



DOCUMENTARY LINGUISTICS I

prof. Nicole Nau, UAM winter 2017/2018

Nineth lecture
05 December 2017

TOPICS OF THE DAY: TRANSCRIPTION AND ANNOTATION

- ❖ Recall: what does a speech event contain, what do we (can we) document and describe
- ❖ Recall: what is annotation
- ❖ First steps after making a recording
- ❖ Types of transcription
- ❖ The problem of segmentation
- ❖ Using ELAN for segmenting and transcribing an audio recording

FROM LAST WEEK'S LECTURE (SLIDE BY MARTA OSTAJEWSKA)

In video documentation of the language it is important that the gestures fit in the frame, do not go beyond the frame, are not cut



Why?

EXAMPLES: GESTURES IN AND OUT OF THE FRAME

1. Recording of a life story in the language **Bakola**, Cameroon (DOBES)
2. Interview with a speaker of **Latgalian** in eastern Latvia

REMEMBER

All **recordings are reductions** – they can never render the total of the speech event, all aspects that may be important for the recorded language use.

All **transcriptions** of a recording **are further reductions**.

You are forced to make a **selection** – make it consciously!

WHAT ARE **THE FIRST THINGS TO DO** WITH A RECORDING?

1. **Back up:** store original at a save place. Use an appropriate, unique file name.
2. **Metadata:** describe the recording and keep the description with the original (and remember where you have the **consent form**)

Only then you should go on and work with a copy of the recording. This may involve:

3. **Cut** the recording into manageable and coherent pieces.

How to cut? Recommended program for audio: «Audacity»

4) ANNOTATE CHOSEN (PIECES OF) RECORDINGS. ANNOTATION MAY INCLUDE:

- ❖ transcription(s)
- ❖ translation(s)
- ❖ grammatical analyses
- ❖ comments on the language/speech/code (use of slang, whispers, code-switching...)
- ❖ comments on the content
- ❖ notes on situational factors (a telephone rings, the speaker pauses to drink coffee, talks to their dog...)
- ❖ video: description of gestures, facial expression etc.
- ❖ ...

TRANSCRIPTION: DIFFERENT SYSTEMS FOR DIFFERENT PURPOSES

- ❖ orthographic transcription(s)
 - using an established writing system
 - using a newly developed / developed writing system
 - using an adhoc-writing «system»
- ❖ phonetic transcription(s)
 - broad / narrow phonetic transcription
 - IPA or a regional/local traditional phonetic script (for example, slavists's system for transcribing Slavic languages)
 - SAMPA
- ❖ phonemic transcription
- ❖ conversation-analytic transcription

WHY TRANSCRIBE? DIFFERENT PURPOSES

- ❖ phonetic and phonemic analysis, (remember: transcription **is** analysis)
- ❖ conversation analysis,
- ❖ other linguistic purposes,
- ❖ making content easily accessible (faster reading than listening),
- ❖ finding words and constructions,
- ❖ other form of “storage”, of having a “record” (but remember a transcript is not primary data),
- ❖ preparing further steps in annotation, e.g. translation

TRANSCRIPTION AS TRANSLATION FROM ORAL TO LITERATE

Sometimes, you see his hand at close up, as he is snatching the pears from the trees, and you hear the sound really strongly.

and... you see his hand sometimes at close up – snatching the pears from the tree. And you hear the sound really strongly.

and (0.4) you see his hand sometimes at close up ;
(1.1) uh snatching the pears from the tree /
(0.8) and you hear the sound really: strongly \



UNITIS IN WRITTEN AND SPOKEN LANGUAGE

Units of written language	Units of spoken language
words	words
	intonation units
sentences	
paragraphs	paragraphs, episodes

NIKOLAUS HIMMELMANN (2006) THE CHALLENGES OF SEGMENTING SPOKEN LANGUAGE: HOW TO DETERMINE INTONATION UNITS?

«In most languages, evidence for intonation unit boundaries is provided by changes in pitch and rhythm.»

1. Pitch

1.a) **boundary tone** at the end of an intonation unit, change in pitch on the last syllable

1.b) new onset at the beginning of the next unit, **jump in pitch**

1.c) reset of the baseline: early pitch peaks in the new unit are higher than the final pitch peaks in the preceding one.



2. Rhythm

2.1 pause

2.2 lengthening of the final segment

2.3 anacrusis

Obligatory in many languages: 1.a) and 1.b)

Most practical diagnostics: 1.a) and 2.a)

TRANSCRIPTION CONVENTIONS USED BY HIMMELMANN

/ rising boundary tones

\ falling boundary tones

; unit ends somewhere in the middle, neither rising nor falling

(.) short pause

(..) longer pause

(0.5) measured pause of 0.5 seconds

: lengthening

= latching

EXAMPLES FROM HIMMELMANN

(1) PEAR-L-36FF

36. and (0.4) you see his hand sometimes at close up ; (1.1)

37. uh snatching the pears from the tree / (0.8)

38. and you hear the sound really: strongly \ (0.8)



(2) PEAR-L-48FF

48. he climbs down the ladder / (0.5)

49. and he puts a couple of the pears— (=)

50. well: (=)

51. as he's standing there ; (=)

52. couple of the pears fall \ (0.4)



ANOTHER EXAMPLE — CAN YOU HEAR INTONATION UNITS IN AN UNKNOWN LANGUAGE?

Breu, Walter. *Le voyage à Palata*, Na-našu corpus. Pangloss Collection, LACITO-CNRS. Programme franco-allemand Euroslav 2010 : ANR(09-FASHS-025) - DFG(BR 1228/4-1), http://lacito.vjf.cnrs.fr/pangloss/corpus/show_text_en.php?id=crdo-SVM_PALATA_SOUND&idref=crdo-SVM_PALATA



LET'S PRACTICE! => HANDOUT

You need ELAN installed and the following files:

- ❖ chapter10_figure1_pear_I_36ff.wav,
- ❖ chapter10_figure2_pear_I_48ff.wav and/or chapter10_figure3_pear_I_88ff.wav,
- ❖ PearStoryExmaralda.wav