26

What Does It Mean to Compare Language and Gesture? Modalities and Contrasts

EVE SWEETSER

University of California, Berkeley

> Perhaps we keep finding iconicity because there is no other way for a semiotic system to be created and used by human beings without a close fit between form and function. After all, is it possible to make a mold for a statue that does not conform to the shape and dimensions and substance of the statue?

Dan I. Slobin (2005, p. 321)

*I*n this paper I would like to reexamine some of the traditional dichotomies between language and gesture. In order to do so, it will be necessary to consider a three-way contrast—spoken languages, signed languages, and gesture. Without this three-way comparison, we risk collapsing contrasts between visual and auditory media with contrasts between linguistic structure and co-linguistic gestural structure. Such a comparison clearly belongs in this volume because Dan Slobin's work on Thinking for Speaking has provided a crucial impetus to the research which feeds my new evaluation—both his own work on spoken and signed language, and the new perspectives on co-speech gesture which have been inspired by that work, not to mention his general intellectual influence on my work for the last 30 years. Dan has never been never afraid to cross boundaries between modalities—or to be skeptical about accepted dichotomies. So I hope readers will see this paper as being in his tradition.

Returning to our topic: what kinds of contrasts are there between language and gesture? First of all, many researchers agree that language is *conventional*, while spontaneous co-speech gesture is *non-conventional and flexible* (though "quotable gestures" are conventional). In gesturing about a tree, for example, a gesturer might trace the fat tube of the trunk with two "C" hands forming a circle, while another gesturer might hold up a forearm as trunk with fingers as branches. Neither would be "wrong," and indeed they might successfully highlight different aspects of the tree's structure. However, Hong Kong Sign Language conventionally and lexically represents TREE in the first way, while the American Sign Language lexical sign for TREE is based on the second strategy. Neither language's word for TREE is correct in the other language.

Similarly, there seems some consensus that language is *compositional, discrete, and analytic*, while gesture is *global and synthetic*. Further, as McNeill (1992, 2005) and others have documented, language is a more *consciously* monitored channel relative to gesture. Speakers are often unaware of gesturing at all, while non-sleeping language-users are aware at least of whether they are producing linguistic forms, even if they cannot be conscious of each lexical and morphological choice involved in production. Overall there are also general feelings that language is somehow *abstract*,

while gesture is "*concrete*" not only in physical performance (we move concrete body parts in space when we gesture) but in the kinds of imagery used and meanings represented.

An obvious question is, to what extent these contrasts are contrasts between modalities (aural/oral vs. visual/gestural), and to what extent they are contrasts between Language (linguistic structure) and Non-Language. It seems plausible, for example, to assume that conscious monitoring is a characteristic of the primary linguistic channel, whatever channel that is. Signers know that they are signing, as surely as speakers know that they are speaking, even though signers and speakers alike may be unaware of ancillary non-linguistic actions—such as scratching their heads, or even iconic co-speech gestures. Full conventionality likewise seems prima facie to be a characteristic of the linguistic channel per se, and not of one modality or the other (though there will be more discussion of quotable gestures below). There may be more than one iconic gesture which can be used to express a particular concept, but (as we saw above with the ASL and HKSL signs for TREE) only one is the conventional sign for some particular meaning in a particular signed language.

Concreteness might initially seem more likely to be a characteristic of the medium rather than of Language vs. Gesture. That is, a linguistic sign is as much a concrete action as a gesture accompanying spoken language. And because of the heavily iconic nature of the visual/gestural medium, in many cases signed languages indicate abstract meanings metonymically or metaphorically. The ASL sign for TIME involves directing the dominant hand toward the back of the non-dominant hand's wrist—this sign is iconic for a wrist-watch, and thus metonymic for the time which a watch measures. The ASL time-line (paralleled in other signed languages as well as in many spoken-language gesture systems) metaphorically maps space onto time: the space in front of the signer maps onto future time, the front of the signer's body onto the present, and the space behind the signer onto the past.

Analyticity is perhaps the most complex issue. The claim is of course that it is a characteristic of Language. Complex linguistic communication depends on the fact that putting signs together produces a meaning which is conventionally related to the conventional meanings of the individual signs. Compositionality in morphology and syntax are generally seen as the most basic requirement of full Language—nor shall I question this. But, as we shall see, gesture can be complex and compositional too.

Let us examine these issues in turn. Part of my argument will be that since gesture is a very heterogeneous category, different subcategories of gesture may in fact relate differently to these parameters. We will also need to keep in mind that speakers' gesture processes are themselves multimodal. Although an English speaker saying *Oh* or *Ow* is using a learned linguistic form, a scream is something a baby does not have to learn; it is a vocal gesture (Liddell 2000, 2003).

Throughout my discussion, I shall be using the Mental Spaces framework and Blending theory of Fauconnier and Turner (Fauconnier, 1998; Fauconnier & Turner, 1996, 2001) as background to this work. This framework underlies work on ASL by Liddell (1995, 1998, 2000, 2003), Taub (2001), and Dudis (2004), and also some recent work on gesture (Sweetser, 1998, 2003; Parrill & Sweetser, 2004; Sizemore & Sweetser, in press). Blending theory has been a crucial tool in furthering our understanding of metaphor and iconicity, in spoken and signed languages and in gesture.

ANALYTICITY AND COMPOSITIONALITY

Let us reexamine McNeill's classic example of the gesture about the tree being bent back (McNeill, 2005, chap. 1; McNeill, 2000), which is represented by a hand apparently grasping an invisible something and bending it backwards. Crucially, says McNeill, it's the *whole* gesture that maps globally onto the whole represented situation, not compositional bits of the gesture that get composed into a whole meaning. When we compare this with the treatment of signed language we can see that similar issues arise there. For example, in the ASL sign for TREE, the entire hand and forearm map onto the shape of a tree, with the extended fingers as branches. On the other hand, it is only the global mapping which determines the match with a tree: in another context, one of those branch-representing extended fingers, the index finger, can represent a standing person; or the index and middle

fingers extended downward can represent a bipedal animal walking (so here the index finger represents a leg). Yet these mappings are not "global" in the sense of being unanalyzed: Taub (2001) enumerates the mappings between the 5 hand and the imagined tree, in describing the TREE sign. The fingers are the branches, the arm is the trunk, the non-dominant arm (which is horizontal below the elbow of the "tree" hand) is the ground, and so on. And it is normal to think of hand shape, location, and motion as independent components of an ASL sign. But this kind of structure happens often in gesture as well. It happens perhaps most in conventional "quotable" gestures; Kendon's (2004) analyses of hand shapes' meanings and Calbris' (1990) analyses of French quotable gestures have shown meanings attributed separately to hand shape, location, and so on. In short, either gesture is (at least sometimes) more compositional than we had thought, or signed languages are less so.

McNeill's major point, however, is that the meaning of the bent-back hand as an agent bending over a tree is dependent on the larger discourse—it is top-down, rather than bottom-up interpretation, "global" in the sense that the meaning of the whole (a tree being bent back) is what gives meaning to the parts (the hand shape and the motion). This is a very important point. However, in itself it does not automatically establish that gesture is more synthetic or more global in its interpretation than language is. One of the most difficult problems for language processing systems has always been how to get the right "top-down" interpretive components into place for understanding texts, which are by no means interpretable entirely by taking the meanings of the pieces and composing them to get a meaning for the whole. Context is crucial in linguistic interpretation. Signed language "classifiers" are a case in point. In ASL, an extended index finger may represent different long thin things, depending on the need—or it may be a deictic point to a locus, rather than iconically representing anything; the signer's addressee will know what is represented from the context.

The question we should be asking, then, is whether the top-down meaning structure seen in gesture is different in kind from the sort of discourse-dependence and top-down interpretation which happens all the time in spoken and signed languages. My feeling is that yes, it is different—and that the difference lies primarily not in what is *present* in gesture, but in what is *absent* from it: the *incompleteness* of co-speech gesture. That is, since co-linguistic gesture presumes that the primary information channel is the linguistic one, there is no pressure for content-expressing gesture to represent the described situation as fully as in language. Of course *all* representations are incomplete: perhaps ironically, one of the few points on which formalist, functionalist and cognitivist linguists have converged in the last 20 years is that language never does more than give small but effective prompts toward construction of much richer cognitive structures. But I mean something specific here, by *incomplete*: I mean that the speaker/signer does not intend co-linguistic gesture to do the "full job" of prompting the desired cognitive representation. Although gestural iconic depiction can be quite rich, it does not attempt to "cover the ground" of the content—there is no reason for it to be complete. Naturally, then, gesture needs context from the primary communicative channel to be understood, while the linguistic channel, though often comparatively sparse in its cueing of representations, is charged with cueing mental structures covering the full content described—and hence needs less *content* context from other channels. Of course, language may often require facial expression and (vocal or non-vocal) gestural cueing to create the relevant *discourse* context—more on this later.

CONCRETENESS

It is not always clear exactly what analysts mean when they say that gestures are more concrete than language. Spoken language is a set of muscularly performed routines, which are physically experienced by the auditory systems of listeners. Signed language is another set of muscularly performed routines, this time performed by most of the same muscles (arms, hands, face, body core) which are involved in gesture. However, one thing that "more concrete" *could* mean in this context is that there is iconic structure involving what Liddell (1995, 1998, 2000, 2003) calls a *grounded blend*—one input to the blend is the physical structure of the Real Space of the speaker's body in

the communicative setting. The English word *drive* does not involve this kind of Real-Space blend, but the ASL sign DRIVE does: it involves two hands grasping an imaginary steering wheel—the real physical presence of the two grasping hands in Real Space is blended with an imagined driving scenario involving a steering wheel. And gestural representations of driving canonically also involve such blends (LeBaron & Streeck, 2000).

In that case, however, there is no absolute line dividing such structures in signed languages from the ones found in gesture—or indeed, dividing those structures from vocal mimicry, where accent or tone of voice iconically represents accent or tone of voice. And perhaps that makes sense—the problem may lie in the fact that such vocal mimicry was not among the phenomena primarily considered by those who labeled gesture more "concrete" than language. Some new questions which need examination, therefore, would seem to be in what ways auditory Real-Space blends are different from visual ones, or how more conventional Real-Space blends differ from more flexible unconventional ones. Very little systematic work has been done on auditory Real-Space blends, which are a potentially rich domain for research, as Liddell has suggested.

It should be noted, in approaching these questions, that deictic aspects of gesture are "concrete" in the sense of involving Real-Space blends: displaced gestural deixis (cf. Hanks, 1990; Haviland, 1993, 2000; Levinson, 2003; Kita, 2003) allows a speaker/gesturer to point in the Real Space *as if* she were in an imagined space. A speaker might say *Go into my office and the light switch is over here on the left, next to the door*—gesturing to the left, although in English there is no need for the Real-Space gesture to be directionally aligned with the actual space of the described office. In this respect, however, deixis is different from iconicity: it is inherently linked to the Real Space, even in spoken language, and even when there is nothing iconic about the forms used (pronouns and words like *this* or *that* do not iconically represent their meaning). One can only interpret deictic forms via knowing who, where, and when a Speaker is—that is, by identifying her Real Space. Of course, spoken-language deictic forms refer to the Real Space more implicitly, in a less profiled manner—they *refer* to the profiled entity (for example, *that* thing) via its place in a network relating it to the Real-Space speaker—while gesture and signed language both seem to explicitly make use of the Real Space, pointing physically in it.

We should also want to know, therefore, how different deictic Real-Space blends are from iconic ones—and what cognitive difference it makes to access Real-Space blends more explicitly, as opposed to more implicitly. None of this seems likely to give us a simple, unified "gesture versus language" picture of the world.

CONVENTIONALITY

Much co-speech gesture is more flexible and less conventional than speech. But not all. The reason why we distinguish conventional co-speech gestures from language, in spoken communities, is the difference in medium, not the difference in conventionality. We know that "emblems" or "quotable gestures" (Kendon, 1990, 2004) are just as conventional as linguistic signs. When an English speaker holds up a hand in the "F" hand shape (thumb and index forming a circle, the other fingers extended), other English speakers know that this conventionally means "perfect, just right". This gesture is even polysemous in a rather predictable pattern paralleling linguistic polysemy: it could mean "I'm fine," "(you did a) Fine job," or "(the situation is) Perfect or precisely right," among other things. No accompanying speech is needed—the form carries conventional meaning independently. Speakers of other languages, I am told, may conventionally ascribe other unrelated meanings to the same gesture—in some cases, objectionable and insulting meanings. Coming to signed languages, of course many analysts have pointed out (Liddell, 2000, 2003 in particular) that signers gesture; but we would not be able to talk about "quotable gestures" as a special category in signed languages, since the language consists precisely of a range of conventional signs performed solely in the visual-gestural medium.

Returning to the distinction between "flexible" non-lexicalized gestures and conventional linguistic forms, it should be noted that conventionality by its very nature limits the relationship of the form to one particular contextual meaning. You can't have a separate conventional form for every shape of every individual object or path that you encounter. So an unconventional iconic form might be uniquely capable of expressing some particular detailed meaning. Also, it is easier to learn conventional form-meaning correspondences when they have some compositional regularity (a regular past tense is simpler to acquire than an irregular one). There is an extra profit to formal analyticity if it gives the learner regularity and productivity. Together with the *incompleteness* of co-speech gestural representation, this may help to account for perceived differences of globality vs. analyticity between more and less conventionalized forms—and hence between flexible co-speech gesture and language.

It remains somewhat of a mystery exactly how conventionalization affects the meaning of human symbolic forms. Our understanding of ritual, for example, would suggest that repeated and conventional Real-Space blends can have cognitive power which derives precisely from their conventionality. Perhaps some of the same forces are involved when a child demands a fourth sequential reading of the same story, despite having heard it dozens of times before. On the other hand, clichés lose the communicative power associated with "freshness" and innovation, as they become conventionalized larger units; they may (Bybee, 2007) show reduced phonological structure in production, and increasingly "subjective" or discourse-related meaning, as well. Linguists don't know how to make sense of both these two facts together; but at any rate they seem to fit into a broader psychological framework that includes both habituation and dishabituation as responses to stimulus repetition.

Flexible co-speech gesture shows strong conventional aspects as well as flexibility, however: there are culture-specific regularities in how people gesture about both concrete and abstract domains. Local catchments arise in any speaking/gesturing group—that is, speakers pick up each other's gestural uses for particular meanings (cf. LeBaron & Streeck, 2000; McNeill & Duncan, 2000) and re-use them. McNeill's well-known Snow White experiment (McNeill, 1992) showed how speedy this process is and how fast it can become "language-like" or symbolic in character; see also Goldin-Meadow (1993, 2003). Metaphor and iconicity can be conventional and culture-specific, both in language and in gesture. One classic iconic example, cited above, is the difference between the ASL sign for TREE (which uses the hand and arm as an iconic symbol of a tree, with the fingers as the branches) and the Hong Kong Sign Language sign for TREE, which uses the two hands to "trace" the form of a round vertical trunk. When we use the term *motivation*, we typically mean that iconicity or productivity is present along with convention, in defining a particular form-meaning relationship.

In metaphor as well, culturally specific metaphors will be conventionally represented in gesture. Núñez and Sweetser (2006) describe how speakers of Aymara (a language of the Andean highlands) gesture forward when referring to the past and backward when referring to the future, gestural structures which reflect their unusual cultural and linguistic metaphors for time. Speakers of English (and the many other languages where the future is metaphorically in front of the speaker) gesture in the opposite directions (cf. Lakoff & Johnson, 1980, 1999, on more general cognitive models of metaphor in language).

We further know from McNeill (1992), McNeill and Duncan (2000), and others that speakers of different languages regularly co-time their gestures with their speech in different language-specific patterns. We even know that they represent or leave unspecified different aspects of meaning, in the gestural medium, depending on their language; Spanish speakers do not tend to represent manner of motion in the same way that English speakers do, in describing motion scenes. This parallels the differences between the relevant linguistic structures, where Spanish speakers tend to omit description of manner (which is not normally part of the lexical verb meaning in Spanish), while English speakers include it naturally because their lexical verbs include manner (see Slobin, 1987, 1996, 2000).

Of course, no care-giver ever says to a child "don't put the stroke of your gesture on the object, put it on the verb" or "put in manner of motion in your gesture"—because of course the care-givers are themselves unconscious of the relevant patterns. (Care-givers also cannot overtly instruct

children in many aspects of linguistic grammar, for the same reason.) Yet we know little about exactly how different culture-specific gestural patterns affect hearers/viewers, outside the world of emblems (which have been a focus for manuals and tourist advice since long before De Jorio or Desmond Morris (cf. Kendon, 2004)). How exactly do an Anglo-American gesturer's co-timing system and manner system communicate *differently* with a Chinese or a Spanish speaker/gesturer than with a fellow English-speaker/gesturer? It does not seem that everyday speakers can consciously put a finger on the differences in patterns at the level of co-timing—though it does seem that a large gesture space makes gesture salient to members of smaller-gesture-space cultures, who then say (e.g.) that Italians "gesture a lot"; conversely, Italians may feel that English speakers don't gesture much.

And where are the universals in gesture structure? To take one specific example, discussed in Sizemore and Sweetser (in press), Anglo-American speakers regularly represent the progress of non-deictic processes either horizontally across the body (left-to-right or right-to-left), or outward from the body, or both (diagonally). (Deictic motion processes are mapped onto directional gestures which fit the deictic viewpoint taken.) Thus, it is possible to gesture toward oneself as one describes deictic motion toward oneself (when saying *She came up to me*, for example). But when describing how to get a degree at Berkeley, or how to dial a telephone number, the gestures will move to locations progressively farther outward from the speaker's body (not toward the body) as the speaker mentions progressively later stages in the process. Is this a culture-specific fact about English gesturers, or is it more general? In my view, it is a potentially more general phenomenon, worthy of comparative crosscultural investigation; I believe it may stem from the fact that gesture is inherently, necessarily viewpointed in a way that speech is not. Not only is it produced *by someone*, as speech is (providing an essential deictic base for speech also), but it is produced in motion which necessarily emanates *from* the trunk of the body—from the unextended rest position of the arms, for example—to farther extended locations. The trunk is a source of bodily gestural motion.

We therefore need more systematic examination of conventional aspects of gesture, as well as of less conventional aspects of language, to distinguish the effects of conventionality from those of (for example) analyticity. And we need to keep in mind that motivation coexists with convention at every level of language and gesture: iconic motivation is not deterministic, but necessarily chooses particular aspects of the two mapped domains to highlight, as in HKSL TREE (outlining the trunk) and ASL TREE (the hand/forearm representing the tree).

ICONICITY, ANALYTICITY, AND CONVENTION

We have said that there are reasons why more iconic structures may indeed be more "global" or "top-down" than less iconic ones. Both gesture and signed languages centrally involve Real-Space iconic mappings in a way that spoken language does not. That is, they both involve capitalizing on interpretation of motion or location in Real Space as meaningful in the domain of expressed content. They therefore cannot have quite as analytic form-meaning relationships as spoken languages, in general. McNeill's "bending back the tree" example, mentioned above, is such a case: as we have said, you have to know the meaning of the whole in order to know the meaning of the parts. For gesture, we have suggested that incompleteness of gestural representation is another reason for globality in interpretation.

However, gesture does seem to have analytic dimensions, or at least, separable parameters of iconic meaning. For example, take Cienki's (1998) examples of gestures accompanying phrases like *good grade* and *do the right thing* versus *bad grade* and *cheat*: the speaker gestured higher for *good grade* and lower for *bad grade*, higher for *do the right thing* and lower for *cheat*. It is clear that the HEIGHT of the location of the gesture corresponds to positive value on an abstract scale, whether of grades or of morality. Clearly this is not the sole meaning of vertical location of gestures, which could also be locating imagined entities in physical space; in language, too, the words *high* and *low* primarily mean literal height, and only secondarily abstract metaphorical moral or grade "height."

But regularly and conventionally, just as with language, the up-down location of a gesture is in itself a parameter distinct from shape of the hand, or even identity of the scale which may be involved.

Force-dynamic structures such as gestural ones are perceived "as wholes"—grasping something, pointing at something. But they are also necessarily parametrically analyzable, presumably by the parameters involved in experientially differentiating them from each other (direction, location, and hand shape, for example). The fact that these parameters are co-performed with each other (a hand has to have both location and hand shape at every moment) long masked their separability as analytic parameters in signed languages, but of course linguists now accept that two signs can differ just by motion, or just by hand shape, or just by location. (Similarly, morphemes with suprasegmental or discontinuous forms in spoken languages were late in being recognized, just because you couldn't neatly assign each of them to a unique temporally adjacent subsequence of the written representation.) But much of the same analytic apparatus developed for cotemporal performance in signed languages is used in description of gesture, where hand shape is regularly noted (and potentially generalized over) independently of hand location.

And on the other hand, we can examine a spoken-language iconic example for comparison. Taub (2001) compares ASL iconicity with that of English sound-symbolic words such as *ding*, *pow, bang*, or *meow*. She notes that the English phonetics of *ding* have an abrupt onset (d) followed by a gradual offset (spelled "ng"); and that the high front vowel I has higher frequency of a crucial formant than, for example, O. Actual ringing of a bell produces a spectrogram with an abrupt onset and a gradual offset. And of course some bells have higher pitches than others —and indeed *ding, ding* represents a higher-pitched bell than *dong, dong*.

Now, that doesn't mean that we can simply say that "d" on its own, in isolation, represents the start of a bell-stroke, or "ng" the latter part of the stroke—these mappings have iconic naturalness *within* the larger form. (The same point can be made with respect to the "m" and "w" of *meow*, and so on.) *Dig* and *dog* both start with the same phoneme as ding and dong, and are contrastive by the same I vs. O vowel contrast: but their initial d does not refer to abrupt sound onset, nor is their I:O contrast interpreted as referring to pitch difference. So the iconic interpretation of the linguistic form *ding* is in a sense necessarily global, in that the individual (phonemic-level) submappings exist by coherence with each other as parts of the broader morpheme-level iconic mapping. We can identify specific parameters and segments and their mappings to specific parameters and aspects of the meaning—but such identification is necessarily relative to the whole mapping of form onto meaning. More of signed language is like this than spoken language; and a great deal of gesture is recognized by analysts to be like this.

For example, take a gesture of handing something to someone—an outstretched palm-up hand, perhaps with the thumb above the palm helping to "hold" the invisible surrogate object. Except as a part of the movement outward from the gesturer's body, it would not necessarily be interpreted as "giving"—it would just be "holding," or perhaps something else entirely. And further, the motion has to occur in the right temporal sequence: that is, as a single global motion from close to the speaker to a location closer to the imagined recipient. Movement in the opposite direction would either invert the meaning (someone giving the gesturer something), or make it meaningless. Location at specific points along the path of giving, without the global motion—and in particular, if the points were not in sequence!—would no longer convey the "giving" meaning. This is as true of gesture as it is of the ASL sign GIVE, which is based on the same iconic structure. *Temporal* and *aspectual* structure, like topological structure, have to be mapped "whole" (globally), in iconic mappings—it is the *entire* motion of the hand in this iconic gesture which maps onto the *entire* motion of giving an object to someone.

So what we are really seeing, in language and in non-language, in visual and in auditory media, is that iconicity has a special character. Iconic structures necessarily involve global synthetic mappings, because both the represented and the representing spaces consist of image-schematic structure which is perceived that way. (A cat's meow, or the sound of a bell, comes as a whole—you never get just the onset or just the end.) That does not of course mean there is not analyzable sub-structure—which may be quite systematically mapped too. This special character of iconicity is, then, in my view necessarily a factor in how analytic structure can be interpreted.

THE HETEROGENEITY OF GESTURE

As many analysts at different times have noted, gesture is extremely heterogeneous. So it doesn't make sense to attribute some particular degree of "iconicity" or "conventionality" to gesture in general. Some gestures are extremely conventional (so-called "emblems"), while others are not. Some are highly decomposable analytically—in particular, those built on broad parametric construals of Real-Space blends, such as GOOD IS UP or ABSTRACT IS UP, or ONGOING PROCESSES ARE FORWARD MOTION—while others may not be.

Different kinds of blends are involved in different classes of gestures. Discourse interactional gestures (including quotables like nods) are distinct in character from representational content gestures (Bavelas, Chovil, Lowrie, & Wade, 1992; Kendon, 1995). Sizemore and Sweetser (in press) argue that interpersonal gestures are inherently structured by the establishment of personal, interpersonal, and extrapersonal spaces, which allow us to interpret (for example) gestures into the interlocutor's personal space as having discourse-regulating functions. This brings a pervasive deictic character into interpersonal gesture. Content-related gestures, literal or metaphoric, often show more depictive and iconic character, though deictic structure is crucial to their interpretation as well (as in realizing that processes are depicted as motion AWAY from the gesturer).

As has frequently been observed, a single gesture may cross these lines. Beats, the prosodic gestures which have hardly been discussed in this paper, are readily superimposed on both content and interpersonal gestures, often without changing the location and hand shape involved in those gestures; speakers may point and "beat" simultaneously, or make beats with a hand which is iconically representing the shape of an object. Similarly, Liddell's (2003) topic-maintenance "buoys" hold loci and may retain hand shape from content gestures. And in Furuyama's (2000) "imaginary origami-folding" task, when speakers reached into their interlocutors' personal gesture spaces to manipulate imagined origami paper, this gesture did indeed clearly combine representational content (depicting origami-folding) with discourse interaction; reaching into the interlocutor's space to do this clearly marked shared focus, in a way that an analogous gesture in the speaker's own gesture space would not have. As Smith (2003) and Engle (2000) have shown, it is not only imagined but real objects in the environment which can be incorporated into gestural structure, so that some gesture involves object manipulation of *visible surrogates* (there is a scene in the movie *Bend it Like Beckham* wherein one character explains the "offside" rule in soccer to another by using table implements such as a salt shaker to represent the relevant soccer field situation).

Sign language also has deictic and iconic features which are not as pervasively present in spoken languages (with the important exception of sound symbolic structures, noted above). As an example of both of these simultaneously, a palm-out hand in ASL can (as in gesture) represent a protective barrier for the speaker, while a palm-inward hand represents a protective barrier on someone else's behalf, perhaps against the speaker. The use of the hand to represent a two-dimensional barrier is iconic, as the hand is saliently two-dimensional when open. The directionality is deictic; the body is a deictic center, and the hand is naturally adapted to push things *away* from that center with the palm outward.

However, signing does *not* to my knowledge incorporate into its linguistic structure actions like those of the origami folders who reached into their interlocutors' space to manipulate the imagined paper developed by those interlocutors. Nor do conventional sign forms "reach into" the interlocutor's space as Sizemore and Sweetser's gesturers consistently did (Sizemore & Sweetser, in press). This is not to say that signers couldn't do such things in co-sign gesture—as has been pointed out, signers clearly do gesture, even though analysts may disagree as to exactly which components of their motions are linguistic and which are gestural. Similarly, signed languages appear to have special constraints with respect to representing motion past a boundary, which are not there in gesture: a gesturer can say *ran into the room* and gesture with the moving hand right past a boundary hand, while a signer has to separately represent the running and the boundary crossing (Slobin & Hoiting, 1994). So another very important area of research is the examination of the precise constraints imposed on linguistic forms—as opposed to those imposed on gesture.

I do not claim to have redefined the characteristics of the Language-Gesture distinction or the Auditory-Visual one in this discussion. Rather, what I hope to have achieved is a rethinking of the problems involved in such definitions. It is important to note potential relationships between primary informational channel and completeness or analyticity—or between iconicity (with its global mappings) and real-space blending—and consider the consequences of such relationships. It is also important to break down traditional binary contrasts such as iconic vs. non-iconic as well as iconic vs. conventional—spectrums need to replace some binary oppositions, and more orthogonality needs to be recognized between classifications such as iconicity and conventionality.

REFERENCES

Bavelas, J., Chovil, N., Lowrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes, 15*, 469–489.

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford, UK: Oxford University Press.

Calbris, G. (1990). *The semiotics of French gesture*. Bloomington: Indiana University Press.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. In J.-P. Koenig (Ed.), *Discourse and cognition* (pp. 189–204). Stanford CA: CSLI Publications.

Dudis, P. G. (2004). Body partitioning and real-space blends. *Cognitive Linguistics, 15*(2), 223–238.

Engle, R. A. (2000). Towards a theory of multimodal communication: Combining speech, gestures, diagrams and demonstrations in instructional explanations. Ph.D. dissertation, School of Education, Stanford University.

Fauconnier, G. (1998). Mental spaces, language modalities, and conceptual integration. In M. Tomasello (Ed.), *The new psychology of language* (pp. 251–279). Mahwah, NJ: Lawrence Erlbaum Associates.

Fauconnier, G., & Turner, M. (1996). Blending as a central process of grammar. In A. Goldberg (Ed.), *Conceptual structure, discourse and language* (pp. 113–130). Stanford: CSLI.

Fauconnier, G., & Turner, M. (2001). *The way we think*. Basic Books.

Fillmore, C. J. (1997 [1971]). *Lectures on deixis*. Stanford CA: CSLI Publications (originally published by Indiana University Linguistics Club).

Furuyama, N. (2000). Gestural interaction between the instructor and the learner in origami instruction. In D. McNeill (Ed.), *Language and gesture* (pp. 99–117). Cambridge, UK: Cambridge University Press.

Goldin-Meadow, S. (1993). When does gesture become language? A study of gesture used as a primary communication system by deaf children of hearing parents. In K. R. Gibson and T. Ingold (Eds.), *Tools, language and cognition in human evolution* (pp. 63–85). New York: Cambridge University Press.

Goldin-Meadow, S. (2003). *Hearing gesture*. Cambridge, MA: Harvard University Press.

Hanks, W. F. (1990). *Referential practice: Language and lived space among the Maya*. Chicago: University of Chicago Press.

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

Haviland, J. B. (2000). Pointing, gesture spaces and mental maps. In D. McNeill (Ed.), *Language and gesture* (pp. 13–46). Cambridge, UK: Cambridge University Press.

Kendon, A. (1990). Gesticulation, quotable gestures and signs. In M. Moerman & M. Nomura (Eds.), *Culture embodied* (pp. 53–77). Osaka: National Museum of Ethnology.

Kendon, A. (1995). Gestures as illocutionary and discourse markers in Southern Italian conversation. *Journal of Pragmatics 23*, 247–279.

Kendon, A. (2000). Language and gesture: Unity or duality? In D. McNeill (Ed.), *Language and gesture* (pp. 47–63). Cambridge, UK: Cambridge University Press.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.

Kita, S. (Ed.). (2003). *Pointing: Where language, culture and cognition meet*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

LeBaron, C., & Streeck, J. (2000). Gestures, knowledge and the world. In D. McNeill (Ed.), *Language and Gesture* (pp. 118–138). Cambridge, UK: Cambridge University Press.

Levinson, S. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge, UK: Cambridge University Press.

Liddell, S. (1995). Real, surrogate and token space: grammatical consequences in ASL. K. Emmorey & J. Reilly (Eds.), *Language, gesture and space* (pp. 19–41). Hillsdale, NJ: Lawrence Erlbaum Associates.

Liddell, S. (1998). Grounded blends, gestures, and conceptual shifts. *Cognitive Linguistics, 9*(3), 283–314.

Liddell, S. (2000). Blended spaces and deixis in sign language discourse. In D. McNeill (Ed.), *Language and gesture* (pp. 331–357). Cambridge, UK: Cambridge University Press.

Liddell, S. (2003). *Grammar, gesture and meaning in American Sign Language.* Cambridge, UK: Cambridge University Press.

McNeill, D. (1992). *Hand and mind.* Chicago: University of Chicago Press.

McNeill, D. (Ed.). (2000). *Language and gesture.* Cambridge, England, UK: Cambridge University Press.

McNeill, D. (2005). *Gesture and thought.* Chicago: University of Chicago Press.

McNeill, D., &. Duncan, S. D. (2000). Growth points in thinking for speaking. In D. McNeill, *Language and gesture* (pp. 141–161). Cambridge, UK: Cambridge University Press.

Nuñez, R., & Sweetser, E. (2006) Aymara, where the future is behind you: convergent evidence from language and gesture in the crosslinguistic comparison of spatial realizations of time. *Cognitive Science, 30*, 410–450.

Parrill, F., & Sweetser, E. (2004). What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture, 4*(2), 197–219.

Sizemore, M., & Sweetser, E. (in press). Personal and interpersonal gesture spaces: Functional contrasts in language and gesture. In A. Tyler, Y. Kim, and M. Takada (Eds.), *Language in the context of use: Cognitive and discourse approaches to language and language learning.* Berlin: Mouton de Gruyter.

Slobin, D. I. (1987). Thinking for speaking. In J. Aske et al. (Eds.), *Proceedings of the 13th annual meeting of the Berkeley Linguistics Society* (pp. 435–445). Berkeley: The Berkeley Linguistics Society.

Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking.' In J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, UK: Cambridge University Press.

Slobin, D. I. (2000). Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 107–138). Amsterdam: John Benjamins.

Slobin, D. (2005). Linguistic representations of motion events: What is signifier and what is signified? In C. Maeder, O. Fischer, & W. Herlofsky (Eds.), *Iconicity inside out: Iconicity in Language and Literature 4.* Amsterdam/Philadelphia: John Benjamins.

Slobin, D. I., & Hoiting, N. (1994). Reference to movement in spoken and signed languages: Typological considerations. *Proceedings of the 20th annual meeting of the Berkeley Linguistics Society* (pp. 487–505). Berkeley: The Berkeley Linguistics Society.

Smith, N. (2003). Gesture and beyond. Honors thesis, Program in Cognitive Science, University of California at Berkeley.

Sweetser, E. (1998). Regular metaphoricity in gesture: boldily-based models of speech interaction. In CD-ROM *Proceedings of the 16th International Congress of Linguists*, Paris, July 1997.

Sweetser, E. (2003). Literal and metaphorical viewpoint in gesture. Paper presented at the 8th International Cognitive Linguistics Conference, University of La Rioja, Spain, July 2003.

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language.* Cambridge, UK: Cambridge University Press.