

Challenges in the Development of the American Sign Language Lexicon Video Dataset (ASLLVD)

Carol Neidle, Ashwin Thangali, and Stan Sclaroff
Boston University



Research partially funded by grants from
the National Science Foundation*

Structure of this presentation

- I. The dataset and tools for annotation and verification
- II. A computer science application

Motivation: Sign Look-up

◆ Problem

How to look up a sign in a (multimedia) ASL dictionary

◆ Proposed solution

Computer-based recognition of a sign produced by a user in front of a Web cam, or identified by a user from a video (by selecting the start and end frames)



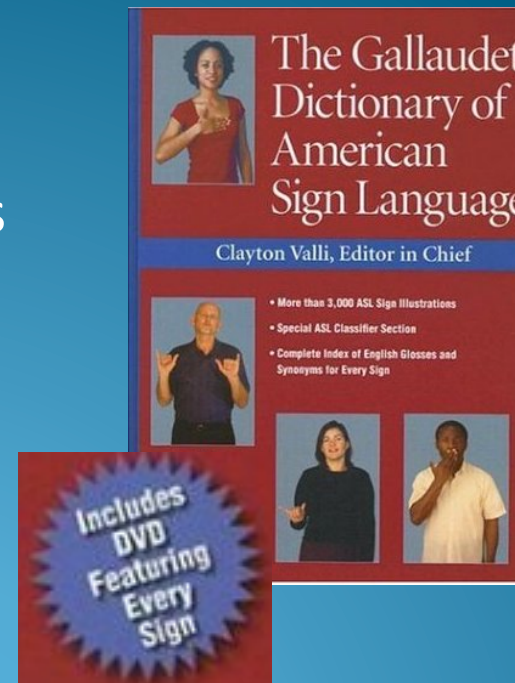
The American Sign Language Lexicon Video Dataset (ASLLVD)

❖ Data collection for this research

Collected over 3,000 signs from up to 6 native signers.

❖ Elicitation

Using as prompts video examples from the Gallaudet dictionary:







Video Characteristics

Four simultaneous views



2 frontal views



side view

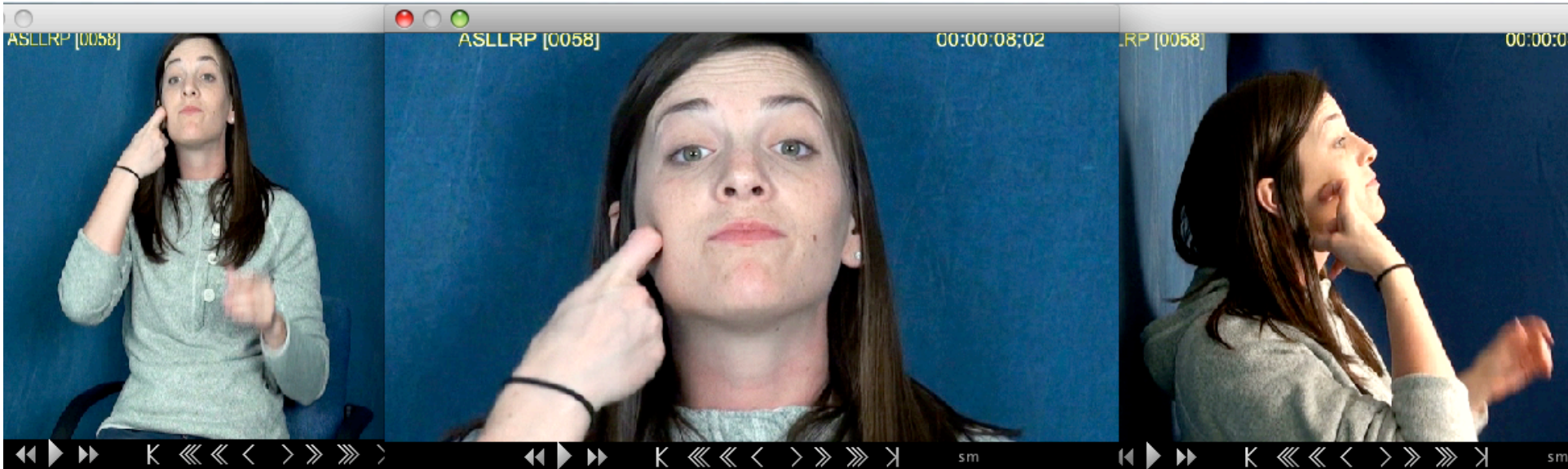


face view

1 half-speed high resolution
1 full resolution

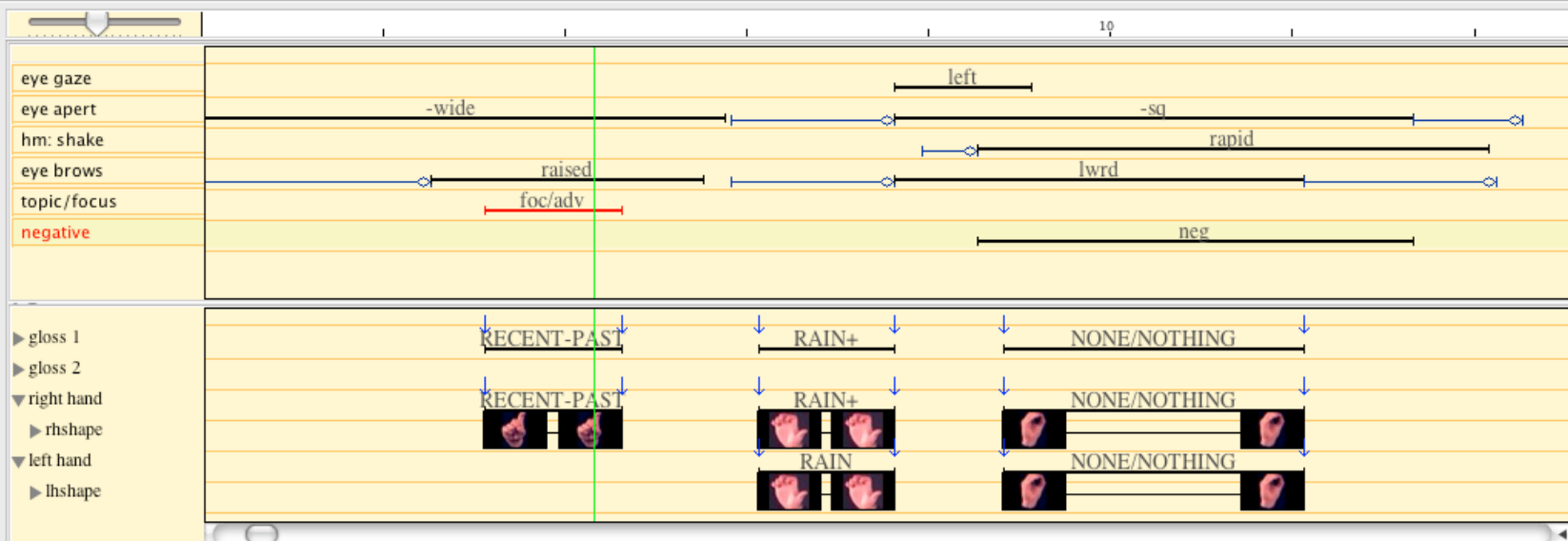
Annotations

- ◆ Conducted first using SignStream[®] version 3 (beta)
 - ◎ Gloss labels of the kind that we have used for all our SignStream[®] annotations since the mid 1990's, consistent with our annotation conventions
(ensuring 1-1 relationship between label and sign production)
 - ◎ Start and end points, onsets and offsets (where relevant) of events, start and end handshapes, and morpho-phonological properties of the sign
(new features in SignStream[®] 3)

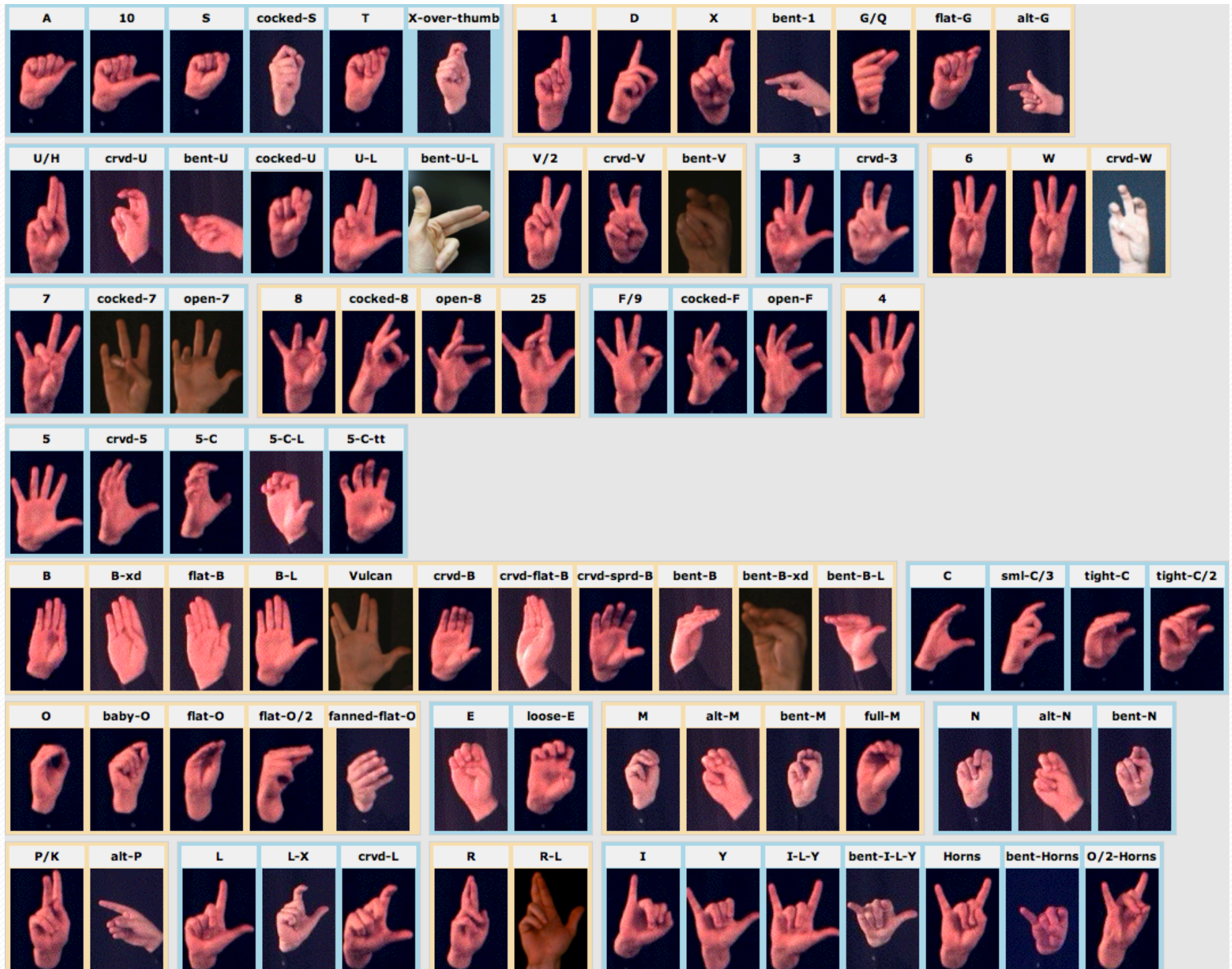


SignStream[®] version 3 beta

hel_2012-02-09_sc58
1 Rachel_2012-02-09_sc58



Temporal Partition 1 Segment 1 U2 Annotator: Indya Participant: Rachel Fields Notes Hide Field Hide Pane



Entry of morpho-phonological information about a sign

The image shows two overlapping software windows used for entering morpho-phonological information about a sign.

Associated Fields Window:

- Gloss:** NONE/NOTHING
- Sign Type:** Fingerspelled, Loan sign, Name sign, Lexical sign, Number, Classifier, Gesture
- Dominant Hand:** L, R
- Hand Shapes:** Location, Orientation, Movement (selected)
- Configuration:** One-handed, Two-handed. Sub-options: L, R. Dropdowns: Same handshapes L & R, Same START/END hs.
- Options:** Marked # of Hands, Passive base arm. Button: Select handshapes
- Selected Handshapes:** Right Hand (Clear), Left Hand (Clear). Each contains two hand images.
- Buttons:** Cancel, Enter

Palette Window:

- Palettes:** American Sign Language, All
- Right and Left Hands:** Search, Enter, Expand all triangles, Collapse all triangles
- Grid:** A grid of hand shapes with labels: A, 1, U/H, V/2, 3, 6, 7, 8, F/9, 4, 5, B, C, ?, Other, O, E, M, N, P/K, L, R, I, Rlxd.
- Buttons:** Clear, Cancel, Enter

Annotation issues

◆ Ensuring consistency

- ⊙ Labeling of handshapes

- ⊙ Glossing

1-1 *consistent* relationship
between production and label

◆ Dealing with lexical variants

Initial groupings based on stimuli from the elicitation.
However, for any given stimulus, there was frequently more than one sign and/or more than one sign variant.

Lexicon Viewer and Verification Tool (LVVT)

Ashwin Thangali

To assist with verifications and sign classifications

Search gloss, press enter for more

brave/recover

- BOY
- BOY+DCL"round flat"
- BOY+FRIEND
- BOY_2+CORRECT+LAW
- BROTHER+IN+LAW
- BRA
- BRA_2
- BRACELET
- BRACELET_2
- WATCH
- WATCH
- BRAG
- BRAKE
- (5-C-L)BRAKE
- ◆ BRAVE/RECOVER**
- ◆ BRAVE/RECOVER
- ◆ BRAVE/RECOVER
- ◆ STRONG
- ◆ STRONG
- ◆ STRONG
- ◆ BRAWL
- ◆ BREAD
- BREAD+DCL"rounded rectangle"
- ◆ BREAK
- BREAK-DOWN
- BREAK-DOWN
- BREAK-DOWN-BUILDING
- ◆ BREASTS
- BREATHE
- ◆ BRIDGE
- ◆ (crvd-V)BRIDGE
- BRING
- BRING-1p
- CARRY
- CARRY

Search gloss, press enter for more

brave/recover

Select signs

Both-hands start ->

Query start	Query end	HS
select	select	

Select signs

Non dominant

Query start	Query end	HS
select	select	

Compound signs

- Excel spreadsheet
- Both-hands start -> end HS
- Both-hands HS variants
- Non-dominant start -> end HS
- Dominant start -> end HS
- Dominant start HS variants
- Dominant end HS variants

AND select

- All signs
- HS variations
- Morphological variations
- Lexical variations
- Star,end HS differ
- One handed signs
- Two handed signs
- Two handed

5		Lexical sign 1	Lexical 1, Variant 1	S	

BRAVE/RECOVER BRAVE/RECOVER Lexical sign 2-handed same hs 2-hands Diff S,E <input type="checkbox"/> Passive base arm Signer video Combined video Old S,E 680, 705 680 705
BRAVE/RECOVER BRAVE/RECOVER Lexical sign 2-handed same hs 2-hands Diff S,E <input type="checkbox"/> Passive base arm Signer video Combined video Old S,E 2636, 2665 2636 2665
BRAVE/RECOVER BRAVE/RECOVER Lexical sign 2-handed same hs 2-hands Diff S,E <input type="checkbox"/> Passive base arm Signer video Combined video Old S,E 2968, 3000 2968 3000
BRAVE/RECOVER BRAVE/RECOVER Lexical sign 2-handed same hs 2-hands Diff S,E <input type="checkbox"/> Passive base arm Signer video Combined video Old S,E 5099, 5123 5099 5123

Annotations

Annotations interface showing video frames, handshaped images, and a list of sign types (Lexical sign, Fingerspelled sign, Loan sign, Number sign, Classifier sign, Compound sign, 2-handed same hs, 2-handed alt. same hs, 2-handed diff. hs, 1-handed). The interface also displays 'BRAVE/RECOVER' labels and video frame controls (680, 705).

Displayed and editable:

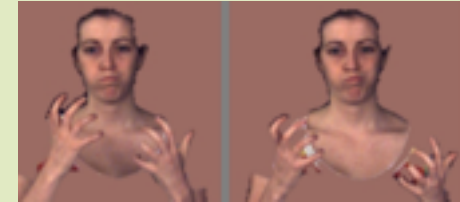
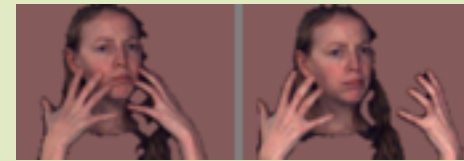
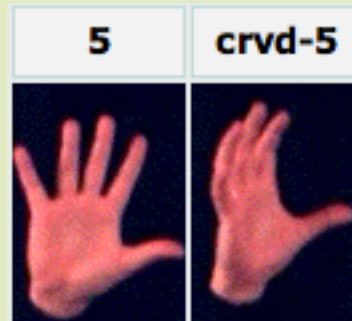
- ⦿ Gloss labels
- ⦿ Start and end frames
- ⦿ Start and end handshapes on both hands
- ⦿ Morpho-phonological properties

Also facilitates grouping signs and sign variants

How to determine what counts as same or different?

- ◆ Where to draw the line between handshapes?

(Reality is gradient.)



- ◆ Same sign/variant or two different variants of the same sign?
- ◆ Two variants: the same lexical item or two distinct lexical items?

ASLLVD

Can search based on a wide range of criteria

Consultants



- Liz
- Tyler
- Naomi
- Brady
- Lana
- Dana

Signs last annotated by:

After mm-dd-yyyy

Before mm-dd-yyyy

- BetsyBeckert
- RebeccaLopez
- JoanNash
- CarolNeidle

Select signs		Select signs	
Dominant sta <input type="text"/>		Non dominant <input type="text"/>	
Query start HS	Query end HS	Query start HS	Query end HS
B-L	bent-B-L	select	select
		<input type="text"/>	<input type="text"/>

AND select

- All signs
- HS variations
- Morphological variations
- Lexical variations
- Star,end HS differ
- One handed signs
- Two handed signs
- Two handed non-alternating same
- Two handed alternating same
- Two handed different
- Passive base arm
- Reduplicated signs
- Lexical signs
- Compound signs
- Fingerspelled signs
- Loan signs
- Numbers
- Classifiers
- Show statistics

Can view statistics of the dataset

Example: Relationship between start and end handshapes in monomorphemic lexical signs.

Compound signs
Excel spreadsheet
Both-hands start → end HS
Both-hands HS variants
Non-dominant start → end HS

Decreasing frequency

START		END			
B-L 16.3		B-L 628	bent-B-L 63	10 23	flat-O 15
1 9.2		1 654	X 48	bent-1 19	5 6
S 6.9		S 339	5 67	1 28	crvd-5 26
5 6.9		5 404	S 69	flat-O 40	5-C 21

Search gloss, press enter for more

- ◆ EASY+
- ◆ EASY++
- ◆ EASY+++
- ◆ EASY+
- ◆ EASY++
- FLAG_2
- ◆ FLY-BY-WINGS
- GENERATIONS-AGO
- ◆ HAVE
- HEAVEN
- HEAVEN_2
- ◆ HOPE
- HYPOCRITE
- ◆ INTRODUCE
- KNOW-THAT
- ◆ LATE
- ALL-NIGHT_2
- ◆ FROM-NOW-ON
- ns-nat-JAMAICA
- ◆ ns-nat-KOREA
- PAINT
- ◆ PANTS+
- ◆ PANTS++
- PAST
- PAST+
- PAST
- PAST_2
- PAST_2
- ◆ PET/SPOILED
- SLOW
- ◆ SOAP

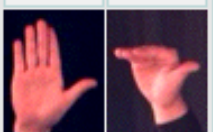
Search gloss, press enter for more


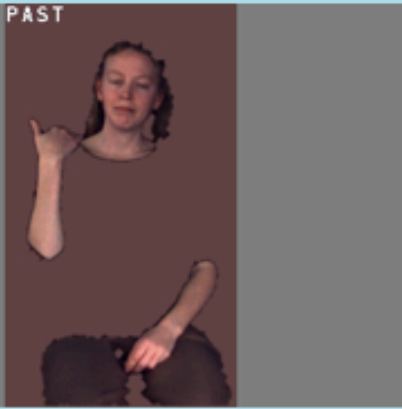


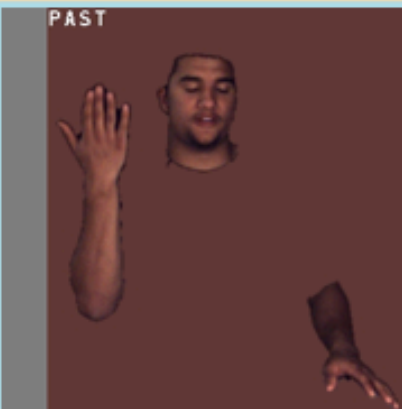
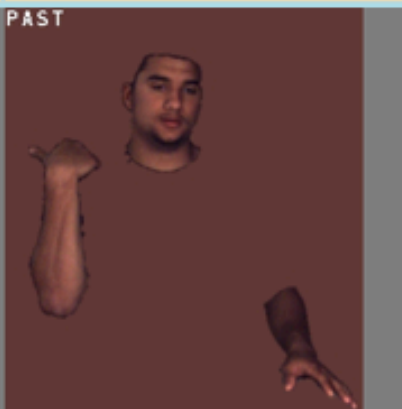


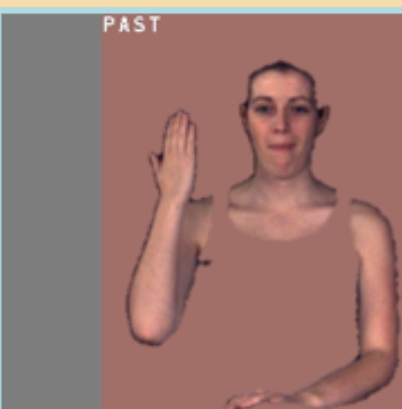
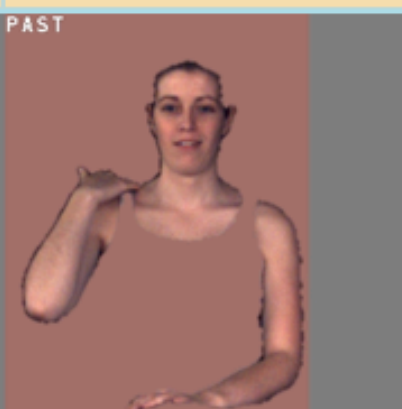


Select signs Select signs

Dominant sta Non dominant

Query start HS Query end HS Query start HS Query end HS

B-L bent-B-L select select



		Lexical sign 1	Lexical 1, Variant 1	
		PAST	PAST	
B-L				<i>bent-B-L</i>
				
		Lexical sign 1	Lexical 1, Variant 1	
		PAST	PAST	
B-L				<i>bent-B-L</i>
				
		Lexical sign 1	Lexical 1, Variant 1	
		PAST	PAST	
B-L				<i>bent-B-L</i>
				
		Lexical sign 1	Lexical 1, Variant 1	
















**Rich source of data for linguistic analysis of variation:
phonological and morphological**

Data can be viewed in various formats,
including a summary spreadsheet

Consultant	Main New Gloss	Gloss Variant	D Start HS	N-D Start HS	D End HS	N-D End HS
Liz	DOUBT	DOUBT	V/2		crvd-V	
Liz	DOUBT	DOUBT	V/2		crvd-V	
Tyler	DOUBT	DOUBT	V/2		crvd-V	
Naomi	DOUBT	DOUBT	V/2		crvd-V	
Brady	DOUBT	DOUBT	crvd-V		crvd-V	
Brady	DOUBT	DOUBT	V/2		crvd-V	
=====	=====	=====	=====	=====	=====	=====
Tyler	DONUT	DONUT	D	D	D	D
-----	-----	-----	-----	-----	-----	-----
Liz	DONUT	(R)DONUT	R	R	R	R
Naomi	DONUT	(R)DONUT	R	R	R	R
Brady	DONUT	(R)DONUT	R	R	R	R
=====	=====	=====	=====	=====	=====	=====
Liz	DOWN	DOWN	1		1	
Liz	DOWN	DOWN	1		1	
Tyler	DOWN	DOWN	1		1	
Naomi	DOWN	DOWN	1		1	
Naomi	DOWN	DOWN	bent-1		bent-1	
Brady	DOWN	DOWN	1		1	
Brady	DOWN	DOWN	bent-1		bent-1	
Lana	DOWN	DOWN	1		1	
Dana	DOWN	DOWN	1		1	
=====	=====	=====	=====	=====	=====	=====
Liz	DEEP	DEEP	bent-1	flat-B	bent-1	flat-B
Naomi	DEEP	DEEP	bent-1	B-L	bent-1	B-L
Brady	DEEP	DEEP	bent-1	B-L	bent-1	B-L
-----	-----	-----	-----	-----	-----	-----
Tyler	DEEP	DEEP_2	bent-1	B-L	bent-1	B-L
=====	=====	=====	=====	=====	=====	=====
Liz	DREAM	DREAM	1		X	
Tyler	DREAM	DREAM	1		X	
Naomi	DREAM	DREAM	1		X	
Brady	DREAM	DREAM	1		X	
=====	=====	=====	=====	=====	=====	=====
Liz	DRESSER	DRESSER	5-C-L	5-C-L	5-C-L	5-C-L
Naomi	DRESSER	DRESSER	crvd-5	crvd-5	crvd-5	crvd-5
Brady	DRESSER	DRESSER	5-C-L	5-C-L	5-C-L	5-C-L
-----	-----	-----	-----	-----	-----	-----

Rich source of data for linguistic analysis of assimilation effects

Compounds in the existing dataset

<p>flat-O</p> 	<p>EAT</p> 	<p>EAT</p> 	<p>flat-O</p> 	<p>EAT+MORNING</p> <p>EAT</p> <p>Lexical sign</p> <p>1-handed</p> <p>Dominant</p> <p>Same S,E</p> <p>Signer video</p> <p>Old S,E 1365, 1372</p> <p>1365 1372</p>				
	<p>Lexical sign 1</p>	<p>Lexical 1, Variant 1</p>		<p>EAT+MORNING</p> <p>MORNING</p> <p>Lexical sign</p> <p>2-handed same hs</p> <p>2-hands</p> <p>Same S,E</p> <p><input checked="" type="checkbox"/> Passive base arm</p> <p>Signer video</p> <p>Old S,E 1390, 1405</p> <p>1390 1405</p>				
<p>B-L</p> 	<p>B-L</p> 	<p>MORNING</p> 	<p>MORNING</p> 	<p>B-L</p> 	<p>B-L</p> 			
	<p>Compound sign 1</p>	<p>Compound 1, Variant 1</p>						<p>EAT+MORNING</p> <p>Compound sign</p> <p>Signer video</p> <p>Combined video</p> <p>Previous S,E 1365, 1405</p>
<p>flat-O</p> 	<p>EAT+MORNING</p> 	<p>EAT+MORNING</p> 		<p>B-L</p> 	<p>B-L</p> 			

Rich source of data for linguistic analysis of assimilation effects

Compounds in the existing dataset

Additional annotated data with same signs
used in continuous signing

Statistics

Not including number signs, fingerspelled and loan signs, and Classifier constructions

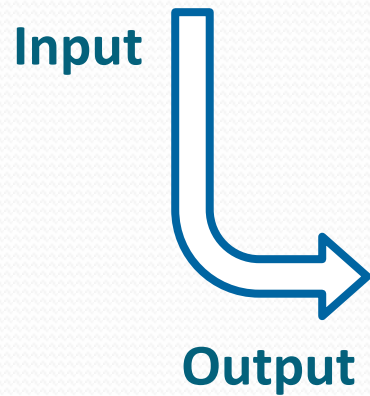
Class of signs	Number of signs	Number of sign variants	# sign variants with { 1, 2, 3, 4... }		# tokens (examples) per sign		Number of sign tokens	
			{ consultants }		{ 1,2,...,6, >6 }			
Monomorphemic lexical signs	2,284	2,793	x1	621	587	x1	8,585	Two-handed
			x2	989	858	x2		5,713
			x3	394	386	x3		67%
			x4	563	491	x4		
			x5	85	142	x5		One-handed
			x6	141	154	x6		2,873
					175	>6		33%
Compound signs	289	329	x1	129	117	x1	749	
			x2	106	107	x2		
			x3	48	46	x3		
			x4	33	33	x4		
			x5	4	11	x5		
			x6	9	13	x6		
					2	>6		

II. A computer science application

Focus: Handshape recognition in monomorphemic lexical signs



Challenge: to recognize 3D hand configurations from 2D images



Approx. 85 handshape labels

A	10	S	cocked-S	T	X-over-thumb	1	D	X	bent-1	G/Q	flat-G	alt-G			
U/H	crvd-U	bent-U	cocked-U	U-L	bent-U-L	V/2	crvd-V	bent-V	3	crvd-3	6	W	crvd-W		
7	cocked-7	open-7	8	cocked-8	open-8	25	F/9	cocked-F	open-F	4	5	crvd-5	5-C	5-C-L	5-C-tt
B	B-xd	flat-B	B-L	Vulcan	crvd-B	crvd-flat-B	crvd-sprd-B	bent-B	bent-B-xd	bent-B-L	C	sml-C/3	tight-C	tight-C/2	
O	baby-O	flat-O	flat-O/2	fanned-flat-O	E	loose-E	M	alt-M	bent-M	full-M	N	alt-N	bent-N		
P/K	alt-P	L	L-X	crvd-L	R	R-L	I	Y	I-L-Y	bent-I-L-Y	Horns	bent-Horns	O/2-Horns		

To improve recognition:

Exploit linguistic constraints on the relationships between

- ⊙ Start and end handshapes

Only certain types of handshape changes are allowed within a monomorphemic lexical sign.

[HS1, HS2]

- ⊙ Handshapes on the two hands

The two hands

Based on Battison's *
(1978) typology

1-handed

2-handed

- **same HS**
- **different HS** (ND = unmarked)

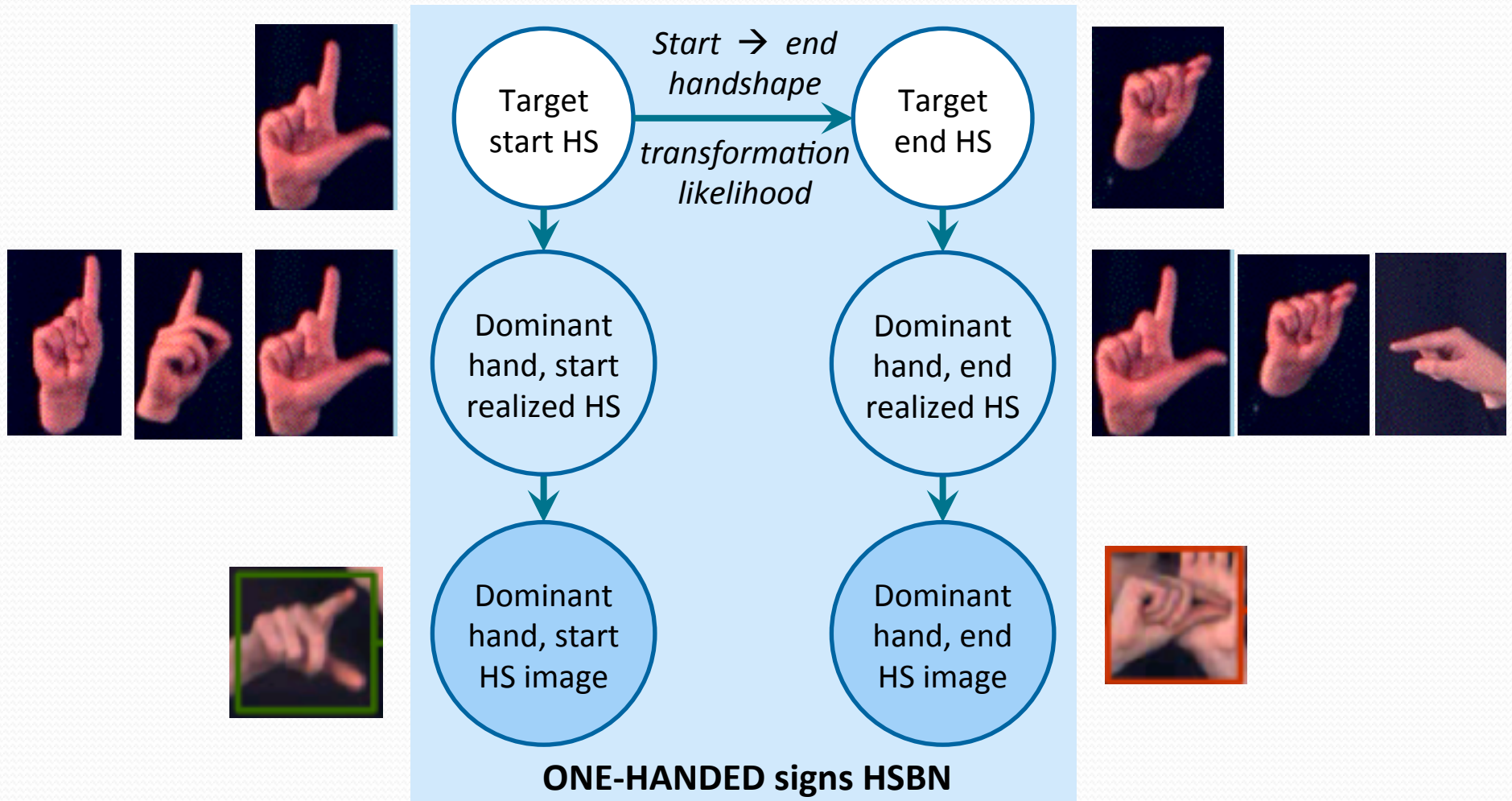
To improve recognition:

Exploit linguistic constraints on the relationships between

- ⊙ Start and end handshapes
- ⊙ Handshapes on the two hands

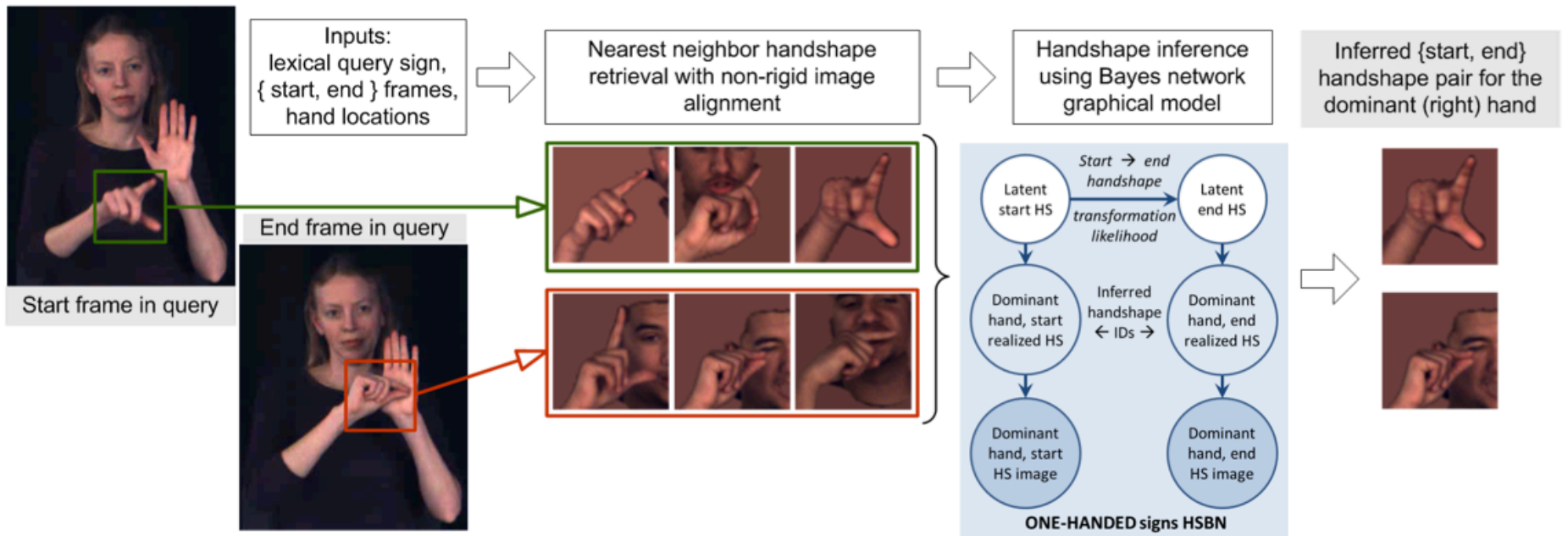
Machine learning of these relationships from the the statistics of the dataset

HMM for one-handed signs

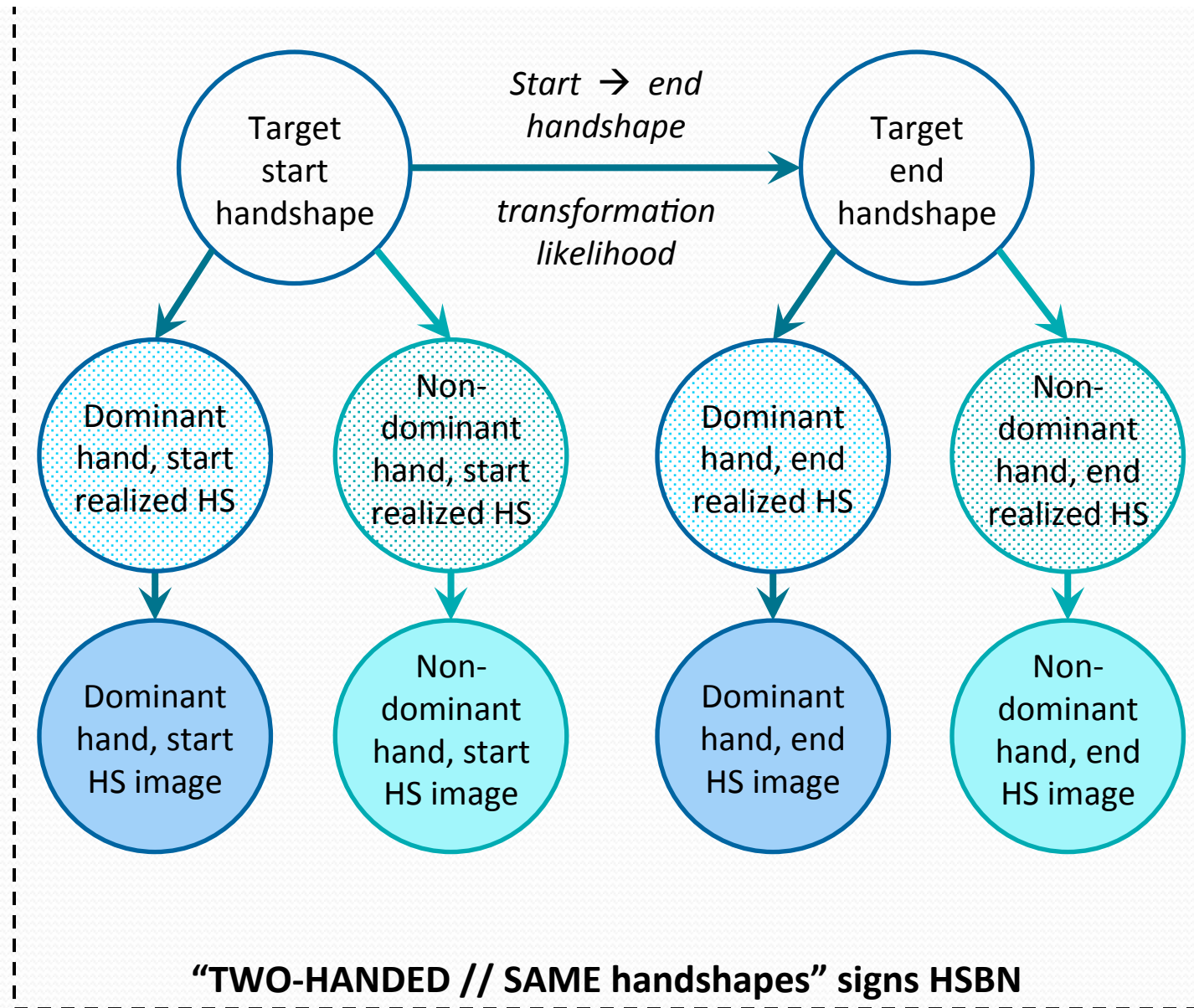


Hand Shape Bayesian Network

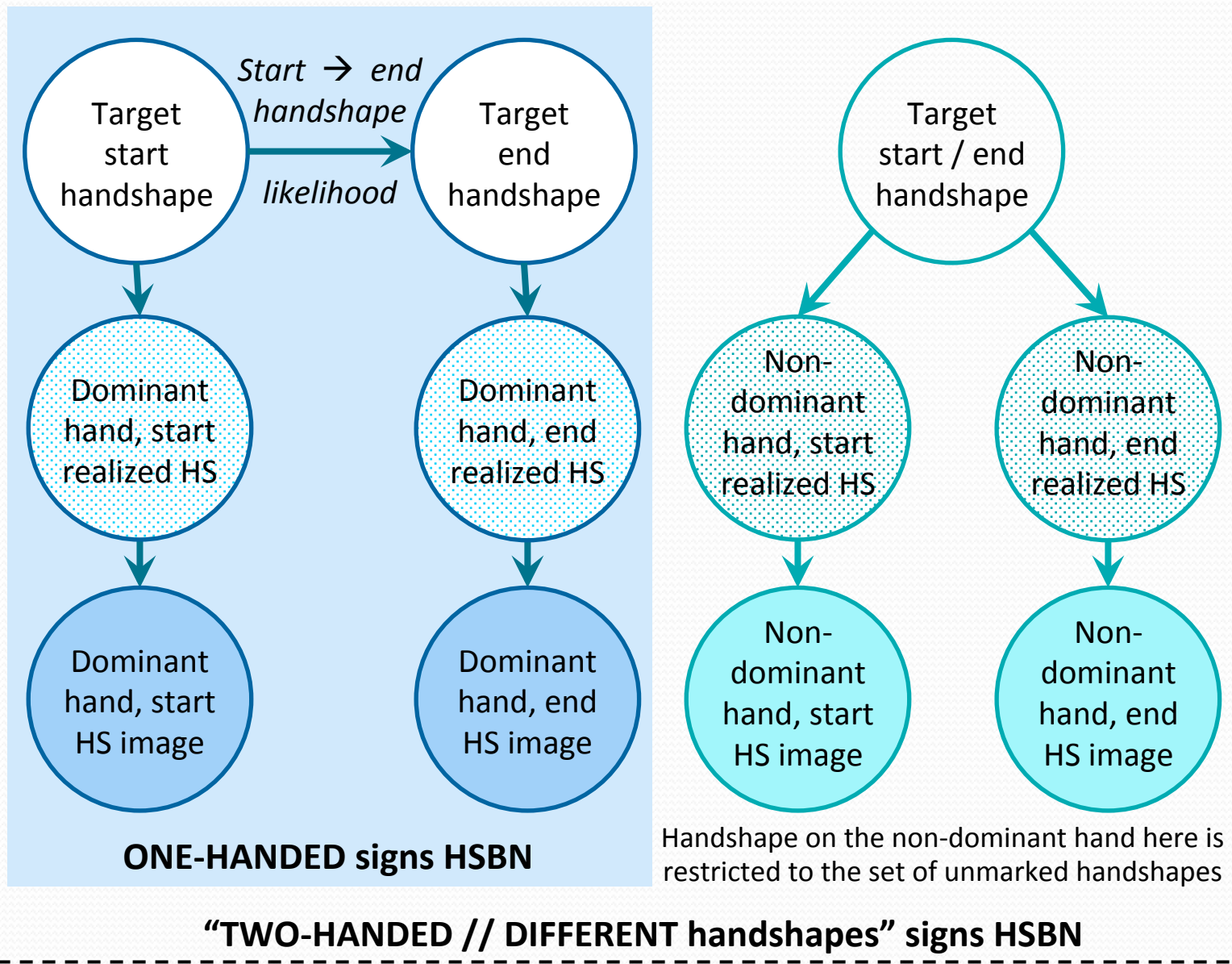
Illustration of the “Hand Shape Bayesian Network” used to infer handshapes on 1 hand

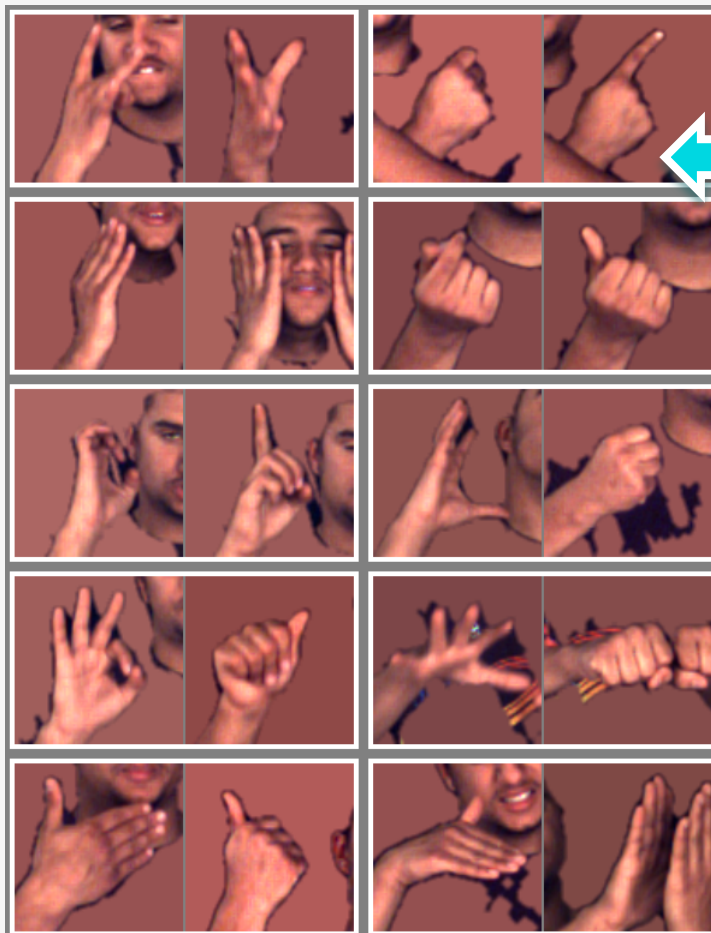


For two-handed signs: same handshapes

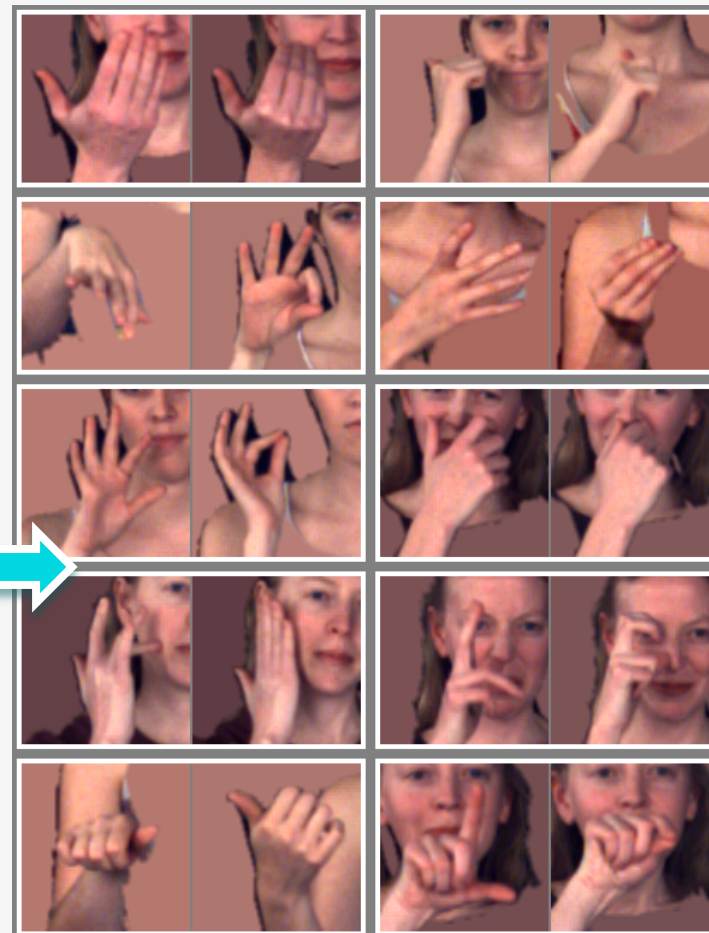


For two-handed signs: different handshapes





Handshape pairs in query signs



Handshape pairs in database signs

Sequestered



1962 handshapes (981 pairs) from
 333 one-handed signs &
 324 two-handed:same handshapes signs
 657 signs used for model evaluation

5232 handshapes from
 1601 one-handed & two-handed:same
 handshapes signs
 6862 signs used for model training

RESULTS:

30.4% (597 of 1962) nearest neighbor handshape recognition accuracy

44.4% (871 of 1962) handshape recognition accuracy using *HSBN model*

Future directions

One major target application

Integration of look-up technology with the DawnSignPress all-ASL multimedia dictionary currently under development

(Ben Bahan, MJ Bienvenu, Beth Benedict, et al., professors at Gallaudet University)

Further extensions for sign search and retrieval, e.g., *Sloogle*)

Related work in progress:

Dataset is

- ◆ **currently being expanded;**
- ◆ **serving as the basis for collaborative research (Dimitris Metaxas, Rutgers University)**
 - ◎ More comprehensive linguistic modeling for recognition of signs based on their morpho-phonological typology
 - ◎ Recognition of signs from continuous signing
 - ◎ Using 3D model-based tracking methods, followed by robust learning methods based on the parameters extracted from the 3D tracking

Dataset can be used for a wide variety of applications











Example: Matt Huenerfauth (CUNY)

- ◆ Using the handshape data as a foundation for building his animation lexicon.
- ◆ Using labels as the gloss-ID standard for annotations of his motion capture data.

Integrate these data and some of the search functionalities with our other publicly accessible corpus (poster earlier) ASLLRP DAI

<http://secrets.rutgers.edu/dai/queryPages/>

2 videos selected [Add to cart](#) **Actions** Video perspective: **Front**

File Name-Utterance	Utterance Video	Sign Video	Full Gloss	Rough Gloss
<input checked="" type="checkbox"/> ncslgr10a.xml-2			Show ...	wh question: _____q/wh main gloss: fs-JOHN FINISH READ BOOK WHEN
<input type="checkbox"/> ncslgr10a.xml-3			Show ...	wh question: _____q/wh main gloss: fs-JOHN FINISH READ BOOK WHEN
<input checked="" type="checkbox"/> ncslgr10a.xml-4			Show ...	negative: _____neg main gloss: fs-JOHN NOT-YET FINISH READ BOOK
<input type="checkbox"/> ncslgr10a.xml-5			Show ...	wh question: _____q/wh main gloss: fs-JOHN FUTURE FINISH READ BOOK WHEN
<input type="checkbox"/> ncslgr10a.xml-6			Show ...	wh question: _____q/wh main gloss: fs-JOHN FUTURE FINISH READ BOOK WHEN

Demo

*Acknowledgments

We are grateful to Vassilis Athitsos, Joan Nash, Donna Riggle, Jessica Scott, Ben Bahan, Rachel Benedict, Naomi Berlove, Elizabeth Cassidy, Lana Cook, Jaimee DiMarco, Danny Ferro, Alix Kraminitz, Indya Oliver, Caelan Pacelli, Braden Painter, Chrisann Papera, Tyler Richard, Tory Sampson, Dana Schlang, Jon Suen, and Iryna Zhuravlova for invaluable assistance with this project. The work reported here has been supported in part by grants from the National Science Foundation (#IIS-0705749, CNS-0855065, CNS 1059218, IIS-0964385).