# Growth Points and Modeling

David McNeill
University of Chicago

# Gestures shed light on thinking-for-(and while)-speaking

- Among the many manifestations of the embodiment of human language and thought, gestures are outstanding - natural and universal.
- Gestures and more broadly imagery are components of speaking, not accompaniments but actually integral parts of it. Much evidence supports this idea, but its full implications have not always been recognized.
- The growth point (GP) hypothesis is designed to explicate the integral linkage of gesture and speech in natural language production.
- My goal is to explain the GP, its motivation and empirical basis, and, at the end, ask if the GP can be modeled computationally and, if not, what is the challenge?

# The GP is so named because it is a distillation of a growth process

- An ontogenetic-like process but vastly sped up and made functional in online thinking-for-speaking.
- The GP is the initial unit of thinking-for/while-speaking (from Slobin 1987, elaborated to include thinking online, during speech).
- Out of it a dynamic process of utterance-level and discourse-level organization emerges.
- Imagery and spoken form are mutually influencing in a GP. It is not that imagery is the input to spoken form or spoken form is the input to imagery. The GP is fundamentally both.

# Some key examples

- The following illustrate an important fact about speech and gesture.
- We can fully understand what motivates any gesture-speech combination only with reference to how a it relates to its context of occurrence. The GP, based on this, is a point of differentiation from the context.
- The combination is what Vygotsky termed a ‘psychological predicate’, the point of contextually newsworthy information. -->

- In a psychological (as opposed to a grammatical) predicate, newsworthy content is differentiated from the context of speech. It cannot exist without this context from which it is differentiated.
- A robust phenomenon is that the gesture form and its timing with speech embody just those features that differentiate the given psychological predicate in its context.
- The next series of slides presents this phenomenon experimentally and observationally.

# Two ascents

- In the cartoon, Sylvester makes two attempts to reach Tweety, climbing a conveniently situated drainpipe. First, he uses it as a kind of ladder, climbing on the outside. He is thwarted and immediately goes up a second time on the inside.
- The distinctive information in Ascent #1 is the pipe itself (its first appearance). In Ascent #2 it is the interiority factor - the pipe now background.

# Timing Reflects Differentiation of Psychological Predicates

TWO 'CLIMBS UP' EXAMPLES - S. Duncan

- Shows that timing is differentiated relative to context
- Climbs 1 "[he climbs up the…]" Climbs 2 "climbs [up in through the]"


# A natural experiment

- Elena Levy and I first used the cartoon as a stimulus for eliciting narrations from subjects. People watched it and told the story from memory. In this situation, they experienced Outside and Inside, in that order, and directly in sequence.
- A natural experiment occurs when some subjects omit the Outside episode while retelling the Inside (no one omitted Inside but remembered Outside).
- So, for these Inside-only narrators, interiority is not contrasted to exteriority. We expect them not to include interiority in their gestures, and they do not,

# Of the ‘original 6’ two omitted the outside-the-pipe episode


- Both show ascent (right speaker with his thumb only). Neither includes interiority.
- It is not that they are unaware of interiority. Both describe the bowling ball and its inside descent, making clear that they have in mind the inside episode, but lacking Outside interiority is not contrastive for them.

# The remaining four described Outside - Inside in the correct order

The first three speakers to be shown highlight interiority in their Inside gestures - either a rising extended index finger, which seems to convey both upward movement and interiority in the form of Sylvester's plump body-compression, or a rising upward cupped hand – an image of 'rising hollowness'. This emphasis differs from the two speakers who did not mention the Outside ascent.

The fourth speaker is the proverbial exception that proves the rule – and she really does, as I will explain when we reach her clip.

# A designed experiment

- If we can manipulate the point of significant contrast in the stimulus, then GPs should form where we tell them to, as it were.
- Such an experiment is being conducted by Sue Duncan and Dan Loehr, and I will show a couple of their examples.

# Outside and inside in isolation

## Duncan & Loehr experiment



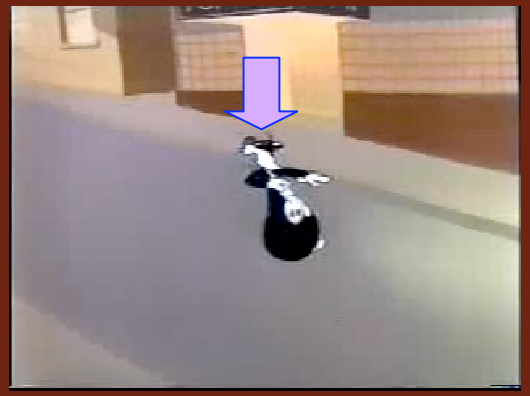
- Unlike the cartoon original, Inside and Outside are shown in isolation. Isolation removes all distinctiveness of interiority:
  - Left saw Outside first, then Inside (the cartoon order). Outside and Inside are consecutive in these clips but in the experiment were separated by 15 unrelated clips.
  - Right saw Inside first, then Outside (reverse order), also separated by 15 unrelated clips. No subject saw any narrative sequence.
- As you will see, neither Inside gesture contained interiority.

# Another condition: In1 --> Out2 removes gesture interiority

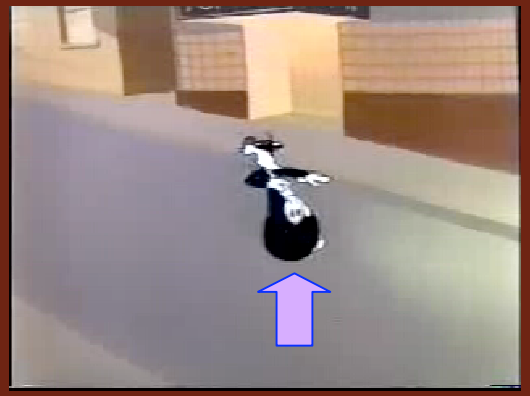
- The subject watched the cartoon in its entirety, as usual but Inside and Outside were reversed. No interiority in either gesture.

# Fey Parrill's thesis manipulated focus

- Cueing discourse focus changes speech and gesture in description of event.
- People see a clip in which Sylvester swallows Tweety's bowling ball, and are prompted in one of two ways:



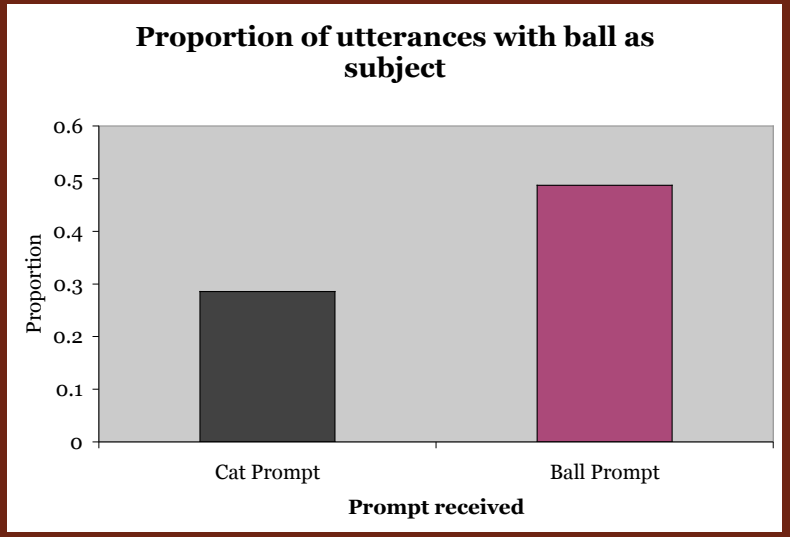
**Cat arrow condition**



**Ball arrow condition**

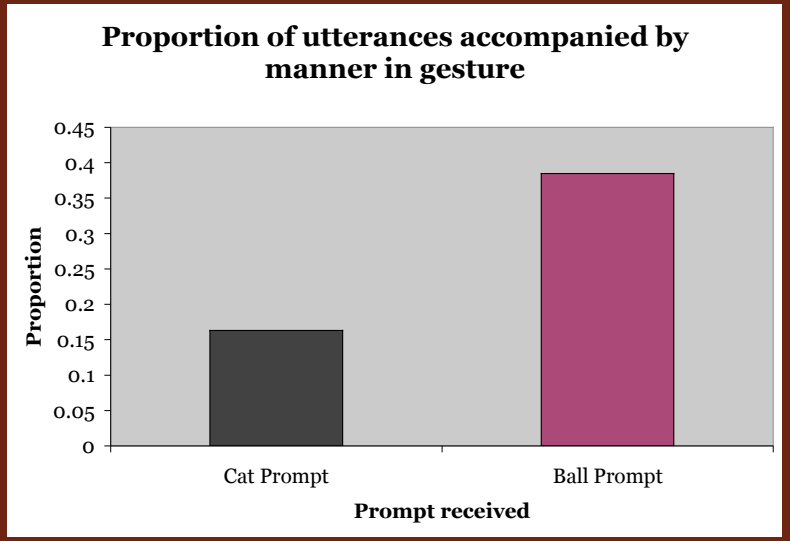
# Experiment 1 results (n=19)

- **Ball prompt** results in more ball-subject utterances - "the ball rolls him down the street" ($t=1.81 > 1.66, p=.03$)




- **Ball prompt** results in more manner in gesture

($t=2.39>1.66$, $p<.01$)



# To summarize

- Gesture-speech combinations select psychological predicates, and are shaped and timed to make differentiation possible.
  - In the Duncan & Loehr experiment and the ‘original 6’, timing and form of gesture reflects newsworthy information in the context of speaking.
  - In the Parrill experiment, focus on the inanimate cause of motion induces manner in gesture - the ball’s rolling being the cause of Sylvester's motion. It also induces speech to use the ball as the subject.
- So, both speech and gesture adapt to the information contour of discourse.

# A ‘minimal unit’ of imagery-language dialectic

- A GP is the smallest packet of an idea unit encompassing unlike semiotic modes - imagery and linguistic encoding.
- A ‘minimal unit’ with the property of being a whole is from Vygotsky:
  - “By a unit we mean a product of analysis which, in distinction from elements, possesses all the basic properties of a whole. Further, these properties must be a living portion of the unified whole which cannot be broken down further…”

Even when the information (the 'semantic content') of speech and gesture is similar, it is formed according to contrasting semiotic modes.

Simultaneous unlike modes create instability. Instability fuels thinking-for-speaking as it seeks resolution. global

The reasons why semiotic opposition creates instability and initiates change include:

1. Conflict (between semiotic modes: analog imagery/analytic categorical), and
2. Resolution (through change: fueling thinking-for-speaking, seeking stability).

The result is an idea unit in which holistic imagery and discrete code are combined, and this drives thinking-for/while-speaking - an inherently dynamic psycholinguistic model.

# Contrasting semiotic modes

- Gesture embodies imagery -
  - It is *global*, meaning that the significance of the elements of the gesture depend on the meaning of the whole. In contrast, the linguistic segment(s), even if a single word or sublexical part, is significant at its own level. Also, the gesture does not gain significance by combining parts, again in contrast to linguistic segments.
  - And the gesture is *synthetic*, meaning that it may embody meanings that are syntagmatically distributed over the surface segments of the utterance.
- In a GP, these two semiotic modes simultaneously embody the same idea, and this is the source of its instability, and its dynamism.

# An example

Gesture embodying ‘rising hollowness’ synchronized with “up thróugh”.

- The temporal and semantic synchronies imply a GP built on the idea of rising interiority.
- We infer the simultaneous presence of the idea of ascent inside the pipe in two unlike semiotic modes – “up thróugh” paired with a uniquely gestural way of packaging meaning – something like ‘rising hollowness’.
- Prosodic emphasis (amplitude, pitch, timing) on “thróugh” is part of the co-expressiveness.
- The shared reference to climbing up inside the pipe makes clear that it is being represented by the speaker simultaneously in two ways—analytic/combinatoric in speech and global/synthetic in gesture.

# The catchment

- As mentioned earlier, a GP is dependent on its context. It is a psychological predicate - a point of differentiation of fresh or newsworthy content in the context, and does not exist without this relation to it.
- The effective contextual background can often be discovered by finding the catchment(s) of which a target gesture is a part.
- Catchments are when space, trajectory, hand shapes, etc. recur in two or more (not necessarily consecutive) gestures.
- The recurring imagery suggests the discourse theme. Given material carrier status, it is more - it *is* this theme, in materialized form.

# Definition

- A catchment is recognized from recurrences of gesture form features over a stretch of discourse
- It's a kind of thread of consistent dynamic visuospatial imagery running through the discourse
- The logic 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 discourse themes in the text with which it co-occurs.

# Viv.'s Battle Plan Catchments

Identified from hand use:

- C1 - 1 handed = Sylvester as a solo force
- C2 - 2 similar handed = the bowling ball as an antagonistic force
- C3 - 2 different handed = the relative spatial positions of the bowling ball and Sylvester inside the pipe

# Context in the Case Study

## Viv's Battle Plan (full episode)

(1) he tries going [[up the insid][e of the drainpipe and]]

(2) Tweety Bird runs and gets a bowling ba[ll and drops it down the drai]npipe

(3) [[and as he's coming up]

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

(5) he ssswallows it]

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

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

(8) [[and he rolls on down] [into a bowling all]][ey

(9) and then you hear a sstri]ke

Look for the catchments

C1 = one hand
C2 = two similar hands
C3 = two different hands

# Catchment Interpretations- 1

- C1 One-handed gestures - items (1) and (6) - ties together references to Sylvester as a solo force.
- C2 Two-handed symmetrical gestures - items (2), (7), (8) and (9) - groups descriptions where the bowling ball is the antagonist, the dominant force. The 2-handed symmetric gesture form highlights the shape of the bowling ball.
- C3 Two-handed asymmetrical gestures - items (3), (4) and (5) - groups items in which the bowling ball and Sylvester are equals differing only in their direction of motion.

- (1) he tries going [[up the insid][e of the drainpipe and]] - (6) [and he comes out the bottom of the drai][npipe
- (2) Tweety Bird runs and gets a bowling ba[ll and drops it down the drai]npipe

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

  (8) [[and he rolls on down] [into a bowling all]][ey

  (9) and then you hear a sstri]ke
- (3) [[and as he's coming up]

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

  (5) he ssswallows it]

# Unpacking - Rest of the Story

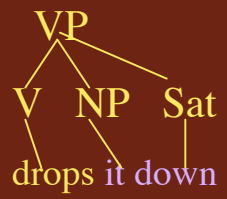
- Unpacking creates structures with which to stabilize the unlike cognitive modes combination in the GP.
  - It is ‘unpacking’ the GP into a grammatical construction (or viable approximation thereto) that preserves its core significance while cradling it in a stable grammatical format.
  - Achieving this often takes additional meaning formulation.
  - The process is regulated by the speaker’s linguistic intuitions—called ‘intuitions-1’ (a sense of well-formedness and contextual appropriateness of the linked semantic frame), in contrast to ‘intuitions-2’, devised for testing grammatical analyses.

# Focus on one GP and how it was unpacked

- “and Tweety runs and gets a bowling bal[l and drops it down the drain] pipe”
- [ ] = gesture phrase - starting with “ball” and ending with “drainpipe”
- ‘it down’ = stroke phase
- Underlining ‘drops’ and ‘down’ = holds (pre- and poststroke)

# The GP

- To understand unpacking, we have to explain the GP --> “it down” + downward image -
  - why it included elements that map into different constituents of syntax?
  - why it did not include the verb? - *actively excluded* by the preparation and a prestroke hold
  - why this image?




- We shall see that an explanation requires reference to the context at the moment of speaking. Unpacking in turn is coordinated with this GP/context relationship.

# Catchment Interpretations- 2

- "and drops it down" was in the symmetrical C2 - this shows that it was part of the various guises in which the bowling ball appeared in the role of an antagonist.
  - The significant contrast was the downward motion of the bowling ball. This downward motion had significance as an antagonistic force. The field of oppositions was Ways of Countering Sylvester.
- We can write this as
  - Bowling Ball Down: Ways of Countering Sylvester
  - 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 of antagonistic force in the field of oppositions.

# Catchment Interpretations -3

- The verb “drops”, therefore, was excluded from the GP. This despite the fact that the gesture showed ‘dropping.’
- We can explain this as follows. The verb describes what Tweety did, not what the bowling ball did (it went down). “Drops” was not part of significant contrast involving the bowling ball - the core idea at (2) was the bowling ball and its action, not Tweety and his.
  - The detailed synchrony of speech and gesture thus incorporated the context at the moment of speaking.

# Two contexts converged

- Viv. construed the episode as a battle of contending forces
  - Tweety foce down
  - Sylvester force up
- The antagonistic forces paradigm was established in the first two sentences:
  - He tries going up …
  - Tweety sees him …
- The viewpoint is that of the upper (Tweety) force.
- The speaker had to find a way to shift the upper force from Tweety to the bowling ball.

# The Full Utterance Belonged to Two Contexts at Once - Why there was "Drops"

Target utterance is product of two contexts:

C2 - with b-ball as antagonistic force, "it down" a unit

C1 - Parallel antagonistic forces, 'Tweety' the subject of (2), matching 'Sylvester' in (1).

| | Force | Direction | |
|---|---|---|---|
| (1) | (Sylvester) | up | in "(Sylvester) goes up" |
| (2) | (Tweety) | down | in "(Tweety) drops it down" |

By choosing "drops" (transitive verb), the speaker shifted the antagonistic force dynamics from Tweety to the b-ball.

"Drops" also captured the force dynamics of using gravity (excluding "throws").

Thus the full structure of (2) - "and (Tweety) drops it down" - came from 2 catchments, C1 and C2, each contributing a part. "Drops" came from the contrast with C1

# Relationship to linguistic form

- As mentioned, this GP on its encoded side was not a grammatical unit. This kind of looseness is an advantage, allowing play in the system without loss of contact with meaning in context.
- Synchronic form is not related to the dialectic in the GP as a formative unit, but as a 'stop-order'. This can be achieved with a construction that does not alter the core meaning, even while providing further meanings of its own.
- In Viv.'s case, the construction provided caused-motion as the means for transferring the conceptual core of Tweety's force to the b-ball, his surrogate against the Sylvester force.

# Constructions

- The GP, being a point of differentiation from a background, must be realized in a surface position consistent with its communicative dynamism (Firbas).
- The contextual weight generated for the initially non-grammatical pair, “it down” plus gesture, was completed by a construction along with its caused-motion meaning that did not undo the GP (Goldberg):

| Subj | V | Obj | Goal |
|---|---|---|---|
| Tweety | drops | it (b-ball) | down the pipe |

- The idea of an antagonistic force in the form of the bowling ball going down was unpacked into encoded concepts of an object caused to move down.

# When Did Caused-Motion Arise? Simultaneously with the GP in this case

Preparation = image shaped by her understanding of antagonistic forces. As Viv. first mentions b-ball, her hands turn down.

- Hands shaping the ball show the b-ball as an antagonistic force
- Hands facing down show T as agent - T thrusts it down

This agent-caused downward push was created by Viv. with a force dynamics differing from the actual launch, where gravity propelled the ball.

Start of Preparation


Start of Stroke


Actual launch

# When things happened - 2

- The gesture phases show that Caused Motion, with Tweety as the agent and the bowling ball as the patient, began at the onset of the idea unit (during the utterance of “ball”). Thus, both the core idea GP, and the extra meanings to unpack it, arose simultaneously.
- “Drops” tied C1 and C2 together even though dropping did not describe Tweety’s action in the animated stimulus.
- “Drops” and caused-motion were non-iconic emergent meanings with which to the unpacking of the “it down” + image GP was brought about.
  - A more iconic description would have been to say “release” - with gravity as the causal agent.

# Some Implications

- The sentence didn't start from the verb - "drops" emerged in the unpacking as a result of new meaning - a caused-motion construction.
  - The verb is not necessarily the core of the dynamic process of evolving meaning and utterance.
- The original GP and the unpacking were both guided by contextual contrasts - each by its own.
- Meaning develops during unpacking - the opposite concept from an 'input' - keep going until a well-formed stopping point is reached.

# Morals

- An utterance, even though seemingly self-contained, contains content from outside of its own structure.
  - This other content ties the utterance to the context at the level of thinking.
  - It is this fact - by no means unique or peculiar - that conflicts with the axioms of modularity yet fits the schema outlined in the GP.
- That multiple contexts collaborate to form one grammatical structure implies that a sense of grammatical form enters into utterances in piecemeal and oblique ways.

# Falsifiability

- Determining a growth point is a kind of hypothesis formulation and test. The GP is empirically recoverable, inferred from speech-gesture synchrony and co-expressiveness of newsworthy content in context.
- The GP is falsifiable in the Karl Popper sense. The chief vulnerable point is that it must be a dialectic of opposite semiotic modes.
  - Exact synchrony is the best evidence but many things can disrupt timing and result in an imperfect sample, but imperfect examples also are meaningful. Slight asynchrony does not matter so long as the linguistic and imagery sides of the GP are co-expressive and differentiate the same idea unit in context.
  - If speech and gesture can't differentiate the same idea unit, or if they differentiate a wrong one, or if they relate to different ideas and contexts, the GP hypothesis is falsified in the Popperian sense.

# ‘Cognitive being’

- The entire conception of speech and gesture is moved to a new level when we draw on Merleau-Ponty (1962) for insight into the duality of gesture and language and what we can expect of gesture in a two-component process. M-P wrote:
  - “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)

- The GP is a mechanism geared to this "existential content" of speech—this "taking up a position in the world". Gesture, as part of the GP, is inhabited by the same "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 we see part of the speaker's current cognitive being, her very mental existence, at the moment it occurs.
- By performing the gesture, a 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 not accidental.

- In Heidegger's emphasis on being, a gesture is not a representation, or is not only such: it is a form of being.
- Gestures (and words, etc., as well) are themselves thinking in one of its many forms—not only expressions of thought, *but thought, i.e., cognitive being, itself.*
- 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.
- 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.

# Theater of the mind

- The H-model avoids anomalies that arise in models where significance is via some form of representation, specifically the 'theater of the mind' problem highlighted by Dennett (1991).
- The theater of the mind is the presumed central thinking area in which representations are presented to a receiving intelligence. The possibilities in this theory of homunculi inside other spiraling downward homunculi are well known.
- In the H-model, there is no theater and no extra being; the gesture is, rather, part of the speakers momentary mode of being itself, and is not watched. In the IW case, 'constructeds' are watched by him, but he avoids the abyss since these gestures are for him external objects of awareness.

# Additional comments to fill out the GP picture

- First, the following question may come to mind: If gesture is ‘part of language,’ how could it and language be ‘semiotically unalike’? There is a certain polysemy in the word ‘language’.
  - When we say gesture is ‘part of language’, we mean language in the sense of Saussure's *langage*, the total semiotic system.
  - When we say that ‘gesture contrasts to language’ we mean it in the sense of his *langue*, the synchronic system of differences.
  - We are analyzing *parole*/performance but in a way broader than this concept is usually understood (Saussure himself, in his recently discovered notes, seems to have had the aim of combining *parole* and *langue*, but had no concept like the GP).

Replies to queries from
Liesbet Quaeghebeur

# More additional comments

- Gestures (and imagery more broadly) lack ‘duality of patterning’.
  - The form of the gesture-signifier is a non-arbitrary product of the signified content (including, via metaphor, abstract ‘non-imagistic’ meanings) – its form doesn't need or get its own level of structure.
  - Speech again contrasts: it has duality of patterning—meaning and sound structured at their own levels, and paired arbitrarily.
- This has to do with the role of convention and where it intrudes. There are conventions of good-form for speech, but none for gesture (apart from the well-known emblem vocabularies in every culture and general kinesic conventions for space and action).

# Still more comments

- In the verbal modality, as in the manual modality, the meaning of the first part ("up" or the spread fingers) remains, as Liesbet Quaeghebeur wrote, 'alive', 'present', or 'active' while the second part is being produced ("thróugh" or the upward movement). There is this kind of continuation in both cases, but the explanation differs—
  - a construction in the verbal case; a global image in the gesture case.
- The continuities differ as well—
  - sequential in the linguistic form, simultaneous in the gesture.
- Between the two means of attaining continuation the difference comes down to whether symbolic actions are organized by syntagmatic patterns or by instantaneous imagery.

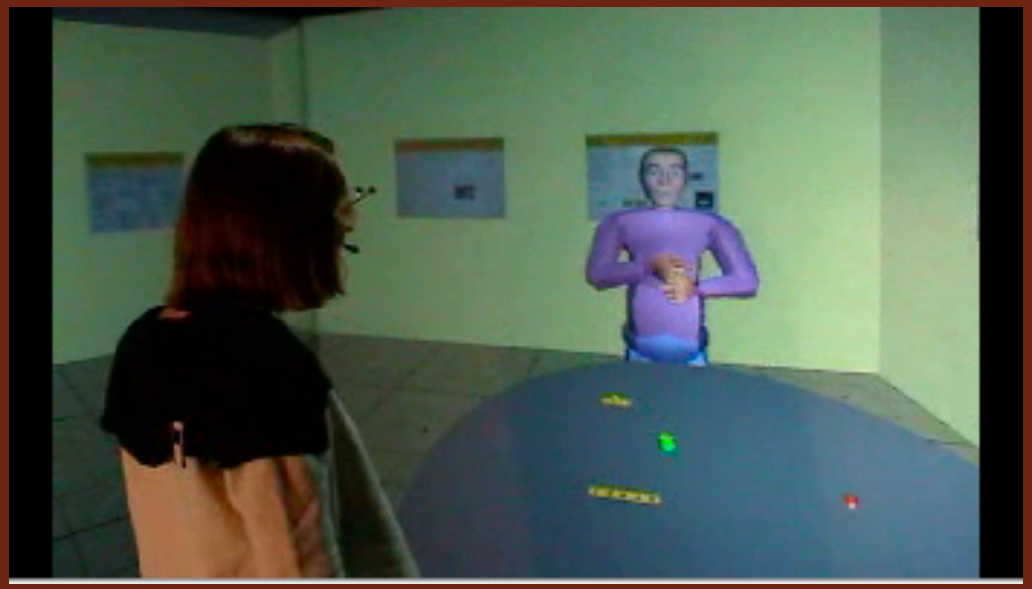
# Prospects, if any, for modeling the GP

In the spirit of an interesting challenge for computer science modeling

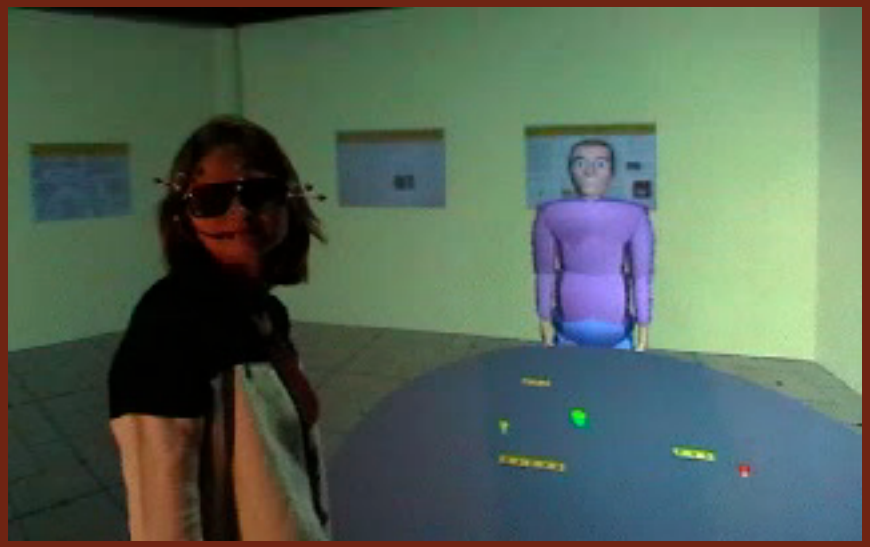
- As we have seen, gestures can be conceptualized as objects of cognitive inhabitance.
- Inhabitance seems utterly beyond current modeling, but here we focus on what is called the global property of gestures: can it be modeled? This would be immensely useful, since then GP experiments with the model could be possible
- A main sticking point appears to be the GP's character as a minimal dialectic unit. One aspect of this is the global character of gesture.
- Being global seemingly cannot be modeled, but listen on.
- The example I shall use is 'Max' - A virtual human developed by Ipke Wachsmuth's group at the University of Bielefeld.

# Max

- Began as testbed for modeling speech and gesture.
- Has since been used in a wide variety of applications - here giving instructions for how to assemble a virtual model airplane from parts.

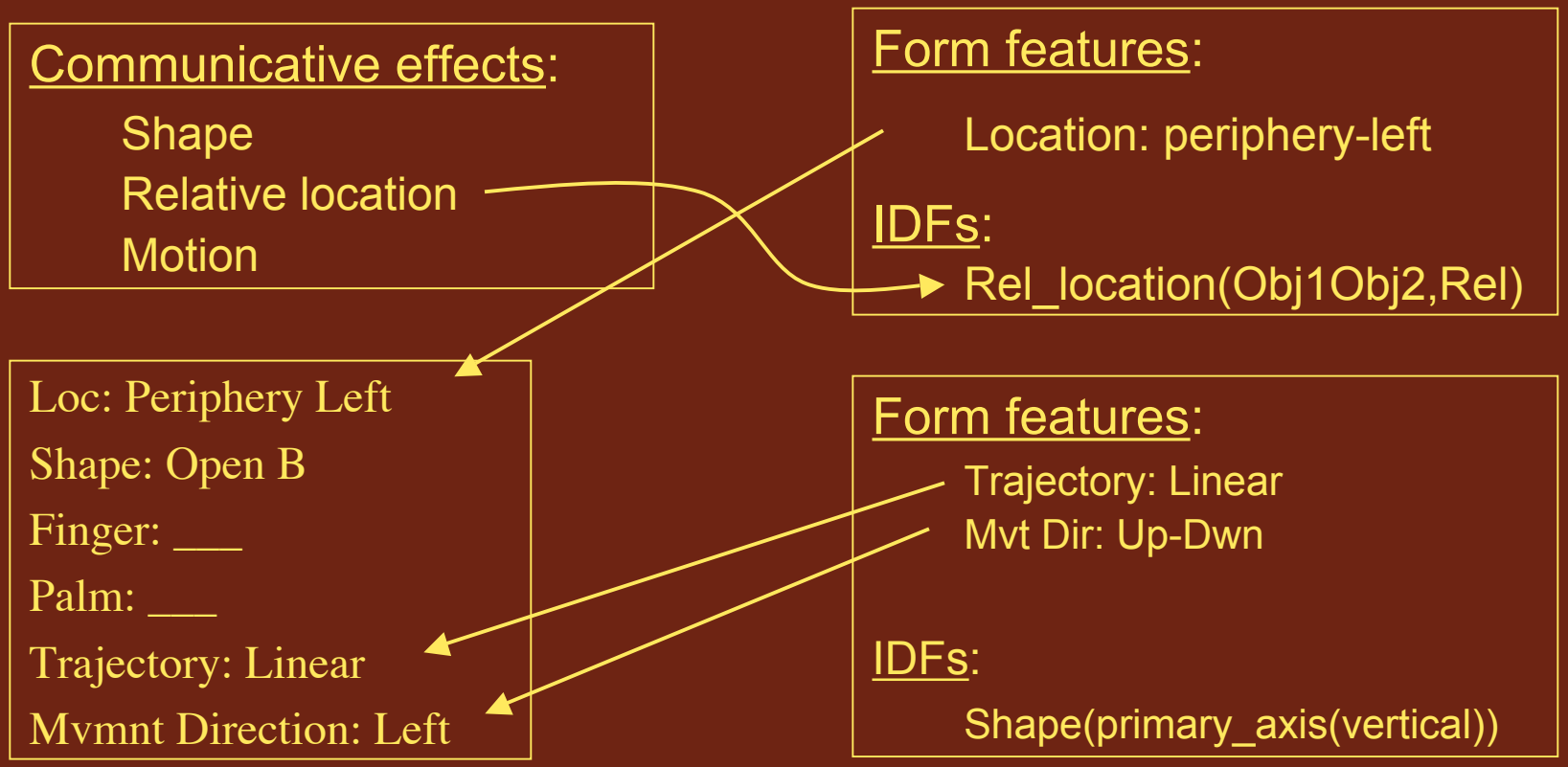


# Max’s gestures



# Gesture planner

- Builds a gesture feature structure by selecting Form Feature Entries whose IDFs (Image Description Features) match the desired communicative effects




Source: Stefan Kopp, Timo Sowa, Paul Tepper & Justine Cassell

# Max does not have GPs - speech-gesture timing is a good test

- Max works as follows – looks ahead, sees what the linguistic resource will be, calculates how far back the preparation will have to be in order for the stroke to coincide with this. Then speech and gesture are generated on their own tracks, and the two assembled into a multimodal utterance.
- In contrast, in the GP the gesture image and linguistic categorization constitute one idea unit, and timing is inherent part of how this thought is created. The start of prep is the dawn of the idea unit, which is kept intact and is unpacked, as a unit, into a full utterance. Timing is inherent.

Source: Sue Duncan

# The problem

- The problem is that use of features forces gesture creation to be combinatoric, thus losing the opposition of semiotic modes essential to the dialectic and the GP.
- To be global, the process wants to work from the overall meaning downward.
- Even if we force a model to proceed in this direction, the form features still need to have their own meanings in order for the model to find them – but do they?

# Is action the solution?

- Suppose that a speaker improvises something that we, the analysts, decide means ‘interior, ‘upward’, and ‘effort’ – what does she need to do to produce this?
- She needs to perform *an action that embodies these meanings*. Does this imply combining form-meaning features? *Or is it enough to ‘act’?* Is thought of the action of rising upward inside the pipe sufficient to generate a gesture with interiority?
- This would be an image both meaning-determined and global-synthetic.

# A resolution

- The idea of *coordinative structures* seems to apply.
- Coordinative structures are not themselves significant forms; they are "flexible patterns of cooperation among a set of articulators to accomplish some functional goal" (anonymous Yale web handout).
- The functional goal is usually thought of as appropriating, affecting, avoiding, etc., some external context whose affordances may select the coordinative structure.

- With the addition of a thought-language-hand link (accessing and steering coordinative structures) use significances rather than external information structures.
  - Ideas or significances are *attractors* of coordinative structures; the coordinative structures zero in on these attractors. Meanings coordinate actions to make gestures (whereas regular actions are coordinated by goals).
  - The existence of a 'thought-language-hand link' in the human brain is revealed by the remarkable case of IW, a man suddenly deafferented from the neck down, who still makes gestures under conditions where he cannot control instrumental actions.
- The properties of the attractor bring out features in the coordinate structures interactively: so *features are outcomes*, not initial conditions, with significances that derive from the action as a whole, and this is the global property.
- This role for coordinative structures under the spell of significance also is compatible with speech-gesture synchrony being inherent to thought.

# Limits

- A weakness of the coordinative structures approach is that it implies a distinction between ‘image’ and ‘gesture’ (the attractor is the image and coordinative structures fashion a gesture to embody it).
- I think this distinction is wrong: the gesture *is* the image, not a copy of it – it is the image in its most material form – so this model creates some distortion.
- Perhaps, as M-P argued,

  “We must recognize first of all that thought, in the speaking subject, is not a representation, that is, that it does not expressly posit objects or relations. The orator does not think before speaking, nor even while speaking; his speech is his thought.”
- While overstated, since to unpack a GP we must often *think further*, his basic point is fundamental: it is error to distinguish sense from signal; moreover, the material carrier concept disavows it.

# To conclude

- In judging modeling, the process is Janus-faced and it's important to distinguish the two agendas:
  - Formatting a theory of human language production on the basis or close systematic observation of natural multimodal language behavior
  - Programming a virtual human to speak and gesture in ways that human observers will feel is 'natural' (Source: Sue Duncan)
- There can be progress on one without the other. I have been arguing that to reach the second goal, which would create a great tool for experimenting with the GP, some way must be found to model global imagery and its dialectic with linguistic form at points of co-expressivity in context.
- The coordinative structures route has promise, especially if a way is found to make attractors out of internal 'significance's, although it still imposes a Cartesian distinction between form ('body') and meaning ('mind').

# References


Freyd, Jennifer J. 1983. Shareability: The social psychology of epistemology. *Cognitive Science* 7: 191-210.

Goldberg, Adele. 1995. *Constructions: A Construction Approach to Argument Structure*. Chicago: University of Chicago Press.

Harris, Roy. 2003. *Saussure and His Interpreters*, 2nd ed. Edinburgh: Edinburgh University Press.

McNeill, David, & Levy, Elena. 1982. Conceptual Representations in Language Activity and Gesture. In R. J. Jarvella & W. Klein (eds.), *Speech, Place, and Action*, pp 271-296. Chichester, Eng: John Wiley & Sons Ltd.

McNeill, David. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.

Merleau-Ponty, Maurice. 1962. *Phenomenology of Perception* (C. Smith, trans.). London: Routledge.

Sowa, Timo, Kopp, Stefan, Duncan, Susan, McNeill, David & Wachsmuth, Ipke. 2008. Implementing a non-modular theory of language production in an embodied conversational agent In I. Wachsmuth, et al. (eds.), *Embodied Communication in Humans and Machines*. Oxford University Press.

Vygotsky, L. S. 1987. *Thought and Language*. Edited and translated by E. Hanfmann and G. Vakar (revised and edited by A. Kozulin). Cambridge: MIT Press.

Wundt, Wilhelm. 1970. The psychology of the sentence. In Arthur Blumenthal (ed. and trans.), *Language and Psychology: Historical Aspects of Psycholinguistics*, pp. 20-33. New York: John Wiley & Sons Ltd.

Zinchenko, V. P. 1985. Vygotsky's ideas about units for the analysis of mind. In James V. Wertsch (ed.), *Culture Communication, and Cognition : Vygotskian Perspectives*, pp. 94-118. Cambridge: Cambridge University Press.

# Many Thanks!

Merleau-Ponty c. 1950

Vygotsky c. 1930

*Supported by grants from NSF*

*Thanks to the 'GP to the Max' group at the University of Chicago*