ESSLLI2025
RUHR-UNIVERSITÄT BOCHUM

Text Technology

ViCom

Spatial Gesture Semantics

1. Introduction: Visuo-Spatial Level of Meaning

Andy Lücking Alexander Henlein

Goethe University Frankfurt

July 28–August 01, 2025

Basic Information

Outline of course

1. Introduction: Visuo-spatial Level of Meaning
2. Spatial Gesture Semantics
3. Extemplification and Informational Evaluation
4. AI and Gesture Detection
5. Frame-based Speech–Gesture Integration

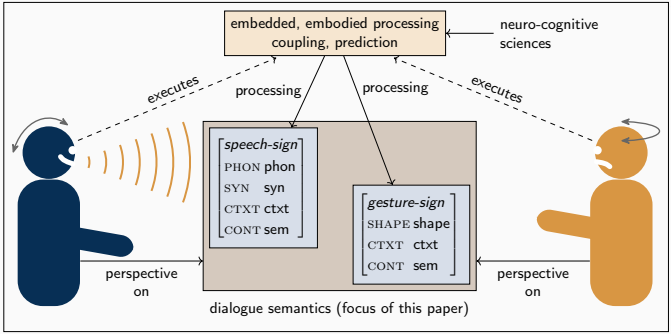
Instructors

Andy Lücking
luecking@em.uni-frankfurt.de

Alexander Henlein
henlein@em.uni-frankfurt.de

Website:https://aluecking.github.io/ESSLLI2025/

Multimodal Dialogue



[1]

Dialogue: research topic *sui generis*

- interaction
- "polyphony"
- alignment
- verbal and non-verbal signals

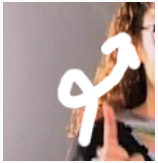
[1] A. Lücking and J. Ginzburg (2023). "Leading voices: Dialogue semantics, cognitive science, and the polyphonic structure of multimodal interaction". In: Language and Cognition 15, 148–172

Overview of nonverbal communication means[2]

- paralinguistic signals / speech phonation
- laughter, sighing / non-speech phonation
- manual gesture (our focus)
- facial expressions
- gaze
- proxemics: spatio-social behavior connected to the interpersonal distances of interlocutors.
- tactile codes
- time behaviour ("chronemics"): length of laughter, respiratory pauses; "kairemics": at the right moment
- Others, hitherto not much studied: clothing, smelling, ...

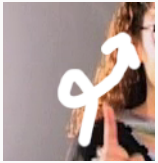
[2] From A. Lücking and T. Pfeiffer (2012). "Framing Multimodal Technical Communication. With Focal Points in Speech-Gesture-Integration and Gaze Recognition". In: Handbook of Technical Communication. Ed. by A. Mehler and L. Romary. In collab. with D. Gibbon, 591–644

Why gesture?

- Gesture can add meaning:
- "Ich g[laube das  sollen TREP]pen sein" /
  I think that should be staircases
  (Capitalization indicates main stress of the first syllable of the noun *Treppen* 'staircases', square brackets indicate the temporal alignment of speech and gesture)
- Obligatory pointing gestures: *I want to have this!*

[3] P. Schlenker (2018). "Gesture projection and cosuppositions". In: Linguistics and Philosophy 41, 295–365, 296

Why gesture?

- Gesture can add meaning:
- "Ich g[laube das  sollen TREP]pen sein" /
  I think that should be staircases
  (Capitalization indicates main stress of the first syllable of the noun *Treppen* 'staircases', square brackets indicate the temporal alignment of speech and gesture)
- Obligatory pointing gestures: *I want to have this!*

- But what exactly is the semantic contribution of a gesture?
- Semantic lacuna: "It should be emphasized that we will not seek to explain how a gesture [...] comes to have the content that it does."[3]

→ truth-conditional visuo-spatial meaning

→ theoretically challenging: informational evaluation (extemplification)

[3] P. Schlenker (2018). "Gesture projection and cosuppositions". In: Linguistics and Philosophy 41, 295–365, 296

Eclecticism[4]

- Gesture studies
- Dynamic semantics
- Grammar theory
- Computer Science
- Model-theoretic semantics
- Philosophy of language
- Cognitive Science / Psychophysics

[4] A. Lücking (11, 2013). Eclectic Semantics for Non-Verbal Signs. Talk at the Conference on Investigating Semantics: Empirical and Philosophical Approaches, Bochum

On gestures and gestures

Ex: Wheel

(1) Stewart Robson on *ESPN FC Extra Time* (first gesture):
You know when they go on that wheel

and throw the dagger would you ever like to see that go wrong?
$\Rightarrow$ the wheel has a circular layout (`https://www.youtube.com/watch?v=CiVS5_HKFY8&t=0h1m18s`)

Ex: Wheel

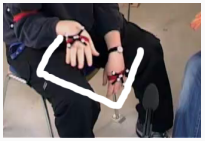
(1) Stewart Robson on *ESPN FC Extra Time* (first gesture):
You know when they go on that wheel

and throw the dagger would you ever like to see that go wrong?
$\Rightarrow$ the wheel has a circular layout
(`https://www.youtube.com/watch?v=CiVS5_HKFY8&t=0h1m18s`)

$\rightsquigarrow$ Hand is drawing a shape

Ex: Like this

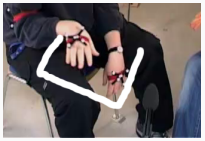
(2) SaGA dialogue V11, starting at 2:32:
dann ist das Haus halt so: / *then the house is*



like this:
⇒ the house has a rectangular, U-shaped layout.

Ex: Like this

(2) SaGA dialogue V11, starting at 2:32:
dann ist das Haus halt so: / *then the house is*



like this:
⇒ the house has a rectangular, U-shaped layout.

⇝ Again, shape drawing

Ex: Right

(3) SaGA dialogue V5, starting at 6:48
und da musst du sofort, scharfer rechter
Winkel,  rechts rein /
and you must enter immediately on the right, at an acute right angle
$\Rightarrow$ the path of the movement runs straight to the right

Ex: Right

(3) SaGA dialogue V5, starting at 6:48
und da musst du sofort, scharfer rechter
Winkel,  rechts rein /
and you must enter immediately on the right, at an acute right angle
$\Rightarrow$ the path of the movement runs straight to the right

$\rightsquigarrow$ pointing/drawing a direction

Ex: Pull out

(4) Keanu Reeves at the *Graham Norton Show*:

a car pulled out in front of me

$\Rightarrow$ the car pulled out in a straight line from the right
(`https://www.youtube.com/watch?v=6VxnceG_eh4&t=0h13m39s`)

$\rightsquigarrow$ Hand draws movement trajectory, from speaker's perspective + represents car

Ex: Tubes

(5) Etienne Gardet:
Was mich wundert is, es gibt ja beim Tennis so Röhren  , ich weiß nich ob ihr die kennt / *What surprises me is that there are such tubes in tennis, I don't know if you know them*
(`https://www.youtube.com/watch?v=lxLPdMU91F8&t=1h23m42s`)

Ex: Tubes

(5) Etienne Gardet:
Was mich wundert is, es gibt ja beim Tennis so Röhren  , ich weiß nich ob ihr die kennt / *What surprises me is that there are such tubes in tennis, I don't know if you know them*
(`https://www.youtube.com/watch?v=lxLPdMU91F8&t=1h23m42s`)

⇝ Hands holding/sculpturing a cylindric volume

Ex: Dagger

(6) Stewart Robson on *ESPN FC Extra Time* (second gesture):
You know when they go on that wheel and throw  *the dagger would you ever like to see that go wrong*?
$\Rightarrow$ throwing exhibits a certain handshape and movement trajectory
(`https://www.youtube.com/watch?v=CiVS5_HKFY8&t=0h1m18s`)

Ex: Dagger

(6) Stewart Robson on *ESPN FC Extra Time* (second gesture):
You know when they go on that wheel and

throw *the dagger would you ever like to see that go wrong*?
$\rightsquigarrow$ Speaker mimes throwing action
$\Rightarrow$ throwing exhibits a certain handshape and movement trajectory
(`https://www.youtube.com/watch?v=CiVS5_HKFY8&t=0h1m18s`)

Ex: Hit

(7) Keanu Reeves at the *Graham Norton Show*:

no one could hit

$\Rightarrow$ fist and straight movement path away from speaker's body

(`https://www.youtube.com/watch?v=6VxnceG_eh4&t=0h1m6s`)

Ex: Hit

(7) Keanu Reeves at the *Graham Norton Show*:

no one could hit

$\Rightarrow$ fist and straight movement path away from speaker's body

(`https://www.youtube.com/watch?v=6VxnceG_eh4&t=0h1m6s`)

$\rightsquigarrow$ Speaker mimes hitting action

Ex: Bowl

(8) SaGA dialogue V21, starting at 4:38:

und 'ne kleinere Schale obendrauf /
and a smaller bowl on top
$\Rightarrow$ layout of hands represents main axis of bowl object

Ex: Bowl

(8) SaGA dialogue V21, starting at 4:38:

und 'ne kleinere Schale obendrauf /
and a smaller bowl on top
$\Rightarrow$ layout of hands represents main axis of bowl object

$\rightsquigarrow$ Hands represent object talked about

Ex: Two towers

(9) SaGA dialogue V24, 6:25:

die rechte Kirche die hat zwei spitze Türme /
the church to the right it has two pointed towers
⇒ Church towers are stretched vertically

Ex: Two towers

(9) SaGA dialogue V24, 6:25:

die rechte Kirche die hat zwei spitze Türme / *the church to the right it has two pointed towers*

⇒ Church towers are stretched vertically

⇝ Each index finger represents a church tower

Ex: Some ...some

(10) Daniel Levitin, *Ted talk*:

some of them are obvious  ,
some of them are not so obvious


$\Rightarrow$ *on the one hand...on the other hand* figure (`https://www.youtube.com/watch?v=8jPQjjsBbIc&t=0h3m55s`)

Ex: Some ...some

(10) Daniel Levitin, *Ted talk*:

some of them are obvious  ,
some of them are not so obvious


$\rightsquigarrow$ Hands engage in discourse structuring

$\Rightarrow$ *on the one hand…on the other hand* figure (`https://www.youtube.com/watch?v=8jPQjjsBbIc&t=0h3m55s`)

Ex: So then

(11) Kathy Griffin:[5]

Anyway um .. so so then *I decided to promote my own shows*
⇒ palm up open hand gesture towards Colbert (host on the right) shows "re-engagement with main discourse" (*op. cit.*, p. 177f)
(`https://www.youtube.com/watch?v=UWKIljRfltM&t=0h7m17s`)

[5] Slightly simplified from S. Laparle (2022). "The shape of discourse: how gesture structures conversation". PhD thesis. University of California, Berkeley, p. 177

Thumbs Up

(12)

Male Hand Giving Thumbs Up Sign by freebie.photography, CC BY 3.0 Unported.

$\Rightarrow$ emblem: *good / like*

Thumbs Up

(12) 

Male Hand Giving Thumbs Up Sign by freebie.photography, CC BY 3.0 Unported.

$\Rightarrow$ emblem: *good / like*

$\rightsquigarrow$ Conventionalized response item

Modes of representation

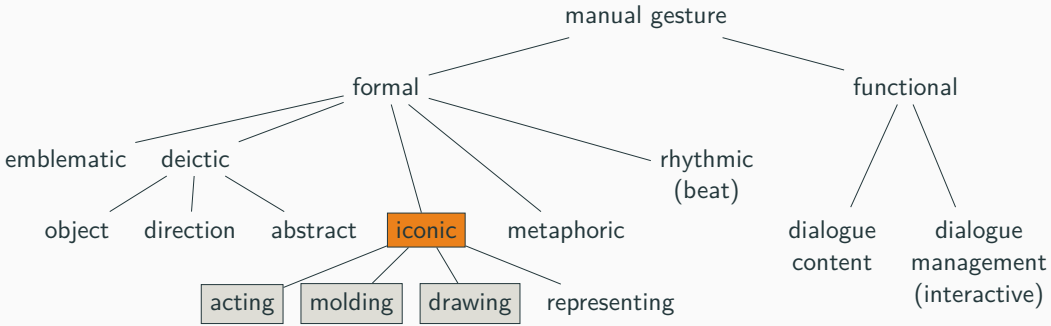
Different iconic motivations of gestures

Acting, molding, drawing, representing

"Müller [...] has distinguished four modes of representation: acting, molding, drawing, and representing. In the acting mode, the hands are used to mime or reenact actual manual activities, such as grasping, holding, giving, receiving, opening a window, turning off a radiator, or pulling an old-fashioned gear shift; in the molding mode, the hands mold or shape a transient sculpture, such as a picture frame or a bowl; in the drawing mode, the hand(s) outline(s) the contour or the form of objects or the path of movements in space; and in the representing mode, the hand embodies an object as a whole, a kind of manual 'sculpture', when, for example, a flat open hand represents a piece of paper and the extended index finger represents the pen used to make notes on that paper." (p. 1691)[6]

[6] C. Müller (2014). "Gestural modes of representation as techniques of depiction". In: Body – Language – Communication. Ed. by C. Müller et al. Vol. 2, 1687–1702

Dimensions of classifying gestures



Kendon's Continuum[7], modified

- To the right: increasing degree of conventionalization, decreasing language dependency
- We are concerned with Gesticulation and recurrent gestures
- Problem: no lexicon (as exist for emblems and sign language)

[7] D. McNeill (1992). Hand and Mind. What Gestures Reveal about Thought. Cambridge UP, 37

Affiliation

- A gesture attaches to a "docking point" in speech, the affiliate[8]
- Hints: Temporal alignment, stressed intonation/information structure, semantic constraints (′ indicates secondary stress)

- Affiliate is a is a lexical item in about 70% of cases[9]

[8] E. A. Schegloff (1984). "On some Gestures' Relation to Talk". In: Structures of Social Action. Studies in Conversational Analysis. Ed. by J. M. Atkinson and J. Heritage, 266–296

[9] A. Mehler and A. Lücking (2012). "Pathways of Alignment between Gesture and Speech: Assessing Information Transmission in Multimodal Ensembles". In: Proc. of the International Workshop on Formal and Computational Approaches to Multimodal Communication under the auspices of ESSLLI 2012, Opole, Poland, 6-10 August

Scope of spatial gesture semantics

- There are various kinds/uses of gestures
- lack of "gestionary" (gesture lexicon) $\rightarrow$ we need a ways to interpret a gesture('s form) "online"
- Gestures integrate with (lexical) affiliate
- There is no presupposition that a spatial semantics captures all of them (alike)
- Stroke is assumed to be the semantically relevant gesture phase

Gesture phases

- By ‘gesture’, we usually refer to the stroke phase
- Preparation often anticipates/precedes the gesture’s affiliate → interesting from a production perspective
- typical alignment: co-speech

Co/Pro/Post/Pre[10]

- Co:
  I think that should be staircases

- Pre:
  I think that should be staircases

- Pro:
  I think that should be

- Post:
  I think that should be staircases

[10] P. Schlenker (2018). "Gesture projection and cosuppositions". In: Linguistics and Philosophy 41, 295–365; T. Slama-Cazacu (1976). "Nonverbal Components in Message Sequence: 'Mixed Syntax'". In: Language and Man. Anthropological Issues. Ed. by W. C. McCormick and S. A. Wurm, 217–227

Co/Pro/Post/Pre

- Co seems to be natural, Pro and Post presumably intentionally produced $\rightarrow$ pragmatic effects
- Conjecture: Pro works the better the more right on Kendon's continuum (conventionalization)
- Pre found with word elicitation difficulties[a]

[a] M. L. Rose (2013). "Releasing the Constraints on Aphasia Therapy: The Positive Impact of Gesture and Multimodality Treatments". In: American Journal of Speech-Language Pathology 22, 227–239

Meaning

- Sometimes gestures add meaning.
- Often, speech and gesture have seemingly identical meanings.
- "Yet they express this meanings in completely different ways." (McNeill 1992:11)[11]

[11] D. McNeill (1992). Hand and Mind. What Gestures Reveal about Thought. Cambridge UP

A visuo-spatial level of meaning

Teasing apart verbal and gestural meaning

[Talking about an entrance door]

zwei .. äh .. Stangen mit nem  Dach drüber halt
(*two .. uh .. poles with a roof over them*)

- Intuitively, the gesture just depicts what is said (*over them*)
- **IF** gestures contribute meaning like words do, we would expect that meaning to be addressable.

Interlude: Parallelism constraint

- Clarification request: categorical and phonological parallelism with their source[12]
  - A: Do you fear him? B: Fear? / #Afraid?
  - A: I phoned him. B: Him? / #He?
  - A: Were you cycling yesterday? B: Cycling? / #Cycled?

[12] J. Ginzburg and R. Cooper (2004). "Clarification, Ellipsis, and the Nature of Contextual Updates in Dialogue". In: Linguistics and Philosophy 27, 297–365

[13] J. Ginzburg and A. Lücking (2021). "Requesting clarifications with speech and gestures". In: Proc. of the 1st Workshop on Multimodal Semantic Representations, 21–31

Interlude: Parallelism constraint

- Clarification request: categorical and phonological parallelism with their source[12]
  - A: Do you fear him? B: Fear? / #Afraid?
  - A: I phoned him. B: Him? / #He?
  - A: Were you cycling yesterday? B: Cycling? / #Cycled?
- Parallelism constraint generalizes multimodally[13]
  - B: You have to move your legs like this [*moves right hand up and down in a wave-like manner*]. A: [*moves right hand up and down in a wave-like manner, raises eye-brows*] (constructed from a kids TV show)
  - A: I hear you're busy ⟨laughter⟩ [= little giggle]. B: ⟨laughter⟩? (= low arousal laughter with rising contour). (attested example)

[12] J. Ginzburg and R. Cooper (2004). "Clarification, Ellipsis, and the Nature of Contextual Updates in Dialogue". In: Linguistics and Philosophy 27, 297–365

[13] J. Ginzburg and A. Lücking (2021). "Requesting clarifications with speech and gestures". In: Proc. of the 1st Workshop on Multimodal Semantic Representations, 21–31

Interlude: Parallelism constraint

- Repair: Repetition of (delayed) reparandum [14]
  - A: I think I'll wear my green dress. Can you bring it to me please?
  - B: OK [leaves to go get dress].
  - Wait, did I say green? / #blue? Sorry, I meant my red dress.

[14] A. Lücking and J. Ginzburg (2022). "How to repair a slip of the tongue?" In: Proc. of SemDial 2022, 35–46

Denial i

(13) [...] with a  roof over them

a. ?No, that's [?] not true. The roof (i) is not ⟨∗⟩ / (ii) actually is ⟨∗⟩

b. ?Wait a minute. The roof (i) is not ⟨∗⟩ / actually is ⟨∗⟩

Denial i

(13) [...] with a  roof over them

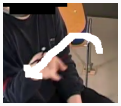
a. ?No, that's [?] not true. The roof (i) is not $\langle * \rangle$ / (ii) actually is $\langle * \rangle$

b. ?Wait a minute. The roof (i) is not $\langle * \rangle$ / actually is $\langle * \rangle$

- What is *that* referring to? It does not seem to be able to pick out the gesture
- What can be said in the places marked with the wildcard "$\langle * \rangle$"? Any verbal answer would violate the parallelism constraint!

→ No gesture meaning accessible

Denial ii

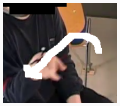
(14) […] with a  roof over them

a. No, that's not true, the roof (i) is not  / (ii) actually is  .

- Multimodal repair is possible: categorical parallelism is preserved.

→ gestures can respond to gestures

Establishing a linguistic interpretation

(15) a. A: […] with a  roof over them

b. B: By  , do you mean that (i) the roof is a flat one / (ii) extends to the right?

c. A: Yes.

d. B: But that's not true. The roof (i) is not flat / (ii) does not extend to the right.

- If interlocutors agree on a linguistic interpretation of a gesture, then this interpretation can finally be questioned.
- "The Linguistic Interpretation of Non-emblematic Gestures Must be agreed in Dialogue"[15]

[15] A. Lücking, A. Mehler, and A. Henlein (2024). "The Linguistic Interpretation of Non-emblematic Gestures Must be agreed in Dialogue: Combining Perceptual Classifiers and Grounding/Clarification Mechanisms". In: Proc. of the 28th Workshop on The Semantics and Pragmatics of Dialogue

Anaphoric potential

(16) [...] two poles with a  roof over them

a. They (= two poles) stand firm.

b. It (= roof) was covered in moss.

c. #It (= *gesture*) extends to the right.

- Noun phrases can be picked out by pronouns
- The gesture does not appear to be anaphorically accessible

Interim conclusion

- The linguistic and visual meanings are not on a par and one cannot simply switch from one to the other.
- Accordingly, it is important to keep the two dimensions apart when trying to formalize the semantic integration of speech and gesture.

Visuo-spatial models

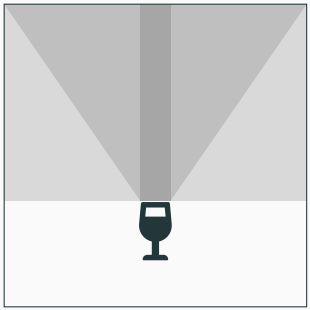
“Visual truth-functions”

- There are at least quite clear semantic intuitions about the circumstances under which a multimodal utterance is true that can guide semantic theorizing.
- Informally, the multimodal PP is true of the modified NP if the extension $[\![\text{NP}]\!]^e$ of the NP in situation e has a roof and the roof “looks like” the gesture.

➔ The only thing left to do: specify what “look like a gesture” means.

[...] two poles with a roof over them

Spatial prepositions

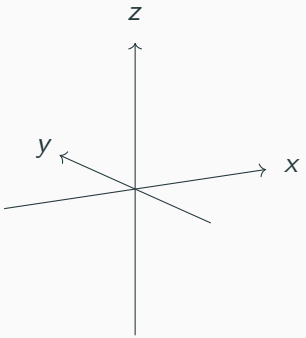


- Three different readings of *above*
- Each is a sub-region within the space surrounding the reference object

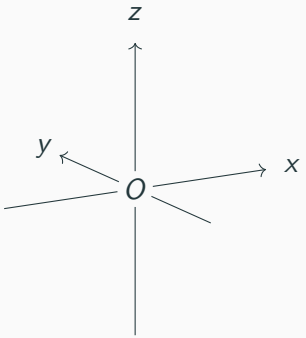
→ (Three-dimensional, Euclidean) Vector Space Semantics[16]

[16] J. Zwarts (1997). "Vectors as Relative Positions: A Compositional Semantics of Modified PPs". In: J. of Semantics 14, 57–86

Euclidean space, idea

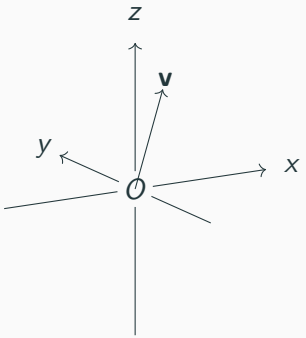


Euclidean space, idea



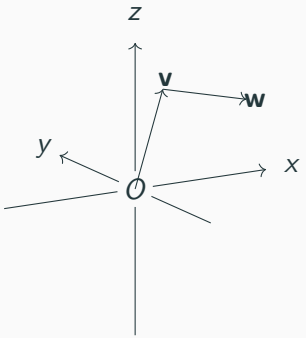
- O: Origin

Euclidean space, idea



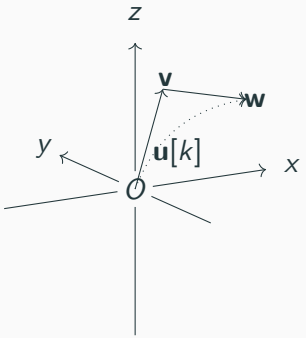
- O: Origin
- $\mathbf{v}$: Vector

Euclidean space, idea



- O: Origin
- **v**: Vector
- **w**: Vector
- **vw**: Sequence of (two) vectors

Euclidean space, idea



- O: Origin
- $\mathbf{v}$: Vector
- $\mathbf{w}$: Vector
- $\mathbf{vw}$: Sequence of (two) vectors
- $\mathbf{u}[k]$: Path: Sequence of k vectors

Model-theory

- Model: $\langle V, D_e, D_s \rangle$, with
  - V: Interpretation function
  - D_e: Domain of entities
  - D_s: Domain of situations/events

Extending the domain of semantic model[17]

- The domain of points: $D_p = V$ (each point is defined by a vector's endpoint)
- The domain of vectors: $D_v = V \times V$ (the Cartesian product of V)

Vector space population: For each element ("point") $w \in D_p$ there is a vector space $V_w \subseteq D_v$ (w is the zero-vector, or origin of V_w).

[17] J. Zwarts and Y. Winter (2000). "Vector Space Semantics: A Model-Theoretic Analysis of Locative Prepositions". In: Journal of Logic, Language, and Information 9, 169–211

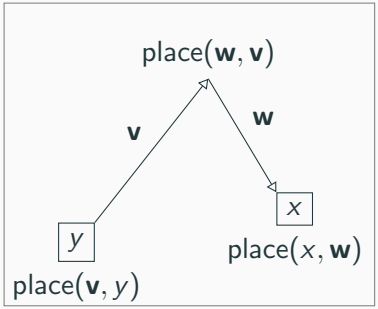
Connecting the extending domain[18]

D_e (entities), D_s (events) and D_v (vector sequences) are related by a couple of functions:

- The vector space located at a concrete object denoted by an NP α is given by 'space($[\![\alpha]\!]^e$)'.
- Place and axis vectors determine spatial relationships:
  - place($x, \mathbf{v}$): x is placed at the end of $\mathbf{v}$; place($\mathbf{v}, x$): the starting point of $\mathbf{v}$ is placed at x; place($\mathbf{u}, \mathbf{v}$): the starting point of $\mathbf{u}$ is placed at the end of $\mathbf{v}$.
  - axis($x, \mathbf{v}$) object x has an axis $\mathbf{v}$.

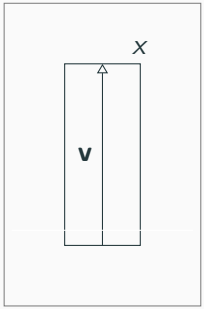
[18] J. Zwarts (2003). "Vectors Across Spatial Domains: From Place to Size, Orientation, Shape, and Parts". In: Representing Direction in Language and Space. Ed. by E. van der Zee and J. Slack, 39–68, frame

Place



The car is in front of the house.
The spoon is near the fork.
...

Axis



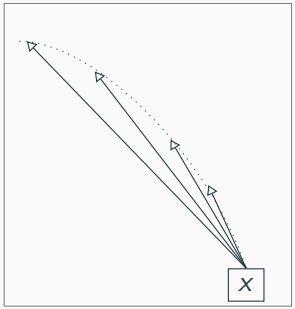
The bar is straight.
The tower is wide.
...
(Very simple model, the use of subsidiary axes would lead to much more detailed object shapes.)

Paths

- Paths: sequences of axis or place vectors. Paths are defined as n-tuples of points, $n > 2$.
- Paths are notated as $\mathbf{v}[k]$.
- We use the letters a and z to denote the first and the last element of the tuple, respectively. The start of a path is indexed as $\mathbf{v}[a]$, its end point as $\mathbf{v}[z]$.
- Paths can be constructed for axis and for place vectors, giving rise to axis-paths and place-paths, respectively.
- Note that paths are non-temporal entities. They receive a temporal interpretation only if index k is mapped to points or intervals in time.

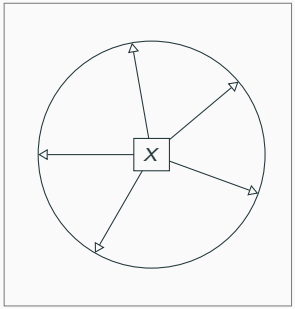
Note: Paths can be implemented in terms of a vector plus one or more transformation matrices.

Place path



The road bends to the left.
She travels along the river bend.
...

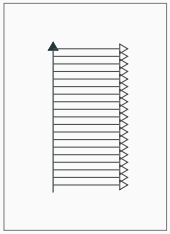
Axis path



A round disk.
A cylindric cup.
...

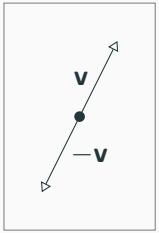
Level

Vectors pointing in the same direction constitute a level.



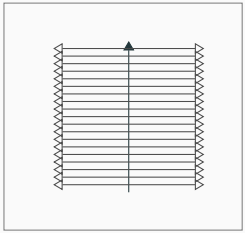
Invers

The inverse $-\mathbf{v}$ of a vector $\mathbf{v}$ points in the opposite direction of $\mathbf{v}$.



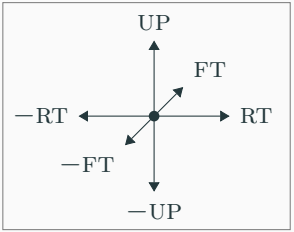
Plane

The set of vectors varying only in the direction of one of their three dimensions (i.e., vectors + their inverses) make up a plane.



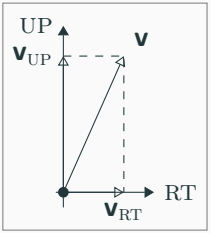
Oriented Vector Space

Each vector space V provides three, mutually perpendicular orienting levels. Intuitively, these levels correspond to the directions *up* (UP), *forward* or *front* (FT), and *right* (RT). These levels in addition to their corresponding inverses, give rise to an oriented vector space.

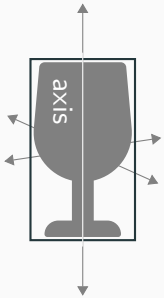


Projections

Given orienting levels A and A', a vector $\mathbf{v}$ can be decomposed into its projections onto the levels, $\mathbf{v}_A$ and $\mathbf{v}'_A$. The figure shows the orthogonal components of $\mathbf{v}$ on the $\textsc{up}$ and the $\textsc{rt}$ levels.

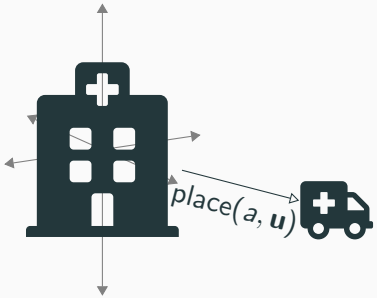


Ex.: Wine glass



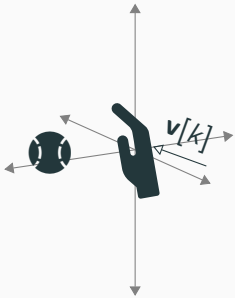
Spatial projection of a wine glass. The glass is located at the center of a vector space, and has a main axis.

Ex.: Near



Spatial projection of two objects being spatially related. The ambulance a is near the hospital if the length of $\mathbf{u}$ is below some threshold τ.

Ex.: Throwing



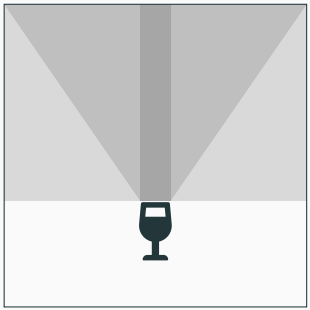
Spatial projection of a throwing event. The movement of the hand "leaves" a place path vector sequence in the domain.

(We will address the issues of handshapes soon.)

Vector spaces in truth-conditional action

- Clear motivation for vectors from modified spatial prepositions such as *3cm above X*.
- But it is reasonable to assume that vectors should extend truth-condition of many more constructions.

Ex.: *above*



- Three readings of *above*(*wineglass*):
- $\{\mathbf{v} \in \text{space}([\![\text{wineglass}]\!]^e) \mid$
  1. $|\mathbf{v_{UP}}| = |\mathbf{v}|\}$ [dark gray]
  2. $|\mathbf{v_{UP}}| > |\mathbf{v_{RT/-RT}}|\}$ [medium gray]
  3. $|\mathbf{v_{UP}}| > 0\}$ [light gray]

$|\mathbf{v}|$: length of vector $\mathbf{v}$.

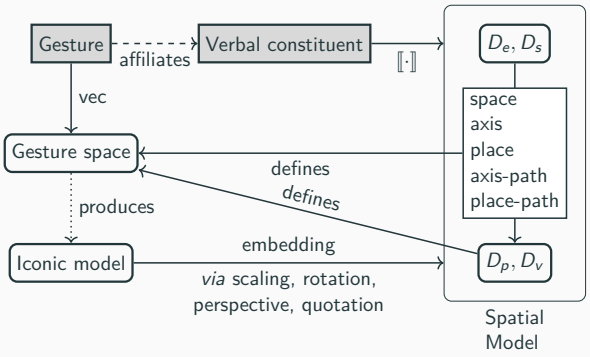
$|\mathbf{v}| = \sqrt{v_1^2 + v_2^2 + v_3^2}$

Vectors are not normalized.

Ex.: Walk

- $\lambda x.\lambda e[\text{walk}(e) \wedge \text{ag}(e) = x \wedge \exists \mathbf{v}[\text{path}(e, \mathbf{v})]]$.
- "Mary walked the shortest route from the university to the capitol." is true if there is an e and an x such that the above holds, and there is no $\mathbf{w}$ which is shorter than $\mathbf{v}$.

Architecture of spatial gesture semantics



Spatial models relate entities (D_e) and events (D_s) onto points (D_p) and vectors (D_v) in terms of vector spaces, axes, places, and paths. The latter are used to define vectorial gesture spaces. Gestures are translated into these gesture spaces, producing formal iconic models that – eventually modified by the operations of scaling, rotation, perspective, or quotation – impose constraints on the spatial domain in which their affiliated verbal expressions are evaluated.

Very Basic Vector Math

Vectors

- A vector $\mathbf{v}$ is a collection of numbers:
  $\mathbf{v} = [x_0, x_1, x_2, \ldots, x_{d-1}]$.
- d is the dimensionality of a vector (note that indexing usually starts with "0").
- Ex.: $\mathbf{v} = [0, 4, 6]$ is a three-dimensional vector.

- Vectors can be "stretched" by multiplying them with a scalar a (a number).
- $a\mathbf{v} = [ax_0, ax_1, ax_2, \ldots, ax_{d-1}]$

- Vectors of the same dimension can be added together by adding their elements ("superposition").
- $\mathbf{v}+\mathbf{w} = [x_0+y_0, x_1+y_1, x_2+y_2, \ldots, x_{d-1}+y_{d-1}]$

Length and Norm

- The length of a vector is the square root of the sum of the squares of its elements.
- $|\mathbf{v}| = \sqrt{x_0^2 + x_1^2 + x_2^2 + \ldots + x_{d-1}^2}$ (also written $||\mathbf{v}||$)

- Normalized or unit vectors are vectors of length 1.
- The unit vector of a vector $\mathbf{v}$ is obtained by dividing $\mathbf{v}$ by $|v|$.
- $\hat{\mathbf{v}} = \frac{\mathbf{v}}{|\mathbf{v}|} = \left[\frac{x_0}{|\mathbf{v}|}, \frac{x_1}{|\mathbf{v}|}, \frac{x_2}{|\mathbf{v}|}, \ldots, \frac{x_{d-1}}{|\mathbf{v}|}\right]$

Linear Transformations

- A linear mapping od a vector is defined in terms of a set of vectors collected in a matrix.
- A 2×2-dimensional matrix is $\mathbf{M} = \begin{bmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{bmatrix}$
- A linear transformation of a two-dimensional vector $\mathbf{v}$ is $\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$
- $\mathbf{Mv} = \begin{bmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} m_{1,1}x_0 + m_{1,2}x_1 \\ m_{2,1}x_0 + m_{2,2}x_1 \end{bmatrix}$

- Ex.: Counterclockwise rotation by $\theta°$.
- $\mathbf{M} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$