Dialogue across the Lifespan

June 2022 | Lecture 4

Jonathan Ginzburg
Andy Lücking

Laboratoire de Linguistique Formelle
Université Paris Cité

23rd June 2022

Yesterday's lecture

- Characterizing the response space of queries
- Non Sentential Utterances and dialogue context

Today's Lecture

1. Partiturs: multimodal input representations
2. Gesture perception as type assignment (by example of manual co-speech gesture)
3. Multimodal integration in multimodal grammar
4. Head shake and 'No'
5. Non-verbal social signals: laughter, smiling, crying

➔ the 'naturalness' of natural language interaction

Where we are now: Today's lecture in context

Desiderata of dialogue competencies (modified Turing test; Lect. 1)

- meanings actually talked about, ex. QNPs (Lect. 2)
- dialogical relevance, response space; constraining coherent behaviour (Lect. 3)
- non-verbal (social) signals; towards expressivity (Lect. 4 [today])
- grammar and dialogical competency as 'organic system': from acquisition to forgetting (Lect. 5)

All in a uniform formal framework (Lect. 1 and throughout)

Clarifying *Frauenkirche* [Church of Our Lady]

SaGA V8 (Lücking et al., 2010)

- route direction dialogue
- left: **router** (R), right: **follower** (F)

About:

tidied up, translated exchange:

R: well, it is a bit cross-like, because of the two towers, the arrangement of the towers is like the Frauenkirche
F: Frauenkirche? The one in Dresden?
R: eh, [...] Munich

Clarifying *Frauenkirche* – closer look

ja, ist auch so'n bisschen (..) kreuzartig gebaut (..) durch

die beiden Türme ähm also von der Anordnung der Türme

sind die

so wie die Frauenkirche 'n bisschen

jeweils am –

die Frauenkirche?

Dresdner?

äh [...] München

Clarifying *Frauenkirche* – closer look

ja, ist auch so'n bisschen (..) kreuzartig gebaut (..) durch

shaping (2×)

die beiden Türme ähm also von der Anordnung der Türme

placing

placing

nod *inhale*

sind die

placing

so wie die Frauenkirche 'n bisschen

hedging

jeweils am –

placing

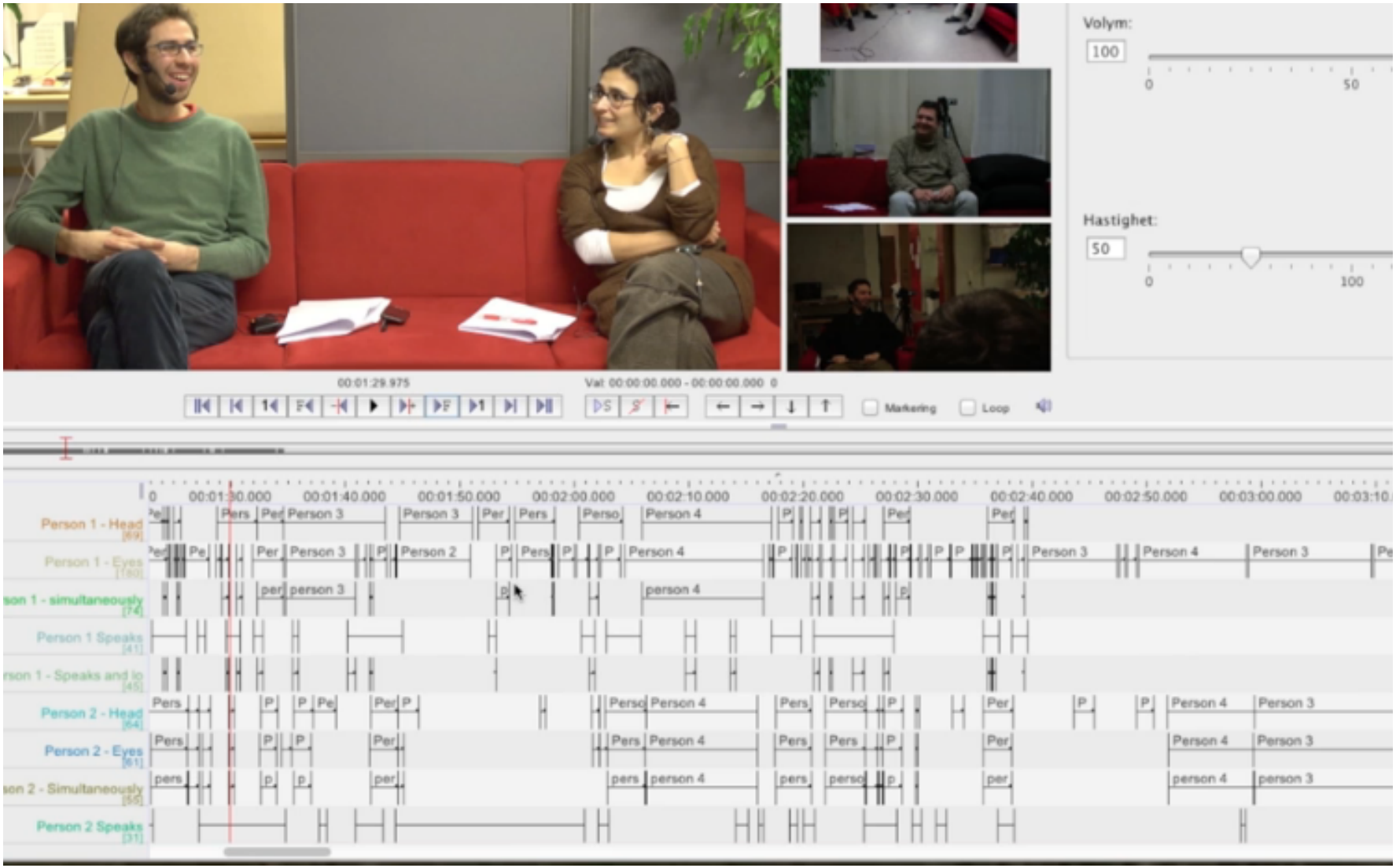
die Frauenkirche?

away

router Dresdner?

äh [...] München

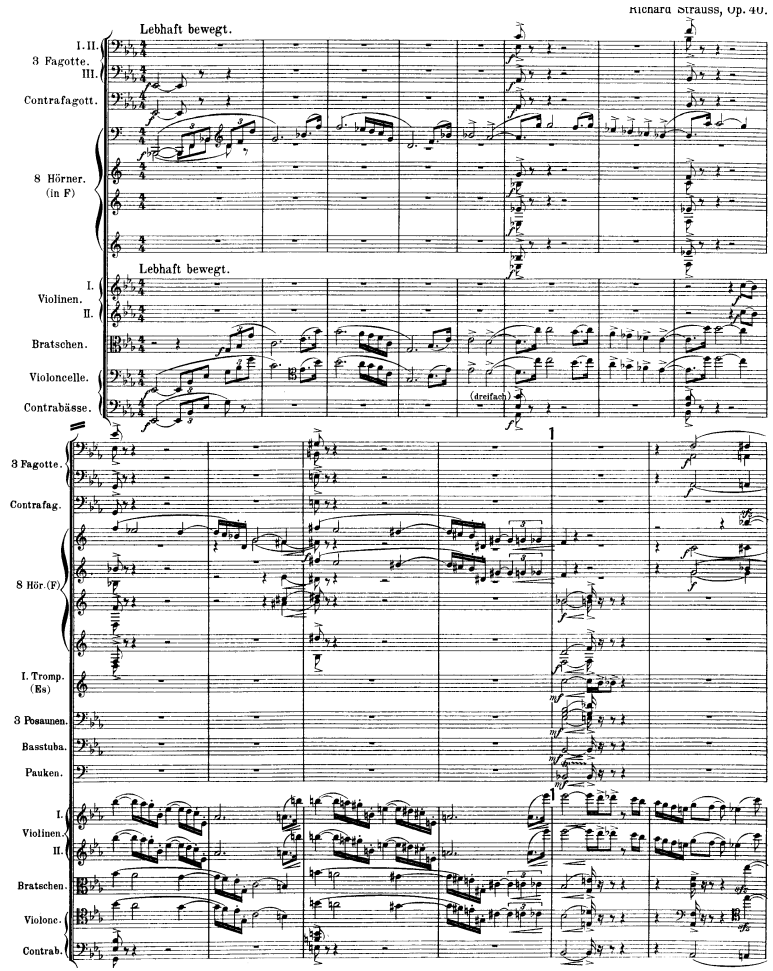
More impressive annotation (Engström, KTH Sweden)



4-party discourse, transcribed for speech, head direction and interlocutor gazed at (https://vimeo.com/84295277)

Scores

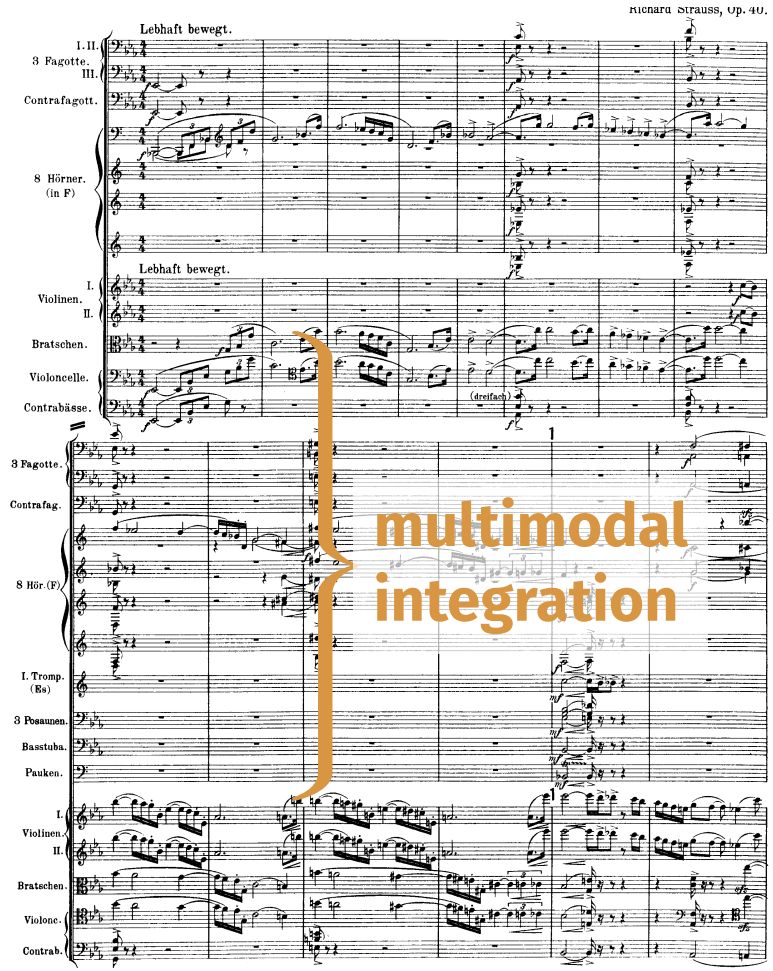
- Natural language interaction looks like a musical score



(*Heldenleben*, CC BY-SA 4.0, IMSLP)

Scores

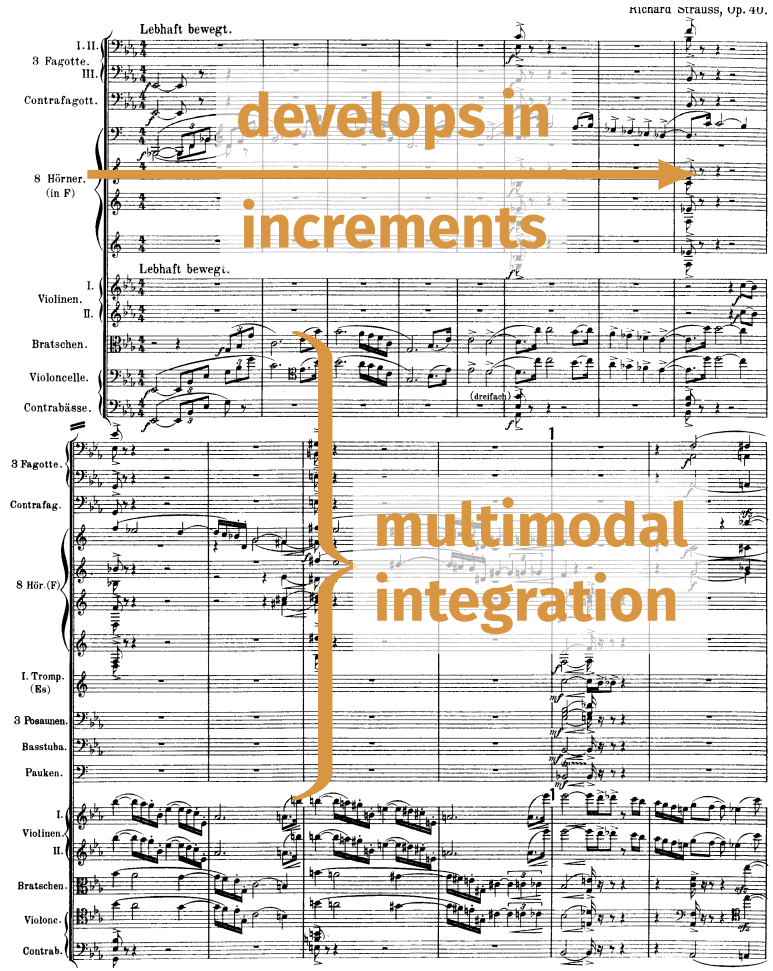
- Natural language interaction looks like a musical score
- ➔ ‘vertically’: heavy multimodal integration, within speakers and between speakers



(*Heldenleben*, CC BY-SA 4.0, IMSLP)

Scores

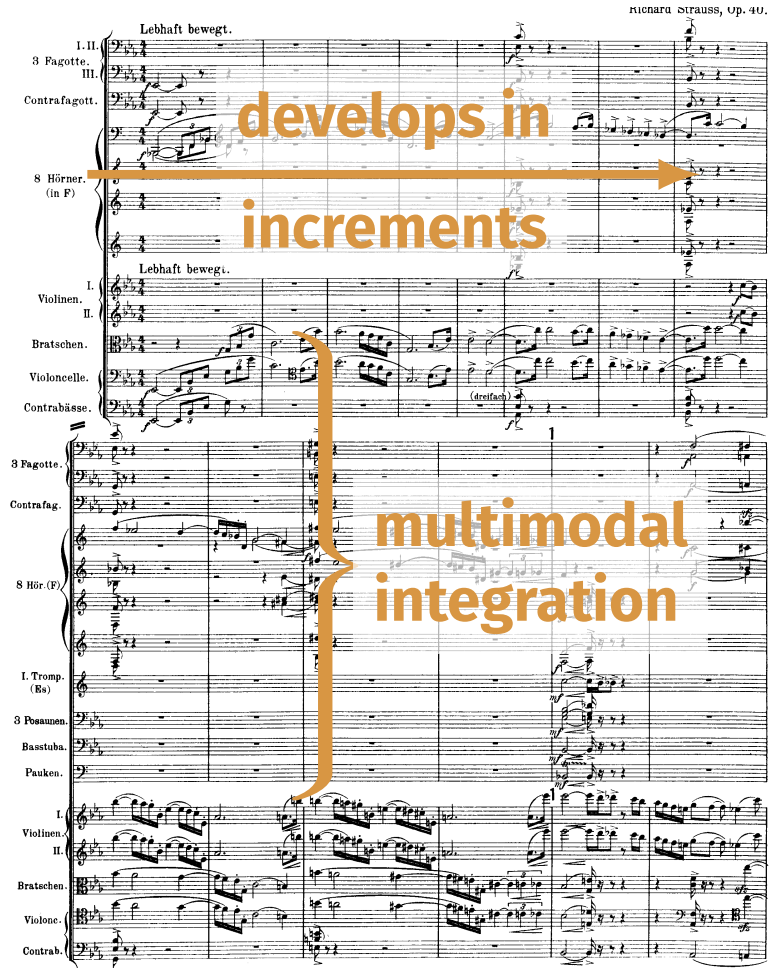
- Natural language interaction looks like a musical score
- ➔ ‘vertically’: heavy multimodal integration, within speakers and between speakers
- ➔ ‘horizontally’: incrementalizing turns



(*Heldenleben*, CC BY-SA 4.0, IMSLP)

Scores

- Natural language interaction looks like a musical score
- ➔ ‘vertically’: heavy multimodal integration, within speakers and between speakers
- ➔ ‘horizontally’: incrementalizing turns
- ➔ technically: incremental, multimodal dialogue theory
- ➔ conceptually: ‘horizontal’ and ‘vertical’ coherence



(*Heldenleben*, CC BY-SA 4.0, IMSLP)

Tiers I

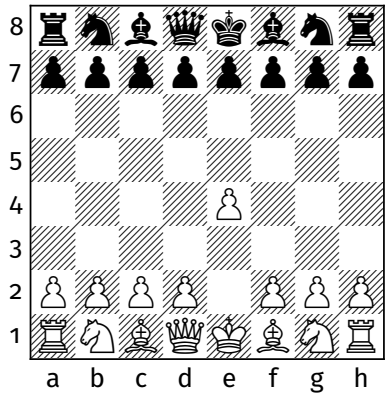
Articulator	*Articulated*	*Mode*	*Signal*
vocal tract	sound waves	auditive	phones
vocal tract	sound waves	auditive	prosody, stress
vocal tract	sound waves	auditive	laughter
arm, hand	movements, shapes	visual	manual gesture
eyes	gaze	visual	focus of attention
facial muscles	face	visual	facial expressions
shoulders	shoulder position	visual	shrug
head	head position	visual	greeting, bending
arm, hand	touch	tactile	handshake

Tiers II

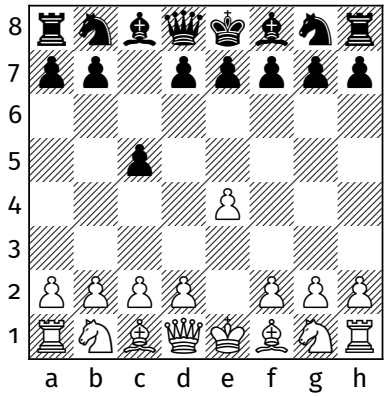
- Partiturs: channels as dimensions of a common communication event *e*

$$partitur := \left[\begin{array}{l} e : (\left[\begin{array}{ll} e_{speech} & : Phon \\ e_{gesture} & : Trajectory \\ e_{gaze} & : RecType \\ e_{head} & : headMove \\ e_{face} & : faceExpr \end{array} \right])^{+} \end{array} \right]$$

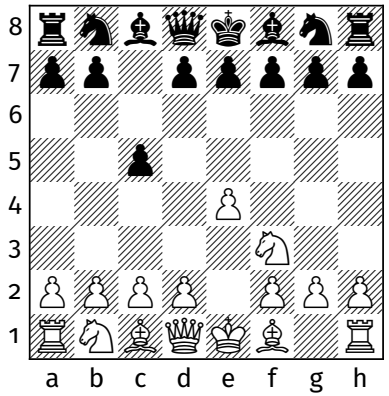
- String types: flip book theory of events (Fernando, 2011)
- Example: Sicilian opening **1 e4 c5 2 ♘f3**



1 e2–e4



1. . . c7–c5



2 ♘g1–f3

Tiers III

- concatentation of three records: pawn move (a_1), pawn move (a_2), knight move (a_3) $= a_1a_2a_3$
- shorthand for temporally indexed record: $\begin{bmatrix} t_0 = a_1 \\ t_1 = a_2 \\ t_2 = a_3 \end{bmatrix}$ where $t_0 \prec t_1 \prec t_2$
- corresponding concatenation of record types: $T_{e4}{}^\frown T_{c5}{}^\frown T_{Sf3}$, indicated by '$^\frown$'
- Judgement: $a_1a_2a_3 : T_{e4}{}^\frown T_{c5}{}^\frown T_{Sf3}$
- Using record types:
  $\left[\text{e} : \text{move(pawn,e2,e4)}\right]^\frown\left[\text{e} : \text{move(pawn,c7,c5)}\right]^\frown\left[\text{e} : \text{move(knight,g1,f3)}\right]$
- Label *e* labels a dimension of the string type.
- String types underlie incrementality/incremental processing

- Lex('Beethoven', NP):

$$\begin{bmatrix}
\text{s-event} : \begin{bmatrix} \text{e} & : \texttt{beethoven} \\ \text{spkr} & : \textit{Ind} \\ \text{addr} & : \textit{Ind} \\ \text{c}_{\text{sp}} & : \text{addressing(spkr,addr,e)} \end{bmatrix} \\
\text{syn} : \begin{bmatrix} \text{phon} & : /\text{Beethoven}/ \\ \text{cat=np} & : \textit{Cat} \\ \text{dtrs=}\langle\rangle & : \text{List}(\textit{Sign}) \end{bmatrix} \\
\text{q-params} : \begin{bmatrix} \text{refind} : \textit{Ind} \\ \text{c}_{\text{nm}} : \text{named(refind, 'Beethoven')} \end{bmatrix} \\
\text{cont=refind} : \textit{Ind}
\end{bmatrix}$$

Strings and incremental processing II

- Start chart parsing:

$$\left[\begin{array}{l} e_1 = \texttt{beethoven} : \textit{Phon} \\ e_2 : \text{Lex}(\text{`Beethoven'}, \text{NP}) \wedge_{\textit{merge}} \left[\text{s-event} : \left[\text{e}=e_1 : /\text{Beethoven}/\right]\right] \\ e_3 : \left(\left[\begin{array}{l} \text{rule}=\text{S}\rightarrow\text{NP VP} : \text{NP}^\frown\text{VP} \rightarrow \textit{Type} \\ \text{fnd}=e_2 : \textit{Sign} \\ \text{req}=\text{VP} : \textit{Sign} \\ \text{e} : \text{required(req,rule)} \end{array}\right]\right) \\ e : \left(\left[\begin{array}{l} e_1 : \text{start}(e_1) \\ e_2 : \text{start}(e_2) \end{array}\right]^\frown \left[\begin{array}{l} e_1 : \text{end}(e_1) \\ e_2 : \text{end}(e_2) \\ e_3 : \text{start}(e_3) \end{array}\right]^\frown \left[e_3 : \text{end}(e_3)\right]\right) \end{array}\right]$$

- But how to include gestures?
  1. within multimodally extended grammar
  2. by means of conversational rules

Gesture perception as type assignment

- TTR: linguistic processing as type assignment
- Schematically: $\begin{bmatrix} \text{sit} & : \text{speech event} \\ \text{sit-type} & : \text{grammatical sign} \end{bmatrix}$
- Extension to gesture: perceptual classification as type assignment

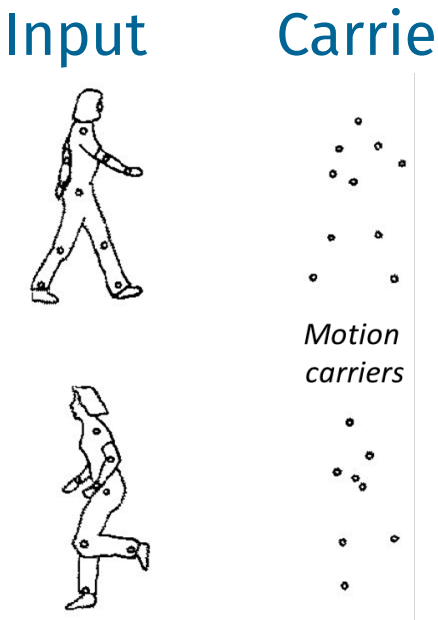
Vector analysis of biological motion

- Motion perception can be captured by means of a vector model (Johansson, 1973).
- Rotation and translation Carriers are the basis for the vector model.

Input

Vector analysis of biological motion

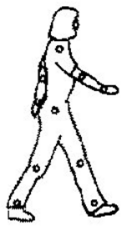
- Motion perception can be captured by means of a vector model (Johansson, 1973).
- Rotation and translation Carriers are the basis for the vector model.

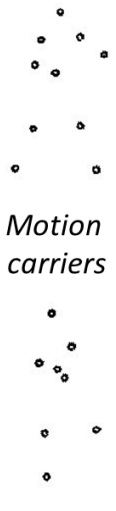


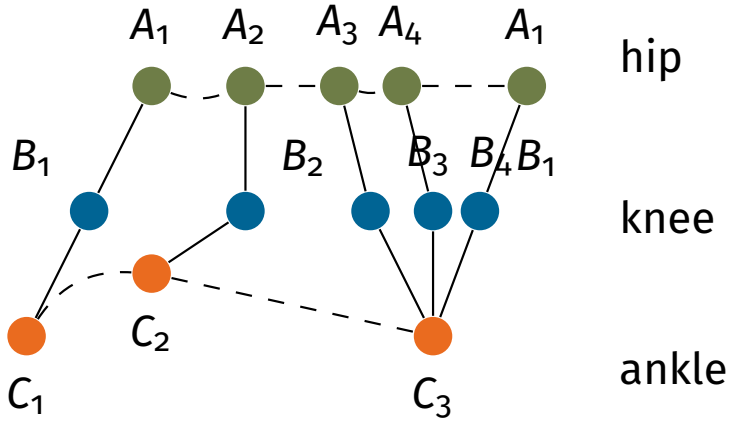
Vector analysis of biological motion

- Motion perception can be captured by means of a vector model (Johansson, 1973).
- Rotation and translation Carriers are the basis for the vector model.

Input Carrier Carrier movement



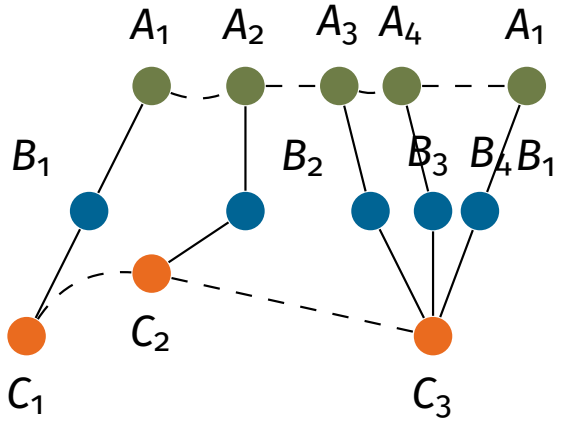


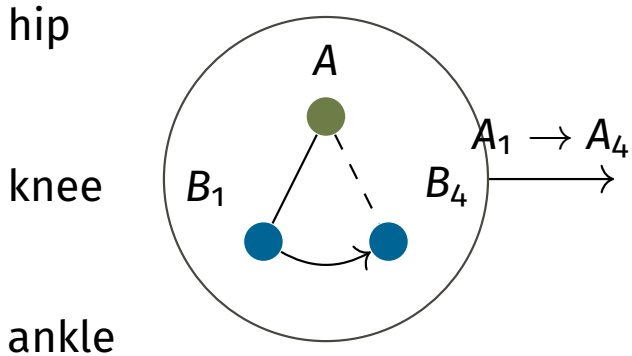


Vector analysis of biological motion

- Motion perception can be captured by means of a vector model (Johansson, 1973).
- Rotation and translation Carriers are the basis for the vector model.

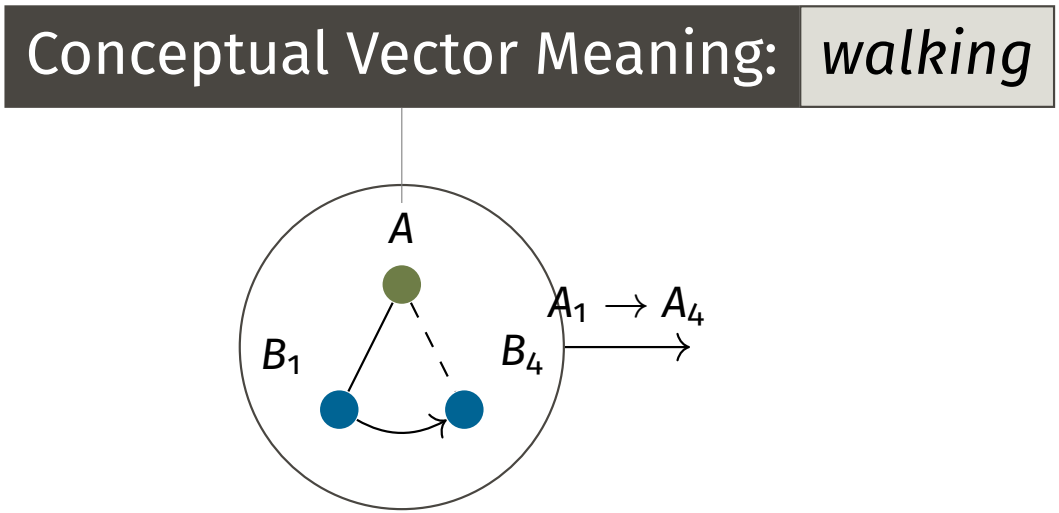
Input Carrier Carrier movement Abstract vector model



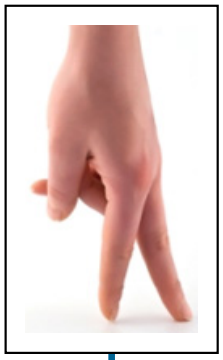


factoring out common movement shares

Gesture as vector model exemplifiers

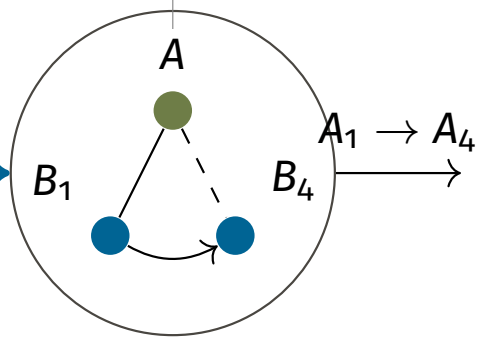


Gesture as vector model exemplifiers

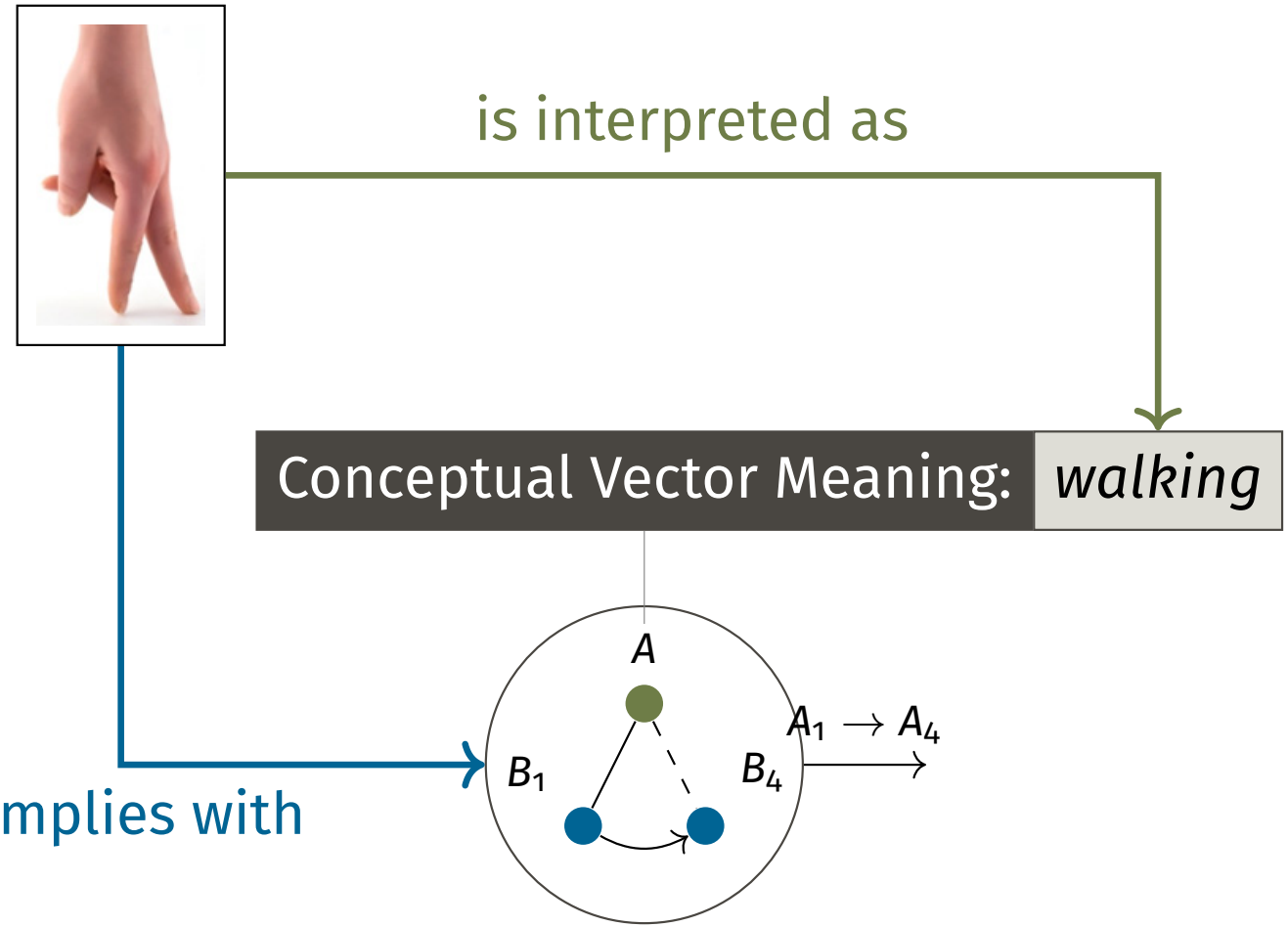


Conceptual Vector Meaning: *walking*

complies with



Gesture as vector model exemplifiers



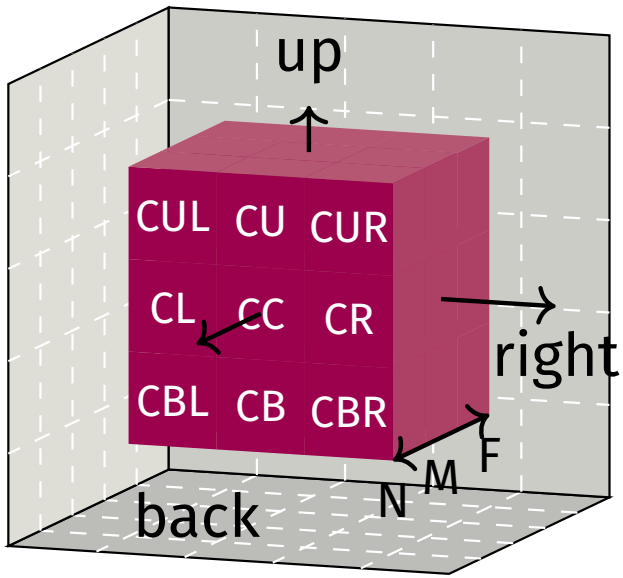
Representing gestures

$$
\begin{bmatrix}
\text{hand} & = \text{right} \\
\text{hs} & = \text{claw} \\
\text{carrier} & = \begin{bmatrix} \text{boh} & = \text{none} \\ \text{plm} & = \text{none} \\ \text{wrst} & = \text{MR>MB>ML} \\ \text{move} & = \text{line>line>line} \end{bmatrix} \\
\text{sync} & = \begin{bmatrix} \text{sloc} & = \text{CBR-F} \\ \text{eloc} & = \text{CBR-N} \\ \text{stime} & = 2{:}32 \\ \text{etime} & = 2{:}33 \end{bmatrix} \\
\text{rel} & = \text{none}
\end{bmatrix}
$$

- Annotation format:
  - handedness (right, left)
  - handshape (modified ASL lexicon)
  - movement carrier (back-of-hand, palm or wrist; path of movement)
  - synchronized info (temporal, local)
  - relation to other hand
- The values of the features are of type *AP* (*annotation predicate*), e.g. [hs : *AP*]

Gesture Space Model

start and end locations of gesture movements are given in terms of three-dimensional **gesture space** (adapted from two-dimensional model of McNeill (1992))

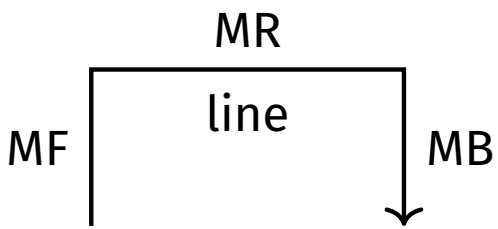


CBL:	center below left
CL:	center left
CUL:	center upper left
CB	center below
CC:	center center
…	…
N:	near
M:	middle
F:	far

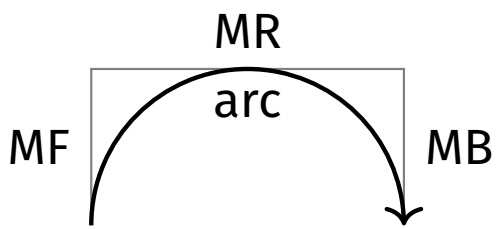
Movements: lines vs. arcs

- A movement is captured in terms of a direction seen from the speaker (e.g. *move forward* (MF)) and
- a concatenation type which distinguishes straight ("line") from roundish ("arc") trajectories.
- Complex movements are built by combinations of directions ('>').

$$\begin{bmatrix} \text{wrst} & = \text{MR>MB>ML} \\ \text{move} & = \text{line>line} \end{bmatrix}$$



$$\begin{bmatrix} \text{wrst} & = \text{MR>MB>ML} \\ \text{move} & = \text{arc>arc} \end{bmatrix}$$

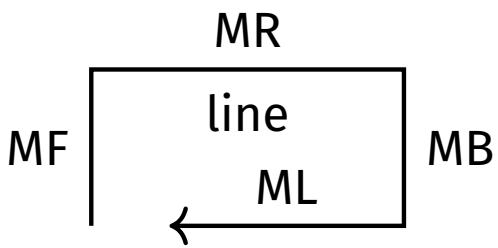


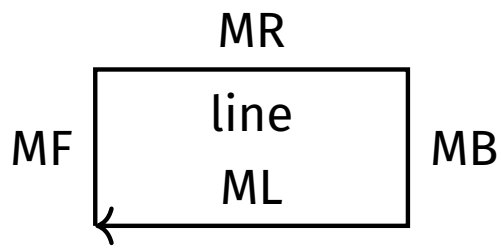
Open vs. closed paths

- Movements are underspecified with regard to the lengths of the movement parts.
- Closed and open paths are discriminated in terms of the *sync*-feature.

wrst	= MF>MR>MB>ML
move	= line>line>line>line
sloc	= CC-M
eloc≠sloc	= CR-M

wrst	= MF>MR>MB>ML
move	= line>line>line>line
sloc	= CC-M
eloc=sloc	= CC-M





Vector types

- Gesture annotations are mapped onto vector sequence representations **p** form spatial vector semantics (Zwarts, 2003): $\mathbf{p} : [0, 1] \mapsto \mathbf{V}$.
- Format:
  - **Type**: axis, place, outline, ...(Zwarts, 2005)
  - **Path**: description of contour (Zwarts, 2003)
  - **Shapes**: shape constraint (cf. Weisgerber, 2006)
- $Vec =_{\text{def}} \begin{bmatrix} \text{vt} : \text{Vtype} \\ \text{pt} : \text{Vpath} \\ \text{sh} : \text{multiset(Vshape)} \end{bmatrix}$
- Rule-based translation from gesture event to vector type: π_v and π_d.

Configuration	=	Vector π_v	→	Constraints π_d
Handshape $\in$ {C, 5, B, O, Y}	=	$\{\mathbf{u}\}$	→	volume
{MF, MR, MB, ML}	=	$\mathbf{u}$	→	translational
$\emptyset$	=	–	→	–
MF>MR + line	=	$\mathbf{u} \perp \mathbf{v}$	→	orthogonal
MR>MB + line	=	$\mathbf{u} \perp \mathbf{v}$	→	orthogonal
MB>ML + line	=	$\mathbf{u} \perp \mathbf{v}$	→	orthogonal
MF>ML + arc	=	$\mathbf{u} \circ \mathbf{v}$	→	quadrant
MF>MR + arc	=	$\mathbf{u} \circ \mathbf{v}$	→	quadrant
...	=	...	→	...
MF + ...+ MB	=	$\mathbf{u}, \mathbf{u}^{-1}$	→	inverse
ML + ...+ MR	=	$\mathbf{u}, \mathbf{u}^{-1}$	→	inverse
sloc = eloc	=	$\mathbf{u}(0) = \mathbf{v}(1)$	→	closed
sloc $\neq$ eloc	=	$\mathbf{u}(0) \neq \mathbf{v}(1)$	→	open
lh.sloc = rh.sloc +	=	$\mathbf{u}(0) = \mathbf{v}(0)$		
lh.eloc = rh.eloc [two-handed]	=	$\mathbf{w}(1) = \mathbf{x}(1)$	→	closed
quadrant + quadrant + invers				semicircle
semicircle + semicircle + closed				circle
orthogonal + orthogonal + invers + open				rectangular
orthogonal + orthogonal + invers + closed				rectangle
...				...

Configuration	=	Vector π_v	→	Constraints π_d
Handshape $\in$ {C, 5, B, O, Y}	=	$\{\mathbf{u}\}$	→	volume
{MF, MR, MB, ML}	=	$\mathbf{u}$	→	translational
$\emptyset$	=	–	→	–
MF>MR + line	=	$\mathbf{u} \perp \mathbf{v}$	→	orthogonal
MR>MB + line	=	$\mathbf{u} \perp \mathbf{v}$	→	orthogonal
MB>ML + line	=	$\mathbf{u} \perp \mathbf{v}$	→	orthogonal
MF>ML + arc	=	$\mathbf{u} \circ \mathbf{v}$	→	quadrant
MF>MR + arc	=	$\mathbf{u} \circ \mathbf{v}$	→	quadrant
...	=	...	→	...
MF + ...+ MB	=	$\mathbf{u}, \mathbf{u}^{-1}$	→	inverse
ML + ...+ MR	=	$\mathbf{u}, \mathbf{u}^{-1}$	→	inverse
sloc = eloc	=	$\mathbf{u}(0) = \mathbf{v}(1)$	→	closed
sloc $\neq$ eloc	=	$\mathbf{u}(0) \neq \mathbf{v}(1)$	→	open
lh.sloc = rh.sloc +	=	$\mathbf{u}(0) = \mathbf{v}(0)$		
lh.eloc = rh.eloc [two-handed]	=	$\mathbf{w}(1) = \mathbf{x}(1)$	→	closed
quadrant + quadrant + invers				semicircle
semicircle + semicircle + closed				circle
orthogonal + orthogonal + invers + open				rectangular
orthogonal + orthogonal + invers + closed				rectangle
...				...

Vectorizing our example

- $$\pi_v\left(\begin{bmatrix}\text{wrst} = \text{MR>MB>ML}\\ \text{move} = \text{line>line>line}\\ \text{sync} = \begin{bmatrix}\text{sloc} = \text{p1}\\ \text{eloc} = \text{p2} \neq \text{p1}\end{bmatrix}\end{bmatrix}\right) = \begin{bmatrix}\text{pt1} : \begin{bmatrix}\mathbf{u} \perp \mathbf{v} \perp \mathbf{w}\\ \mathbf{u}(\text{o}) \neq \mathbf{w}(\mathbf{1})\end{bmatrix}\end{bmatrix}$$
- $$\pi_d\left(\begin{bmatrix}\text{pt1} : \begin{bmatrix}\mathbf{u} \perp \mathbf{v} \perp \mathbf{w}\\ \mathbf{u}(\text{o}) \neq \mathbf{w}(\mathbf{1})\end{bmatrix}\end{bmatrix}\right) = \begin{bmatrix}\text{sh} : \{\text{rectangular, open}\}\end{bmatrix}$$

Perceptual contents

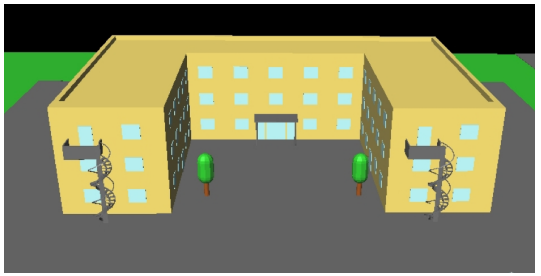
- The intensions of some predicates have a **Conceptual Vector Meaning (CVM)**, representing their perceptual impression in terms of **vector sequences** (Lücking, 2013).
- $[\![\text{U-shaped}]\!] =$

$$\begin{bmatrix} x : \text{Ind} \\ c_u : \text{U-shaped}(x) \\ \text{cvm} = \begin{bmatrix} \text{vt} : \text{axis-path}(x, \text{pt}) \\ \text{pt} : \begin{bmatrix} \mathbf{u} \perp \mathbf{v} \perp \mathbf{w} \\ \mathbf{u}(0) \neq \mathbf{w}(1) \end{bmatrix} \\ \text{sh} : \{\text{rectangular}, \text{open}\} \end{bmatrix} : \textit{Vec} \\ c_{\text{shape}} : \text{shape}(x, \text{cvm}) \end{bmatrix}$$

Demonstration

‘dann ist das Haus halt so’

‘then the house is like this’



Annotation:

$$\begin{bmatrix} \text{wrst} = \text{MR>MB>ML} \\ \text{move} = \text{line>line>line} \\ \text{sync} = \begin{bmatrix} \text{sloc} = \text{p1} \\ \text{eloc} = \text{p2} \neq \text{p1} \end{bmatrix} \end{bmatrix}$$

Vector representation:

$$\begin{bmatrix} \text{pt1} : \begin{bmatrix} \mathbf{u} \perp \mathbf{v} \perp \mathbf{w} \\ \mathbf{u}(0) \neq \mathbf{w}(1) \end{bmatrix} \\ \text{sh} : \{\text{rectangular, open}\} \end{bmatrix}$$

Processing *house*

- Lexical entry: $[\![\text{house}]\!] =$

$$\begin{bmatrix} \text{bg} = \begin{bmatrix} \text{x} : \text{Ind} \end{bmatrix} \\ \text{f} \quad = \lambda r : \text{bg} \,.\, \left(\begin{bmatrix} \text{c}_{\text{hs}} : \text{house}\ (r.\text{x}) \\ \text{cvm} : \textit{Vec} \\ \text{c}_{\text{shape}} : \text{shape}(r.\text{x}, \text{cvm}) \end{bmatrix} \right) \end{bmatrix}$$

- Information state after processing the noun:

$$s_{t+1} = \begin{bmatrix} \text{x} & : \text{Ind} \\ \text{c}_{\text{hs}} & : \text{house}\ (\text{x}) \\ \text{cvm} & : \textit{Vec} \\ \text{c}_{\text{shape}} & : \text{shape}(\text{x}, \text{cvm}) \end{bmatrix}$$

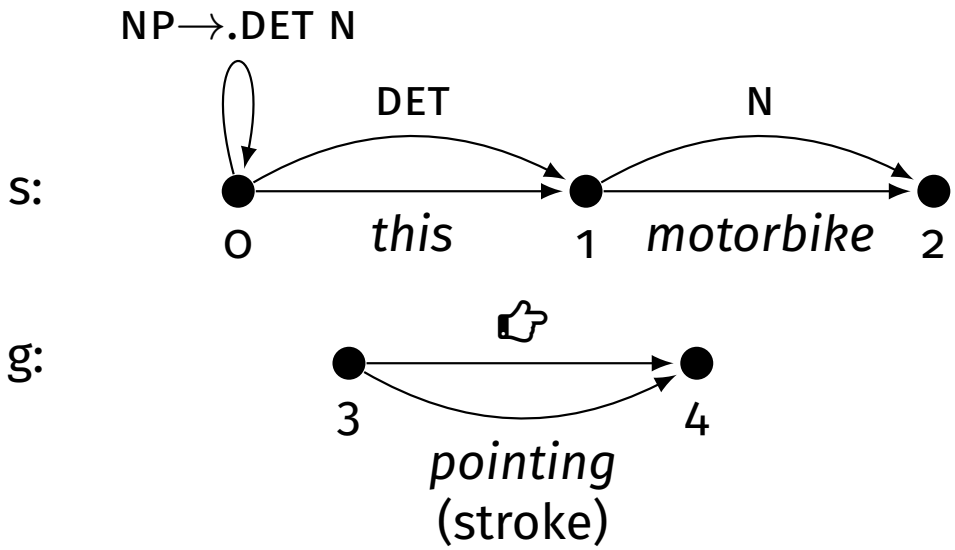
Adding gesture

- Gesture updates cvm of s_{t+2} and introduces additional predicate *U-shaped* via perceptual linking:

$$
s_{t+2} = \left[\begin{array}{ll}
x & : \text{Ind} \\
c_{hs} & : \text{house}(x) \\
\text{cvm=dp} & : \textit{Vec} \\
c_{shape} & : \text{shape}(x, \text{cvm}) \\
\text{dp} & = \left[\begin{array}{l}
\text{pt} : \left[\begin{array}{l} \mathbf{u} \perp \mathbf{v} \perp \mathbf{w} \\ \mathbf{u}(0) \neq \mathbf{w}(1) \end{array}\right] \\
\text{sh} : \left\{\text{rectangular, open}\right\}
\end{array}\right] : \textit{Vec} \\
c_{u} & : \text{U-shaped}(x)
\end{array}\right]
$$

- $\approx$ 'U-shaped house'

Multimodal chart parser

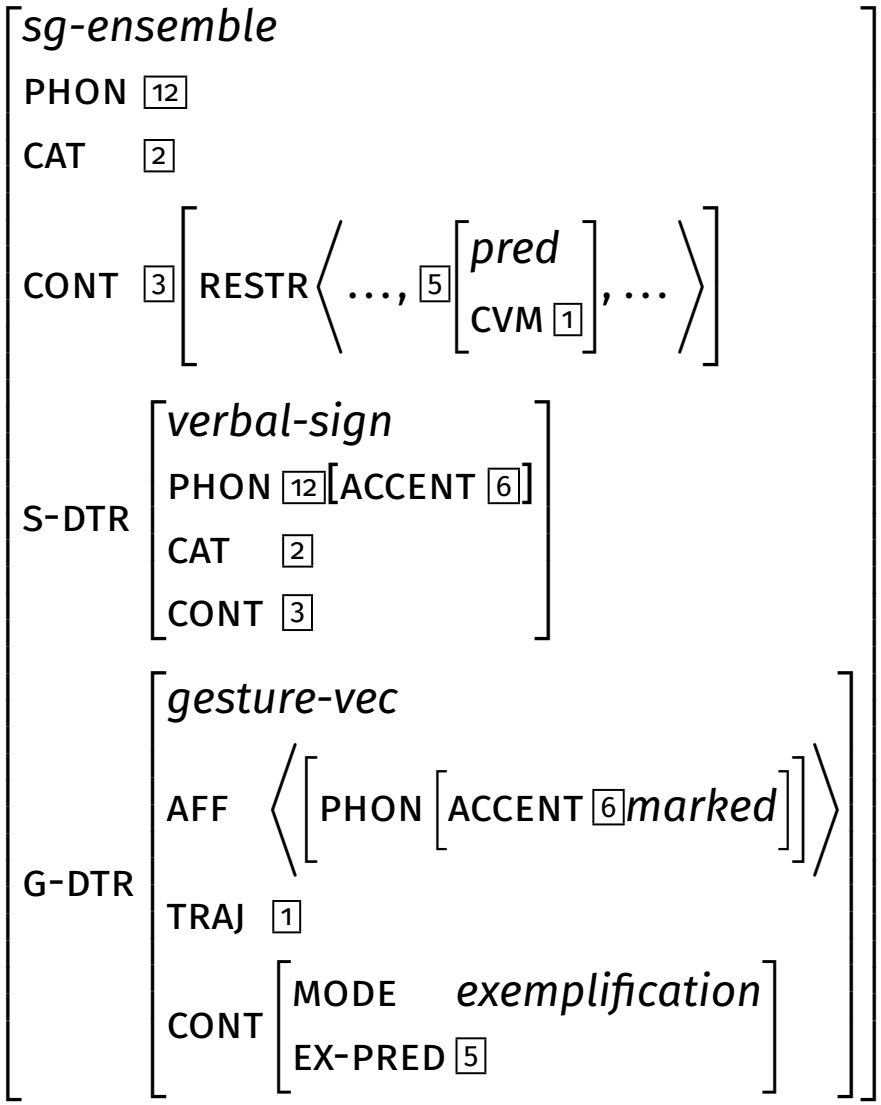


Possible **multicharts**, licensed by tier-crossing grammar rules (Johnston, 1998):

- $\{(s, 0, 1), (g, 3, 4)\}$,
- $\{(s, 1, 2), (g, 3, 4)\}$,
- $\{(s, 0, 2), (g, 3, 4)\}$

MM integration scheme in grammar I

- ‘Ensembles’ (Lücking, 2013)



MM integration scheme in grammar II

- other approaches:
  - assigning underspecified semantic descriptions to gesture morphology (instead of perceptual processing) (Alahverdzhieva, Lascarides and Flickinger, 2017)
  - speech and gestures as mutually communicating channels (instead of grammar) (Rieser and Lawler, 2020)
- various approaches needed since ensembles not appropriate for any kind of gesture ➔ head shake

Head shake and 'No' I

- Claim: headshake is a form variant of verbal 'No'.
- Initial support:

(1) a. A: (1) Do you want some coffee? / (2) You don't want some coffee?

b. B:  (= head shake)

- The crucial observation here is that depending on whether A produced a negative or a positive propositional kernel in the question, B's head shake is either a denial of the positive proposition (1) or a confirmation of the negative one (2). That is, a head shake behaves like q/a 'No'.

Head shake and 'No' II

- This is one of the meanings of *no* discussed by Tian and Ginzburg (2016) as '"No" with explicit antecedent', a simplified lexical entry for which is given as in (2):

(2)

$$\begin{bmatrix} e_{head} : \text{no} \; / \; \text{  } \\ \text{dgb-params} : \begin{bmatrix} \text{spkr} : \textit{Ind} \\ \text{addr}: \textit{Ind} \\ \text{u-time}: \textit{Time} \\ \text{c1} : \text{addr(spkr,addr,u-time)} \\ p : \textit{Prop} \\ \text{MaxQUD} = p? : \text{PolarQuestion} \end{bmatrix} \\ \text{content} = \text{Assert(spkr,addr, u-time, NoSem}(p)) : \textit{IllocProp} \end{bmatrix}$$

NoSem

- NoSem negates p if p is a positive proposition, and confirms p if p is a negative proposition:

$$\text{NoSem}(p) = \begin{cases} \neg p & \text{if } p : \mathit{PosProp} \\ p & \text{if } p : \mathit{NegProp} \end{cases}$$

- Note that the result of 'NoSem(p)' is always of type *NegProp* (if p : *NegProp* then $p = \neg q$, which remains unchanged).

Excursus: negative propositions

- We need/want to distinguish positive and negative propositions.
- But what could a negative proposition be?

Propositions as possible worlds?

- if a proposition p is modeled as a set of possible worlds, then $\neg p$ is its complement set
- but: how to distinguish between positive and negative sets of worlds?

Propositions

Following Austin (1950) and Barwise and Etchemendy (1987), propositions in KoS are individuated in terms of a situation and a situation type:

- $Prop := \begin{bmatrix} \text{sit} & : Rec \\ \text{sit-type} & : RecType \end{bmatrix}$
- A proposition $p = \begin{bmatrix} \text{sit} & = s_0 \\ \text{sit-type} & = ST_0 \end{bmatrix}$ is true iff $s_0 : ST_0$

  (If we had negative types, negative propositions could be defined right away)

Type-theoretical negation

- Negative types: if T is a type, then $\neg T$ is a type
- $a : \neg T$ iff there is some T' such that $a : T'$ and T' *precludes* T
- T' *precludes* T iff:
  - $T = \neg T'$, or
  - T and T' are non-negative and there is no a such that $a : T$ and $a : T'$ ($[^{\vee} T]$ and $[^{\vee} T']$ have no overlap)

➔ Type-theoretical negation captures non-realization of a situation (*via* preclusion) and provides negative types ($\neg T$, licensed by some negative particle in speech: *no*, *n't*, *not*, ...)

Type-theoretical negation I

- T and $\neg\neg T$ are equivalent, but the former is a positive, the latter a negative type
- some object a need not be of type T, and there need not be a type T' that precludes T; in other words: $a : T \vee \neg T$ is not a tautology.
- If I observe Jo cutting onions, the situation I observe neither tells me if Johnson is smoking a cigar, nor that he is not smoking a cigar.
- Hence, $s_{visual} : \mathit{Cutting}(a, o)$, $s_{visual} \not: \mathit{CigarSmoke}(\mathit{johnson})$, hence: it is not the case that $s_{visual} : \mathit{CigarSmoke}(\mathit{johnson})$, but neither is it the case that $s_{visual} : \neg\mathit{CigarSmoke}(\mathit{johnson})$

Type-theoretical negation II

- negative proposition (finally!):
  $$\mathit{NegProp} :\!= \begin{bmatrix} \text{sit} & : \mathit{Rec} \\ \text{sit-type} & : \mathit{RecType}^{\neg} \end{bmatrix}$$

Head shake and 'No' I

- The other uses of "No" discussed by Tian and Ginzburg (2016) are called '"No" with exophoric antecedent' (3) and '"No" with implicit antecedent' (4).

(3) a. *(A child is about to touch a socket)* Adult: No!

b. *(A discovers smashed beer bottle in freezer)* A: No!
(Both uses of 'No' indicate that the speaker does not want a certain situation type to happen or to be realized)

(4) a. A: How's your girlfriend?

b. B: She is no longer my girlfriend.

c. A: Ah, I'm sorry.

d. B: No, she is my wife now.

Head shake and 'No' II

- The occurrences of *No* in (3) and (4) can be replaced by the head shake  without a change in meaning. (Though speaking requires auditory, shaking visual attention.)
- Hence, there is evidence that the head shake and the particle 'No' are both form variants of the same lexical resources (this in cultures where the head shake is associated with negation and not with affirmation, as it is in Bulgaria and, with some modifications, Greece, Turkey, and Southern Italy) (Jakobson, 1972).

Simultaneous head shake

- Simultaneous head shake can be used by a speaker to emphasize negative utterances, as in a famous speech given by Bill Clinton in (5) [6:29].
- Note that three chunks of head shake gestures are produced, one for each of the negated verbal sub-utterances (*never … not … never*).

→ Repetition seems to be used as a temporal means of aligning head movements and the scope of negation, as observed in manual gesture (Harrison, 2010)

(5) *I never told anybody to lie (.) not a single time (..) never*

[repeated ] (.) [repeated ] (..) []

Simultaneous head shake

- Simultaneous head shake seems to presuppose a negative particle in speech:

(6) a. I don't believe you.

b. ?I believe you.

- (6a) provides a negative proposition, ¬believe(A,B), which by *NoSem* the headshake affirms.
- (6b) provides a positive proposition, believe(A,B), which by *NoSem* the headshake negates, hence a contradiction ensues.

Head shake and dissociated contents

- However, the contradiction can be ameliorated:

(7) *(Context: Claims that B stole 500 Euro)*

a. B: They say I stole the money. But I didn't.

b. A: I believe you.

- One can understand A's headshake as
  1. affirming the negative proposition B makes, **or**
  2. expressing that A is upset about 'their' accusation.
- In either case, this requires us to assume that the head shake can be **disassociated from speech** that is simultaneous with it.

Head shake and dissociated contents

- Dissociated gesture and speech is an assumption argued for in some detail with respect to speech laughter. (Mazzocconi, Tian and Ginzburg, 2020)
- Such observations are of great importance for a multimodal theory.
- This is because it has been claimed that multi-tier interpretation is guided by the heuristic 'if multiple signs occur simultaneously, take them as one'. (Enfield, 2009, p. 9)
- The semantic and pragmatic synchrony rules stated by McNeill (1992) are even more explicit: '[…] speech and gesture, present the same meanings at the same time', p. 27; '[…] if gestures and speech co-occur they perform the same pragmatic functions', p. 29

Back to example

(8) *(Context: Claims that B stole 500 Euro)*

a. B: They say I stole the money. But I didn't.

b. A: I believe you.

- One can understand A's headshake as expressing that A is upset about 'their' accusation—*can't believe it / can't get one's head around it*: CBI

Appraisal

- incorporate the effect of pos/neg valenced signals (laughs, smiles, frowns, sighs) on an interlocutor's public face in the DGB in terms of the *Mood* field. (Ginzburg, Mazzocconi and Tian, 2020b)
- two-dimensional Component Process Model (Scherer, 2009; Russell, 2003): Pleasantness and Power:

$$\mathit{Appraisal} := \begin{bmatrix} \text{pleasant} & : \begin{bmatrix} \text{pred} = \text{pleasant} : \mathit{EmotivePred} \\ \text{affect} : \begin{bmatrix} \text{pve} : \mathbb{N} \\ \text{nve} : \mathbb{N} \end{bmatrix} \end{bmatrix} \\ \text{responsible} : \mathit{Ind} \vee \mathit{RecType} \\ \text{power} & : \begin{bmatrix} \text{pred} = \text{powerful} : \mathit{EmotivePred} \\ \text{control} : \mathbb{N} \end{bmatrix} \end{bmatrix}$$

Changing mood

- δ: increment, ϵ: weight (diff between new and existing appraisal)
- PositivePleasantnessIncr(δ, ϵ) $=_{def}$

$$\left[\begin{array}{l}\text{pre: } \left[\text{LatestMove.cont : IllocProp}\right] \\ \text{effect : } \left[\begin{array}{l}\text{Mood.pleasant.arousal.pve =} \\ \quad \epsilon(\text{preconds.Mood.pleasant.arousal.pve}) + (1-\epsilon)\delta : \mathit{Real} \\ \text{Mood.pleasant.arousal.nve =} \\ \quad \epsilon(\text{preconds.Mood.pleasant.arousal.nve}) : \mathit{Real}\end{array}\right]\end{array}\right]$$

- NegativePleasantnessIncr(δ, ϵ) $=_{def}$

$$\left[\begin{array}{l}\text{pre: } \left[\text{LatestMove.cont : IllocProp}\right] \\ \text{effect : } \left[\begin{array}{l}\text{Mood.pleasant.arousal.pve =} \\ \quad \epsilon(\text{preconds.Mood.pleasant.arousal.nve}) + (1-\epsilon)\delta : \mathit{Real} \\ \text{Mood.pleasant.arousal.nve =} \\ \quad \epsilon(\text{preconds.Mood.pleasant.arousal.pve}) : \mathit{Real}\end{array}\right]\end{array}\right]$$

Back to example

'I believe you.'

- $$\begin{bmatrix} \text{form}: \text{  } \\ \text{dgb-params}: \begin{bmatrix} \text{spkr} : \textit{Ind} \\ \text{sito} : \textit{Rec} \\ \delta : \textit{Int} \text{ (negative)} \\ \text{c2} : \text{Arousal}(\delta, \textit{form}) \\ \text{Q} : \textit{Type} \text{ (= what they did)} \\ \text{p} = \begin{bmatrix} \text{sit} = \text{so} \\ \text{sit-type} = \text{Q} \end{bmatrix} : \textit{Prop} \end{bmatrix} \\ \text{cont} = \text{CBI}(\text{spkr},\text{p},\delta) : \textit{Prop} \end{bmatrix}$$
- NegativePleasantnessIncr:

$$\begin{bmatrix} \text{pre}: \left[\text{LatestMove.cont} = \text{Assert}(\text{spkr}, \text{CBI}(\textit{spkr}, p, \delta)) : \textit{IllocProp}\right] \\ \text{effect} : \left[\text{NegativePleasantnessIncr}(\delta, \epsilon)\right] \end{bmatrix}$$

Throat whistling

Contestant throat whistles while playing the guitar in talent show)

Judge: You're such a talent. Incredible + 

(Simplified from `https://languagelog.ldc.upenn.edu/nll/?p=50436`)

[show on YT, images clickable]

- The head shake expresses amazement concerning the artistic achievement.

→ It expresses positive appraisal.

Noetic head shake

- ‘Negation of situation’ is expressed via negative *Mood* ≈ *don’t want a situation to be realized*
- CBI triggered by both positive and negative mood

→ Head shake as a noetic signal: an expressive phenomenon (mood, emotion) that influences thinking and knowing (semantics) [inspired by William James]

→ common pattern underlying multimodal communication (?)

- further evidence: laughter

The meaning of laughter I

- Laughter has meaning akin to what words and phrases possess. (Ginzburg, Mazzocconi and Tian, 2020a)
- It involves reference to external real world events, quite analogously to event anaphors (Plessner, 1970).
- It has stand alone meanings:

(9) a. (Context: Bayern München goalkeeper Manuel Neuer faces the press after his team's (*Dreierkette*) defense has proved highly problematic in the game just played (3-2 against Paderborn).)

b. Journalist: (smile): Dreierkette auch 'ne Option? (*Is the three-in-the-back also an option?*) Manuel Neuer: fuh fuh fuh (brief laugh) $\rightsquigarrow$ The three-in-the-back is not an option!

The meaning of laughter II

- Laughter participates in semantic and pragmatic processes like scare quotation, repair, implicature, and irony:
- (10a-c) exemplifies intra-utterance laughter, where the laughter has the effect of scare-quoting ((Predelli, 2003) the sub-utterance it precedes.

(10) a. A : well I I'm interested in it in a **(. laughs) ((comfortably)) re:laxed** way, you know, I mean I . I do keep, I have kept up with it (London Lund Corpus)

b. (i) A: Jill is John's, (laugh) long-term friend. (ii) A: She is John's long-term (laugh) friend.

c. (i) A: Jill is John's, (wink) long-term friend. (ii) A: She is John's long-term (wink) friend.

The meaning of laughter III

- Two basic meanings for laughter (cf Kundera's devilish and angelic laughter in *The Book of Laughter and Forgetting*):

(11) a. Pleasant($p, \delta,$ *spkr*) given: a context that supplies a laughable *p* and speaker *spkr*, content: the laughable is pleasant for the speaker to a contextually given degree δ.

b. Incongr(p,δ,τ) given: a context that supplies a laughable *p* and topos τ, content: the proposition that *p* is incongruous relative to τ (to extent δ).

The meaning of laughter IV

c. Conversational rule (inspired by (Morreall, 1983)):

   Positive affect incrementation of Mood (the speaker's public emotion display): given the LatestMove being an incongruity proposition by the speaker, the speaker increments the (positive) pleasantness recorded in Mood to an extent determined by the laughter's arousal value.

The meaning of laughter V

- From pleasantness, we can derive three functions of laughter: **affiliation**, **empathetic acknowledgement**, and **superiority**.
- Affiliative laughter arises by resolving the laughable as the state where the speaker and addressee are *co-present*.
- We abbreviate the laughable

$$\begin{bmatrix} \text{sit} = l \\ \text{sit-type} = \begin{bmatrix} \text{A:Ind} \\ \text{B:Ind} \\ \text{t: TIME} \\ \text{c1:addressing(A,B,t)} \\ \text{c2: CoPresence}(\{\text{A,B}\},\text{t}) \end{bmatrix} \end{bmatrix}$$ as CoPresence(A,B).

The meaning of laughter VI

- Affiliation then involves the following sequence:
  1. A laughs at B; content: Pleasant(A,δ,CoPresence(A,B)) bringing about an update: A's Mood.pleasant.arousal is positively incremented by δ.
  2. This can give rise to a similar Mood update for B, signalled by laughter at A with content Pleasant(B,δ',CoPresence(B,A)).
- (Common in parent–child interaction)
- This does not rule out the possibility one would like to distinguish the two "functions" (expressing pleasure and affiliation) if there were systematic reasons for so doing—say, a laugh/smile incontrovertibly dedicated to the latter function and positing a "precompiled" lexical entry therefor (cf Ekman (1992) and Wood and Niedenthal (2018).
- Nonetheless, absent such a demonstration, we need not assume affiliation *requires* a distinct laughter.

The meaning of laughter VII

- Empathetic laughter: Empathetic acknowledgement of A's utterance by B laughing requires the defeasible assumption (more on this soon) *If it's pleasant for me that you said that p, then I agree that p*—A's utterance is the event pleasant for B.
- Superiority/mocking laughter: A observes an event *e* which affects B negatively. Laughter can then be taken to reflect A's appraisal of *e* as pleasant. If, in addition, A has control over the event, the added element of superiority or even sadism can emerge.

The meaning of laughter VIII

- Building on work in humour theory[1], we explicate incongruity as a notion that relates a contextually salient entity l with a defeasible rule (a *topos*[2]) in case there exists a contextually salient characterization of l that is incompatible with τ.
- The topos is not explicitly introduced into the context; the most plausible assumption is to assume it requires access from Long Term Memory.

[1] Raskin, 1985.
[2] Breitholtz and Cooper, 2011.

A lexical entry for pleasant laughter I

- We can now formulate a lexical entry for pleasant laughter, as in (1a): the content we posit is that the laughable is pleasant for the speaker to a contextually given degree δ.
- The *effect* of such laughter on the speaker is captured in terms of an update rule that increments the (positive) pleasantness recorded in Mood to an extent given by the weight ϵ, as described earlier.

A lexical entry for pleasant laughter II

$$\left[\begin{array}{l}\text{phon} : \texttt{laughterphontype} \\ \text{dgb-params} : \left[\begin{array}{l}\text{spkr} : \text{Ind} \\ \text{addr} : \text{Ind} \\ \text{t} : \text{TIME} \\ \text{c1} : \text{addressing(spkr,addr,t)} \\ \delta : \text{Int} \\ \text{c2} : \text{Arousal}(\delta, \textit{phon}) \\ \text{s} : \text{Rec} \\ \text{p} = \left[\begin{array}{l}\text{sit} = \text{l} \\ \text{sit-type} = \text{L}\end{array}\right] : \text{prop}\end{array}\right] \\ \text{content} = \left[\begin{array}{l}\text{sit} = \text{s} \\ \text{sit-type} = \left[\text{c4: Pleasant}(p, \delta, \textit{spkr})\right]\end{array}\right] : \text{Prop}\end{array}\right]$$

$$\left[\begin{array}{l}\text{preconditions:} \left[\begin{array}{l}\text{LatestMove.cont} = \\ \text{Assert(spkr, Pleasant}(p, \delta, \textit{spkr})) : \text{IllocProp}\end{array}\right] \\ \text{effect} : \left[\text{PositivePleasantnessIncr}(\delta, \epsilon)\right]\end{array}\right]$$

References I

- Alahverdzhieva, Katya, Alex Lascarides and Dan Flickinger (2017). 'Aligning speech and co-speech gesture in a constraint-based grammar'. In: *Journal of Language Modelling* 5.3, pp. 421–464.
- Austin, John L. (1950). 'Truth'. In: *Proceedings of the Aristotelian Society. Supplementary*. Vol. xxiv. Reprinted in John L. Austin: *Philosophical Papers*. 2. ed. Oxford: Clarendon Press, 1970., pp. 111–128.
- Barwise, Jon and John Etchemendy (1987). *The Liar: An Essay on Truth and Circularity*. Oxford: Oxford University Press.
- Breitholtz, Ellen and Robin Cooper (2011). 'Enthymemes as Rhetorical Resources'. In: *SemDial 2011 (Los Angelogue): Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue*. Ed. by Ron Artstein et al.
- Ekman, Paul (1992). *Facial expressions of emotion: New findings, new questions*.

References II

- Enfield, Nick J. (2009). *The Anatomy of Meaning: Speech, Gesture, and Composite Utterances*. Language, Culture and Cognition 13. Cambridge, UK: Cambridge University Press.
- Fernando, Tim (2011). 'Constructing Situations and Time'. In: *Journal of Philosophical Logic* 40.3, pp. 371–396. DOI: 10.1007/s10992-010-9155-1.
- Ginzburg, Jonathan, Chiara Mazzocconi and Ye Tian (2020a). 'Laughter as Language'. In: *Glossa* 5.1, 104. DOI: 10.5334/gjgl.1152.
- — (2020b). 'Laughter as language'. In: *Glossa: a journal of general linguistics* 5.1.
- Harrison, Simon (2010). 'Evidence for node and scope of negation in coverbal gesture'. In: *Gesture* 10.1, pp. 29–51. DOI: 10.1075/gest.10.1.03har.
- Jakobson, Roman (1972). 'Motor signs for 'yes' and 'no''. In: *Language in Society*, pp. 91–96.

References III

- Johansson, Gunnar (1973). 'Visual Perception of Biological Motion and a Model for its Analysis'. In: *Perception & Psychophysics* 14.2, pp. 201–211. DOI: doi.org/10.3758/BF03212378.
- Johnston, Michael (1998). 'Unification-based Multimodal Parsing'. In: *Proceedings of the 36th Annual Meeting on Association for Computational Linguistics – Volume I*. Montreal, Quebec, Canada, pp. 624–630.
- Lücking, Andy (2013). *Ikonische Gesten. Grundzüge einer linguistischen Theorie*. Zugl. Diss. Univ. Bielefeld (2011). Berlin and Boston: De Gruyter.

References IV

- Lücking, Andy et al. (2010). 'The Bielefeld Speech and Gesture Alignment Corpus (SaGA)'. In: *Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality*. LREC 2010. 7th International Conference for Language Resources and Evaluation. Malta, pp. 92–98. DOI: 10.13140/2.1.4216.1922.
- Mazzocconi, Chiara, Ye Tian and Jonathan Ginzburg (2020). 'What's your laughter doing there? A taxonomy of the pragmatic functions of laughter'. In: *IEEE Transactions on Affective Computing*. DOI: 10.1109/TAFFC.2020.2994533.
- McNeill, David (1992). *Hand and Mind – What Gestures Reveal about Thought*. Chicago: Chicago University Press.
- Morreall, John (1983). *Taking laughter seriously*. Suny Press.
- Plessner, Helmuth (1970). *Laughing and crying: a study of the limits of human behavior*. Northwestern University Press.

References V

- Predelli, Stefano (2003). 'Scare Quotes and Their Relation to Other Semantic Issues'. In: *Linguistics and Philosophy* 26, pp. 1–28.
- Raskin, V. (1985). *Semantic mechanisms of humor*. Vol. 24. Springer.
- Rieser, Hannes and Insa Lawler (2020). 'Multi-modal meaning – An empirically-founded process algebra approach'. In: *Semantics and Pragmatics* 13.8, n/a. DOI: 10.3765/sp.13.8.
- Russell, James A. (2003). 'Core affect and the psychological construction of emotion'. In: *Psychological Review* 110.1, pp. 145–172. DOI: 10.1037/0033-295X.110.1.145.
- Scherer, Klaus R (2009). 'The dynamic architecture of emotion: Evidence for the component process model'. In: *Cognition and emotion* 23.7, pp. 1307–1351.
- Tian, Ye and Jonathan Ginzburg (2016). 'No I Am: What are you saying "No" to?' In: *Sinn und Bedeutung 21*.

References VI

- Weisgerber, Matthias (2006). 'Decomposing Path Shapes: About an Interplay of Manner of Motion and 'The Path''. In: *Proceedings of Sinn und Bedeutung*. Ed. by Christian Ebert and Cornelia Endriss. SuB 10, pp. 405–419. DOI: 10.18148/sub/2006.v10i2.741.
- Wood, Adrienne and Paula Niedenthal (2018). 'Developing a social functional account of laughter'. In: *Social and Personality Psychology Compass* 12.4, e12383. DOI: 10.1111/spc3.12383.
- Zwarts, Joost (2003). 'Vectors Across Spatial Domains: From Place to Size, Orientation, Shape, and Parts'. In: *Representing Direction in Language and Space*. Explorations in Language and Space 1. Oxford, NY: Oxford University Press. Chap. 3, pp. 39–68.

References VII

Zwarts, Joost (2005). 'Prepositional Aspect and the Algebra of Paths'. In: *Linguistics and Philosophy* 28.6, pp. 739–779. DOI: 10.1007/s10988-005-2466-y.