

Gesture and Thought presents a new conception of language: language as an imagery-language dialectic, in which gestures provide imagery. Gesture is an integral component of language in this conception, not merely an accompaniment or ornament. Such gestures are synchronous and co-expressive with speech, not redundant, and are not signs, salutes, or emblems. They are frequent—about 90% of spoken utterances in descriptive discourse are accompanied by them. The synchrony of speech forms and gestures creates the conditions for an imagery-language dialectic. A dialectic implies:

- a) a conflict or opposition of some kind, and
- b) resolution of the conflict through further change or development.

The synchronous presence of unlike modes of cognition, imagery and language, which are co-expressive of the same underlying thought unit, sets up an unstable confrontation of opposites. It is this very instability that fuels thinking for speaking as it seeks resolution. Instability is an essential feature of the dialectic, and is a key to the dynamic dimension. The concept of an imagery-language dialectic extends a concept initiated by Vygotsky, in the 1930s.

This new conception also recaptures an insight lost for almost a century, that language requires two simultaneous modes of thought—what Saussure, in recently discovered notes composed around 1910, termed the ‘double essence’ of language (although he expressed this without reference to gestures).

Gesture and Thought focuses on the real-time actualization of thought and language, regarding language multimodally and in context—what in the book is called its dynamic dimension. On the dynamic dimension, language appears to be a process, not an object. On the crosscutting static dimension, it looks to be an object but not a process. In this book, both dimensions are considered, as both are indispensable to a full theoretical explication of utterances. An important question, answered in the course of the book, is how they combine in realtime utterances.

The smallest unit of the imagery-language dialectic is posited to be a ‘growth point,’ so named because it is theoretically the initial unit of thinking for speaking out of which a dynamic process of organization emerges. A growth point combines imagery and linguistic categorial content, and the theory is that such a combination initiates cognitive events. A growth point is an empirically recoverable idea unit, inferred from speech-gesture synchrony and co-expressiveness. An example recorded in an experiment (offered in part because of its ordinariness) is a description by one speaker of a classic Tweety and Sylvester escapade, which went in part as follows: “and Tweety Bird runs and gets a bowling ba[ll and drops it down the drainpipe].” Speech was accompanied by a gesture in which the two hands thrust downward at chest level, the palms curved and angled inward and downward, as if curved over the top of a large spherical object. At the left bracket, the hands started to move up from the speaker’s lap to prepare for the downward thrust. Then the hands, at the very end of “drops,” held briefly in the curved palm-down position, frozen in midair (the first underlining). Next was the gesture stroke—the downward thrust itself—timed exactly with “it down” (boldface). Movement proper ceased in the middle of “down,” the hands again freezing in midair until the word was finished (the second underlining). Finally, the hands returned to rest (end of second underlining up to the right bracket). The two holds reveal that the downward thrust was targeted precisely at the “it

down” fragment: the downward thrust was withheld until the speech fragment could begin and was maintained, despite a lack of movement, until the fragment was completed. Significantly, even though the gesture depicted downward thrusting, the stroke bypassed the very verb that describes this motion, “drops,” the preparation continuing right through it and holding at the end—an explanation for this seeming overshoot is provided later.

The growth point was thus the fragment, “it down,” plus the image of a downward thrust. Both sides of the growth point are essential, and are opposed dialectically in that the linguistic components have combinatoric potential and categorize the image; the imagery component embodies these categories in an instantaneous whole; the different modes are simultaneously active (for the speaker and the listener, who is trying to recreate the growth point). That one idea exists in two such different modes is the motive force for the utterance and its linked meaning formation.

The growth point is resolved by unpacking it into a more stable form, with a grammatical construction being the most stable possible outcome. Intuitions of good form (called ‘intuitions-1,’ in which language appears to the individual as an unchanging object) arise and are the *stop orders* for the dialectic. Once the speaker sensed a well-formed construction, she resolved the conflict by distributing the imagery and categorial content of the growth point into its prepared slots, and this stopped the dialectic process (how this might work is illustrated below). In this way, the dynamic intersects the static, as expected by Saussure’s double essence insight—intersects it in fact in several ways: in the growth point, in the unpacking, and in the stop order. It is not that unpacking invariably reaches a full grammatical construction. It proceeds until some threshold of stability is reached, which may often be less than a complete outcome; or it may just break down and stop if stability proves unattainable in the time spans attainable at socially realistic speech rates, e.g., because of an inappropriate construction attempt. Thus pauses and grammatical approximations, rife in daily discourse, can be explained as products of the dialectic resolution and the speaker’s efforts toward it within limited time spans.

In this example, nonetheless, the growth point smoothly unpacked into a construction, the causative, “someone drops (=causes to move by dropping) something down some landmark object.” Intuitions-1 of the caused-motion construction arose and became the stop order, the construction plausibly resolving the dialectic by providing slots for the growth point image and its categorial content.

Subj	V	Obj	Obl
↓	↓	↓	↓
Ø (Tweety)	drops	it (b-ball)	down

(boldface for the slots that distributed the pieces of the growth point; the Tweety subject and the verb “drops” are explained below).

Context is a second source of dynamism. Theoretically, a growth point is a psychological predicate in Vygotsky’s sense, a significant contrast within a specific context. While context reflects the physical, social and linguistic environment, it is also a mental phenomenon; the speaker constructs this context, in order to make the intended contrast meaningful. The growth point is thus not fixed and implies the context from which it is differentiated. Finding this context in actual data is an essential part of validating the growth point empirically. The mental construction of the context is modeled

as a field of oppositions; what the speaker creates is a field of oppositions to make the psychological predicate differentiable within it. This is a model in which meaning is a relationship between a point of contrast and the background or field of oppositions from which it is being differentiated

A further concept, the catchment, provides an empirical route for finding this field of oppositions. A catchment comprises multiple gestures with recurring form features, and reveals the discourse segment to which the growth point belongs. More than one catchment can be simultaneously active for the same growth point. The full complement of catchments can suggest the particular set of oppositions from which the growth point is being differentiated.

To identify the catchment in the “it down” case, we look for other gestures in which the hands are shaped and/or move similarly to the target gesture, and see if all these gestures comprise a family with thematic continuity. We find such a family; in the speaker’s rendition, similar two-handed gestures had to do with the bowling ball in the role of an antagonistic force, contra-Sylvester. We can thus further specify the “it down” growth point: it was a psychological predicate specifying the bowling ball *as* this antagonistic force. Antagonistic forces against Sylvester was the field of oppositions; the differentiated opposition was this force as the bowling ball moving downward. The growth point and this context provide a richer picture of the speaker’s idea unit than a purely referential reading of the phrase, “drops it down the drainpipe,” suggests:

Antagonistic Force: Bowling Ball Downward

Also, we can now explain the timing of the gesture: the downward thrust coincided exactly with the linguistic categorial content with which it formed a growth point, or psychological predicate. It skipped the verb “drops,” despite the fact that this verb described the bowling ball’s motion down, precisely because the verb does not describe the bowling ball in its role as an antagonistic force; it describes what Tweety did, not the bowling ball, and thus could not have categorized the image with the intended meaning. The speaker’s core idea was not dropping but the idea of the bowling ball moving down as an antagonistic force. Hence, the details of how gesture and speech combined, including timing, can be explained as aspects of the speaker’s construction of the psychological predicate in the context, which is to say her thought process in context. (Other psychological predicates in the same catchment also conveyed the antagonistic force theme, specifying its effects on the unfortunate Sylvester: how he became a kind of living bowling ball, rolled down a street, into a bowling alley, and knocked over all the pins. Each of these can be analyzed in turn as psychological predicates differentiating further contrasts within the Antagonistic field of oppositions.)

The growth point was unpacked into a caused-motion construction, as noted, and we can analyze this and explain where the remaining pieces of the utterance, Ø (Tweety) and “drops”, came from as well. Unpacking is more than just finding a construction in which to house a growth point; it includes the differentiation of further meanings with their own contexts and integrating them with the growth point so that the construction, including its semantic frame, can resolve it. The unpacking took place in a second catchment, also active during the speaker’s representation of the bowling ball episode. The immediately preceding utterance was, “he tries going [**up**] the insid][e of **the drainpipe**],” which segued

directly into our target utterance. The three gestures (in bold) were made the same way, with one hand rising upward, the first finger extended. Although this may include pointing, the gestures occurred with the theme of Sylvester acting as a force of his own (for many speakers an extended first-finger gesture conveys compression: Sylvester, inside the pipe, squeezes his plump body down to about half size). So, for this speaker, the utterance and the target utterance comprised a paradigm of opposed forces. Opposed forces was her way of construing the episode: not merely the bowling ball and Sylvester colliding, but Sylvester, a force moving up, versus the bowling ball, a force moving down—each force with its own gesture imagery. The bowling ball moreover was not the original antagonistic force; the sentence was “(Tweety) drops it down,” which starts out with Tweety in the subject slot as the force. The speaker understood from the cartoon that she had to make the bowling ball into this force. The verb “drops” plus the caused-motion construction neatly achieved the shift from Tweety to the bowling ball. This is the growth point account how the verb and the Tweety subject made their way into the utterance.

The whole target utterance was thus the product of two contexts: 1) the growth point in the context of the bowling ball as an antagonistic force: this was the core idea unit; and 2) caused-motion with “drops” and Tweety as subject: the further meanings in the paradigm of opposed forces that resolved the imagery-language dialectic, and shifted the antagonistic force to the bowling ball. The target utterance, although a single grammatical construction, grew out of two distinct contexts and gained oppositional meaning from each.

The linguistic side of a growth point is not necessarily grammatical. The “it down” growth point is not grammatical but nonetheless formed a growth point with the downward image in the context of thwarting Sylvester.

Metaphoricity also is present. The abstract idea of an antagonistic force is presented as something else, a downward moving bowling ball. The metaphor embodied the abstract meaning of antagonistic force in an image, the importance of which is that an unimageable meaning could combine in an imagery-language dialectic with “it down”. In other words, metaphoricity was an essential part of forming this growth point (in this case and numerous others). The bowling ball metaphor was an impromptu creation but other gesture metaphors are culturally established and play the same role of enabling imagery-language dialectics with abstract meanings that are unimageable. An example of a cultural metaphoric gesture is the ‘palm up open hand’, in which the hand(s) appear to present a discursive object. The metaphor is recognizable as the so-called ‘conduit’ metaphor, an image in the general European metaphor culture (and in other cultures, but not universally), in which an abstract idea is presented as if it were a substance in the hand or a container (cf. verbal examples like “the movie had a lot of meaning,” where the movie is a container, or “she handed him that idea,” where an idea is on or in the hand).

Finally, as Vygotsky famously argued, human thought is fundamentally social in character. While the case study highlights a single speaker recounting a story, the growth point concept is not restricted to the classical individual mind-in-isolation. Among the shaping factors in a field of oppositions is the speaker’s social interactive context. In her dissertation Asli Özyürek showed that changing the number and the spatial loci of listeners has an effect on the speaker’s gestural imagery. Plugging this result into the growth point model, we infer that an imagery-language dialectic can be altered by changes of the social context. One also can find two-party growth points, gestures from one person

synchronizing with a second person's speech, and vice versa—someone's speech accompanied by another person's gestures. Conversations are dynamically affected by the participant's gestures, even decisively altering its direction when a conflict arises over the meanings metaphorized as the shared gesture space. Such conflicts produce diverging imagery-language dialectics, which speakers attempt to realign.

A final, final point places this entire discussion on a different plane, and provides an answer to the question: what becomes of an imagery-language dialectic when gestures do *not* appear? We get a deeper understanding of the imagery-language dialectic by introducing the concept of a 'material carrier'. A material carrier is the embodiment of meaning in a concrete enactment or material experience. A material carrier appears to enhance the symbolization's representational power. The concept implies that the gesture, *the actual motion of the gesture itself*, is a dimension of meaning. Such is possible if the gesture *is* the very image; not an 'expression' or 'representation' of it, but *is* it. The gesture *itself* is a component of the dialectic. From this viewpoint, a gesture is an image in its most developed—that is, most materially, naturally embodied—form. The absence of a gesture is the converse, an image in its least material form. The greater the felt departure of being from the immediate context, the more likely 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 being. Absence of gesture is then the predictable result of a minimal departure from context; in repetitive or denatured contexts, imagery fades and, Cheshire Cat like, only imageless form remains. Merleau-Ponty expressed a similar view of language in *The Phenomenology of Perception*: "The link between the word and its living meaning is not an external accompaniment to intellectual processes, the meaning inhabits the word ... 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). The "it down" growth point was this speaker's taking up of a position in the world of her cartoon narration, her momentary state of being materialized in the image of the bowling ball as an antagonistic force.

I apply this theoretical framework to a very wide range of situations—discourse and gestures in different languages (Turkish, Spanish, Mandarin, as well as my primary source, English); the gestures of children at the earliest stages of development; the Whorfian hypothesis, arguing that the impact of language on imagery is often a dynamic dimension effect that has been concealed by concentration on the static dimension; linguistic impairments (aphasia; right-hemisphere damage, which impairs discourse cohesion; and the split-brain state, all of which were described in *Hand and Mind* but are now integrated into a new neurogestural model). An important new source of observations is the case of IW. This is a man who suffered, as a young adult, complete deafferentation from the neck down and has relearned movement control, including gestures with speech, which he can do to perfection even without vision, a condition where it is difficult for him to carry out nongesture actions. His case suggests a partial dissociation in the brain of the organization of gesture from the organization of instrumental action, and the existence of a dedicated thought-language-hand link that would be the common heritage of all humankind.

The book accordingly ends with an attempt to provide ‘the ultimate answer’ to the question of an imagery-language dialectic and why it exists at all, by proposing a theory of language evolution, focusing on this thought-language-hand link. I develop the hypothesis that what made us human crucially depended at one point on gestures. Without gestures, according to this hypothesis, language could not have evolved; some of the brain circuits required for language could not have evolved in the way they apparently have. The integration of gesture with language we observe in ourselves today is an essential part of the machinery that evolved. Gesture is not a behavioral fossil but an indispensable part of our current day ongoing system of language. The theory is based on the IW case and the neurogestural model, and employs recently discovered ‘mirror neurons’, supplemented with what I am calling ‘Mead’s loop’, to explain how and under what conditions a thought-language-hand link could have evolved. According to the Mead’s loop theory, what had to be selected is a capacity, not present in other primate brains, for the mirror neuron circuit to respond to one’s own gestures as if they belonged to someone else (this produces the apparent social framing of gestures, even when they are invisible—talking on the phone, a blind person talking to another blind person). This reconfiguration of circuitry provided the thought-language-hand link and a way for language to co-opt the machinery of Broca’s area. Contrary to the gesture-first theory, a model that has become popular with Corballis (2002) and others, I am arguing that evolution selected the ability to *combine* speech and gesture under a meaning other than the meaning of the action itself, and that speech and gesture therefore emerged in evolution together. This combination was the essential property that evolution chose; there would not have been a gesture-first step. If gestures had emerged alone, they would have been a kind of pantomime, and it is possible that pantomime, in ourselves, had such an origin. Different kinds of gestures (all called gestures but not all the same) may have had different evolutionary trajectories, and not all played a part in the origin of language. Pantomime does not combine with speech, and must be separated from the speech-synchronized gestures that fulfill an imagery-language dialectic. Just as speech could not have evolved without these gestures, these gestures could not have evolved without speech.

Gesture and Thought is to be published by the University of Chicago Press in the fall of 2005.