

Odense Summer School on
Corpora and Multimodality

Alignment tools: using ELAN for CA

Lorenza.Mondada@univ-lyon2.fr
(ICAR research lab, CNRS & University of Lyon)

Alignment tools: general features 1

- *Timed association* between the transcript (text) and the recording (audio or video file)
 - Distinction between primary data and secondary data is relativized: They are not autonomous, not separated anymore.
 - Although it has always recognized that THE data were the recordings, people often worked on transcripts only, because of the difficulty to retrieve exact positions on tapes.
 - For multimodality, immediate and precise access to the video is indispensable

Alignment tools: general features 2

- Generally, alignment tools have two components and functions:
 - They assist and facilitate transcription
 - They allow for coding, searches and statistics
- Transcription and representation of the flow of talk and multimodal conducts:
 - Transposition from time to space
 - Representation of time is crucial
 - Two formats exist :
 - The list format (ex. CLAN, Transana,...)
 - The partition format (ex. Praat, ELAN, ANVIL,...) --> based on an infinite timeline
 - For a CA perspective on talk, the list format is more adequate for the representation of sequentiality; however, for a multimodal analysis of various simultaneous lines of action, the partition format is very useful
 - These formats have analytical implications

Alignment tools: general features 3

- In the process of analysis, we deal with different kinds of transcripts and different versions of a transcript:
 - (usually raw) transcripts of the whole of the recording
 - (usually detailed) transcripts of a particular segment
 - This is namely the case of multimodal transcripts in CA,
 - since we don't code systematically a whole recording
 - Multimodal transcription does not annotate a fixed, pre-defined set of features but the features made relevant by the participants (and focus of the analysis)
 - Edited (.doc or .txt) working transcripts
 - Edited (.doc or .txt) transcripts for publication
- Thus, transcripts are plural, hybrid, and dynamic object...

ELAN

- <http://www.lat-mpi.eu/tools/elan>
- Developed at the MPI Nijmegen
- Works on PC, MAC, Linux
- Formats: wav, mpeg

ELAN: some useful features

- The perspective adopted here
 - User's perspective
 - CA's perspective
 - As many *tiers* as we need
 - Various segmented (you can have different hierarchical and dependent relationships between tiers - but also keep them autonomous)
 - Which can be variously displayed (flexible order)
 - Easy to change the scope, length, etc. of the segments
 - A very fine *player*
 - Allows to transcribe both on the basis of the visualization of audio and the video
 - frame per frame, speed control
 - Synchronization of multiple sources
- very interesting for complex co-occurrent phenomena (ex. dinner conversation)
- Very interesting for the timed representation of 'non-verbal' actions coordinated to talk (ex. surgery)

ELAN: some problems for the (CA) user

- Readability: In CA, talk is segmented in turns or TCUs --> difficult to see where the simultaneous segments do overlap talk exactly (only possible to hear)
- Export to .txt and to excel - but not to an edited transcript respecting/reproducing the spatialization of the timed coordination between participants
- Possibility to have various audio and video sources, but once you have chosen your source(s), you cannot add others

- For an example of aligned data in different softwares - and especially in CLAN and ELAN, see:
- http://icar.univ-lyon2.fr/projets/corinte/documents/Projets/incorporation_exemplaire_analyses.htm
- This page contains :
 - Video sources (multi-scope video: 3 views edited in split screen)
 - Corresponding aligned transcripts + their dynamic capture (.mov) + their word version
 - Analysis of this fragment (in French)

Odense Summer School on
Corpora and Multimodality

Corpus annotations

Transcribing, describing, coding

Lorenza.Mondada@univ-lyon2.fr
(ICAR research lab, CNRS & University of Lyon)

aims

- Explicitate some of the specific demands CA has for transcribing multimodality
- Give an example : a specific phenomenon
- The way in which it is transcribed -->
conventions for transcribing multimodality
- The way in which it can be analyzed in a systematic way (**collections** / generalizations)
- Difference between *coding*, *describing*, and *transcribing*

CA demands for multimodal transcripts

- *Relevance* of multimodal phenomena: given by participants' orientations
 - VS. transcribed a priori, uniformly (coding)
 - Virtually everything can be made relevant and turned into a situated resource within/for interaction
 - Broad definition of multimodality: gesture, gaze, facial expressions, body postures, mouvements, manipulations of objects, ...
- *Temporal* coordination of these phenomena with other concurrent resources (talk but not only)
 - Importance of timed precision (beginning/end)
 - Why that now?
- *Sequential* importance of these phenomena in the forming of turns and series of actions
 - Sensitivity to the detailed organization of the ongoing action
 - Observable in the organization of the next action (i.e. sequentially consequential)

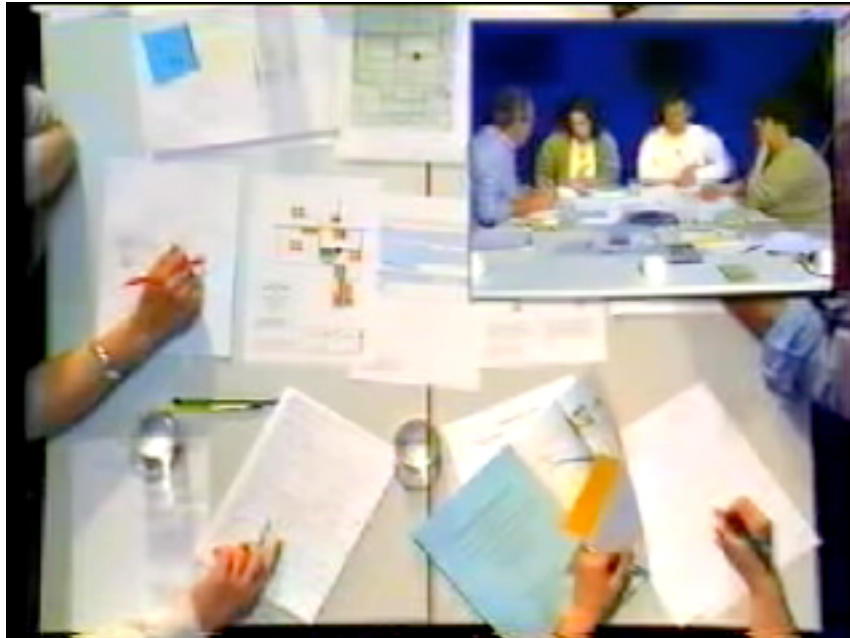
An example

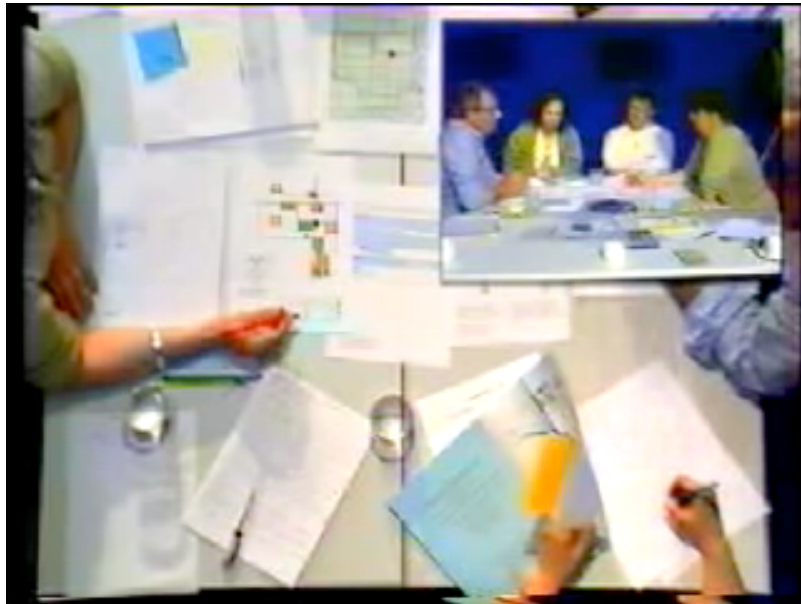
- Pointing in pre-beginnings as a resource projecting self-selection by the pointer
- Cf. :
- Mondada, L, 2007, Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, 9 :2, 195-226.
[reprinted in Dijk, T. van (ed.). *Discourse Studies*, London : Sage, vol. IV, 126-157].

Ex1: pointing at turn's beginning

(1) (e9/ag1-47.00)

- 1 PAL ben suivant le cas euh: ben on tra- (.) on est là que pour le
well in certain cases ehm well one wor- (.) one is there only for the
2 champ/ et puis à d'autres moments:/ ben on va échouer/ (0.3)
field/ and at other moments:/ well one finish/ (0.3)
3 en pâturage\ .h sur l'assemblage sans parcours\ .h je pense
in pasture\ .h on an assembling without any path\ .h i think
4 que dans le cas du gaec du pradou/ .h c'est tout l'un/
that in the case of the ((Region's name))/ .h it's either one/
5 tout l'autre\
or the other case\
6 VIV +.hh oui\ parce que:# i'm'sem+ble: eh i- ici c'était s::- ce
.hh yes\ because: it seems to me eh he- here it was w::- what
+.....+ppp with pen-->>
7 qui: ce que ça voulait représenter/ [c'était
this wanted to represent/ [was
8 LAU [*c'est les am*andes ça\
[are these almonds\ ((=kind of field))
*.....*ppp with finger-->>





Exc. 1

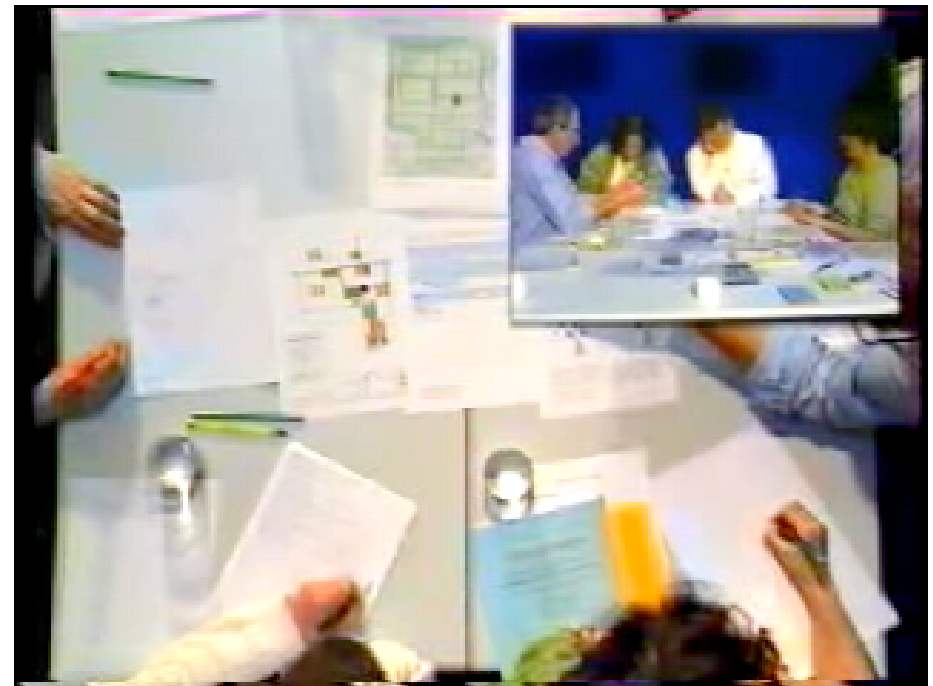
6 VIV +.hh oui\ parce que: i'm'sem+ble: eh i- ici c'était s::- ce
 .hh yes\ because: it seems to me eh he- here it was w::- what
 +.....+ppp with pen-->>

- pointing's maximal extension is reached when a deictic (*ici* = here) is uttered
- BUT also:
- pointing begins before, at the beginning of the turn
- (note on the importance of transcribing the finely tuned temporality of these gestures)
- --> looking after other occurrences

Exc. 2 — pointing anticipates the end of previous turn

(2) (e3/ag1-16.48)

- 1 PAL et donc on voit la logique/ avec cet cet aménagement du: de
and so one sees the logic/ with this settlement of: of
2 l'espace/ .hh qui revient/ à obtenir/ (.) euh: des des de- des
space/ .hh which allows to obtain (.) ehm
3 pâturages/ (.) RElativement plus productives/ qu'elles ne
pasturelands/ (.) RElatively little productive/ than what they
4 l'étaient avant/ grâce notamment à la: (.) <la
were before/ thanks among other things to the: (.) <the
5 re*distribution des biens communs\ ((decrecendo))> *
redistribution of common goods\ ((decrecendo))>
la *.....*
6 LAU *et donc le: ce qui est en (.) orange ici là/ terre assolée
and so the: what is coloured in orange there/ farmland
*ppp->>
7 PAL hum
8 LAU c'est/ (.) des prairies/
it's/ (.) grassland/

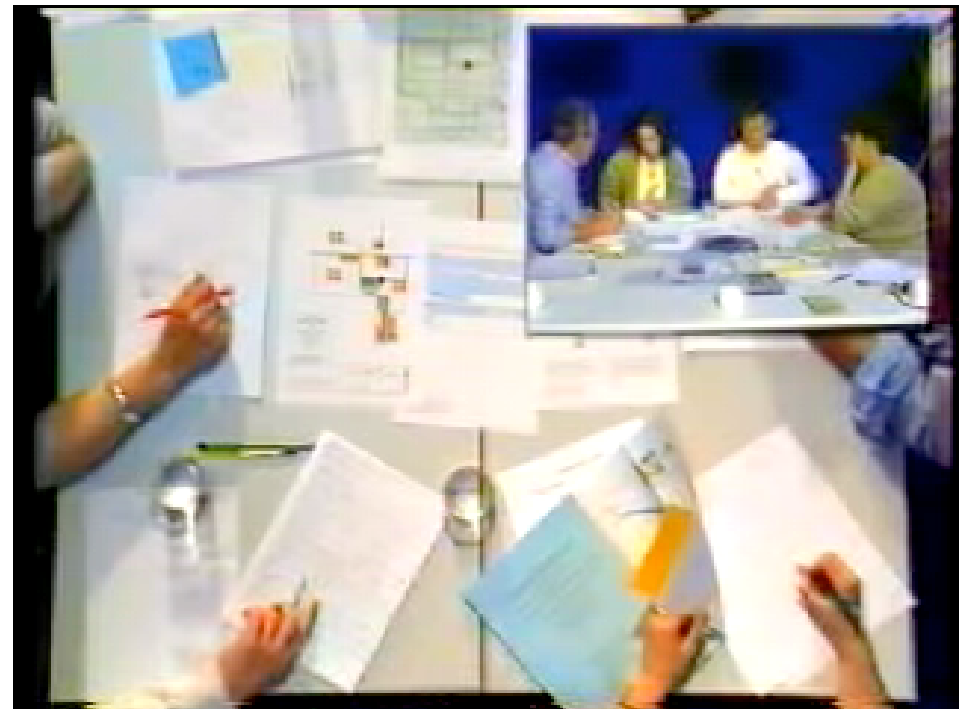


Exc1bis — other anticipating movements

(1bis)

1 PAL ben suivant le cas euh: ben on tra- on est là que pour
2 le champ/ et puis à d'autres moments:/
3 ben on va échouer/ (.) en pâturage/ .h
4 sur l'assemblage +sans parcours/ .h +je pense que+
viv +.....+moves paper+
5 +dans le cas du gaec du pr+adou/ .h c'est tout l'un/
viv +turns wrist, comes in----+
6 tout l'autre\
7 VIV +.hh oui\ parce que: i'm'sem+*ble: eh i- *ici c'était
+.....+ppp with pen-->>
lau *opens folder*

- VIV anticipates the end of the ongoing TCU (« parcours » 4) but
- PAL expands his turn with a new TCU (« .h je pense... » 5)



projections

- Summary of the preceding excerpts
In exc. 2 pointing occurred in a *pre-terminal* position
In exc. 1 pointing is prepared by other gestures
and occurs well before, in a *pre-terminal position*
transformed by current speaker into a middle position
- A general feature of gestures: anticipations
Pointing gestures + other gestures or movements (preparatory phase) exhibit the online member's interpretation of the ongoing TCUs and TRPs and projects their completion
The span of this anticipation can vary
- General features of TCUs: their completion is locally defined by the participants, in a flexible, dynamic way - with the possibility of redefine it retrospectively
(cf. work by Auer, 1996, 2005; Selting, 2000; Ford, Fox, Thompson, 1996, Schegloff, 1996, Ford, 2004...)

Exc. 3 — revisions of the ongoing interpretation

(5) (agro 1 / 19.26-19.56 = 20.15 persp)

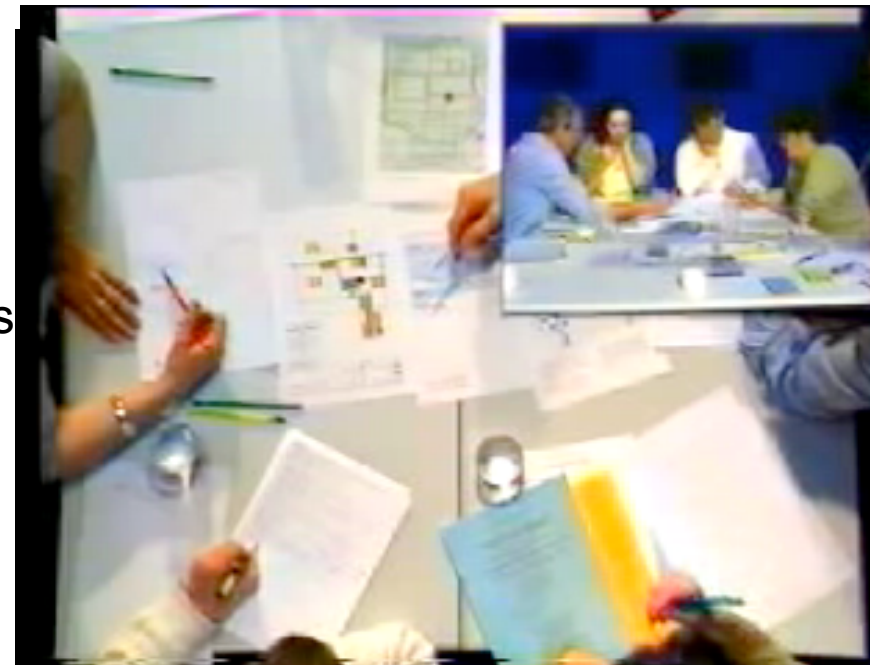
1 P&L .h et donc elles rentrent en production à::m: c'est-à-dire elles
 .h and so they go into production at m: that is to say they
 2 son- elles mettent bas, .h (.) à quatorze mois, c'est-à-
 are- they give birth, .h (.) at fourteen months, that is to
 3 *dire ou trei- oui qua*torze mois, c'est-à-dire deux mois après
 say or thir- yes fourteen months, that is to say two months after
 lau *.....*ppp-->
 4 les brebis, .hl et donc, (.) pour arriver à cette performance
 the ewes, .h and so, (.) in order to secure this performance
 lppp--->
 5 c'est-à-dire être aptes à mi- *à être mises à la reproduction
 that is to say be able to re- to be put to reproduction
 lau ----->*suspends pointing--->
 6 .h à leur premier automne, .hh euh elles sont à un régime assez soigné,*
 .h for their first autumn, .hh ehm they are given a careful diet
 lau ----->*
 7 *et elles sont notamment *hébergées là.
 and they are specifically housed there.
 lau *.....*ppp----->>
 8 LAU c'est ça qu'il est r'présenté ici/
 pal ---->l,,, ,
 9 P&L voilà\

LAU points projecting the end of the TCU

The end is delayed by a series of reformulations
 (« c'est-à-dire »)

Pointing is suspended after the initiation of a
 subordinate clause (« pour arriver... »)
 and a reformulation

Suspension ends at the end of the main clause



Transcript conventions

- Back to demands addressed to CA conventions
- Describing vs transcribing
(cf. Jefferson, 1985 on laughter)

extrait 1 (e9/agro1-47.00) transcr. simplifiée

1 PAL ben suivant le cas euh: ben on tra- on est là que pour le
2 champ/ et puis à d'autres moments:/ mais on va échouer/ .
3 comme pâturage\ .h sur l'assemblage sans parcours/ .h je pense
4 que dans le cas du gaec du pradou/ .h c'est tout l'un/
5 tout l'autre\
6 VIV ((pointe)) .hh oui\ parce que: i'm'semble: eh i- ici c'était s::- ce
7 qui: ce que ça voulait représenter/ [c'était
8 LAU [((pointe)) c'est les amandes ça\

extrait 1 (e9/agro1-47.00) transcription revue

1 PAL ben suivant le cas euh: ben on tra- on est là que pour le
2 champ/ et puis à d'autres moments:/ mais on va échouer/ .
3 comme pâturage\ .h sur l'assemblage sans parcours/ .h je pense
4 que dans le cas du gaec du pradou/ .h c'est tout l'un/
5 tout l'autre\
6 VIV +.hh oui\ parce que: i'm'sem+ble: eh i- ici c'était s:- ce
+.....+point-->>
7 qui: ce que ça voulait représenter/ [c'était
8 LAU [c'est les am*andes ça\
*.....*pointe-->>

Transcript conventions: some principales

- Every gesture or action is timely synchronized with talk
- Every phenomenon is *delimited* (left and right)
 - Use of one symbol repeated on both lines of talk and action
- The *trajectory* is described (cf. Kendon, Schegloff, Goodwin...)
 - Preparation:
 - Hold: ----
 - Retraction: ,,,,,,
- *Description* of the action is not standardized (no constraints on the descriptors to be used, in this sense, *no coding*) - it intends to capture the relevances for participants (emic description)
 - Ex. « points », « points with the pen », « comes in with the hand »...
 - In this sense, it is a '*praxeological description*' (vs a functional description, vs a physiological description).
- Possibility to think about an (extra) annotation of the phenomenon (such as « early pre-beginning pointing »)
(nota: phenomena are defined in CA by a) sets of resources, b) in sequential positions, c) doing a certain action)

Transcript conventions



- Exc 1 again

1 PAL ben suivant le cas euh: ben on tra- on est là que pour
2 le champ/ et puis à un autre moment:/
3 ben on va échouer/ . en pâturage/ .h
4 sur l'assembla:†ge †sans parcours\ .h †je pen†se que+
>>>reg. lau&jm†reg. viv -----†reg.bas
viv +.....+dépl.feuille+
+buste en av+
5 +dans le cas du gaec du pr†adou/ .h c'est: † tout l'un/
†reg. lau-----†reg. viv->
viv †tourne le poignet, avance†
6 tout l'autre\
7 VIV †.hh oui\ par†ce que: i'm'sem†ble: eh i- *ici c'était
†.....†*ppp av >lylo-->>
lau †laisse retomber page*
pal †reg. feuille pointée-->

TRANSCRIPT CONVENTIONS

Multimodal details have been transcribed according to the following conventions:

* *	delimitate Laura's gestures and actions descriptions.
+ +	delimitate Viviane's gestures and actions descriptions.
⊥ ⊥	delimitate Pierre-Alain's gestures and actions descriptions.
Glances are described by the same conventions, but in a light-coloured character.	
*--->	gesture or action described continue across subsequent lines.
*--->>	gesture or action described continue until and after excerpt's end.
--->*	gesture or action described continue until the same symbol is reached.
>>--	gesture or action described begins before the excerpt's beginning.
....	gesture's preparation.
-----	gesture's apex is reached and maintained.
.....	gesture's retraction.
lau	participant doing gesture is identified when (s)he is not the speaker.
fig	the exact point where screen shot (figures) has been taken is indicated,
#	with a specific sign showing its position within turn at talk.
com	commentary
ppp	pointing gesture
w	with (pen/finger)
H	hand

How to generalize these observations?

In CA: How to built *collections* ?

- This anticipation of turn beginnings and of self-selection of the next possible speaker
 - Is a recurrent phenomenon (*systematicity*)
 - Can be achieved by different gestures
 - Depending on the ecology of action (*indexicality*)
- => Towards an analysis of *collections* (= a possibly large number of occurrences documenting the same *practice* for achieving an action, characterized by a set of *resources* arranged in a precise *sequential environment*)

Ex: pointing: Examples taken from other corpora



- Different ways of pointing (with pen / with index) but within the same sequential positions : anticipation of turn-beginnings
- Other gestures can be made in the same positions, for achieving the same action, but with different resources (the type and shape of these resources is clearly related to the ecology of the activity)

Multimodal collections

- A complex *cluster* of different, multimodal, concurrent phenomena in the same sequential environment for achieving the same action
- In this case, what is important is less the detailed description of the gesture than
 - the detailed position within turns at talk
 - And the trajectory of gesture (its preparation, then its retraction), related to talk
- What kind of automatic searches are thinkable on the basis of this analysis?
 - possible automatic searches on *preparations*
 - Searches associating preparation + sequential positioning within the last / pre-last TCU

summary

- *Transcribing vs describing* (Jefferson 1985): « .hh ah hh Hh » vs ((laughter))
- For multimodality:
 - Praxeological description (relevance < *emic*)
 - Integration of transcription requirements: temporality (position + trajectory)
- *Description* is not *coding* (set of predefined *etic* searchable categories)
- Extra *annotation* level (analytical categories such as third turn other-initiated repair)
- A more general issue for corpus data bases:
- What are the objects on which to run searches?
 - Surface features of talk VS coded categories
 - surface features + annotations associated to them (scope)