

Sign Language Corpora Survey* (last update: September 2010)

Sign Language	Research project	Raw data	Metadata	Primary data	Resources and references	Contact	mailto																		
Name (abbreviation)	English translation (Ethnologue code)	Project name (time frame)	Research objectives	Time span of data collection	Amount of raw data (h:min)	Digitized	Accessibility	Kind of data	Number of informants	Age (year:month)	Sign language proficiency	Sociolinguistic background: elicitation setting	Transcribed	Transcribed with ...	Lemma-ized	Lemma-ized using ...	Amount of transcribed data	Annotated	Kind of annotation	time-aligned	Annotation tools				
Deutsche Gebärdensprache (DGS)	German Sign Language (gsg)	DGS Corpus Project (2009-2023)	documentation (open archive), general purpose corpus of DGS, lexicography (lexical variation, semantics, morphology, phonology)	winter 2009-2011	scheduled: 350-400 h	yes (until sept. 2010: 45 sessions; 92 informants, 5 regions)	scheduled: partial corpus of about 50 h (online)	dialogues free conversation elicited narratives elicited lexical items	scheduled: 328	18-7	fluent DGS signers (native signers and others)	yes	(not yet)	-	-	lexical database, web-based voting is scheduled	-	(not yet)	-	-	ILex	Nishio et al. 2010, König/Langer (2009), Prillwitz et al. (2008) project homepage: http://www.dgs-korpus.de	Thomas Hanke Susanne König Reiner Konrad Gabriele Langer Christian Rathmann	info (at) dgs-korpus.de	
		Classifying sign language texts by genres based on a digitized corpus	digital library of st texts, archive; linguistic features of st text genres	digitizing data from 1980 and later, new data from 2009	scheduled: more than 200 h	yes	scheduled for 2010: small section online, for researchers (with permission)	dialogues narratives free conversation	more than 100	all ages	native signers CODA interpreters	no	partial	glosses	scheduled	-	n/a	yes	orientation/direction or location	yes	ELAN	project homepage: http://www.sgw-hs-magdeburg.de/dgstatyp/index.html	Jens Hessmann	Jens_Hessmann (at) hs-magdeburg.de	
		6 st dictionaries of technical terms: Psychology, Caparity, House economics, Social work/Social pedagogics, Health and Nursing, Nursery and Landscaping	corpus based selection of DGS equivalents of technical terms	1994-2008	21:30 h interviews, 67:30 h conversation, 247 h elicited data (spontaneous answers for technical terms stimuli)	yes	no	interviews prestructured conversation elicited data	79:20:70	20-70	native signers early late learners	yes	partial (reviewed elicited answers and selective in conversations)	glosses and HamNoSys	yes	lexical database	about 88 h (152.000 tokens)	yes	technical term, mouthing, facial expression, meaning	yes	ILex	Arbeitsgruppe Fachgebärdensprache (1996, 1998) Konrad et al. (2000, 2003, 2007, 2010) online version of dictionaries: http://www.sign-lang.uni-hamburg.de/projekt/abg_projekte.html (see "Fachgebärdensprache")	Reiner Konrad	Reiner_Konrad (at) sign-lang.uni-hamburg.de	
		JEHÖRLOS SOF Corpus	natural dialogues and annotated corpus for teaching DGS	may 1993	about 21 h	partial	partial: 10 x 5-7 min (total: 57 min)	spontaneous dialogues (themes are chosen of a list of subjects, e.g. school, sports, events, etc.)	20:20:78	20-78	native signers	yes	partial	glosses	yes	lexical database	10 x 5-7 min (total: 57 min)	yes	mouthing, facial expression, mime german translation, interpretation and comments	yes (time code)	FileMaker, ELAN	Heßmann (2001a, b) http://corpus1.mpi.nl/ids/mdl_browser/0/penpath=MP184302%23 (nodes: ECHO-ECHO-Sign Language-DGS-Dialogue)	Jens Hessmann	Jens_Hessmann (at) hs-magdeburg.de	
Deutsch-schweizerische Gebärdensprache (DSGS)	Swiss-German Sign Language (sgg)	A Multimedia Online Lexicon for Technical Terms in German and Swiss German Sign Language (DSGS)	dictionary of technical terms (economy, nutrition)	nov. 2007- aug. 2009	economy: 18 h (elicitation), 2:30 h (definitions and example sentences); nutrition: in progress	yes	group discussions: only via database access. Definitions and example sentences: online	group discussions: meaning of technical terms and their translation in DGS	8:25:45	1 native signer 7 early learners with 1 hard of hearing	yes	partial: group discussions and definitions	glosses	yes	lexical database, web-based feedback function is scheduled for online dictionary	yes	-	-	yes	ILex	project description (in German): http://www.hfh.ch/projekte_detail-n70-674-sd.html	Penny Boyes Braem	boyesbraem (at) gmail.com		
		DSGS CD-ROM for learners, level 4	natural texts for teaching DGS	may-nov. 2005	2:56 h (32 video clips)	yes	published as CD-ROM (translations only)	dialogues and monologues, spontaneously signed but on a given general theme, different registers (information, interviews, debate, signing with children, storytelling)	33:4:40	4-80	early learners	yes	yes	glosses	yes	lexical database	totally transcribed	yes	productive signs (classifiers), role shift, nonverbal communication	yes	XML-Format	DSGS-Kurs (2005)	Penny Boyes Braem	boyesbraem (at) gmail.com	
		Religious signs and bible texts in Swiss German Sign Language	signs for religious terms and translations of bible texts in DGS	mar. 2002- may 2005	42 min (22 video clips)	yes	published as CD-ROM (translations only)	formal register: translation of Bible texts into DGS, signed by one model	10:35:55	35-55	early learners 2 late learners	no	yes (DSGS translations)	glosses	yes	lexical database	totally transcribed	no	-	yes	XML-Format	Christliches DSGS-Lexikon (2005)	Penny Boyes Braem	boyesbraem (at) gmail.com	
		A Multimedia Bilingual Database for Swiss German Sign Language	lexicography; building a lexical database	1996-2000	about 60 h (about 30 video clips)	in progress	only for researchers (with permission)	informal group discussions on given themes (medicine, sports, schooling, etc.)	14:21:76	21-76	early learners	yes	yes	glosses, only handform in HamNoSys	yes	lexical database	3000 (now extended to 9000)	type information (lexical database)	form, meaning, POS, modification, synonyms, antonyms, homonyms, lexical variants, loan signs, semantic fields, mouthing, mouth gesture	yes (time code on VHS-video)	FileMaker Pro	Boyes Braem (2001a)	Penny Boyes Braem	boyesbraem (at) gmail.com	
Österreichische Gebärdensprache (ÖGS)	Austrian Sign Language (asg)	SignLex: Formation and further education, st dict. of technical terms German-Austrian SL	st dictionary of technical terms	sept.-nov. 2006	5616 video clips (about 2 sec)	yes	no	elicited lexical items (word list)	16:20:30 (8 persons) 31-40 (5 persons) 41-60 (3 persons)	20-30	native signers	yes	no	-	-	-	-	-	-	-	-	Grbic (2007) Grünbichler/Hofstätter (2008)	Barbara Andree Sylvia Grünbichler	barbara_andree (at) uni-graz.at sylvia_gruenbichler (at) uni-graz.at	
		SignLex: Labour & Social Affairs, st dict. of technical terms German-Austrian SL	st dictionary of technical terms	may-oct. 2004	5776 video clips (about 2 sec)	yes	no	elicited lexical items (word list)	16:21:30 (5 persons) 31-40 (8 persons) 41-50 (2 persons) 51-60 (1 person)	21-30	20-30	native signers	yes	no	-	-	-	-	-	-	-	-	Grbic (2005) Grünbichler/Hofstätter (2008)	Barbara Andree Sylvia Grünbichler	barbara_andree (at) uni-graz.at sylvia_gruenbichler (at) uni-graz.at
Nederlandse Gebarentaal (NGT)	Dutch Sign Language (dse)	Corpus NGT (2006-2008)	documentation (open archive) and sociolinguistic variation, annotated corpus	2006-2007	72 h (2375 video clips, from 5 sec to 16 min each)	yes	yes (open access with Creative Commons license BY-NC-SA); http://www.corpusngt.nl ; for researchers: http://corpus1.mpi.nl/ids/mdl_browser/0/penpath=MP184302%23	dialogues, narratives elicited narratives (comics, tv spots, NGT video clips)	92:17:86	17-86	native signers	yes	yes	glosses	no	-	10 h	partial	mouth movements (phonetic description and classification)	yes	ELAN	Ormel et al. 2010, Crasborn et al. (2008) Crasborn/Debesse (2008a) Crasborn/Zwiers (2008b) corpus homepage: http://www.ru.nl/corpusngt/	Orno Crasborn	o_crasborn (at) let.ru.nl corpusngt (at) let.ru.nl	
		European Cultural Heritage Online (ECHO)	documentation (open archive) and sociolinguistic variation, annotated corpus	2003-2004	about 30 min per st	yes	yes (open access with Creative Commons license BY-NC-SA); http://corpus1.mpi.nl/ids/mdl_browser/0/penpath=MP184302%23	elicited narratives (written texts) and lexical items (word list) (see IMDI metadata)	4 for NGT 2 for STS	20-40	20-40	native signers	yes	yes	glosses	no	-	about 30 min per st	partial	direction and location, head movement, eye gaze, eye aperture and eye brows, mouth gesture, cheeks, role shift, comments, translation (English, Dutch, Swedish)	yes	ELAN	Crasborn et al. (2007, 2004) Bergman/Mesch (2004) Woll et al. (2004) project homepage: http://www.let.ru.nl/sign-lang/echo-corpus ; http://corpus1.mpi.nl/ids/mdl_browser/0/penpath=MP184302%23	Orno Crasborn	o_crasborn (at) let.ru.nl
Langue des Signes Française (LSF)	French Sign Language (fs)	LS-COLIN (Langues des Signes – Cognition, Linguistique et Informatique; 2000-2002)	general purpose corpus and development of annotation tools, testing linguistic hypothesis (esp. iconic structures)	jan. 2002	2:12 (80 video clips)	yes	yes (video clips and translations); no transcripts	elicited narratives free monologues on given themes (recipes, linguistics, news)	13:22:61	22-61	fluent LSF signers	yes	yes	glosses	n/a	-	1:05 min (39 video clips, 3 per informant)	yes	highly iconic structures (see Cuxac 2000; Sallandre 2003a); facial expression, mime, eye gaze	approximate time codes in transcripts	MS-Excel	LS-COLIN (2002) Sallandre (2003a-50-104) video clips: http://corpusdeslangues.in2p3.fr ["> recherche"; "Recherche par langue"; enter: "LSF"]	Marie-Anne Sallandre Ivni Fusellier Christian Cuxac Christian Cuxac	marie-anne_sallandre (at) uni-paris8.fr ivni_fusellier (at) uni-paris8.fr christian_cuxac (at) uni-paris8.fr christian_cuxac (at) uni-paris8.fr	
		CREAGEST (2007-2011)	documentation (open archive) of LSF of young signers	about 100 h	yes	scheduled	elicited dialogues free conversation explanation of technical and metalinguistic subjects; peer-to-peer (d-h, d-d, d-child-d adult)	57:3:12:0	3-12:0	3 x 10 pairs	yes	partial	n/a	-	-	-	-	-	-	yes	ELAN	Balvet et al. 2010 ProjetBeschreibung: http://www.umr7023.cnrs.fr/Realisation-de-corpus-de-donnees.html	Christian Cuxac, Brigitte Garcia	christian_cuxac (at) uni-paris8.fr brigitte_garcia (at) uni-paris8.fr	
Langue de Signos Española (LSE)	Spanish Sign Language (ssp)	General purpose corpus, variation	corpus for research (transcription, grammar, language acquisition, linguistic universals), st teaching supported by natural data, language planning	1995-2000, 2008... (in progress)	454 min	in progress	scheduled	monologues prestructured conversation free conversation (formal, colloquial and scientific register)	85:21:35 (majority) 36-50 (less than 10) over 50 (less than 5)	21-35	native signers late learners L2-learners	no	yes	glosses	no	-	about 25 %	no	-	yes	MS Word	Álvarez Sánchez et al. (2010, 2008) Báez Montero/Cabeza Perero (1999a, b, 1995)	Patricia Álvarez Concha Báez Ana Fernández	patiavarez (at) unigo.es cbaez (at) unigo.es anafae (at) unigo.es	
		British Sign Language (BSL)	British Sign Language Corpus Project	general purpose corpus (open archive), sociolinguistic variation and language change	aug. 2008-... (in progress; 2 stages)	180 h (in progress)	yes	scheduled in 2011 partial (online); further access for researchers (with permission)	stage 1: personal experience narratives, free conversation, interview, lexical elicitation task; stage 2: elicited sentences, elicited narration, discussion/interview	scheduled: 240	18-7	native signers early learners (under 7 years old)	yes	no	-	scheduled	-	-	-	-	yes	n/a	project homepage: http://www.bsllcorpusproject.org/	Adam Schembri	a_schembri (at) ucl.ac.uk
Dansk Tegnsprog (DTS)	Danish Sign Language (dsl)	The Danish Sign Language Dictionary Project	lexicography (lexical variation, semantics, morphology, phonology)	feb. 2004 - apr. 2007	73 h (207 video clips)	partial: 13 h	no	13 h of monologue without any prompts 60 h of dialogue, partly elicited sentences, discussions and free conversation	25:19:62	19-62	native signers	yes	partial: example sentences for the dictionary	glosses	n/a	-	see "Transcribed"	no	-	no	MS Access	Troelsgård/Kristoffersen (2008) Kristoffersen/Niemelä (2008) dictionary: www.tegnsprog.dk	Jette Hedegård Kristoffersen	jthk (at) ucc.dk	
		Irish Sign Language (ISL)	Signs of Ireland Corpus	general purpose corpus for teaching and learning and for ISL research. Focus on the morpho-syntax of ISL, with some information noted regarding idiosyncratic behaviours; gesture used, variants (gendered, regional, age related), information on mouthing, eye gaze, eye aperture, point of view, and manual simultaneity	2004	40 monologues (2-10 min each) 40 frog stories (8-10 min each; see Mayer 1969) 40 elicited utterances	yes	intern for research and teaching and for ISL research (with permission)	monologues (narratives) elicited narratives (frog story) elicited utterances (drawings, see Volterra et al. 1984)	40:18:70+	18-70+	native signers early learners (under 6 years old)	partial	partial: 40 monologues, 10 frog stories (Mayer 1969)	glosses (for ECHO transcription conventions see Nonhebel et al. 2004a)	n/a	-	40 monologues (2-10 min each) 10 frog stories, 8-10 min each	scheduled	-	yes	ELAN	Leeson (2008) Leeson/Nolan (2008) Leeson et al. (2006)	Lorraine Leeson	leeson (at) tod.ie
Islensk Tákmal, Dansk Tegnsprog (DTS)	Icelandic Sign Language (idl), Danish Sign Language (dsl)	A Lexical Comparison of Icelandic and Danish Sign Languages	lexical and phonological similarities	sept.-oct. 2006	6000 video clips about 5-10 sec each	yes	some examples as part of the online publication of Alderson/McEntee-Atalians (2008)	elicited lexical items (word list)	je 10 pro Gebärdensprache	25-75	ISL: 7 native signers 3 interpreters with 1 CODA DTS: native signers	yes	no	-	-	-	-	-	yes	only classified in same, similar, different sign	no	-	Alderson/McEntee-Atalians (2008) download-version (Word file with video-clips (online)): http://www.bsll.bb.ac.uk/publications/volume2/pdf/articleword	Russell Alderson	Russellalderson (at) gmail.com
		American Sign Language (ASL)	American Sign Language Lexicon Video Dataset (ASLLVD)	see below: ASLLRP	2007... (in progress)	about 90 min in total (8400 video clips)	yes	yes	isolated signs	n/a	19-28	1 native signer 2 CODA 2 L2-learners	yes	scheduled	glosses	yes	gloss list	about 50 % of raw data	scheduled	-	yes	SignStream (Mac OS 9); scheduled: http://csr.bu.edu/asl/aslvid/annotate/index.html new version summer 2009)	Athitos et al. (2008) project description: http://csr.bu.edu/asl/aslvid/annotate/index.html	Vassilis Athitos Carol Neidle Stan Scaroff	athitos (at) uta.edu carol (at) bu.edu scaroff (at) bu.edu
American Sign Language (ASL)	American Sign Language (ase)	American Sign Language Linguistic Research Project (ASLLRP; Boston University)	linguistic analysis of ASL with focus on syntax; automated sign recognition	2007... (in progress)	n/a	yes	yes	more than 400 elicited sentences 15 narratives in ASL	4:20:45	20-45	native signers	yes	yes	glosses	yes	gloss list	n/a	yes	POS, non-manual gestures (form and function), English translation	yes	SignStream (Mac OS 9); scheduled: http://www.aslproject.org/ new version summer 2009)	Drew et al. (2008) Neidle (2007, 2002, 2001) Neidle et al. (2001) McLaughlin et al. (2000) Neidle (without year) corpus: http://www.bu.edu/asllrp/calgr/ Database Access Interface (DAI): http://ling.bu.edu/asllrp/data/queryPages/CD-ROM/order/ ; http://www.bu.edu/asllrp/pdf/	Carol Neidle	carol (at) bu.edu	
		ASL variation project	sociolinguistic variation	62 h group discussions 32 h interviews and elicitation	yes	partial	62 group discussion (2-7 persons) interviews elicited lexical items	207	207	yes	yes	n/a	-	-	-	-	-	-	yes	ELAN	Valli et al. (2005) Lucas et al. (2001a)	Ceil Lucas	ceil_lucas (at) ucl.ac.uk		
Australian Sign Language (Auslan)	Australian Sign Language (asf)	Endangered Languages Documentation Project (2004-2007)	documentation (open archive) general purpose corpus; lexicography, syntax, semantics, morphology	2004-2005	300 h 150 h reviewed for transcription (1200 video clips)	yes	for other researchers (with permission); scheduled in 2012: open access (with registration)	interviews elicited narratives group conversations elicited lexical items	100:18:85	18-85	native signers	yes	yes	glosses (D-gloss)	yes	-	less than 5 % (dec. 2008)	yes	POS, constructed action, semantic role (macro role), sentence segment, spatial modification, mouthing, mouth gesture, aspect, literal and free translation	yes	ELAN	Johnston (2008a) Johnston/Schembri (2006a) project outlines: http://www.hrep.org/grants/projects/index.php?projid=48 http://www.auslan.org.au/about/corpus/	Trevor Johnston	trevor_johnston (at) mq.edu.au	
		Sociolinguistic variation in Auslan project (2003-2004)	sociolinguistic variation (grammar and lexicon)	2003-2004	140 h	in progress	for researchers (with permission)	interviews free conversation elicited lexical items	211:15:89	15-89	native signers early learners (under 7 years old)	yes	yes	glosses	yes	-	small section	yes	POS, grammatical aspects, phonological features	small section	ELAN	Schembri/Johnston (2004)	Trevor Johnston Adam Schembri	trevor_johnston (at) mq.edu.au a_schembri (at) ucl.ac.uk	
New Zealand Sign Language (NZSL)	New Zealand Sign Language (nzs)	Sociolinguistic variation in New Zealand Sign Language	sociolinguistic variation in NZSL	2005-2008	100 h	yes	no, internal only	free conversation (70 %) interviews (30 %)	150:16:91	16-91	early learners (90 % under 7 years old, 10 % under 12 years old)	yes	yes	glosses	no	-	2 h (about 2 mins each of 60 signers)	yes	POS, mouthing, variation information, verb types, demographic information, English translation	yes	ELAN	project description: http://www.victoria.ac.nz/lares/research/projects/dsu.aspx	David McKee	david_mckee (at) vuw.ac.nz	
		NZSL-Dictionary Project (1992-1998)	comprehensive dictionary of NZSL	1992-1994	about 25 h	no	no	guided conversational data collected with 2-3 informants	Core group of 13, validated by 90	19-60	19-60	early signers	yes	partial	glosses	no	-	only target lexical items, not in full	no	-	no	MS Word, WordSmith	Kennedy et al. (1998) Kennedy (1998) electronic version (prototype): http://homepages.mcs.vuw.ac.nz/~pond/nzsl/	David McKee	david_mckee (at) vuw.ac.nz
Korean Sign Language (KSL)	Korean Sign Language (kvk)	Agreement verbs in KSL	morphology and phonology of agreement verbs in KSL	july 2003- oct. 2005	21 h (100 video clips)	yes	scheduled	dialogues, monologues elicited signs, sentences and narratives discussions prestructured interviews free conversation	36:20:70	20-70	native signers CODA late learners	yes	selective: agreement verbs	glosses and HamNoSys	n/a	-	see "Transcribed"	yes	morphological features	yes	ILex	Hong (2008)	Sung-Eun Hong	Sung-Eun_Hong (at) sign-lang.uni-hamburg.de	
		Animal classifiers in KSL	morphology and phonology of classifier constructions in KSL	jan. 2000	more than 68 min	68 min	scheduled for researchers (with permission)	elicited descriptions (jumping course model as stimulus)	8:17:65	17-65	native signers late learners	yes	selective: classifier constructions	glosses and HamNoSys	n/a	-	see "Transcribed"	yes	phonological and morphological features	yes	FileMaker	Hong (2003, 2002, 2001)	Sung-Eun Hong	Sung-Eun_Hong (at) sign-lang.uni-hamburg.de	

*This survey originates from the doctoral thesis of Reiner Konrad: *Die lexikalische Struktur der DGS im Spiegel empirischer Fachgebärdensprachographie. Zur Integration der Ikonizität in ein korpusbasiertes Lexikonmodell*. Universität Hamburg. [Unveröffentl. Dissertation].
[The lexical structure of German Sign Language (DGS) as mirrored by empirical sign language lexicography of technical terms. A corpus-based lexicon model of DGS taking into account the iconicity of signs. University of Hamburg. [Unpublished dissertation].]

**Ethnologue: http://www.ethnologue.org/show_family.asp?subid=90008

Legend
n/a = not applicable
sl = sign language
d = deaf
h = hearing