Housekeeping

- **Workshop dinner**
  - If you have booked, see card in your name tag for details of restaurant
  - If you have booked but have changed your mind (or haven’t booked but want to go), see registration desk

---

**Digging into signs**

Developing standard annotation practices for cross-linguistic, quantitative analysis of sign language data

- Digging into Data Challenge
- UK: Funded jointly by Economic & Social Research Council (ESRC) and Arts & Humanities Research Council (AHRC), managed by JISC (£100,000)
- Netherlands: Dutch Science Foundation (NWO) (€100,000)
- June 2014 to May 2015
• Collaborations between US-Canada-UK-NL-FR
• Three rounds of funding so far
• Funds for working with big data
• Big data: ‘size forces us to look beyond the tried-and-true methods that are prevalent at a given time’ (Jacobs 2009)

• Pararchive: Open Access Community Storytelling and the Digital Archive → ‘oral history online’ beta.pararchive.com
• UK AHRC Connected Communities funding scheme, 18 months

BSL
(DCAL, UCL)
Kearny Cormier
Sandra Smith
Sannah Gulamani
With thanks to:
Jordan Fenlon

NGT
(CLS, Radboud University)
Onno Crasborn
Richard Bank
Merel Van Zuilen
Anique Schuller
Aarthy Somasundaram

Pararchive Conference, 2015

Participatory design as a driver for long-term community engagement [a theatre archive]
Memories on film: digital storytelling with people in residential dementia care
Creating and using ‘big rich’ data for voluntary sector communities
The PoetryFilm archive 2002-2015
Box stories: from a digital inclusion project to a global digital storytelling app
Digital outreach and digital preservation: case studies from parliament
Sign language narratives online: how to make them accessible
What came before

- Series of 6 LREC workshops *Representation and Processing of Sign Languages* (2004-2014)
- Stockholm, June 2010: Sign Linguistics Corpora Network workshop on ‘Standardising annotations for sign language corpora’
- Hamburg summer school on corpus sign linguistics, 2012
- Many smaller events
no annotations?

no corpus!
Why do we need sign language annotation standards?

- Prevent reinventing the wheel in every project: standards will be based on experiences from multiple researchers and research areas
- Annotation standards will support consistency within corpus projects
- Many corpus projects underway; a window of opportunity for the field
- Ensure possibility of data exchange between current and future projects

ThisThat workshop

- Learn about what others do
- Exchange practices from various research groups
- Get a first impression of the amount of overlap

And then

- Go home and use what you have learned
- Document and publish our annotation guidelines and templates
- Think about becoming a member of the ISOcat Thematic Domain group

Quantitative analysis of (sign) language data relies on corpora

What we mean by “corpus”

- Modern linguistic corpus
  - Large collection of spoken, written or signed language data, with associated metadata
  - Maximally representative (as far as is possible) of the language and its users
  - Machine-readable form
- Can be consulted to study the type and frequency of constructions in a language
  - E.g. British National Corpus of English
- Not just any language dataset
Why corpora for sign languages?

- Sign language research since the 1960s
- Traditional reliance on limited datasets, small number of signers
- Sociolinguistic situation
  - High degree of variation
  - Interrupted transmission between generations
  - Only ~5% native signers
- Native signer intuitions are problematic
- Small, limited datasets are problematic

Why no sign language corpus until now?

- Difficulties in annotating a sign language corpus
  - No standard, widely-used writing system for any sign language
  - Some notation systems that represent sign language data at phonetic/phonological level (though nothing standard like IPA)
- Recent technological advancements have made sign language corpora possible
  - Time aligned video annotation software (e.g., ELAN)
  - Improved capacity for the capture and storage of digital video

(Johnston, 2009)

19 sign languages represented here

Researchers working on these sign languages are at various stages of corpus creation.

Some more advanced than others.

However: Even the most advanced of these datasets are not true corpora. Not yet.
Modern linguistic corpus

✔ Large collection of spoken, written or signed language data, with associated metadata
✔ Maximally representative (as far as is possible) of the language and its users
✖ Machine-readable form

Machine-readability requires annotation.

Problems with achieving machine readability

• Inconsistencies when signs are annotated via spoken/written language
• Many parts of signed discourse are not composed of fully lexical signs (equivalent of words)
• Some standards are beginning to emerge (e.g., Johnston 2014)
• But no attempts to standardise annotation practices across sign language corpora
• This project is the first, using BSL and NGT Corpora

Project aims

• To create clear standards for addressing problems with sign language annotation
• Using BSL Corpus and Corpus NGT we aim to:
  – Develop annotation standards
  – Test their validity & reliability
  – Improve current software tools that facilitate a reliable workflow (ELAN, Signbank)
  – Publish two data sets (BSL Corpus & Corpus NGT in a shared location (The Language Archive @MPI)

Validity & reliability

• Validity: does the annotation convention make sense (w.r.t. a specific theory)?
• Reliability: will annotators use the convention consistently? (across files, across annotators)
• Small reliability study planned for the coming month
• Quantitative & qualitative analysis, aimed at improving the guidelines
Software development

- Targeted at displaying lexical information in ELAN

Publishing data sets

The Language Archive: corpus1.mpi.nl
**Impact**

Open-access annotated corpora & lexicons

Expands capacity for quantitative analysis of sign languages by linguists

Allow quantitative comparisons across signed and spoken languages

**Impact**

Applications for language technology and language teaching/learning

- Machine translation, signing avatars, automatic sign language recognition: *Cutting edge language technology*
- *Increased usefulness of corpora for:
  Sign language teaching and learning materials, interpreter training*

---

**Digging into Signs: BSL and NGT Corpora**

Kearsy Cormier & Onno Crasborn

- Minimum requirements for a machine-readable sign language corpus (Johnston 2010):
  - Translation
  - Type-token matching at lexical level
    - For lexical signs
      - ID glossing which requires lexical database
ID glossing

- Lexical signs
  - Two variants: same or different?
  - How to choose the ID gloss given possible translation equivalents?
  - These issues relate to lexical database building and lemmatisation
  - To be discussed at another time and place

Other things we’re not talking about here

- Annotation standards for other types of constructions above and below the lexical level
  - Phonological
  - Morphological
  - Syntactic
- See Johnston (2014, Auslan guidelines)
- Shared guidelines for translation

Minimum requirements for a machine-readable sign language corpus (Johnston 2010):

- Translation
- Type-token matching at lexical level
  - For lexical signs
    - ID glossing which requires lexical database
  - For other types of manual activity aside from lexical signs (e.g. partly lexical signs – classifier/depicting signs, pointing signs, etc)
    - Annotation practices vary. No clear standards emerging yet.
Aims

- Aim of Digging into Signs project
  - To develop standards for annotation, particularly with partly-lexical signs given that these vary so much
- Aim of Digging into Signs workshop
  - To compare and contrast annotation practices across sign language corpus projects
  - To get feedback on our draft Digging annotation standards

What should we be aiming for?

- Translation and ID glossing are necessary minimally for a machine-readable sign language corpus. This is clear (Johnston 2010, JCL)
- Beyond this, what is the minimal annotation needed for other types and subtypes of signs (e.g. partly-lexical signs)?
  - Classifier/depicting signs
  - Pointing signs
  - Buoys
  - Name signs
  - Fingerspelling
  - Gestures
  - Etc

Annotation guidelines: motivations

- BSL annotation guidelines roughly follow those for Auslan (Johnston, various iterations since 2010)
  - Cormier, Fenlon, Gulamani & Smith. 2015. BSL Corpus Annotation Conventions v2.0. (Available on Digging into Signs Workshop website)
- Motivation: identification of lexemes and other symbolic units (e.g. partly lexical signs)
  - Type-component – general category (e.g. classifier/depicting sign, pointing sign etc)
  - Type-like components for depicting and pointing signs (subtypes)
  - Token-like components for non-lexical material (gestures and constructed action)
- Prioritise annotation over transcription
  - Relationships between form and meaning: Type/token relationships

(Johnston 2014, Reluctant Oracle, Corpora)

Annotation guidelines: motivations

- NGT annotation guidelines developed from the ECHO ‘corpus’ guidelines, moved towards the Auslan guidelines
  - Crasborn, Bank, Zwervelood, van der Kooij, Meijer & Safar. 2015. Annotation Guidelines for Corpus NGT. (Available on Digging into Signs Workshop website)
- Motivation
  - ID-glossing of lexicalised items
  - No pre-existing lemmatised lexicon: created as we go
  - Sometimes: focus on form first, function later; form we know (more) about (phonology), function (often) requires (more) analysis
- Gloss
  - Meaning
  - Reference
- Mouth
  - MouthType
- Translation
  - (of everything; at sentence level)
Joint annotation guidelines

- Focus on Gloss tiers (i.e. primary manual activity)

Where is the standard?

- Agreement of classification between BSL–NGT, no pre-processing needed (requires common language of annotation)
- Agreement, but some pre-processing of data sets needed in order to perform easy queries across data sets. Some detail may get lost.
  - Houston, we have a problem.

Sign types compared

- Basic gloss
  - All lexical signs are annotated using an identifying gloss (ID-gloss), written in upper case. This gloss corresponds to 'Annotation ID-gloss' in SignBank

- Variants
  - BSL: Lexical variants (same/similar meaning, unrelated phonologically) use number suffixes
  - NGT: Dash followed by letter; lexical and phonological variants are not distinguished.

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIGN</td>
<td>SIGN-A</td>
</tr>
<tr>
<td>SIGN02</td>
<td>SIGN-B</td>
</tr>
<tr>
<td>SIGN03</td>
<td>SIGN-C</td>
</tr>
</tbody>
</table>

Sign types compared

- Two-handed signs
  - Two handed signs annotated on both right and left hand tiers
### Sign types compared

- **Buoys**
  - Buoys get their own ID gloss

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>List, Pointer, Fragment and Theme buoys. Start and end of buoy follows start and end of co-occurring sign on dominant hand.</td>
<td>Only List buoys (others considered perseveration)</td>
</tr>
<tr>
<td>LBUOY</td>
<td>COUNTING-HAND-1</td>
</tr>
<tr>
<td>PBUOY</td>
<td>COUNTING-HAND-2</td>
</tr>
<tr>
<td>FBOUY</td>
<td></td>
</tr>
<tr>
<td>TBUOY</td>
<td></td>
</tr>
</tbody>
</table>

- **Repetition**
  - Repeated lexical signs are annotated separately

<table>
<thead>
<tr>
<th>BSL/NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOY SHOUT WOLF WOLF</td>
</tr>
</tbody>
</table>

- **Compounds**
  - Lexicalised compounds in SignBank

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARENTS</td>
<td>PARENTS</td>
</tr>
</tbody>
</table>

- **Possible compounds**
  - With ^ for BSL. Not annotated for NGT

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRAPHIC^ART</td>
<td>GRAPHIC ART</td>
</tr>
</tbody>
</table>

- **Manual negative incorporation**
  - ID-gloss with negative suffix

<table>
<thead>
<tr>
<th>BSL/NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNOW-NOT</td>
</tr>
</tbody>
</table>
Sign types compared

- **Directional verbs**
  - ID glossed
  - BSL: Grammatical/modification info on separate project-specific tiers
  - NGT: Prefixed or suffixed with ‘1’ if 1SG location is used

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASK</td>
<td>ASK:1</td>
</tr>
<tr>
<td>TAKE-OVER</td>
<td>1:TAKE-OVER</td>
</tr>
</tbody>
</table>

Sign types compared

- **Plurality**
  - BSL: Grammatical/modification info on separate project-specific tiers (except for pointing signs)
  - NGT: Suffix with .PL

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHILD</td>
<td>CHILD.PL</td>
</tr>
</tbody>
</table>

Sign types compared

- **Numbers**
  - ID glossed

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>1-A</td>
</tr>
<tr>
<td>ONE02</td>
<td>1-B</td>
</tr>
</tbody>
</table>

- **Number sequences**
  - ID glossed

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>NINETEEN^EIGHT^NINE02</td>
<td>1989</td>
</tr>
</tbody>
</table>

Sign types compared

- **Numbers incorporation**
  - ID glossed with suffix for information about incorporated number

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOUR-FOUR02</td>
<td>HOUR-4</td>
</tr>
</tbody>
</table>

- **Ordinal numbers**
  - BSL: ID glossed as lexical signs
  - NGT: Numeral with -ORD suffix

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRST</td>
<td>1-ORD</td>
</tr>
<tr>
<td>RANKING</td>
<td>RANKING02</td>
</tr>
<tr>
<td>RANKING02-THREE</td>
<td></td>
</tr>
</tbody>
</table>
Sign types compared

• Sign names

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SN:FIRST</td>
<td>*FIRST-LAST</td>
</tr>
<tr>
<td>SN:PETER(FS:P-PETER)</td>
<td>*PETER-SMITH</td>
</tr>
<tr>
<td>SN:OSAMA-BIN-LADEN(BEARD)</td>
<td>*OSAMA-BIN-LADEN</td>
</tr>
</tbody>
</table>

• Fingerspelling

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS:WORD</td>
<td>#WRD</td>
</tr>
<tr>
<td>Meaning tier: word</td>
<td></td>
</tr>
</tbody>
</table>

• Pointing signs

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT:LOC</td>
<td>PT</td>
</tr>
<tr>
<td>PT:DET</td>
<td>PT</td>
</tr>
<tr>
<td>PT:PRO1SG</td>
<td>PT:1</td>
</tr>
<tr>
<td>PT:LOC/PT:PRO3</td>
<td>PT</td>
</tr>
<tr>
<td>PT:LOC/PT:BPUY</td>
<td>PT:W</td>
</tr>
</tbody>
</table>

• Classifier/depicting signs

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSEW(2)-MOVE</td>
<td>MOVE+2</td>
</tr>
<tr>
<td>Meaning tier: cat walks to and fro</td>
<td></td>
</tr>
<tr>
<td>DSEP(1)-PIVOT</td>
<td>PIVOT+1</td>
</tr>
<tr>
<td>Meaning tier: person turns around</td>
<td></td>
</tr>
<tr>
<td>DSEW(2)-AT</td>
<td>AT+2</td>
</tr>
<tr>
<td>Meaning tier: bird is here</td>
<td></td>
</tr>
</tbody>
</table>

• Type-like classifier/depicting signs

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSEW(1-VERT)-MOVE:HUMAN</td>
<td>MOVE+1</td>
</tr>
<tr>
<td>DSEW(FLAT-LATERAL)-AT:VEHICLE</td>
<td>AT+flat</td>
</tr>
</tbody>
</table>

• Shape constructions

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSS(CYL)</td>
<td>SHAPE+cylinder</td>
</tr>
<tr>
<td>Meaning tier: drain pipe</td>
<td></td>
</tr>
</tbody>
</table>
Sign types compared

- **Gestures**

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>G:HOW-STUPID-OF-ME</td>
<td>% Meaning tier: how stupid of me</td>
</tr>
<tr>
<td>HOW-STUPID-OF-ME</td>
<td></td>
</tr>
</tbody>
</table>

- **Manual constructed action**

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>G:CA:HOOD-HANDS-UP-IN-FRIGHT</td>
<td>% Meaning tier: frightened</td>
</tr>
</tbody>
</table>

Some other conventions

- **Uncertainties**
  - False starts
  - Doubt as to whether this is a sign (vs. ‘self-touching’ e.g.)
  - Invisible (out of video frame, behind other hand)

- **‘Workflow issues’**
  - new sign yet to be added to lexical database (ELAN)
  - doubt about whether the gloss is chosen correctly
  - suggestions for an alternative gloss

Where is the standard?
Where is the standard?

Goal of the workshop for us
• Can we be greener?
• Or are there good reasons for doing things differently right now?
• More agreement may be achieved as we come to learn more about the languages

Finally: Other conventions that will impact cross-corpus research

• How (approximate or fine-grained) is the time-alignment of the glosses?
• File naming conventions (not for iLex?)
  – annotation files
  – video files
  – metadata files

Workshop presentations

• Paper and poster presenters have been asked to prepare a presentation that compares their annotation practices with those of the NGT and BSL Corpora and to provide feedback on draft annotation standards
• Paper presentation format
  – 30 minutes: main presentation
  – 10 minutes: Q&A
  – 20 minutes: discussion of key issues

Extra slide
What should we be aiming for?

• Translation and ID glossing are necessary minimally for a machine-readable sign language corpus. This is clear (Johnston 2010, IJCL)

• Beyond this, what is the minimal annotation needed for other types and subtypes of signs (e.g. partly-lexical signs)?
  – Classifier/depicting signs
  – Pointing signs
  – Buoy signs
  – Name signs
  – Fingerspelling
  – Gestures
  – Etc

• Is it enough to (minimally) identify categories at this level (i.e. ‘type component’, Johnston 2014)? What about subtypes of each?

• Also, issues with identification of each of these as distinct from each other and from other constructions (e.g. some types of buoys versus perseveration)

What’s on the gloss tier?

• Glosses should be a basic level of information, pointers to your lexicon – no details

• But: what you (can) put on the gloss tier, is related to the structure and progress of your lexicon
  – BSL: lemmas
  – NGT: phonological variants; lemmatisation yet to be done, but that level may well remain in the lexicon (as we need the ‘phonological subtype’ information for research)

What’s on the gloss tier?

• Glosses should be a basic level of information, pointers to your lexicon – no details

• But: what you (can) put on the gloss tier, is related to the structure and progress of your lexicon
  – DGS & Polish SL: two simultaneous levels of glossing
  – Polish SL: lemma and phonological variant glosses
  – DGS: lemmas and ‘mother lemmas’ (containing the shared iconic image)
Minimal needs beyond lexical signs

- Identification of these kinds of categories (whether by prefix, suffix, abbreviation, symbol, whatever):
  - Pointing signs
  - Classifier/depicting signs
  - Fingerspelling
  - Gestures
  - Sign names
  - Buoys
  - Etc

Where is the standard?

Goal of the workshop for us
- Can we be greener?
- Or are there good reasons for doing things differently right now?
- More agreement may be achieved as we come to learn more about the languages

Where is the standard?

Goal of the workshop for us
- Can we be greener?
- Or are there good reasons for doing things differently right now?
- More agreement may be achieved as we come to learn more about the languages

(1. Actually, black would be good enough)

2. Sometimes, yes: the extent to which you are interested in the form will determine how much elements of the form (phonetic transcription) you allow to sneak into the gloss tiers

(1. Actually, black would be good enough)

For example: morphosyntactic analysis of pointing

For example: handedness of signers
Where is the standard?

Goal of the workshop for us

1. Can we be greener?
2. Or are there good reasons for doing things differently right now?
3. More agreement may be achieved as we come to learn more about the languages

Where is the standard?

1. ID glossing via lexical database is important for a machine-readable corpus
2. Contextual glossing is not enough to achieve this
   - ...and one can wonder whether it is efficient as a first step
3. Workshop/training may be needed
   - In the meantime:

Sign types compared

- Variants
  - BSL: Lexical variants (same/similar meaning, unrelated phonologically) use number suffixes
  - NGT: Dash followed by letter; lexical and phonological variants are not distinguished.

BSL          NGT
SIGN         SIGN-A
SIGN02       SIGN-B
SIGN03       SIGN-C

To be addressed as lemmatisation in NGT Signbank progresses

Sign types compared

- Two-handed signs
  - Two handed signs annotated on both right and left hand tiers.

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start and end follows dominant hand.</td>
<td>Start and end of two-handed signs is determined independently for each hand</td>
</tr>
</tbody>
</table>
| Perseveration not annotated except if intentionally meaningful (see buoys) | Will not change: one team more interested in phonetics than the other.
### Sign types compared

#### Buoys
- Buoys get their own ID gloss
  - BSL: List, Pointer, Fragment and Theme buoys. Start and end of buoy follows start and end of co-occurring sign on dominant hand.
  - NGT: Only List buoys (others considered perseveration).

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBUOY</td>
<td>COUNTING-HAND-1</td>
</tr>
<tr>
<td>PBUOY</td>
<td>COUNTING-HAND-2</td>
</tr>
<tr>
<td>FBOUY</td>
<td></td>
</tr>
<tr>
<td>TBUOY</td>
<td></td>
</tr>
</tbody>
</table>

#### Directional verbs
- ID glossed
- BSL: Grammatical/modification info on separate project-specific tiers
- NGT: Prefixed or suffixed with ‘1’ if 1SG location is used

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASK</td>
<td>ASK:1</td>
</tr>
<tr>
<td>TAKE-OVER</td>
<td>1:TAKE-OVER</td>
</tr>
</tbody>
</table>

Reduce the amount of morphological information in the gloss tier → move to child tiers

#### Numbers incorporation
- ID glossed with suffix for information about incorporated number

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOUR-FOUR02</td>
<td>HOUR-4</td>
</tr>
</tbody>
</table>

#### Ordinal numbers
- BSL: ID glossed as lexical signs
- NGT: Numeral with -ORD suffix

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRST</td>
<td>1-ORD</td>
</tr>
<tr>
<td>RANKING</td>
<td></td>
</tr>
<tr>
<td>RANKING02</td>
<td></td>
</tr>
<tr>
<td>RANKING02-THREE</td>
<td></td>
</tr>
</tbody>
</table>

### Sign types compared

#### Sign names
- BSL: ID glossed
- NGT: *FIRST-LAST

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SN:FIRST</td>
<td>*FIRST-LAST</td>
</tr>
<tr>
<td>SN:PETER(F3,P-PETER)</td>
<td>*PETER-SMITH</td>
</tr>
<tr>
<td>SN:OSAMA-BIN-LADEN(BEARD)</td>
<td>*OSAMA-BIN-LADEN</td>
</tr>
</tbody>
</table>
Sign types compared

- **Pointing signs**
  - BSL
    - PT:LOC
    - PT:DET
    - PT:PRO1SG
    - PT:PRO3
    - PT:LOC
    - PT:B:upward
  - NGT
    - PT

Should become more uniform after an ongoing research collaboration with Johnston/Schembri. Again: morphological modification to child tiers?

- **Classifier/depicting signs**

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSEW(2)-MOVE</td>
<td>MOVE+2</td>
</tr>
<tr>
<td>DSEP(1)-PIVOT</td>
<td>PIVOT+1</td>
</tr>
<tr>
<td>DSEW(2)-AT</td>
<td>AT+2</td>
</tr>
</tbody>
</table>

- **Type-like classifier/depicting signs**

<table>
<thead>
<tr>
<th>BSL</th>
<th>NGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSEW(1-VERT)-MOVE:HUMAN</td>
<td>MOVE+1</td>
</tr>
<tr>
<td>DSEW(FLAT-LATERAL)-AT:VEHICLE</td>
<td>AT+flat</td>
</tr>
</tbody>
</table>

Where is the standard?

- **Outlined in our annotation guidelines**
- **Reflected in our published data sets**
  - especially the annotation data (tokens)
  - but also the Signbanks (types)
  - and the related open source software
    - ELAN: tla.mpi.nl/tools
    - Signbank: see links on Digging into Signs website
Where will ‘the Digging standard’ be?

- 2 more months of work ahead
  - PT
  - annotations: revising current annotations to updated guidelines
  - reliability study
  - software development
- Output
  - Updates to the BSL and NGT guidelines
  - Document outlining the differences between BSL-NGT, incl. some discussions we had here
  - Software update to ELAN
  - Updated publication of annotation files for both corpora
  - Continuous work on the two Signbanks (both the systems and the content)
  - Website with links to all of your online guidelines

Thank you!

- All presenters
- Interpreter teams
  - International ASL
  - BSL
  - VGT
  - LSFB
  - SSL
- Volunteers & others from DCAL
- Our discussant Jordan
- Chairs Richard & Trevor

Thank you.2PL!

- All presenters
- Interpreter teams
  - International ASL
  - BSL
  - VGT
  - LSFB
  - SSL
- Volunteers
- Our discussant Jordan
- Chairs Richard & Trevor