

Exploring Lexical Variation in a Growing Corpus of DGS



Sabrina Wähl, Gabriele Langer, Anke Müller, Julian Bleicken, Thomas Hanke, Reiner Konrad University of Hamburg, Institute of German Sign Language and Communication of the Deaf

AKADEMIE DER WISSENSCHAFTEN IN HAMBURG

DGS Corpus recorded 2010-2012

- Number of informants 330
- Controlled sample balanced for
- gender
- 13 regions
- 4 age groups: 18-30, 31-45, 46-60, 61+
- Native and near-native signers



Data

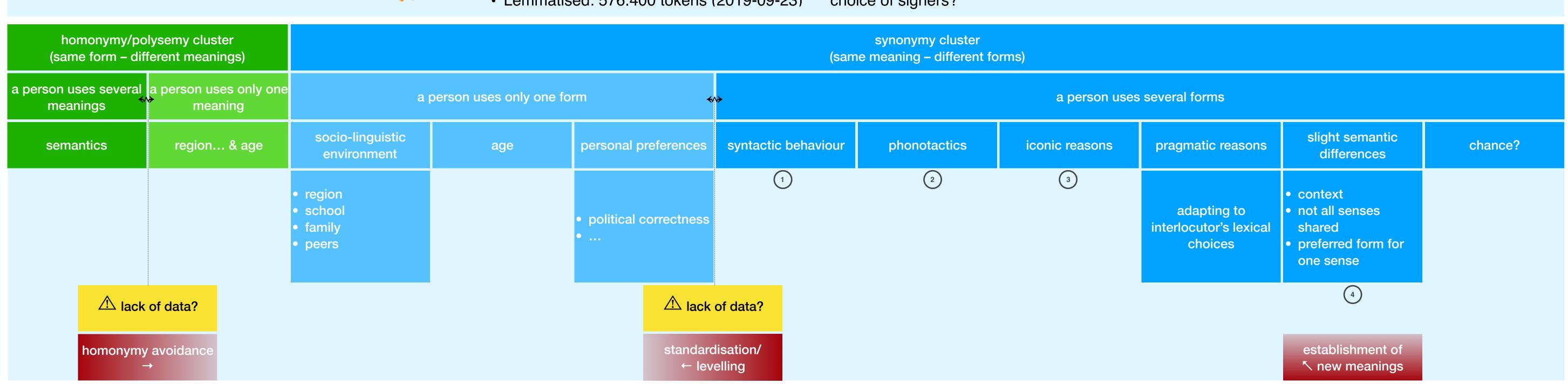
- Filmed conversations and staged
- communicative events (Nishio et al. 2010) Multi-modal corpus, lemmatised and
- accessible through iLex (Hanke/ Storz 2008)
- About 560 h footage of natural signing
- Lemmatised: 576.400 tokens (2019-09-23)

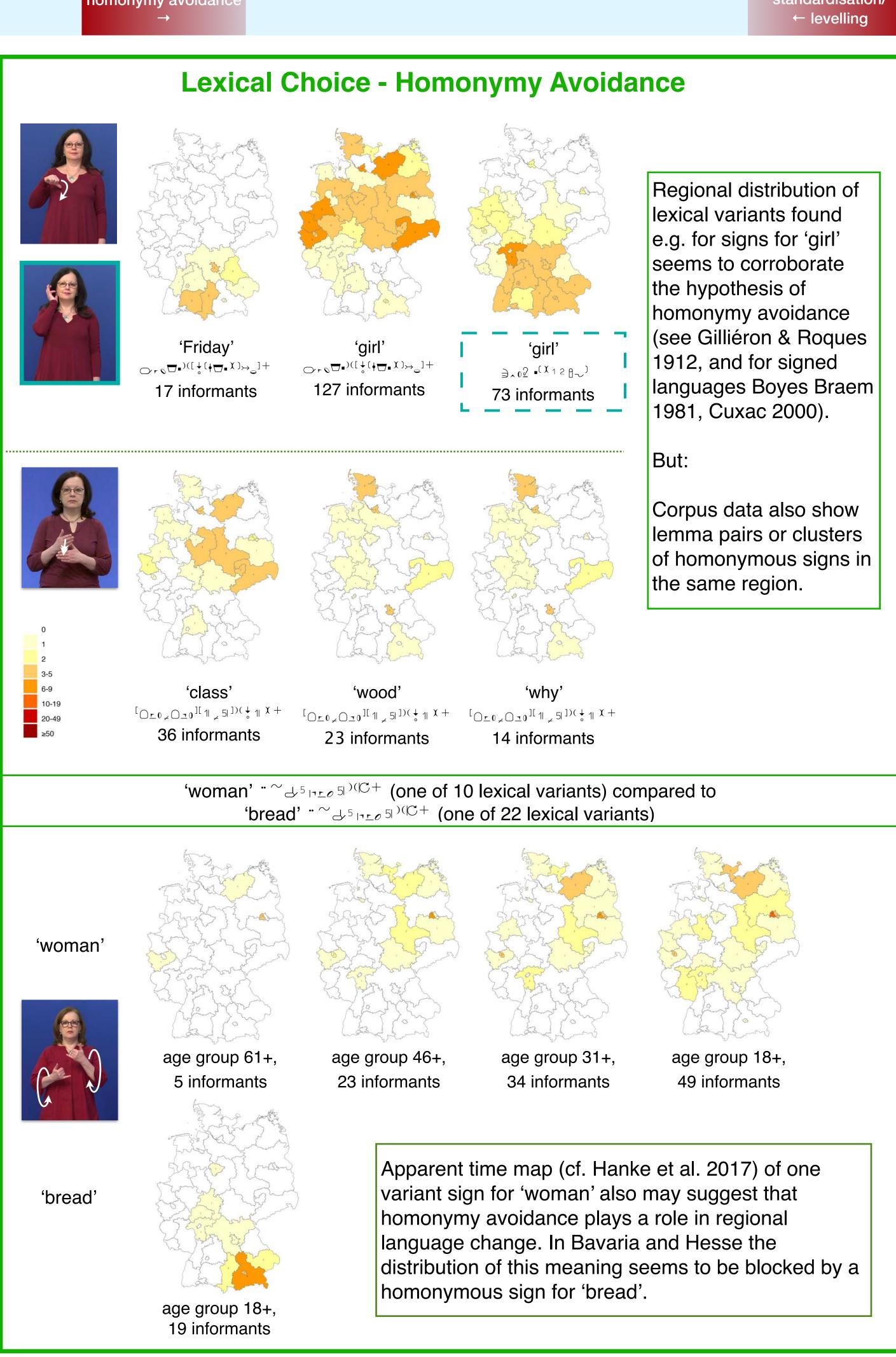
Starting Point

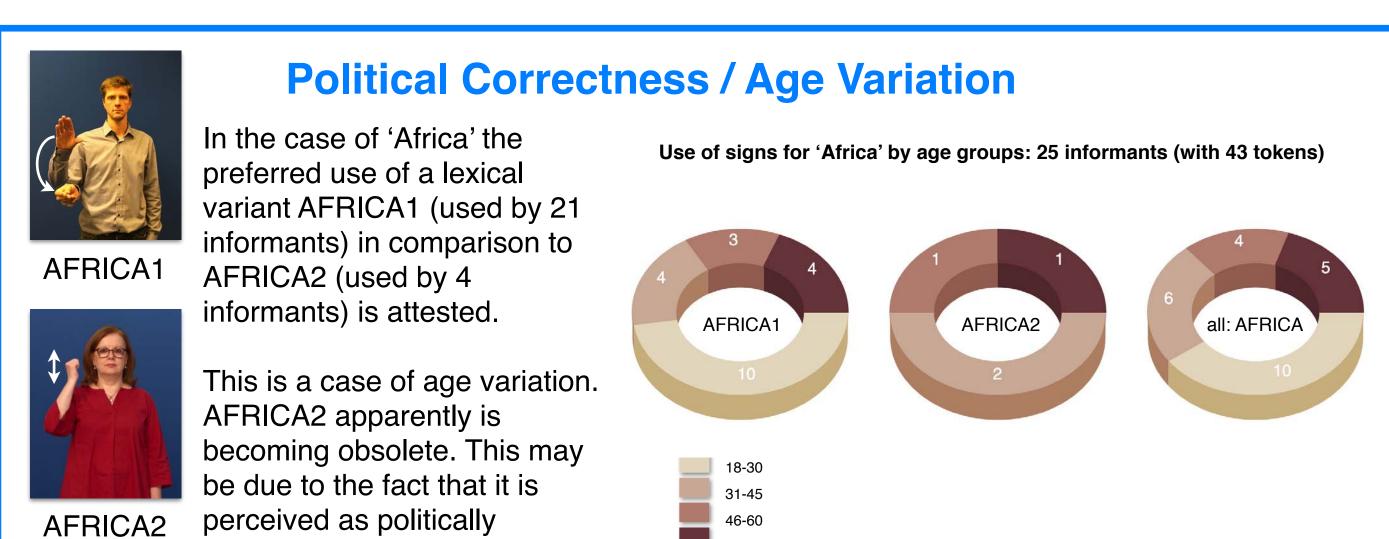
- The size of our corpus supports analyses of regional variation. Regional distribution of lexical variants of roughly synonymic sign clusters can easily be visualised on maps (cf. Hanke et. al. 2017).
- However: Often several competing signs of a sign cluster are used within the same region and even by a single individual.

Question

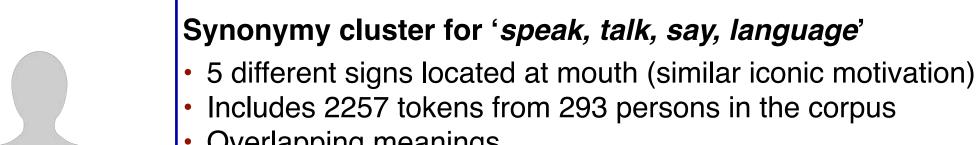
• Looking beyond regional and sociolinguistic background: What other factors influence the lexical choice of signers?







Lexical Choice - within Semantic Clusters



- Includes 2257 tokens from 293 persons in the corpus
- Overlapping meanings
- No clear regional distribution: several signs used in each region
- Assumption: slight meaning differences
 - investigation: closer look at use by one person "signer X" reveals use of different forms for different meaning aspects for this one person



TO-SAY1

 $\exists r 0 \cup \chi[P \rightarrow \underline{r} 0]$

Signer X from the

Hamburg area,

age group: 60+

35 tokens of 4 forms



TO-SPEAK1

-9r\^00)(>-0+



LANGUAGE1

9/9·00/0)(C+



LANGUAGE2

corpus

191

sprechen,

sprache

₩₩⟨\^≎○•)([**1**₩]+



	Corpus	Signer A	corpus	Signer A	corpus	
okens	1529	17	372	5	296	
nouthings nclude orms of	sagen	sagen	sprechen, sagen	sprechen	sprache, sprechen, sagen	3
redominant neaning as ised by igner X	independen status or lar used: introducing a of an utteral or opinion o	nguage the content	hearing pers speaking in language to person(s), s especially a movements	a spoken Deaf ome using rticulated lip	reference to (spoken) lar Deaf person spoken lang focus on all speak	16 1 :

sprechen speaking by gestures,

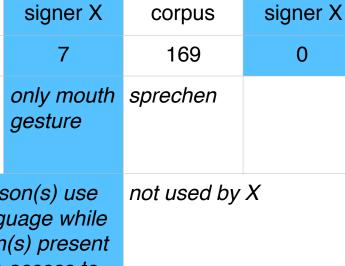
focus on manner of

of loan compounds)

hearing person(s) use spoken language while Deaf person(s) present do not have access to the content (usually in group situations such as school or in mixed groups); focus on (also used as element

inaccessibility of

content



LANGUAGE3

₩,0 ∪ X[N→|¬]+



focus on content





(visible) articulation





	TOGETHER1A + 2 variants	TOGETHER3A	TOGETHER6	TOGETHER- PERSON1
token count	552	230	39	126
semantic difference	together: in a group	together: two persons	together: two persons/ two parties (abstract)	together: two persons
polysemy	'group' (328 corpus tokens), 'community' (43 corpus tokens)	'with' (1075 corpus tokens)	'with' (31 corpus tokens), sign becoming obsolete	morphologically related signs: TO-ACCOMPANY1A, TO-SEPARATE4B
syntactic behaviour	spatial modification (1 locus)	deictic use, spatial modification (2 loci)	spatial modification (1 locus)	spatial modification (1 or 2 loci)
iconicity	depicting handshape: size & shape	no depicting	depicting handshapes: '2 persons'	
		Strategy of individual info	Strategy:	

Contrastive analysis of meaning

- Individuals using several items of a cluster tend to distinguish between different senses and functions
- Strategy of individual informants using both: difference in semantic roles indicated (person vs. place)
 - style: two-handed form allows for enlarging visual context: build coherent units using same handshape (handshape harmony)
- use in context of two specific person referents, add depicting constructions ² in context ³

Conclusion:

cluster

- Looking into individual language use can help to find usage distinctions in semantic clusters. Results need to be validated against a larger number of corpus informants: Across informants, more meaning overlaps can be observed.
- Apparent time only allows a rather coarse diachronic view on the data, competing processes like establishment of new meanings and levelling would need a finer granularity on the timeline to be separated. Exact synonyms (lexical variants) are rare, if not regionally distributed.
- Homonymy avoidance cannot be claimed as a general rule, but we find data fitting the pattern.

Poster presented at Theoretical Issues in Sign Language Research. TISLR13 2019, Hamburg, Germany. September 26-28, 2019.

incorrect.

Boyes Braem, P. (1981): Features of the Handshape in American Sign Language. Ph.D. thesis, University of California, Berkeley.

Conference on Language Resources and Evaluation (LREC 2008), Marrakech, Morocco. Paris: ELRA, pp. 64-67

Cuxac, C. (2000): La langue des Signes Française (LSF). Les voies de l'iconicité. Paris: Ophrys. Gilliéron, J. / Roques, M. (1912): Études de géographie linguistique d'après l'Atlas linguistique de la France. Paris: Champion. Hanke, T. / Storz, J. (2008): iLex – A Database Tool for Integrating Sign Language Corpus Linguistics and Sign Language Lexicography. In O. Crasborn, T. Hanke, E. Efthimiou, I. Zwitserlood, & E. Thoutenhoofd (Eds.), Construction and Exploitation of Sign Language Corpora. Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages. 6th International

Hanke, T. / Konrad, R. / Langer, G. / Müller, A. / Wähl, S. (2017): Detecting Regional and Age Variation in a Growing Corpus of DGS. Poster presented at the workshop "Corpus-based" approaches to sign language linguistics: Into the second decade", Birmingham UK, 24 July, 2017. Nishio, R. / Hong, S. / König, S./ Konrad, R. / Langer, G. / Hanke, T. / Rathmann, C. (2010): Elicitation methods in the DGS (German Sign Language) Corpus Project. Poster presented at the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, following the 2010 LREC Conference in Malta, May 22-23, 2010. Workshop Proceedings. W13. 4th Workshop on Representation and Processing of Sign Languages: Corpora and Sign Language Technologies. May 22/23, 2010. Valetta – Malta. Paris: ELRA. pp. 178-185. This publication has been produced in the context of the joint research funding of the German Federal Government and Federal States in the Academies' Programme, with funding from the Federal Ministry of Education and Research and the Free and Hanseatic City of Hamburg. The Academies' Programme is coordinated by the Union of the German Academies of Sciences and Humanities.