

Annotation manual EcoGest: Gestures

Olga Abramov Ulrich Mertens Anne Németh Katharina Rohlfing
Friederike Kern Stefan Kopp

12.09.2021

Contents

1	Rough Annotation Scheme	5
1.1	Annotation hierarchy	5
1.2	Annotation overview	6
1.3	Explanation	6
1.3.1	Sequence	6
1.3.2	Phrase	7
1.3.3	Phase	9
1.3.4	Practice	10
1.4	Gesture Practice	11
1.4.1	Object related practices	11
1.4.2	Action related practices	12
1.4.3	Other practices	14
1.5	Example Annotation	14
2	Semantic Features	17
2.1	Annotation hierarchy	17
2.2	Annotation procedure	17
2.2.1	Entity	18
2.2.2	Action	18
2.2.3	Manner	19
2.2.4	Shape	19
2.2.5	Direction	19
2.2.6	Amount	19
2.2.7	Relative Position	19
2.2.8	Size	20
2.2.9	Property	20
2.2.10	Examples	20
2.2.11	Semantic Feature IDs	21
3	Actionsemantics	24
3.1	Annotation hierarchy	24
3.2	Redundancy	24
3.3	Combined	25
4	Semantic Relations	29
4.1	Reinforces	29
4.2	Integrates	29

4.3	supplements	29
4.4	Complements	30
4.5	Contradicts	30
5	Annotation manual of gesture morphology	32
5.1	Annotation hierarchy	32
5.2	Rules of thumb	33
5.3	Hand Shape	34
5.3.1	Hand-Shape	34
5.3.2	Path of Hand-Shape	37
5.3.3	Hand-Shape Movement Direction	37
5.3.4	Hand-Shape Movement Repetition	37
5.4	Hand Orientation Features	38
5.4.1	Palm Direction	38
5.4.2	Path of Palm Direction	38
5.4.3	Palm Direction Movement Direction	38
5.4.4	Palm Direction Movement Repetition	39
5.4.5	Back of Hand Direction	39
5.4.6	Path of Back of Hand Direction	40
5.4.7	Back of Hand Direction Movement	40
5.4.8	Back of Hand Direction Movement Repetition	40
5.5	Wrist Location	40
5.5.1	Wrist Position	40
5.5.2	Wrist Distance	42
5.5.3	Path of Wrist Location	43
5.5.4	Wrist Location Movement Direction	43
5.5.5	Wrist Movement Repetition	43
5.6	Movement Features	44
5.6.1	Extent	44
5.6.2	Temporal Sequence	44
5.7	Hand Combination Features	44
5.7.1	Two-handed configurations	44
5.7.2	Movement relative to other Hand	45
5.7.3	Two-handed Form	45
	References	47

Rough Annotation Scheme

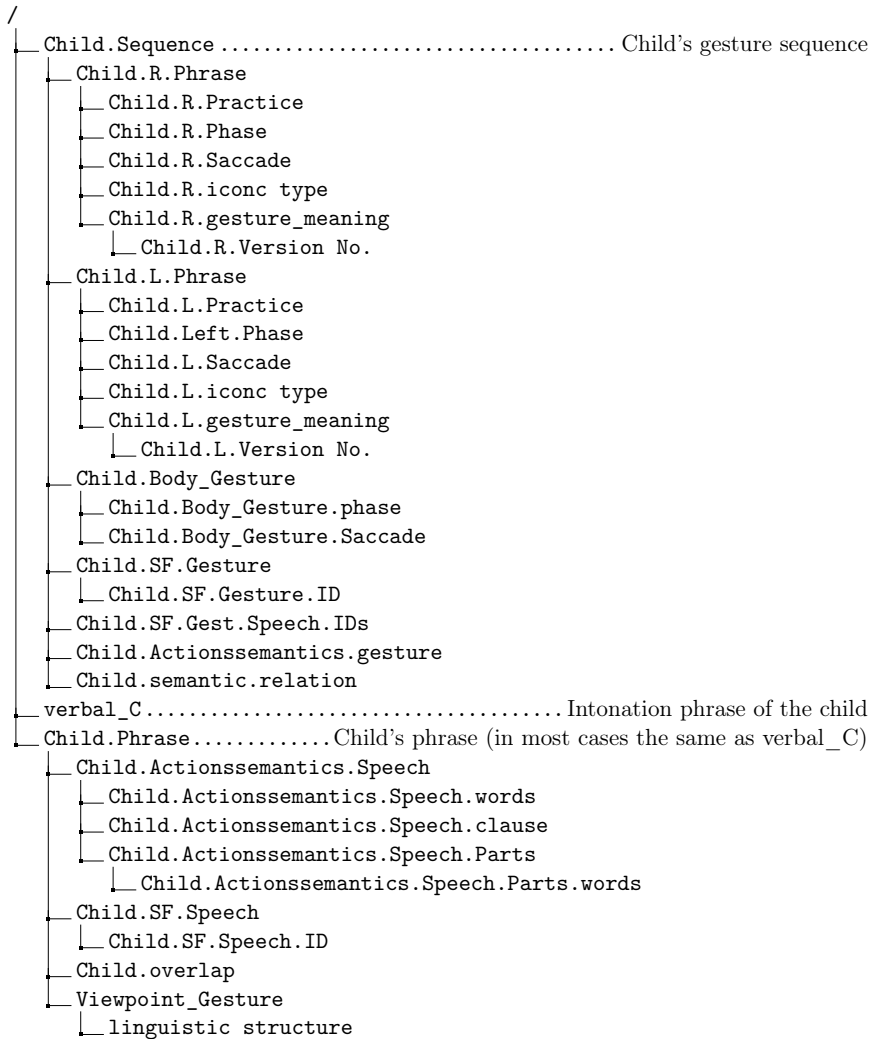
Olga Abramov Ulrich Mertens Anne Németh Katharina Rohlfing
Friederike Kern Stefan Kopp

Version 1 – 16.04.2018

1 Rough Annotation Scheme

1.1 Annotation hierarchy

The following labels are used: Child: *Child*, Mother: *Mother*, R.: *Right* for the right and L.: *Left* for the left hand, SF: *Semantic Features* related to the gesture sequence or speech.



1.2 Annotation overview

Tier	Values	Description
SEQUENCE	–	Gesture unit, which consists of at least one gesture phrase
PHRASE	deictic, iconic, beat, iconic-deictic, iconic-beat, deictic-beat, iconic-deictic-beat, move, discourse, unclear	The gesture phrase's values determine the gesture type; forms which are separated by hyphen are mixed forms. Move is the type of movement which are not expressions of communicative relevance. One gesture phrase consists of gesture phrases.
PHASE	preparation, pre-stroke hold, stroke, post-stroke hold, retraction, unclear	gesture phases, where stroke is the expressive, content bearing part.

1.3 Explanation

The *building blocks* of an annotation are the *annotation elements*. Each annotation element is specified by four parameter:

1. Annotation tier;
2. start point (time);
3. end point (time);
4. content (annotation value).

Those four parameters are to be fixated by the annotator (Because of the dependency between tiers and the specific translation with other annotation tools, the decision of time stamps and content is maybe restricted).

Many values for the annotation elements are chosen from standard literature. In addition to those values, some helpful values are established by Andy Lücking, Kirsten Bergmann and Stefan Kopp. Those values are denoted accordingly. *

1.3.1 Sequence

A gesture sequence is a gesture unit, which consists of multiple gesture phrases. The gesture phrases of a gesture sequence are identified by kinetic cohesiveness, this means that movement, with which the single gestures are performed, are blend into each other. More prosaic: There is no pause between consecutive gesture phrases in terms of a resting posture. This characteristic is the formal identification criteria of a gesture sequence.

*For newly added values exists a higher demand of explanation and justification. The token should highlight this responsibility.

When does a gesture phrase start a new gesture sequence and when does it continue the existing gesture sequence? A sequence is characterised by multiple gesture phrases without a breaking resting posture. Resting posture means:

- (a) to pose in a resting position (resting hands in the lap or on the upper legs, or
- (b) a retraction from a expressive configuration from a previous gesture through relaxation without posing in a resting position.

The pose in a resting position (a) should be easy to identify and there should be no difficulty annotating them. More difficult is case (b). If there is a relaxation of the hand's expressive posture, it counts as relaxation. For example: Fixating the hand configuration, but arms are arranged aside the body; fixating the arm position, but finger configuration is relaxed. These relaxations are called retraction phase ('retraction'), the resting position does not have to be part of every gesture. In those cases criteria (b) is achieved. Therefore, the gesture starts a new sequence. If there is no relaxation, which means the new gesture starts from a Hold, the existing gesture sequence has to be continued. To distinguish relaxation (b) from "physiological unease" the relaxation has to be held for a distinct timeframe. For a rule of thumb this timeframe should be at least one second long.

Because of the person's ambidexterity there might occur the case, that one hand is already in a resting position while the other hand is retracting, therefore stopping the actual sequence while the first hand starts a new sequence. This case results in overlapping sequence elements. This can't be modelled by ELAN). Convention: The overlap will be cut from the 'retraction'.

ELAN-notice: The sequence tier is mandatory. Therefore, a gesture sequence can also consist of exactly one gesture phrase. Elements in the sequence tier don't have an annotation value but mark a specific timeframe.

1.3.2 Phrase

A gesture phrase includes the whole movement, which is used to perform the gesture, therefore - in a prototypical case - from the onset's preparation phase until the renewed posing in a resting position.

For elements in the phrase tier, new annotation values with class names are introduced. The three basic classes are:

1. 'deictic';
2. 'iconic';
3. 'beat'.

Additionally, these two classes

4. 'move' und

5. 'discourse'

are included. Furthermore, these mixed forms are possible:

6. 'deictic-beat';
7. 'iconic-beat';
8. 'iconic-deictic';
9. 'iconic-deictic-beat'.

If an annotation decision can't be put into those gesture classes, there is also the annotation value:

10. 'unclear'.

deictic

The value 'deictic' is assigned for pointing gestures, this means gestures, commonly static, which are identified by - typically - stretching the index finger and which have no other function than pointing at an entity (for example, a concrete object or a position in the gesture space). Especially, deictic gestures have no descriptive function. They point at something but don't illustrate anything by themselves.

iconic

An iconic gesture, annotation value 'iconic', is every non-deictic movement in the gesture space, which describe a narrative item. Iconic gestures are typically dynamic and it is often the case that they resemble the object they illustrate in some way.

An iconic gesture ALWAYS represents a meaningful unit that can be, for example, an action, an object or a manner to perform an action. In order to decide, whether different strokes represent one iconic gesture or are two, three or more iconic gestures, ask: are the strokes part of the same action/object description? Yes? >iconic, No? >iconic, iconic, etc.

beat

Beats, value 'beat', are rhythmic gestures. They always have two phases (back and forth, up and down). They don't represent an item but enhance the spoken language either emphatic or structural. Single beats may emphasize a (prominent) word. Iterative beats may follow and structure the spoken words.

Another usage of beats is overlapping. At that iconic or deictic or iconic-deictic gestures are accompanied with multiple beats, so they are overlain by those beats. Mixed forms, which contain a beat, have at least two strokes, which means, that the stroke is rhythmically repeated. Therefore, all iterative gestures are gestures overlapped by a beat. Beats appeal like a Kleene operator. +

Mixed forms

The three basic classes don't always appear in pure form. For example, a deictic gesture may be overlapped by an iconic gesture (pointing to the table + mimicing the movement of the salt from the salt shaker that was strewed on the table); Every gesture may be overlapped by a single or multiple beats (s. 1.3.2). In those cases, the values should be taken from the upper list of mixed forms.

discourse (AL)

The value 'discourse' is taken for gestures which help organizing or structuring the dialog, typically turn-taking signals. In contrast to the three basic classes the discoure gesture are not used for narration but for interaction. Typical discourse gestures are, for example, a lifted hand to signal the need to talk or a hand stretched towards the dialog partner to tell the partner to stop talking.

emblems

Emblems are conventionalised gestures with a fixed established form and meaning within a community. They can often replace speech. Some emblems differ across cultures, others are globaly used. Example: 'thumb up' sign for 'all right' (Bergmann, 2012).

move (AL)

Some movements are no expressive significant gesture (therefore no real gesture), but an unarticulated, small, twitching, etc. movement. Those movements don't contribute to the narration; Their appearances are explained by speech/gesture production, which means on an psycholinguistic layer as well as with psychological states like nervousness, unease, etc. Those movements are annotated with the value 'move'.

1.3.3 Phase

Arms and hands, which aren't part of a gesture are held in a so called resting position. The exact resting position may vary from speaker to speaker but typically is a relaxed posture in which the arms and hands don't have to be held in that position by force, for example the resting of both arms and hands in the speaker's lap. Active arms and hands are expressing a *gesturephrase*. A gesture phrase can roughly be divided in three *phases*. The emphasizes lays on the part of the gesture movement, which describes or means something. Those phases have to be distinguished from the movement which is made to bring the arms and hands in a position from where the central phase of the gesture can be performed and also the movement which leads arms and hands back to the resting position.

Remark: In a sequence, where gesture phrases follow each other without interruption, the retraction phase isn't always performed completely. This means, that after a stroke or post-stroke hold follows no retraktion phase, instead a preparation phase begins for the stroke of next gesture.

preparation

Movements, which are starting with the hands and arms in the resting position. Those preparation phase, value ‘preparation’, ends with reaching the stroke phase.

stroke

The stroke of a gesture is its central, meaningful part. It is physiologically identified as the part with the highest muscle tone. The stroke may be dynamic or static. The stroke is dynamic when it’s realized by movement. A static stroke is realized by holding hands in a specific configuration without any movement. Both cases of strokes are identified over the whole time span by its annotation value ‘stroke’ (*independent hold* by Kita).

retraction

Both arms and hands are retracted from the stroke area to the resting position. Its annotation value is ‘retraction’.

pre-hold

If in a hold phase the arms and hands are fixated in one position, constituting the next movement, it’s called a pre-stroke hold. The annotation value is ‘pre-hold’.

post-hold

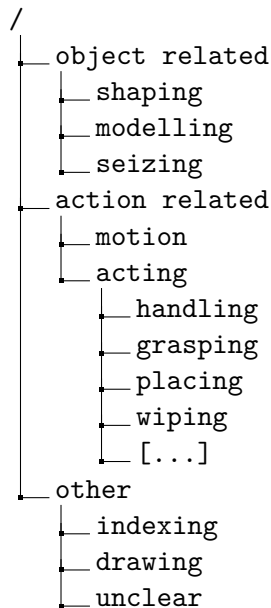
If a hold phase is subsequent to a dynamic stroke and the posture of the hold is constituting to the stroke, it is called a post-stroke hold. The annotation value is ‘post-hold’. Static strokes can’t be followed by hold.

1.3.4 Practice

There are several ways of displaying iconic gestures. Those define methods of depictive use of hands and are annotated in the PRACTICE tier. Those methods are not mutually exclusive, which means there may be overlaps and combinations. For example: movement alongside a path. One hand represents the object (‘modelling’), the path is painted into the gesture space (‘drawing’). The whole gesture is performed by the whole practice of ‘modelling’ and + ‘drawing’. Mixed forms have to be added to the practice-dictionary in ELAN.

Moves, beats and discourse gestures are not of interest for the current annotation - the annotation tier remains empty at this point. The different practices used here are described in detail in the following section.

1.4 Gesture Practice



The gesture practices described below can be roughly divided into three categories. OBJECT RELATED practices are mostly static describing the shape of the object or simply modelling the object. ACTION RELATED practices are dynamic and focus on actions with or without objects where the hand is the subject / object ('acting') or simply depicts a movement ('motion'). If the action is performed with an object, mixed forms like 'acting'+ 'shaping' are annotated. The last category OTHER is related to more general practices applicable either to actions or objects.

1.4.1 Object related practices

Object related practices are mostly static and are used to represent entities and their spatial characteristics.

shaping

The hands form something in the gesture space. One can imagine the 'shaping' method like producing a sculpture in the air. The meaningful form is the shape formed in the gesture space, not the hands themselves. This method of iconic display is typically three dimensional and dynamic. Mostly, objects are formed by this method, therefore the standard type of referent is 'object'.

Helpful question: Is this shape three dimensional?

→ No: drawing or sizing

→ Yes: shaping

sizing (Kirsten & Stefan)

Displaying a size, distance or diameter. Sizing may be performed with one or two hands, for example through posture ("imposition of hands"). A sizing gesture is static. 'Sizing' refers to one axis of a body, therefore to its extension in one dimension. The display of additional attributes of the depictive item is only possible with one other simultaneous practice, for example, 'shaping'.

modelling

The hands are in a static configuration and represent something narrative. The hands themselves represent something.

Helpful question: Hand is object?

→ No: shaping

→ Yes: modelling

1.4.2 Action related practices

Action related practices are dynamic and model a particular action or movement.

motion

Motion mimics any kind of action. Example 1: the hand draws the movement of the car driving along a path way ('drawing'+ 'motion'). Motion can also be represented via body-gestures, beat gestures. In 'motion' the hands display the movement not the object or the actor. It is difficult to decide between acting (without any object) or motion, thus, check first whether 'acting' is applicable, otherwise use 'motion'. Example 2: Dann bin ich aufgestanden + both hands go up ('motion').

Helpful question: Does the hand depict movement without specific object or subject?

→ yes: motion

→ no: acting

acting

An action is imitated, the hand or fingers play the role of the subject of the action. Rule of thumb: the hand is the subject of the performed action. An action can also be performed with an object, e.g., forming a bottle and imitating drinking from the bottle ('acting' + 'shaping'). In this example the hand represents the hand itself performing the drinking action and the handshape mimics the form of the bottle. Another example is when the hand itself is the object: the fist is clenched forming a salt shaker and performs the salting gesture. Here, the hand does not shape the object, the hand is the object (salt shaker). Thus, we annotate ('acting' + 'modelling').

Helpful question: Hand-is-hand? Hand is subject? Hand is object?

→ yes: acting

→ no: motion

Acting can be further specified (if visible in the gesture) into particular types of action (e.g. handling, grasping, placing) which are described below.

placing

An object is placed somewhere in the gesture space. The relative position of an object in gesture space relates to the actual position of this object in the described situation. 'placing' typically looks like placing a virtual object. In contrast to pointing gestures the place is not referred to but rather formed by the hand(s).

Helpful question: Do the hands model the placing of an object (or do they only refer to the position)?

→ no: Indexing

→ yes: Placing

handling

A manual action with an object (Hand-is-Hand). Examples: turning the key, holding a spoon.

1. Helpful question: Does the gesture reflect a manual manipulation of an object?

→ yes: handling

→ no: motion

2. Does the action contain grasping or placing?

→ yes: grasping / placing

→ no: handling

grasping

The hands or one hand grasps an item. The shape of the object is not clearly visible. The gesture displays the grasping action. If the shape of the object is visible too, then annotate ('acting'+ 'shaping').

Helpful question: Is an item only being touched and not completely covered?

→ yes: grasping

→ no: shaping

wiping

The hand mimics the wiping movement, e.g., the hand moves on the table as if wiping back and forth.

1.4.3 Other practices

This category includes practices not initially attributed to either action or objects. They can be applied to both of them.

drawing

Normally, exactly one finger (mostly the index finger) is used as drawing tool. Therefore, one finger does not represent the referent. The finger draws outlines on an imaginary ("into the air") or an actual surface (for example on the other hand).

Helpful question: Is the outline, not the surface of an object being drawn?

→ yes: drawing

→ no: shaping

indexing

Pointing gesture, which refers to a specific position in the gesture space. The position is placeholder for an entity of the narration (compare to *remote indexing* according to Streeck). It is not the hand itself, which refers to an item or position but refers to something somewhere else.

Hilfsfrage: Ist es die Hand selbst, die die Position eines Gegenstands darstellt?

→ Nein: Indexing

→ Ja: Placing

unclear

On every tier it is possible to declare the annotation value as 'unclear'. This value should be used as rare as possible and only if the annotator is positive, after consulting every annotation guideline, that there is no other annotation value possible.

1.5 Example Annotation

Figure 1.5 illustrates an annotation of a gesture sequence. The corresponding verbal utterances are: 1. *gucke, der hat so eine Flasche genommen.* 2. *falsche Seite gucke.* 3. *dann hab ich so.* During the utterances the child produces three iconic gestures. First, forming a shape of a bottle with both hands on the floor and moving the shaped bottle to the mouth. Second, depicting the wrong side of the bottle with both hands. Third, moving the shaped wrong side of the bottle to the mouth while pretending to drink. In the following, we describe the steps of the annotation in 4 steps as shown in Figure 1.5.

1. First, define the gesture sequence (**Child.Sequence**) between two resting positions.
2. Identify the gesture phrases (**Child.L.Phrase**, **Child.R.Phrase**, **Child.Body_Gesture**).

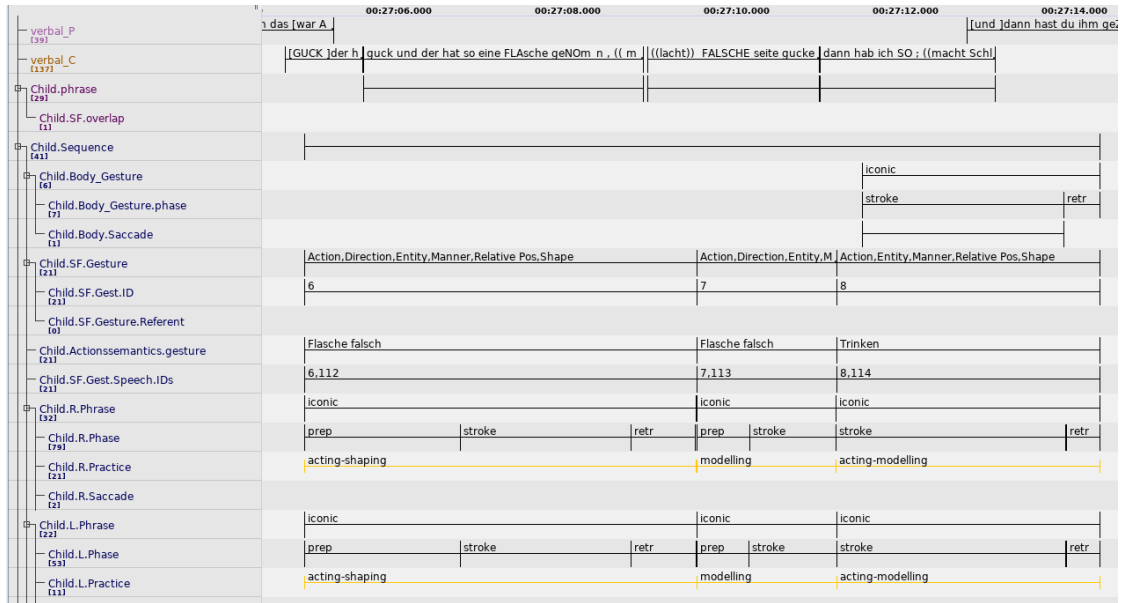


Figure 1.1: Example: annotation of one gesture sequence.

3. select the appropriate gesture phases (prep, stroke, etc.)
4. This step is relevant for 'iconic' gesture phrases as well as for mixed forms containing 'iconic'. Which practice is used to produce the gesture? The gesture practice describes how the meaning of the gesture is presented. Alongside the three utterances, three gestures are produced. In the first gesture we have a 'shaping' gesture and an action (i.e., moving the hands to the mouth). Thus we annotate: 'acting-shaping'. The second gesture represents the wrong side of the bottle, is static, hands are object of the verbal utterance, thus, we annotate: 'modelling'. In the third gesture, the modelled wrong side of the bottle remains, but the movement towards the mouth is added. Consequently, we annotate: 'acting-modelling'.

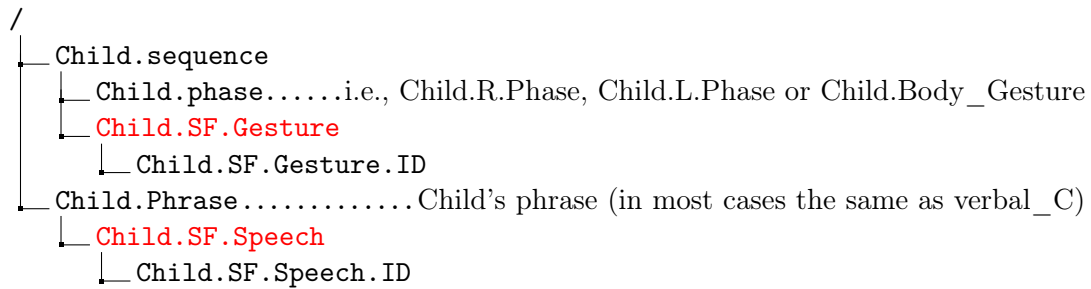
Semantic Features Annotation - Gesture

Olga Abramov Ulrich Mertens Anne Németh Katharian Rohlfing
Friederike Kern Stefan Kopp

Version 1 – 16.04.2018

2 Semantic Features

2.1 Annotation hierarchy



2.2 Annotation procedure

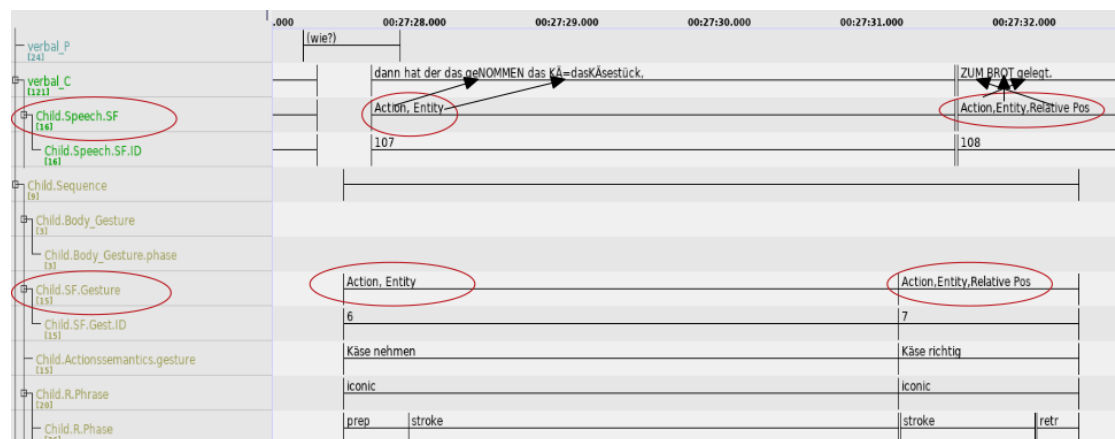


Figure 2.1: Example: annotation of *Child.SF.Speech* and *Child.SF.Gesture*.

Semantic features *Child.SF.Gesture* and *Child.SF.Speech* represent the meaning transmitted via speech or/and gesture. The set of SFs is taken from (Bergmann & Kopp, 2006) and extended to action based dialogues. Children often code different kinds of information in speech vs. gesture. Thus, we annotate the semantics separately for speech and for gesture sequences.

Semantic Features are annotated only with respect to ‘iconic’ (or mixed forms like ‘iconic-beats’, ‘iconic-deictic’ etc.) gestures. The *Child.SF.Gesture* can be applied to

single ‘gesture phases’ or the whole gesture sequence. Each *Child.SF.Gesture* becomes a unique *Child.SF.Gesture.ID* starting from 1 upwards. Semantic features can be assigned to 1-hand, 2-hand configurations or/and the body-gesture.

The semantics of the gesture should be determined semi-separately from speech. That means:

1. watch the gesture and listen to the (intonation) phrase related to the gesture on the time scale
2. watch the gesture again without listening and try to determine what kinds of information (see Table 2.2 for the list of possible meanings) is present in gesture.

Then, the corresponding gesture phase is assigned one or more values from the *Child.SF.Gesture* lexicon described below.

Generally, the gestures are mostly combinations of semantic features listed in Table 2.2. Consider an example: The child shapes with the four fingers and a thumb a half-circle on the floor and utters: "*dann hat er eine Flasche genommen*". Then we ask, what kind of information is present in the gesture? We have a shaping gesture, and an object formed by the gesture. Thus, we annotate > **Shape, Entity**.

There are cases when intonation phrases consist of multiple SFs of the same kind (e.g., ENTITY: [*he put salt in a soup*.] or ACTION: [*he took the salt and put it in the pot*.]). Convention: if more than one SF is present, annotate the second SF of the same kind with a number (i.e., ENTITY, ENTITY₁). Accordingly, if the gesture mimics only one of the multiple objects mentioned in the IP, refer to it with ENTITY or ENTITY₁ respectively. Doing so, we can account for the actual correspondance of meanings in speech and gesture.

2.2.1 Entity

The semantic feature ENTITY is annotated on a gesture phase if there are objects formed by the gesture. Accordingly, any object occurring in the intonation phrase is assigned this semantic feature.

Helpful question: Do the hands represent an object?

→ yes: Entity

Examples: *Löffel, Brot, Käse, Teebeutel* etc.

2.2.2 Action

The semantic feature ACTION represents any kind of action reported by the participant. Actions are characterized by a movement of hands or body or described verbally (he took, he put, etc.). Code ACTION only if the gesture mimics the action and not simply moves.

Examples: "*nehmen*", "*legen*", "*drauf-tun*" etc.

2.2.3 Manner

MANNER characterizes the way an action is performed (*"dann hat er es so gemacht" + gesture*). Children often refer to MANNER verbally (*"gucke, so.."* followed by a gesture). Convention: in such cases annotate MANNER only on the gesture. But sometimes *gucke, so ...* refers only to an ACTION, ENTITY, etc., thus, always judge what the gesture actually expresses.

Verbally, the MANNER can be expressed by adverbs like *richtig rum, ganz*, by the verb itself *reinschmeissen*, or by expressions like *platsch* etc.

Helpful question: Does the gesture show HOW the action is performed? Does the sentences answer the question: how something was done?

→ yes: Manner

Examples: *"richtig drauf"*, *"ganz falsch"*, *"er hat das reingeschmissen"*, gestures

2.2.4 Shape

SHAPE is usually related to an object (ENTITY). Use SHAPE when the hands model the form of the object.

Helpful question: Is the shape of the object visible in the gesture?

→ yes: Shape

Examples: *"rund, zickzackförmig"*, forming a bottle with the hand.

2.2.5 Direction

DIRECTION is used when words like directional information are present in the utterance or a gesture showing the direction of the action. In contrast to RELATIVE POSITION, DIRECTION can be unspecified and does not require two (or more) objects to be present. Note that if directional information among two objects is present, both DIRECTION and RELATIVE POSITION have to be coded.

Examples: *in den Topf, aus, rein*

2.2.6 Amount

This category is only applied if more than one object of the same kind is mentioned. AMOUNT is verbalized by numerals or expressions like 'several', 'multiple', 'many' or by two, three etc. finger counts.

2.2.7 Relative Position

RELATIVE POSITION is annotated when two objects are involved, e.g., gestures placing the objects in relation to each other or verbally: *"Salz in den Mund"*. Note that if directional information is present, both DIRECTION and RELATIVE POSITION are annotated.

Helpful question: Is more than one object modelled?

→ yes: Relative Position

→ no: Direction

Examples: *Salz in den Mund*", *Käse aufs Brot*, *Käse neben das Brot gelegt*, left hand stands for the bread, right hand for cheese.

2.2.8 Size

SIZE is coded when information regarding the size of objects is given.

Examples: [*The stone was very big.*], hands showing the width, height of an object.

2.2.9 Property

This feature is coded when properties of objects other than SIZE, SHAPE are present. Properties are for example related to color (e.g. *green star*, *sparkling stone*). They can also specify a particular feature of the object (e.g., [*the back of the spoon*]).

2.2.10 Examples

We apply the same lexicon of semantic features to speech and to gesture (see Table 2.2). The starting point of the verbal annotation is the (intonation) ‘Child.phrase’. Note, that only phrases related to the described actions (See Sec. 3 for the list of these actions) are considered here. Looking at the phrase we ask: What kind of information is transmitted verbally? Verbally means, which words form the semantics? Each verbal SF annotation is assigned a unique *Child.SF.Speech.ID* starting from 100. For a deeper understanding of the verbal semantic annotation consider the examples listed in Table 2.2.10.

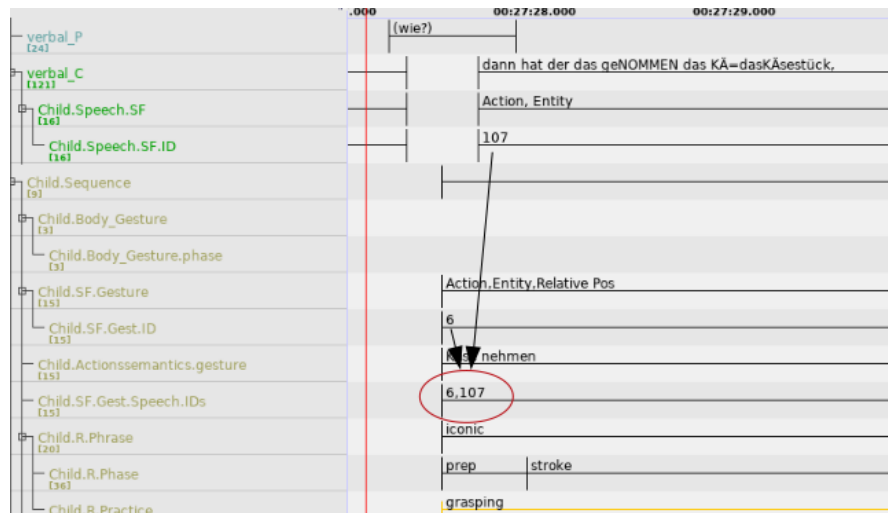
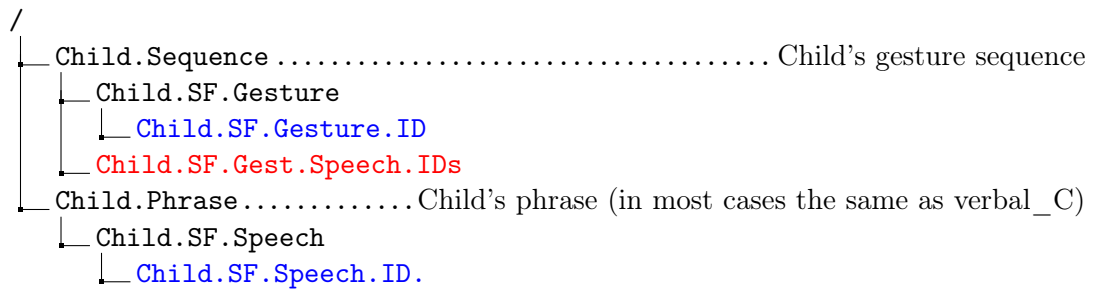


Figure 2.2: Example: annotation of *Child.SF.Gest.Speech.IDs*.

2.2.11 Semantic Feature IDs



While SFs of gesture and speech have their unique IDs, an additional ELAN-annotation brings the two parts together. The field *Child.SF.Gest.Speech.IDs* is dependent on the *Child.sequence* and contains the SF.IDs of the iconic gesture and the *Child.SF.Speech.ID* of its lexical affiliate (a (intonation) phrase related to the gesture). It is possible that multiple lexical affiliates correspond to a gesture, in this case all *Child.SF.Speech.IDs* are listed separated by a "," in the field.

SF	Description	Examples
ENTITY	refers to objects	spoon, bread, tea, chesee, etc.
ACTION	all kinds of action, action verbs	take, put, drink, eat, moving gestures
MANNER	the rule of thumb: in most cases where an action is present in the gesture we annotate ‘manner’	verbally: richtig rum, so, falschrum, CONVENTION: if we have a referential expression like, e.g., "Ne, so rum+gesture" or "so + gesture" then manner is coded only on the gesture.
DIRECTION	descriptions of directions	prepositions conveying spatial information, in, on, left, right, up, down, gestures, e.g., "Ganz nach unten"
SHAPE	object descriptions	forming an object with hands, shaping a bottle.
PROPERTY	object descriptions	other properties of objects than SHAPE, e.g., color.
AMOUNT	verbalized by numerals, annotated if more than 1 things are mentioned or referred to	serveral,multiple, two fingers shown
SIZE	related to the size of objects	so big, small, hands showing the width, height of objects
RELATITVE POSITION	relation between 2 objects, two hands placing an object	spatial relations, Löffel in den Mund, gestures.

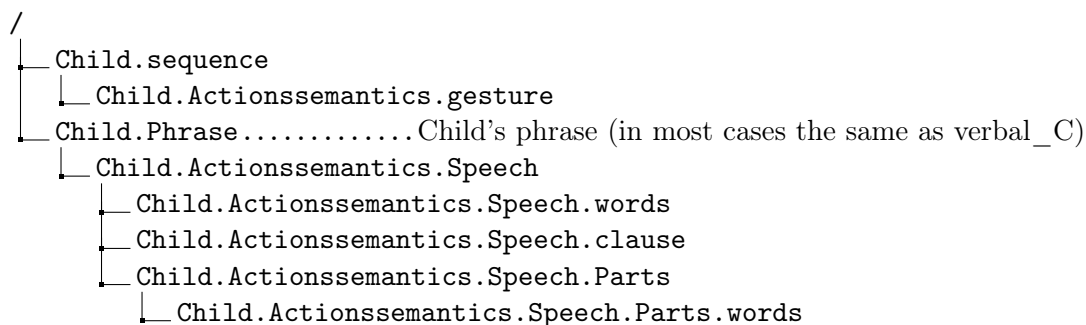
Table 2.1: Semantic Features used for gesture and speech.

Sentence	Description	Annotation
"er hat da so eine flasche genommen"	Here, an entity (=flasche) and an action (=genommen) are present	> Entity, Action
"den Käse drauf gelegt"	entity (= Käse), action (= gelegt), direction (=drauf)	> Entity, Action, Direction
"er hat es falschrum genommen"	entity (=es), action (=genommen), manner (=falschrum)	> Entity, Action, Manner

Table 2.2: Examples of SF-annotation of speech.

3 Actionsemantics

3.1 Annotation hierarchy



The annotation of actionsemantics is inspired by the work of (Alibali, Evans, Hostetter, Ryan, & Mainela-Arnold, 2009) and aims to measure the redundancy of speech and gestures on the level of single words or (intonation) phrases. Actionsemantics asks which action is performed by the particular gesture or phrase. In the ELAN-lexicon we defined a range of actions. If applicable, please select one of the listed actions with respect to gesture and speech separately.

Helpful questionos:

- What kind of action is described by this (intonation) phrase?
- What kind of action mimics this iconic gesture?

More specifically, for each iconic gesture and the corresponding intonation phrase we ask what kind of action (from the Table 3.1) is represented by this gesture/phrase? ‘Actionsemantics’ can be assigned to one or more intonation phrases combined into a *Child.phrase*. See Figure 3.1 for an illustration.

3.2 Redundancy

Redundancy is annotated with respect to the *Child.phrase*, but taking also the corresponding gesture into account. If a verbal phrase is uttered without gesture the fields *Child.Actionsemantics.words* and *Child.Actionsemantics.clause* remain empty.

Child.Actionsemantics.words Do the **exact words** accompanying the gesture express the meaning assigned to the gesture in the field of actionsemantics? (cf. (Alibali et al., 2009))

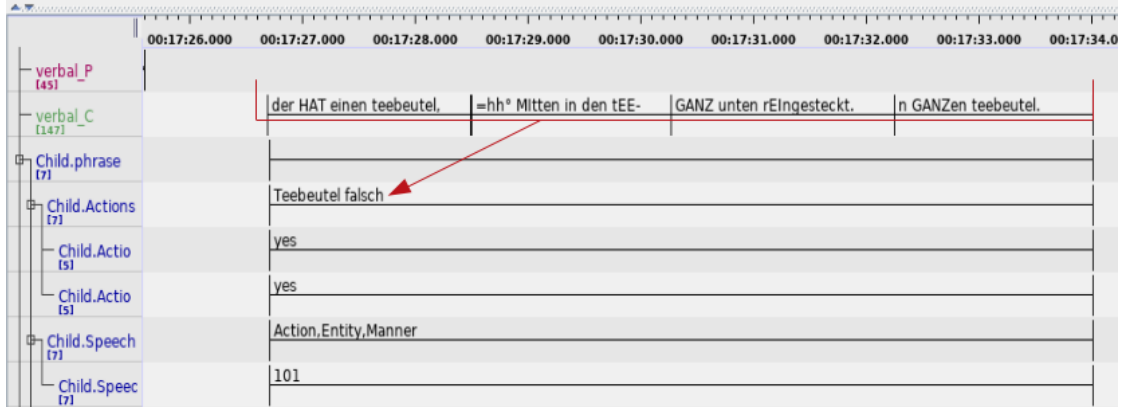


Figure 3.1: One gesture sequence corresponds to four intonation phrases.

For example, "*dann hat er das Käsestück genommen*", assigned actionsemantics is "*Käse nehmen*" and the corresponding gesture mimics the grasping of cheese. This results in the annotation of "yes" in the field `Child.Actionsemantics.words`.

Child.Actionsemantics.clause Does the clause(i.e., `Child.phrase`) express the same action as the gesture?

For example, "*dann hat er die andere Seite genommen*", assigned actionsemantics is *Flasche falsch* and the gesture mimics the bottom of the bottle. Here, we don't have the exact words representing the meaning, thus, the redundancy on the word level would be `Child.Actionsemantics.words` = "no". However, the meaning is still present in the clause, thus, we annotate `Child.Actionsemantics.clause` = "yes"

See Figure 3.2 for an example annotation of redundancy.

3.3 Combined

There are some rare cases when a child mixes two different actions in one utterance. In order to account for such cases we use the actionsemantics category `combined`. Consider the example from Figure 3.3. The child says "*ALso der wollte [..] tee ESSen,*" and uses the gesture related to grasping a spoon and leading it to the mouth. Obviously, two different situations are mentioned, whereby the gesture reflects only one of them. In this case, the lexical affiliate of the gesture is not the intonation phrase but a single word (i.e., "*essen*").

In order to account for this case, we assign to the `Child.Actionsemantics.Speech` the value `combined` (see Table 3.1). In addition, we assign to single parts of the IP the actionsemantics in the field `Child.Actionsemantics.Speech.parts`. Respectively, this layer contains only one level of redundancy - the word level.

	00:27:28.000	00:27:29.000	00:27:30.000	00:27:31.000	00:27:32.000	00:27:33.000
verbal_P (24)	e?)					
verbal_C (122)	dann hat der das geNOMMEN das KÄ=dasKäsestück,					ZUM BROT gelegt.
Child.Actionsemantics.Speech (17)	Käse nehmen					Käse richtig
Child.Actionsemantics.words (17)	yes					no
Child.Actionsemantics.clause (17)	yes					yes
Child.Speech.SF (16)	Action, Entity					Action,Entity,Relative Pos
Child.Speech.SF.ID (16)	107					108
Child.SF.overlap (4)	6.107					7.108
Child.Sequence (9)						
Child.Body_Gesture (3)						
Child.Body_Gesture.phase (3)						
Child.SF.Gesture (15)	Action, Entity					Action,Entity,Relative Pos
Child.SF.Gest.ID (15)	6					7
Child.Actionsemantics.gesture (15)	Käse nehmen					Käse richtig
Child.R.Phase (20)	iconic					iconic
Child.R.Phase (18)	prep	stroke			stroke	retr

Figure 3.2: Word level redundancy is defined as the presence of the exact words representing the meaning of the gesture. Clause level redundancy is the meaning of the gesture but produced using different words.

In our example, the part "tee" is assigned the actionsemantics "Teebeutel richtig" and the `Child.Actionsemantics.Speech.parts.words="no"`. To the part "essen" we assign the actionsemantics "Essen" which is the same for the gesture and results in the redundancy on the level of words `Child.Actionsemantics.Speech.parts.words="yes"`.

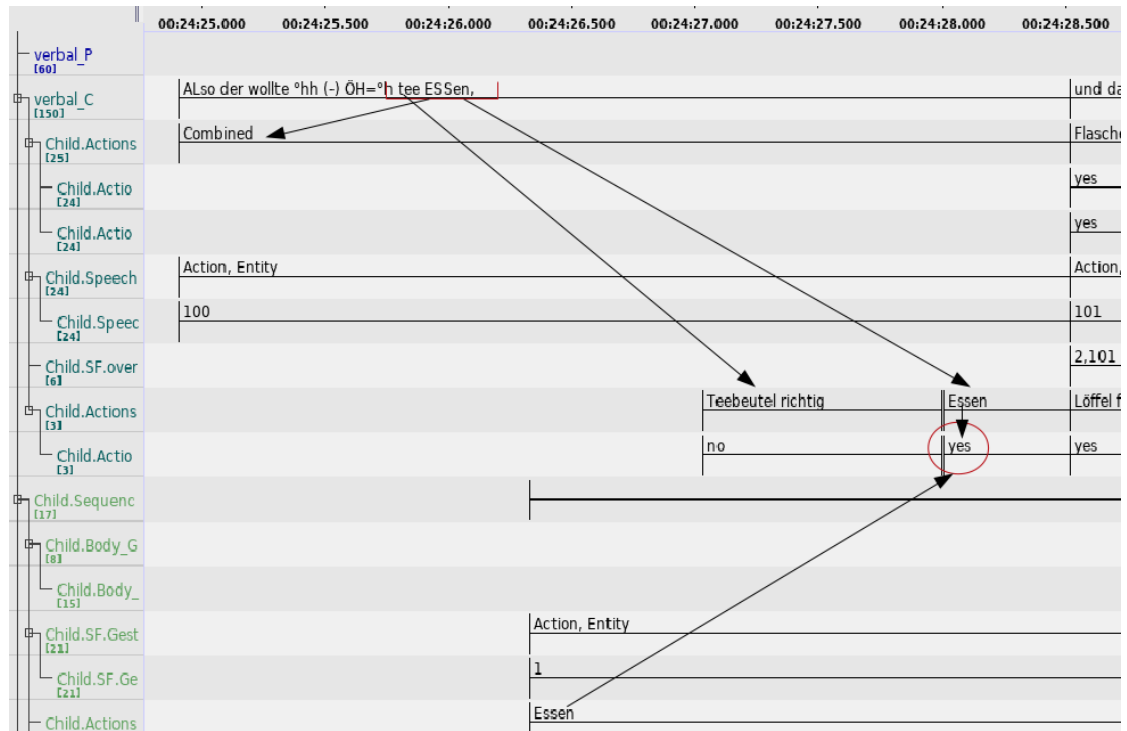


Figure 3.3: Combined actionsemantics.

Action	Speech	Gesture
BROT NEHMEN	<i>das Brot hingelegt</i>	hands performing a placing and/or shaping action
BROT UND KÄSE NEHMEN	<i>er hat so den Käse und das Brot genommen</i>	one hand = cheese, other = bread
ESSEN	<i>dann hat er so gegessen</i>	hands or body gesture representing eating
FADEN RAUS	<i>man sollte den so raushängen lassen</i>	modelling the action of putting the tea bag into the mug
FLASCHE	<i>er hat so eine Flasche genommen</i>	hands modelling the action of taking a bottle
FLASCHE RICHTIG	<i>so muss man trinken</i>	modelling the action of drinking from a bottle
FLASCHE FALSCH	<i>dann hat er so getrunken, das war nicht richtig</i>	modelling the action of drinking from a bottle
KÄSE FALSCH	<i>der hat den Käse auf die Ecke gelegt</i>	hands performing the action of placing the cheese on the side of the bread
KÄSE NEHMEN	<i>dann hat er den Käse genommen</i>	hands modelling the grasping of cheese
KÄSE RICHTIG	<i>ganz aufs Brot gelegt</i>	hands modelling the placing of cheese on the bread
LÖFFEL	<i>er hat den Löffel genommenn</i>	gesture showing a grasping of a spoon
LÖFFEL FALSCH	<i>der hat den so genommen</i>	hands showing the grip of the back side of the spoon
LÖFFEL RICHTIG	<i>ne, richtig rum</i>	gesture showing a grasping of a spoon
SALZ TISCH	<i>er hat das auf den Tisch gestreut</i>	a hand forms a salt shaker and shakes up and down modelling the action of salting
SALZ AUFSAMMELN	<i>dann hat er das mit den Pfoten genommen</i>	hands modelling the action of collecting salt from the table
SALZSTREUER	<i>er hat den Salzstreure genommen</i>	a hand grasps the salt shaker
TEE	<i>er wollte Tee kochen</i>	-
TEEBEUTEL FALSCH	<i>er hat den ganz reingeschmissen</i>	modelling the action of throwing the tea bag into the mug
TEEBEUTEL	<i>er hat den Teebeutel genommen</i>	the pointing finger and the thumb grasp the tea bag
TRINKEN	<i>dann hat er so getrunken</i>	one/both hands modell the drinking action
COMBINED	<i>also er wollte tee essen</i>	the hand models grasping of a spoon

Table 3.1: List of possible actions and examples of use cases occuring in speech or gesture.

4 Semantic Relations

Semantic relations describe the role of gesture with respect to the verbal utterance ('verbal_C'). Based on the work of (Colletta et al., 2014) we distinguish between five different categories describing the gesture.

4.1 Reinforces

The information provided by the gesture is the same as the linguistic information. The gesture emphasizes the verbal utterance, often expressed by 'beat-gestures'.

This feature is very similar to the next one and sometimes it is difficult to decide which one to use. Rule of thumb: use 'reinforces' only if an emphasize is visible in the gesture, in other cases use 'integrates'.

Example: nodding + *yes!*

<shouting loudly:> *den Salzstreuer!!!!* + beat gesture.

Helpful question: Is the gesture a 'beat' or 'iconic-beat'?

→ No: integrates

→ Yes: reinforces

4.2 Integrates

This feature is assigned to a gesture when the gesture adds precision to the verbal phrase.

Example: "*dann hat er die andere Seite genommen*" + the hand forms the back of the spoon. This feature does not add new information.

Helpful question: Is the gesture iconic?

→ Yes: integrates

4.3 supplements

The information present in the gesture adds new information to the content of the verbal phrase.

Example: "*aufgeleckt*" + showing with the body how the salt was collected. Supplements adds additional information, brings more details about the content of the verbal phrase.

Helpful question: Is the information provided by the gesture mandatory?

→ No: supplements

→ Yes: complements

4.4 Complements

the information provided by the gesture is the necessary complement to the verbal phrase.

Example: *Mit dem Maul SO in Mund* + gesture showing how the collecting of salt was performed. The rule of thumb: when referential expression like "so" is used, then the gesture is a necessary complement.

Helpful question: Is the verbal phrase incomplete (i.e., the gesture complements the content)?

→ Yes: complements

4.5 Contradicts

The information provided by the gesture contradicts the content of the verbal phrase. See Section 3.3 for an example.

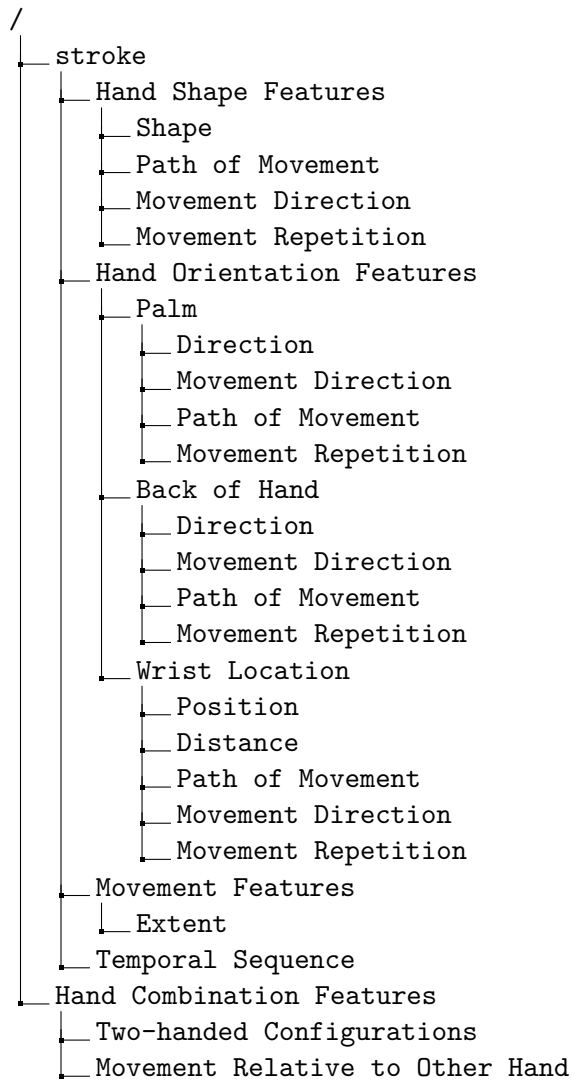
Annotation manual of gesture morphology

Kirsten Bergmann Oliver Damm Christian Fröhlich Stefan Kopp
Andy Lücking Florian Hahn Hannes Rieser Nick Thomas

Version 2 – 31.01.2014

5 Annotation manual of gesture morphology

5.1 Annotation hierarchy



5.2 Rules of thumb

1. To keep clarity, at the beginning of the first annotation of a video the dictionaries should be reduced to the most basic predicates. Therefore, the annotation should be made with the standard ELAN-template.
2. Always choose the most specific annotation predicate, for example, choose "FTT" instead of "BHA".
3. Always choose the smallest movement primitive, for example, "LINE>LINE>LINE>LINE" instead of LINE. Therefore, it is possible to uniquely identify the path through succession of the movement predicates.
4. The intermediate values of "palm direction" and "back of hand direction" (for example, "BTR/BAB"), which are expressed by "/" should be used when there's a deviation of 45 degrees.
5. The intermediate values of "palm direction" and "back of hand direction" are allowed to have three direction predicates at most.
6. "Movement directions", for example "MU>MD", and movement paths, for example "ARC", are only annotated with the greatest indicator. A change of "back of hand direction" and "palm" directions should always be annotated. For example:

	Hand Shape	Palm
Hand Shape	G>XI	—
Path	o	ARC
Movement Direction	o	MD
Direction	—	BUP>BDN

If the maximal cause can't be determined, a change of movement should be annotated in all tiers. There can be multiple maximal causes in one gesture, for example a change of wrist and back of hand direction can appear simultaneously.


7. Deictic-beats and iconic-beats have to be fully annotated. This means, if no repetition of movement can be annotated every movement direction has to be annotated.
8. If multiple predicates have to be merged together, it should happen in alphabetical order, for example: MF/ML/MU
9. A inverse repetition may only be annotated if it is *complete*. There's only one inverse repetition for each gesture.

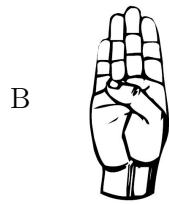
10. Slash predicates are only annotated as start or end values. The atomic predicates are also used as intermediate values. For example: PUP>PTL>PDN.
11. For every movement path stands: For each LINE value a movement value should be annotated. In the ARC case (exception: half circle, three quarters circle), there will be two annotation values; a segment consisting of n ARC values needs $n + 1$ movement values. The shape of a curve is always be approximated by a series of straight lines. For example, a quarter circle from bottom left to top right is annotated by MU>MR. In case of a half circle or three quarters of a circle, but with the whole number of movement values.
12. For every repetition stands: Despite it being possible to have multiple repetitions of one movement, there may be only one inverse repetition. The repeated movement segment starts always with the beginning of the gesture. This means, the repetition refers only to the first movement or to a sequence of movements, which includes the first movement.
13. Moves, discourse and hedging gestures mustn't be annotated. Therefore, every two handed configuration is assigned to the value 0, even when one hand performs another gesture.

5.3 Hand Shape

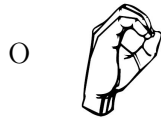
5.3.1 Hand-Shape

If one gesture token does not match the listed hand shapes, the one hand shape, which resembles the given gesture token the most, has to be chosen. A change of hand shapes during a gesture is annotated by successive ASL-symbols, which are linked by ">". There are additions for form "C" (grasping gesture), "medium", "small" and "large", to determine the distance between index finger and thumb. "Medium" is the default distance. To hint a weak hand tension while performing B,C or G, there are the values "loose-V", "loose-C" and "loose-G". If the index finger, while forming a G-shape, does not point in the back of hand direction, the value "Bent-G" should be annotated. "5" is annotated as "loose B-spread". Y is an additional hand shape. To cover a orthogonally held finger (relative to "back of hand direction"), the value "bent-B" is introduced.

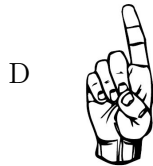
Predicate	Shape	Comment
A		In case of a fist with splayed out thumb, A-spread should be annotated



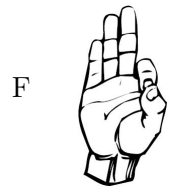
If the thumb isn't clearly folded, "B-spread" should be annotated. In this case, the thumb may be applied sideways. With spread out fingers, "loose" should be annotated.



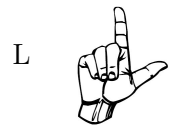
If the fingers don't form a circle and are distanced from the palm, "tapered O" should be annotated. If only the thumb and index finger form a circle and the other fingers don't, "baby-O" should be annotated. In case of thumb and middle finger forming a circle, "O" should be annotated.



Thumb is positioned on top of the tip of the middle finger

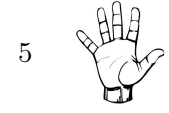
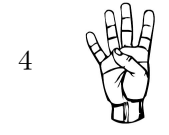
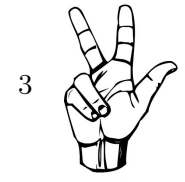
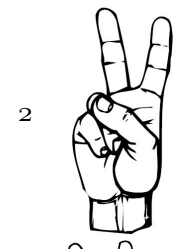


Annotate "bent-G" in case of folded index finger.





The difference to "bent-G" is the bent index finger and the folded thumb.



"Bent-5" requires a slightly bent finger and a curved palm.



5.3.2 Path of Hand-Shape

All following values may be combined at will; the single values should be separated by ">". Possible values for hand-shape movement are:

value	description
LINE	Movement in a straight line. A gesture may include a series of line segments; every segment requires one movement value.
ARC	Movement along an arc or curve; one ARC requires two movement values. A gesture may include a series of Arc segments. Two segments require three movement values; three segments require four movement values, and so on.
WIGGLE	Slight wiggling movement. In the hand shape, this is a wiggling of the fingers. In the orientation, this is a wiggle of the hand by moving the wrist. This seems to be uncommon in location.

5.3.3 Hand-Shape Movement Direction

For every LINE-value, a movement value should be annotated. There are two values for every ARC; a segment out of n ARC values requires $n + 1$ movement values. The shape of a curve will always be approximated by shapes of lines (e.g.: a quarter circle from bottom left to top right is annotated by "MU>MR"). From the following values, categories with up to three candidates may be formed (e.g.: MU/ML/MF). Possible values for hand-shape movement directions are:

value	description
o	no movement
MU	Up
MD	Down
ML	Left (right to left)
MR	Right (left to right)
MF	Forward (straight ahead, away from body)
MB	Backward (opposite of forward, towards body)

5.3.4 Hand-Shape Movement Repetition

Only the number of repetitions and not the first performance is annotated. If there is a repetition of movement, only the first item is described.

value	description
#N	Repeated N times
i#N	Repeated invert N times

5.4 Hand Orientation Features

5.4.1 Palm Direction

Intermediate values are described by the successive sequence of its extreme values and are separated by slashes. There may be only three intermediate values at the most.

Bsp.: PTL/PUP/PDN>PDN

value	description
PTL	palm facing toward left
PTR	palm facing toward right
PUP	palm facing upwards
PDN	palm facing downwards
PTB	palm facing towards body
PAB	palm facing away from body

5.4.2 Path of Palm Direction

value	description
LINE	(default Wert)
WIGGLE	Slight wiggling movement. In the hand shape, this is a wiggling of the fingers. In the orientation, this is a wiggle of the hand by moving the wrist. This seems to be uncommon in location.

5.4.3 Palm Direction Movement Direction

From the following values, one may be create mixed categories with up to three tokens (e.g.: MU/ML/MF).

As movement direction for the palm and back of hand only two rotation axis and therefore four direction of movement are sufficient. The "back of hand" and "palm"-direction may be described by a ray with its starting point at the middle of the back of the hand. For rotations of that ray, the middle of the back of the hand is the hub. A rotation of a line in a three dimensional space is only possible in two direction (MU or MD, and MR or ML). Starting from the original posture of the hand, it is possible to determine the exact direction in which the hand was rotated.

Possible values for direction of palm movement are:

value	description
o	no movement of palm orientation
MU	Up
MD	Down
ML	Left (right to left)
MR	Right (left to right)

5.4.4 Palm Direction Movement Repetition

Only the number of repetitions and not the first performance is annotated. If there is a repetition of movement, only the first item is described.

value	description
#N	Repeated N times
i#N	Repeated invert N times

5.4.5 Back of Hand Direction

The expression of intermediate values is analogous to "palm directions". The "back of hand" direction correlates to the axis, which is formed by expanding the back of the hand.

For example: BTL/BUP>BUP

value	description
BTL	Back of Hand pointing left
BTR	Back of Hand pointing right
BUP	Back of Hand pointing upwards
BDN	Back of Hand pointing downwards
BTB	Back of Hand pointing towards body
BAB	Back of Hand pointing away from body

5.4.6 Path of Back of Hand Direction

value	description
LINE	Movement in a straight line.(default Wert)
ARC	Movement along an arc or curve
CIRCLE	A circular movement. This does not necessarily need to be a complete full circle, but should be close to a circle
WIGGLE	Slight wiggling movement. In the hand shape, this is a wiggling of the fingers. In the orientation, this is a wiggle of the hand by moving the wrist. This seems to be uncommon in location.

5.4.7 Back of Hand Direction Movement

There may be formed categories with slashes with three of the following values at most (e.g.: MU/ML/MF). There are two annotated values for ARC and four for a CIRCLE. Possible values for the back of the hand direction movement are:

value	description
o	no movement inside the back of the hand
MU	Up
MD	Down
ML	Left (right to left)
MR	Right (left to right)

If the movement of the back of the hand is identical to the movement of the palm, only one movement in direction of the back of the hand is being annotated.

5.4.8 Back of Hand Direction Movement Repetition

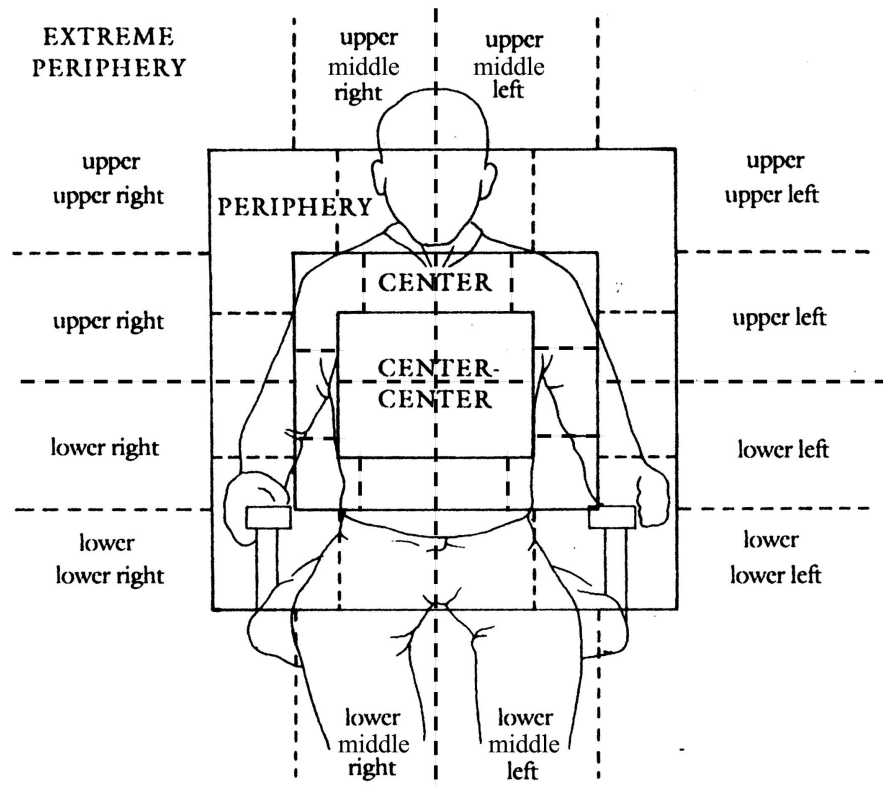
Only the number of repetitions and not the first performance is annotated. If there is a repetition of movement, only the first item is described.

value	description
#N	Repeated N times
i#N	Repeated invert N times

5.5 Wrist Location

5.5.1 Wrist Position

Hand location is coded using the McNeill schema plus a distance symbol (values illustrated in the figures below). For movements, only the starting position is coded, the rest can be reconstructed from the movement features.



value	description
CC	center center (@chest)
C-UP	center-upper (@neck)
C-UR	center-upper-right (@R-shldr)
C-UL	center-upper-left (@L-shldr)
C-RT	center-right (@R-arm)
C-LT	center-left (@L-arm)
C-LW	center-lower (@stomach)
C-LR	center-lower-right
C-LL	center-lower-left
P-UP	periphery upper (@face)
P-UR	periphery upper right (@abv R-shldr)
P-UL	periphery upper left (@abv L-shldr)
P-RT	periphery right
P-LT	periphery left
P-LW	periphery lower (@lap)
P-LR	periphery lower right
P-LL	periphery lower left
EP-UP	extreme periphery upper
EP-UR	extreme periphery upper right
EP-UL	extreme periphery upper left
EP-RT	extreme periphery right
EP-LT	extreme periphery left
EP-LW	extreme periphery lower
EP-LR	extreme periphery right
EP-LL	extreme periphery left

5.5.2 Wrist Distance

value	description
D-C	Hand in contact with body
D-CE	Hand between body and elbow's length away
D-EK	Between elbow and knee
D-KO	Between knee and length of outstretched arm in front away
D-O	Length of outstretched arm in front away

5.5.3 Path of Wrist Location

value	description
LINE	Movement in a straight line. A gesture may include a series of line segments
ARC	Movement along an arc or curve
CIRCLE	A circular movement. This does not necessarily need to be a complete full circle, but should be close to a circle
WIGGLE	Slight wiggling movement. In the hand shape, this is a wiggling of the fingers. In the orientation, this is a wiggle of the hand by moving the wrist. This seems to be uncommon in location.

5.5.4 Wrist Location Movement Direction

There may be formed categories with slashes with three of the following values at most (e.g.: MU/ML/MF). There are two annotated values for ARC and four for a CIRCLE. Possible values for the hand shape movement directions are:

value	description
o	wenn keine Bewegung des Wrist
MU	Up
MD	Down
ML	Left (right to left)
MR	Right (left to right)
MF	Forward (straight ahead, away from body)
MB	Backward (opposite of forward, towards body)
MBB	Behind body

5.5.5 Wrist Movement Repetition

Only the number of repetitions and not the first performance is annotated. If there is a repetition of movement, only the first item is described.

value	description
#N	Repeated N times
i#N	Repeated invert N times

5.6 Movement Features

5.6.1 Extent

value	description
SMALL	Within one region
MEDIUM	Across two regions
LARGE	Across three or more regions
o	statische Gesten

5.6.2 Temporal Sequence

Temporal sequence stance for the order in which the movements are performed. Therefore, the hand stages are listed in order of occurrence and are separated by ">". The default for synchronous gestures is "o". Cause of movements are: back of hand, wrist, palm direction and hand shape.

Wrist>Back of Hand

5.7 Hand Combination Features

5.7.1 Two-handed configurations

If the two-handed configuration at the beginning of a gesture is different from the ending, the predicates are put together in the right order and separated by ">". A break up of a FT is being annotated with "FT>BHA". All predicates are allowed to be annotated during a *stroke* or a *hold*. Contact with single fingers is annotated with FT or FTT. New predicates aren't introduced for this case. The annotation is performed by overlapping the stroke or hold. It is also possible to see a stroke in one hand and simultaneously a hold in the other hand. In that case, the annotation should be made respective to the hand which performs the stroke.

value	description
o	no stroke or hold is performed
FTT	Tips of fingers and thumbs touching
FT	Tips of fingers touching (thumbs not touching)
TT	Thumbs touching (tips of fingers not touching)
FTF	Tips of fingers facing
PT	Palms together
PF	Palms facing
LPBH	left palm faces right BOH
RPBH	right palm faces left BOH
LFTH	Tips of left hand finger touching right hand
RFTH	Tips of right hand finger touching left hand
LFTA	Left hand touching right arm
RFTA	Right hand touching left arm
BHA	Both hands active, without any facing or touching

5.7.2 Movement relative to other Hand

Mirror predicates SYNC may only be annotated for the duration of the overlap of the strokes. All other predicates (RHH, LHH, NOSYNC, o) can also be annotated in presence of a "hold"; the length of this annotation depends on the hand which performs the stroke.

value	description
o	Both hands perform no movement, both hands are static or one hand is not participating
Mirror-Sagittal	
Mirror-Frontal	
Mirror-Transversal	
RHH	Right hand is held in a stable position, anchoring the frame of reference, while the left hand is active
LHH	Left hand is held in a stable position, anchoring the frame of reference, while the right hand is active
NOSYNC	Use when actions of the left and right hand are not in synch with each other in some way that is significant to understanding the gesture
SYNC	Both hands move in the same way. Die Mirror Prädikate der Hände müssen identisch sein.

5.7.3 Two-handed Form

As additional tier, a two-handed form is introduced for router and follower. In this tier, the dynamic complex forms, which are performed by both hands together, are annotated.

For example, when both hands form a semi circle, CIRCLE is annotated as the two-handed form. The following list of possible predicates is open and should be expanded by the annotator if needed:

value	description
o	Both hands do not move or the movement does not form a recognizable complex shape
U-shape	
V-Shape	
RECTANGLE	
QUADRAT	
CIRCLE	
Tropfen	
Edge	

References

- Alibali, M., Evans, J., Hostetter, A., Ryan, K., & Mainela-Arnold, E. (2009, 12). Gesture-speech integration in narrative: Are children less redundant than adults? , 9, 290-311.
- Bergmann, K. (2012). The Production of Co-Speech Iconic Gestures: Empirical Study and Computational Simulation with Virtual Agents. Retrieved 2018-08-17, from <https://pub.uni-bielefeld.de/publication/2460005>
- Bergmann, K., & Kopp, S. (2006). Verbal or Visual? How Information is Distributed across Speech and Gesture in Spatial Dialog. In D. Schlangen & R. Fernandez (Eds.), *brandial 06. proceedings of the 10th workshop on the semantics and pragmatics of dialogue (semdial-10)* (pp. 90–97). Universitätsverlag.
- Colletta, J.-M., Guidetti, M., Capirci, O., Cristilli, C., Demir-Lira, O. E., Kunene, R., & Levine, S. (2014, 02). Effects of age and language on co-speech gesture production: An investigation of french, american, and italian children's narratives. , 42, 1-24.