DATA CODING OF, AND TRANSITION TIME IN, ASL FINGERSPELLING

Jonathan Keane

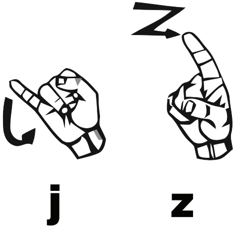
University of Chicago

Outline

Background – Fingerspelling

All handshapes are static except for -J- and -Z-.

Fingerspelling makes up anywhere from 12–35% of ASL discourse.
(Padden, 1991; Padden and Gunsauls, 2003)



The phonetics of Fingerspelling

Wilcox (1992) looks at about 7 words and describes some of the dynamics of hand motion.

Tyrone et al. (1999) looks at fingerspelling by parkinsonian signers from a phonetic perspective.

Brentari and Padden (2001); Cormier et al. (2008) both look at the nativization process for fingerspelled words.

Quinto-Pozos (2010) described the rate of fingerspelling for two signers within fluent discourse.

What data looks like

data.mpg

Data coding needs to be

Accurate

Accurate, detailed data is necessary for any linguistic analysis.

Reproducible

Coding should be able to be reproduced, and individual coders should form some sort of consensus.

Quick

Coding time is often directly related to the amount of data we can collect.

Easy

A coding system that requires little specialized training is better than one that requires experts to use. (All else being equal)

Recording specifications

Signers

2 signers, both are deaf of deaf parents, and native ASL users.

Video

2 video cameras recording at 60 FPS.

We collected 2 sessions for each signer
1 at a normal, conversational speed, and 1 at a careful speed.

There were 100 non-English words, 100 names, 100 nouns.

Each word was fingerspelled twice in each speed.

The video was then post processed and compressed for coding.

Session details

Careful elicitation and data collection allowed us to maximize the data we started with.

1. Output a logfile with the order of the words as they were presented to the signer (although this could be improved further)
2. Segmentation – green button
3. First pass error detection – red button

3–4 humans hand coded apogees

Using ELAN, coders watched the videos at 20–40% speed.

Told to press a button whenever they thought there was an apogee.

- Described as the point where the hand was maximally or minimally open.
- Could be described as the minimum instantaneous velocity of all of the articulators.
- Use discretion when coding apogees with movement, but be consistent.
- Not defined as the canonical form

The position of each apogee was algorithmically determined.

Minimized the mean absolute distance between the points for each word.

We accounted for errant, and missing presses by assigning a violation cost for every apogee that was deleted or added.

The coders were already fairly close together.
- Mean absolute deviation:
  - 27.93 msec for all letters
  - 62.52 msec for letters with movement

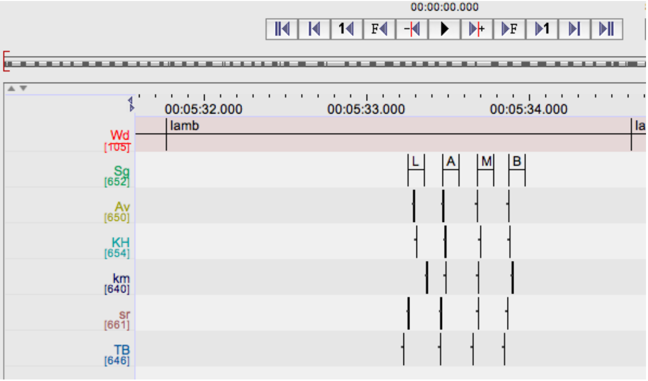
Leveraging known data

A first guess at the letter of each apogee was added using left edge forced alignment.

Although the letters it assigns are not 100% accurate, they are close.

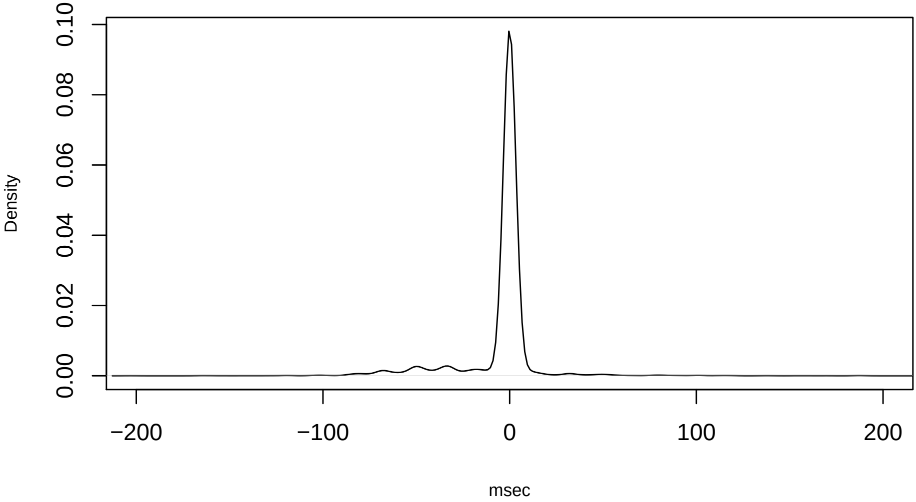
Finally someone trained in fingerspelling went through and verified the location, and letter of each apogee. The vast majority of apogees are unchanged.

Example


00:00:00.000
00:05:32.000
00:05:33.000
00:05:34.000
lamb
la
Wd [105]
Sg [652]
Av [650]
KH [654]
km [640]
sr [661]
TB [646]
L
A
M
B

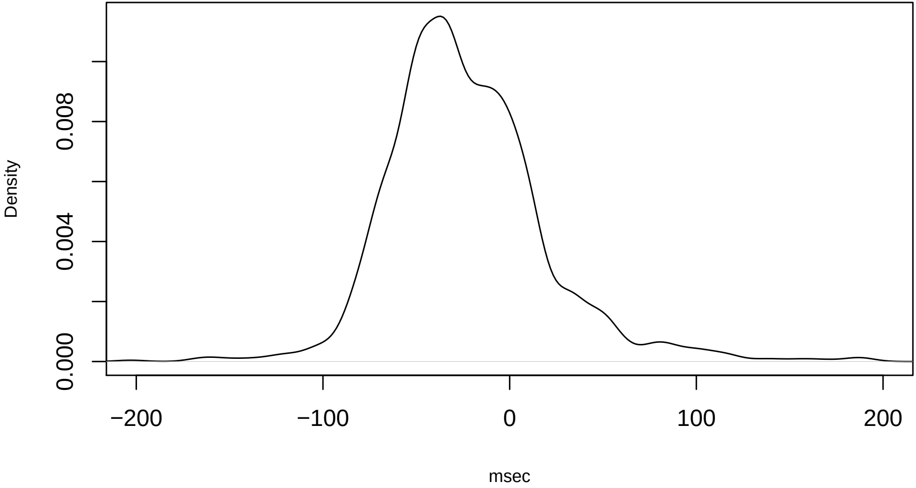
Verification

5404 apogees of 6594 in normal are unchanged. (~ 81%)



Verification

Of the changed apogees, they are often shifted back by 2–3 frames.



How does our method look?

Accurate

Much less chance for transcription or other errors compared to traditional methods.

Reproducible

The first stage of coding is incredibly reproducible, to a high degree of accuracy. The second step might be (we're working on testing this)

Quick

Varied, but each clip chunk (5–10 min of data) took 105 min for 3 coders, 75 min of which can be done simultaneously by 3 coders, so 55 min at the fastest.

Easy

Our initial pass for coding can be done with very little training. The verification task requires a bit of training in fingerspelling, but since it's more confirmation, it's much easier and quicker than

Example

staxi.mp4

Figure: T-A-X-I, normal speed.

Example

staxi2.mp4

Figure: T-A-X-I, normal speed, slow motion

Example, revisited

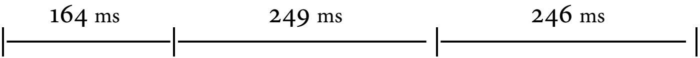


Figure: durations for T-A-X-I

T-A-X-I, L-A-M-B, F-R-E-D, C-A-R-P, and P-U-H-U

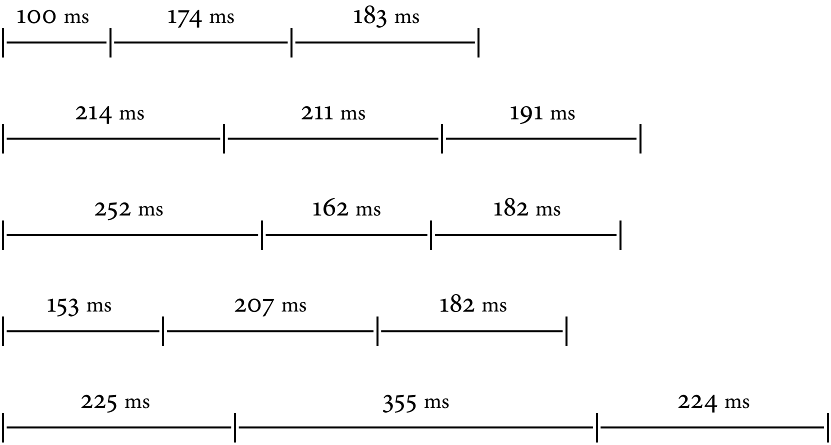


Figure: durations for T-A-X-I, L-A-M-B, F-R-E-D, C-A-R-P, and P-U-H-U (signer: s1 speed: **normal**)

Questions

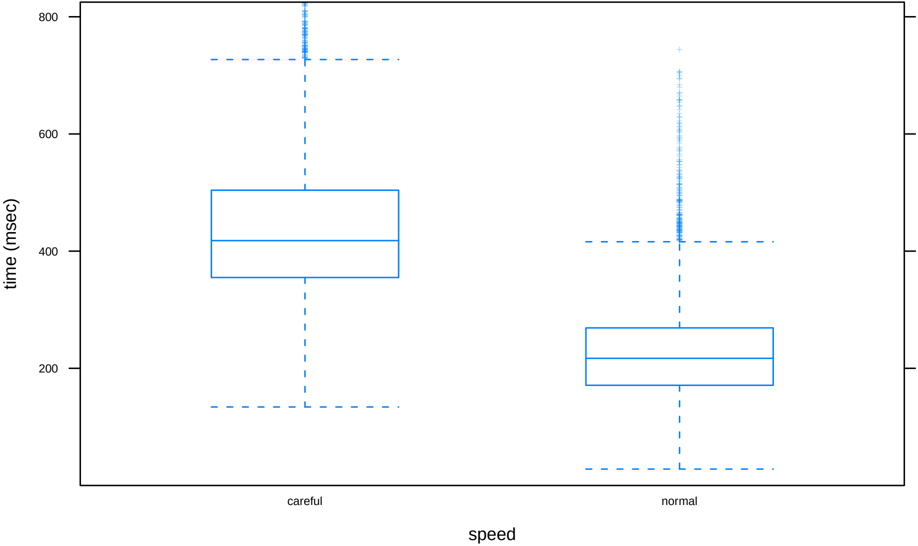
1. When asked to sign at two different speeds, how much of a difference is there between them?
2. Is there individual variation?
3. Does the type of word affect the speed of fingerspelling?
4. Does the position of a letter in a word affect transition time?
5. Do letters with movement take longer to execute?
6. Does articulatory complexity change transition time?

ANOVA table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
wordtype	2	33.19	16.59	220.34	0.0000
speed	1	1425.48	1425.48	18927.22	0.0000
signer	1	13.90	13.90	184.55	0.0000
wordtype:speed	2	13.04	6.52	86.58	0.0000
wordtype:signer	2	0.20	0.10	1.36	0.2570
speed:signer	1	232.93	232.93	3092.84	0.0000
Residuals	12267	923.87	0.08		

Table: ANOVA table for log(time)

between speeds

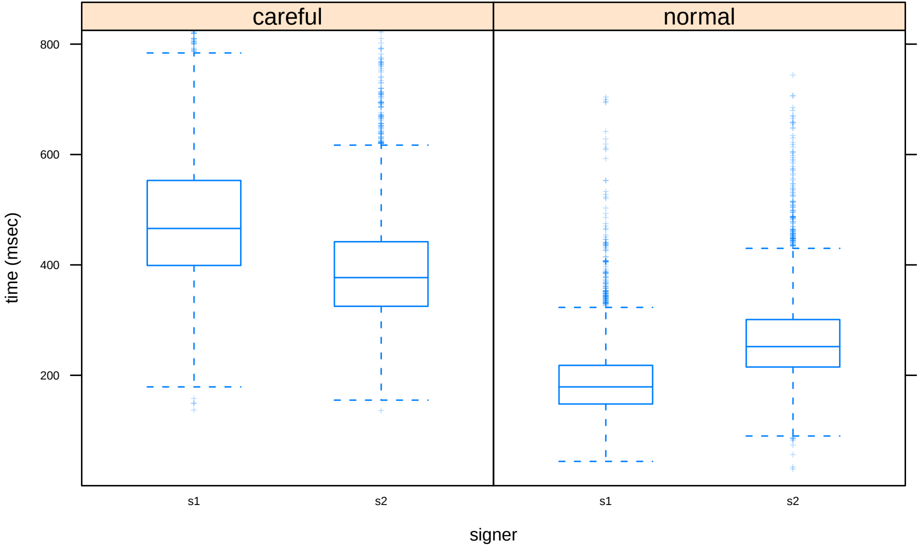


ANOVA table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
wordtype	2	33.27	16.64	219.11	0.0000
speed	1	1424.43	1424.43	18760.33	0.0000
signer	1	13.96	13.96	183.82	0.0000
wordtype:speed	2	13.57	6.79	89.38	0.0000
wordtype:signer	2	0.24	0.12	1.60	0.2020
speed:signer	1	234.20	234.20	3084.53	0.0000
Residuals	12354	936.53	0.08		

Table: ANOVA table for log(time)

between signers, by speed

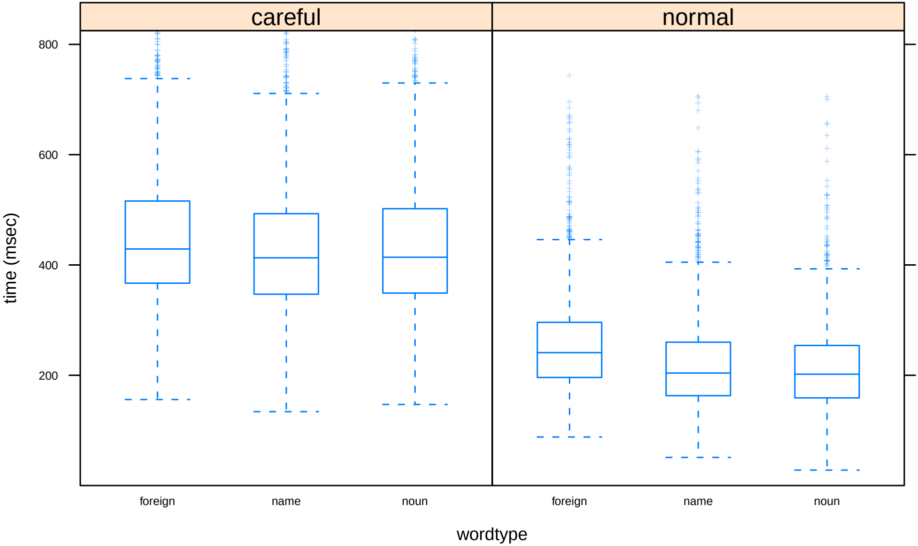


ANOVA table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
wordtype	2	33.27	16.64	219.11	0.0000
speed	1	1424.43	1424.43	18760.33	0.0000
signer	1	13.96	13.96	183.82	0.0000
wordtype:speed	2	13.57	6.79	89.38	0.0000
wordtype:signer	2	0.24	0.12	1.60	0.2020
speed:signer	1	234.20	234.20	3084.53	0.0000
Residuals	12354	936.53	0.08		

Table: ANOVA table for log(time)

between wordtypes, by speed

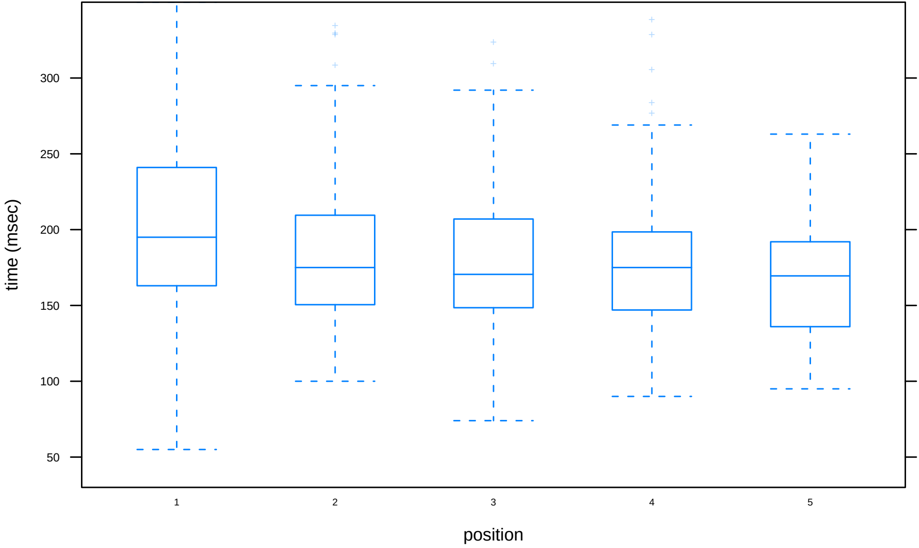


ANOVA table

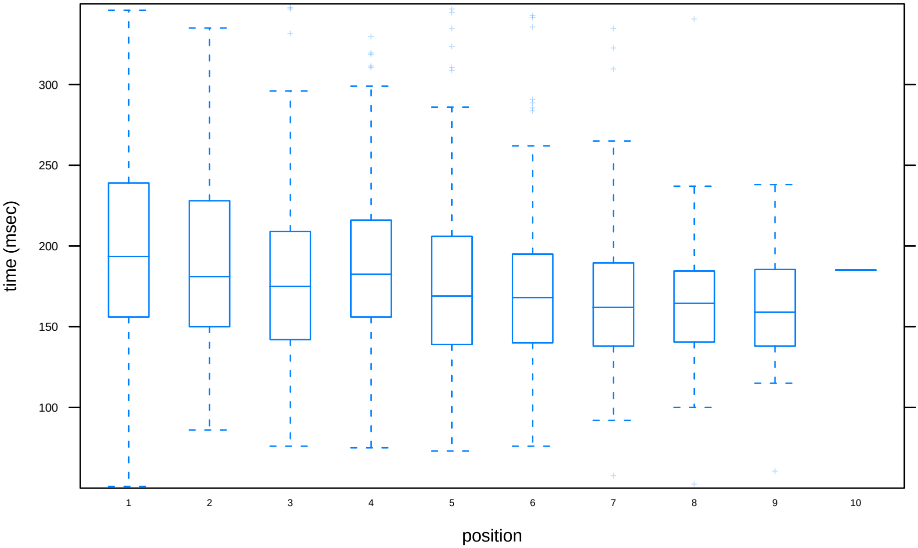
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
wordtype	2	33.27	16.64	219.11	0.0000
speed	1	1424.43	1424.43	18760.33	0.0000
signer	1	13.96	13.96	183.82	0.0000
wordtype:speed	2	13.57	6.79	89.38	0.0000
wordtype:signer	2	0.24	0.12	1.60	0.2020
speed:signer	1	234.20	234.20	3084.53	0.0000
Residuals	12354	936.53	0.08		

Table: ANOVA table for log(time)

Short words (3 – 6 letters) – s1, normal



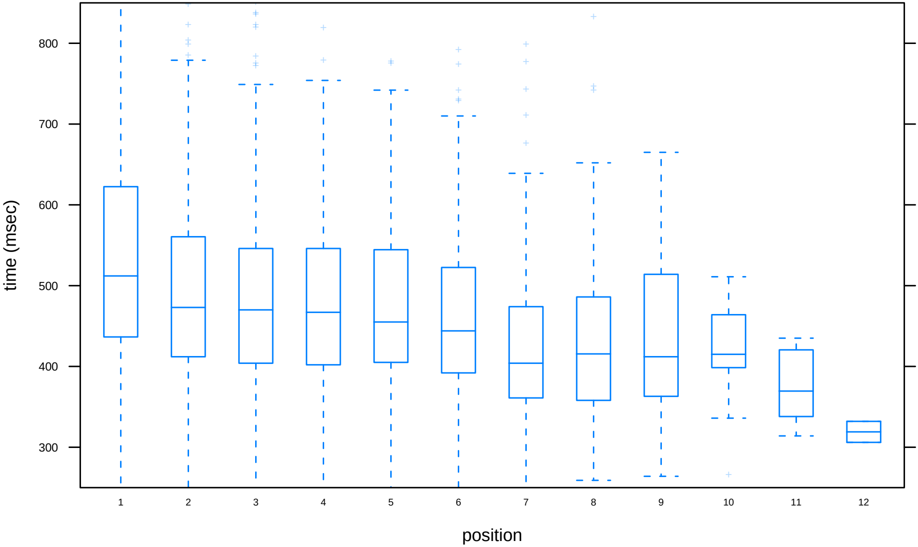
Long words (8 – 10 letters) – s1, normal



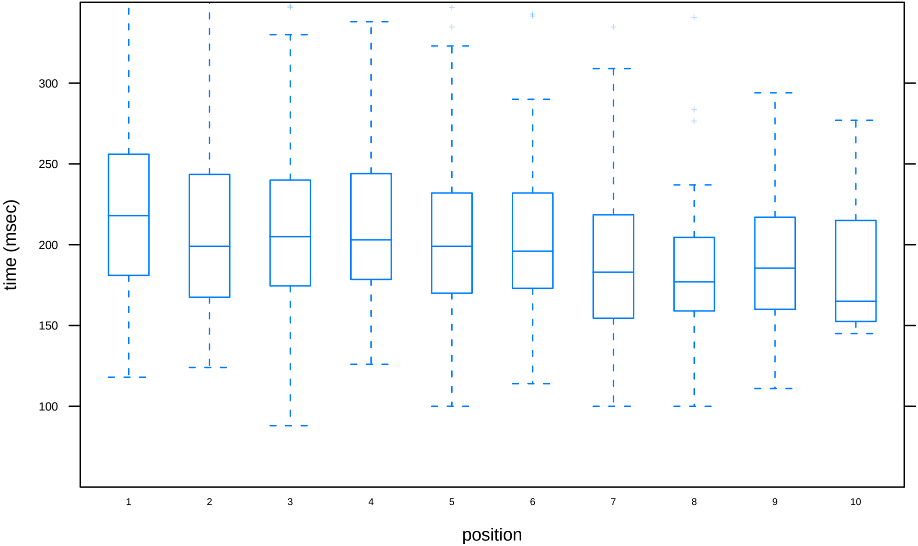
Possible explanations

1. memory limitations
2. articulation limitations
3. phonological chunking
   4 letters ~= 3 movements ~= 1 ASL sign

careful is different



non-english is different



T-A-X-I, L-A-M-B, F-R-E-D, C-A-R-P, and P-U-H-U

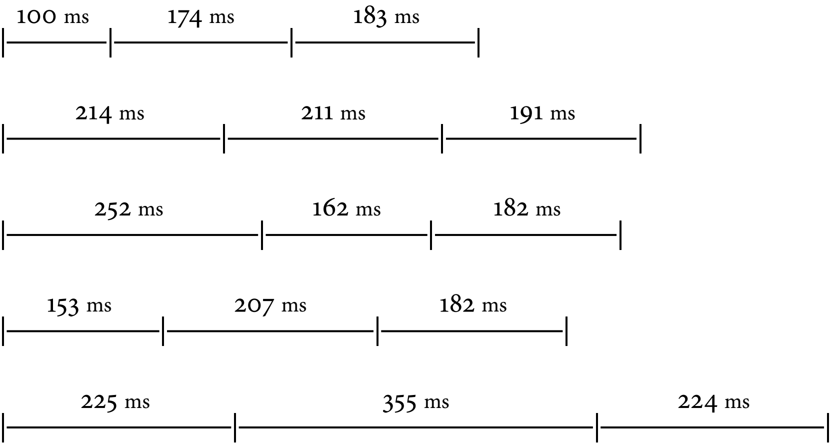


Figure: durations for T-A-X-I, L-A-M-B, F-R-E-D, C-A-R-P, and P-U-H-U (signer: s1 speed: **normal**)

T-A-X-I, L-A-M-B, F-R-E-D, C-A-R-P, and P-U-H-U

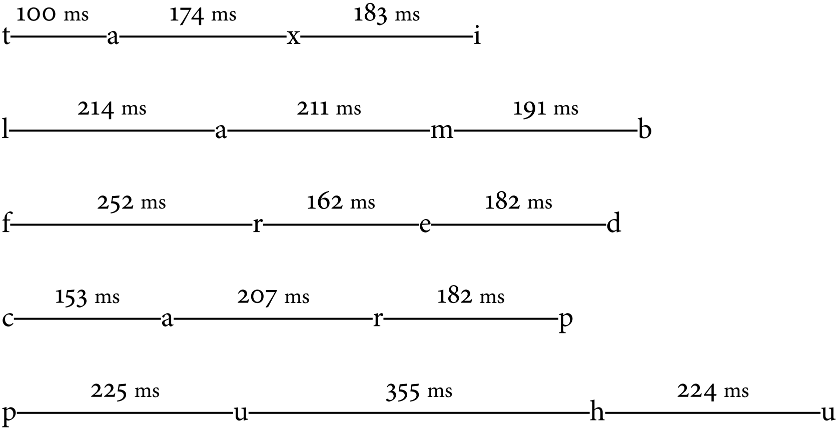
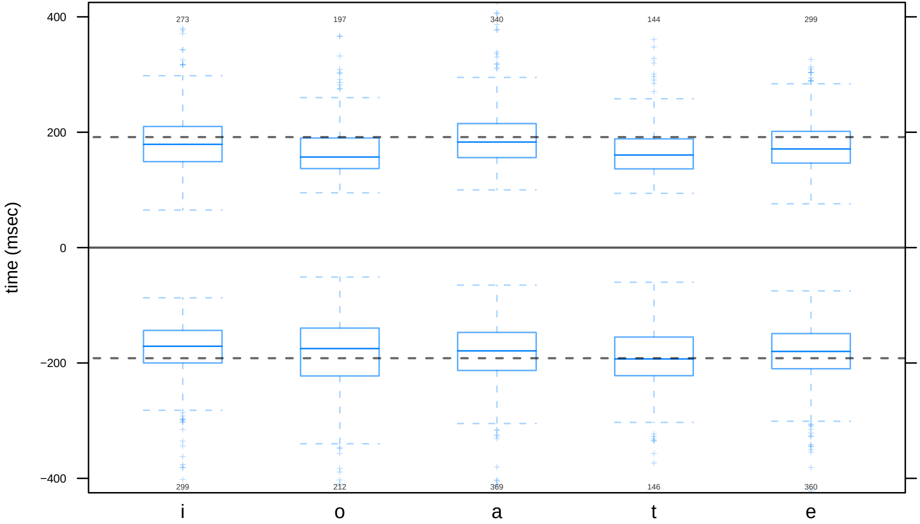
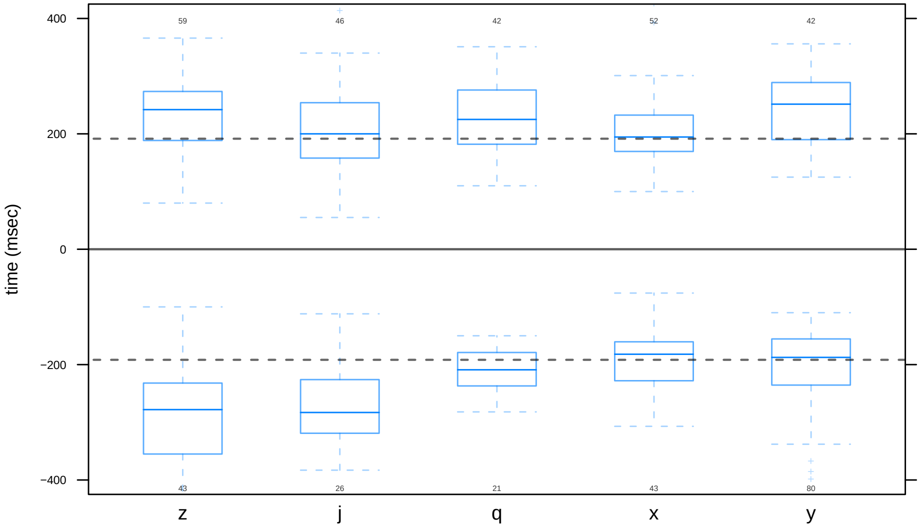


Figure: durations for T-A-X-I, L-A-M-B, F-R-E-D, C-A-R-P, and P-U-H-U (signer: s1 speed: **normal**)

transitions for high frequency letters



transitions for for low frequency letters

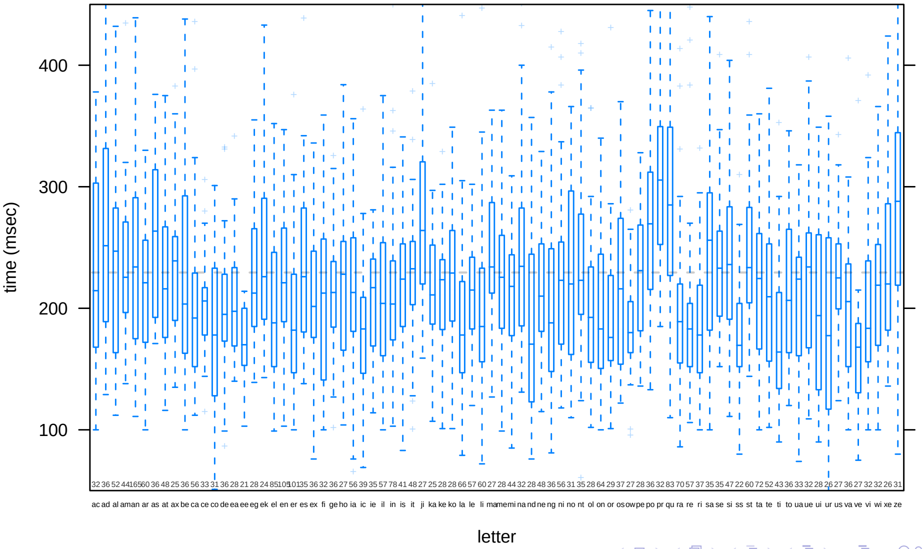


Movement in -Y-

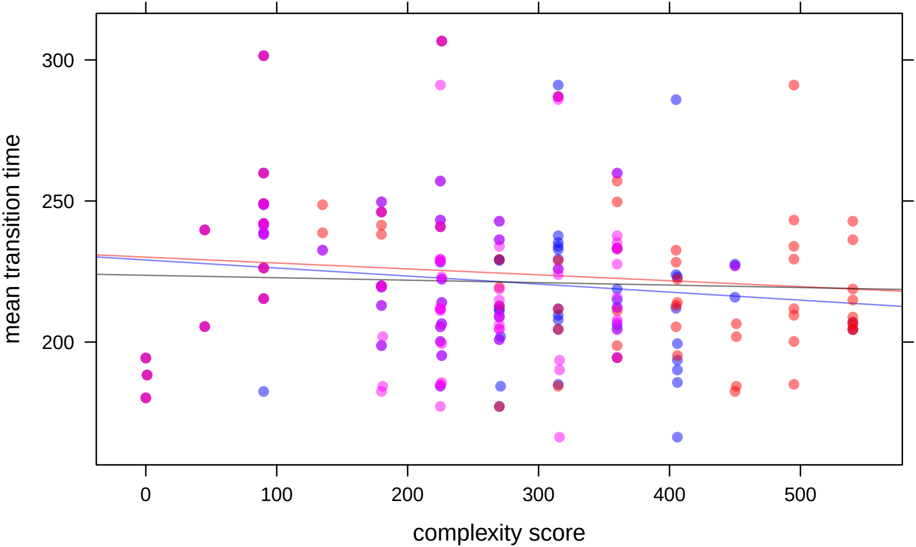
sNormal.mp4

Figure: The first 10 instances of -Y- – not at the end of the word

82 bigrams with more than 50 instances



Transition time by articulatory complexity



Conclusions

1. When asked to fingerspell at different speeds, the spread is significant.
2. There is individual variation overall and in speed, but not wordtype.
3. Signers fingerspell slower on non-English words.
4. Signers seem to chunk their production into 3-4 letter chunks with longer words.
5. Letters with movement take longer to execute.
6. The class of letters that have movement might need redefining: -Y- and possibly -Q-.
7. Transition time does not seem to correlate with articulatory complexity.

Future directions

1. More sophisticated modeling
2. Quantification of other articulatory features
3. Recognition related tasks
4. More signers (in progress!)

Thank you for coming.

I must also acknowledge the contributions of many who contributed in ways big and small:

Fingerspelling data
Drucilla Ronchen and Andy Gabel

Main advisors
Jason Riggle and Diane Brentari

Other researchers
Susan Rizzo, Karen Livescu, Greg Shakhnarovich, Raquel Urtasun, Erin Dhalgren, and Katie Henry.

Brentari, Diane, and C. Padden. 2001. Foreign vocabulary in sign languages: A cross-linguistic investigation of word formation, chap. Native and foreign vocabulary in American Sign Language: A lexicon with multiple origins, 87–119. Mahwah, NJ: Lawrence Erlbaum.

Cormier, Kearsy, Adam Schembri, and Martha E. Tyrone. 2008. One hand or two? nativisation of fingerspelling in asl and banzsl. Sign Language & Linguistics 11.3–44.

Padden, Carol. 1991. Theoretical issues in sign language research, chap. The Acquisition of Fingerspelling by Deaf Children, 191–210. The University of Chicago press.

Padden, Carol, and Darline Clark Gunsauls. 2003. How the alphabet came to be used in a sign language. Sign Language Studies 4.10–33.

Quinto-Pozos, David. 2010. Rates of fingerspelling in american sign language. Poster at TISLR 10.

Tyrone, M.E., J. Kegl, and H. Poizner. 1999. Interarticulator co-ordination in deaf signers with parkinson's disease. Neuropsychologia 37.1271–1283.

Wilcox, Sherman. 1992. The phonetics of fingerspelling. John Benjamins Publishing Company.

between wordtypes, by speed and by signer