

## GESTURE AND LANGUAGE DIALECTIC

David McNeill

University of Chicago

A classic psycholinguistic problem is *parole*; in its modern formulation, performance. This paper looks at aspects of *parole* or performance from a new direction, arguing that it is organized as a dialectic of unlike cognitive modes, a dialectic of gesture imagery and language. This dialectic is the key to the evocation, organization, and ultimate execution of meaningful actions shaped to take linguistic form in discourse — imagery, as embodied in the gestures that accompany speech, and the categorical-sequential thinking of speech itself. Both modes are essential components of psycholinguistic performance. A dialectic moreover naturally incorporates the context of speaking in an integral way and we shall develop a theoretical description of context as part of the analysis of performance. Such a dialectic model is self-motivating, self-directed and does not need a guiding hand (hidden or otherwise) of an executive.

### GESTURES

This initial section aims to explain something about gestures themselves and lays the groundwork for subsequent sections concerned with the combination of gestures and speech dialectically. For more detail, see McNeill (1992).

#### **The Problem of Definition.**

The word *gesture* is imprecise. It covers a host of phenomena, with different functions and plausibly different underlying processes of evocation and organization. Gestures of the kind referred to in this paper were called *gesticulations* by Kendon (1972). Gesticulations are distinct from *emblems* — codified hand signals such as contacting the index finger and thumb with the remaining fingers extended for OK. Emblems have established standards of form — contacting the second finger and thumb for example may show precision but it is not the OK sign. In contrast, speakers create gesticulations on the fly, as they are speaking. Speakers can do this precisely because such gestures are *not* constrained by conventions of form. Gesticulations are driven by current meanings, not by conventions. To ask if gesticulations are well-formed is missing an essential point about them, which is that they are idiosyncratic and created at the moment of speaking. Since only gesticulations are under consideration in this paper, I shall continue to refer to them simply as *gestures* except when confusion might arise.

#### **An Example.**

The following gesture, together with speech, shows an individual's thinking at a specific moment. The speaker was describing an episode from an animated cartoon that she has just seen. A character is entering a drainpipe on the side of a building and climbing up it on the inside. The speaker described this with, and he goes **up through** it this time. Synchronously, in the boldfaced section, she raised her right hand upward, her palm up and fingers and thumb spread apart: a kind of open basket shape moving up (see Figure 1). The gesture embodies several ideas — the character (the hand itself) rising

up (the trajectory) and interiority (the open shape). Such a combination of meanings in a single symbol is imagistic in that ideas that in speech require separation in time are concentrated and instantaneous in the gesture. The gesture does not seem to have been a conscious communicative effort. The speaker's gaze was not directed at it and there was no temporal separation or any other indexing of it. It was an unconscious movement with semantic significance that was produced with the co-expressive speech conveying the same significance.



Fig. 1. Gesture of the kind discussed in this paper. The gesture depicts a story event in which a character is climbing up a drainpipe on the inside.

When we see a gesture like this, what are we actually seeing? Motion; something going on while the person is speaking to be sure, but I shall propose that the gesture and its synchronous speech are also components of a dialectic and merge into minimal units termed growth points. Growth points (GPs) combine opposite modes of thinking — image and language. The combination creates instability: instability, in turn, propels thinking and speaking forward, and seeks a state of repose.

Insight into what gesture might contribute to linguistic performance can be derived from a description a century ago by Wilhelm Wundt. Wundt observed how sentences include two modes of cognition simultaneously; there is the simultaneous and the sequential :

From a psychological point of view, the sentence is both a simultaneous and a sequential structure. It is simultaneous because at each moment it is present in consciousness as a totality even though the individual subordinate elements may occasionally disappear from it. It is sequential because the configuration changes from moment to moment in its cognitive condition as individual constituents move into the focus of attention and out again one after another. (Translation by Blumenthal 1970, p. 21).<sup>1</sup>

When co-expressive speech and a gesture synchronize, we see something that is both simultaneous and sequential, as Wundt envisioned. There is a combination of two semiotic frameworks for the same underlying idea unit each with its own expressive potential. Speech and gesture are co-expressive but semiotically non-redundant.

How do the semiotic frameworks differ? One way, obviously, is how they distribute information in time. In language, ideas are separated and arranged sequentially but in gesture they are instantaneous — the ideas of upward motion and interiority were distributed across time in the words, up and through but were concentrated at a single

---

<sup>1</sup> I am grateful to Zenzi Griffin for drawing my attention to this passage.

instant in the gesture. As Wundt described, these contrasting modes are themselves simultaneous, and exactly so.

A fuller picture of how gesture and speech differ as semiotic modes is conveyed by Kendon's Continuum (McNeill 1992), a series of distinctions among different kinds of gestures but also distinctions that contrast gesticulations to linguistic codes. Gestures of the gesticulation type differ from linguistic codes (including ASL and other signed languages) on a number of levels:

- *In gestures, meanings are global*: the parts of the gesture derive their meanings from the meaning of the gesture as a whole. In Fig. 1, the hand signified the character but the hand could have been something else in another gesture. It was the character in this gesture because the meaning of the gesture as a whole was the character climbing up the inside of a pipe. In speech, the presentation is analytic and combinatoric. Separately meaningful morphemes are combined according to grammatical rules to make up the whole.
- *The form of the gesture lacks form-meaning conventions*. Form and meaning are created at the moment of speaking. Speech is socially-constituted and held to conventions of form-meaning pairings.
- *The gesture is imagery*. Speech is an arbitrary signifier-signified combination.<sup>2</sup>

The upshot is that the synchrony of gestures and speech puts different semiotic modes together at the same moment. The synchronized modes are opposites in multiple ways, and the combination sets up the conditions of a gesture-language dialectic — global meaning with analytic meaning; idiosyncratic symbolic forms, created on the fly, with prespecified form-meaning pairings; forms motivated by imagery with forms motivated by conventions. Fig. 1 illustrates all of these dualities: the gesture parts (the hand, the trajectory, the shape) are meaningful only because of the meaning of the whole, while in speech the words, *goes*, *up* and *through* are meaningful on their own, and by combining grammatically compose the meaning of the whole; the gesture is idiosyncratic, had just been created, and was original both as a form and as a conceptual object (a new idea of *upward hollowness* or some such); and the gesture is motivated by imagery, in contrast to the largely arbitrary morphology of words.

### **Unsplitability.**

The dialectic fostered by the *sequential* and the *simultaneous* and the dualities in Kendon's Continuum, is reinforced by the tightness of the bond between speech and gesture: they are *unsplitable*. Unsplitability shows itself in many psycholinguistic facts. In each of the following some force affects speech or gesture but fails to split them apart. Either they continue to be produced together or are suppressed together: either way, they remain glued together. (a) Delayed auditory feedback does not disrupt the synchrony of speech and gesture, despite a major impact on speech itself (McNeill 1992, first DAF experiment). (b) Clinical stammering is incompatible with gesture strokes, the beginning

---

<sup>2</sup> Elisabeth Engberg-Pedersen points out that gestures can also be partly conventional. Metaphoric gestures, e.g., *eidola* emanating from the head, embody cultural conventions, here, mental states as radiating beams (cf. McNeill 1992). On the language side, ideophones are partly iconic — *ban* in Japanese conveys impact iconically (cf. Kita 1993). Thus there may be a continuum of conventionality and one can say that language and gesture typically express the same underlying meaning at different points on the scale, and this is why the combination of speech and gesture is more than mere repetitiveness.

of the stroke inoculating against stammering, and if a stroke is ongoing and stammering commences the stroke immediately stopping (Mayberry & Jaques 2000). (c) Congenitally blind speakers perform gestures even when speaking to blind listeners (Iverson & Goldin-Meadow 1998). And, finally, (d) speech and gesture remain bound together in the total absence of proprioceptive feedback, even without vision (Gallagher et al 2002).

### **Gesture Inhabiting Language.**

Finally, we can draw on Merleau-Ponty (1962) for a deep insight into the duality of gesture and speech and what we can expect the role of gesture in a dialectic to be. Gesture, the instantaneous, global, non-conventional component, is not an external accompaniment of speech, the sequential, analytic, combinatoric component, but inhabits it:

The link between the word and its living meaning is not an external accompaniment to intellectual processes, the meaning inhabits the word, and language is not an external accompaniment to intellectual processes.<sup>3</sup> We are therefore led to recognize a gestural or existential significance to speech. Language certainly has inner content, but this is not self-subsistent and self-conscious thought. What then does language express, if it does not express thoughts? It presents or rather it *is* the subject's taking up of a position in the world of his meanings (p. 193; emphasis in the original).<sup>4</sup>

The dialectic unit, the GP, which makes static language spring into motion, is a mechanism for this existential content of speech — this taking up a position in the world. Gesture in the GP carries out the living meaning that inhabits the word (and beyond, the discourse). A deeper answer to the query therefore — when we see a gesture, what are we seeing? — is that it is part of the speaker's current cognitive being, her very mental existence.

### **STATIC AND DYNAMIC VIEWS**

Historically, the view of language most prevalent in linguistics traces back to Ferdinand de Saussure in the early 20<sup>th</sup> C (Saussure 1959 [originally 1915]). In this view language is a static system. Many areas of linguistics remain essentially Saussurian, static, to the present day. To quote one authority:

Ferdinand de Saussure is the father of modern linguistics, the man who reorganized the systematic study of language and languages in such a way as to make possible the achievements of twentieth-century linguistics (Culler 1976, p. xiii).

Saussure's goal was to capture the *fundamentals of language form* in terms of a set of dichotomies or interlocking axioms:<sup>5</sup>

- To see *langue* (competence) a static view is crucial.

---

<sup>3</sup> Merleau-Ponty's quotation is from Gelb and Goldstein, 1925, p. 158.

<sup>4</sup> I am indebted to Jan Arnold for this quotation.

<sup>5</sup> This section draws on work by Katherine Greeno at the University of Chicago in the late 1970s.

- All the Saussurian features of *langue* — *langue* rather than *parole*, synchronic rather than diachronic, oppositions, the arbitrariness of signs, and language as a social fact — imply a static view.
- The concept of competence, in which *langue* is attributed as a property of the individual, also implies the same static view.

The most important axiom is the first: the synchronic view that is necessary to observe *langue*, the entire system of language. This axiom is the source of much linguistic methodology but it is not just a methodological requirement. The synchronic view is also a substantive claim: the static dimension of language, it says, has ontological reality. At any moment, the system of language *is* a stable, listable collection of entities realized as socio-cultural facts in the Saussurian conception. Somewhat paradoxically, a dynamic view of language can absorb the ontological reality of the static dimension. If we restore the lost social fact axiom of Saussure, omitted in the transmutation of *langue* into competence as a psychological phenomenon, we shall discover a functional role of static *langue* in a dynamic model of an imagery-language dialectic. If the Saussurian tradition is reaching an impasse, a breakthrough might come over two routes. One would attempt to dissolve the competence-performance distinction and replace it with a new paradigm in which features of *langue* and *parole* are unified. The other would be to retain the notion of competence as distinct from performance, but make it into a special case. Scott Liddell (2000) has been developing an approach of the first type (also see Armstrong et al. 1995). The dialectic is an approach of the second type.

In a dialectic model, a dynamic view is of the essence. A historical figure who well articulates this view is Vygotsky (1986, 1978, based on texts written in the 1930s). Some quotes will give the general idea. In the first, meaning is identified with acts of thought ; in the second, the dynamic dimension of language itself is brought in:

*Meaning is an act of thought*

[M]eaning is an act of thought in the full sense of the term. But at the same time, meaning is an inalienable part of word as such, and thus it belongs in the realm of language as much as in the realm of thought (1986, p. 6).

*Meaning is dynamic*

[W]ord meanings develop. Word meanings are dynamic rather than static formations. They change as the child develops; they change also with the various ways in which thought functions. This insight must replace the postulate of the immutability of word meanings. (1986 p. 212, 217).

A critical further point is the distinction between what Vygotsky called units and something else he termed elements . In modern terms, confusingly, Vygotsky's elements are what, in information processing models, would be units — the ultimate primes to which the model reduces or deconstructs the process that it is modeling (cf. Palmer and Kimchi 1986). To Vygotsky, a reduction to such ultimate elements destroys the very phenomenon to be modeled. A psychological unit, he insisted, is the smallest component that *retains the property of being a whole*:

The first method analyzes complex psychological wholes into elements. In essence, this type of analysis, which leads us to products in which the properties of the whole are lost, may not be called analysis in

the proper sense. It is generalization, rather than analysis. [I]nstead of enabling us to examine and explain specific instances and phases, and to determine concrete regularities in the course of events, this method produces generalities pertaining to all speech and all thought. It leads us moreover into serious errors by ignoring the unitary nature of the process under study. Psychology, which aims at a study of complex holistic systems, must replace the method of analysis into elements with the method of analysis into units. (1986 pp. 4-5).

A GP is a minimal psychological unit in Vygotsky's sense. It is the smallest unit of an imagery-language dialectic with the essential property of the whole dialectic. The dialectic model below describes acts of meaning in a moving system and resolves these into minimal units ( growth points ). In the dialectic model, the static dimension is a special case — the point of repose toward which dynamic change tends—the end-point of the process.

## IMAGERY-LANGUAGE DIALECTIC

### The Dialectic

A dialectic implies a) *a conflict or opposition* of some kind, and b) *resolution of the conflict through change*. A dialectic is thus inherently dynamic and presupposes Vygotsky's concept of a unit — the smallest whole. Vygotsky invoked a dialectic in the following:

The relation of thought to word is not a thing but a process, a continual movement back and forth from thought to word and from word to thought. In that process, the relation of thought to word undergoes changes that themselves may be regarded as development in the functional sense. Thought is not merely expressed in words; it comes into existence through them. (1986 p. 218).

The combination of imagery and language in a dialectic is inherently unstable. Instability implies repose, but this repose is also a new beginning. A dialectic therefore also implies *cycles* in which instability is followed by repose, this by new instability, and the process repeating indefinitely (cf. the idea of a cycle in McNeill 1992). Such cycles echo William James' metaphor of the objects of consciousness and the semi-conscious transitions between them.

### Relation to Linguistic Form.

How does a dialectic relate to the formal aspects of language? Formal linguistic properties, as they emerge, define the *outcome* of the dialectic. This outcome has psychological reality in the form of growing linguistic intuitions of completeness. A grammatically complete sentence (or its approximation) corresponds to intuitions of stability: grammatical form is the socially-constituted state of repose *par excellence*, reached after dialectic instability. Synchronic linguistic description is an analytic reconstruction of such intuitions. It displays the standards that a speaker's speech actions are held to. A different metaphor is needed for this role of grammar as outcome. Clearly, grammatical form is not input; linguistic properties instead function as a kind of stop-order. These stop-orders arise out of the speaker's growing intuitions of grammatical completeness (without significant meaning distortion) as the dialectic moves forward. In

this model, imagery animates syntactic form: forms (and their intuitions) say when the generation process closes down.<sup>6</sup>

**Example stopping point effectiveness.** By effectiveness I mean the emergence of intuitively complete structures that assert themselves as stopping points even when the process of change can go further. Elena Levy (pers. com.) is carrying out a research project in which she collects (from child speakers) repeated narrative retellings of an animated cartoon stimulus. A child recounts the same story on successive days (each day to the same listener, a feature that encourages increasingly compact descriptions). Looking across the days, we see different stopping points, with each day's stopping point a well-formed structure (intuitively recognizable) but not the structure the child is ultimately able to reach.

Day 1 found a stopping point in two non-successive sentences that contained all the relevant information. Day 2 compacted this into two successive sentences, and Day 3 into a single sentence. In each case, we can interpret the stopping point as an intuitively recognized well-formed structure. The rule seems to be that the dialectic stops if it yields an intuitively complete form. Day 2 is the clearest illustration. It shows compactness but is not a complete transition. The following is the record (a girl, 7 years old):

Day 1	Day 2	Day 3
1) and there was this um there was this big box	1) and there was this big box	1) and there is a big box full of ice
2) and he sat in there		
3) and it was nice and cold		
4) and there was ice in there for him	4) and then um and it was full of ice	

**Goals of linguistic description.** What kind of linguistic form description meshes most gracefully with imagery and best defines the endpoint of a dialectic?

First, it must transparently exhibit the speaker's intuitive sense of well-formedness. Ever since Chomsky's breakthroughs (cf. Chomsky 1965), most synchronic models of the generative variety attempt to systematize the linguistic intuitions of native speakers as data. Such models are directly applicable to the dialectic analysis insofar as they are interpreted as descriptions of intuitions, since they define stop-orders for the dialectic. The field of candidate synchronic models however is drastically narrowed (essentially to zero) by two further goals:

Second, each derivational step must also exhibit well-formedness in that:

- a) it is not a violation of good form and
- b) it opens a commitment to a constituent that is also well-formed.

A third constraint is that each constituent, as it appears, must be ratable as a stopping point in terms of:

<sup>6</sup> A dialectic sifting through a series of possible stopping points and finally achieving an intuitively acceptable solution, is described in Dray and McNeill (1990) (the nurturing example).

- a) how well it corresponds to an intuition of completeness and
- b) how well it corresponds to the opposite, an intuition of incompleteness — leading to a continuation of the dialectic and search for good form.

No current analysis meets all of these requirements, especially the second and third. However, construction grammars (Goldberg 1995, 1997), together with frame semantics (Fillmore 1982), can do the next best thing, which is to offer a way to by-pass them. By treating constructions/frames as basic units, sentence derivation is flattened, and the second and third requirements are effectively by-passed. I shall make use of construction grammar in the case study below to illustrate how an opposition of imagery and linguistic categorial content could have been resolved.

### **MATERIAL CARRIERS.**

We can deepen our description of an imagery-language dialectic further by adding the idea of a material carrier. A material carrier — the phrase was used by Vygotsky (1986)<sup>7</sup> — is the embodiment of meaning in concrete enactments or material experiences and appears to enhance the symbolization's representational power. The concept of a material carrier implies that the gesture, *the actual motion of the gesture itself*, is a dimension of thinking. From this viewpoint, a gesture is an image in its most developed — materially embodied — form. While the material carrier effect has interesting implications for how the brain links speech, gesture and online thinking (cf. Gallagher et al. 2002), the following provides an abstract model of how materialization has an effect on representational power.

### **The H-Model.**

This model is a realization of Merleau-Ponty's existential content of speech (and gesture). It gives the material carrier concept an interpretation on the level of cognitive being. To the speaker, gesture and speech are not only messages or communications, but are a way of cognitively existing, of cognitively being, at the moment of speaking. By performing the gesture, the core idea is brought into concrete existence and becomes part of the speaker's own existence at that moment. The Heideggerian echo in this statement is intended. Gestures (and words, etc., as well) are themselves thinking in one of its many forms — not only expressions — *but thought, i.e., cognitive being, itself*. The speaker who creates a gesture of Sylvester rising up fused with the pipe's hollowness is, according to this interpretation, embodying thought in gesture, and this action — thought in action — was part of the person's being cognitively at that moment. To make a gesture, from this perspective, is to bring thought into existence on a concrete plane, just as writing out a word can have a similar effect. The greater the felt departure of the thought from the immediate context, the more likely is its materialization in a gesture, because of this contribution to being. Thus gestures are more or less elaborated depending on the importance of material realization to the existence of the thought.

There are, however, deep and hitherto unexplored issues here, and possibly some contradictions. If to the speaker the gesture and linguistic form are themselves forms of being cognitively, there would seem to be no room in this process for the presentation of

---

<sup>7</sup> Pointed out by Elena Levy.



symbols; the signifier-signified distinction that constitutes semiosis is lacking. A semiotic relation appears when an observer is taken into account — a listener who participates or a coder who looks at the communicating individual. Dreyfus (1994), in his invaluable exposition of Heidegger, explains Heidegger's treatment of symbols in a way that suggests a rapprochement. To cope with signs is not to cope just with them but with the whole interconnected pattern of activity in which they are embedded (this still has the viewpoint of a listener/observer; from the speaker's viewpoint, we should say that producing a sign *carries the speaker* into a whole interconnected activity). Heidegger, according to Dreyfus, says that signs point out the context of a shared practical activity — and this is the key to the rapprochement. To have your thoughts come to exist in the form of signs is to cause them to exist in a context of shared practical activities. A sign signifies only for those who dwell in that context. In this we can recognize a recipe for a dialectic: sign and image are inseparable and jointly form a context to dwell in. This description brings the GP and the social Other together as joint inhabitants of the context (and it is the speaker who always is the one dwelling there the best). The communication process is then getting the Other to dwell there too. Heidegger spoke of language as the house of being,<sup>8</sup> just as Merleau-Ponty spoke of inhabiting it. Their metaphors dovetail, in that the imagery-language dialectic inhabits this house and brings it to life.

**Werner and Kaplan's Organismic Foundations.** A related dimension of the material carrier is its *organismic* grounding in symbol formation. In an organismic approach, the symbol's material qualities (sound image, gesture action) is an integral part of its semantic qualities (cf. Armstrong et al.'s semantic phonology). It is organismic in that meaning is never dissociated from body; indeed, it is impossible (unthinkable) without body.

Werner & Kaplan (1963) wrote specifically of the organismic foundations of symbol formation; of our capacity to represent the world symbolically by capturing the world in a kind of imitation carried out in bodily movements. Symbols are made possible by:

... this *transcendence of expressive qualities*, that is, their amenability to materialization in disparate things and happenings ... (p. 21, emphasis in the original).

According to Werner & Kaplan, the development of an individual child is a process of adding semiotic *distance* between movement and the expressive qualities that the movement can have:

The act of denotative reference does not merely, or mainly, operate with *already formed* expressive similarities between entities. Through its productive nature, it brings to the fore latent expressive qualities in both vehicular material and referent that will allow the *establishment of semantic correspondence* between the two entities. It is precisely this productive nature of the denotative act that renders possible a symbolic relation between any entity and another. (pp. 21-22, emphasis in the original).

We see this kind of creation of meaning in a gesture-language interaction. As a material carrier a global, instantaneous gesture is a further creation, beyond infancy, of the organismic foundations of which Werner & Kaplan spoke. In the case of gesture, the

---

<sup>8</sup> I am grateful to Shaun Gallagher for calling my attention to this phrase

material carrier interacts with the fully distanced symbols of the linguistic code.<sup>9</sup> The growth point formalizes this interaction of image and linguistic code.

**Embodied Cognition** . Finally, the concept of a material carrier links to the idea of embodied cognition, as described by Johnson (1987) and Lakoff & Johnson (1999). Johnson, though not mentioning gesture, links the embodiment of conception to modes of meaning organization similar to gesture — non-propositional and figurative, with provision for imagery and analogic meaning construction. More recently, Lakoff and Johnson (1999) cite the phenomenon of gesture, but restrict it to gestures embodying metaphors. Nonetheless, a far more wide ranging materialization by gesture appears to be the essence of embodiment.

## THE GROWTH POINT

The *growth point* the minimal unit of an imagery-language dialectic. In terms of the H-model, the growth point concentrates one s momentary cognitive being in a package that has both a linguistic and a gestural component. Imagery is always implied, even when not fully realized in a motor movement.<sup>10</sup> GPs are inferred from the totality of communicative events with special focus on speech-gesture synchrony and co-expressivity. It is called a growth point since it is meant to be the initial form of a thinking-for-speaking unit out of which a dynamic process of organization emerges. It is also called a GP since it is a theoretical unit in which the principles that explain the mental growth of children — differentiation, internalization, dialectic, and reorganization — also apply to realtime utterance generation by adults (and children). A final reason for calling it a GP is that it addresses the concept that there is a specific starting point for a thought. Although an idea unit continues out of the preceding context and has ramifications in later speech, it does not exist at all times, and comes into being at some specific moment; the formation of a growth point is this moment, theoretically.

**Some semiotic observations of growth points.** Semiotically, such a unit combines, as defined, imagery and linguistic information. And this combination implies others. These combinations serve to summarize the dialectic oppositions within growth points. For example, the individually contrived with the socially-constituted, since imagery arises from the individual and language, while emanating from the individual, is

---

<sup>9</sup> The Werner & Kaplan idiom suggests that an interaction of imagery and language is not a given, that young children must develop it. Until there is distancing, there is a single mode of meaning creation, organismic imitation of the world. The other principal implication of this framework is that an organismic mode of meaning creation does not disappear but continues along side the symbolic mode organized in an imagery-language interaction.

<sup>10</sup> A model sharing goals with the GP — that language and gesture should be analyzed together — is Sotaro Kita s information packaging hypothesis (IPH) (see Kita 2000). Despite the IPH name, there are no units in the IPH. Speech and gesture are described as independent streams simultaneously running and intertwining as continuous threads. A speaker can use gesture to prepare information for linguistic encoding, and blockage of language can accordingly be circumvented by gestural imagery. In the GP, imagery and encoded meanings are packaged into bundles (GPs) that have definite onsets and offsets (the GP, while bundled, is far from isolated from context as explained below). It is conceivable that both IPH and GP are valid but under different conditions. Many of Kita s examples arise from situations where a dialectic of imagery and language seems to have broken down or not yet started. Thus the IPH may apply to thinking outside the overlapping Venn diagrams of Vygotsky s verbal thought — In their overlapping parts, thought and speech coincide to produce what is called verbal thought. Verbal thought, however, does not by any means include all forms of thought or all forms of speech. There is a vast area of thought that has no direct relation to speech. Nor is there any reason to derive all forms of speech activity from thought (Vygotsky 1986, pp. 88-89). The GP applies to the overlapping parts. This analysis implies a yet-to-be analyzed continuum of language-thought relationships from the separate to the integrated.

not separate from a cultural tradition. In terms of the static-dynamic distinction, a growth point is the joining of what is realized statically with what is emerging in the speaker's consciousness at the moment. Distinct modes of semanticity combine the global (whole to parts) semantics of imagery with the analytic-combinatoric (parts to whole) semantics of language. The GP is the locus of the duality of which Wundt spoke—the instantaneous and the successive: information in an image is instantaneous but in language information spreads out in time. The GP, for all these reasons, is an inherently unstable combination of unlike cognitive modes. Instability provokes cognitive movement that seeks repose in the form of intuitively complete linguistic forms, with this closure corresponding to the focal, and the movement toward it the transitory phases of Jamesian consciousness. The process repeats cyclically.

Two further properties of the GP are as important. One is that a GP exists in relation to a context; the other is the form taken by this relationship. The context is the subject of the next section. Briefly, a GP models the act of speaking as a kind of perceptual differentiation. The slogan could be *synthesis by analysis* (inverting the old analysis by synthesis slogan, which similarly tried to bring the perceptual and the generative together, but focused upon perception rather than generation). As we shall see, the analysis in the GP is a perception-like differentiation: differentiation forms the GP. Differentiation in turn requires a background, and this is the context. It is clear therefore that a GP *always* implies a context: differentiation demands it. The context must have been one in which the proposed GP could take form. This, finally, has a methodological implication, for a context that is inappropriate falsifies a hypothesized GP (McNeill & Duncan 2000). The conditions that contexts must meet in order for a GP to take form are explained below.

### **Growth Point Differentiation and Context.**

To understand the relationship of a GP to context, it is first necessary to understand how the GP is a product of differentiation. Growth points emerge as the newsworthy elements in the immediate context of speaking. They are, to adopt another concept from Vygotsky (1986), *psychological predicates* (not always grammatical predicates).

The concept of a psychological predicate illuminates the theoretical link between the GP and the context of speaking. Defining a psychological predicate (and hence a GP) requires reference to the context; this is because the psychological predicate and its context are mutually defining. The GP:

1. marks a significant departure in the immediate context; and
2. implies this context as a background.

We have in this relationship the seeds for a model of realtime utterance generation and coherent text formation.

Regarding the GP as a psychological predicate suggests a mechanism of GP formation in which differentiation of a focus from a background plays an essential part. Such differentiation is validated by a very close temporal connection of gesture strokes with the peak of acoustic output in speech. Nobe has documented this connection instrumentally: The robust synchrony between gesture strokes and the peaks of acoustic aspects suggests that the information the gesture stroke carries has an intrinsic relationship with the accompanying speech information prominently pronounced with

these peaks. The manifestation of the salient information seems to be realized through the synchronization of these two modalities (Nobe 1996, p. 35).

Regarding the GP as a psychological predicate also clarifies the sense in which we use the term *context*. This term has a host of meanings (cf. Duranti & Goodwin 1992), but for our purposes *context* is the background from which a psychological predicate is differentiated. This background indexes and is constrained by external conditions, both social and material, but an essential fact is that the background is also under the control of the speaker; *it is a mental construction, part of the speaker's effort to construct a meaning*. The speaker shapes the background in a certain way, in order to make possible the intended significant contrast within it. Background and contrast are both necessary and are constructed together.

I will use the terms *field of oppositions* and *significant (or newsworthy) contrast* to refer to this constructed background and the differentiation of the psychological predicate. All of this is meant to be a dynamic, continuously updated process in which new fields of oppositions are formed and new GPs or psychological predicates are differentiated in ongoing cycles of thinking for speaking.

**Showing differentiation of significant contrasts in fields of oppositions.** That growth points are significant contrasts in fields of opposition can be demonstrated with pairs of examples in which gesture-speech synchronizations differ systematically, dependent upon the field of oppositions, even though the linguistic segments themselves are the same or nearly so. Example (A) was at the end of a series of references to the bowling ball, where what it was doing would have been highlighted. The crucial fact about the gesture is that it included bowling-ball imagery (rolling) and was synchronous with a verb describing it, *rolls*:<sup>11</sup>

and he drops a [**bowl**]ing ball [into the rain spout]  
 [and it **goes** down]  
 and it\* [/] ah\*  
 you [**can t tell**]  
 [if the **bowling ball** /]  
 [is un\* /]  
 [is **under Sylvester**]  
 [or **inside** of him]  
 [but it **rolls him out** ]\* (=A)

Example (B), from a different speaker, appears with the same verb but differs in that it is not about the bowling ball — it lacks manner and is synchronized with the path particle *down*, the direction equally followed by Sylvester. The gesture in (B) suggests a focus on Sylvester and the path he followed, instead of the bowling ball and its manner of motion. The different GPs took form in fields of opposition that can be inferred from the preceding discourse context. Whereas (A) was the culmination of a series of references to the bowling ball, the (B) series began similarly but then shifted to Sylvester and his path, creating a contest in which the bowling ball and its manner of motion were downplayed:

<sup>11</sup> **Notation:** square brackets mark the start and end of a gesture phrase. Boldface is the gesture stroke — the phase with semantic content and the quality of effort. RH, LH and BHs are right, left, and both hands. Asterisks mark self-interruptions. # is an audible breath pause. / is a silent pause. Repeated letters mark lengthenings. Each line contains a gesture phrase. *NumberX* (e.g. 3X) designates the repetition of a gesture stroke.

[the canary] #  
 [throws\*] #  
 [puts a]  
 # [bowling]  
 [ball] #  
 [into # the drain spout as the]  
 [cat is climbing up /  
 and it goes into his]  
 [mouth] / (topic switch to Sylvester)  
 [and of course] #  
 [into his sto]mach #  
 [and he rolls # down the drain spout] (=B)  
 [and across the street]  
 [into the bowling] alley #

### CASE STUDY OF A GROWTH POINT

Thus the GP is a mechanism of verbal thinking in which imagery-language dialectics take place. The narrative text in (1) - (9) below will be used to illustrate a number of features important for explaining the GP, especially the critical role played in this dialectic by the context of speaking. The narrative is by an adult native speaker of English recounting from memory the story of an animated color cartoon that she had just watched (a c. 1950 Warner Brothers Tweety and Sylvester saga). The speaker and her listener were told that the listener would have to retell the story to a third person. There was no mention of gestures. The scene she is describing is outlined in the attached note.<sup>12</sup>

#### Vivian s Battle Plan

(1) he tries going [**up the inside of the drainpipe** and]

*Ihand: RH rises up 3X*

(2) Tweety Bird runs and gets a bowling ba[ll and drops **it down** the drainpipe #]

*Symmetrical: 2 similar hands move down (see Fig. 2)*

(3) [**and / as he s coming up**]

*Asymmetrical: 2 different hands , LH holds, RH up 2X*

(4) [**and the bowling ball s coming d**]

*Asymmetrical: 2 different hands , RH holds, LH down*

(5) [own **he ssswallows it**]

*Asymmetrical: 2 different hands , RH up, LH down*

(6) [ # and he comes **out the bottom of the drai**]

*Ihand: LH comes down*

(7) [npipe and he s **got this big bowling ball inside h**]im

*Symmetrical: 2 similar hands move down*

---

<sup>12</sup> Sylvester crawls in a drainpipe at the bottom; Tweety Bird dumps a bowling ball into the top of the drainpipe; scene shifts to the side of the building, central portion of drainpipe; bulge moves down drainpipe; bulge explodes; Sylvester falls into street with bowling-ball shaped bottom; Sylvester rolls down street on bottom, feet rotating above ground; Sylvester rolls down different part of street, looks upset; scene of empty street; scene shifts to end of street; bowling alley; Sylvester rolls into entrance of bowling alley; entrance of bowling alley without Sylvester; sounds of bowling pins crashing (see McNeill 1992, Appendix, for full transcript; created by Elena Levy).

(8) [and he **rolls on down**] [into a **bowling all**]  
*Symmetrical: 2 similar hands move forward 2X*

(9) [ey and then **you hear a sstri]ke #**  
*Symmetrical: 2 similar hands move apart*

To focus on one item for analysis, consider the utterance and gesture in (2) — the words *it down* plus a downward thrusting gesture. See Fig. 2. My purpose will be to show how this utterance-gesture combination can be explained utilizing the GP model. I have chosen this example in part because the linguistic elements *it* and *down* do *not* form a grammatical entity. While *it* is the direct object of *drops* and *down* is a satellite of this verb (cf. Talmy 1975, 1985), *it* and *down* together do not comprise a grammatical unit. Nonetheless, in this example they are the inferred growth point. The challenge is to explain this GP and show why *it* and no other took form in the context of speaking.

First, to explain the GP itself. The gesture in (2) was made with two symmetrical hands — the palms loosely cupped and facing downward as if placed on top of a large spherical object, and the hands moved down during the linguistic segments *it do(w)n*. The inferred GP is this image of downward movement *plus* the linguistic content of the *it* (i.e., the bowling ball) and the PATH particle *down*. The GP is both image and linguistic categorial content: an image, as it were, with a foot inside the door of language. Such imagery is important, since it grounds sequential linguistic categories in an instantaneous visuospatial context. It may also provide the GP with the property of chunking, a hallmark of expert performance, (cf. Chase & Ericsson 1981), whereby a chunk of linguistic output is organized around the presentation of an image. The downward content of the gesture is a specific case of *down*, the linguistic category a specific visualization of *it* in which imagery is the context of the category and possibly the unit of performance. The linguistic categorization is also crucial, since it



Fig. 2. Downward stroke synchronized with *and drop s it down the drainpipe*.

brings the image into the system of categories of the language, which is both a system of classification and a way of patterning action. The speech and its synchronized gesture are the key to this theoretical unit.

### **Incorporating Context - Catchments.**

A GP is a psycholinguistic unit based on contrast, and this concept brings in context as a fundamental component. As mentioned above, I use the terms field of oppositions and significant contrast to analyze the role of context in thinking while speaking. A significant contrast and the field of oppositions within which it is drawn are linked meaning structures, under the creative control of the speaker at the moment of speaking. Control by the individual ensures that GPs establish meanings true to the speaker's intentions and memory. The formation of a GP is the highlighting or differentiation of what is novel in such a field. The field defines the significance of the contrast, it establishes what is meaningful about it; the contrast itself is the source of the GP.

To explain the utterance in (2) and why it has the growth point we infer, we must consider the complete context above and see how the parts of the utterance came together with contributions from the context. A useful approach to this analysis is by means of *catchments* a phenomenon first noted by Kendon in 1972 (he did not use the term *catchment*). See McNeill et al (2001) for a detailed discussion of the catchment.

**Definition.** A catchment is recognized when one or more gesture features recur in at least two (not necessarily consecutive) gestures. The logic is that the recurrence of an image in the speaker's thinking will generate recurrent gesture features. Recurrent images suggest a common discourse theme. In other words, a discourse theme will produce gestures with recurring features. These gesture features can be detected. Then, working backwards, the recurring features offer clues to the cohesive linkages in the text with which it co-occurs. A catchment is a kind of thread of visuospatial imagery that runs through a discourse to reveal the larger discourse units that emerge out of otherwise separate parts.

By discovering the catchments created by a given speaker, we can see what this speaker is combining into larger discourse units — what meanings are being regarded as similar or related and grouped together, and what meanings are being put into different catchments or are being isolated, and thus are seen by the speaker as having distinct or less related meanings. Individuals differ in how they link up the world into related and unrelated components. Catchments give us a way of detecting these individual grouping patterns, which are a version of one's cognitive style.

To summarize the definition of a catchment:

¥ A catchment is recognized from recurrences of gesture form features over a stretch of discourse. Catchments are recognized from 2 or more gestures (not necessarily consecutive) with partially or fully recurring features of shape, movement, space, orientation, dynamics, etc.

¥ A catchment is a kind of thread of consistent dynamic visuospatial imagery running through the discourse segment that provides a gesture-based window into discourse cohesion.

¥ The logic of the catchment is that discourse themes produce gestures with recurring features; these recurrences give rise to the catchment.

¥ Thus, working backwards, the catchment offers clues to the cohesive linkages in the text with which it co-occurs.

We can scrutinize Vivian s Battle Plan for catchments and in this way recover the effective fields of opposition and hence the psychological predicates/GPs and their semiotic values in her discourse.

Table 1 Catchment Structure of Vivian s Battle Plan.

Ln	Catchment	Utterance	Gesture Feature
1	C1	he tries going [up the inside of the drainpipe and]	1-hand (right)
2	C2	Tweety Bird runs and gets a bowling ball and drops <u>it down</u> the drainpipe	2-similar hands
3	C3	[and as he's coming up]	2-different hands
4	C3	[and the bowling ball's coming d]	2-different hands
5	C3	[own he swallows it]	2-different hands
6	C1, C3	[and he comes out the bottom of the drainpipe]	1-hand (left)
7	C2	[drainpipe and he's got this big bowling ball inside him]	2-similar hands
8	C2	[and he rolls on down into a bowling ball]	2-similar hands
9	C2	[ey and then you hear a strike]	2-similar hands

**Vivian s catchments.** Table 2 lays out 3 catchments, recognizable in this case from hand use and hand shape/position — right hand or left; one hand or two and when two hands, same hand shape and position or different. Each of the gesture features embodies a certain thematic content and this content is what motivates it: C1 is about a single moving entity and its recurring gesture feature is a single moving hand; C2 is about the bowling ball and what it does, and its recurring feature is a rounded shape (in gesture transcription terms, 2 similar hands with shape details added); C3 is about the relative positions of two entities in a drainpipe and its recurring feature involves two hands in the appropriate spatial configuration ( 2 different hands ).

**C1.** The first is the catchment of one-handed gestures in items (1) and (6). These gestures accompany descriptions of Sylvester s motion, first up the pipe, then out of it with the bowling ball inside him. Thus C1 ties together references to Sylvester as a solo force. This one-handed catchment differs from the two-handed gestures, which in turn divide into two other catchments.

**C2.** Two-handed symmetrical gestures in (2), (7), (8) and (9). These gestures group descriptions where the bowling ball is the antagonist, the dominant force. Sylvester becomes what he eats, a kind living bowling ball, and the symmetric gestures accompany the descriptions where the bowling ball asserts its power. In (2) the bowling ball is beginning its career as antagonist. The rest of the catchment is where it has



achieved its result. A two-handed symmetric gesture form highlights the shape of the bowling ball or its motion, an iconicity appropriate for its antagonist role.

**C3.** Two-handed asymmetrical gestures in items (3), (4) and (5). This catchment groups items in which the bowling ball and Sylvester mutually approach each other in the pipe. Here, in contrast to the symmetric set, Sylvester and the bowling ball are equals differing only in their direction of motion.

With these catchments, we can analyze the realtime origins of the utterance and gesture in (2) in a way that incorporates context as a fundamental component. The occurrence of (2) in the symmetrical catchment shows one of the factors that comprised its field of oppositions at this point was the various guises in which the bowling ball appeared in the role of an antagonist. This catchment set the bowling ball apart from its role in C3 where the bowling ball was on a par with Sylvester. The significant contrast in C2 was the downward motion of the bowling ball toward Sylvester. Because of the field of oppositions at this point, this downward motion had significance as an antagonistic force. We can write this meaning as:

Antagonistic Force: Bowling Ball Downward.

This was the context and contrast. Thus, it down, unlikely though it may seem as a unit from a grammatical point of view, was the cognitive core of the utterance in (2)—the it indexing the bowling ball and the down indexing the significant contrast itself in the field of oppositions.

The verb drops, therefore, was *excluded* from this GP. Exclusion is evidenced in the fact that the stroke did not synchronize with the verb; in fact, it was withheld from the verb by continued preparation concluding in a brief prestroke hold. We can explain this as follows. The verb describes what Tweety did, not what the bowling ball did (it went down), and thus was not a significant contrast in the field of oppositions involving the bowling ball. The core idea at (2) was the bowling ball and its action, not Tweety and his (an ergative relation).<sup>13</sup>

### **One Utterance, Several Contexts.**

That drops was excluded from the GP yet was included in the utterance points to a second context at play in the origins of (2). The utterance, though a single grammatical construction, grew out of two distinct contexts and gained oppositional meaning from each.

The first context we have already analyzed; it was the C2 theme in which the bowling ball was an antagonistic force. The second context can be seen in that the gesture at (2) also contrasted with C1 the preceding one-handed gesture depicting Sylvester as a solo force. This significant contrast led to the other parts of the utterance in (2) via a partial repetition of the utterance structure of (1), a poetic framework within which the new contrasts were formed (cf. Jakobson 1960). Contrasting verbal elements appeared in

---

<sup>13</sup> Line (6) is simultaneously in C1 and C3, a fusion that neatly captures the fact that Sylvester, at this point, has turned into a kind of living bowling ball — the gesture is a single hand, which in C1=Sylvester in motion, but it is the *deft* hand, which in C3 was the bowling ball; here the gesture belongs to both catchments or fields of opposition, Sylvester as a solo force and the bowling ball heading down. The overlap also neatly captures the irony of the turn of events.

close to equivalent slots (the match is as close as possible given that the verb in (2) is transitive while that in (1) is intransitive):

- (1) | (Sylvester) | up | in he tries going up the inside of the drainpipe  
 (2) | (Tweety) | down | in and (Tweety) drops it down the drainpipe

The thematic opposition in this paradigm is counter forces — Tweety-down vs. Sylvester-up. Our feeling that the paradigm is slightly ajar is due to the shift from spontaneous to caused motion with *drops*. This verb does not alter the counter forces paradigm but transfers the counter force from Tweety to the bowling ball, as appropriate for the gesture with its downward bowling ball imagery.

The parallel antagonistic forces in (1) and (2) made Tweety the subject of (2), matching Sylvester as subject of (1). The contrast of (2) with (1) thus had two effects on our target utterance. It was the source of the verb, *drops*, and was also why the subject was Tweety, rather than bowling ball. The verb expressed Tweety's role in the contrast and shifted the downward force theme to the field of oppositions about the bowling ball. The identity of subjects and the similar syntactic frame expressed the antagonistic forces paradigm itself. The prestroke hold over *drops* is thus also explained: the verb, deriving from an antagonistic forces context, was propaedeutic to the GP, and the stroke was withheld until the way had been prepared for it.<sup>14</sup>

Let's summarize how (2) came into being:

1. The field of oppositions in which the significance of the downward motion of the bowling ball was that of an antagonistic force the contrast of (2) with (3), (4), (5): this gave the growth point a core meaning centered on *it down*. It's noteworthy that the preparation for the gesture in (2) began in the preceding clause, concurrent with mentioning the bowling ball for the first time ( *Tweety Bird runs and gets a bowling ba[ll and drops it down the drai]npipe* ). That is, the new growth point embodying the idea of the bowling ball in its role as the antagonist to Sylvester began to take form as soon as the bowling ball itself entered into the discourse.

2. The field of oppositions in which the significance was the counter forces of Sylvester-up vs. Tweety-down. This gave a sentence schema that included the words *drops*, *down*, *drainpipe*, and the repetition of the sentence structure with Tweety in the subject slot.

The choice of verb in (2) was *drops*, rather than *throws*, *thrusts* or some other caused-motion option for a downward trajectory from a manual launch, possibly because it, among these options, corresponds most closely to the force-dynamics of how Tweety made use of gravity to launch the bowling ball.<sup>15</sup> Thus, a further aspect of the context of (2) is this force-dynamics. If this is the case, we have a further argument for the analysis above in which *drops* and *it down* are contrasts in different contexts. This comes from the hand shape of the gesture in (2). The speaker made the gesture with her hands facing *down*, in a thrusting position. See Fig. 3. They were not positioned for

<sup>14</sup> The poststroke hold on the second half of *down* derives differently, as Kita (1990) observed. At one level, pre- and poststroke holds are opposites — a prestroke hold signals that the stroke is being withheld, a poststroke hold that it is being extended. In both kinds, however, the hold ensures the synchrony of the stroke with specific speech segment(s).

<sup>15</sup> Pointed out by Karl-Erik McCullough.

a simulation of Tweety's hands when he exploited gravity to launch the bowling ball. Tweety in the cartoon stimulus held the bowling ball from the bottom and chucked it into the pipe, allowing gravity to do the rest. The GP image altered this force-dynamics by making the launch into a thrust. The gesture, that is, reflected the speaker's reconceptualizing of the cartoon. The new force-dynamics is not appropriate to Tweety, but it does fit the field of oppositions that concentrated on the force-dynamics of the bowling ball in its persona as antagonist.<sup>16</sup>

The question of how a GP leads to a grammatically allowable surface form can be answered in part with filters or templates. The idea unit in a GP, being a point of differentiation from a background, must be realized in a surface position consistent with its communicative dynamism, but this fact does not, by itself, specify a position in a sentence framework. The concepts of construction grammar (Goldberg 1995) and of emergent grammar (Hopper 1987) may apply here. The contextual weight generated for the initially non-grammatical pair, *it down plus gesture*, could be completed by

accessing a caused-motion construction,

Subj	V	Obj	Obl
------	---	-----	-----

Tweety drops it (b-ball) down

This construction type, learned from childhood and able to function as a unit, could ease the GP, despite its double contextual base, into a single grammatical format and guide the speaker to a well-formed output. This is *emergent* in the sense that the grammatical form was not part of the input (this was the GP, the non-grammatical *it down plus gesture*), yet the output was this grammatical pattern. Thus, constructions, retrievable as wholes, can provide part of the template that the speaker needs.

## UNPACKING

Unpacking is the process of articulating the implications of the core idea and using these implications as a guide to a well-formed surface structure. The sentence is an action to present the GP. A key process is the generation of new meanings during this unpacking process.

In the case study, the emergence of *drops* is an illustration of the emergence of a further meaning. The details of the gesture already described suggest an answer to the following question: When was the semantic meaning of caused-motion activated? The answer is that it emerged when the speaker first activated the bowling ball image. That is both sets of contrasts, involving C1 and C2, were activated at this moment. Thus, the surface order of words in the speaker's speech did not coincide with the order of emergence of ideas in her thinking — she had two separate ideas (one, the GP idea of the bowling ball as an antagonistic force and two, caused motion, each with its own context), but finding a well-formed stopping point in the caused-motion construction led these

---

<sup>16</sup> Sotaro Kita has pointed out that the first-person viewpoint inherent in the gesture at (2), in which the speaker's hands depict Tweety's hands, is deducible from the agentivity tier (in Jackendoff's 1990 terminology) in the utterance. While the Sylvester-Tweety opposition is not a new contrast at this point and so cannot be a psychological predicate, the agentivity tier does carry new information in that Tweety now has done something. And this fact is embodied in a first-person character viewpoint gesture.

ideas to emerge in a certain order in speech that did not coincide with this cognitive sequence.

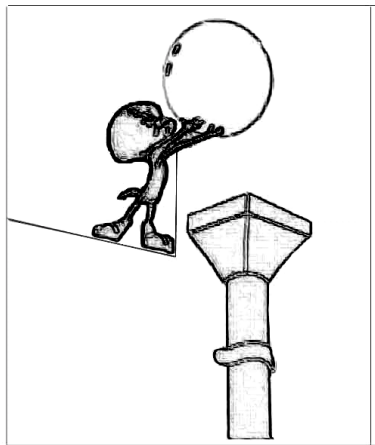
The evidence for the actual moment of emergence of the content of the sentence is that gesture preparation revealed an image already shaped to embody the speaker's idea of an antagonistic force. As she first mentioned the bowling ball her *hands turned down*, preparing to thrust the ball downward. As noted above, this caused-motion meaning implies a different force dynamics from the actual launch, but it was a natural expression of the caused-motion meaning. In addition, the hands were already Tweety's hands, so the agentive role of Tweety was also established. All these details are shown in Fig. 3.



Gesture preparation (start)



Gesture stroke (start)



Character s actual launch

Fig. 3. Comparison of the crucial bowling ball gesture (in 2 phases) to its iconogenic original from the stimulus.

Unpacking drew on multiple sources — the idea of an antagonistic force, which was the GP; the idea of caused-motion, which was the unpacking frame; the referential appropriateness of *drops* for describing the ball's descent (but not the caused-motion construction); and the poetic replication of the initial sentence format (1) in (2), giving a sentence format with Tweety as the subject (cf. Furuyama 2001). The impetus for unpacking (i.e., finding a stable construction) is the inherently unstable and incomplete combination of all this information, which creates a drive to find an intuitively complete structure, or format of repose.

The gesture phase analysis shows that the Caused Motion meaning, with Tweety as the agent and the bowling ball as the patient, began in the previous clause (during *ball*). That is, the core idea was the bowling ball and its role as an antagonistic force, but to unpack this meaning a further meaning was necessary - this was caused motion. Caused motion tied together the two catchments, C1 and C2. *Drops* was integrated in this caused motion construction and performed the crucial job of shifting the antagonistic force from Tweety to the bowling ball.

Reconciling the various components led to the intuitively complete caused-motion sentence, *drops it down the X* — an intuitively complete structure is the construction of a state of repose *par excellence*. Thus a grammatically complete surface structure was the stop-order for the unpacking process. The process presses, if need be, on through interruptions and re-starts until a well-formed surface structure or locally acceptable approximation to one is reached; then it ceases. Unpacking is the completed materialization of the speaker's meaning in the context of speaking with new meanings generated to achieve a well-formed pattern — it is thought in systematic form (in the Saussurian sense) and comprises the final evolved form of the speaker's current thought.

## MORALS OF THE STORY

The GP captures Wundt's insight that in speech there is something instantaneous and something sequential. The combination of opposite semantic modes is an inherent quality of the GP and is necessary for a dialectic of imagery and language.

In saying *drops it down* Vivian was taking up a position in the world. Her living meaning in the form of an image of a downward moving bowling ball inhabited the words and, from our analysis of the gesture and her speech (and beyond, of the discourse exhibited in the relevant catchments), we can see something of what causes static language to spring into motion. The case study example is of course only an illustration but the processes involved are general and appear across situations, speakers, languages, and ages (down to age 3 at least; McNeill, McCullough & Duncan in press).

An important further moral of the story is that an utterance, even though grammatically self-contained, contains content from outside of its own borders. This other content ties the utterance to the context at the level of thinking. The incorporation can occur from multiple places — in the case study, one context (C2) was the backdrop to the GP and the core meaning of a downward antagonistic force; the other context (C1) supplied the semantic framework out of which the caused-motion construction was evoked. In the latter case, caused-motion was sufficient for the context. In the GP case, the context was part of the core meaning, the indispensable background of the psychological predicate. This in fact is an important distinction in the course of a dialectic — a core meaning that gels into the GP and commences a dialectic vs. that which

is sufficient merely and guides the unpacking and resolves the dialectic. All of these factors are concurrent and play out in realtime, over temporal intervals typical of thinking in context.

The GP model moreover *predicts* a context in the sense that the context is necessary for the GP to be a psychological predicate and for the construction choice to be the endpoint and resolution of the dialectic. The GP model treats context as an integral part of thinking for speaking rather than as an external parameter (as, for example, in Levelt 1989, de Ruiter 2000). By making a prediction of the context, a GP analysis meets the Popperian standard of falsification: the analysis of a GP would be rejected should the required context not be observed. This falsifies the claim made by de Ruiter and Wilkins (1998) that the GP is not falsifiable.

That two contexts could collaborate to form a single grammatical structure also implies that a sense of grammatical form enters into utterances in piecemeal and oblique ways that do not necessarily follow the rule-governed patterns of the utterance's formal linguistic description. Analysis of grammatical structure is not a recipe for performance. Grammatical form is the point where the process stops, and this could be variable, dependent upon a variety of factors, including the threshold for completeness set at the moment of speaking.

## REFERENCES

- Armstrong, D., Stokoe, W. & Wilcox, S. 1995. *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.
- Blumenthal, A. J. 1970. *Language and Psychology*. New York: Wiley.
- Chase, W.G. and Ericsson, K.A. 1981. Skilled memory. In J. R. Anderson (ed.), *Cognitive Skills and Their Acquisition*, pp. 227-249. Hillsdale, NJ.: Erlbaum.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Culler, J. *Ferdinand de Saussure*. 1976. Harmondsworth, England: Penguin.
- De Ruiter, J. P. 2000. The production of gesture and speech. In D. McNeill (ed.), *Gesture and Language*, pp. 284-311. Cambridge: Cambridge University Press.
- De Ruiter, J. P. and Wilkins, D. 1998. Arrernte-Dutch comparison of speech-gesture synchronization and gesture space. In S.Kita and L.W. Dickey (eds.), *Max-Planck Institute for Psycholinguistics Annual Report*, pp. 53-54. Nijmegen, Holland: MPI.
- Dray, N. L. and McNeill, D. 1990. Gestures during discourse: The contextual structuring of thought. In S. L. Tsohatzidis (ed.), *Meanings and Prototypes: Studies in Linguistic Categorization*, pp. 465-487. London: Routledge.
- Dreyfus, H. 1994. *Being-in-the-World: A Commentary on Heidegger's Being and Time, Division I*. Cambridge, MA: MIT Press.
- Duranti, A., and Goodwin, C. 1992. *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press.
- Fillmore, C. J. 1982. Frame semantics. In Linguistic Society of Korea (eds.), *Linguistics in the Morning Calm*, pp.111-138. Seoul: Hanshin.

- Furuyama, N. 2001. De-syntacticizing the Theories of Reference Maintenance from the Viewpoint of Poetic Function of Language and Gesture: A Case Of Japanese Discourse. PhD Dissertation, University of Chicago.
- Gallagher, S., Cole, J. & McNeill, D. 2002. Social Cognition and Primacy of Movement Revisited. *Trends in Cognitive Sciences*. 6:155-156.
- Gelb, A. and Goldstein, K. 1925. ber Farbennamenamnesie, *Psychologische Forschung* 6:127-186.
- Goldberg, A. E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, A. 1997. Making One's Way Through the Data. In A. Alsina, J. Bresnan & P. Sells (eds.), *Complex Predicates*, pp. 151-173. Stanford, CA: CSLI Publications.
- Hopper, P. 1987. Emergent grammar. In J. Aske, N. Beery, L. Michaelis & H. Filip (eds.), *Proceedings of the 15<sup>th</sup> Annual Meeting of the Berkeley Linguistics Society*, pp. 139-157. Berkeley, CA: Berkeley Linguistics Society.
- Iverson, J. & Goldin-Meadow, S. 1998. Why people gesture as they speak. *Nature* 396:228.
- Jackendoff, R. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.
- Jakobson, R. 1960. Concluding statement: Linguistics and poetics. In T. Sebeok (ed.), *Style in Language*, pp. 350-377. Cambridge, MA: MIT Press.
- Johnson, M. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Kendon, A. 1972. Some relationships between body motion and speech: an analysis of an example. In A. Siegman & B. Pope (eds.), *Studies in Dyadic Communication*, pp. 177-210. Elmsford, NY: Pergamon.
- Kita, S. 1993. Language and thought interface: A study of spontaneous gestures and Japanese mimetics. PhD Dissertation, Department of Psychology, The University of Chicago.
- Kita, S. 1990. The temporal relationship between gesture and speech: a study of Japanese-English bilinguals. Masters Thesis, Department of Psychology, University of Chicago.
- Kita, S. 2000. How representational gestures help speaking. In D. McNeill (ed.), *Gesture and Language*, pp. 162-185. Cambridge: Cambridge University Press.
- Lakoff, G. & Johnson, M. 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.
- Levelt, W. 1989. *Speaking*. Cambridge, MA: MIT Press.
- Liddell, S. 2000. Blended spaces and deixis in sign language discourse. In McNeill, D. (ed.), *Gesture and Language*, pp. 331-357. Cambridge: Cambridge University Press.
- Mayberry, R. & Jaques, J. 2000. Gesture production during stuttered speech: Insights into the nature of gesture-speech integration. In D. McNeill (ed.), *Gesture and Language*, pp. 199-214. Cambridge: Cambridge University Press.



- McNeill, D. 1992. *Hand and Mind: What Gestures Reveal About Thought*. Chicago, IL: University of Chicago Press.
- McNeill, D, McCullough, K.-E., & Duncan, S. in press. An ontogenetic universal and how to explain it. In C. Miller and R. Posner (eds.), *The Semantics and Pragmatics of Everyday Gestures*. Berlin: Verlag Arno Spitz.
- McNeill, D., Quek, F., McCullough, K.-E., Duncan, S., Furuyama, N., Bryll, R., Ma, X.-F., & Ansari, R. 2001. Catchments, prosody and discourse. *GESTURE*, 1:9-33.
- McNeill, D. and Duncan, S. D. 2000. Growth points in thinking for speaking. In D. McNeill (ed.), *Gesture and Language*, pp. 141-161. Cambridge: Cambridge University Press.
- Merleau-Ponty, M. 1962. *Phenomenology of Perception* (C. Smith, trans.). London: Routledge.
- Nobe, S. 1996. Representational gestures, cognitive rhythms, and acoustic aspects of speech: a network/threshold model of gesture production. PhD Dissertation, Department of Psychology, The University of Chicago.
- Palmer, S. & Kimchi, R. 1986. The information processing approach to cognition. In T.J. Knapp & L.C. Robertson (eds.), *Approaches to Cognition: Contrasts and Controversies*, pp. 37-77. Hillsdale, NJ: Erlbaum.
- Saussure, F. de. 1959 [1915]. *Course in General Linguistics*. New York: Philosophical Library.
- Talmy, L. 1975. Syntax and semantics of motion. In J.P. Kimball (ed.), *Syntax and Semantics, Vol. 4*, pp. 181-238. New York: Academic Press.
- Talmy, L. 1985. Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (ed.), *Language Typology and Syntactic Description. Vol. III: Grammatical Categories and the Lexicon*, pp. 57-149. Cambridge: Cambridge University Press.
- Vygotsky, L. S. 1978. *Mind in Society* (M. Cole, V. John-Steiner, S. Scribner & E. Souberman, eds.). Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. 1986. *Thought and Language* (revised and edited by A. Kozulin). Cambridge, MA: MIT Press.
- Werner, H. & Kaplan, B. 1963. *Symbol Formation*. New York: Wiley [reprinted in 1984 by Erlbaum].
- Wundt, W. 1900. *Die Sprache*. Leipzig: Engelmann.

David McNeill  
 Department of Psychology  
 University of Chicago  
 5848 S. University Ave.  
 Chicago, IL 60637 USA  
 dmneill@uchicago.edu