What we mean by meaning

Conceptual integration in gesture analysis and transcription

Fey Parrill and Eve Sweetser
University of Chicago / University of California, Berkeley

Determining what a gesture "means" is an intuitive inferential process, which can profit from the application of a formalism which guides analysis. This paper uses a detailed exploration of a single example to illustrate the efficacy of the *conceptual integration* framework (Fauconnier & Turner, 2002) in organizing and understanding the process of meaning construction, and particularly in helping the analyst to understand the relationship between iconicity and metaphor.


The electronic edition of this article includes audio-visual data.

Keywords: gesture, blending, conceptual integration, meaning

A colleague recently mentioned to one of us that he discourages his students from studying gesture because it is such a Herculean task; he was quite unconvinced by the claim that gesture is not inherently more difficult to study than, say, attention. While we suspect that his students will prevail — researchers from many disciplines have begun to feel they cannot do without the study of gesture — we also sympathize with his perspective. Gestures contain enormous amounts of information; this can be rather intimidating. And the sorts of gestures which we (the authors) study occur with speech, speech to which they are intimately linked. This fact requires investigating the relationship between how the hands move in producing a gesture, and whatever mental representation underlies it, as inferred both from the gesture and the accompanying speech. We refer to this relationship as the *meaning* of a gesture, and it is this relationship that we are concerned with here.

Gesture 4:2 (2004), **197–219**.
ISSN 1568–1475 / E-ISSN 1569–9773

We sometimes encounter the claim that coming to understand gestural meaning is an essentially subjective process, particularly in the case of abstract gestures. We take the (uncontroversial, we hope) perspective that the analyst's claims about gestural meaning can be explicitly laid out, and furthermore, that this enterprise can profit from the use of a framework from within the cognitive linguistics tradition,[1] the mental spaces framework (Fauconnier, 1994 [1985], 1997; Fauconnier & Sweetser, 1996) and its later extension as *conceptual integration* or *blending* (Fauconnier & Turner, 2002). Blending theory provides extremely general mechanisms for stating and analyzing mappings, or correspondences, between domains; it can therefore be used to express both correspondences between physical forms in space (hands, e.g.) and meanings (ideas, e.g.) — as well as the ways in which speech-expressed meaning content combines with gesturally expressed content to create a dynamically developing whole. Liddell (1998, 2000, 2003), for example, has made use of the blending model to describe the correspondences between a signer's construal of the physical space surrounding him or her, and meaning in American Sign Language. This paper will attempt to show, by means of a relatively detailed explication of a single example, that employing the blending formalism enables the analyst to be explicit about claims regarding how a gesture relates to the discourse in which it is embedded. We also claim that greater explicitness leads to a more detailed understanding of the data.

Meaning and gesture

Meaning is here assumed to be a cognitive construct built up during production or interpretation of a discourse (Lakoff & Johnson, 1980, 1999) — that is, we are not talking about objective meaning that exists independent of how a given person understands it, and needs only to be encoded (by a speaker) or decoded (by a hearer), but of something more like a mental model. This model changes dynamically as the discourse proceeds — as entities are introduced into the discourse, for example, they are also added to the model. Gestures are as much a part of this meaning as speech — they are part of a speaker's communicative goals (whether or not they are communicative themselves) because they are part of language production. The kinds of meaning present in gesture and speech are, of course, quite different: gesture is more global and image-based, while language is more analytic (McNeill & Duncan, 2000). Furthermore, gesture (unlike speech, but like signed languages) involves iconic

mapping as a central mechanism of meaning construction. *Mapping* here refers to a correspondence of some kind between one thing and another. The critical point about this notion of mapping is that it is a cognitive process of perceiving similarity. When we see a hand in a certain configuration and conclude that the hand represents a physical entity being talked about in the accompanying discourse, we are engaging in this process.

To take a concrete example, in the following stretch of speech, the speaker is referring to a ball which has been tossed into the air, and produces a gesture in which her hand is in the shape of a fist, as shown, with the gesture occurring during the bracketed speech.

Figure 1. when [it stops]

Movie 1.

We know *it* refers to a ball, and we know some things about features of a ball — we can form a mental representation of a one, thus adding it to our mental model. Balls, we know, tend to have an outer curve and inner solidity. These features are preserved in the hand shape of the speaker's gesture. Our knowledge of the shape of a (prototypical) ball enables us to extract these features in order to fill out a mental model whereby the hand can represent the ball, the motion of the hand can be construed as the motion of the ball, and so on. Thus we know the gesture represents the ball, and our knowledge comes from the mappings we have constructed between features of the hand and of the ball.[2]

There are obviously two distinct issues in play here: why does the speaker's hand take on a certain shape when she is speaking of a ball, and how do we, as analysts, recognize the shape of her hand as having something to do with a referent (*it*) in her speech? Our discussion focuses on the second issue, but we hope it will become clear that we assume some correspondence between the

two — that in analyzing the gestures we see we are attempting to gain insight into why and how they are produced.

Sarah Taub (2001) gives the topic of mappings across the domains of physical space and linguistic meaning a detailed treatment in her discussion of conceptual metaphor and iconicity in American Sign Language. Conceptual metaphor (Lakoff & Johnson, 1980), which we will have occasion to discuss later, is an approach to the study of metaphor in which metaphorical utterances are explained in terms of mappings from a concrete domain of experience to an abstract one. For example, an utterance like *I'm getting to my next point* maps the concrete domain of goal-directed motion in space onto the abstract domain of discourse production or discussion. Nothing is actually being *gotten to* — the "point" is not spatial. Such metaphors are conventionally represented with capital letters: in this case, DISCOURSE IS MOTION (Sweetser 1992; cf. Lakoff & Johnson, 1980 — AN ARGUMENT IS A JOURNEY).

Taub begins by expanding Mandel's (1977) catalogue of sign-language iconicity in order to describe the ways in which the correspondences we perceive between the sign and its referent enable us to fill out a schematic mental model of the scene. Some of these constraints on iconicity are as follows. In cases where the hands depict a physical entity, the shape of the hands may map onto the shape of a physical referent (*shape-for-shape* mapping). Similarly, the motion of the hands may map onto the motion of the referent (*motion-for-motion* mapping). Motion can also be used to represent shape, as when the path of the hand traces the shape of a referent (*path-for-shape*). A location in physical space can represent a location in some abstract or metaphorical space, and so on. We are concerned with many of the same kinds of mapping when we study gesture, but in speech-accompanying gesture, unlike in sign, we have an additional information channel which informs our construction of iconic mappings. That is, gesture is always contextually embedded, the context deriving from speech.

The blending model

In describing the blending model, let us return to the case in which the speaker is using a hand in the shape of a fist to represent a ball. According to the model, the interpreter of this discourse fragment will have a partially structured mental model which contains the ball, since the speaker has been talking about a ball. Such partially structured mental models are called *mental spaces* (Fauconnier,

1994 [1985], 1997). The interpreter will also have a mental space in which the speaker's physical surroundings are represented, including the gesture she's just produced. These partially structured mental models then act as inputs to a *blended space*, or a space which contains emergent (or new) structure based on the structure from its inputs. Simply put, one input to the blend is the fist, the other is the imagined ball, and the blended space is the conceptualization of the fist as a ball — the understanding wherein, without losing the ability to see a human hand, the interpreter simultaneously construes it as the ball which it represents. This cognitive process is part of our interpretation as analysts, and is a matter of creating correspondences between domains.

Before describing this example in more technical detail, we will provide a little background on the theory. The construct of the mental space was originally developed to account for lexical ambiguities in reference. Fauconnier (1994 [1985], 1997) began constructing his theory of mental spaces because of the observation that linguistic reference is pervasively achieved by "indirect" means — speakers use a description or label appropriate to one referent in order to refer to some counterpart of that referent in another mental space. For example, movie reviewers regularly say things like, *At the end of Gone with the Wind, Clark Gable leaves Vivian Leigh*, meaning simply that the character Rhett Butler (played by Gable) leaves his wife Scarlett (played by Leigh). That is, there is a correspondence between entities in a "movie" mental space (Rhett and Scarlett) and entities in our model of the real world (Gable and Leigh). In this example there is no blending — there is no emergent structure, just a counterpart mapping which lets us identify a movie character by reference to the relevant actor counterpart. Blending theory was later developed to account for cases where something new emerges as a result of mappings across domains.

The application of blending to gesture and signed languages was first noted by Scott Liddell, who introduced the notion of *Real Space*, or "the mental representation of the physical elements of one's immediate physical environment" (2000, p. 342). The importance of Real Space where gesture is concerned is that physical space — both the body and its surroundings — becomes a resource for meaning construction. Liddell points out that a speaker's use of physical space or surroundings to represent some entity in the discourse creates a *blend* (Liddell calls this a *grounded blend*, because it involves Real Space).

This construct allows us to present the hand-as-ball gesture described above with greater technical detail. Real Space acts as one input: a second input is the mental space in which the ball is evoked (with its corresponding properties of roundness, internal solidity, etc.). In addition to the inputs, another

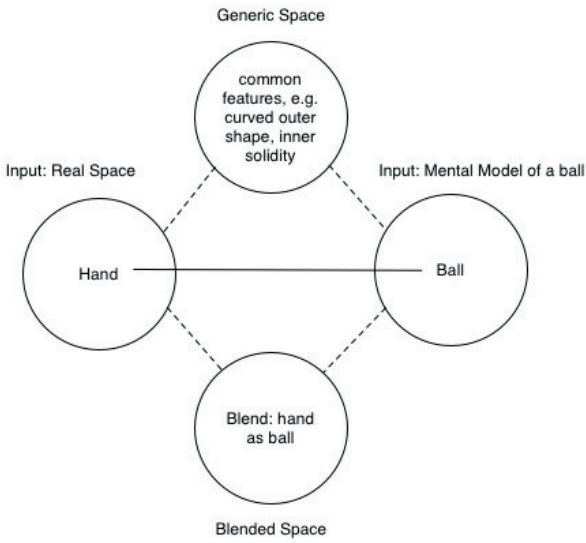

Figure 2.

important element of the blend is the *generic space*, which contains shared features (perceived similarities) that themselves permit cross-space mappings. That is, whatever structure is shared between the two inputs is mapped into the generic space. This example can be depicted as in Figure 2, using the conventions of the blending model, where circles show mental spaces, and lines show correspondences, or mappings across domains.

One of the benefits of this model is that it provides an easily formalized explanation of which things become part of the mapping, that is, of which elements are *projected* into the blend. In Real Space there will be an arm extending out of the hand, but this is not mapped into the blend. The arm has no analogue in this particular mental model of the ball, and therefore is not extracted as a shared feature and added to the generic space (i.e. projection into the blend is *selective* — Fauconnier & Turner, 2002). This model may seem an unwieldy system for the expression of such simple relationships. In reality, however, few gestures have such simple correspondences, and it is in describing the mappings between elements in complex discourse that this framework can be seen to be invaluable, as it allows the analyst to build a coherent representation of the unfolding discourse structure.

For instance, rather than representing a moving entity such as a ball, a hand may instead represent something in the ongoing speech interaction at a meta-narrative level. Let us examine another example. During the following

Figure 3.

speech fragment, the speaker produces a series of 15 small circles by rotating his wrist away from his body and in again, at chest level, within the same area repeatedly. (Each circle is enclosed in a pair of brackets.)

> [and / [um /] [the reason I] [want to] [do that] [is I want] [to talk]
> [about] [a non] [-na][tural] [behavior] [that evolves] [o][ver] [time]]

This circular motion in one location does not correlate with speech about a rolling object, but marks the speech segment as merely a backgrounded part of a longer, ongoing discourse. This, then, is a metaphoric gesture.[3] In our earlier remarks about conceptual metaphor, we provided an example of mapping the concrete domain of traveling onto the abstract domain of discourse production or discussion, and this gesture is also an example of the same conceptual metaphor. The difference here is that the metaphor is not simultaneously manifested in the accompanying speech, which is not unusual (Cienki, 1998). In the case of metaphoric gestures, iconic mappings alone are not sufficient to set up the correspondence between the speech and the accompanying gesture. Speech interaction does not literally engage in spatial motion or "occupy a space" in a larger physical progression, thus it cannot be the case that the rolling motion of this gesture iconically represents the speech.

Taub (2001) lays out for ASL signs the way in which a second layer of mappings, metaphoric mappings, can be layered onto iconic mappings, and this approach works nicely for gesture as well. Taub's analysis involves describing the physical, or Real Space features, then the mappings between those features and analogues in the source domain of the metaphor, which is always concrete (motion, in this case). These mappings are iconic. The next step is to relate the source domain features to the target domain of the metaphor, which is always abstract (discourse, in this case). These mappings are metaphoric. In this example, we might say that the rolling hand iconically represents a "rolling in place" motion. Metaphorically, discourse structure is mapped onto traversal

Table 1.

Iconic Mappings		Metaphoric Mappings
Real Space	**Source (motion)**	**Target (Discourse)**
Location in space	Location on path	Discourse topic
Motion in space	Motion in space	Presentation of argument
Origin of motion	Source of path	Beginning of argument
Endpoint of motion	End of path — Goal	End of argument
Circular motion over same area	Re-traversing same area on path	Presenting background information
Motion forward	Moving forward on path	Adding information to argument

of a path from source to goal, and background material which does not in itself advance the argument structure is mapped onto delays along that path. Again taking DISCOURSE IS MOTION as the basic metaphor behind this example, Table 1 (following the convention of Taub, 2001) elaborates on some of the key double mappings.

These same mappings can also be represented through the formalism of the blending model, as below in Figure 4. In this case, however, there are multiple blends to be considered. The first involves the iconic mappings which permit us to perceive the hand as representing some *other* moving entity, that is, the mappings from Real Space, in which the hand performs some action, to the mental model of a moving object, which here is the source domain of the metaphor. As with the example in which the hand is representing a ball, the perceived similarity between the two inputs — such as the fact that there exists

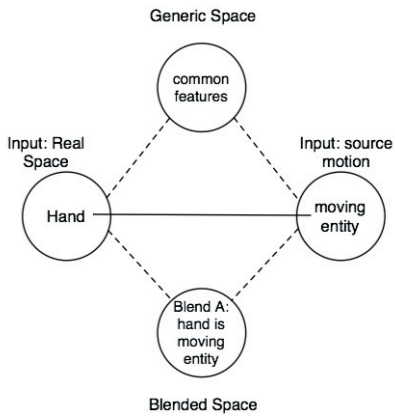

Figure 4.

some entity, that it is in motion, that the motion is of a particular sort, and so on — is captured by the generic space. What results is a blend, represented here in Figure 4 as blend A for convenience,[4] in which the hand is perceived as a generic moving entity. This blend looks very much like our hand-as-ball example, but thus far we have only mapped a hand onto a object rolling in place.

The second blend (which we will refer to as blend B) makes explicit the mappings between the source and target domains of the metaphor. These domains act as inputs to a second blended mental space, in which the movement of an object is construed as a discourse process. Thus, the moving object maps onto what we have called a *discourse entity* (or *discourse object* — McNeill, 1992; Kendon, 1995; Müller, 2004), that is, an idea or a proposition metaphorically conceived of as an object, as shown in Figure 5. Both blend A and blend B share the source domain (motion) of the metaphor as an input.

These blends can then be combined (that is, the blended spaces can serve as inputs to a new blend) in order to represent the construal of the moving hand as a discourse entity, as represented in Figure 6. In this way, real entities (the speaker-gesturer's hands) can give us information about something as abstract as discourse.

Blending is proposed to be a cognitive operation by which partial structure from input spaces is combined to yield emergent structure. However, the blended-spaces formalism is clearly not literally representing neural structures engaged in a dynamic cognitive processes. Mental "spaces" are not literal spaces — or literal direct representations of cognition. One way of construing them is simply as effective representations of some real generalizations about language and other human behavior; as ways of talking about the complex

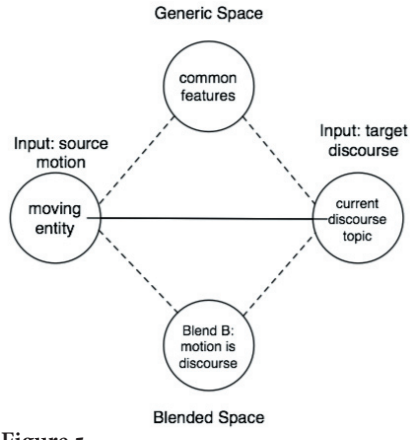

Figure 5.

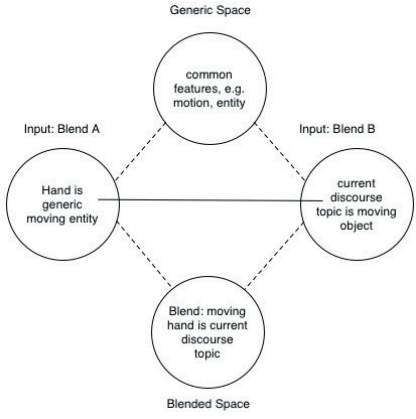

Figure 6.

correspondences between (for example) Real Space and conceptual structure which humans are uniquely qualified to build up. (So far as we know, even other primates don't engage in such complex construals.) In more recent work, Fauconnier and Turner (2002) state that the higher-order phenomena of projection and mapping can be thought of as physically instantiated in neural coactivation or binding. The spaces themselves are representations in short-term memory, while the completed blend is an active neural network (a set of neurons which are firing at the same time). Relating these formal devices to neural structures about which relatively little is definitively known (e.g. the exact behavior of neural networks or the neural representation of short-term memory) may appear to have little theoretical significance. It is important, however, to take a clear stance on this point; the sorts of predictions made — for example, that there may be ordering effects in the construction of complex cross-modal blends — will ultimately serve to improve the conceptual integration model. And since this is ultimately a model of the cognitive processes underlying both production and perception/ interpretation of speech and gesture, improvement to the model can benefit the student of gesture.

Finally, mental spaces and blends are dynamic structures, although a given diagram is necessarily static. Fauconnier and Turner explicitly claim that the constructs represented are fluidly evolving cognitive structures. Gestural representation is no different in this respect from the linguistic representation of rapidly changing content structure.

Discourse cohesion and mental space mappings

We now turn to the discussion of an example which spans multiple utterances, in order to illustrate more fully the complexity of a speaker-gesturer's typical use of physical space. One of the realms in which the blending framework is most clearly useful is in describing how speakers maintain cohesion during a complex discourse. With respect to gesture, discourse cohesion is often apparent in the form of repetition in the features of gestures produced, a phenomenon which McNeill refers to as the *catchment*, or "a kind of thread of consistent dynamic visuo-spatial imagery running through a discourse segment…" (McNeill, 2000). The following example comes from an academic lecture. The speaker is talking about dynamic programming and using it as a metaphor for reinforcement learning. That is, a program which solves a problem via recursion is being compared to an animal learning to find a reward in a maze. We

will first present the speech as a complete chunk, then a transcription of the speech and gesture. The gestures on which we will be focusing are numbered 1–5 below, though there are some intervening gestures which we will not discuss. Gestures are enclosed in brackets, gesture stroke is in bold, / is a speech pause, * is a self-interruption, and underlining indicates a gesture hold. The prosodic peak is marked with larger font.

> the way reinforcement working it does* [you just wander around idly / okay /] **1** [and then when you bump into the reward]**2** [you know what to do]**3** from the [penultimate state]**4**. You're not going to have a policy, you're clueless, but the penultimate state, you know what to do / and you know how to discount it right? [Then you're wandering around idly a* again] **5** [you might bump into that penultimate state] **6**

Movie 2.

Gesture and speech transcript:

1 the way reinforcement working [[it does* **you just**] [**wander**] [**around**] [**id**][**ly**] [/] [/] [/] [/] <u>okay / </u>]
Both hands (BH) in spread B-hands with palms facing chest start in the center of chest with finger tips touching, R a little behind L. While L holds at chest, R moves over L, makes a series of 8 hopping arcs, moving forward until arm is fully extended. R holds, returns to position of L.

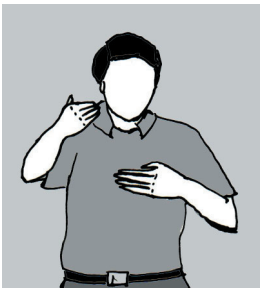

Figure 7.

2 [and then when <u>you bu</u>[**mp into the re**]<u>ward</u> you know what to do]
BH in spread B-hands with palms facing chest start in the center of chest with finger tips touching, R in front of L. While L holds at chest, R moves out from chest to fully extend again, makes beat (Repeats gesture 1 without hops).

Movie 3.

Figure 8.

3 [o from the [**pen**][**ul**][**ti**]mate state]
BH come apart in relaxed C-hands with palms facing center and fingers away from body, to center of chest. Make three up-down beats.

Figure 9.

The following series of gestures will not figure in our analysis, but are included for completeness.

[you're **not** going to have a policy]
BH come out in palm-up open handed presenting gesture.

[you're **clu**]
BH come out in palm-up open handed presenting gesture.

[eless but the **pen**] [**ultimate state** you know what to do]
BH come towards center in relaxed C-hands, with palms facing center and fingers away from body.

Hands return to end position of gesture 1 — L in spread B-hand with palm facing chest and R fully extended away from body, palm facing chest.

[**and you know**] [how to [**dis**][**count**] [**it**] / right /]
R returns to chest in 5 hand.
BH come towards center in relaxed C-hands, with palms facing center and fingers away from body, make three up-down beats.

These final two gestures will be analyzed below, in addition to those numbered 1–3, above.

4 [**then you're wan**dering **around** idly a*] again
Full body Character Viewpoint gesture — the speaker's arms move up and down in a loose motion, with hands in loose 5 hands. Hands first come up with palms down, then palms turn up and arms move down, then palms turn down. The speaker simultaneously takes two backward steps while head waggles back and forth.

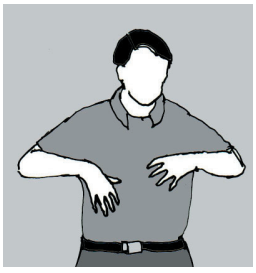

Figure 10.

5 [**you might bump** [into that] [pen**ul**]timate state]
BH come out in relaxed C-hands, arms full extended, fingers towards center, hands relatively close to one another in space in front of chest. Makes two up-down beats.

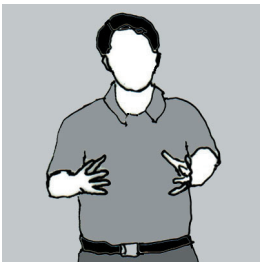

Figure 11.

This discourse segment mixes language from the abstract programming domain (dynamic programming) and language from the concrete learning domain (really, the rat-learning-a-maze domain), making it a complex task to sort out the relationships between speech and gesture. There is no non-metaphorical "penultimate state" in a maze, although there is a location in space before the reward's location. Nor is there any non-metaphorical "wanderer" nor notion of "idleness" (which presupposes volition), nor "reward" in the programming domain, instead there is something more like a circuit's particular electrical state at a given moment,[5] which can be construed as an entity's location. The conflation of these two domains *can* be unpacked, and we will now attempt to do so in our analysis of these gestures.

During 1 (*you just wander around idly okay*), the speaker sets up an axis moving outward from his body with a series of hopping gestures. We claim that this is the axis along which the process of "reinforcement learning" proceeds. Reinforcement learning here involves an animal learning to navigate a maze, and the speaker describes this process as *wandering around idly*. We claim the right hand's motion outward is a path-for-path mapping: the path of the hand is mapping onto the path of a moving entity. But the entity wandering around idly is not moving in a series of hopping motions, so why does this gesture have the form it does? This is a metaphoric gesture in which a complex process is represented as iterative motion through space. This gesture is not really about the maze, but rather is about the programming domain. By picking out the relevant shared structure between the two domains, namely, that both involve a process, the speaker can built an analogy between the two. The left hand holds a few inches from the speaker's chest, where it serves to represent some other entity. Based on the rest of the discourse, we identify this entity as the location before the reward in the maze domain or the penultimate state in the programming domain. Some of the mappings we postulate are shown in Table 2

Table 2.

Iconic Mappings		Metaphoric Mappings
Real Space	**Source** (maze)	**Target** (program)
Right hand	Location of wanderer in maze	Electrical state of circuit
Motion	Activity of wandering	Program operating
Left hand	Location in maze before reward	Penultimate state
Circularity of motion	Lack of directed motion through maze	Program operating without explicit training

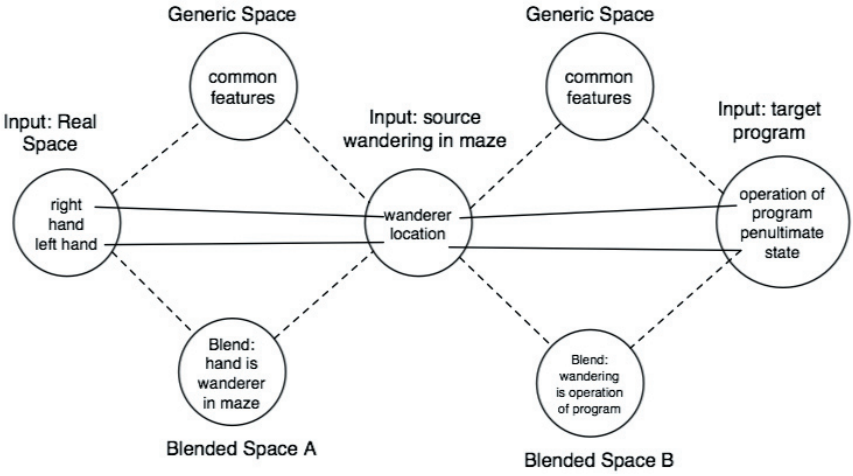

Figure 12.

and Figure 12. For the sake of space we have here represented the two blends as one diagram, but these diagrams should be understood as functioning just like those in Figures 4 and 5 above. That is, one blend represents the mappings from the Real Space articulators (hands, etc.) to the source domain of the conceptual metaphor, while the second represents the mappings from the source domain to the target domain. These blends have simply been collapsed so that the metaphor's source domain is represented only once.[6]

At 2 (*when you bump into the reward)*, the speaker moves his right hand away from his body until the arm is fully extended again, but this time in a single, smooth motion. This motion acts to situate *the reward* at the location defined by the final location of the hand. We base this claim on the relationship between speech and gesture: the indexing of a location in space is synchronized with an expression identifying a referent. The mappings in this case are similar to those described above, but the virtual axis extending outward from the speaker's body (as defined by the space between his right and left hands) now represents not the path of the process of wandering, but the "step" between the penultimate state and the reward. A change in mapping is inferred both from the speaker's overt action of identifying a spatial location with an entity, and from the fact the motion is smooth rather than circular. In this fashion, coherence of the model being constructed is maintained, and we assume that speakers maintain coherence unless we see evidence to the contrary. This kind of dynamic shifting is not in any way unusual, and we claim the blending model is

Table 3.

Iconic		Metaphoric
Real Space	**Source**	**Target**
Right hand	Location of reward	Final state of program
Left hand	Location before reward	Penultimate state

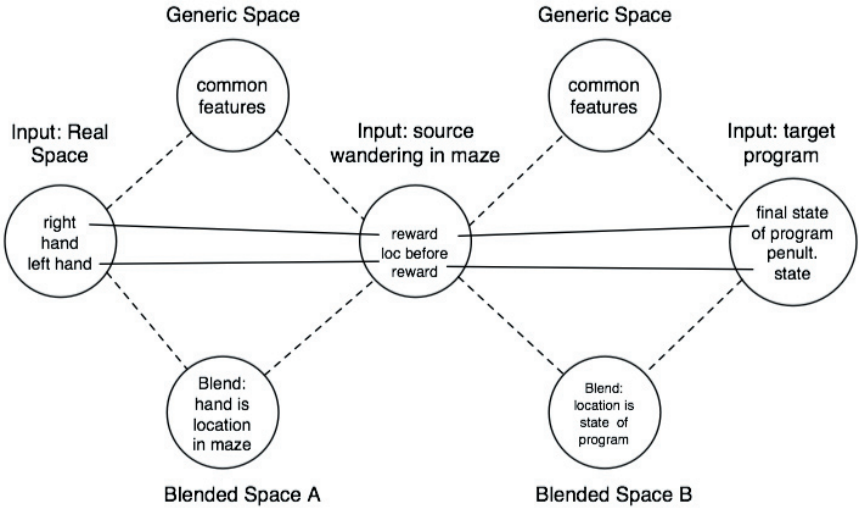

Figure 13.

well-suited to capturing it. The mappings which differ from the segment above, and a diagram of the blend, are shown in Table 3 and Figure 13.

As noted above, one of the interesting features of this discourse segment is the blending of language from the programming domain and the maze domain (there is no penultimate state in a maze and one does not reward a program). With respect to the gesture, however, it is important to keep in mind that there is no way of representing the programming domain itself — that is, gestures about a program can only be metaphorical. It is our ability to construe physical motion as sharing structure with the abstract domain of the operation of a computer program which allows us to create these mappings, and to understand the gestures as having something to do with programming: as previously emphasized, blending is a cognitive operation.

Turning now to the third gesture (*from the penultimate state*), we see that the speaker uses both hands to define a space in front of his body. In this case, it is important to note that not all elements from the iconic domain have mappings in the target domain of the metaphoric blend. As pointed out before, projection is selective. There is no equivalent in the programming domain to the

Table 4.

Iconic	Metaphoric	
Real Space	**Source**	**Target**
Hands	Boundaries of location in maze	
Space between hands	Location in maze	State of program (penultimate state)

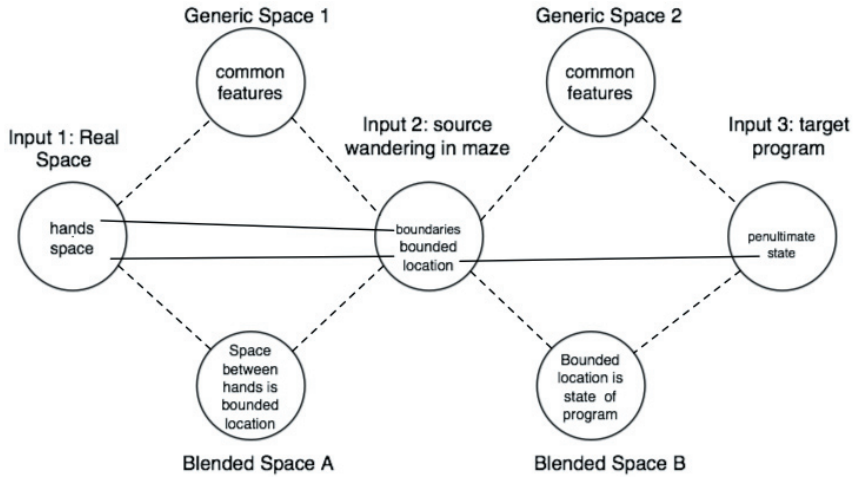

Figure 14.

boundaries of a physical space in the reinforcement learning domain, shown in Table 4 and Figure 14.

In the fourth gesture the speaker says *wandering around idly* for the second time, while waving his arms and assuming a bewildered expression. Here he is assuming *character viewpoint* (McNeill, 1992), or creating a blend in which he "becomes" the wanderer. In this case, his Real Space body maps onto the body of the wanderer, an entity which has been projected into the mental space inputs to this blend, shown in Table 5 and Figure 15. That is, the initial use of the phrase *you're wandering around idly* cues the interpreter to introduce a wanderer into the maze mental space, as described above, and this gesture simply profiles that entity at a different level of granularity. When watching this gesture, we don't wonder why the speaker is suddenly flailing about: there is an already existing notion of an entity wandering, and because there is coherence in this blended model of the discourse, we can interpret the action as referring to that entity in the same way we normally distinguish between observer and character viewpoint gestures.

Table 5.

Iconic	Metaphoric	
Real Space	**Source**	**Target**
Location of Body	Location of wanderer in maze	State of program

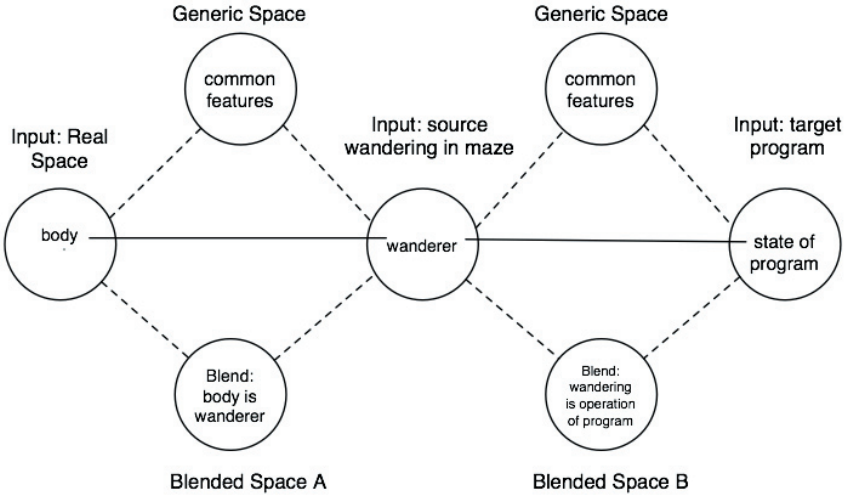

Figure 15.

While this scenario is quite complex already, the speaker complicates it further with the final gesture, occurring with *you might bump into that penultimate state*, in which his hands define a space and his arms are fully extended. This gesture serves to assign a new spatial location to the penultimate state. Rather than representing the "step" between the penultimate state and the reward, the space between the hands and body now represents the "step" between the wanderer's location and the penultimate state.

These five gestures depict elements in the discourse from dynamic perspectives, entailing changes in the mappings which are being profiled, which we will attempt to summarize. In the original gestural representation (gesture one), the penultimate state, although it was not explicitly labeled as such, is located next to the body (the holding left hand of gestures 1 and two). The focus during gestures one and two is the process which leads to the reward, and the reward itself (represented by the right hand). The penultimate state is linguistically introduced into the discourse with gesture three, where it is represented from the perspective of the wanderer who has suddenly stumbled upon it, with the two handed gesture. This representation is at a new level of granularity: it

Table 6.

Iconic	Metaphoric	
Real Space	**Source**	**Target**
Hands	Boundaries of location in maze	
Space between hands	Location in maze	State of program (penultimate state)
Space between hands and body	Locations between current location and loc. before reward.	Actions of program "between" current and penultimate state

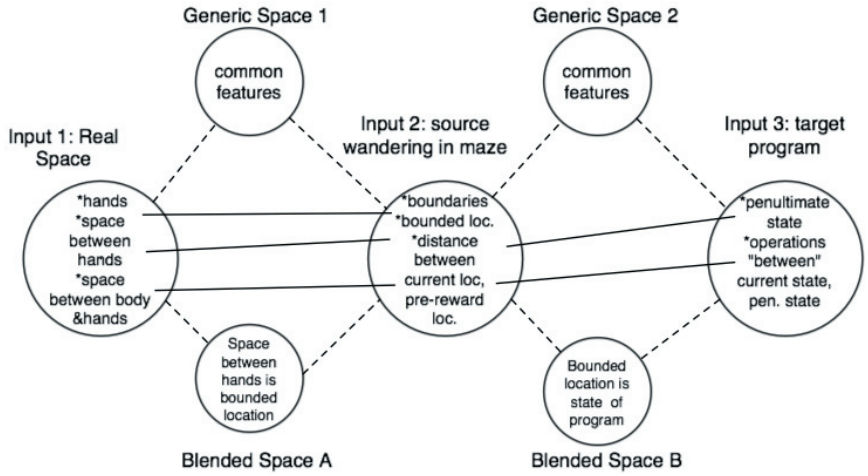

Figure 16.

constitutes "zooming in" on the mental space in which the penultimate state is a physical location in a maze. The speaker has moved from observer viewpoint to character viewpoint, and creates a blend with his body as part of the Real Space input. In the fourth gesture (and also in the gestures which intervene but are not analyzed) the speaker maintains character viewpoint, again with his body standing in for the wanderer, this time representing the wandering action rather than any encountered object. With the fifth gesture, the axis extending from the speaker's body, established with the first gesture, is used again, with the penultimate state now mapping onto the former location of the reward. That is, the speaker is claiming that the program has learned, because the wanderer now can now find its way to the reward from the penultimate state, so the penultimate state takes on the features, and therefore the location, of the reward. In addition, the gesture is a mix of character and observer viewpoint.

The penultimate state is still being represented with the two-handed gesture used in its first linguistic introduction (character viewpoint), but rather than being where it would be if the wanderer were actually bumping into it (as it is in gesture three, and as one might expect from the co-occurring speech — *bump into*), it's extended away from his body, maintaining the axis set up with the initial two observer viewpoint gestures. This is what coindexes the penultimate state with the previous location of the reward.

Because this discourse is so abbreviated, it may not be appropriate to speak of a *catchment* (McNeill, 2000) here, but the speaker is clearly employing multiple different coherent representations in structuring his discourse. Recurring gestural features may represent mappings which are being reintroduced into the discourse, or they may represent different entities entirely. That is, if a speaker uses the hand-as-ball gesture (Figure 1) when talking about a ball, and then produces the same fist shape later, it may be the ball again. But if it occurs with "power to the people!" we will not assume so. The blending model will capture the phenomenon of the catchment in an explicit fashion, whether the discourse involves simple iconic mappings or something like the remarkable — yet quite typical — manipulation of mental space blends we have discussed above.

Conclusions

Whether or not the reader agrees with the details of this analysis, we hope to have shown how the blending formalism can be employed to the benefit of the analyst. One of the advantages of this approach, we believe, is that it offers an elaborated treatment of the role of iconicity in representational gestures, deictics, and beats. While the classic McNeill gesture categories (McNeill, 1992) were never intended to be mutually exclusive, providing mappings forces the transcriber to be explicit about his or her judgments. That is, if one decides that a gesture is deictic, being obliged to transcribe mappings may prevent one from ignoring the fact that the referent is some abstract, metaphoric entity. To put it another way, gestural metaphor can only exist by being layered upon iconicity, deictic referents may be abstract (even metaphoric, again), and the emphasis inherent in a beat is not meaningless: attempting to decide where on a continuum from beat to metaphoric a given gesture lies may be facilitated by considering the conceptual mappings.

Finally, as Liddell's work (1998, 2000, 2003) and our own experience make clear, there must be a means of transcribing discourse uses of physical space. Speakers use space to represent concepts and entities. Differences in physical space are used to express contrast, (metaphorical) temporal ordering, and so on. Given that the blending framework is specifically designed to depict the structure of mental representations, use of space in discourse may be easier to understand with this framework.

The more an analyst examines gesture, the more surprising it becomes that people effortlessly produce and process these complex sets of conceptual mappings between spaces. The same, of course, has been said of language — and special built-in hardware has been postulated to account for such amazing abilities. But most linguists, including those who see language as a modular system, acknowledge that linguistic signs are no more than the tip of the iceberg. They are but prompts to the evocation of a meaning which is very far from being determined by a particular linguistic sequence. This is even more strongly true of gesture. The vast indeterminacy of iconic mappings means that one particular sequence of hand configurations could be iconic for many different possible literal physical referents; the additional possibilities brought up by cross-space mappings ensure that the referential possibilities are multiplied again. In general, access to accompanying linguistic production ensures that the mapping possibilities are appropriately constrained. We have tried to argue here that, when confronting the complexity such as this — complexity which makes some fear the very idea of studying gesture — the blending model can help the analyst organize her thoughts for the task. After all, it is this discovery procedure which is for some the most fascinating part of the enterprise.

Notes

1. Cognitive linguistics is an approach to the study of language which focuses on the relationship between language and the mind, and is sometimes contrasted with "formal" linguistics, which generally abstracts away from non-linguistic information.

2. This example was selected because we know the gesture does not represent a hand holding a ball, because the ball has been tossed into the air.

3. We base the claim that this is not a gesture about *evolving over time* — though this would still be metaphorical — on the fact that the gesture begins and continues over the preceding segment, and also on observations about this speaker's patterns of gesture use (i.e. he uses this meta-narrative gesture frequently).

4. This is not meant to imply that there is any temporal ordering in the construction of these two blends, although this may be a question worth exploring, as will be discussed below.

5. The structure of the blend between these two domains is a fascinating subject, but too complex to be described here except as it pertains to the example being analyzed. That is, rather than beginning by laying out exactly how binary encoding of information can come to be construed as a location in space, we will simply discuss how the speaker makes use of this construal.

6. It is worth noting that the imagery for this gesture is in place before the speaker has determined how to encode it in language. That is, the positions for the reward and the penultimate state are already established before they have been introduced in speech. For gesture to precede speech may well be a feature of this speech genre, which is relatively informal while still being well-rehearsed.

References

Cienki, Alan (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. In Jean-Pierre Koening (Ed.), *Discourse and cognition: Bridging the gap* (pp. 189–204). Stanford: CSLI Publications.

Fauconnier, Gilles (1994 [1985]). *Mental spaces*. Cambridge, MA: Cambridge University Press.

Fauconnier, Gilles (1997). *Mappings in thought and language*. Cambridge, MA: Cambridge University Press.

Fauconnier, Gilles & Eve E. Sweetser (Eds.) (1996). *Spaces, worlds, and grammar*. Chicago: University of Chicago Press.

Fauconnier, Gilles & Mark Turner (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.

Kendon, Adam (1995). Gestures as illocutionary and discourse structure markers in southern Italian conversation. *Journal of Pragmatics 23*: 247–279.

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

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

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

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

Liddell, Scott K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge, MA: Cambridge University Press.

Mandel, Mark (1977). Iconic devices in American Sign Language. In Lynn A. Friedman (Ed.), *On the other hand: New perspectives in American Sign Language* (pp. 57–107). London: Academic Press.

McNeill, David (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
McNeill, David (2000). Catchments and contexts: Non-modular factors in speech and gesture production. In David McNeill (Ed.), *Language and gesture* (pp. 321–328). Cambridge, MA: Cambridge University Press.
McNeill, David & Susan D. Duncan (2000). Growth points in thinking-for-speaking. In David McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge, MA: Cambridge University Press.
Müller, Cornelia (2004). *Metaphors. Dead and alive, sleeping and waking: A cognitive approach to metaphors in language use*. Habilitationsschrift, Free University Berlin.
Sweetser, Eve E. (1992). English metaphors for language: Motivations, conventions, and creativity. *Poetics Today*, 13(4), 705–724.
Taub, Sarah F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge, MA: Cambridge University Press.

Author's address

Fey Parrill
University of Chicago
Department of Psychology
Green Hall, 208
5848 S. University Ave
Chicago, IL 60615
U.S.A.

feyparr@uchicago.edu

Eve Sweetser
University of California, Berkeley
Department of Linguistics
1203 Dwinelle Hall
Berkeley, CA 94720-2650
U.S.A.

sweetser@cogsci.berkeley.edu

About the authors

Fey Parrill is a doctoral student in Linguistics and Psychology at the University of Chicago; her research focuses on pragmatics and gesture, particularly on information packaging in speech and gesture.

Eve Sweetser is Professor in the Department of Linguistics and Director of the Program in Cognitive Science, at the University of California at Berkeley. Her linguistic research includes work on metaphor, mental spaces, and the semantics of grammatical constructions. She is the organizer of the Berkeley Gesture Project, whose work centers on examining the metaphoric and iconic character of speech-accompanying gesture which represents abstract subject-matter, including discourse-structure.