

Sign Language Text Transcriptions and Analyses Using ‘Microsoft Excel’

Penny Boyes Braem
Center for Sign Language Research, Basel

ABSTRACT

This is a short explanation of how we have used the commercial software program ‘Microsoft Excel’ for transcriptions of Swiss German Sign Language monologues and conversations. Although the transcription is not directly linked to video, it has the advantage that various kinds of searches and correlations can be carried out on it, that the same basic transcript can be expanded or reduced to be used for different kinds of analyses, and that the analyses can be easily represented and printed out in different forms (charts, graphs, etc.).

Type and Purpose of the Sign Language Corpus

The first study of Swiss German Sign Language (Deutschschweizerische Gebärdensprache, DSGS) which involved large texts was carried out in 1990 - 1995¹. The purpose of this study was to compare the signed communication of three deaf early learners of DSGS and three deaf late learners. Videotapes signed data consisted of 3 to 5 minute monologues from all six subjects, as well as five 10-minute conversations between different combinations of the subjects.² (Additional information about the methodology and subjects in this study can be found in Boyes Braem, 1995, 1999, 2000.)

Why Excel was Initially Chosen

The transcriptions of the signed data include German glosses for the manual signs, information about nonmanual signals used for grammatical and discourse purposes, ‘mouthings’, and a German translation of the signed text. At the time that these transcriptions were begun (1990), the computer software which would allow direct notation of videoclips on the computer screen, such as can now be done with programs such as SyncWriter, SignStream and MediaTagger, were in the process of being developed. At that time, there was no possibility to make queries and search for correlated data in any of these programs incorporating videoclips. We therefore decided to make all our transcriptions using a Macintosh version of the software ‘Microsoft Excel’.

In view of the limited resources of our very small research team (one hearing and two deaf linguists), the use of Excel offered the distinct advantage that we did not have to do any programming ourselves and that constant updates of the program would be available to us. As the program is relatively easy to use, our transcribers would

¹ The transcription method described here was developed for the Swiss National Science Foundation-sponsored projects Nrs. 11-28770.90 and 11-36347.92 (1991-1995): „An exploratory study of how age of acquisition affects forms of sign language used by the deaf in German Switzerland“. These projects were carried out by the Forschungszentrum für Gebärdensprache in Basel by P. Boyes Braem, T. Tissi and C. Jauch. (cf. Boyes Braem, 1995) The data and transcripts from this study were the basis for additional studies of mouthings (Boyes Braem 2000) and of prosody in DSGS (Boyes Braem 1999).

² There conversational situations were as follows:

- (a) One group with all three early learners;
- (b) One group with all three late learners;
- (c) Three groups in which one late learner was paired with one early learner.

require no special schooling. An additional consideration was the fact that because Excel runs on both PC and Macintosh computers, we can share our transcriptions with other researchers, even though their computer platforms might be different from ours.

Due to the complex nature of noting the several different kinds of simultaneously occurring components of sign language, transcriptions are often reworked and added to, according to the objectives of individual analyses. Such changing and elaborating of the transcript is much more easily done with computer-stored than with hand-written transcriptions. Full transcriptions can then be printed out only as needed.

Format of Excel Worksheet for Sign Language Transcriptions

The computer program allows a transcription of temporally successive linguistic units (usually manual signs) in separate cells. Different kinds of co-occurring linguistic elements (for example, the notation of individual non-manual components of facial expression) are transcribed on separate rows, co-ordinated with other temporally co-occurring components in the same column. The result is a format similar to that of a musical orchestral score. The advantage of the computer program is that elements can be added or deleted within any cell without destroying the vertical synchronisation between rows.

The horizontal rows in Excel worksheets are limited to 250 cells. At the beginning, we entered all our data along the horizontal axis, as it seemed easier to read from left to right. This meant that the transcriptions had to be broken up into files that were not longer than 150 signs.

The transcription can, however, be easily transposed to read vertically from top to bottom. In this format, thousands of cells are available and a whole story or conversation can be stored in one file. We quickly discovered that the vertical format was most useful in many of our analyses. In the end, we found it was just as easy to 'read' the transcripts vertically as horizontally (see Fig. 2).

Figure 1 is an example of the transcription format for the purposes of this study; a coding of 18 different kinds of information was made. (See the Appendix for a more detailed explanation of the information coded.)

1	Grouping	10					
2	Time-Beginning	0.549					
3	Time-End						0.593
4	Head		sl	Koo			Kor
5	Trunk						
6	Eye gaze	o	#	ad	# ad	ad	r
7	Eyes/Eyebrows						
8	Mouth	mt	mt	lv			
9	Cheeks						
10	Nose						
11	Expression						"schaut nach rechts"
12	Role	Narrator					Participant-1
13	Nondom. hand						
14	Both hands		[WAS?]	KURZ	ZUFALL (plötzlich)		
15	Dom. hand	SbwB (v, gemütlich)				ICH	SCHAUEN (vl>vr)
16	Handshape					☺	
17	Mouthing		ä	ku-r_	_z Zu_	_fa_	_a-ll
18	Translation	(10) Ich bin zu Fuss gegangen und habe zufällig geschaut, stimmt!					
19	Commentary					ICH = signer	

Fig. 1 Example of transcription format for one signed sentence

The transcription of conversations between two or three signers required the co-ordination of the 18 tiers for each person with all the tiers of the transcription of the other persons participating in the conversation. The temporal co-ordination of the individuals' transcripts is necessary for the analyses of the regulators of the conversation, overlapping comments, feedback, etc. This co-ordination was done by manually lining up the cells in the different transcripts according to the times in the time code tier, a process which proved to be especially time-consuming.

Printouts of the Transcriptions. Printouts can be made of all or of selected parts of the transcripts. Often only a small portion of the transcript was printed out, for use as an example in a lecture or article. Figure 2 shows an extract of only the manual and mouthing tiers of an early learner transcript, selected for the purpose of showing how a mouthing is stretched over two signs:

German Gloss for manual sign	SCHNELL-MACHEN	UNGEFÄHR	EIN	MINUTE	ICH	SICH-BEWEGEN-Bein
German Mouthing	<i>schne_</i>	<i>_ll</i>	<i>ei_</i>	<i>_n Mi_</i>	<i>_i_</i>	<i>_nute</i>

Fig. 2: Example of German mouthings stretched over 6 manual signs (early learner data).

Adapting the Transcription Format for Specific Analyses

New tiers and columns can be added at any time to the transcript, or to copies of the transcript for the analyses of special topics. For example, in a vertical version of the transcriptions, new columns were added for different types of analyses of the mouthings used with the manual signs. As Excel is a spreadsheet, quantitative totals can quickly be made. (See Fig. 3)

Gr.	Time Begin	Non dom hand	Both hands	Dom. hand	Mouthing	Total Nr. Mth	Number of Mouthings per Sign				Stretched mouthings		
							Signs with no Mth	Mth with no Signs	1Mth per Sign	2 or more Mth per Sign	over 2 signs	over >2 signs	over Gr. Bound
28	1.47			BIS	<i>bi_</i>	1					1		
	1.471		LEHRER		<i>_s Lehr</i>	1			1				
	1.475			Sbw1 (or-vm)			1						
	1.485			GLAUBEN	<i>glaub</i>	1			1				
	1.488		BRECHEN		<i>ge_</i>	1						1	
	1.49		FERTIG		<i>bro_</i>								
	1.493				<i>_ochen</i>			1					
29	1.496			ICH			1						
	1.498		SbwB (v, 'aufstehen')				1						
	1.504			ICH			1						
	1.505												
	1.508		PRIMA		<i>gu_</i>	1	1				1		
TOTALS						11	10	1	6		2	3	
% of Total Mouthings									55%		18%	27%	

Fig. 3: Excerpt of basic transcript with columns added for coding co-ordination of mouthings with manual signs

Computer-aided Analyses

'Macros' for simple queries and correlations

Once stored in a computer program, the sign language transcriptions can be used for a wide variety of computer-aided analyses. For example, simple queries and correlations could be made of the data (sometimes using macros that we wrote ourselves). The 'macro' programs are written in the script language of Excel. They are available through a special menu appended to Excel's standard menu. There are two kinds of program: service programs and analysis programs. The service programs prepare transcript files, in Excel format, for analysis and write the analysis results to a cumulative results file for each transcript.

The analysis programs provide simple statistical analysis of a transcript. The basic operations of the analysis programs are: (1) to extract various components or combination of components from the transcripts, (2) to search for specific signs or types of signs, (3) to compare how often various components occur and (4) to measure timing. These are described in more detail below:

(1) Extraction program:

For example:

- An alphabetised list of all signs which occur in the data. This can be done for all unique signs occurring in the transcript and for all occurrences of all signs.

(2) Search for specific signs, types of signs or combination of signing elements:

For example:

- List and/or total number of the different types complex polymorphemic verbs;
- A list of all 'mouthings' together with the manual components which might accompany them (with and without repetitions).

(3) Comparison of how often various components occur:

For example:

- The ratio of 'all signs' to 'signs accompanied by a mouthing';
- The use of the indexical pronominal sign 'I' to all uses of indexical deictic signs (INDEX)
- The ratio of INDEX signs with and without 'mouthings'

(4) Temporal Calculations:

For example:

- Of the duration of each "sign grouping" (= larger discourse/prosodic units);
- The speed of signing per "grouping" and a running average of the number of signs per "grouping".

Graphic Representations

The program Excel offers a large number of virtually automatic charting tools, making it easier to visualise otherwise complex data relationships. For example, the information extracted by the macros for the analyses of the mouthings could very easily be converted to graphic representation (See figure 4). The results of the additional analyses done on the expanded transcripts can be shown in pie graphs (Figures 5). Figure 6 is a graphic representation of the duration of pauses between sentences, data that was extracted by a macro

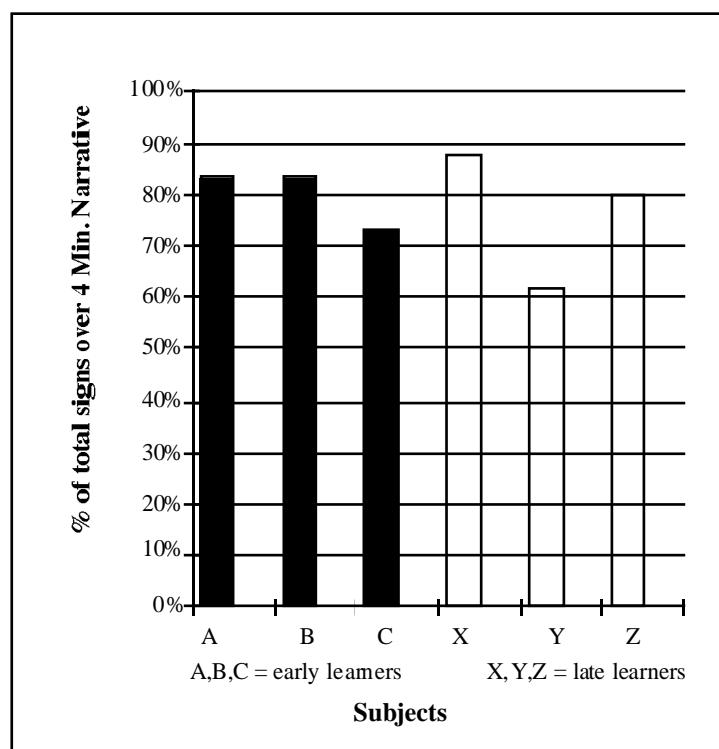


Fig. 4: Percentage of manual signs accompanied by mouthing over 4-minute stretches of spontaneous signed narratives of early and late learners

Coordination Techniques used by Late Learners

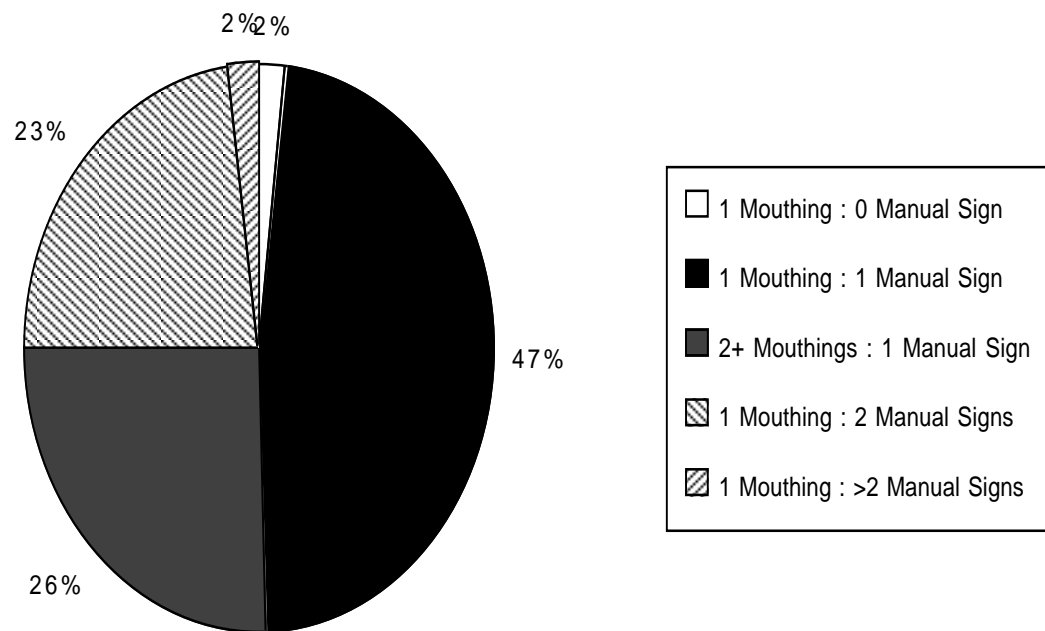


Fig. 5: Techniques for co-ordinating mouthings and manual signs used by deaf late learners of Swiss German Sign Language (DSGS)

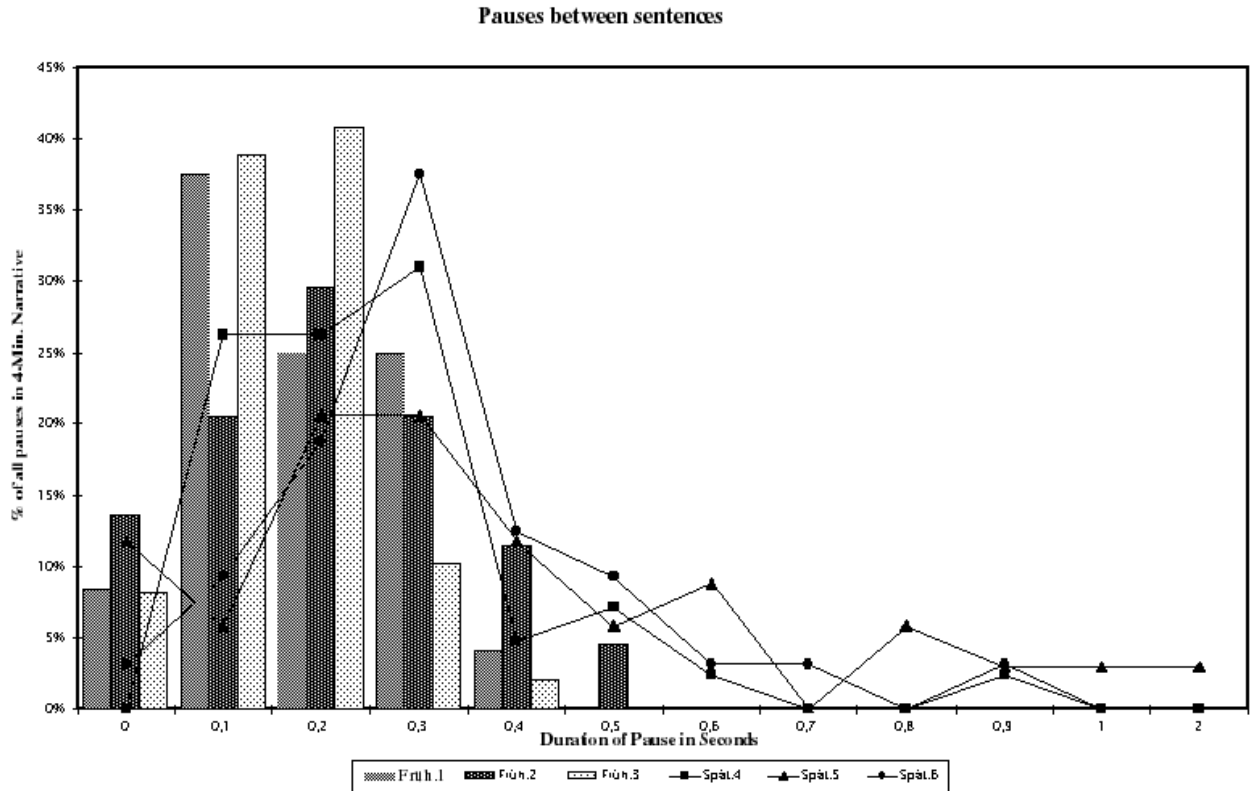


Fig. 6. Duration of pauses between sentences in signing of deaf early and late learners of DSGS.

Using the Excel transcriptions for other projects

In the project for which the Excel transcripts were made, the assigning of a German gloss to each sign turned out to be not as straightforward as was originally assumed. As there no standardised form of Swiss German Sign Language and also no existing dictionaries of the language to which to refer, the deaf transcribers, in their first transcriptions, sometimes using the same German word gloss for several different signs. Or, alternatively, a single sign was given different German glosses by the transcribers. When this problem became evident, we began compiling a separate file of glosses we began to use in the transcripts to ensure that every unique sign form in our data was given a separate and consistent German gloss. This served as the basis of a following project, a multimedia lexicon databank for the language (1996 – 2001). In addition, the Excel transcripts can be searched for sentence examples of signs, which can then be entered as ‘example sentences’ in the lexical database.

References

- Boyes Braem, P. (2000). Functions of the Mouthing Component in Swiss German Sign Language. In Brentari, D. (Ed.) *Foreign Vocabulary in Sign Languages*. Mahwah, N.J.: Lawrence Erlbaum and Associates.
- Boyes Braem, P. (1999). Rhythmic temporal patterns in the signing of early and late learners of German Swiss Sign Language. *Language and Speech* (Special issue on prosody in spoken and signed languages, edited by W. Sandler and M. Nespore.). Vol 42 (Parts 2 & 3 April – Sept. 1999) pp. 177-208.
- Boyes Braem, P. (1995). *Eine Untersuchung über den Einfluss des Erwerbsalters auf die in der deutschsprachigen Schweiz verwendeten Formen von Gebärdensprache: Ein Überblick zu einem vom Schweizerischen Nationalfonds unterstützten Projekt des Forschungszentrum für Gebärdensprache, Basel 1991-1995*. Informationsheft Nr. 27. Zürich: Verein zur Unterstützung der Gebärdensprache der Gehörlosen.
- Engberg-Pedersen, E. (1991). *Space in Danish Sign Language*. Hamburg: Signum Verlag.

APPENDIX

Explanations of Elements Coded in the Transcription

Row 1: Groupings. The deaf persons doing the transcribing (all native signers or early learners of DSGS) were asked to look at the videotaped data and to note on the transcripts 'major breaks' in the signing strings. The resulting groupings between major breaks turned out to contain one or several underlying semantic propositions and seem, upon first analysis, to correspond sometimes to sentences, other times to larger discourse or prosodic units. These major groups were numbered and their ends marked with //. Subgroups could also be noted in this row, as needed for further analyses.

Rows 2 & 3: Time Beginnings and Ends. These are notations (to 0.1 sec), entered by hand, of where the groupings occur on the videotape, for the purposes of cross-referencing with the original tapes. However, the beginning and end times is sometimes noted for much smaller units, depending on the analysis. For the analysis of upper body movements which were thought to be some kind of prosodic marking, for example, the on and offset of each sign was entered. Analysis programs ('macros') can use the data from these two tiers to calculate temporal aspects of different kinds of linguistic organisation, including stress, rhythm as well as the shortening or prolonging of individual signs for semantic purposes.

Rows 4 – 10: Nonmanual components. These are rows for more detailed notations of selected non-manual behaviours (movement/orientation of the head, body, eye gaze; specific changes in shape of eyes/eyebrows, mouth, cheeks and nose.) Combinations of specific non-manual behaviours are used, among other things, to signal sentence type (questions, conditionals, relative clauses) or give specific adverbial/adjectival information. These non-manual elements were noted with a system of German abbreviations, but other kinds of notation symbols could be used (Facial Action Coding System, SignWriting, etc.)

Row 11: Expressive elements: This row is used by the deaf transcriber for a subjective labelling of combinations of face and body signals which seem to signal different signer 'identities' (e.g. a facial and body expression of 'being agitated' which correlates with the identity of the referents or characters in the signed text.) The transitions between signer identities can often be correlated with elements in the non-manual signals.

Rows 12: 'Role': Here, in addition to clear signer or participant roles, role mixtures are also notated: Narrator/Participant where the signer identify predominates and Participant/Signer where the participant predominates.

Rows 13 - 15: Manual signs. Three separate rows are used to separately code signs made with the dominant, the non-dominant or with both hands. This separation is especially important for syntax and discourse analyses in which separate morphemes or lexical items can occur simultaneously on the dominant and non-dominant hand.

The manual signs are, following usual sign language transcription conventions, using spoken language (here, German) words written in all capital letters. Grammatically important spatial information (where the sign is made, or the direction it is orientated or moves in space) is noted in parentheses before and after the gloss. The notation system used for these spatial indices utilises some of the notation for direction developed by Engberg-Pedersen (1990) for Danish Sign Language. Also in parentheses following the gloss is information about the semantically significant manner in which the sign is produced (Exp.: CHANGES-LOCATION ("leisurely")) as well as whether a plural form has been used. Signs which seem to have a predominantly discourse function are enclosed within square brackets [].

Row 16: Handshape. Notations are made in this row when non-canonical handshapes are used in signs. This information is important not only for indicating 'wrong' handshapes (for example, in some of the signs in our data produced by the late learners of DSGS), but also for notating precisely which handshape is being used as a 'pro-form' in polymorphemic verbs. This row is formatted with the HamNoSys computer font.

Row 17: Mouthings: 'Mouthings' are mouthings with no voicing of words or parts of spoken language words. Although the use of mouthings is typical of all Swiss German signing, the early and late learners seem to use these verbal gestures differently. Therefore, care has been taken to note this component as precisely as possible. The deaf transcriber has transcribed what, to her, is readable from the lips. The dashes indicate when verbal gestures are prolonged to accompany the following manual sign. Sometimes (especially in the late learners' data), a verbal gesture will be the main linguistic component defining the column, unaccompanied by a manual sign.

Row 18: Translation. A trained sign language interpreter, bilingual in German and Swiss German Sign Language, made a translation into German from the videotapes. This translation was then typed entered into the computer transcript. A full spoken language translation can be helpful for the analysis in several ways. The content of the German sentence can stem from several different kinds of sign language components or signals (for example, information carried on the hands or on the face). Particularly for the analysis of the information carried

by the non-manual components, it is sometimes helpful to look at the German translation and then try to determine which non-manual components conveyed this information. The translations can also be helpful for discourse level analyses, for which the entire text is broken down into discourse episodes, events, etc.

Row 19: Commentary: This row is where diverse observations, explanations and comments on the signed text can be noted. For example, an explanation can be put here of why an unconventional form of a sign (as noted by an * before the gloss) might have been used. The linguistic referent of pro-forms indicated by the handshape in polymorphemic verbs can also be identified in this row.