

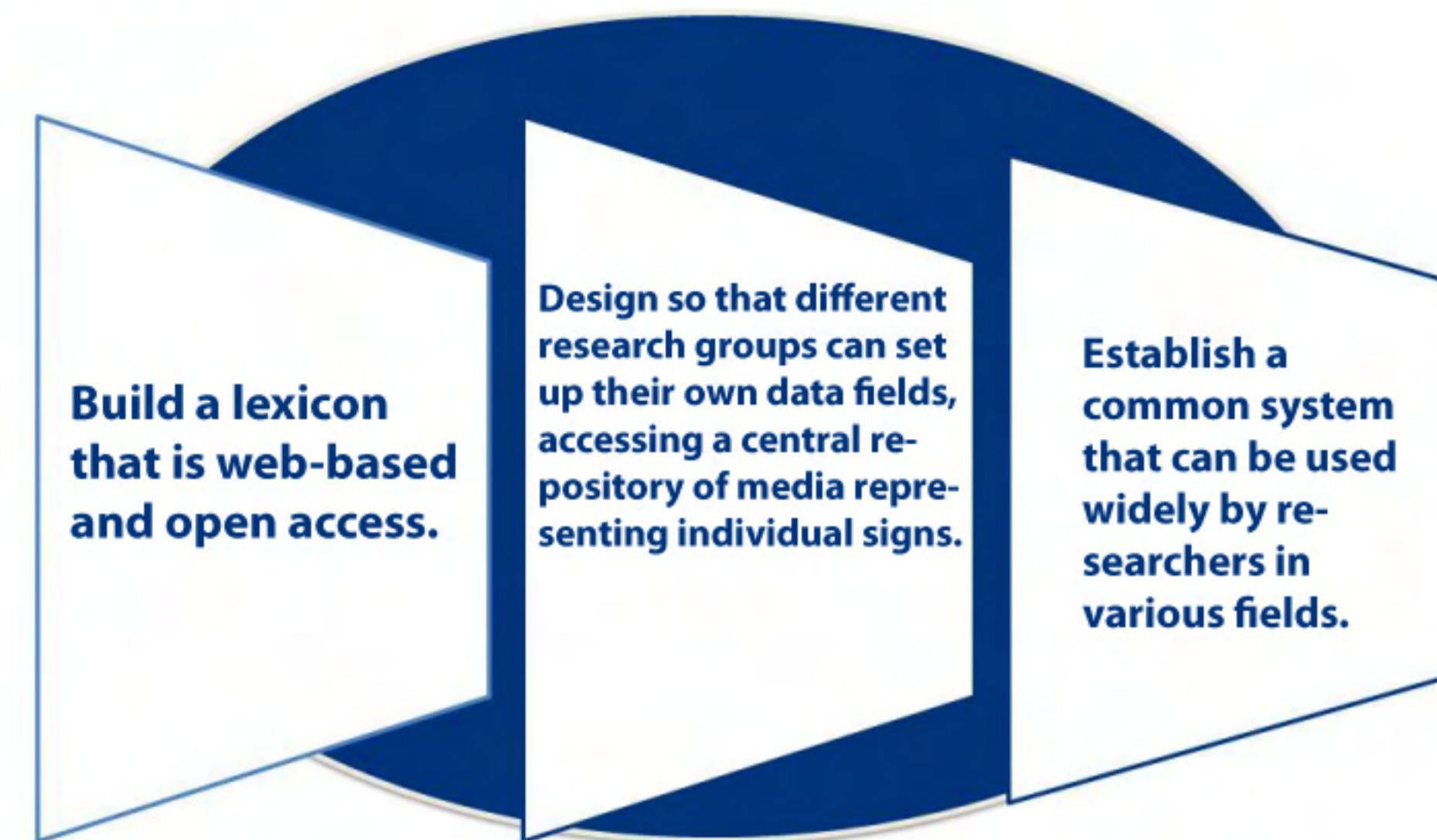
Construction of an ID Gloss Database

Gallaudet University & University of Connecticut

Karen Alkoby, Julie A. Hochgesang, Gene Mirus & Pedro Pascual-Villaneuva



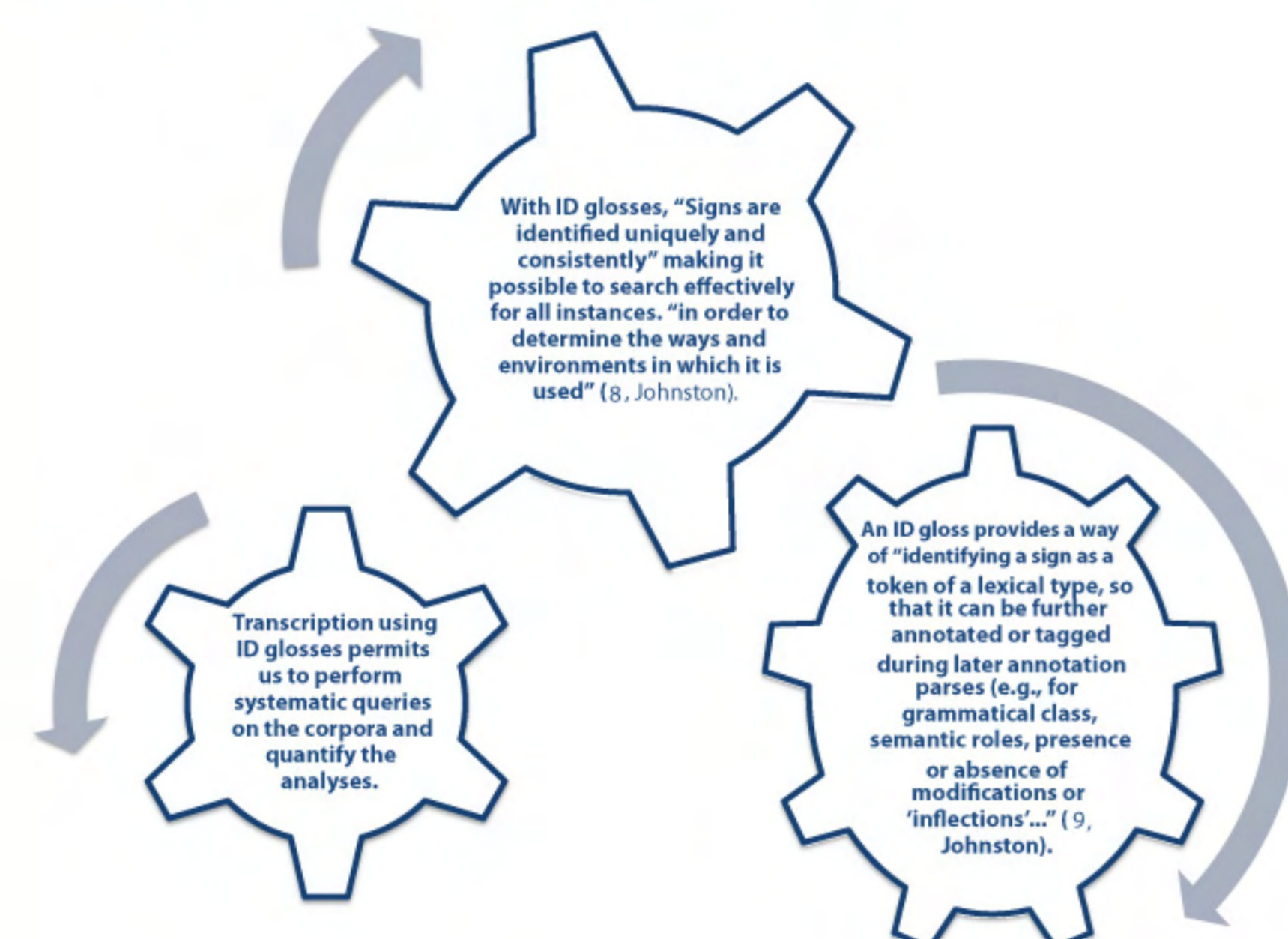
Goals



ID Gloss is...

"the common identifier for each lexical sign" (6, Johnston and Beuzeille). A sign lemma, abstracting away from actual usage such as non-target pronunciation or spatial modification

Why use ID Glosses?



Application & Benefits

- * Contains the vocabulary items used most frequently in ordinary conversation.
- * Each local group determines principles for assigning glosses to signs.
- * Database structure allows local groups to select which fields to complete.
- * Incorporates notation systems usable by various transcription tools such as ELAN, CHILDES, and etc.
- * To be maximally usable, includes at least the following two components:
 - * moving images (not only still photos or drawings),
 - * information about sign, phonological or other (e.g., handshape), by which the sign can be searched
- * Designed for use with software which allows the video to be time-aligned to each annotation. In this way, analyses can be based on the actual form of the utterance, rather than the transcript.

Template

Develop a systematic annotation system for maximum comparability across research.

Optimize the production of transcripts facilitating efficient searches and analyses by incorporating an ID gloss for each sign drawn from a lexicon having information about each sign's phonological and grammatical characteristics.

Gloss Information <ul style="list-style-type: none"> ID-Gloss Alternate glosses Semantic information English translations Example in sentences Example in pictures Example in movies Data entry specialist 	Phonological Information <u>Notation-System Free</u> Handedness Type of two-handed signs Joints movement NMS <u>Notation-System Specific</u> Stokoe Prosodic Model HamNoSys Johnson and Liddell
Morpho-Syntactic Information <ul style="list-style-type: none"> Type of sign/Part of speech Inflection Spatial information 	Sociolinguistic Information <ul style="list-style-type: none"> Usage notes Age group(s) Gender group(s) Region(s) Racial/ethnic group(s) Etymology of sign

Videoclip of a Sign

Examples

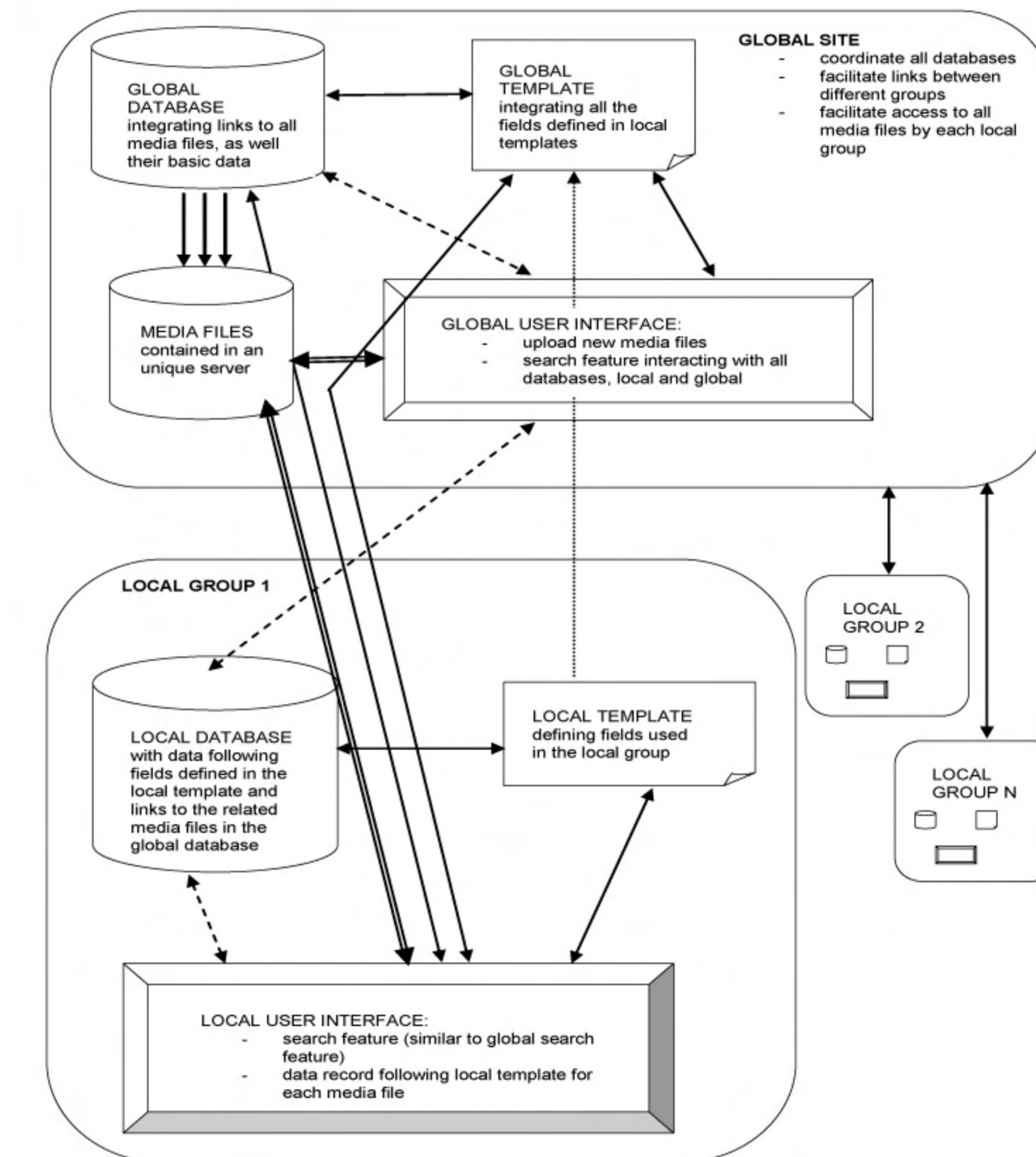
Gloss Information <ul style="list-style-type: none"> NOT-YET have not yet begun have not; haven't; not yet not yet done; begun NOT-YET FINISH fs (ID_GLOSS) 	Phonological Information <u>Notation-System Free</u> Both one-handed & two-handed Same handshape & same movement wrist flexion tongue sticking out <u>Notation-System Specific</u> Stokoe handshape: B to B'' (both) location: movement: in Johnson and Liddell timing information: D movement: none (local movement) hand configuration: UFF: 1RE= 2RE= 3RE= 4RE to UFF: 1FR= 2FR= 3FR= 4FR (both) placement: hands at M1CH
Morpho-Syntactic Information <ul style="list-style-type: none"> Adverb (none) (none) 	Sociolinguistic Information <ul style="list-style-type: none"> used by all age groups used by all gender groups used in most regions used in all racial/ethnic groups

Gloss Information <ul style="list-style-type: none"> A-LITTLE -BIT little, few 	Phonological Information <u>Notation-System Free</u> one-handed shoulder, fingers eyebrow squint <u>Notation-System Specific</u> Stokoe handshape: :T location: movement: Johnson and Liddell timing information: PTP movement: straight hand configuration: UFF: 1aCb 1 1 EF+ 2FF= 3FF= 4FF (then) Ueh = 1FF= 2FF= 3FF= 4FF placement: hand at M1CH
Morpho-Syntactic Information <ul style="list-style-type: none"> Adjective can be more tense to reflect a tinier amount can be located in space to indicate small amounts of different objects 	Sociolinguistic Information <ul style="list-style-type: none"> considered informal used by all age groups used by all gender groups used in most regions used by all racial/ethnic groups

Design

The design of the following two major parts.

- * A Global site
- * Multiple local sites



Acknowledgements

The project described in this poster was supported by Award Number R01DC009263 from the National Institute on Deafness and Other Communication Disorders. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Deafness and Other Communication Disorders or the National Institutes of Health.

This presentation is part of a larger project whose members for the past year have included Jeffrey Bernath, Paul Dudis, Diane Lillo-Martin, and Gaurav Mathur, along with the authors of this poster.

Selected Bibliography

- * Bird, S., & Simons, G. (2003). Seven dimensions of portability for language documentation and description. *Language* 79.3, 557-582.
- * Chen Pichler, D., Hochgesang, J., Lillo-Martin, D. & Quadros, R.M. (2010). Conventions for sign and speech transcription of child bimodal bilingual corpora in ELAN. *Language, Interaction and Acquisition* 1(1), 11-40.
- * Hochgesang, J., Pascual Villanueva, P., Mathur, G., & Lillo-Martin, D. (2010). Building a Database While Considering Research Ethics in Sign Language Communities. *Proceeding of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies; 7th Language Resources and Evaluation Conference.*
- * Johnston, T. (2010). From archive to corpus: transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics* 15(1), 106-131.
- * Johnston, T. (2001). The lexical database of Auslan (Australian Sign Language). *Sign Language & Linguistics*, 4(1/2), 145-169.
- * Johnston, T., & Schembri, A. (1999). On defining lexeme in a sign language. *Sign Language & Linguistics*, 2(1), 115-185. Nathan, David. 2005. *Developing Multimedia Documentation*. In P. Austin (Ed.), *Language Description and Documentation Vol 2*, 154-168. London: The Hans Rausing Endangered Languages Project.