

# Nonverbal Behavior

Not in textbook

[hannes@ru.is](mailto:hannes@ru.is)

T-725-MALV

# Multimodal Communication

- During face-to-face conversation, the entire body is in **constant motion**. Body movement is **seamlessly integrated** with **other modalities** without much thought.
- Beyond speech, we see for example:
  - Gesture
  - Head Movement
  - Gaze Shifts
  - Posture Changes

# Serving Communicative Purpose?

- Is this symphony of observed behavior actually **part of the communication?**
- Or is it superfluous and simply part of some unrelated need for the body to stay in motion?
- Does the spoken text perhaps completely carry the communication?

# Face-to-Face Dialog is Significant

- Compared to text, **face-to-face dialog** is considered the basic and **primary** use of language.



- We can therefore imagine that the verbal is most naturally in the context of the nonverbal, and that the **two modes evolved together**.

# Communicative Act

- Face-to-face dialog is **not simply:**  
verbal + nonverbal communication
- The verbal and nonverbal belong to the **same seamless and comprehensive communicative act.**
- Many prefer the term **Dialog Act** to over the more verbally oriented **Speech Act** when describing intension in the discourse of dialog.

# Communicative Purpose: Evidence?

- A range of behavior can be seen when people have a conversation.
- **How do we know** that some of these behaviors are actually **communicative**?



# Communicative Purpose: Evidence?

- **Social:** They correlate with the availability of a receiver.
- **Form:** They can take on an unnatural form to convey concrete or abstract things.
- **Synchrony:** They are very tightly synchronized with the verbal stream during conversation.
- **Context:** Their meaning/function becomes apparent in the current discourse context.

# Devices Carrying out Functions

- If the nonverbal behaviors communicate, they are carrying out a **Discourse Function**.
- They can therefore be considered a set of available **Discourse Devices**, expanding the range of devices provided by text-only communication.



# Studying the Correlation

- The **possible connection** between nonverbal behavior as a discourse device and a discourse function can be **formally studied**.
- Such study begins with audio-visual **data collection** from a face-to-face encounter, typically involving a conversation.



# Method

1. The speech is transcribed and the **discourse functions** arising from **linguistic devices** identified.
2. All **observed behavior** of interest annotated in isolation from the speech and text.
3. The **temporal correlation** between identified discourse functions and annotated behavior is examined and high predictability considered a discovery of a useful mapping.

# Tools

- Audio transcription, behavior annotation and the examination of correlation can take place in a specialized **multimodal annotation tool** where multiple channels can hold information about each modality.

# Multimodal Annotation Tool: Anvil

The screenshot displays the Anvil multimodal annotation tool interface, divided into two main windows.

**Top Window (Video Player):**

- Left Panel (Console):** Contains a list of system messages such as "BUG MODE IS ON", "open file iq1-kara.anvil", "closing annotation", "read anvil file: anvil.Annotation@34263a", "create player for: video.quicktime:", "FrameRate = 0.0", "Movie duration: 02:27:28", "No. of frames: 3681", "open player (first)", and "creating annotation window". Below this, it shows "Frame number: 458", "Media time: 00:10:31", and "Modified: true".
- Center:** A video player showing a scene with three men. A vertical slider on the left indicates the current frame position.
- Right Panel (Track Management):** Shows a track named "gesture.metaphoric" with "Start: 00:10:23" and "End: 00:10:47". Below this, there are fields for "Attributes" (phase: stroke, location-height: chest, location-side: outer-right, handedness: right) and a "Comment" field containing "copy gesture" and "-> elaboration". At the bottom of this panel are buttons for "start", "edit", "end", "unlink", "<-link", and "del".

**Bottom Window (Timeline):**

The timeline displays a sequence of frames from 16 to 24. It features several horizontal tracks for different types of annotations:

- ling (Linguistic):** Shows text segments like "d [brea, das, der", "Erzähler", "d [sch-Erzähl...", "Schl. verb, es g.l.a. a. ande Geschi", "[brea, Liebesgeschl. zwisc. [breath] einer Jüdin".
- posture:** Includes sub-tracks for "head" (with "fa" and "lower" annotations) and "upper".
- beat:** Shows colored blocks representing different beats or segments.
- gesture:** Includes sub-tracks for "deictic", "emblem", "iconic", and "metaphoric" (with yellow, red, and green blocks).
- adaptive:** A track at the bottom with a few colored blocks.

# Some Results

- These kinds of studies have uncovered various interesting relationships.
- One can find them in a myriad of **published** science articles typically categorized as dealing with **Communicative Behavior, Multimodal Behavior, Social Behavior** or **Psychosocial Behavior**.

# Gaze

- Where people look, i.e. eye gaze, can indicate various things like:
  - Speakers often **select next speaker** with gaze near the end of their own turn (Kendon, 1990).
  - Speakers usually break eye-contact at **turn beginning** (Argyle and Cook, 1976).
  - Speakers perform **grounding** by looking at listener and/or listener looks up at speaker from task if explanation is needed (Nakano et al., 2003).

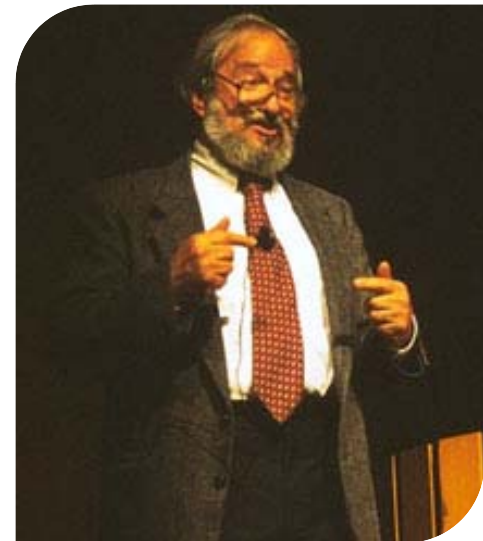


# Gesture

- Conversational gesture is a very complex phenomenon. The general gesture activity has some clear correlations:
  - Gestures are almost exclusively **made by speakers**, while listeners do not gesture unless asking for the floor.
  - Gestures are more likely to occur when **new discourse entities** are created, than when older ones are referred to (Cassell et al. 1994).

# Gesture

- Main Categories of Conversation Gesture
  - **Beats:** Emphasis
  - **Deictics:** Referring Expressions
  - **Iconics:** Concrete Shapes
  - **Metaphorics:** Abstract Concepts
  - **Emblems:** Conventional Signs





# Posture

- People often change posture when they **change the topic** (Cassell et al. 2001).
- The initiator will lead with a shift and if the **listener accepts**, the listener will shift as well, mirroring the speaker.
- This is an example of a **silent negotiation** that takes place nonverbally.



# Computer Applications

- **Speech Understanding:** Looking at both verbal and nonverbal behavior may provide redundancy which is helpful if the environment is noisy. One could say that with more discourse devices available behind each function, the more likely the function will correctly be decoded.

# Computer Applications

- **Animated Characters:** If the computer is represented by a graphical character, it is very important that the character's nonverbal behavior is appropriate and well synchronized with the character's verbal behavior, otherwise inconsistent communication results.



# Computer Applications

- **Computer Mediated Communication:** A narrow communication channel, like text chat, can be augmented with nonverbal behavior if the discourse functions from the textual devices are known.

