Lorenza.Mondada@unibas.ch

Conventions for multimodal transcription

(initial version : 2001 ; current version : 3.0.6, July 2016)
The French version is available under
http://icar.univ-lyon2.fr/projets/corinte/documents/convention_transcription_multimodale.pdf
The English version is available under
https://franz.unibas.ch/fileadmin/franz/user_upload/redaktion/Mondada_conv_multimodality.pdf

Multimodal transcript conventions

(short version)

Embodied actions are transcribed according to the following conventions
developed by Lorenza Mondada (for a full version see
 https://franz.unibas.ch/fileadmin/franz/user_upload/redaktion/Mondada_conv_multimodality.pdf):

* *	Gestures and descriptions of embodied actions are delimited between
+ +	two identical symbols (one symbol per participant)
Δ Δ	and are synchronized with corresponding stretches of talk.
*--->	The action described continues across subsequent lines
--->*	until the same symbol is reached.
>>	The action described begins before the excerpt's beginning.
--->>	The action described continues after the excerpt's end.
.....	Action's preparation.
----	Action's apex is reached and maintained.
,,,,,	Action's retraction.
ric	Participant doing the embodied action is identified when (s)he is not the speaker.
fig	The exact moment at which a screen shot has been taken
#	is indicated with a specific symbol showing its position within the turn at talk.

Detailed description of the conventions with examples

(long version)

1. Principles

These conventions are conceived to annotate all possibly relevant embodied actions, such as
gesture, gaze, body posture, movements, etc. that happen simultaneously to talk or during
moments of absence of talk.
The convention is based on two principles:
a) *Characterization of the temporal trajectory*: each embodied action is precisely temporally
located within the course of the multimodal activity and it is delimited by two symbols –
notating on the left its emergence and on the right its completion.
b) *Characterization of the embodied action*: each embodied action is shortly described.

The conventions are based on economy and consistency: by putting no more annotations than necessary, amd by choosing annotations that are interpretable univoquely.

2. Identification of the participants doing the embodied action

Every embodied movement is attributed to a participant, identified by his/her pseudonym and by a symbol consistently used for the same participant through the transcription.
(1) Example:

```
    * delimits gestures done by LAU
    ⊥ delimits gestures done by PAL
    Δ delimits gestures done by BRU
    + delimits gestures done by VIV
```

Sometimes, it might be relevant to create several lines for different embodied actions made by one participant at the same time. In this case, different symbols will be used – though it might be good to choose similar symbols.
(2) Example:

```
    * for gestures made by LAU
    • for gaze by LAU
    + for gestures made by VIV
    † for gaze by VIV
    etc.
```

If the embodied action is made by the current speaker, then its description is *not* preceded by her/his identification in the margins; if it is done by another participant, s/he will be identified in the margins. It might be useful to use capitals for the identification of the speaker and the normal style for the identification of the participant doing the embodied action.
(3) Example:

```
1  LAU      +°ah oui° je+ sais
   viv      +grasps pen+
```

3. Delimitation of the beginning and the end of an embodied action

Every embodied action has a temporal trajectory that is delimited by two identical symbols, one indicating when the action begins and the other one when it ends. These two symbols are inserted in the line of talk, in order to allow a synchronization of the verbal and the embodied conduct.
These two symbols are spatially aligned, one above the other, in order to represent their simultaneous unfolding.
The description of the action is inserted between these two symbols.
(4) Example:

```
1  BRU      la maison qui Δ se retrouve ici:,Δ
                           Δpoints to doc-----Δ
```

If an embodied action begins on a line and continues either the next line or some lines later, its description is followed by an arrow pointing to the direction of the next symbol/landmark that indicates its end. In this way, the arrow works as an instruction for the reader to search, in the following lines, for the next arrow pointing at the same symbol, closing that annotation.
(5) Example:

```
1  LAU      +°ah oui°
   viv      +points--->
2           (0.5)
3  VIV      mais alors+ i- il y va pour les dé- pour les sortir
                  --->+
```

If an embodied action begins in the middle of a pause, then the pause has to be segmented into smaller temporal fragments in order to insert the landmark.

(6) Example (line 2, the pause is 0.5 seconds long):

```
1  LAU     °ah oui°
2          (0.3) + (0.2)
   viv           +points--->
3  VIV     mais alors+ i- il y va pour les dé- pour les sortir
                   -->+
```

Note that if an embodied action is synchronized with(in) a pause, there is always an identification of the participant doing it in the margin.

In some cases, when the end of the embodied action is located several lines after its beginning, it is possible to help the reader to find when it ends by indicating the number of the line after the first arrow. Nonetheless, be careful in not over-using this type of indication: if it is used too often, it can blur the readability of the transcript by adding too many redundant indications.

(7) Example:

```
1  LAU     +°ah oui°
   viv     +points--->1.9
((transcription continues))
9  VIV     on garde une agnelle+ pour: cinq brebis euh
                          -->+
```

If an embodied action continues until the end of an except and afterwards, its description is followed by a *double arrow*. In this case, there will be no second landmark closing the temporal span of the gesture (and this is precisely indicated by the double arrow):

(8) Example:

```
1  PAL     par rapport au cas de figure de:: (.) des causses, en général,
2          euh Δcette attribution de: de terrain est bien moins claire.
   bru         Δlooks at PAL---->>
```

If an embodied action begins before the beginning of the excerpt, this is indicated by an *initial double arrow*, like in the example below:

(9) Example:

```
1  PAL     @par rapport au cas@ de figure de:: (.) des causses, en général,
           @>>points----------@
2          euh cette attribution de: de terrain est bien moins claire
```

If embodied actions of more than one participant are described, the first line is devoted to the speaker, and the following ones to the other co-participants:

(10) Example:

```
1  PAL     @par rapport +au Δcas@ de figure de:: (.) des +causses, en général,
           @>>points-----------@
   viv                  +points-------------------------+
   bru                        Δlooks at VIV------>
2          euh cette attribution de: de terrainΔ est bien moins claire
   bru                                 ---->Δ
```

The order of the annotated lines of the 2d and 3d participants can vary: however, it might be good, for the readability of the sequential relations between different embodied conducts, to insert first the ones coming early, and the ones coming late on the next line (e.g. in ex. 10, viv comes first, beginning earlier than bru).

In case of overlaps, the annotations of embodied conducts are placed *after* the overlapped talk:

(11) Example:

```
1  LAU     °ah [+oui,+   *ici  c'est vrai.°
2  VIV         [+mais+ al*ors i- il y va p*our les dé-* pour les sortir
   lau                 *points---------------------*
   viv         +nods+
```

4. Trajectories of embodied actions

Embodied actions have a temporal trajectory, which can be roughly described by distinguishing
a) a preparatory phase, b) a recognizable shape of the action, c) a retraction or withdrawal
phase. Their annotation is inspired by conventions used by Kendon for gestures and Goodwin
for gaze:

.....   small dots indicate that the embodied action is emerging,

,,,,,   commas indicate that the embodied action is withdrawing, retracting.

The embodied action itself is described when it has reached its recognizable shape, which can
also be maintained for some time.

(12) Example:

```
1  VIV     sur des terres assez *euh ass- a:*:- assez *bonnes*
   lau                          *............*points---*,,,,,,*
```

The description of action and its segmentation is an analytical decision, depending on the
precision and the granularity of the transcription.

For example the following transcripts (13a, 13b) offer two contrasted versions: the first
transcribes Jean's movements (his walk) more globally than the second (in which different
steps are distinguished, showing the coordination between the steps and the organization of the
progression of his talk):

(13a) Example (less granularity):

```
1  JEA     IL EST, *il est de bar*celo:ne, i ramène  Δle so[leil Δca]talan,
           HE IS,  he is  from Barcelona, he brings back the catalan sun,
                   *moves--------*
2  ELI                                                   [le soΔleil]
                                                         [the sun]
   eli                                                Δpivots fwd--Δ
3  YAN     exacte[ment.]
           exact[ly.]
4  JEA          [la me]:r, le *bleu:, voilà.
                [the se]e:, the blue:, that's it.
                               *begins to walk slowly-->
5  ELI     mhm.
6          (0.3)
7  JEA     voilà. (.) c'est *tout.
           that's it. (.) that's all.
                            --->*walks away-->>
```

(13b) Example (more granularity):

```
1  JEA     IL EST, *il est* de bar*celo:ne, i ramène  Δle so[leil Δca]talan,
           HE IS,  he is  from Barcelona, he brings back the catalan sun,
                   *Lfoot-*Rfoot--*
2  ELI                                                   [le soΔleil]
                                                         [the sun]
   eli                                                Δpivots fwd-Δ
3  YAN     exacte[ment.]
           exact[ly.]
4  JEA          [la me]:r, le *bleu:, voi*là.
                [the se]e:, the blue:, that's it.
                               *1 step fwd*1 more step-->
5  ELI     mhm.
6          (0.3)*
   jea       -->*1 step fwd----->
7  JEA     voilà. (.) c'est *tout.
           that's it. (.) that's all.
                            --->*walks away-->>
```

If a participant makes different relevant embodied movements at the same time, they will be
described in different lines, delimited with different symbols:

```
 (14) Example:
1  PAL  .hh ∠ c'est des ∠ c'est le le∇ schéma euh: ∇ordinaire, .hh
   lau       ∠looks at PAL∠
   lau                                    ∇............∇points----->>
```

5. Description of embodied actions

The description of embodied actions is not straightforward.
It is important to avoid physicalist descriptions (e.g. 'arm moves 45° down to the floor') as well as intentionalistic descriptions (e.g. 'wants to grasp the bottle')!
Some categories tend to standardize und homogeneize actions that might be quite different (e.g. 'points', 'gazes'), some tend to be very general and vague (e.g. 'gesticulates'), while some can be very specific (e.g. 'points with the top of the pen'). This depends on the granularity of the transcript/analysis, as well as on the local relevancies governing the action.

Note that a CA transcription is very different than a standardized coding (the former favors 'emic' descriptions, the latter 'etic' categories).
(15) Example:
```
1  LAU   °ah [oui, *ici c'est +vrai.°*
2  VIV       [mais *alors i-il+ y va p*our+ les dé- pour les sortir
   lau             *points with her pen*
   viv                   +two big nods+
```
Although sometimes 'points' might be a good enough annotation, further expanded in the analytical text, in other cases 'points with his pen', 'points with open horizontal palm' etc. might be more accurate – as well as the description of what is pointed at ('points at the statue').

Another constraint for these annotations is their lenght: it is better to choose short descriptions, fitting within the transcript's spatialization, in a readable and intelligible way.
Sometimes annotations can be abbreviated. Different solutions are possible, as here below:
(16) Example:
```
1  LAU   °ah [oui, *ici c'est vrai.°*
2  VIV       [mais *alors i- il   y* va pour les dé- pour les sortir
   lau             *LH points w pen*
L = left
H = hand
```
(17) Example:
```
1  LAU   °ah [oui, *ici c'est vrai.°*
2  VIV       [mais *alors i- il   y* va pour les dé- pour les sortir
   lau             *------1--------*
1: Left hand points with the upper part of the pen
```

6. Screen shots

Multimodal transcripts make use of text descriptions as well as images, which are integrated in the transcription.
It is very important to always specify the exact moment in the video to which the image refers to and to synchronize it with respect to the lines of talk and of embodied conduct. This is done by inserting a symbol (#) both on the line of the talk and on the line dedicated to the image (*fig* in the margins).

(18) Example:
```
1  GEO   alors, y a quand même u:ne autre chose à vérifier, c'est
2        que: ce: ces chambres-là n'ont que une issue de secours
3        (0.7)#
   fig       #fig.1
```

```
4  LUC    *non:, on# a on a*# touj- ah oui, (.) parce qu- (.) parce
          *...............*points--->>
   fig            #fig.2    #fig.3
```


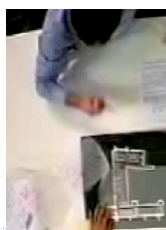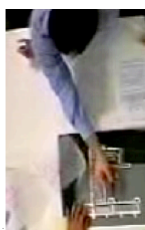figure 1   figure 2   figure 3

```
5         que là i bloquerait cette issue
6  GEO    ben oui.
```
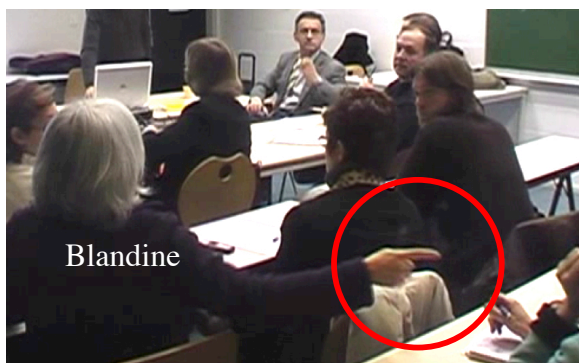
Alternatively, each image can be followed by its number on the same line (see below ex. 19, fig. 1)

Often, it is useful to insert several images one beside the other, one consecutive to the other, contrasted enough to show the trajectory of a movement, a gesture, or an embodied action (see figures 1-2-3 in ex. 18 above).

In order to enhance the readibility of the images, it might be useful to add circles and arrows highlighting a relevant detail.

(19) Example:
```
1  BLA                                     [yeah %i mean,
                                            %turns back->
2         (0.3) % (1.9)
   bla       -->%extends her arm backwards-->
3  BLA    anne ram%beau# is [ehm the person who:
             -->%points back-->>
   fig             #fig.1
```


7. Line numbering

The lines referring to talk and silences are numbered. They constitute the relative temporal metrics to which gesture and other embodied conducts are synchronized.
The "appended" lines relative to embodied conducts are *not* numbered.
Note that this numbering is manually done – automatic numbering does not work.

(20) Example:
```
1  PAL    devient une unité annexe, (.) qui: sert au renouvellement.
2         *(0.5)
   la     *....-->
3  PAL    elles y restent jusqu'à toussaint au moment de la: lutte.
```

```
4  LAU    et *ç- ça, là. (.) y a une différence entre ça et ça?*
              ->*points---------------------------------------*
5  PAL    non.
```

In order to enhance the readability of the transcripts various types of font can be used. In the example below, bold is used for talk, italics for translation and grey for embodied conducts.

(21) Example:
```
1  PAL    ben suivant le cas euh: ben on tra- on est là
          well depending on the case ehm: well one wo- one uses
2         que pour le champ, ou bien on va sur le pâturage, .h
          only the field, or one goes on the pasture, .h
3         sur l'assembla:‡ge +sans parcours. .h +†je pen‡se que+†
          on assembled parcels without roads. .h I think that
           >>gazes at lau‡gazes at viv------------------‡looks down->
   viv                    +..................+moves sheet+
                                               †leans forward†
4         +dans le cas du gaec du pr+‡adou, .h c'est: ‡ tout l'un,
          in the case of the ((name of the farm)), .h it's either one,
                            -->‡gazes at lau----‡gazes at viv->
   viv    +moves RH forwards-------+
5         tout l'autre.
          or the other.
6  VIV    +.hh oui. par‡ce que: i'm'sem+ble: eh i- ici c'était
          .hh yes. because: it seems to me: eh he- here it was ((cont.))
          +..........................+points-->>
   pal              -->‡gazes at the pointed at map-->
```

8. Note about the difference between multimodal transcript and description

Descriptions are often put between two double parentheses:
(22) Example:
```
1  PAL    une unité annexe, ((coughs)) qui: sert (.) uniquement,
2         et en continu, (0.3) à la génération de renouvellement.
```
This convention might be useful for some comments related to vocal conducts that are not transcribed – although its limitations have been pointed at (see Jefferson on transcribing laughter).

This convention is not useful for *transcribing* embodied conducts, because it is not temporally precise enough and it does not describe neither the boundaries nor the lenght of the embodied movement (see the principles above; see the contrast between the following two examples, ex. 23a being an insufficient description of the gesture, vs. 23b being a precisely temporalized description of the pointing gesture).

(23a) Example:
```
1  BRU    la maison qui se retrouve ici:,((points))
          the house that is located here:
```
(23b) Example:
```
1  BRU    la maison qui ∆ se retrouve∆ ici:,∆
          the house that is located here:
                        ∆points to doc∆,,,,,∆
```

Bibliography

Some articles in which these notations have been used/explained/discussed.

Mondada, L. (2003). Working with video: how surgeons produce video records of their actions, *Visual Studies*, 18/1, 58-72.

Mondada, L. (2004), Temporalité, séquentialité et multimodalité au fondement de l'organisation de l'interaction: Le pointage comme pratique de prise du tour. *Cahiers de Linguistique Française*, 26, 169-192.

Mondada, L. (2005). La constitution de l'origo déictique comme travail interactionnel des participants : une approche praxéologique de la spatialité. *Intellectica*. 2/3, 41-42, 75-100.

Mondada, L. (2006) 'La pertinenza del dettaglio: registrazione e trascrizione di dati video per la linguistica interazionale', in Y. Bürki and E. de Stefani (eds). *Trascrivere la lingua. Dalla filologia all'analisi conversazionale*, pp. 313–44. Bern: Lang.

Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies, 9 :2,* 195-226.

Mondada, L. (2007). Transcript variations and the indexicality of transcribing practices. *Discourse Studies*, 9/6, 809-821.

Mondada, L. (2008). Documenter l'articulation des ressources multimodales dans le temps: la transcription d'enregistrements vidéos d'interactions. In: M. Bilger (éd). *Donnees orales, les enjeux de la transcription*. Perpignan: PEP, 127-155.

Mondada, L. (2009a). The Embodied and Negotiated Production of Assessments in Instructed Actions, *Research on Language and Social Interaction*, 42 : 4, 329–361.

Mondada, L. (2009b). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics. 41,* 1977-1997.

Mondada, L. (2011). Understanding as a embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*. 43, 542-552.

Mondada, L. (2012). Video analysis and the temporality of inscriptions within social interaction: the case of architects at work. *Qualitative Research*, 12, 3, 304-333.

Mondada, L. (2012). The conversation analytic approach to data collection. J. Sidnell, T. Stivers (Eds.). *The Handbook of Conversation Analysis*, Blackwell-Wiley, 32-56.

Mondada, L. (éd). (2014). *Corps en Interaction. Participation, spatialité, mobilité*, Lyon : Editions ENS.

Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics, 65,* 137-156.

Mondada, L. (2014). Bodies in action: multimodal analysis of walking and talking, *Language and Dialogue* 4:3, 357–403.

Mondada, L. (2013). Embodied and spatial resources for turn-taking in institutional multi-party interactions: the example of participatory democracy debates. *Journal of Pragmatics, 46,* 39-68.

Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics, 65,* 137-156.

Mondada, L. (2014). Bodies in action: multimodal analysis of walking and talking, *Language and Dialogue* 4:3, 357–403.

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*. 20:2, 2-32.

Mondada, L. (2016). Going to write: Embodied trajectories of writing of collective proposals in grassroots democracy meetings. *Language and Dialogue, 6:1,* 140-178.

Mondada, L., Svinhufvud, K. (2016). Writing-in-interaction: Studying writing as a multimodal phenomenon in social interaction. *Language and Dialogue, 6:1,* 1-53.