Blending across modalities in mathematical discourse

Nathaniel J. Smith

To appear in: L. D. Edwards, F. Ferrara, & D. Moore-Russo, *Emerging perspectives in gesture and embodiment in mathematics*. Charlotte, NC: Information Age Publishing.

When people engage in face-to-face discourse about mathematical topics, they deploy a variety of communicative modalities, including at least speech, gesture, and various forms of symbolic and diagrammatic marking on paper or blackboards. These modalities are not parallel streams: to be parallel means to never meet, but speakers mix these streams constantly, coordinating distinct representational resources in a fine-grained way within a single, shared, and meaning-laden space, and it's in this space that communication and thinking are performed. There are even cases where the apparent boundaries between modalities break down entirely. To understand mathematical discourse, then, it's not enough to examine individual meaningful actions – single utterances, gesture strokes, and so forth. We also need to understand the structure of the space in which they occur, and how this structure is built, maintained, and manipulated through communicative acts. In this chapter, I'll examine this question through a series of examples drawn from a corpus of naturalistic video recordings of upper-division and graduate mathematics students working together in small groups to complete their assigned homework in topology and abstract algebra (Smith, 2003).

1 Real-space blends

The theory of *real-space blending* was originally developed for the analysis of sign language (Liddell, 2003), but has proven useful in the study of both gesture (Parrill & Sweetser, 2004) and complex conceptual artifacts like maps and compasses (Hutchins, 2005). This theory builds in turn on the general theoretical framework of *conceptual integration theory* (Fauconnier, 1997; Fauconnier & Turner, 2002). The core idea of conceptual integration theory is that humans constantly engage in the creation of new conceptual structures, and that they very often do this by following a particular recipe: they take two or more existing concepts, identify mappings between their parts, and then *blend* them together to create a new, indivisible concept whose structure is drawn from the input concepts, adjusted as necessary to produce a coherent whole. Such a process seems to be needed to explain phenomena as diverse as adjective/noun combinations (Sweetser, 1999), *if* statements (Dancygier & Sweetser, 2009), moral judgements (Coulson, 2001), and the development of abstract mathematical concepts (Lakoff & Núñez, 2000; Edwards, 2009). Here, though, we will focus on real-space blends, i.e., those which instantiate a mapping between abstract conceptual

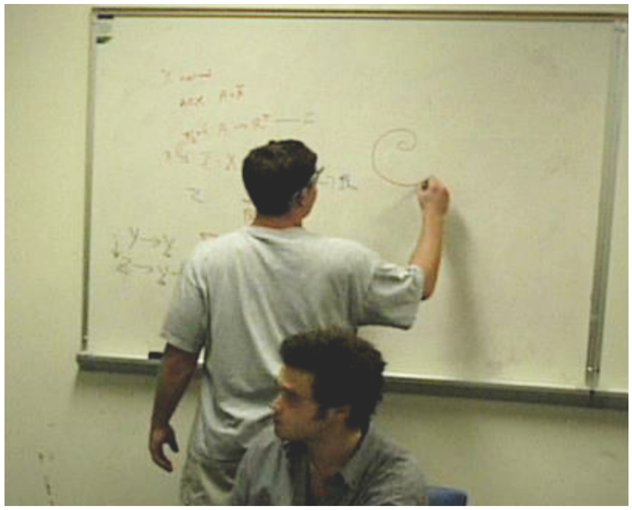


Figure 1: *The logarithmic spiral...*

structure, and the physical structure of the real environment.[1] For example, consider the diagram drawn in (1):[2]

(1) the logarithmic spiral...
draws a spiral on the board, Fig. 1

In physical terms, all the speaker is doing is placing some bits of colored dust on the surface of a whiteboard. But of course the reason he does this is because he is setting up a coherent mapping between this dust's particular spatial arrangement, and a certain mathematical object. This object – the *logarithmic spiral* – consists of those complex numbers which can be written as $e^{\theta}e^{i\theta}$ for some real number θ, and if we conceptualize the set of all complex numbers as forming a two-dimensional plane (itself a complex many-input blend), then the points which satisfy this formula will form the shape of a spiral around the origin. It is this conceptual spiral which is then mapped onto the physical markings on the whiteboard – a second blend, and the one that is the focus of our attention here. Of course a number of accomodations must be made: the whiteboard is not infinite in extent; his mark does not extend infinitely in either direction; his drawing is not accurate in its proportions. Nonetheless, a great deal of internal structure does map between these domains: the various locations on the drawn spiral correspond to different points on the abstract spiral, the blank space around the drawn line maps to other complex numbers, and the center of the spiral maps to the plane's origin. The speech provides a critical cue for what the marking is intended to mean, and together with the shape of the drawn spiral, allows us to construct the appropriate mapping.

[1]Since co-present interlocutors have perceptual access to their shared physical environment, we can treat their cognitive representation of this environment as an especially detailed and concrete kind of concept, and use it as an input to the blending process; see Liddell (2003) and Hutchins (2005) for further discussion of this point.

[2]For transcripts, I use ellipses to represent silent pauses, and "<XXX>" to denote inaudible speech. When two speakers interact, they are labeled 'L' and 'R', corresponding to their position on either the left or right of the video frame, according to the reader's viewpoint. Gestures are represented by [brackets] to indicate the timing of pre- and post-stroke holds, **bolding** to mark the gestural stroke itself, and italicized text describing the motions which is horizontally aligned to the start of the stroke. Underlining marks periods in which the speaker is writing or drawing simultaneously with their speech.

But once the blend is established, it remains indefinitely, lingering long after the speech is gone; in fact immediately after this spiral is drawn, the group turned to discuss a tangential topic for several minutes before mentioning or interacting with this drawing again.

We'll follow their example, and return to the spiral below, after examining a gesture example.

2 The life-cycle of real-space blends

In (2), the speaker on the right describes how she constructed a new set (let's call it S_2) by performing a series of operations on some original set (S_1, say), with reference to a third set she dubs U. Overall, she has two points she wants to make: the first is that S_2 has a simple relationship to S_1; it is obtained by taking all elements of S_1 that are not in U, or in set arithmetic terms, $S_2 = S_1 - U$. The second is that the new set S_2 has a certain topological property. In topology, sets can be *open*, *closed*, both, or neither; and these properties are defined in part by how they propagate through the two operations of set complement and set union. Here, she demonstrates that if S_1 is closed and U is open, then S_2 is also closed. To accomplish this, she describes a complex series of complement and union operations, which together show that S_2 is closed; simultaneously, she uses gesture to build a complex real-space blend to track these operations, and show that this sequence of operations is equivalent to a simple set difference operation.

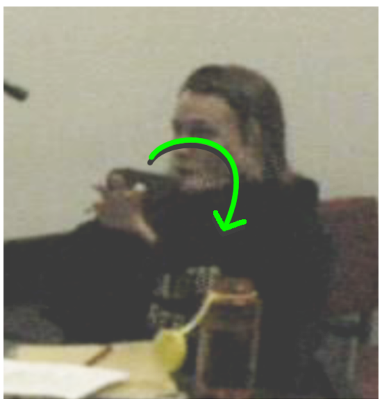
Because they are doing topology, the sets are conventionally conceptualized as sub-regions of some larger region of space. She begins with the set S_1, blended onto her right hand (Fig. 2a). Notice that this is essentially similar to the process described above, with the physical structure of her hand replacing the whiteboard marker pigment[3]:

(2) so [$_1$**my** original closed set,
    puts right hand up in air, holds, Fig. 2a
  right, was [the comp**lement of some** open set.]
      scoops left hand around to indicate open set, Fig. 2b
  now I'm [taking **that** open set,]
      repeats left hand scoop, quicker this time, Fig. 2c
  union [another open set... right?]
    left hand hold, Fig. 2d
  [**that open set**]
  left hand scoop again, Fig. 2e
  union [$_2$**U**,
    left hand hold again, Fig. 2f
  and taking the]$_1$[**complement of that**
      right hand moves away from left, Fig. 2g
  ]]$_2$...
  both hands drop

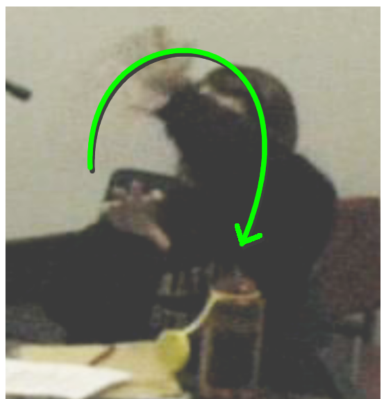
First we see that by initially placing the right hand, she has created a blend that extends beyond it – she is also structuring the space around her hand, so that left hand gestures she performs in the

[3]Subscripts in brackets indicate the beginning and ending of two overlapping gestures.

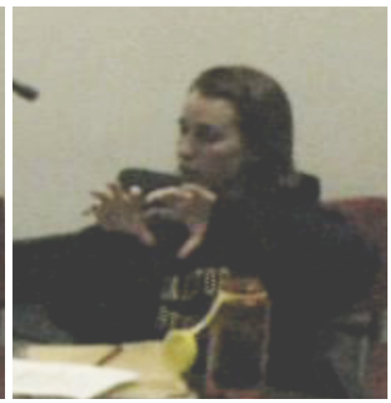
(a) *so [**my** original closed set*



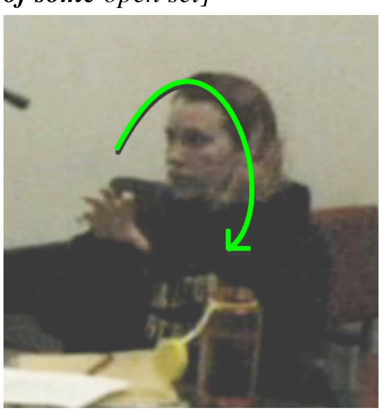
(b) *right, was [the comp**lement of some** open set]*



(c) *now I'm [taking **that** open set,]*



(d) *union [another open set... right?]*



(e) *[**that open set**]*

(f) *union **U***

(g) *and taking the] [**complement of that**]*

Figure 2

physical region where her right hand is *not* located refer conceptually to the points which are not in the set S_1 – i.e., the set complement of S_1 (Fig. 2b, c).

She then introduces the set U, which she conceptualizes as partially overlapping the original set S_1.[4] This set is also blended onto the left hand (Fig. 2d), and the set overlap is represented by placing the corresponding gestural hold physically into the region between her right hand index finger and thumb, so that the convex hulls of the two hands representing S_1 and U physically overlap.

In Fig. 2e and f, she alternates between invoking the complement set and the new set U; since she has only two hands, she cannot perform them simultaneously, while also maintaining the right-hand hold as a reference point anchoring the blend. The speech co-timed with both of these gestural pairs c-d and e-f tells us what the relation between these sets is: she is taking the union between them (i.e., the set which contains all points that are in either or both of the original sets). Interestingly, her syntax here does not follow the conventional rules of English grammar; while it is embedded in a standard English sentence, the phrasing *X union Y* instead follows the conventions for reciting mathematical formulas like $X \cup Y$. According to these conventions, we read symbols from left to right, and this rule is what ultimately provides the temporal ordering on her gestures – the order of her gestures, first the complement, second the new set U, make clear that the two gestures correspond to the two parts of such a formula.

Once this is established, yet another new set enters the discourse: the set which is the union of the complement and of U, and which we know by the definition of a union should, in the blended space, consist of the region which overlaps precisely with the sets which were used to construct it. Starting in Fig. 2g, then, the left hand hold becomes ambiguous: it's been established that it occupies the same space as U, and U occupies the same space as the union set; therefore, the left hand hold also gives us the blended position of (part of) the union set.

Finally, she takes the complement of this union set. For depicting this final set she uses the right hand, finally releasing it from the position where it has been held since the beginning of this description – it moves slightly from the S_1 position, to a new position which is almost identical, except that it no longer overlaps with the left hand representing U. It's this motion that encodes what the speech does not: that this set is exactly the set we originally described, $S_2 = S_1 - U$, and that this sequence of operations together has performed a set difference.

This motion would not be meaningful without real-space blend that it occurs within, and this blend required the whole sequence of preceding speech and gesture to construct and elaborate. But once the rich blended space has been established, the simple motion in Fig. 2g is enough to convey a proof of an interesting theorem: that a closed set minus an open set is a closed set.[5] The purpose of the preceding speech and gesture is, in some sense, to put this blend in place, and allow this final gesture.

If we look at this sequence of speech and gesture, we see that it constantly interacts with the

[4] Interestingly, the overlap occurs only in the gesture, not in the speech. This corresponds to the fact that formally, the degree of overlap between S_1 and U is irrelevant: the proof goes through just as well even if S_1 is a subset of U, or if they don't intersect at all. But in either of these cases, the result is trivial (because either S_2 is the empty set, or else $S_2 = S_1$). Her speech describes the general case, just as a formal written presentation would; her gesture not only helps by tracking the entities involved, but also by depicting the particular case where the proof is most interesting and requires real thought.

[5] In fact, immediately before the portion of the video analyzed here, she uses exactly this theorem to justify an argument she is making; this sequence follows when she realizes that she must first justify the theorem.

blended space in complex ways. First, the blended space is first established in Fig. 2a. Then, the left-hand gestures and the final right hand gesture all depend on the existing blend to depict new content (e.g., the complement set gestures, which are interpretable only by comparing their spatial position to the spatial position of S_1), while simultaneously elaborating it further (e.g., by placing U so that it partially overlaps S_1). Simultaneously while this is occuring, she devotes some energy to maintaining the blend. This is accomplished through an elegant dance of hold gestures: whichever hand is not currently elaborating the blend is held steady, anchoring whichever portion of the blend it is currently assigned – first she uses the right hand, then both (briefly) in Fig. 2f, and finally the left hand takes over for Fig. 2g (Enfield, 2003, 2009). These holds both signal the continued relevance of the blend, and provide a visible spatial reference point for it. And, finally, once the blend has served its purpose, she drops both hands, signaling the end of this blend's utility, and freeing up the physical space to whatever content is needed next.

These operations – construction, maintenance, exploitation/elaboration, and destruction – form the fundamental life-cycle of real-space blends in discourse, which interlocutors must systematically manage in order to accomplish their goals. Different strategies may be available and used depending on the specific details of how each blend is instantiated, the current goals, and trade-offs between conflicting ones. For example, a whiteboard drawing remains until erased, but gestures are more transient, and thus ambiguity is more likely to arise about whether an earlier blend is still relevant; the speaker in (2) finds one solution to this problem using hold gestures, but this solution creates a new problem of how to simultaneously maintain and elaborate the blend given her limited number of hands, which she in turn solves with the switch in Fig. 2g. The point of this schema is not to allow us to predict exactly what interlocutors will do in any particular situation, but to provide a conceptual framework to understand these competing constraints on the structure of the persistent, meaningful spaces that discourse participants navigate to succeed – and to encourage us to consider all the work going on outside of the individual gestural strokes.

3 Real-space blends as shared spaces

Real-space blends, by definition, map physical structure and relationships onto conceptual structure and relationships, and vice-versa. The previous examples demonstrated real-space blends in two different modalities separately – drawing and gesture – together with a third modality, speech. But, crucially, real-space blends are not modality specific. While these modalities each have their own unique conventions and affordances (Edwards & Robutti, this volume), their physical components all take place within the same physical environment and a shared spatial and temporal frame. A real-space blend lives in this shared physical space, and is thus accessible to all modalities simultaneously. At this level of analysis, hand positions, drawn lines, and written text are simply different ways of locating meaning in space; and hand motions, acts of drawing and writing, and speech utterances are ways of locating meaning in time (Hiraga, 2005; Sweetser, 2006). Spatial and temporal relationships can hold between any and all of them, and are used to link together these disparate representational tools into a coherent, meaningful blended space.

Familiar examples of this phenomenon include the co-timed production of speech and gesture – a temporal relationship indicating a shared meaning – or a pointing gesture directed at a diagram – a spatial relationship used to indicate some conceptual entity. In (3) and (4), our interlocutors return to the diagram drawn in (1), and produce an elaborate complex of representations around it involving speech, gesture, writing, and drawing.

The problem they are working on requires that they find a function that sends all the points on the complex plane onto the logarithmic spiral, and that also satisfies some other conditions.[6] One participant begins by proposing a candidate function:

(3) I'm thinking we're gonna send all the [**circles**]
both hands up and rounded to form circle, Fig. 3b
...like e to the t e to the i theta are all going to go to these
writes "$e^t e^{i\theta}$" *writes "$\longrightarrow$" pointing at "$e^\theta e^{i\theta}$", which is already on the board, Fig. 3c*
you know what I'm sayin?
moves to other side of board
like... all the circles[**...**]
traces a circle on board over spiral with pen, but doesn't draw, Fig. 3d
are gonna go to[**...**]the **point** on the circle that
points at "$e^\theta e^{i\theta}$", Fig. 3e
points in the air with the tip of the pen
is [**e to the theta times e to the i theta**]
jabs the air four times, approximate 2-d positions shown in Fig. 3f
In 3-d, it's clear the positions are laid out $1^2 3^4$, schematically similar to $e^\theta e^{i\theta}$.

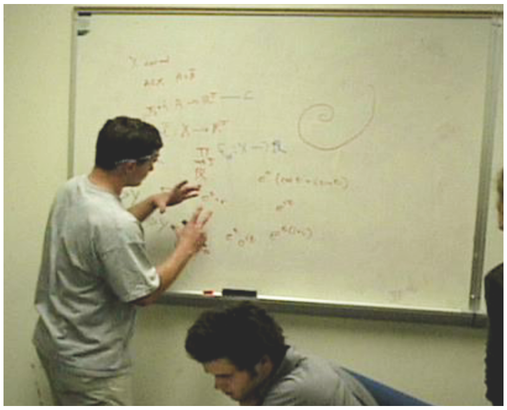
His proposed function maps all the points in each circle centered on the origin to the point where that circle intersects the logarithmic spiral. To explain this idea, he produces a flurry of different representations. I count six representations of the circles:

- the phrase *all the circles*;
- an iconic circle gesture (Fig. 3b);
- the phrase *e to the theta e to the i theta*, co-timed with:
- writing $e^t e^{i\theta}$, the formula for a circle with radius t;
- the phrase *all the circles* again;
- tracing a circle in the air over the spiral (Fig. 3d, again taking advantage of a visible entity to anchor a larger blend; here the presence of the spiral implies the complex plane around it).

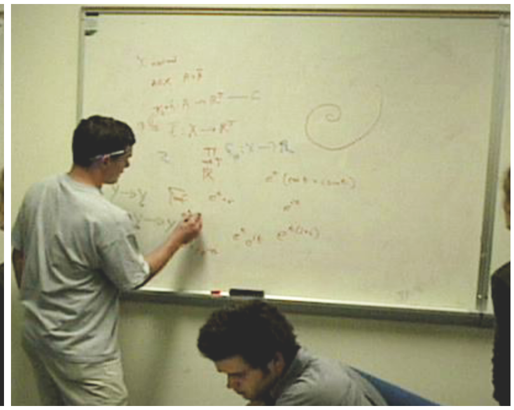
Three representations of the mapping:

- the phrase *are all going to go to*,
- co-timed with an arrow drawn on the board;
- the same phrase again but reduced, *are gonna go to*.

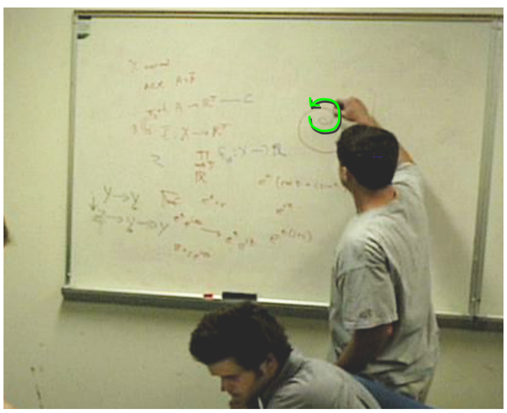
[6] Specifically, it must be a "retraction"; they eventually prove that no such function exists.



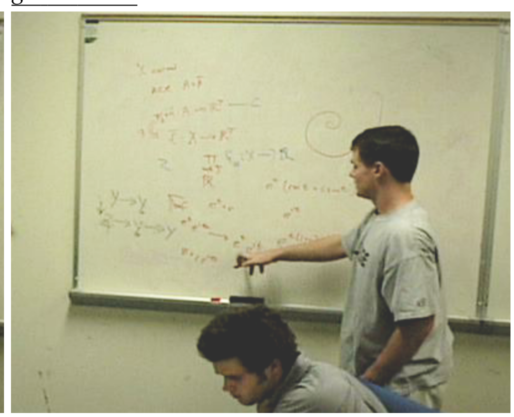
(a) *all the* ***circles***



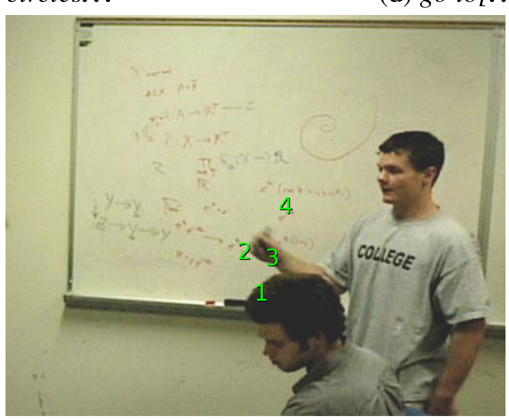
(b) *e to the t e to the i theta* *are all going to go to these*



(c) *all the circles*...



(d) *go to[...]the point*



(e) ***e to the theta times e to the i theta***

Figure 3

And seven representations of the intersection point:

- the anaphoric phrase *these* co-timed with:
- the drawing of an arrow pointing at the text $e^{\theta}e^{i\theta}$;
- a gestural point at this text (Fig. 3e);
- an iconic gesture co-timed with:
- the word *point*;
- the phrase *e to the theta times e to the i theta* co-timed with:
- a schematic representation of the spatial layout of this formula gestured into the air in front of him (Fig. 3f).

Several of these actions depend on pre-existing real-space blend structure in his environment: the writing of the formula and its arrow pointing to an existing formula (Fig. 3b), and the gestural point to this same formula (Fig. 3d); and the gesture over the existing spiral diagram (Fig. 3c). Notice that these are performed without any regard for the apparent boundaries between modalities. It's possible that for some purposes we might want to analyze the lower portion of the whiteboard, with its equations, as a different space from the upper portion, with its diagrammed spiral, and different again from the iconic circle gesture performed in peripersonal space (Fig. 3a). But such a division is entirely orthogonal to the division between modalities: in all of these spaces we see the same principles of spatial and temporal relationships being used to bind meaning across modalities that include speech, gesture, writing, and drawing (J. B. Haviland, 2000; J. Haviland, 2007; Goodwin, 2007).

This explains some key parts of how the many representations we counted above gain their individual meanings. But how is it that he produced sixteen representations, and yet communicates a single coherent idea of circles mapping to points according to a single rule? How does he bind these different representations together? (Nathan & Alibali, 2011) Some of the linking is done locally by co-timing, and by the word *like* which indicates syntactically that the meaning of the phrase *e to the theta e to the i theta* is an expansion/clarification of the meaning of the phrase *all the circles*. But then overlaid on this local structure is a larger repetition (Du Bois, 2010). The phrase *you know what I'm sayin?* and his movement from the left to the right half of the frame divide the explanation into two parts. Both parts first describe the circles, and then their mapping, and then the points they map to. This structural repetition, together with individual repeated fragments (*all the circles* versus *all the circles*, *are all going to go to* versus *are gonna go to*, the repeated use of the written $e^{\theta}e^{i\theta}$) make clear that the two overt descriptions in fact refer to the same underlying concept, and similarly for the corresponding pieces of the two descriptions. Note that these are repetitions of physical form – when he repeats a phrase, we notice the repetition because he literally makes similar sounds with his mouth. That is, this is yet another relationship between forms which is mapped into relationships between meanings via real-space blending. It's by tracing multiple such formal relationships that we come to understand that the circle gestures in Fig. 3a and c refer to the same entity, and the formula in Fig. 3b likewise.

Another interesting possible example of form repetition used to link concepts occurs purely within the mathematical notation used in this example. The scope and quantification of the variable

θ is quite ambiguous. For $e^t e^{i\theta}$ to be a circle, t must be fixed while θ varies; for $e^{\theta} e^{i\theta}$ to be a point, θ must be fixed. As a result, there's no mathematical formalism that lets us define θ so that the formula $e^t e^{i\theta} \to e^{\theta} e^{i\theta}$ makes sense as a mapping of a circle to a point. Yet this doesn't seem to bother the participants. Perhaps they note that the formulas are identical except for first exponent, and understand this as defining an identity constraint, just like the speech's syntactic repetition creates identity mappings between analogous elements: for each circle with its varying θ, pick the value of θ where $t = \theta$.

Returning to the episode illustrated in Fig. 3, thirty seconds later, the original speaker and one of his collaborators begin to find reasons why the mapping he's suggested cannot fulfill the requirements set out in their homework problem:

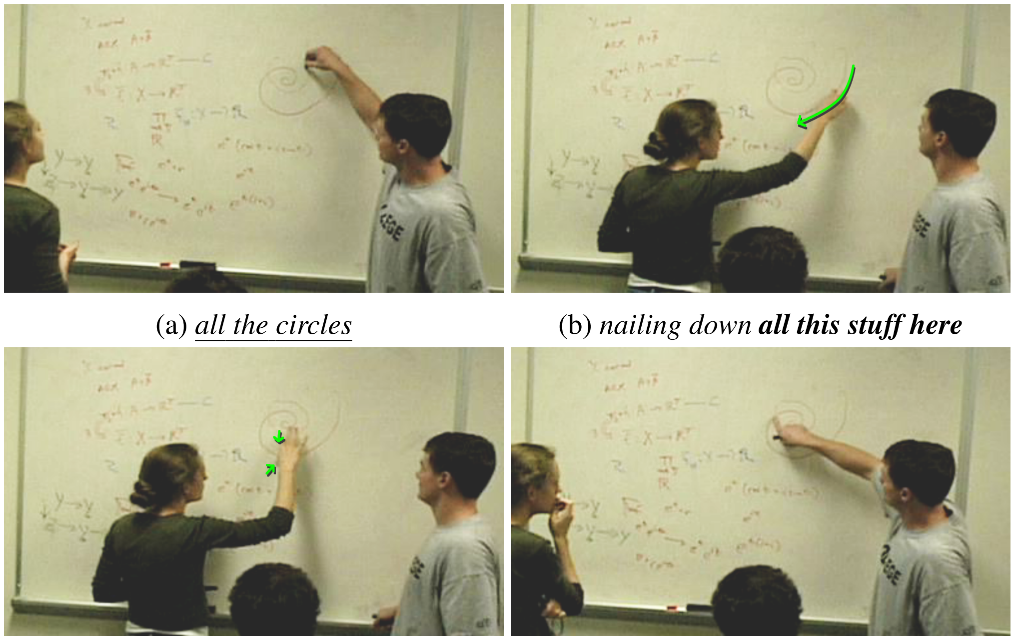
(4) **R** the pre-image of the [**origin can't just be** the] origin
points with pen at center of spiral on board, multiple beats
...
so that can't be right, like
we can't– [we can't be sending <u>all the circles</u>] to–
draws a circle on the board, along same path as had gestured a circle earlier, Fig. 4a

L cause it seems like you're like [nailing down **all this stuff here**
O hand taps down along spiral, Fig. 4b
and then you're gonna have to **break something apart too**] <XXX>
hand stops moving, uses fingers as pincers to "pull"
bits of space toward the spiral, Fig. 4c

R well I was thinking like we could send [all of the **circles**]
index finger traces previously drawn circle
[to the unique **point** that this thing intersects]
index finger stops on intersection of drawn circle and drawn spiral, Fig. 4d
cuz [the **spiral**] intersects a circle at one and only one point
traces spiral in air in front of him with index finger

Let's compare Fig. 3c to Fig. 4a. Both occur co-timed with the initial portion of schematically similar utterances, *all the circles…to the point*. (In the second case he is interrupted half-way through this utterance, but we can see what he intended by how he resumes the utterance at the end of the example.) Both involve very similar motions: tracing the pen in a circular motion over the spiral diagram. The difference is that in the first case, the pen is held just above the surface of the whiteboard, so that no mark is made; in the second, the pen is lowered slightly, so that a trace is left behind, and the motion is somewhat slower and more deliberate. Why is this?

Consider that for the students' purposes, the precise size of the circle is irrelevant – it's a generic representative of the class of all circles that share the same center, so the motions could just as easily have been a bit bigger or smaller. And a gesture, being both schematic and transient, leaves vague the exact size of the circle that is intended. So while the gesture establishes that there is *some* point where the circle and spiral intersect, it does not establish which precise physical, spatial location that point should be blended onto, and in particular, leaves no way to later determine whether any particular location corresponds to this special point. The first time, this is irrelevant: while the student immediately produces a reference to this point, his elaboration uses the formula



(a) *all the circles* (b) *nailing down* ***all this stuff here***

(c) ***break something apart too*** (d) *the unique* ***point***

Figure 4

(pointing at $e^{\theta}e^{i\theta}$, saying *e to the theta times e to the i theta*), not the diagram. The second time, though, when he eventually gets a chance to finish his sentence and say *to the unique point*, he now elaborates by directing a gesture at the particular location in the diagram where the circle and spiral lines intersect. This reference succeeds only because the circle he has drawn has elaborated the spatial structure of the blend, creating a stable and unambiguous perceptual anchor for this intersection point, in a way that the gesture did not.

Clearly his reasoning did not proceed as follows: first, deciding which modality to use (gesture or drawing); second, deciding what kind of gesture or drawing to make. In fact, the drawing has more canonical 'gesture-like' features than the actual gesture, including both a pre-stroke hold and simultaneous production with a co-referential syntactic constituent. Instead, it appears as if he decided that he wanted to exploit this real-space blend to spatially describe an abstract circle, and then compared the affordances of these two different methods of tracing the circle: pen up is quicker, more vague, and more transient, thus cluttering the diagram less; pen down is slower, but having the side-effect of leaving a precise and stable anchor for the elaborated structure. Which choice is optimal in any particular situation depends not on the *current* referential goal – they both refer to the same circle via the same real-space blend – but on whether having a stable anchor will be useful for *future* actions. And in each case, he does in fact choose the option which matches the future actions which he does in fact perform. It seems unlikely that this is an accident. Rather, I would argue, it demonstrates the strategic selection between different strategies for manipulating the blended space on the basis of current discourse goals, and plans for achieving those goals.

Finally, consider the interruption. For our purposes, the interesting thing about this is that speaker **L** simply moves into the real-space blend spaces and uses speaker **R**'s diagram as a substrate for making her own statements, in a way reminiscent of what Furuyama (2000) refers to as "collaborative gestures", in which one person reaches into another's gesture space to manipulate

Figure 5: Example (5): ***pshew pshew pshew***

the structure of their blend directly. These blends are shared not just between modalities, but also between interlocutors.

In (5) we see something that blurs the line between gesture, writing, and drawing to invisibility — what one might call 'gestural drawing'. Here the speaker starts with markings on her paper corresponding to a group T that is 'acting' on a point. The idea of a group acting on a point is that each element of the group somehow modifies the point, so that from one point arises a whole collection of differently modified points:

(5) **L** you're taking it [**up...**]
draws an arrow showing the movement of a point

R you're moving it somewhere else

L acting on it with T, [**...pshew pshew pshew pshew pshew pshew**]
repeated strokes across paper from T to point, Fig. 5
and then bringing it back.

Speaker **L** draws multiple lines from the place corresponding to the group towards the place corresponding to the point. These lines are drawn in a very fast and haphazard fashion, which makes no sense in terms of drawing, for a line is the same whether drawn quick or slow – but if one conceptualizes the process as a series of group elements hurtling out of T and smacking into the point, then the motion does code appropriate force dynamics and imagistic structure. Furthermore, they are drawn on top of or nearly on top of each other, which again makes little sense if one's motivation is simply to draw a particular picture, but does accurately represent the motion of multiple elements whose paths share start and end points. It's even unclear whether this is drawing or writing, since arrows are such a common conventionalized mathematical notation, and notational arrows can do all the normal things arrows can do in diagrams. Is this gesture or goal-directed action? In this instance, the distinction doesn't seem useful.

Figure 6: ...[...]...

Meanwhile, she uses speech to produce co-timed sound effects, which iconically represent repeated fast metaphoric motion. Here the overall timing is driven by an imagined activity, and both the hand movements and speech are aligned to and represent it. The standard terminology for real-space blends seems particularly inadequate here, since 'space' is so clearly just one aspect of the physical environment that is participating in the blend (a residue of the theory's sign-language heritage). Perhaps it's a real-space-and-sound blend.

4 Thinking in a real-space blend

Interlocutors share a physical environment, but, lacking telepathy, do not directly share a mental environment. This makes real-space blends, which map rich mental constructs into the shared physical environment, very useful for accomplishing communicative goals. Theoretically, though, this is not a requirement; real-space blends can and do exist outside of communicative contexts. In some ways, this is a commonplace observation: of course people draw diagrams when solving problems alone, and there are many cultural artifacts whose entire purpose is to encode meaningful structure into their physical structure: clocks, maps, slide rules, protractors, calculators, and so forth (Hutchins, 2005).

Because real-time interactional communication uses the same cognitive tool of real-space blending, it is perhaps not a surprise that people also use ad hoc, iconic/metaphoric gestures to support their own thinking (Chu & Kita, 2008; Rodríguez & Palacios, 2007). In (6), the two students are working independently in their notebooks. Then one of them pauses her writing, raises her hands in a hold gesture, and remains motionless for thirty seconds, before resuming writing:

(6) **R** ...[...]...
30 second silent hold, Fig. 6

At no point during this does her partner look up or interact in any way. This appears to be an anchoring gesture, like those we saw in (2), used to maintain a blend that she is using for some cognitive purpose of her own. While the absence of speech makes it difficult for us as analysts to know what exactly this gesture means to her, it's intriguing to observe that this gesture is very similar to one that she produced approximately one minute earlier, while they were still discussing the problem:

(7) **R** yeah it's [basically it's like taking an **open interval**

Figure 7: *it's [basically it's like taking an* ***open interval***

forms a two-handed hold, Fig. 7
rotates both hands back-and-forth, making a beat

While this occurred a minute earlier, it is the last-but-one utterance before they turned to their notebooks to work independently, and highly suggestive that the real-space blend that she uses for cognitive purposes in (6) is actually the same real-space blend that she using earlier for communicative purposes in (7). If the very same blend can be used for both purposes, though, then we have a new analytic puzzle: when we observe people using such representations within a problem-solving discourse, then how can we tell whether they are being used for communicative (other-directed) or cognitive (self-directed) purposes? Many meaningful actions might even serve both purposes simultaneously.

5 Conclusion

A real-space blend is a structure-preserving map between physical and conceptual realities, and is a generic cognitive tool which underlies many different modalities of meaningful action, including at least speech, gesture, writing, and drawing. These blends often persist through time, and allow for rich, fine-grained interaction between modalities, between interlocutors, and support both communicative and cognitive activity. To be useful, though, they must be managed: blends must be constructed, maintained, elaborated, and deconstructed, and many discourse actions are systematically chosen to accomplish these tasks simultaneously with accomplishing more immediate communicative goals.

When looking at a single gesture in isolation, it is the stroke that catches our attention. But in the context of an ongoing discourse, every word, gesture, and stroke of the pen is performed into a physical environment whose structure has been richly overlaid with meaning, and every such action also further modifies this environment, affecting future actions. In many cases, it may be more important to understand the space in which a gestural stroke occurs, and the way in which it modifies that space, than it is to understand the stroke itself.

References

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, *137*(4), 706–723.

Coulson, S. (2001). *Semantic leaps: Frame-shifting and conceptual blending in meaning construction*. Cambridge: Cambridge University Press.

Dancygier, B., & Sweetser, E. (2009). *Mental spaces in grammar: Conditional constructions*. Cambridge: Cambridge University Press.

Du Bois, J. W. (2010). *Towards a dialogic syntax.* Unpublished manuscript, UC Santa Barbara.

Edwards, L. D. (2009, March). Gestures and conceptual integration in mathematical talk. *Educational Studies in Mathematics*, *70*(2), 127–141. doi: 10.1007/s10649-008-9124-6

Enfield, N. J. (2003). Producing and editing diagrams using co-speech gesture: Spatializing nonspatial relations in explanations of kinship in Laos. *Journal of Linguistic Anthropology*, *13*(1), 7–50.

Enfield, N. J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press.

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

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

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

Goodwin, C. (2007). Environmentally coupled gestures. In S. D. Duncan, J. Casssell, & E. T. Levy (Eds.), *Gesture and the dynamic dimension of language* (pp. 195–212). Amsterdam: John Benjamins Publishing Company.

Haviland, J. (2007). Master speakers, master gesturers: A string quartet master class. In S. D. Duncan, J. Casssell, & E. T. Levy (Eds.), *Gesture and the dynamic dimension of language* (pp. 147–172). Amsterdam: John Benjamins Publishing Company.

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

Hiraga, M. K. (2005). *Metaphor and iconicity: A cognitive approach to analysing texts*. Basingstoke, Hampshire: Palgrave Macmillan.

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics*, *37*(10), 1555–1577.

Lakoff, G., & Núñez, R. (2000). *Where mathematics comes from*. New York, New York: Basic Books.

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

Nathan, M. J., & Alibali, M. W. (2011). How gesture use enables intersubjectivity in the classroom. In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 257–266). Amsterdam: John Benjamins Publishing Company.

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

Rodríguez, C., & Palacios, P. (2007). Do private gestures have a self-regulatory function?: A case study. *Infant Behavior and Development*, *30*(2), 180–194. doi: 10.1016/j.infbeh.2007.02.010

Smith, N. J. (2003). *Gesture and beyond*. Undergraduate thesis, Program in Cognitive Science, UC Berkeley.

Sweetser, E. (1999). Compositionality and blending: semantic composition in a cognitively realistic framework. In G. Redeker & T. Janssen (Eds.), *Cognitive linguistics: Foundations, scope, and methodology* (pp. 129–162). Berlin: Mouton de Gruyter.

Sweetser, E. (2006). Whose rhyme is whose reason? sound and sense in cyrano de bergerac. *Language and Literature*, *15*(1), 29–54. doi: 10.1177/0963947006060553