

# Sign language transcription conventions for the ECHO Project

Annika Nonhebel, Onno Crasborn & Els van der Kooij  
*University of Nijmegen*

*Version 9, 20 Jan. 2004*

URL: [http://www.let.kun.nl/sign-lang/echo/docs/transcr\\_conv.pdf](http://www.let.kun.nl/sign-lang/echo/docs/transcr_conv.pdf)

## 1. Introduction

The main goal of the case study ‘a comparative study of European sign languages’ is to create a database with comparable multi-media recordings documenting sign languages as used in three European countries (Sweden, UK, and the Netherlands). This study should be regarded as a demonstration of how multi-media data can be documented and used on the internet to make them accessible for the whole scientific community, for deaf people themselves or for anyone interested in sign languages. These publicly available data collections are expected to contribute significantly to the systematic study of individual languages, but also the study of the similarities and differences between languages.

This new technology for presenting sign language data and transcriptions poses the following question: to what extent should we use standard transcription conventions? If all the raw material (the video sources) is available, do we need full transcriptions? In principle, one can look at the video source for all kinds of information that are traditionally included in various transcription system, such as eye gaze, head nods, etc. On the other hand, the great strength of computer tools such as ELAN is that it allows for complex searches in large data domains and for the immediate inspection of the video fragments relating to the search results; this is typically very time-consuming when using paper transcription forms or even computerized transcription forms that are not directly linked to the original data. We wanted to establish an annotation system that could be useful for any researcher, with a focus on the syntactic and discourse domains. We tried to be careful not to impose too much analysis on any tier by saying that a specific phonetic form is an instance of ‘person agreement’, for example. The relatively small set of transcription tiers allows for the coding of a large data set, which can be further expanded by researchers according to their specific needs.

ELAN will see many updates in the coming year (2004); one of the future functions will be the possibility to expand a publicly available transcription file with ones own additions, including extra tiers, and storing these additions in a local file while maintaining the link to the original transcription that will be stored on a remote server.

## 2. Annotation tiers

Below we describe the different tiers used in this case study. A tier is a set of annotations that share the same characteristics, e.g. one tier containing the Gloss RH or another tier containing the Dutch Translations. For ELAN two types of tiers exist; “parent tiers” and “child tiers”. Parent tiers are independent tiers, which contain annotations that are linked directly to a time interval of the media frame. Child tiers or referring tiers contain annotations that are linked to annotations on another tier (the parent tier). (See also ELAN manual, available at <http://www.mpi.nl/tools/elan.html>.)

ELAN gives the opportunity to select one or more video frames and assign a specific value to this selected time span. For example, when the eye brows are first up and then down (neutral) in the same sign, one would only select that part of the video in which the eyebrows are up for the *brows* tier, and mark that time-domain with a specific code (for instance ‘up’). This is possible for all tiers that one creates.

### 2.1 Gloss tiers

The information in the gloss tier is bound to information about the hands. In contrast to conventional glossing, the gloss line is divided in *Gloss right hand tier* and *Gloss left hand tier*. The gloss of the sign will be notated in the tier line corresponding with the visual information. So, a gloss of a one-handed sign produced with the right hand is added to the right hand tier and a gloss of a one-handed sign produced with the left hand is added to the left hand tier. When a sign is two-handed the gloss will be added to both tiers. *Deviations* of handedness are also included. When a sign that is normally made with one hand is made with two hands, then the code ‘2h’ is included before the gloss. A two-handed sign made with one hand receives the code ‘1h’ before the gloss. When there is a clear hold at the end of the sign, this is notated as ‘-h’ on the gloss tier.

The Gloss tiers are given in the first language and form the parent tier for repetition and direction and location. Next to that an English gloss line is given. This gloss line should be time aligned but has no child tiers.

A classifier is a handshape that refers to a few objects with the similar properties, such as size and shape. Objects can be placed in the syntactic signing space by using a classifier. A classifier shows a fixed relation between the handshape and the object you refer to. It is marked as ‘(p-)’ (p followed by a dash). The code ‘p’ stands for ‘poly’, meaning many meaning components.

*Gloss Right Hand values (same values as Gloss Left Hand):*

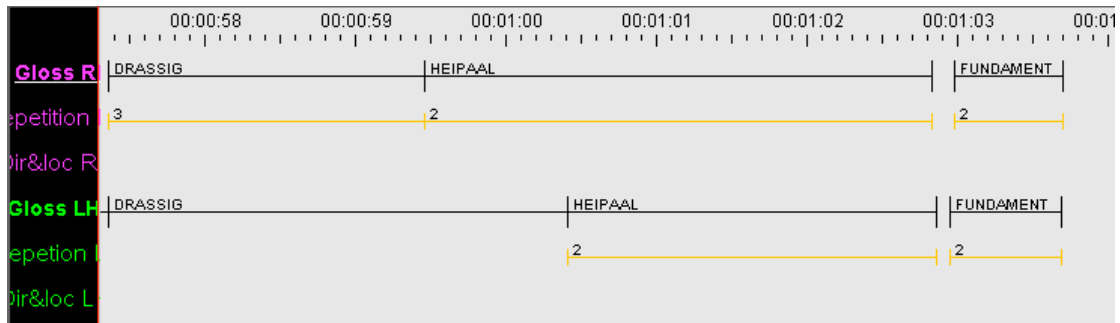
<b>GLOSS</b>	Use hyphen (-) if the gloss has more words: LOOK-AT
<b>(fs-) GLOSS</b>	Fingerspelled words are preceded by (fs-): (fs-) BUICK ((fs-) is not used for annotating initialized signs)
<b>IND</b>	Pointing sign, index: IND (for English and Dutch, PEK for Swedish)
<b>(2h)</b>	normally one-handed sign is made two-handedly: (2h)KIEZEN
<b>(1h)</b>	normally two-handed sign is made one-handedly: (1h)GEBAREN
<b>(p-)</b>	P stands for poly, meaning many meaning components. For example involves a classifier: (p-)vehicle-be-located

- (-h) identifies a final hold of the sign e.g. RUN (-h)
- (g-) the ‘meaning’ of a gesture is notated in lower case and placed between quotes (g-)“well”. When palms up is used we describe this as (g-)pu.

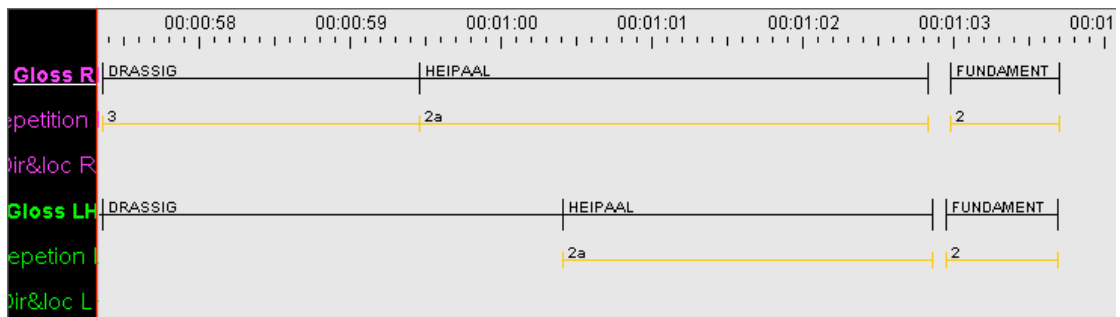
## 2.2 Repetition tiers

Repetition of a sign will be described in a separate tier: the repetition tier. The repetition tier is a child tier of the *Gloss hand tier*. It is tied to the repetition of the gloss in the right-hand tier or the left-hand tier. Thus, there is a *Repetition RH tier* and a *Repetition LH tier*. Number of repetitions will be marked with a number (2, 3, etc). When there are uncountable repetitions such as wiggling, trilled movement or oscillation the code ‘u’ (‘uncountable’) is included.

It is necessary to make a difference between alternating repetition and repetition of both hands at the same time (synchronic repetition).



When notated in the above way it appears that the repetition of HEIPAAL and FUNDAMENT are the same. But this is not the case. The repetition of the latter is made with both hands at the same time, so there are a total of 2 movement phases. The repetitions of the first are alternating (r,l,r,l), so there is a total of 4 movement phases. To make a difference between those two forms of repetition an extra code is required. When the repetition is made in an alternating fashion we will notate the number of movement phases for each hand, so in the case of HEIPAAL, ‘2’ for right and ‘2’ for left (r,l,r,l) and add the code ‘a’.



Note that repetition of a sign in different directions is glossed as different signs.

*Repetition RH or Repetition LH:*

<b>2, 3, 4 etc.</b>	number of movement phases
<b>u</b>	Uncountable repetitions (wiggling, trilled movement or oscillation)
<b>a</b>	repetition is alternating

### **2.3 Direction and spatial location tiers**

There is a spatial information tier for both the right hand and the left hand that can describe location and direction of movement. It is important to notice that we make a distinction between a *manual direction tier* and a *gaze direction tier*. The *direction and spatial location tier* is a child tier of the *gloss hand tier* (one for the right hand and one for the left hand). Note that this tier does not describe the lexical direction and location of the signs, but rather the actual direction and location of signs as they appear in the video source (which can be determined by morphological, syntactic, gestural, or other factors).

Categories for different directions and locations in space can be described with categories such as ‘to the right or left, close to 90 degrees’ (*r-90*, *l-90*) or ‘to the right or left, close to 45 degrees’ (*r*, *l*), to the left hand (*lh*) (when in Direction and spatial location RH tier), to the right hand (*rh*) (on the *direction and spatial location LH tier*). Many signs have a lexical default height. Height should not be described unless it is different from normal/predicted. When that is the case we use the categories ‘upward’ (*u*), ‘downward’ (*d*), ‘ahead, more to the front than lexical default proximity’ (*a*), ‘closer to the signer than lexical default proximity’ (*s*) and ‘toward a person present’ (*p*).

*Direction and spatial location RH and LH:*

<b>r-90</b>	to the right, close to 90 degrees (of midsagittal plane)
<b>r</b>	to the right, close to 45 degrees(of midsagittal plane)
<b>l-90</b>	to the left, close to 90 degrees(of midsagittal plane)
<b>l</b>	to the left, close to 45 degrees(of midsagittal plane)
<b>lh</b>	to the left hand (for RH tier) / <b>rh</b> : to the right hand (for LH tier)
<b>u</b>	upward, higher than lexical default height
<b>d</b>	downward, lower than lexical default height
<b>a</b>	ahead, more to the front than lexical default proximity
<b>s</b>	towards the signer, close to the signer
<b>p</b>	toward a person present

Combinations of the different features are possible e.g. **ru** (to the right and upward), **ld** (to the left and down) etc. So first notated left or right, then up or down.

#### 2.4 *Head tier*

In this tier a clear nod (*n*), shake (*s*) or tilt (*t*) (tilt = ear to shoulder) of the head is notated.

*Movement of the head:*

<b>n</b>	nod
<b>s</b>	shake
<b>t</b>	tilt

#### 2.5 *Eye brows tier*

In this tier the position of the brows is transcribed. We distinguish the positions: brows raised (*r*) and brows furrowed (*f*).

*Position of the eye brows:*

<b>r</b>	raised
<b>f</b>	furrowed

#### 2.6 *Eye aperture tier*

In the eye aperture tier we notate eye blinking, widely opening, narrowing and closing of the eyes with the following codes:

*Eye aperture:*

<b>b</b>	blink: eye blink (repetition should be included by - and number) e.g.: b-4
<b>w</b>	wide: eyes opened widely
<b>s</b>	squint: eyes narrow
<b>c</b>	closed: eyes closed

#### 2.7 *Eye gaze tier*

We make a distinction between a manual direction tier and a gaze direction tier. In the tier *eye gaze* one can notate the direction in which the signer looks. The gaze can be on the hands (left, right or both), to the left, to the right, upward, downward or toward a person present. For the material at hand (poetry and narratives) we also distinguish the direction toward the camera (*c*). The different values show some resemblance with the categories for the direction and spatial location, but there are some differences.

*Eye gaze:*

<b>l-90</b>	left, close to 90 degrees(of midsaggital plane)
<b>l</b>	left, close to 45 degrees(of midsaggital plane)
<b>r-90</b>	right, close to 90 degrees (of midsaggital plane)
<b>r</b>	right, close to 45 degrees(of midsaggital plane)
<b>u</b>	upward
<b>d</b>	downward

<b>lh</b>	to the left hand
<b>rh</b>	to the right hand
<b>bh</b>	to both hands
<b>p</b>	toward a person present
<b>c</b>	toward the camera

Combinations of the different features are possible e.g. **ru** (to the right and upward), **ld** (to the left and down) etc.

## 2.8 *Mouth tier*

For the notations of the mouth pattern in SSL we used the description proposed by Bergman & Wallin (2001). The annotation of the BSL and NGT material is annotated with slightly different labels and codes. Below you find a short description of the mouth conventions used for SSL and of the mouth annotation conventions for BSL and NGT. We refer to the mouth pattern documents for more information on both systems:

[http://www.let.kun.nl/echo/docs/transcr\\_mouth.pdf](http://www.let.kun.nl/echo/docs/transcr_mouth.pdf)

[http://www.let.kun.nl/echo/docs/transcr\\_mouth\\_SSL.pdf](http://www.let.kun.nl/echo/docs/transcr_mouth_SSL.pdf)

### 2.8.1 *SSL mouth annotations*

Bergman & Wallin (2001) describe a subset of mouth gestures, namely lexically determined and non-morphemic mouth gestures, referred to as mouth components. Mouth components do not carry a meaning on their own (oral/facial adverbs are excluded). Bergman & Wallin analyze these gestures as sequentially ordered combinations of a small set of open and closed segments. The description of the form of mouth patterns in terms of features is based on visible contrast. The different mouth patterns used in the Bergman & Wallin system are illustrated by means of pictures of mouth positions, the features involved in these positions (e.g. [-open]) and the label that is used in the annotation (e.g. /BILABIAL/). We refer to [the mouth pattern document](#) for more information on these features and labels.

Bergman & Wallin propose a small set of segments: three *closed* (/BILABIAL/, /CHEEKS/ and /LABIODENTAL/) and seven *open* (/STRETCHED/, /FORWARD/, /ROUND/, /PURSED/, /OPEN/, /AIRSTREAM/ and /TONGUE/).

A set of labelled segments allows us to represent mouth components of signs with monosegmental, bisegmental and trisegmental mouth components. Mouth components with four segments are actually reduplications; i.e. they do not contain four different segment types, only two, and may therefore be described as reduplicated bisegmental forms.

### ***Monosegmental Mouth Components***

<b>-open</b>	/BILABIAL/ /CHEEKS/
<b>+open</b>	/STRETCHED/ /FORWARD/ /ROUND/ /PURSED/ /OPEN/

### ***Bisegmental Mouth Components***

<b>-open, +open</b>	/BILABIAL, STRETCHED/ /FORWARD, BILABIAL/ /BILABIAL, ROUND/ /BILABIAL, OPEN/ /BILABIAL, AIR/ /BILABIAL, TONGUE/ /LABIODENTAL, ROUND/ /LABIODENTAL, FORWARD/ /LABIODENTAL, OPEN/ /LABIODENTAL, TONGUE/
<b>+open, -open</b>	/FORWARD, BILABIAL/ /FORWARD, LABIODENTAL/ /OPEN, CHEEKS/ /OPEN, LABIODENTAL/ /OPEN, BILABIAL/
<b>+open, +open</b>	/STRETCHED, PURSED/ /STRETCHED, OPEN/ /OPEN, STRETCHED / /OPEN, PURSED / /OPEN, TONGUE /

### ***Trisegmental Mouth Components***

<b>-open, +open, -open,</b>	/BILABIAL, ROUND, LABIODENTAL / /BILABIAL, OPEN, LABIODENTAL /
<b>+open, -open, +open</b>	/OPEN, CHEEKS, AIR/

### ***Reduplicated Bisegmental Mouth Components***

<b>-open, +open,-open,+open,</b>	/BILABIAL ,STRECTCHED, BILABIAL ,STRECTCHED /
<b>+open ,-open, +open, open</b>	/OPEN ,STRECTCHED, OPEN ,STRECTCHED /

Sometimes there are mouth components which should not be defined as /AIRSTREAM/, but do have a clear presence of air sucked in or breathed out. This can be represented by ‘-air in’, ‘-air out’, e.g. /ROUND-air in/ or /FORWARD-air out/.

#### ***2.8.2 BSL and NGT mouth annotations***

The annotation of the BSL and NGT material is annotated with slightly different labels and codes. After using the Swedish system for a while we felt that some mouth positions were missing. We thus started from scratch and tried to decide what the dimensions of the different mouth positions are (for instance lip opening, jaw opening, lip rounding, tongue protrusion) and what the extreme (or relevant) positions and possible combinations of these dimensions were. However the overall segmental

approach is the same. The positions are always related to the neutral position of the person's personal rest position of the mouth (so called 'neutral' mouth). We used different labels than the Bergman & Wallin system in order to reduce confusion between the two annotation systems. We refer to the mouth pattern document for more information on these features and labels.

*Lip aperture:*

<b>closed</b>	lips are closed
<b>open</b>	lips are not closed

*Lip position:*

<b>round</b>	lips are round (corners pulled toward the centre)
<b>forward</b>	lips are pushed forward
<b>stretched</b>	lips are stretched (corners pulled toward the sides of the face)

*Air:*

<b>ai</b>	air in; the impression that air is sucked in can be gained from the position and movement of the cheeks (sucked in) and/or the raising chest
<b>ao</b>	air out; the impression that air is coming out of the mouth can be gained from the position and movement of the cheeks (puffed) and/or the sinking chest

*Corners:*

<b>up</b>	corners of the mouth are pulled up
<b>down</b>	corners of the mouth are lowered

Some extra features are needed to distinguish the shape and position of the tongue and the touching of the lips by the teeth.

*Tongue shape:*

<b>pointed</b>
<b>relax</b>

*Tongue position:*

<b>10%</b>	10% out of the mouth
<b>30%</b>	30% out of the mouth
<b>60%</b>	60% out of the mouth
<b>100%</b>	100% out of the mouth

The percentages of tongue protrusion are arbitrary and can be changed according to the need for more or less distinctions.



*Teeth:*

**labiodental-up** the teeth touch the upper lip

**labiodental-low** the teeth touch the lower lip

The dimensions of mouth positions and the extra features can (physically) be combined in the following ways:

If the mouth movement consist of a transition from one mouth segment to another this is indicated by a '>'.  
>

Repetitions of mouth movements can be annotated as : -2, -3, -n.

For example : closed >open-3.

### **2.9 Cheeks tier**

There are only two values for the *cheeks tier* that are distinguished here, namely *in* and *puffed*. 'In' stands for cheeks sucked in and receives the value 'i'. When either the right, left or both cheeks are puffed this will be marked with 'p'. Repetition of puffed cheeks can be notated by adding a number e.g. 'p-3'.

*Cheeks:*

<b>i</b>	in
<b>p</b>	puffed

### **2.10 Role tier**

This tier shows when the participant uses role shift or takes role during production. By this we means the character from which the narration is given. So this can be one of the characters in the story or the narrator himself telling the story. If the story is told from the narrator's perspective we do not notate this. If the signer is telling the story from the perspective of one of the characters in the story we show this by selecting the right time interval and notating which role the signer takes e.g. hare, tortoise, dog, boy, etc.

### **2.11 Comments/notes tier**

The *comments/notes tier* gives the opportunity to add comments or notes about a segment of the video source. It is therefore important to specify and align the field before notating any notes or comments.

### **2.12 Translation tier**

The translation tier is divided in an *English Translation tier*, a *Swedish Translation tier* and a *Dutch Translation tier*. The Swedish and Dutch translation tiers are child tiers of the English translation tier. The translation tier will implicitly add information about modifications to the lexical meaning.

### 3. References

- B. Bergman & L. Wallin (2001) A preliminary analysis of visual mouth segments in Swedish Sign Language. In P. Boyes Braem & R. Sutton-Spence (eds.) *The hands are the head of the mouth*. Hamburg: Signum-Verlag. Pp. 51-68.
- B. Hellwig (2003) *EUDICO Linguistic Annotator (ELAN) version 1.4: manual*. Nijmegen: <http://www.mpi.nl/tools/elan.html>.
- A. Nonhebel, O. Crasborn & E. van der Kooij (2004) *Sign language transcription conventions for the ECHO project: BSL and NGT mouth annotations*. ECHO project, University of Nijmegen. [http://www.let.kun.nl/sign-lang/echo/docs/transcr\\_mouth.pdf](http://www.let.kun.nl/sign-lang/echo/docs/transcr_mouth.pdf).
- A. Nonhebel, O. Crasborn & E. van der Kooij (2004) *Sign language transcription conventions for the ECHO project: SSL mouth annotations*. ECHO project, University of Nijmegen. [http://www.let.kun.nl/sign-lang/echo/docs/transcr\\_mouth\\_SSL.pdf](http://www.let.kun.nl/sign-lang/echo/docs/transcr_mouth_SSL.pdf).