Coding and annotation manual of multimodal feedback events

by Anastasia Bauer
(University of Cologne)

last updated: 20 March 2025

The following annotation scheme is based on Bauer & Gipper et al (under revisions)[1].

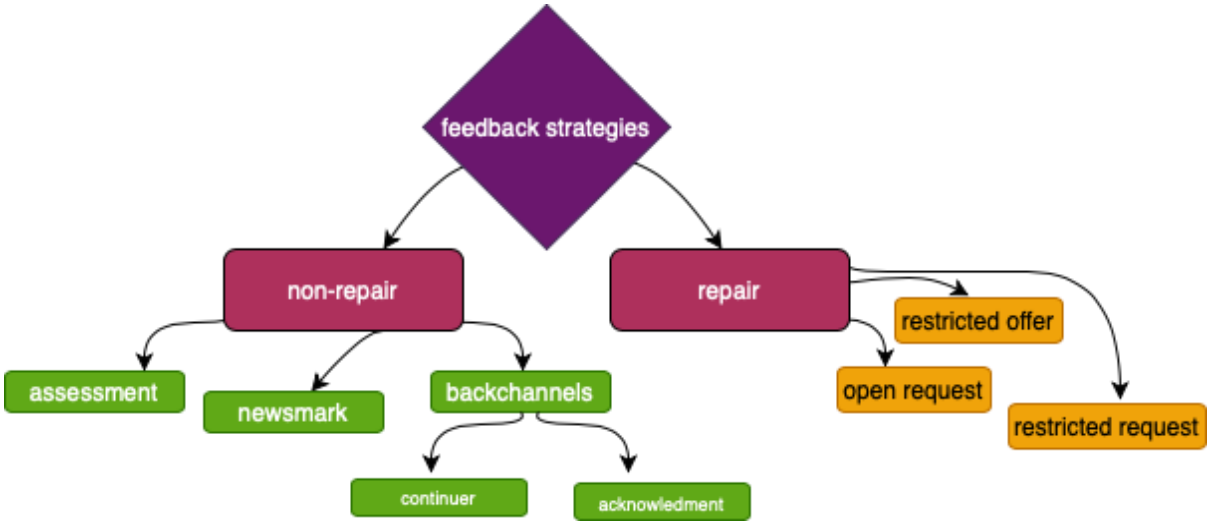

Figure 1: *Feedback strategies based (Dingemanse & Enfield 2015; Gardner 2001; Clark 2020)*

This manual includes three types of other-initiated repair (i.e., we exclude self-repair) and four types of non-repair feedback[2].

In ELAN annotations, the speaker/signer to the left is speaker/signer 1, the speaker/signer to the right is speaker/signer 2. This should be done consistently with ALL manual and non-manual cues: e.g. head_1, mouth gesture_2 etc.

We annotate the following:

i. The feedback signals itself on various ELAN tiers (speech, head, eyebrows etc. see section I for annotation scheme of non-manuals)
ii. The feedback category (whether it is a continuer or a newsmark or other) (see Figure 1)
iii. The type of the feedback event (whether it is lexical such as the word 'ja' only or non-manual only such as a head nod or a combination)
iv. Turn-taking behaviour of the feedback event

[1] Bauer, Anastasia & Sonja, Gipper, Tobias-Alexander Herrmann, Jana Hosemann. (submitted). Rethinking linguistic feedback: A modality-agnostic and holistic approach to multimodal addressee signals in signed and spoken dyadic interaction. Submitted to GLOSSA, Nov. 2024.

[2] Continuer and acknowledgments are difficult to differentiate by semantics or their form and are best distinguished by their sequential position, therefore are treated here as backchannels. Continuers come in the 2nd position, acknowledgments in the 3rd position).

I. Annotation scheme for non-manual signals

As we are collecting cases of multimodal (verbal as well as non-verbal, manual and non-manual) feedback, we annotate all possible feedback signals on the following tiers. Note that the tier "eyes" includes different types of eye blinks. **For first the comparative study, we are not going to look at eyeblinks, therefore the annotation of those is not obligatory for that investigation**:

Abbreviation	Meaning
Tier eyebrows	
bf	eyebrows furrowed (=eyebrows are pulled together)
br	eyebrows raised
brd	eyebrows lowered
Tier eyes	
mbl	multiple eye blinks
sbl	short blink (no longer than 410[3] ms)
lbl	long blink (longer than 410 ms)
esc	eyes squinted
ew	eyes wide opened
Tier eye gaze	
gaz	looking/starring at the addressee
gaw	looking away/gaze aversion
gob	looking at the object (e.g. when they are manipulating)
gf	eye gaze straight
gfp	eye gaze up
gl	eye gaze left & straight
gld	eye gaze left & down
glp	eye gaze left & up
gr	eye gaze right & straight
grd	eye gaze right & down
Tier mouth gesture	
lbt	biting of the lower lip
ldn	corners of the mouth lowered down
ldr	lips sucked in
lo	lips rounded
lpd	lower lip pushed forward
lpr	lips stretched and pressed together
lpf	lips pushed forward
ls	lips stretched, mouth slightly open
lp	lips pressed together
lvb	lips tremble
mbl	blowing out air

[3] Inspired by Hömke et al. (2017) who categorized blinks into short and long listener blinks, based on the distribution of blink durations observed in their corpus of face-to-face conversations, using a threshold of 410 ms. In their data the long blink comprised 25%.

Abbreviation	Meaning
mcl	mouth is closed
mo	mouth open
msc	sucking in the air
tch	tongue against the cheek
tll	tongue against the lower lip
tt	tongue out
cms	closed mouth smile (s1)[4]
oms	open mouth smile (s2)
woms	wide open mouth smile (s3)
lgh	laughing smile, smiling with jaw dropped (s4)

Tier head	
hnn	many short head nods
sn	small (shallow) head nod
ln	large head nod[5]
lnn	many large nods
mn	mixed nod (e.g. one large nod followed by small nod(s))
hb	head tilt back (the head is moving up, without a turn to the left or to the right) – a head movement backwards, sometimes the upward position is held in an intensified way, looks like a reversed nod, first going up and then sometime after down to the default position
hbn	head tilt back (the head is moving up, without a turn to the left or to the right) – a head movement backwards, sometimes the upward position is held in an intensified way, looks like a reversed nod, first going up and then sometime after down to the default position with a subsequent head nod
hs	head shake
hmb	head move backward- is a movement of the head backward (the head is pushed backwards along the saggital axis) this can either be a movement of the head only or can be movement of the whole trunk
hmf	head move forward (without the turn to the right or to the left) head move forward- is a movement of the head forward (the head is pushed forward along the saggital axis) this can either be a movement of the head only or can be movement of the whole trunk
ht	head turn to the left/right/forward/backward
hth	head lowering
hths	head lowering & head shake
htl	head tilted to the right or left shoulder or back/forward

[4] Description of the Smiling Intensity Scale (SIS) of Gironzetti et al. (2016) for annotating smile activity.
[5] Large nods are large in amplitude (trajectory), going almost all the way up and then down to the neutral position.

Abbreviation	Meaning
cu	no head back tilt, but the chin is going up and stays there for a while
wig	head wiggle (neither shake nor nod, something like ∞)
Tier nose	
nw	nose wrinkled
nbl	nose blows out air
Tier cheeks	
chp	cheeks blown out/puffed
chs	cheeks sucked in
Tier shoulders	
shf	shoulders curved forwards
shh	shoulders raised
shs	shoulder shrug (raising and lowering of shoulders like „I don't know!“)
Tier body	
bb	body leaned backward
bf	body leaned forward
bu	body moves/raises up
bt	body turned to the left/to the right
bl	body leaned to the left/to the right (without turning)
Gesture[6]	various manual gestures

[6] Most of the labels are taken from Bressem & Müller 2014 as visualized here:

Abbreviation	Meaning
adapter	involve either self-adapters[7] or other-adapters.
Brush-away	
Throw-away	
Palm-up	=PUOH (palm up open hand)
First	
Vague	
Cyclic	
IX	Index finger
Ring	
Hold	
Drop	

Tier "speech": No controlled vocabulary. For speech, write down what the person says/signs.

II. THE FEEDBACK CATEGORY (available as controlled vocabulary in ELAN template)

The name of the tier is feedback category. For the first comparative study, not all categories are obligatory (see table below). Crucially, the annotation in feedback category defines the length of the overall feedback event. A feedback event may consist of a single or of multiple signals (e.g., of a head nod only, or of a head nod combined with speech, or with a head nod combined with an eyebrow raise.). **The length of a feedback event is defined by the start of the first signal involved and the end of the last signal.** To separate one feedback event from the next feedback event, there are two options. Either, two subsequent feedback events were separated by **300ms** of no movements or talk related to giving feedback. Or, two different feedback events could occur consecutively, distinguished by a noticeable change in movement, shape, or direction—for instance, a head nod transitioning into a head shake. We then added annotations on separate tiers for each articulator involved (see Bauer et al submitted). Crucially, all annotations for all articulators have to be under the respective annotation of feedback category in order for the data extraction script to work.

The articulator "gaze" is excluded from this, as, e.g., gaze to the addressee may persist for a long time which may obscure the presence of different feedback events. That is, continuous gaze does not contribute to the length of our feedback event and may be annotated independently from the length of that tier.

Category	Explanation	Example

[7] Self-adapters are inwardly oriented movements, involving self-touching of some kind. Other-adapters are externally oriented and entail touching some object, such as rubbing the desk. Adapters may consist of discrete, one-off actions, such as quickly scratching oneself or pushing back one's hair, or sustained actions, such as rubbing one's fingers together for an extended time. Note that this category of adaptive functions is excluded from many gesture studies because of researchers' focus on referential and pragmatic functions of gesture (Cienki 2024:4)

Continuer	brief signals during the turn or short after the turn of the primary speaker/signers	Verbal signals: mm, ja Non-verbal: Head nod,
Acknowledment	signals claiming to show agreement or understanding of the prior turn; come in the 3rd position (B: have you seen the doctor? A: yes, I have seen him B: ah, okay	Verbal signals: mm, ja, yeah, okay Non-verbal: Head nod
Newsmaker	or newsmarker-like objects mark the prior speakers's turn as newsworthy in some way	Verbal signals: really? Non-verbal: head back,
Assessment	assess the prior utterance	Verbal signals: i*nteresting, wow, cool,* Non-verbal: *laughing, smiling*
Repetition		I won the lottery - you won the lottery? I lost my key yesterday. - your keys?
Open request		Verbal signals: what? Huh? Non-verbal: Freeze look **(Not obligatory for initial comparative study)**
Restricted request		Where? **(Not obligatory for initial comparative study)**
Restricted offer		The boy? **(Not obligatory for initial comparative study)**

III. THE TYPE OF FEEDBACK EVENT (controlled vocab available; associated to feedback category)

lexical	*yes*
lexical_nonmanual	*yes + head nod*
nonlexical	*mm*
nonlexical_nonmanual	*mm + head nod*
nonmanual	*head nod*
manual	*YES, gesture*
man_lap	*manual elements occurring in signer's lap*
man_lap_nonmanual	*manual elements occurring in signer's lap + head nod*
manual_nonmanual	*YES + head nod*

IV. Turn management behaviour (controlled vocab available; associated to feedback category) – here we annotate whether the feedback event initiates a turn by the speaker giving feedback or not:

- full take-over (TT)
- partial take-over (length under 2000ms)
- overlap (2000ms)
- take-over attempt
- passive recipiency (PR)