Gestural Origins Coding Protocol

This guide is intended to be used when coding with the Gestural Origins gesture coding scheme set up by the Wild Minds Lab and is part of the supplementary material of the article Grund et al. 2023: 'GesturalOrigins: A bottom-up framework for establishing systematic ape gesture data across species' (check here for more theoretical background). It is written for our ape coding scheme; we have a modified version for coding elephant communication and will be adding other species - for example, corvids - in the future. This means that it is a living scheme and you can check the Gestural Origins GitHub repository of the Wild Minds lab (https://github.com/Wild-Minds/GesturalOrigins) for updates and new additions (such as, for example, new species-specific information). If you use our coding scheme and/or its related open-access material please cite our article (Grund et al., 2023).

The Gestural Origins project compares ape gesture use across individuals, communities, generations, populations, and species. To read more about our work check our websites: www.greatapedictionary.com and www.wildminds.ac.uk.

0. Video labelling

We label all videos using the following format [Field site, initials of researcher, yyyy-mm-dd, Group, Clip X, Time, Individuals]. Doing so allows us to easily tie a file to the site, individuals present, and the person who filmed it. For example:

> Bwindi CG 2019-12-31 MK Clip 02 10:10:02 MK, BW, GTE (BS)
> Budongo LS 2018-06-13 Clip 07 16:09:37 NR, HW (MS) (KL, KO)

Video files are numbered numerically in chronological order for the day, i.e. Clip 01, Clip 02, etc. Individuals are labelled to indicate who is on the screen from left to right, individuals who join the clip once it has started are indicated in brackets. For example in the 2nd file above NR and HW were on screen at the start of the clip, with NR to the left of HW. MS then joined. KL and KO then joined with KL to the left of KO as they appeared on the screen.

We also describe basic metadata (e.g., date, video length, site, behaviour present) for every video. An example of the data in our East African chimpanzee database is available here.

For the purposes of comparison with earlier work we retained a basic coding structure that distinguishes each case of gesture (an individual use), from a Sequence (a series of gestures made with less than a 1-second interval between their minimum units), from a Bout (a series of gestures made with intervals of any duration, but without a gestural response by the recipient) from a Communication (all gestures made by one individual during the communicative event, this includes all the gestures made by one individual even when they are exchange back-and-forth between individuals). Note that you are not restricted to using this 1-second interval-based structure, sequence structures can be constructed for gestures combined at any interval.

[*] indicates a variable that changes for every Gesture case
[**] indicates a variable that is held constant for every gesture within a Sequence
[***] indicates a variable that is held constant for every gesture within a Bout
[****] indicates a variable that is held constant for every gesture within a Communication

Throughout: Unclear refers to where we know there was a signal, action, etc. but we don't know what type. Unknown refers to where we don't know if a behaviour happened at all.

1. Coding with the Gestural Origins template in ELAN (first time coders)

We decided to implement our coding scheme in the open-access linguistic video annotation software ELAN.

If you've never used ELAN before we'd recommend checking out their webpages or Emilie Genty's online ELAN coding tutorial (https://greatapesgestures.github.io/). Genty and colleagues pioneered the use of ELAN to code (ape) communicative interactions (e.g., Heesen et al. 2020; Genty & Fuchs, 2023) and their online materials are an excellent resource for exploring the software's functions to this purpose.

This guide is from the Wild Minds Lab and focusses on explaining how to code ape gestural interactions with our group's specific coding scheme (The Gestural Origins coding template). We recommend reading our manuscript (Grund et al., 2023) before starting to code as it outlines the conceptual decisions behind the way we code. The templates, controlled vocabularies, definition sheets etc. can be downloaded from the paper's online supplementary material (GitHub). For a short ELAN video tutorial on how to use the template to code gestural interactions, please click here (video example of a mountain gorilla gestural communication being coded with our coding scheme). For any updates on templates, definitions, etc. please check here.

We take a bottom-up approach to coding, so in many cases the variables that we use for analyses (for example 'gesture type' or 'response latency') are calculated or built up afterwards using data from different tiers. In addition, we import variables such as age, rank, species, etc. afterwards into our databases so these do not need to be directly coded from the video itself. All our data are combined in a single database, currently hosted in Filemaker (we are exploring open-access alternatives).

Before you start, you will need to add the specific individuals for your site (ELAN controlled vocabulary: "Individuals"), and add the controlled vocabulary lists for species-specific vocalisations and facial expressions to the template (see excel file GOv1.0_Elan_controlled_vocabulary.xlsx for mountain gorillas and chimpanzees). You can export the relevant sheets as .csv files and directly load them into the GesturalOrigins ELAN template's (GOv1.0_ELAN_template.etf) controlled vocabulary (for instructions see below under point 1.2). All other aspects of coding remain the same across ape species. If you are not yet comfortable with ELAN and would like to skip this step, it makes sense to check the

Gestural Origins GitHub repository as there may be species-specific templates there that can be downloaded (however they are also site-specific, so you will have to change the IDs in case you work with a 'new' group).

Please make sure you download both template files (GOv1.0_ELAN_template.etf and GOv1.0_ELAN_template.pfsx). If you do not download the .pfsx file, the formatting is lost and you will likely have to re-order the tiers. You can find our default here. Note that the order does not impact coding and you may prefer a different order, but we find this one the most efficient. You can drag and drop tiers into any order, but you can also sort the tiers by the hierarchical structure of the tier groups first, to reduce the amount of drag-drop actions needed. Once you've saved this version of your template you won't have to do this again (unless you want to), and you're ready to start coding.

Open a new ELAN project and upload a media file (the video to be coded) and an annotation file (the template coding scheme where the gestural data are stored). To upload the media file and annotation file select *New…* from the *File* drop down menu in the top left corner of the ELAN project (Figure 1a). A new window will open up (Figure 1b). Select *Add Media File…* and *Add Template File (.etf)…* to choose the media and annotation file from the computer.

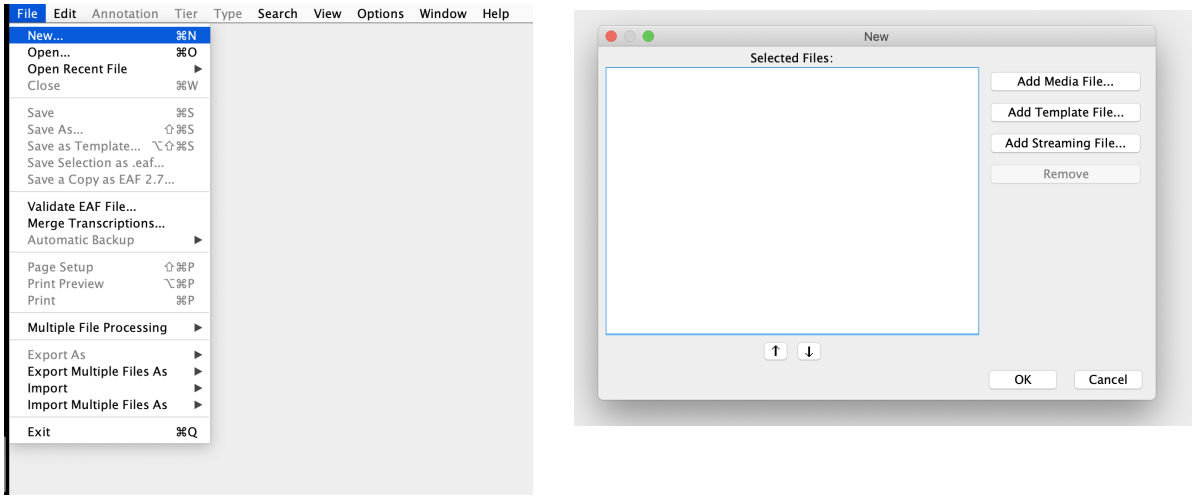


Figure 1: *Uploading media and annotation file: a) how to reach the pop-up window b) the pop-up window used to upload these files.*

Select the *Control* tab in the top right. Using this panel, the volume and playback speed of the video can be adjusted. The playback speed of the video can be sped up to double speed or slowed down. The slow down feature is useful to watch gestures or behaviours in more detail, but we don't recommend only watching videos this way as it can be harder to judge how salient actions are (Figure 2).

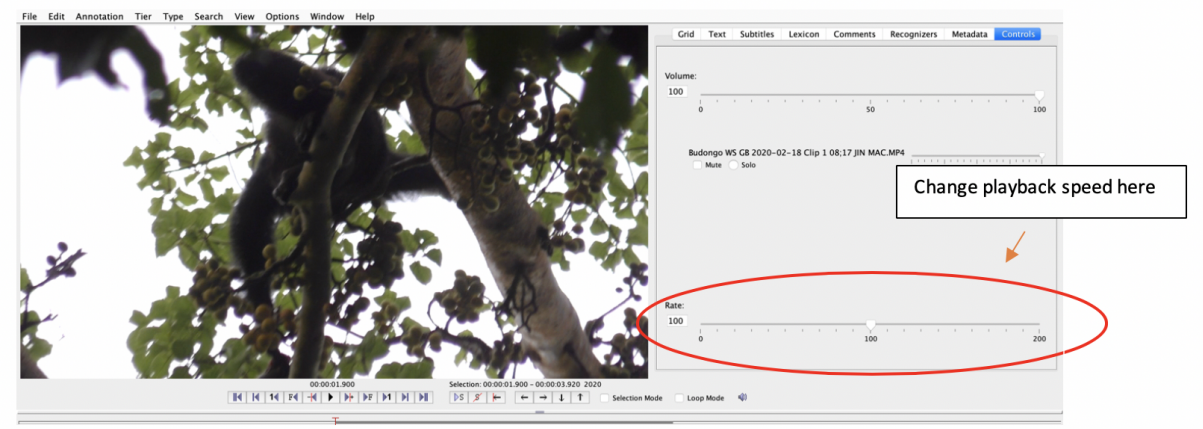


Figure 2: *Control tab in ELAN*

The control panel below the video viewer (Figure 3) can be used to navigate within the video.

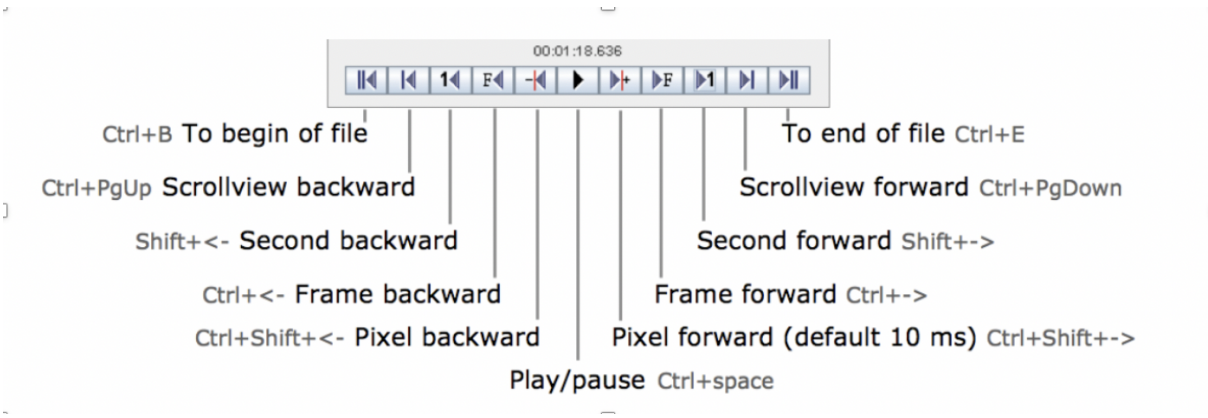


Figure 3: *Video control panel: shortcuts and useful buttons*

To start coding, a selection must be created for the communication bout. A selection in ELAN refers to a time interval over which a behaviour takes place. For example, a selection is used to mark the duration of a communication event or of specific gestures. A selection is made by dragging the mouse across an empty viewer area under the time ruler (Figure 4) and double clicking on the highlighted area. A selection becomes an annotation when information is entered in it. This information may be free text that you enter directly (e.g. in our case things like the Com_number, Clip_name, Rec_number) or selection from a drop down list (e.g. Signaller, Gesture_record). We use drop down lists wherever possible to reduce typos and other errors. To confirm an annotation press *ESC* or *ENTER* and you can then make a new one. A selection can also be made by pressing *SHIFT + ENTER* where you want the annotation to start and again *SHIFT + ENTER* where you want the annotation to end. You can move the crosshairs freely as you wish whilst doing this. Remember first to select the tier where you want the annotation to be made.

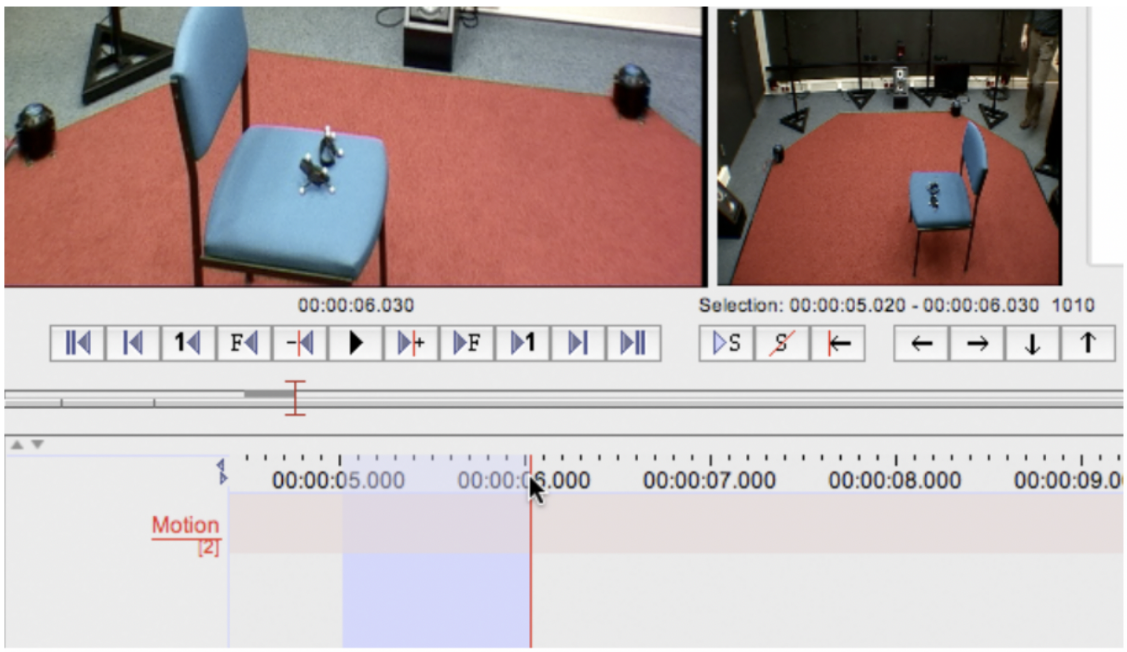


Figure 4: *Making a selection*

The selection controls (Figure 5) can be used to control playback of specific selections.

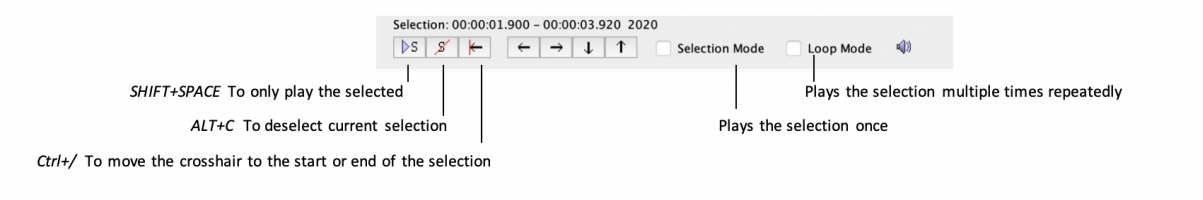


Figure 5: *Selection control panel: shortcuts and buttons.*

A selection/annotation can be deleted with the shortcut *OPTION+D* on Mac.

Once an annotation is created the duration of this annotation can be changed more accurately by using the *Modify Annotation Time (cntrl + shift + M)…* option in the *Annotation* drop down menu in the top left (Figure 6) The duration can also be modified by clicking on the annotation and pressing *OPTION* on mac and dragging the start or end point of the

annotation forwards or backwards.

Figure 6: *Changing the duration of annotations using the pop-up window/*

Tiers can be set up to have other tiers 'tied' to them - the main tier is called a 'Parent tier' and the other tiers are 'Child tiers'. You have to enter an annotation in the parent before you can add anything to the child tiers. All annotations within child tiers must be contained within the annotation of the parent tier (they can not start or end before or after it).

Annotations in the child tiers can be easily added by double clicking on the tier row under the selection of the parent tier (Figure 7). They are filled by either selecting the variable from a drop-down list or using free text.

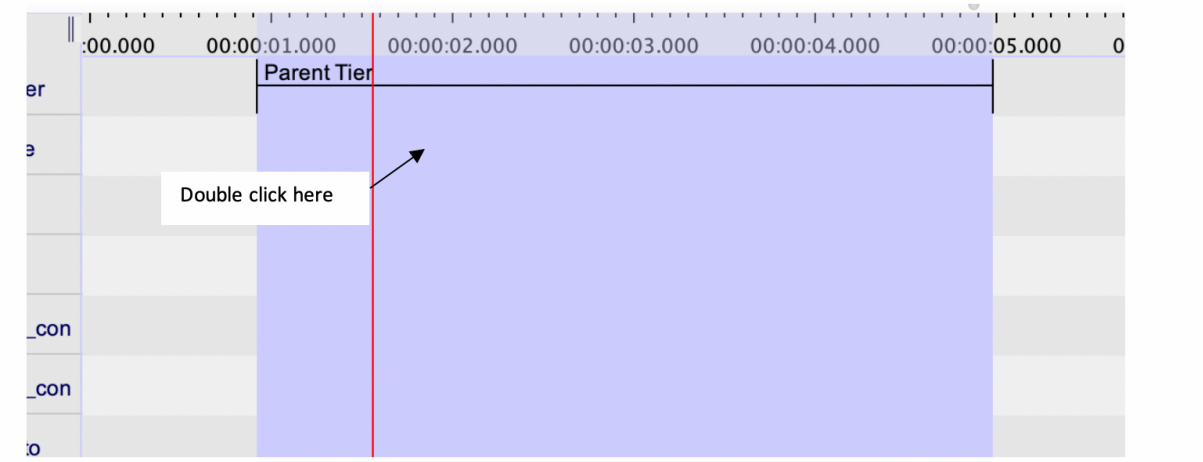


Figure 7: *Adding annotations to child tiers*

There are a lot of useful in-built options to ease navigation in ELAN and adjust it to your specific coding needs. For example, you can edit the default shortcuts to set these up to your own preferences (*Edit…Preferences…Edit shortcuts…)*, search for and edit specific annotations (Search…), and extract summary statistics from your coded .eaf files (*File…Multiple File Processing…Statistics for Multiple Files…*). This is just to name a few of its features. We focus here on how to code with our template, but it will ultimately save you time to first get properly acquainted with ELAN before diving into coding, as there is a lot of flexibility that can help make your coding life easier in the long run. Again, if you've never

used ELAN before we'd recommend checking out their webpages or Emilie Genty's online ELAN ELAN coding tutorial (https://greatapesgestures.github.io/).

1.1 Tier types

Most of our child tiers are linked to a parent tier with either 'Symbolic Association': here the annotation must be the same length as the parent tier, or as 'Included in' where the annotation can have its own length but must be contained within the parent tier.

1.2 Changing Controlled Vocabularies (CVs)

Select *Edit Controlled Vocabularies…* in the Edit drop down menu or use the *SHIFT+CTRL+C* shortcut to open the controlled vocabulary page (Figure 8). On this page, the controlled vocabularies (drop-down lists) can be modified. The controlled vocabulary that needs changing can be selected from the *Current CVs* drop down menu. Variables from a controlled vocabulary can be modified using the controls on the bottom left.

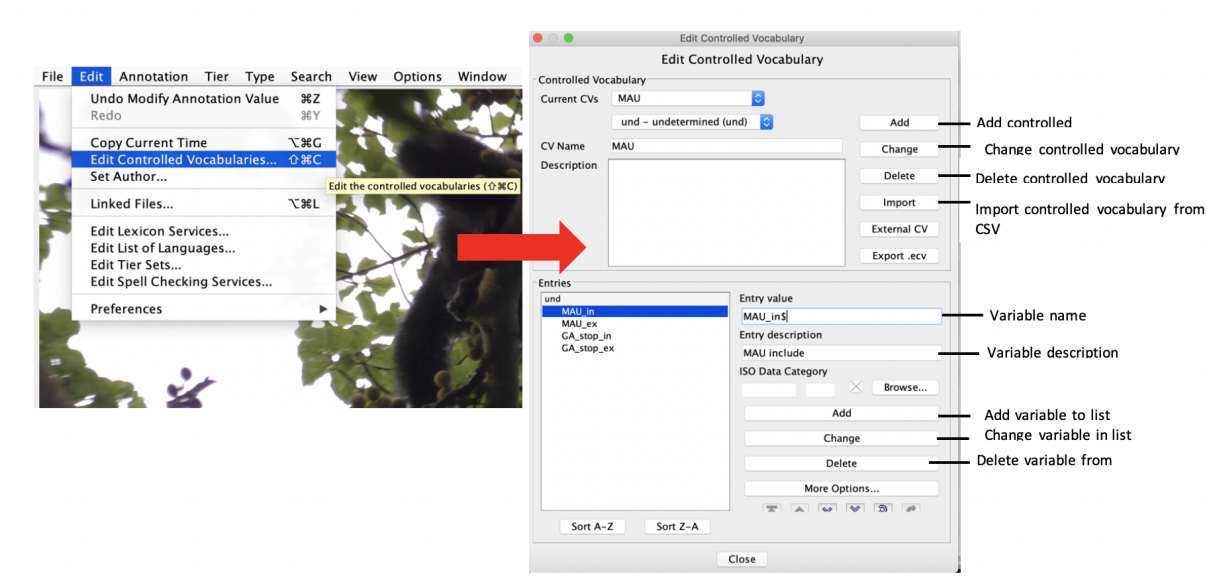


Figure 8: *Changing controlled vocabularies*

New CVs can be added manually in this page, or they can be imported from a CSV file using the *Import* button. To manually add a CV, enter the name of the CV in the *CV Name* text box, add all the variables in the Entries box at the bottom.

To add variables to a list, enter the variable name and description in the *Entry value* and *Entry description* (optional) in the text boxes in the Entries section. Click the *Add* button to add the new variable to the list.

To delete a variable from the list, highlight the entry by clicking on it and then click the *Delete* button. Multiple variables can be highlighted at the same time by holding down the *CTRL* key.

To edit a single variable, highlight the variable and make the appropriate changes using the *Entry value* and *Entry description* text boxes on the right. To save the changes made to an entry click the *Change* button in the Entries section.

1.3 Coding overlapping gestures

ELAN does not allow for overlapping annotations on the same tier. Where a signaller produces two gesture actions simultaneously, we include the second gesture on a separate set of duplicated Gesture_record tiers (Gesture_record2: all child tiers have G2 in front of them, e.g. G2_MAU_duration, G2_Rec_number, etc). For the same purpose we have a duplicated set of Bout_part tier (Bout_part2: all child tiers have B2 in front of them, e.g. B2_Persistence, B2_Sgn_location, etc.).

In the (very very) rare cases of three overlapping gestures from the same signaller the two most relevant gestures are coded and the third gesture should be added in the comments. Use your common sense to assess relevance, e.g. the gesture that is least likely to be received by the recipient (perhaps one that was out of sight), the gesture that is shortest and contained within the other gestures, the gesture that is made with the least effort by the signaller.

Note: If there are e.g. three gestures in a sequence (1_3, 2_3, 3_3) of which only gestures 1 and 2 overlap, the second gesture will be annotated in the Gesture_record2 tier (i.e. the duplicated section, sequence part: 2_3) whereas the third gesture that is not overlapping will be marked again in the main Gesture_record1 section (Gesture_record1, G1_sequence part: 3_3).

1.4 Coding gesture exchanges

Gestural exchanges (Individual A gestures to individual B, individual B gestures back to individual A) are coded successively. First all gestures of individual A are coded and the ELAN file is labelled and saved (e.g. BwindiCG20190809BTClip13Com102**a**). Subsequently the gesture record annotations for individual A are deleted and those for individual B are coded and the file is renamed and saved (e.g. BwindiCG20190809BTClip13Com102**b**). Thus all communications that are gestural exchanges result in two ELAN files (**a** for individual A's gesture records, **b** for individual B's gesture records) that together describe the same communication (Com102). It is also necessary to add the letters to the Com number tier - **see 2.1**

1.5 Using Free text tiers

We use R as part of the export-import process from ELAN to filemaker. As R is case sensitive so when we use codes like 'NV' and 'Unk' these need to have the same letters capitalised. We use NV and Unk (instead of nv and unk) to keep coding consistent across tiers.

2. The Gestural Origins coding scheme: ELAN tiers

OK, now you're ready to code! The variables below are listed roughly following the suggested tier order for coding. You can change the order of the tiers to suit your coding style here.

2.1 Com_number : Communication number [****]

Parent tier: NA
Variable: free text (consecutive numbers, no repeats allowed)
Duration: Whole communication (not exported)

Communication numbers are consecutive and serve to link all gestural signals within a communication – this includes those exchanged between the signaller and recipient.

The annotation should include the whole communication interaction between the signaller and recipient with a little space either side of it (for example it can include the approach before the first signal and then the start of the behavioural outcome afterwards). We don't export the duration of this annotation.

When there are several people coding on the same project it might be useful to have a code that indicates the coder-identity in the communication number. For example, we use a prequel set of two digits that indicates the coder and the site on each communication.

NB: Gestural exchanges
When coding gestural exchanges we add a letter suffix to the communication number to discriminate the gestures made by the original signaller (a) from those made by the original recipient when they gesture back (b). Specifically, for a gestural exchange we would create two ELAN files (one for the signaller and one for the recipient). In the signaller's ELAN file the communication number would be followed by "a" (e.g., 1001a), whereas in the recipient ELAN file the communication number would be followed by "b" (e.g., 1001b). All the gestures produced by the original signaller in the exchange would be coded under the 1001a com_number annotation, whereas all the gestures performed by the original recipient would be coded under the 1001b com_number annotation.

2.2 C_Clip_name [****]

Parent tier: Com_number
Tier info: free text; Symbolic Association
Duration: NA

This tier serves to link to the video file name you are coding from. In ELAN we copy paste the exact text from the video file name up to the clip number (so time, individuals etc. not included here).

2.3 C_Signaller [***]

Parent tier: Com_number
Tier info: controlled vocabulary (site specific vocabularies here); Symbolic Association
Duration: NA

Signaller name. New controlled vocabulary lists should be set up for each research site or group. Our central db has a list of all signallers but individual site-specific ELAN and Filemaker templates don't for convenience when coding to avoid long lists.

2.4 C_Recipient [***]

Parent tier: Com_number
Tier info: controlled vocabulary (site specific vocabularies here); Symbolic Association
Duration: NA

Same controlled vocabulary as for Signaller.

2.5 C_Sgn_prior_context : Signaller prior context [****]

Parent tier: Com number
Tier info: controlled vocabulary (here); Symbolic Association; definitions tbc.
Duration: NA

The behavioural context of the signaller prior to the signaller starting the communication. These contexts are considered from a species-specific perspective - so, for example, travel looks a little different in a gorilla than in a chimpanzee. Where an individual could be in more than one context use the following guidelines:

1. Social interactions take priority over non-social activities (e.g. *Agonistic* over *Feeding*)
2. Social interactions with individuals outside the mother-infant pair take priority over social interactions within the mother-infant pair (except if -> 3 is true within the mother-infant pair).
3. Where an individual exists in two social contexts with two separate individuals (e.g. nursing and grooming) then the context most relevant to the individual with whom they are communicating is marked.

2.6 C_Rcp_prior_context : Recipient prior context [****]

Parent tier: Com number
Tier info: controlled vocabulary (here); Symbolic Association; definitions tbc
Duration: NA

The behavioural context of the recipient prior to the signaller starting the communication. As for **2.5 C_Sgn_prior_context.**

2.7 C_Directed_to [***]

Parent tier: Com number
Tier info: controlled vocabulary (here); Symbolic Association
Duration: NA

Indicates whether or not there is evidence (usually via gaze or physical contact) that the gesture was directed towards a particular recipient. Note - this is decided holistically across the interaction, it does not have to be clear from the start of the interaction as gaze/check backs etc. may occur part way through.

Please use of the following options:

- ***No***: not directed
- ***Sev_pot***: The communication is potentially directed to several individuals.
- ***Sev_dir***: Several individuals are present but the communication is clearly directed to one specific individual.
- ***One_pot***: Potentially directed to one individual (no other individuals present).
- ***One_dir***: The communication is clearly directed towards one individual and the recipient is the only individual present. NB: this option should also be used when contact gestures are present regardless of how big the audience.

NB: for contact gestures this tier should be coded with the One_dir.

2.8 C_Goal [****]

Parent tier: Com number
Tier info: controlled vocabulary (here); Symbolic Association; definitions (here)

What was the behaviour that stops the signaller from signalling? Excludes behaviour that do not represent a plausible goal, for example: Aggressive attack. In some cases we will get this wrong or it will seem like you are coding something that you 'know' is not the goal, but we have to go with the behaviour we can see. It's always possible that there is miscommunication that the signaller settles for, or just gives up after etc. Outcomes describe what happened in a single case. Gesture meanings are only determined by the repeated pattern of use of a gesture, not from single cases. If you have intuitions about a gesture that are not based on the exact behaviour you see use the free text comment tiers e.g. **2.11**

2.9 C_Sgn_post_context : Signaller post context [****]

Parent tier: Com_number
Tier info: controlled vocabulary (here); Symbolic Association; definitions tbc
Duration: NA

The behavioural context of the signaller after the communication is completed. Same controlled vocabulary as 2.5 C_Sgn_prior_context.

Where the signaller both continues a behaviour and changes a behaviour, the change in behaviour gets the priority during coding. For example: a mother may be feeding, and then her infant requests nursing, which she starts but while continuing to feed. The mother's behaviour before the communication would be feeding, and after the communication would be nursing.

2.10 C_Rcp_post_context : Recipient post context [****]

Parent tier: Com number
Tier info: controlled vocabulary (here); Symbolic Association; definitions tbc
Duration: NA

The behavioural context of the recipient after the communication is completed. As for **2.9**, any change in behaviour takes priority over continued behaviour when describing the context afterwards.

2.11 C_Comment : Communication comment

Parent tier: Com_number
Tier info: Free text
Duration: NA

Any comment about the communication as a whole. For example, information about relevant or rare events that happened before, during, or after the communication but are not clearly seen in the clip e.g. inter-community encounter/stranger encounter earlier that day, or individual X has been ill/absent.

2.12 Exchange_part : Part of exchange [***]

Parent tier: NA
Tier info: controlled vocabulary (here); 1_1 1_2 2_2 1_3 2_3 3_3 1_4 2_4 (etc.)
Duration: mark whole gesture duration

Sequences or Bouts of gesturing may be exchanged by two signallers. Where there is no gestural response from the recipient all gestures have an Exchange number 1_1, where a recipient produces a gestural response the first signaller's gestures get an Exchange number of 1_2 and the original recipient's gestures are given an Exchange number of 2_2 when they are the signaller.

All gestures in a sequence (series of gestures separated by <1s) are given the same Exchange number, even where the recipient's response occurs half-way through. E.g. A(1) A(2) B(3) A(4) where A(1,2,4) are a sequence with less than one-second between gestures – are all 1_2 in a two-part Exchange. If there was more than a second between A(2) and A(4) then there would be three parts to this Exchange.

The exchange part can in principle be annotated independently of other tiers. However, it is logically linked to the Gesture_record1 and is not meant to be analysed independently. It gets the exact same annotation selection as the Gesture_record1 (it is treated as if this was its parent tier and symbolically associated). If there are overlapping gestures (i.e. annotations in the tiers "Gesture_record1" and "Gesture_record2", the "Exchange_part" selection should start with the start of the first gesture annotation and end with the end of the second gesture annotation, thus spanning the two gestures)

Mark the exchange parts that are relevant to the signaller being coded, this may mean skipping numbers (e.g. *1_3 3_3*; where *2_3* would be marked for the individual who is currently the recipient on the sheet where they are the signaller).

As for **Sequence_part1** and **Bout_part1**, mark as *Unclear* where there is any uncertainty about missing parts in the exchange.

2.13 G_rcp : Gesture action of recipient

Parent tier: NA
Tier info: controlled vocabulary (Gesture Action list)
Duration: mark whole gesture duration

Independent tier that allows for the annotation of the gesture actions made by the recipient. Note that this may not be the same as a gesture 'type' - something that can be constructed from an action, body part, and different modifiers such as repetition. The start time of the G_rcp annotation is used to calculate the latency from when the signaller starts to gesture to when the recipient starts to gesture in response.

While coding gestural exchanges it can be helpful to use the G_rcp tier to keep an overview over the different exchange parts (gets a bit tricky at times as both individuals are annotated successively, see **1.4** Coding gesture exchanges). If the start time of the first gesture is unclear or unknown then use the Time_ex (Time exclude) option to indicate that a gesture was used but we cannot measure the latency in this case. Do not leave this tier blank if no gesture is produced by recipient or recipient is out of sight, use *None* or *Unk* options.

2.14 Voc_rcp : Vocalisation of the recipient

Parent tier: NA
Tier info: controlled vocabulary (Species specific vocalisation lists)
Duration: mark whole duration

Independent tier that allows for the annotation of the vocalisations of the recipient (with start and end time). The start time of the Voc_rcp annotation is used to calculate the latency from when the signaller started to gesture to when the recipient produces any vocalisation.

If the start time of the first vocalisation is unclear or unknown then use the *Time_ex* option to mark that a vocalisation was used but we cannot measure the latency in this case. If the recipient was already vocalising before the signaller started to gesture and the recipient continued the same vocalisation after the gesture started use *Voc_cont* and start the annotation from the first unit of vocalisation that starts after the gesture: For example, the next grunt starting after the gesture if the recipient was grunting.

Do not leave this tier blank if no gesture is produced by a recipient or the recipient is out of sight (use *None* or *Unk* options).

Whether there was a vocalisation by the recipient but the type of vocalisation is not noted in this tier (e.g. because using Time_ex or Voc_cont) its type can be noted under **BP1_Rcp_Voc**, a child tier of the **Bout_part**.

2.15 FE_rcp : Facial expression of recipient

Parent tier: NA
Tier info: controlled vocabulary (Species specific facial expression lists)
Duration: mark whole duration

Independent tier that allows for annotation of the facial expressions of the recipient (with start and end time). This tier is not currently coded by our lab but is available for use as needed.

Whether there was a facial expression by the recipient or not, as well as its type, is currently noted under **Rcp_FE**, a child tier of the **Bout_part**, but we are unable to capture the start and end time of the facial expressions with this method.

2.16 Outcome : Start of communication outcome
Parent tier: NA
Tier info: controlled vocabulary (here)
Duration: mark start time

This tier marks the end of what we typically consider response waiting and allows the duration of several different latencies to be measured. We mark who is responsible (signaller or recipient) for ending the communication. Note that the recipient must have started to fulfil the goal - not just the start of their movement (e.g. they must have started to groom, have climbed on etc. not just approached in order to start grooming). There is only one start of behaviour change annotation per exchange part.

We only mark this annotation at the end of a communication, because latencies between other sections (e.g. between bout parts, or exchange parts) can be measured from the points between the end and start of different points in the gestures. Even if we have to mark an annotation here as *Exclude* because we can't measure the complete timing, it is OK to mark response waiting as present in the relevant Bout_part1 tier as long as there was at least 1-sec of visible response waiting.

For affiliative contact gestures this tier is coded with an annotation starting at the end of the MAU.

If it is unclear if the goal is met by the change in behaviour that stops response waiting mark *Exclude* or *NV*

2.17 BhvCh1: Behavioural change 1
Parent tier: NA
Tier info: controlled vocabulary (here)
Duration: mark start time

Use this tier to mark the first change in behaviour by the recipient following the start of the first gesture in a communication. We want to be able to do so from the first point of a reaction - e.g. starting to move closer, or to fulfil a goal because it will allow us to explore how often recipients are reacting before a gesture is over or even fully recognisable, or how and when persistence happens, etc. Note that for calculating latency we will often take the

MAU as the measure from point, so changes that occur prior to this are going to be 'negative' latencies.

Here we mark only behavioural changes (no vocalisations/gestures). We mark gestures in the G_rcp tier and vocalisations in the Voc_rcp tier.

2.18 BhvCh2: Behavioural change 2

Parent tier: NA
Tier info: controlled vocabulary (here)
Duration: mark start time

Use this tier to mark the first change in behaviour **after** the MAU of the ***last*** gesture in a communication. Here it would not be possible to have negative latency values (where calculated from the MAU), if there are no further changes after the MAU of the last gesture we mark it as none. If there is only one change in behaviour and it occurs after the last MAU then copy and paste the same information as **BhvCh1** tier.

2.19 Vocalisation_record: Vocalisation of the Signaller

Parent tier: NA
Tier info: controlled vocabulary (Species specific vocalisation lists)

Independent tier that allows for the annotation of the vocalisations by the signaller (with start and end time). This tier is currently not coded but available for later use if needed.

Note: Whether there was a vocalisation of the signaller or not as well as its type is currently noted under **G1_Sgn_Voc**, a child tier of the **Gesture_record1**. Currently we don't code the start and end time of these vocalisations.

2.20 Facial_expression_record: Facial expression of Signaller

Parent tier: NA
Tier info: controlled vocabulary (Species specific facial expression lists)

Independent tier that allows for the annotation of the facial expressions by the signaller (with start and end time). This tier is currently not coded but available for later use if needed.

Note: Whether there was a facial expression of the signaller or not as well as its type is currently noted under **G1_Sgn_FE**, a child tier of the **Gesture_record1**. Currently we don't code the start and end time of these facial expressions.

2.21 Gesture_record1 : Gesture action signaller [*]

Parent tier: NA
Variable: controlled vocabulary (here)

Duration: mark whole

The gestural action produced by the signaller (see Gesture Action list with definitions). The Gesture_record1 tier is an independent tier and the parent tier of all the variables described under points **2.22 - 2.45**.

The annotation marks the duration of each individual gesture token, extending from the very start to the very end of the gesture's movement (*i.e.,* including the movement towards gesture production - preparation - and the movement back to a neutral position - recovery). Definitions for the start and end of each gesture action can be found here.

Unclear is used where the gesture action was seen but the type couldn't be determined, this is also used for apparently 'failed' gestures. For example, the signaller approaches to grab, but the recipient moves away and the gesture isn't completed. In this case, it would not be possible to distinguish, for example, a *Grab* from a *Grab hold* here and the recipient did not experience physical contact, so the gesture action is marked as *Unclear*.

Unknown on this list is only used in **2.56 BP1_Rcp_G** whenever it is unknown whether or not the recipient made a gestural response. Whenever there is a gestural response but the gesture type is unknown it should be marked as Unclear. There would be no annotation if it was 'unknown' whether the signaller produced a gesture.

Colons in Gesture_action names (full names not name codes, e.g. Bite: kiss) indicate that we have recently split a gesture into separate forms, but it is unclear whether these action forms are distinguished by the apes and should be checked for potential lumping during later analysis.

Notes:
Repetition gestures (Hitting, Stomping, Tapping, etc.) are distinguished from multiple single gestures (Hit, Stomp, Tap, etc.) by rhythmic repetition. Where this is asynchronous it might lead to exaggerated sequence length, but using this measure provides a consistent coding point. Thus, use rhythmic repetition to decide whether or not to keep the same annotation but mark it with repetitions, or to mark several annotations (Hit-r2 vs Hit-r1 + Hit-r1).

Directed gestures = behaviour is directed to the body part presented afterwards. Note that it is impossible to distinguish failed directed gestures from the non-directed forms of those gestures (e.g. failed Push directed from a 'normal' Push). This needs to be taken into account when analysing.

All Spin gesture actions should include a full 360 degrees of turn, as this indicates that the movement is more than would be required for locomotion

2.22 G1_Rec_number : Record number [*]
Parent tier: Gesture_record1
Variable: free text
Duration: NA

Record number is consecutive and is a unique numeric identifier for every gesture case. The numbers should be consecutive within an individual coders data set.

2.23 G1_Sequence_part : Part of sequence [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here); 1_1 1_2 2_2 1_3 2_3 3_3 1_4 2_4 etc.
Duration: NA

All gestures that are produced by one individual with **< 1 second** of time between the MAU_stop of the first gesture record (= end point of the annotation of the MAU_duration) and the MAU_start of the second gesture record (= start point of the annotation of the second Gesture_record) are considered to be part of the same sequence (same Bout_part) and are given consecutive sequence numbers (e.g. 1_2 and 2_2 : two gestures in a sequence, same Bout_part 1_1). Any two gestures with **> 1 second** of time between the MAU_stop of the first gesture record (= end point of the annotation of the MAU_duration) and the MAU_start the second gesture record (= start of the annotation of the second Gesture_record) are considered as being part of different sequences (two single gestures, each given a sequence number 1_1, each in a different Bout_part 1_2 2_2).

Some gestures may be coded where their sequence number is unclear (for example where gesturing started before the clip and sufficient information was not observed or dictated to know if this was the first gesture). If this is the case mark as Unclear (do not put 1_1 for where a single gesture is captured on video but might have been, for example 2_2, in a sequence where earlier gestures were missed).

2.24 G1_Rep_count: Repetition count [*]

Parent tier: Gesture_record1
Tier info: free text; Symbolic Association
Duration: NA

NV (Not Valid) : All gestures that can only have a single action unit because they either a) only exist in single form, and the presence of a second action unit would be coded as a second gesture (e.g. Jump), or b) because they are the single-specific form of a gesture (e.g. Hit, Stomp, etc.). Note two exceptions: Dangle and Rake, where it might be possible to repeat the action unit are coded as NV because it would be too difficult/subjective to discriminate the number of repetitions.

1-n : All gestures where it is possible to have a repetition of the action unit, we count the number of repetitions. Note that for some of the repetition specific forms (e.g. Hitting, Stomping, etc. it would be impossible to have less than 2 action units).

Note that this is not necessarily the same as the number of repetitions of the Minimum Action Unit (which itself might include repeated contacts or movements, for example hitting obj/ground MAU = 3 audible beats; BLS MAU = two complete scratch actions). Here we count the number of times an action unit is repeated, e.g. how many contacts in a contact gesture, how many back and forths in an object shake, how many scratches in a BLS.

2+ : notation where we are certain that the action unit was repeated but we are not sure how many times

Unk : it was unknown if it was repeated or not (or where it has not been possible to code repetition for other reasons).

A full list of when and how to code repetition for each action type can be found in the Gesture Action list.

2.25 G1_MAU_duration: Minimum action unit and gesture action duration [*]
Parent tier: Gesture_record1
Variable: controlled vocabulary (here); Included In
Duration: Mark one or two: The first include the whole MAU and second from the end of MAU to end of the gesture action (unless MAU = gesture action)

Two separate aspects of the duration of the gesture are added here. Not all gestures will have both. It is probably useful to read the terminology tab on this sheet before continuing. For more theoretical background on why we code different gesture phases we recommend reading the manuscript Grund et al. 2023.

All gestures have a Minimum Action Unit (MAU). The MAU describes the smallest possible section of the gesture movement in which the recipient can sufficiently distinguish the specific gesture action being produced. The MAU annotation has the same start as the gesture action marked in the Gesture_record and ends at the defined MAU_stop of the given gesture action type (MAU_stops are defined for each gesture action type and definitions are available here). The MAU_stop is the point from which we assume “Response waiting” starts.

Annotate either as MAU_in (MAU_include) for where the whole MAU is visible and the duration should be included in analysis. Annotate as MAU_ex where the start or end is unclear and the duration should be excluded from analysis.

- MAU_in : where the end of the MAU is visible and the duration can be included in analyses
- MAU_ex: the end of the MAU is unclear and the duration should be excluded from analyses.

If the MAU matches the full gesture action use:
- MAU_GA_in: where the end of the MAU is visible and timing can be included in analyses
- MAU_GA_ex: where the end of the MAU is unclear and timing should be excluded from analysis

For some gesture types and in some gesture cases the MAU is the same as the full gesture action, but in most cases the gesture continues after the MAU_stop. For example, after the MAU a gesture might be held in place, or the action repeated etc. After the MAU we indicate

a second annotation within the MAU_duration tier (still the same Gesture_record), the Gesture Action stop (GA_stop). This annotation starts at the MAU_stop and ends at the stopping point of the Gesture Action (definitions of GA_stops here) covering the rest of the gesture movement. The GA annotation does not include the period of time during which the body is moving back to a resting position (the recovery). For some gesture cases the GA_stop point will be at the same time as the stopping point of the full gesture record.

The MAU and GA together indicate the amount of time the signaller has invested in producing the gesture, but do not include the recovery (extra time taken by the movement back to rest etc.).

Annotate:

- **GA_stop_in** for where the end of the GA is visible and the duration can be included in analyses.
- **GA_stop_ex** for where the end of the GA is unclear and the duration should be excluded from analyses.

2.26 G1_Body_sgn : Body part signaller [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

The body part with which the signaller signals. Where a gesture makes contact with the Signaller's own body, note the limb/body part used to produce the action here, and the body part contacted in **2.29** (e.g. in a Big Loud Scratch, note the hand scratching here and the body part scratched in **2.29**). Sometimes there are potentially two options for the body part (e.g. in a BLS you could either choose the hand as the enacting body part or the fingers). Guidelines to decide on which body parts are typically used with different gesture actions are in the definitions here - but note that these are intended to be guidelines and not restrictive.

*Usually the underscore is used to indicate a subtype of a broader category (e.g. the "wrist" as part of the broader category "Hand" in "Hand_wrist"). "Hand_foot" is an exception to that, here the underscore indicates that hand and foot are both used to perform the gesture action. Note that we cannot differentiate if 2hands/1foot or 1hand/1foot etc. are involved in the gesture action.

2.27 G1_Sgn_lat : Signaller laterality [*]

Parent tier: G1_Body_sgn
Variable: controlled vocabulary (here)
Duration: NA

Laterality of body part that produces the signaller's gesture. Only code *L* or *R* for body parts we have two of, i.e. we don't code *L* on Head.All other body parts are *NA*.

2.28 G1_Lat_bim : Laterality bimanual gestures [*]

Parent tier: G1_Body_sgn
Variable: controlled vocabulary (here)
Duration: NA

Where both hands, feet, arms, legs are involved (so where Sgn_lat is either “B” or “Ba”, compare above) how is the gesture action performed? Where it is Ba chose the limb that makes the first action (e.g. the first hit, stomp, etc.) Note that this is not always straightforward but can also use Equally or Unknown if unsure.

2.29 G1_Body_con : Body part contact [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

The body part of the recipient for contact gestures, or of the signaller for gestures where the signaller makes contact with their own body (e.g. Big Loud Scratch, Chest beat).

2.30 G1_Con_Lat : Contact laterality [*]

Parent tier: G1_Body_con
Variable: controlled vocabulary (here)
Duration: NA

Laterality of the body part that receives the gesture (as **2.27**: recipient body part where contact gesture; or signaller body part where signaller makes contact with their own body). Only code *L* or *R* for body parts we have two of, i.e. we don’t code *L* on Head when there is a touch on the left side, but we do code *L* on Foot.

2.31 G1_Object_contact: Object contact [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Indicates which object the signaller makes contact with when performing a gesture that involves the contact of a substrate (eg., Hit object). These objects are not manipulated in any way to produce the gesture but during the gesture the signaller makes contact with the object using the body part or another object used to produce the gesture (check gesture action list for relevant gestures).

2.32 G1_with_Object: Object used [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Indicate whether an object was manipulated to produce the gesture, this doesn't include contact with a substrate (for example Hit object), only gestures in which an object forms part of the gesture action (check gesture action list for relevant gesture types).

2.33 G1_Audible: Audible information [**]

Parent tier: Gesture_record1
Tier info: controlled vocabulary (here)
Duration: NA

Does the gesture action produce any audible information? This is marked for all audible gestures or gestures where the gesture action produces a noise. We do not mark this tier for audible information that is not the result of the gesture action. For example, if during a fling gesture the signaller's arm (used for the gesture) makes audible contact with an object we mark *Yes*, however, if the signaller is moving through dry leaves while gesturing and the leaves make a noise but the gesture itself does not (e.g. fling where arm makes no contact with any object) then we mark *No.*

2.34 G1_Direct_G : Directionality of gesture [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here); Symbolic Association
Duration: NA

Taken from the perspective of the signaller, e.g. *Self* indicates the signaller. Directionality of the gesture, where the gesture is an extension of the body indicates the direction that the gesture is produced. Any gesture can be directed, and we code this quite 'generously' - for example as well as more 'obvious' cases where a reach is extended towards an object or location of interest, a bounce or a punch ground, could be performed above a location of interest. We can only recognise directionality in successful gestures (based on behaviour that happens afterwards) so many gestures, including things like Poke, Bite, etc. could be directed. That means gestures can be 'apparently undirected' but there is no 'not applicable' option.

Where the gesture makes contact with the signaller's own body, indicate Directionality *Self* if it results in physical contact by the recipient, and Directionality *Self: location*, if it results in physical contact by the recipient at the point where the gesture made some contact with.

So, for example, Big Loud Scratch that results in grooming (somewhere other than the scratch location) = Directionality *Self*; if it results in grooming at the location that the scratch made some contact with = Directionality *Self: location.*

Where an individual is already on the signaller (e.g. infant on mother), and the signaller indicates they should move to a new position, use Directionality *Self: location.*

Where the signaller makes contact with the recipient's body and the recipient reacts by exposing a specific location on their body (e.g. directed push to position recipient during grooming) use *Location.*

Big Loud Scratch gestures – when there is more than one Big Loud Scratch in a sequence and the location of the BLSs moves none of them can be marked as Directionality *Self: location*, even where grooming matches one of the locations. Directionality *Self: location* only indicated where the grooming happens on a body part that the Big Loud Scratch made some contact with. Big Loud Scratches that successfully initiate grooming can be marked as Directionality *Self*.

2.35 G1_Emphasis [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Used to indicate whether or not a gesture is produced with more or less energy invested into it than previous examples of the same gesture type. Emphasis can only be coded where another gesture action of the same type was produced earlier in the same communication (not across communication). Increased emphasis would be indicated by a gesture that was larger in movement size and/or faster in speed. Note that if the same gesture action is performed by another limb of a similar type (e.g. left foot -> right foot; or left hand -> right foot) then you can record emphasis if this is clearly different.

An exception is the 'jiggle' code which can be used when this is the only gesture of its type in the communication, Here the fingers or limb is jiggled or wiggled while the limb is held in obvious tension. Typical examples include wiggling the fingers or small rapid shaking of the limb during a *Reach* gesture, or jiggling the fingers during a *Touch* or *Touch: long other* gesture.

Not Applicable means that this was either the first, or only gesture of that type in the communication by the same signaller, so it was not possible to assess emphasis.

2.36 G1_Flexion [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Only coded for Hand-Arm, Foot-Leg gestures, all other body, head, etc. gestures = NV (Not Valid). Note that while descriptions below are based on hand-arm parts these are considered to apply to foot-leg in the same way.

Only coded for silent-visual-non-contact gestures. For both contact and audible gestures, the shape of the hand/foot is partly determined by the location and orientation of the surface/partner.

Where part of the limb is held or moved into flexion while the gesture is produced, for example when curling the fingers, bending the wrist, etc. flexion (and orientation) are marked as they appear at the point of the completed gesture stroke (end of the MAU), they

do not change if the hand position changes while a gesture is held/repeated. Mark whether the fingers, the wrist, and/or the elbow are flexed.

Finger flexion is marked when the knuckles or fist finger joint is held at 90 or more (taken as a measure of non-natural position across most actions).

Wrist flexion is marked when the wrist is held in a 'non-natural' position.

When the hand is held palm down, this would include a wrist angle of 90 degrees,
When the hand is held palm sideways it has to be held at more than 45 degrees in either direction.
When the palm is held facing upwards it has to be held at more than 45 degrees in either direction.

Elbow flexion is marked when the elbow is held at more than 45 degrees.

Note that for different gesture actions there may be variation in what makes a 'natural' position. But this should be consistent across either a gesture action or species etc. so for example, the majority of *Raise* arm/hand gestures will likely have wrist flexion, and it is those that don't that indicate a 'choice' to vary.

For individuals with an injury to that limb, which may impact their opportunity to use flexion mark as 'injury'.

2.37 G1_Orientation [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Only coded for Hand-Arm, Foot-Leg gestures, all other body, head, etc. gestures = NV.

Only coded for silent-visual-non-contact gestures. For both contact and audible gestures, the shape of the hand/foot is partly determined by the location and orientation of the surface/partner.

The orientation in which the hand, arm, foot, leg is held once the gesture stroke is completed (i.e. end of MAU). Orientation and Flexion are coded independently of each other, so - for example - changing the Flexion of the wrist in a Reach action so that the palm orientation doesn't change the orientation of the whole gesture action

Orientation is typically determined from the position of the palm or foot - mark where this is **sideways, down, or up**. If you let your arm hang naturally by your side then the palm is sideways, it stays sideways as you bring it up to a reach, and again to an arm raise. If down - then the palm is facing the wall behind you when the arm is down, facing the ground in a reach, and facing the recipient in a raise. etc. Note: that it would be very unusual/uncomfortable for the foot orientation to be 'up'.

2.38 G1_Sgn_Voc : Signaller Vocalisation [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (Species specific vocalisation lists)
Duration: NA

Signaller vocalisation is marked **per gesture**. Does the signaller vocalize during the gesture action? If so which type (from species specific list, see below).

2.39 G1_Sgn_FE : Signaller Facial Expression [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (Species specific facial expression lists)
Duration: NA

Signaller facial expression is marked **per gesture**. Does the signaller produce a facial expression during the gesture? If so, which type (from species specific list). Note that where there is an pant-hoot, whimper, or scream vocalisation the accompanying facial movements are marked here as Vocalising. These are only marked as facial expressions where the face is moved into this position but there is no vocalisation.

2.40 G1_Distance_rcp : Distance to recipient [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Distance to recipient estimated between the signaller and recipient in meters. Distances are measured from the approximate centre of the body in each case, rather than the closest point between the two individuals.

Distance to recipient is measured immediately at the start of the gesture action, but before the gesture is completed. Where the signaller has been in physical contact with the recipient, this is marked as being in contact, even if the signaller has to briefly move apart in order to be able to produce the gesture.

2.41 G1_Rcp_VA : Recipient visual attention [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Does the recipient have a view of the gesture action while it is being produced, including while it's being held in place. Coded as for the coder, e.g. if you can see that the recipient

has visual attention, not whether or not you think the signaller can see that the recipient has attention (although in most cases over short distances it is likely to be the same).

VA clear requires that either eye gaze or head tracking of action, indicates you are certain that the recipient saw the gesture action.

VA possible is where the head direction makes it very likely that the recipient saw the gesture action clearly (but, for example, eye gaze couldn't be checked, so they might have their eyes closed).

VA 90 is coded generously and can be used if there is any chance that the recipient could visually perceive the gesture action even peripherally.

Out of sight is for where the recipient can not be seen, for example in long-distance drumming, pant-hoot, chest-beat. It is relatively rare to code for intentional gestures because of the challenge of determining intentional use across long-distance.

2.42 G1_Effectiveness : Effectiveness of gesture [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Indicates whether the gesture action may be able to achieve its own goal or is congruent with the goal.

For example: a push may effectively move the recipient whether or not they want to be moved (Yes, in which case this is not a gesture and should usually be excluded). Note that this doesn't include all mechanical actions - just those where the physical movement could achieve the intended goal. Or a push may be a gesture that indicates the direction of movement, but isn't used with any effective force so is congruent with the desired outcome but can't achieve the goal (No_mech i.e. contact is made but it is mechanically ineffective).

One way to distinguish whether contact gestures are mechanically effective or ineffective is to see whether or not the recipient starts or continues to move themselves after the force is removed. E.g. I give a short push, and you move only after I stop applying the push force.

The shape of some gestures' action may be congruent with the desired goal, for example I fling my hand towards you so that you move backwards away from me in the direction indicated, but I don't use any force (No_cng). Finally there may be nothing about the shape of the gesture action that matches the subsequent behaviour, for example pirouette to ask someone to stop, or arm raise to ask someone to move closer (No).

A subset of gestures are themselves very similar to the movements and actions of the goal they want to achieve, for example: embrace or touch can be used as gestures and as actions to achieve affiliation. Kick can be used as a gesture or an action in an agonistic way etc.

These should be marked with caution, only where it is clear that they are used as a gesture (Other).

2.43 G1_Dur_ana : Gesture duration include in analysis [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Indicates whether the gesture duration should be included in the analysis, this refers to the duration of the full gesture action rather than the MAU or G-stop measures described above.

2.44 G1_Cod_status : Gesture coding status [*]

Parent tier: Gesture_record1
Variable: controlled vocabulary (here)
Duration: NA

Indicates whether or not the gesture coding for this communication is complete, useful to mark if you have to stop coding part way through and may not remember if you were finished when coming back to it.

2.45 G1_Comment_rec: Comments on gesture action

Parent tier: Gesture_record1
Variable: free text
Duration: NA

Place to write any comments on the gesture token. For example, if something happens off camera during the gesture or if unsure about one of the tiers.

2.46 Bout_part1 : Part of Bout [**]

Parent tier: NA
Tier info: controlled vocabulary (here); 1_1, 1_2, 2_2, 1_3, 2_3, 3_3, 1_4, 2_4, (etc.)
Duration: Whole bout; from the start of first gesture in sequence to the end of last gesture

Single gestures or gesture sequences may be produced multiple times by a signaller when they show persistence. Persistence is indicated by additional gesture(s) that follow after at least 1 second of response waiting after the original gesture(s). All gestures with < 1 second in between them are in the same sequence and get the same Bout number. For example: A signaller produces a gesture A, waits, and then produces two more gestures B and C in a sequence. The first single-gesture (A) sequence would get the Bout number 1_2 and the two gestures (B,C) in the second sequence would get the same Bout number (2_2, 2_2). For more details see also description of “G1_Sequence_part” (section **2.23**).

Note that the 1 second is measured from the end of the MAU, as this is when we allow response-waiting to start. In some cases where gesture duration is very extended we may

have more than 1 second between the end of the MAU and the end of the gesture, even though the start of a second gesture is less than one second after the end of a gesture action. For example: An object shake gesture - starts at 1.5s, MAU ends 1.8s, gesture duration ends 3.8 seconds, and a stomp is added at 3.7 seconds (or even at 3.9 seconds). Essentially you have a gesture shortly after, or even included within the duration of an earlier but very long gesture. In this case the second gesture is still marked as starting a new bout part (because response waiting since the end of the MAU is over 1 second before the next gesture started). Common sense exclusions are where response waiting could not have reasonably occurred even though 1 second has passed since the MAU ended, e.g. in pirouetting, or a galloping display the gestures could all be included in the same bout part.

All gestures which have a Bout number different to 1_1 are associated with persistence (see **2.59**) having been marked for an earlier/subsequent gesture. As for Sequence number, mark Bout number as Unk if there is any uncertainty about earlier/later missing gestures or gesture sequences.

Note that the duration of bout part annotations is not extracted - only the annotation (e.g. 1_2). So in a simple sequence it is possible to use a single bout part annotation that covers all the gestures within a bout part, rather than replicating the same information under each gesture annotation. However, in complicated sequences where there are overlapping gestures and bout parts it may be more convenient to mark a separate bout part annotation for each gesture annotation. Note that as there is a set of Gesture_record2 tiers for overlapping gestures, there is a set of Bout_part2 tiers for overlapping bout parts.

2.47 BP1_Sgn_location : Signaller location [**]

Parent tier: Bout_part1
Tier info: controlled vocabulary (here)
Duration: NA

Provides an indication of the terrain on which the signaller is located at the time of producing the gesture. Designed so that we can consider the impact of the immediate physical location on the choice of signal.

2.48 BP1_Sgn_arousal : Signaller arousal [**]

Parent tier: Bout_part1
Variable: controlled vocabulary (here)
Duration: NA

Any sign of physical arousal from the signaller immediately prior to or at the time of producing the gesture(s). These may include piloerection, strongly affect-associated signals as e.g. cry, screams, bare-teeth face. There are certain events typically associated with high arousal (e.g. intercommunity encounters).

At present we don’t code mild affect-associated signals like whimper or pant because we can not reliably assess these across videos and species. We don’t include penile erection because it is too unreliable to code (both across species, and across contexts).

2.49 BP1_Gaze before : Signaller gaze before sequence [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (here)
Duration: NA

Did the signaller look in the direction of the recipient before starting the Sequence.

2.50 BP1_Gaze_during : Signaller gaze during sequence [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (here)
Duration: NA

Did the signaller look in the direction of the recipient while producing the Sequence.

2.51 BP1_Rcp_location: Recipient location [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (here)
Duration: NA

Provides an indication of the terrain on which the recipient is located at the time of receiving the gesture. As for **2.47 BP1_Sgn_location**.

2.51 BP1_Rcp_arousal : Recipient arousal [**]
Parent tier: Bout_part1
Variable: controlled vocabulary (here)
Duration: NA

As for **2.48** but any signal of physical arousal from the recipient immediately prior to or at the time of the signaller producing the signal.

2.52 BP1_Vis_during : Visibility during [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (here)
Duration: NA

This variable indicates whether the environment (physical space) between the individuals (Signaller, Recipient) allows both to perceive each other clearly or not. It is coded as for each other (i.e. from the perspective of the two interactants not from the coder's perspective) and it is coded rather broadly for the whole body not the body part performing a gesture action (i.e. it does not ask whether the individual can perceive the specific body part producing a gesture action but rather whether it has clear visibility on the other individual as a whole).

2.54 BP1_Rcp_Voc : Recipient Vocalisation [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (Species specific vocalisation lists)
Duration: NA

This variable is coded **per sequence** and indicates whether the recipient produced a vocalisation at some point during the sequence and which vocal signal he used. Any vocalisation that is produced from the end of the first MAU of the sequence until the next bout (if a vocalisation has started before the MAU_end – some kind of complete unit of that vocalisation must still happen after the MAU_end, e.g. a bark, a section of whimper, not a tailing off of a pant-hoot let down or similar).

2.55 BP1_Rcp_FE : Recipient Facial Expression [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (Species specific facial expression lists)
Duration: NA

This variable is coded **per sequence** and indicates whether the recipient produced a facial expression at some point during the sequence and which facial signal he used.

2.56 BP1_Rcp_G : Recipient gesture [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (here)
DurationL NA

This variable is coded **per sequence** and indicates whether the recipient produced a gesture at some point during/after the sequence and which gestural action he used. Where the recipient produced more than one gesture in response, mark the first one used.

2.57 BP1_Rcp_reaction : Recipient reaction [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (here)
Duration: NA

This variable is coded **per sequence** and indicates whether the recipient reacted at some point during/after the sequence to the signaller's gestures. Choose the type of reaction from the reactions list (same as Goals list). Choose the reaction "Stay same" for the case that the recipient did not change his behaviour.

2.58 BP1_Response_waiting: Response Waiting [**]
Parent tier: Bout_part1
Tier info: controlled vocabulary (here)

Duration: NA

Is there any evidence of response waiting by the signaller (after the multi- or single- gesture sequence)? Response waiting must last at least **1 second** to count as RW = yes and should be accompanied by some visual monitoring or checking-back to the recipient (with the exception of contact gestures for which visual monitoring is not a requirement). Signaller can not start a new behaviour or activity while Response waiting (*Note: there is an exception for Gorillas, who are allowed to put food in their mouths and eat (but not start processing food) while response waiting*).

To determine whether there is "Response waiting" or not we look at the time elapsed between the MAU_stop of the (last) Gesture record (i.e. once the smallest possible section of the gesture movement in which the recipient can sufficiently distinguish the specific gesture action has been produced) and the start of a) following gesture(s), b) other behavioural changes of the signaller that indicate they terminated the communication or c) the recipient satisfying the signaller's goal. MAU_stops are described and defined here.

2.59 BP1_Persistence [**]

Parent tier: Bout_part1
Tier info: controlled vocabulary (here)

Did the signaller produce further gestures after response waiting? These gestures should be separated by at least 1 second from the previous gesture(s) and would be coded as a new sequence. Note that in some cases rather than choosing not to produce more gestures the signaller may be Unable to, for example: their mother arrives and carries them off, or they are displaced by another individual.

For example: aaa : *rw* : bb : *rw* : ccc

Gesture sequences aaa and bb, persistence = yes (more gestures are produced after response waiting, *rw*). Gesture sequence ccc, persistence = no (no more gestures are produced).

2.60 BP1_Goal_met: Goal met [**]

Parent tier: Bout_part1
Tier info: controlled vocabulary (here) (*Yes, Partial, No, Unknown, Yes_forced)*
Duration: NA

Did the signaller stop signalling after the production of an Apparently Satisfactory Response by the recipient? If *Yes,* no further gestures should be produced, if *No* or Unknown further gestures may or may not be produced. *Partial* includes where the recipient changes their behaviour in a way that is congruent with the eventual behaviour that stops the signaller from signalling, but does not fully complete it. For example: this may include Pays Attention, Moves closer, where the successful Goal was: Climb on me, or Groom). It could also include Stops behaviour, where the successful Goal was: Move away.

Yes_forced is for where the signaler signals, there is no or an uninterested response and then the signaler essentially achieves their own goal forcefully but that outcome is tolerated – for example asks to play, there is no clear response, but they then pounces on the recipient and forces play, or asks to climb on, refused/ignored, and then forces a climb on.

3. Saving and naming coded communications (ELAN annotation files, .eaf)

You've just finished coding a communication, yey! Now you must save it. Click the save button and name the file after the clip file name up to the clip number (so without the time and ID info) and add the respective Communication number at the end (with no spaces in between). We add the Com_number because there may be more than one communication for a given clip and it is easier to organise the coded files if each communication gets its own .eaf file.

Here is an example ELAN annotation file (.eaf) name (Communication number C200067):

- BwindiCG20190829BTClip25C200067

Note, where there is an exchange (so there is more than one Communication file) use letters to indicate this e.g.

- BwindiCG20190829BTClip25C200067a
- BwindiCG20190829BTClip25C200067b

Whenever you save an Elan annotation file (.eaf) another file-type (.psfx) with the same name is created (automatically). You cannot open the .pfsx file but it saves information about the formatting of the corresponding .eaf file. Note, when you would like to change ELAN file names you have to rename the corresponding .pfsx too in order to save the formatting.

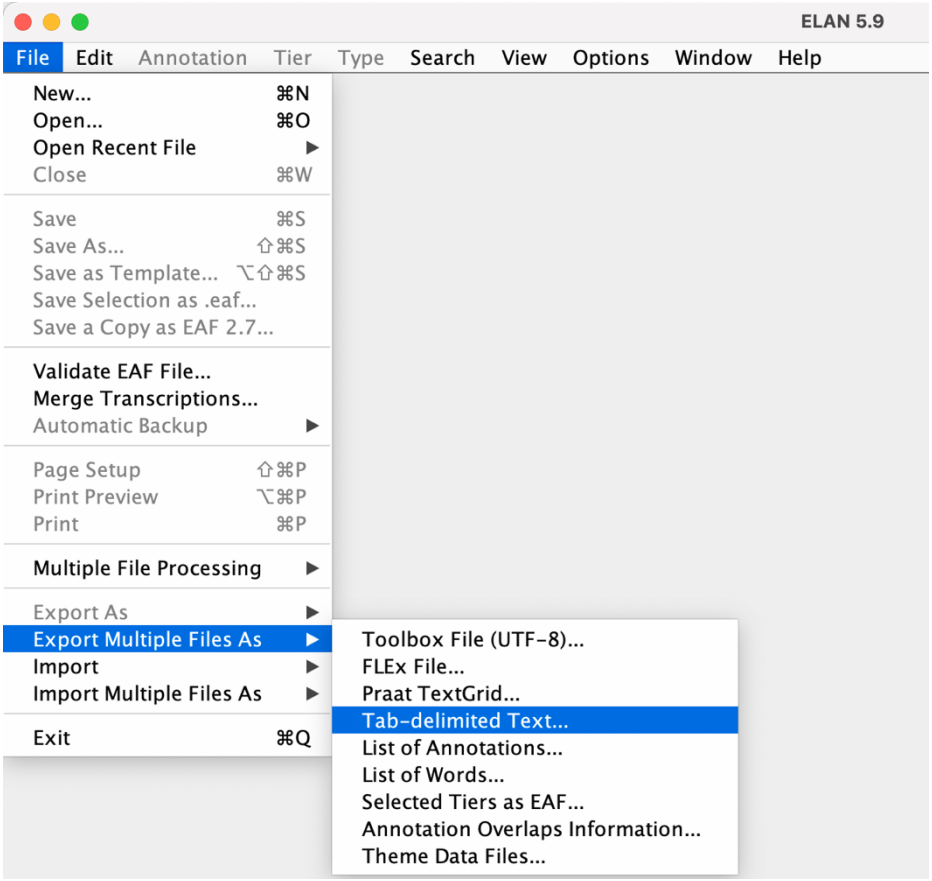
4. Extracting the coded 'raw' data

4.1 Exporting ELAN files (.eaf files)

The Gestural Origins template has several tiers with independent timings and the ELAN export of .eaf files coded with it can look messy. Alex Mielke (GitHub: https://github.com/AlexMielke1988/) has programmed a function in R (R Core Team, 2021) that transforms the rather messy output into a correct, database friendly format, a .csv file where each row contains all the information of the respective coded communication (see how one communication looks like in the example .csv here). If you use Filemaker, you can check our database template here.

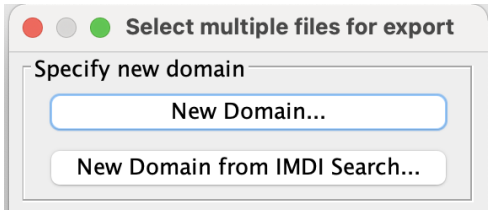
Now to the export. The only steps to follow are the following:

- Open ELAN
- Click File → Export Multiple Files as → Tab-delimited Text…



This will open a pop-up window. In this pop-up window you will have to select the domain (files) that you want to process. You can either:

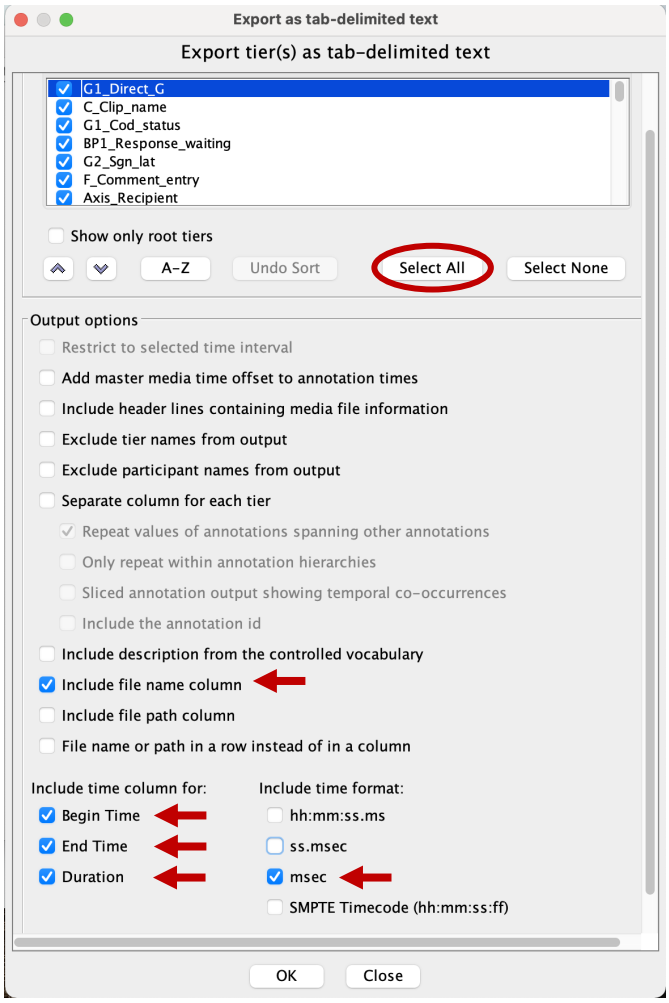
- Create a new domain
- Use an existing one



If you are creating a new domain, ELAN will ask you to upload the ELAN files or folder to the domain. Select all your coded .eaf files and click them over to the right. Then click 'OK'. Name the new domain you have just created.

Once you have selected or created the domain, ELAN will open a second pop-up window with all the tiers from your .eaf files (see below).

- Click 'Select All' below the Tier names
- Unclick everything except: 'Include file name column', 'Begin Time', 'End Time', 'Duration', 'Msec'
- Click 'OK' and save the file in any path as a .csv file



Done. See next step on how to use the extract_ELAN.R function

4.2 Transforming the ELAN export .csv in R

A link to the extract_ELAN.R function is available here. You just have to call the function and enter the path where you saved the ELAN export .csv.

extract_ELAN (path = '/Users/cvg1/Elan_export.csv')

There are more details documented in the function itself. The data list can be saved as a .csv-file and can then either be directly used for analysis or loaded into the database (e.g. filemaker, see here).

5. Analysing gestural data

Here we describe how to code and collect gestural 'raw' data with the Gestural Origins bottom-up approach. Ideas and worked examples of how to analyse data generated this way can be found in the coding methods paper itself (Grund et al., 2023), in Mielke et al. (in prep) and future publications of the Wild Minds lab (www.wildminds.ac.uk)

6. Version remarks and acknowledgements

Our coding approach builds on decades of (great ape) gesture research. This specific scheme is the product of several years and involved the contribution of several people in the Wild Minds lab (in alphabetical order: Gal Badihi, Kirsty Graham, Charlotte Grund, Catherine Hobaiter, Alex Mielke, Daniela Rodrigues, Alexandra Safrighyn). The scheme was finalised in 2023 and is part of the online supplementary material of Grund et al., 2023 'GesturalOrigins: A bottom-up framework for establishing systematic ape gesture data across species'. We thank Emilie Genty for her initial advice to use ELAN as a means to implement our coding scheme. Although complete, there will be additions to our coding scheme and protocol in the future that could be of interest to you. Please check the Wild Minds lab GitHub page for all available versions. Definition sheets, templates, protocols, and other instruction material that we provide to code gestural communications are open-access and can be used by anyone who is interested. We would be grateful if you would cite the paper if you make use of our work. Please don't hesitate to contact Charlotte Grund (grund.charlotte(at)gmx.de) or Catherine Hobaiter (clh42(at)st-andrews.ac.uk) for questions regarding this guide.

Happy coding!

7. References

Genty, E., and Fuchs, M. (2023) GApS: A coding Scheme for Great Apes Signals in ELAN (Version v1.0.2). https//greatapesgetsures.github.io/; doi.org/10.5281/zenodo.7573664

Grund, C., Badihi, G., Graham, K., Safryghin, A., Hobaiter, C. (2023) 'GesturalOrigins: A bottom-up framework for establishing systematic ape gesture data across species'

Heesen, R., Bangerter, A., Zuberbühler, K., Rossano, F., Iglesias, K., Guéry, J. P., & Genty, E. (2020). Bonobos engage in joint commitment. *Science Advances*, *6*(51), eabd1306.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Mielke, A., Badihi, G., Graham, K. E., Grund, C., Safryghin, A., Hobaiter, C. (in prep). Many morphs: establishing great ape gestural repertoires from the bottom-up.