis to the lexicon, and of the speaker's body to his or her linguistically mediated position in physical and conceptual space.

Notes

* Some of the ideas put forward in this paper were germinated in conversation with the members of the Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics, in particular Stephen C. Levinson and David P. Wilkins. My thanks also go to two anonymous reviewers, who helped to make this a better paper, and to Jürgen Bohnemeyer for his discussion of the ideas and terminology. Much of the discussion of deixis and demonstration is presaged by Bühler 1934.

1. I would like to stress that the critique presented here of the three-frame paradigm is also a critique of my own earlier thinking, and should not be taken as a failure to appreciate the seminal achievements of those who have used the three-frame typology over the past two decades.

2. In principle, other distinctions among locutions could of course have been chosen as well. Terrill and Burenhult (2008), for example, choose to elaborate a spatial reference typology based on the question of whether or not reference is made in terms of a part of the Ground object. So "north of the spout of the kettle" and "at the spout of the kettle" are classified together, while "north of the kettle" is classified separately. Such a classification however does not yield the differential sensitivities to rotation that have been so important in forming a link between linguistic and conceptual frames of reference (Pederson 1998, Levinson 1996).

3. In this example, the Anchor is located at the Speaker. Note that an Anchor located at the Addressee also gives rise to a Relative Frame of Reference.

4. It is clear for example (see Levinson et al. 2002) that a failure to distinguish between what are here called Binary and Ternary versions of the Allocentric led Li and Gleitman (2002) to misinterpret their otherwise interesting experimental results.

5. For "Ternary" and "Binary", for example, we could perhaps substitute "Projective" and "Topological" from developmental psychology (Piaget and Inhelder 1963 [1948]), "Orientation-Bound" and "Orientation-Free" from perceptual psychology (Just and Carpenter 1985, see Danziger 1996a), or "Field-Independent" and "Field-Dependent" from the literature on cognitive style (Witkin et al. 1977).

6. For this illustration it is important to stipulate that if the speech participant Anchor is based in the speaker, the speech participant Ground should be located in the addressee. If both Ground and Anchor are located in the same speech participant (as in 'The milk is in front of me.' (redundantly 'from my perspective')) then Anchor and Ground are found in the same entity and the Frame of Reference is by definition Binary and not Ternary (see example (9)). At this stage in the argument we are only interested in exemplifying Ternary Frames of Reference.

7. One reviewer objects to the use of the preposition "at" in this kind of example, claiming that it specifies no vector from Ground to Figure and therefore that the locution cannot be considered

in the Frame of Reference discussion. But the usage is in that respect strictly parallel to that of "at the spout of the kettle", a type of locution which has been widely accepted to exemplify the Intrinsic Frame of Reference, precisely in that the specified part of the Ground object provides the vector information which narrows the search space from Ground to Figure.

8. It should be clear that since it is Anchor and not Ground that defines the Frame of Reference type, there is no need to review the Frame of Reference assignment of examples (4)–(7). These examples illustrate usages in which Anchor is independent of Ground (Ternary Frames of Reference). The proposed new distinction only concerns cases where Anchor is conflated with Ground. By definition, this occurs only in Binary and not in Ternary Frames of Reference.

9. Nature and location of Ground as distinct from Anchor may of course be of great interest in discussions of spatial language that are concerned with issues other than those of Frame of Reference assignment (cf. Bickel 2000).

10. Relative rebuilders in effect obey the instructions by neutralizing the effect of their own rotation. They do so by rebuilding a version of the original that is also rotated, thereby preserving the spatial relationships of scene components to their own body. Absolute rebuilders do not need to adopt such a strategy, since their own rotation has no effect on the mental representation which they use to represent the scene.

11. Lest "front" cause confusion due to its near-grammaticalization, consider the unlikely but semantically similar possible sentence "The milk is at my belly-button". Then turn the speaker 180 degrees (without moving the milk), and note that the description has been rendered false.

12. Distinguishing the right and left sides of a non-speech-participant Ground depends upon applying a Ternary Frame of Reference, since an Anchor other than the Ground (namely, the speech participant's own right and left) must be brought to bear. But distinguishing the right and left sides of one's own body does not require departure from the Binary mode, since one's body serves as both Ground and Anchor of the computed relation. Li et al. (2005) report experiments that require Tzeltal consultants to compute only the latter (Direct) and not the former (Relative) kinds of relations.

13. Thanks to an anonymous reviewer for cogently reinforcing this point.

14. Languages may exhibit forms of linguistic demonstration that provide more semantic information than English does about the search vector required from Ground to Figure. In such cases, we might expect pairing with a pointing gesture to be less obligatory, although a greater degree of precision in vector specification might be achieved with a pointing gesture.

15. ~~This example concerns a deictic locative.~~ Pointing gestures are also deeply associated with other shifter forms such as demonstrative pronouns. As a reviewer points out, this raises the possibility that Frames of Reference could be nested within the same sentence ("The milk is behind this kettle."). This is of course also true for non-deictic locators: ("The milk is behind the kettle that is beside the fridge."). In some languages (cf. Danziger 1994), demonstrative pronouns are in fact transparently formed as relative clauses based on the deictic locatives (i.e. "this kettle" is phrased as "the kettle that is here" and so on).

16. The complexities of deictic transposition are many and intricate, and a significant literature exists in which they are discussed (see for example Bickel 2000). My aim here is not to pursue

that discussion, but to point out that in the application of the spatial Frame of Reference typology to pointing gestures, an adequate understanding even of the *non-transposed* case has so far been absent.

17. If Anchor remains speaker under such transposition, then the Frame of Reference is Relative (Ground separate from Anchor; Anchor in speech participant). If Anchor is removed from the speaker without being conflated with the transposed Ground, then the Absolute Frame of Reference applies. If Anchor is instead identified with the transposed Ground, the gesture can be classified as Object-Centered (see Kita et al. 2001, Danziger et al. ms.).

References

Baayen, R. H. & E. Danziger (eds.). 1994. *Max Planck Institute for Psycholinguistics Annual Report 14 (1993)*. Nijmegen: Max Planck Institute for Psycholinguistics.

Bickel, Balthasar. 2000. Deictic transposition and referential practice in Belhare. *Journal of Linguistic Anthropology* 10(2): 224–227.

Bühler, Karl. 1990 [1934]. *Theory of Language: The Representational Function of Language* (English version translated by D.F. Goodwin). Amsterdam, Philadelphia: John Benjamins. [Foundations of Semiotics Vol. 25].

Clark, Herbert H., R. Schreuder & S. Buttrick. 1983. Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior* 22: 245–258.

Danziger, Eve. 2001. Cross-cultural studies in language and thought: Is there a metalanguage? *The Psychology of Cultural Experience*. Carmella C. Moore and Holly F. Mathews (eds.), 199–222. Cambridge: University Press. [Publications of the Society for Psychological Anthropology].

Danziger, Eve. 1999. Language, space and sociolect: Cognitive correlates of gendered speech in Mopan Maya. *Language Diversity and Cognitive Representations*. Catherine Fuchs and Stéphane Robert (eds.), 85–106. Amsterdam: Benjamins.

Danziger, Eve. 1996. Parts and their counter-parts: Social and spatial relationships in Mopan Maya. *Man* 2(1): 67–82.

Danziger, Eve. 1994. Out of sight, out of mind: Person, perception and function in Mopan Maya spatial deixis. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 32(4/5): 885–907.

Danziger, Eve, S. Kita & C. Stolz. Submitted. The conversational gestures of culture-specific cognition: Language, space and semiotic relativity in two Mayan speech communities. *Current Anthropology*.

De León, Lourdes. 1994. Exploration in the acquisition of geocentric location by Tzeltal children. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 32 (4/5): 857–884.

De Ruiter, Jan Peter & D. P. Wilkins. 1998. The synchronization of gesture and speech in Dutch and Arrernte (an Australian aboriginal language): A cross-cultural comparison. *Oralité et Gestualité: Communication Multimodale, Interaction*, Serge Santi et al. (eds.), 603–607. Paris: Harmattan.

Enfield, Nicholas J., S. Kita & J.P. de Ruiter. 2007. Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics* 39: 1722–1741.

Haviland, John B. 1998. Guugu Yimithirr cardinal directions. *Ethos* 26(1): 25–47.

Haviland, John B. 1993. Anchoring, iconicity, and orientation in Guugu Yimithirr pointing gestures. *Journal of Linguistic Anthropology* 3: 3–45.

Johnston, Judith R. & Dan Slobin 1979. The Development of Locative Expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language* 6:529–545.

Just, M. A. & P.A. Carpenter. 1985. Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review* 92(2): 137–172.

Kita, Sotaro, E. Danziger & C. Stolz 2001. Cultural specificity of spatial schemas as manifested in spontaneous gestures. *Spatial Schemas in Abstract Thought.* Meredith Gattis (ed.), 115–146. Cambridge: MIT Press.

Levelt, Willem J. M. 1984. Some perceptual limitations on talking about space. *Limits in Perception,* A. van Doorn, W. de Grind, and J. Koenderink (eds.), 323–358. Utrecht: VNU Science Press.

Levinson, Stephen C. 2003. *Space in Language and Cognition.* Cambridge: University Press.

Levinson, Stephen C. 1996. Frames of reference and Molyneux's question: Crosslinguistic evidence. *Language and Space,* Paul Bloom, Mary A. Peterson, Lynn Nadel & Merrill F. Garrett (eds.), 109–169. Cambridge: MIT Press.

Levinson, Stephen C., S. Kita, D. B.M. Haun & B. H. Rasch. 2002. Returning the tables: Language affects spatial reasoning. *Cognition* 84: 155–188.

Li, Peggy, L. Abarbanell & A. Papafragou, A. 2005. Spatial Reasoning Skills in Tenejapan Mayans. *Proceedings from the 27th Annual Meeting of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Li, Peggy and L. Gleitman. 2002. Turning the tables: Language and spatial reasoning. *Cognition* 83: 265–294.

Majid, Asifa, M. Bowerman, S. Kita, D. B. M. Haun & S.C. Levinson. 2004. Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8: 108–114.

Pederson, Eric, E. Danziger, S. C. Levinson, S. Kita, G. Senft & D. Wilkins. 1998. Semantic typology and spatial conceptualization. *Language* 74(3): 557–589.

Piaget, Jean & B. Inhelder. 1963 [1948]. *The Child's Conception of Space* (English version translated by F. Langdon and J. Lunzer). London: Routledge and Kegan Paul.

Rock, I. 1990. The frame of reference. *The Legacy of Solomon Asch: Essays in Cognition and Social Psychology,* I. Rock (ed.), 243–271. Hillsdale, NJ: Lawrence Erlbaum Associates.

Talmy, Len. 1983. How language structures space. *Spatial Orientation: Theory, Research, and Application,* H. Pick and L. Acredolo (eds.). New York: Plenum.

Tanz, Christine. 1980. *Studies in the Acquisition of Deictic Terms.* Cambridge: Cambridge University Press.

Terrill, Angela & N. Burenhult. 2008. Orientation as a strategy of spatial reference. *Studies in Language* 32(1): 93–136.

Whorf, Benjamin L. 1939 [1956]. The Relation of Habitual Thought and Behavior to Language. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf,* John B. Carroll, (ed.), 134–159. Cambridge: MIT Press.

Witkin, H.A., C. A. Moore, D. R. Goodenough & P. W. Cox. 1977. Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research* 47(1): 1–64.

Author's address

Eve Danziger
Department of Anthropology
100 Brooks Hall
University of Virginia
Charlottesville, VA 22903
United States

ed8c@virginia.edu