Using N-Gram Analytics to Improve Automatic Fingerspelling Generation

John McDonald[*], Rosalee Wolfe, Souad Baowidan, Ningshan Guo, Sarah Johnson, Robyn Moncrief

School of Computing, DePaul University, United States

Abstract Fingerspelling recognition is one of the last skills acquired, due to the complex nature of fingerspelling and a lack of display technology that is sufficiently natural for recognition practice. This paper describes a corpus-based study utilizing an n-gram extension to ELAN to gain a deeper understanding of deletion and coarticulation in fingerspelling. The analysis shows that coarticulation and deletion increase with fingerspelling speed and that deletions form an increasing percentage of the modifications at shorter durations. Insights from the study informed strategies to improve current avatar-based fingerspelling generation.

Keywords Fingerspelling, Corpus Analysis, Sign Language animation, Interpreter Training

1. Introduction

In 2004 researchers at The Language Archive introduced ELAN ("Eudico Language Annotator") (Max Planck Institute for Psycholinguistics, 2013), an annotation tool that features synchronized video, audio and annotations. Its major applications include gesture studies, documentation of endangered languages and analysis of sign languages (Brugman & Russel, 2004). In subsequent years, researchers added support for synchronized motion capture (Crasborn, Sloetjes, Auer, & Wittenburg, 2006), broadened the scope of video support including enhanced time accuracy (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006), and explored automated techniques (Dreuw & Ney, 2008) to reduce the time required to annotate media (Auer, Russel, Sloetjes, Wittenburg, Schreer, & Masnieri, 2010). In 2008, new extensions allowed users to create references from annotation systems defined in the central International Standards Organization Data Category Repository (ISO DCR) (Sloetjes & Wittenburg, 2008). The goal is to continue to foster greater data sharing among language researchers.

Part of ELAN's appeal derives from its powerful and diverse search tools (Stehouwer & Auer, 2011). These provide an immense gamut of search granularity, ranging from finding individual annotations in local files, to accessing web-based corpora. Users can also search for *n-grams* within a single tier of annotation codes or for phenomena that co-occur on multiple tiers (Crasborn, Hulsbocsh, Lampen, & Sloetjes, 2013). The term n-gram refers to a contiguous sequence of n letters from a word. The two main formats for search results are the concordance and frequency views. In either view, users can elect to "Show hit in transcription", which cues the linked media and annotation to the position where the hit occurs in the ELAN annotation file.

Statistical services available include frequency counts for search queries. The "Statistics for multiple files" search also includes basic descriptive statistics for hits within a tier, including a duration's minimum, maximum, mean and median. For further analysis, ELAN provides capability to export the raw search results for further study.

The remaining sections describe the application of ELAN's n-gram analysis to improve technology for acquisition of fingerspelling recognition skills. The Materials and Methods section reviews the properties of fingerspelling that contribute to the challenges for sign language learners, and discusses the shortcomings of current technologies for self-study. The Results section describes a new corpus of fingerspelling examples, and the use of the ELAN n-gram software module to examine certain properties of fingerspelling. The Discussion and Conclusions section considers the incorporation of the newly-discovered properties into an improved technology for automatically generating fingerspelling via an avatar which will provide a better tool for self-study.

2. Materials and Methods

Fingerspelling is difficult for hearing L2 learners to master. The next two subsections describe contributing factors to this phenomenon.

2.1. Uses and Properties of Fingerspelling

Fingerspelling is the process of spelling words using signs that represent individual letters of a language alphabet (Sutton-Spence & Woll, 1998). These signs are known collectively as a *manual alphabet*. Figure 1 illustrates the manual alphabet for American Sign Language (ASL).

Fingerspelling is an integral part of ASL and also many other signed languages around the world. In these languages, it is a necessary skill for complete communication (Battison, 1978). In ASL, fingerspelling is used to spell proper nouns, technical terms that lack a universally-accepted sign, places without a sign name and loan words. It can also convey words from a spoken language that a signer does not know in the signed language, and to clarify signs that are unknown to others. Signed communication systems based on English, such as Signed English (SE), Signing Exact English (SEE) and Pidgin Signed English (PSE), also rely on fingerspelling as an essential means of communication (Strong, 1988).
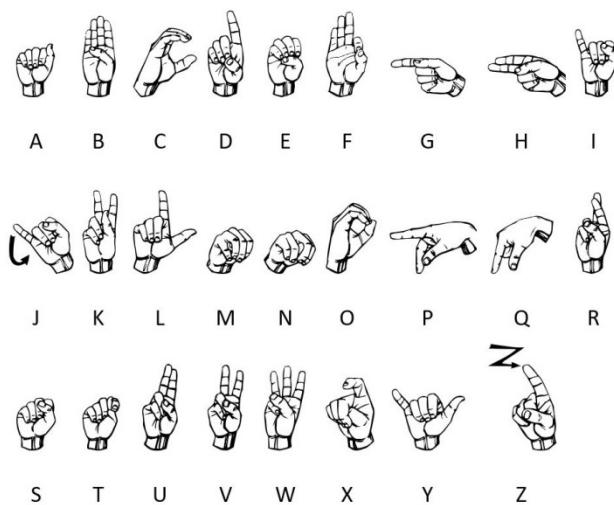

Figure 1. The manual alphabet of American Sign Language (ASL)

Padden (Padden, 1991) found that on average, fingerspelling comprises approximately six percent of the signs produced in everyday ASL conversations, but that in certain contexts, fingerspelling can form as much as twelve percent of the signs present in an ASL discourse.

2.2. Factors Contributing to the Difficulty of Fingerspelling Reception

American Sign Language learners find fingerspelling receptive skills to be one of the most difficult aspects of sign language to master. In interpreter training programs it is the first skill taught, but the last skill mastered (Grushkin, 1998). Patrie notes that "hearing people who are learning ASL as adults tend to have great difficulty in correctly recognizing fingerspelled words" (p. 19). There are two contributing factors to this difficulty.

The first challenge is the marked difference between how the handshapes actually appear in fingerspelling production and the idealized manual letters shown in textbooks. These differences arise from 1) speed of production, 2) the motion

of the transitions between fingerspelled letters and 3) the precision in forming individual fingerspelled letters in a word. The second challenge is the lack of practice opportunities, which will be discussed in the following subsection.

The speed of production can vary, depending on the context of the fingerspelling. In *careful fingerspelling,* fluent signers produce fingerspelled letters at a rate of four per second, and in *rapid fingerspelling,* production speed can rise to a rate of six letters per second (Patrie & Johnson, 2011). This is in contrast to a sign in ASL, which has at most two hand shapes on the dominant hand (Battison, 1978). Thus, a person observing fingerspelling needs to comprehend a larger number of handshapes being produced at a faster rate.

The transitions between fingerspelled letters also pose challenges to fingerspelling reception. Fingerspelling is more than the production of a sequence of static hand configurations. Studies as early as 1971 (Zakia & Haber, 1971) suggest that a person fluent in ASL does not read individual letters, but rather the total pattern of the motion. Particularly in rapid fingerspelling, it is a smoothly-flowing motion that does not come to rest until the last letter. Akamatsu (Akamatsu, 1982) called this the "motion envelope'. Wilcox (Wilcox, 1992) posits that learning to fingerspell involves learning both the static hand configurations and the set of possible transitions. He created a model of targets and transitions suggesting that fingerspelling can be seen as a series of movements.

Lastly, discerning the individual letters in a fingerspelled word also depends on the degree to which any instantaneous hand configuration in a fingerspelled word will match one of the 26 canonical fingerspelled letters. In *careful fingerspelling* the signs representing individual letters are "produced fully and completely" (Patrie & Johnson, 2011). Careful fingerspelling typically occurs when a word first appears in a discourse. It also occurs in response to such questions as "What is the English word for _____" or "What is your name?" When a word appears in subsequent occurrences, a signer will spell it more rapidly. *Rapid fingerspelling* also occurs in informal settings. Of the two styles, rapid fingerspelling appears more frequently. In rapid fingerspelling, the hand movement is a smoothly-flowing organic whole.

While forming the movement comprising a fingerspelled word, the individual handshapes can influence each other in a manner similar to the way that spoken words can influence each other (Wilcox, 1992). The effects of such *coarticulation* can cause a blending of one fingerspelled letter into the next so that the forms produced in the fingerspelled word differ from the canonical forms of the fingerspelled letters (Armstrong, Stokoe, & Wilcox, 1995).

The study of coarticulation and compression processes that occur in rapid fingerspelling has been an area of active research. Battison (Battison, 1978) examined the process of how a fingerspelled word becomes a loan word. From interviews with nineteen prelingually deaf informants, he

identified a total of 40 fingerspelled words which generally became accepted as loan words in ASL. The words ranged in length from two to five letters. He found nine separate categories of potential change. These included 1) deletion of a fingerspelled letter, changes in 2) location, 3) handshape, 4) movement and 5) orientation when comparing the produced fingerspelled letter to the idealized fingerspelled letter, 6) reduplication of a movement, 7) addition of a second hand when producing the loan sign, 8) morphological involvement such as inflection of the loan sign to show the addition of grammatical information and 9) a change in the semantics where the loan sign now has a meaning substantially different from the original fingerspelled word. One of the first changes that typically occur is the deletion of a letter or letters in the fingerspelled word. An example is the deletion of both medial letters from the fingerspelled word B-A-N-K as compared to the loan word #BK (bank).

Jerde (Jerde, Soechting, & Flanders, 2003) studied coarticulation as a question of assimilation or dissimilation between hand configurations in a selected series of fingerspelled letters. From an interpreter service, he recruited four participants fluent in ASL. Each participant donned a Cyberglove before producing a series of 40 fingerspelled sequences. Each sequence was either 1) an English word, 2) a pronounceable non-word or 3) an unpronounceable, non-word. All of the fingerspelled sequences contained either the letter sequence I-S-C or N-T-R. In each case, a vowel followed the three-letter sequence. He then examined the velocity profiles and movement times of individual joints while producing these sequences. He found that in the index and middle fingers, the proximal interphalangeal joints show dissimilation. Jerde posits that the dissimilation may serve to enhance visual discrimination among handshapes of fingerspelled letters.

To search for patterns of coarticulatory processes of anticipation and perseveration in fingerspelling, Channer (Channer, 2013) chose ten words based on their likelihood to exhibit coarticulation and then recorded five hearing ASL signers who each fingerspelled the words. She found that anticipation occurred more often than perseveration. She also found that coarticulation occurred more often in medial letters of a word, and that it occurred quite frequently. In this study, 53 percent of all fingerspelled letters exhibited some type of coarticulation. Deletion was also most prevalent in medial locations, and occurred in five percent of all medially- located fingerspelled letters.

Wager (Wager, 2012)found different rates of occurrence in coarticulation and deletion during a recorded address by a native Deaf signer. Her search for occurrences of careful fingerspelling and rapid fingerspelling resulted in the identification of 45 fingerspelled words. Although the discourse was a public address in a formal register, nearly half of the fingerspelled words displayed characteristics of rapid fingerspelling. Among her measurements was a *coarticulation Index* metric, which consisted of the average number of coarticulatory processes identified per fingerspelled letter. Of the fingerspelled words, 44 percent

exhibited deletion, and approximately 40 percent of all fingerspelled letters exhibited coarticulatory processes.

Thumann (Thumann, 2009) also explored a comparison of careful versus rapid fingerspelling by analyzing a recording of a conversation between two native ASL users discussing the city of Mobile, Alabama. In the conversation, the women fingerspelled the word "Mobile" 23 times. Thumann found occurrences of both deletion and coarticulation, which resulted in shorter durations.

Keane (Keane, 2014)reported on a first step to analyze a newly-established fingerspelling corpus for coarticulation by considering the spread of pinky extension across multiple fingerspelled letters. He found that the spread was more prevalent in rapidly-fingerspelled words. As a result of letter deletion and coarticulation, the letters in a fingerspelled word may appear differently from the idealized form shown in a textbook illustration, and it can be difficult for novice signers to recognize it in the context of a word.

2.3. Technologies for Practicing Fingerspelling Recognition

As mentioned in the previous subsection, one of the challenges to acquiring fingerspelling recognition skills is the lack of opportunities for self-study. When learning a spoken language such as English, students have access to a rich and varied supply of materials for self-study, including newspapers, video recordings, learning software and entire libraries of written material. There are far fewer opportunities for a student wanting to practice fingerspelling recognition.

Previous technologies used for self-study include 1) video recordings of fluent signers, 2) flash card technology and 3) 3D avatar technology. The 1980s marked the appearance of videotaped recordings of fingerspelling produced by fluent ASL signers. In the 1990s, CDs and DVDs designed for fingerspelling practice became available (Jaklic, Vodopivec, & Komac, 1995). These media showed skilled signers demonstrating words in careful and rapid styles of fingerspelling production. Because these are fluent signers, the fingerspelling naturally exhibits both coarticulation and deletion. However, in media of this type, the vocabulary words are fixed at the time of recording. Adding new vocabulary required more recording sessions at an additional cost. Since the videos were recorded at low frame rates, motion blur was a problem, as was a lack of variation in the presentation order. As a student viewed and reviewed the same vocabulary presented in the same order, it was not clear if the student was improving their recognition skills or merely memorizing the recording.

The advent of Internet-based technologies paved the way for browser-based applications, such as (Vicars, 2005), which offer fingerspelling practice. When using one of these applications, a student can view a word as a succession of static snapshots or flash cards, each showing a single letter, see Figure 2. Once the spelling is complete, students can

guess the word and receive feedback. The advantage of these applications is their flexibility. The application can spell any word by simply shuffling the flash cards and can introduce new vocabulary without incurring costs for additional recordings. However, there is a drawback due to the static nature of the snapshots. There is no connective movement between the static images in these practice tools. As previously noted, linguistic research has revealed that the transitions between fingerspelled letters are not only important, but vital to fingerspelling recognition. Students need to view the movement envelope that is intrinsic to fingerspelling.
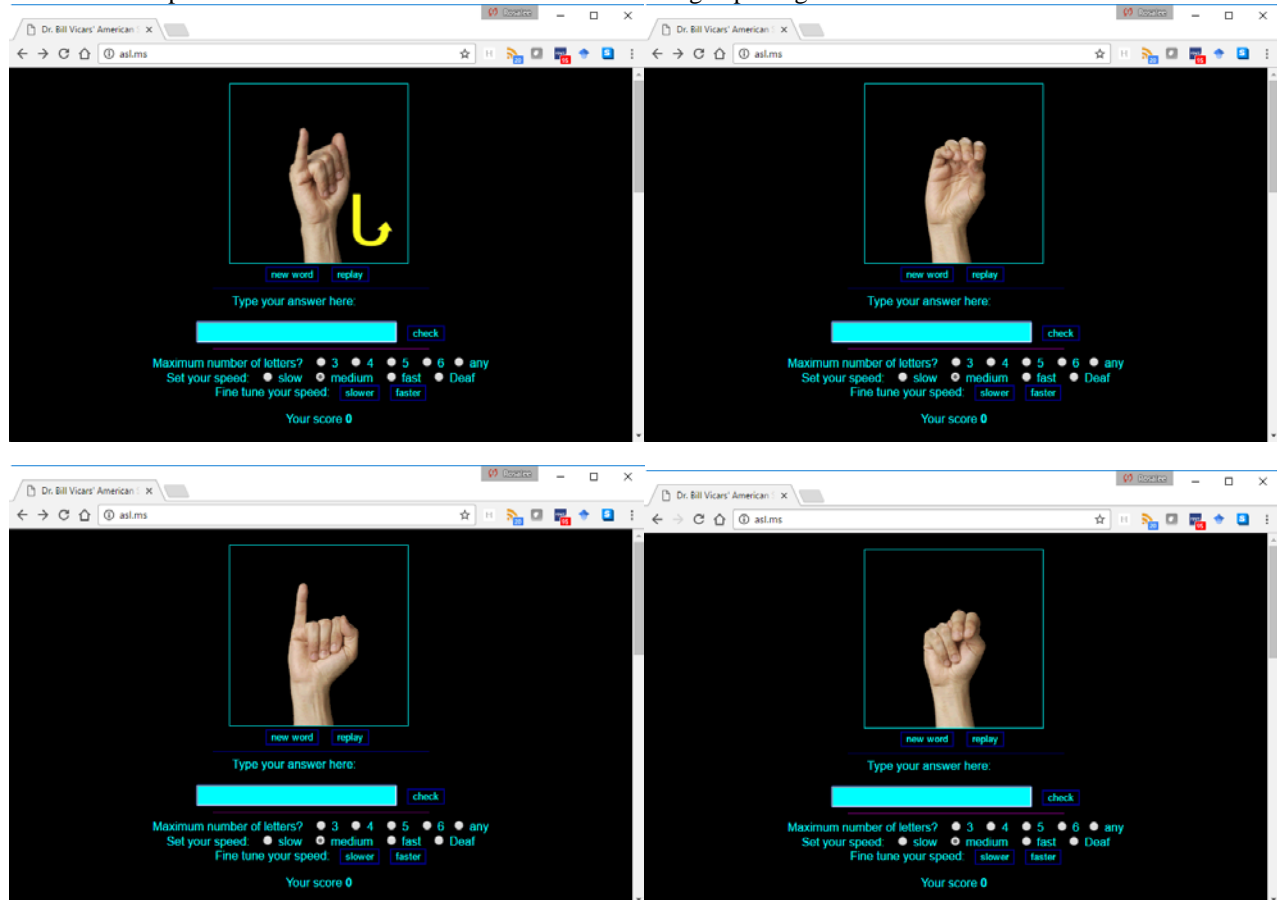

Figure 2. "Flash-card" style of fingerspelling presentation (Vicars, 2005)

A third alternative is 3D avatar technology, which promises extensibility for the addition of new vocabulary words while producing smoothly flowing motion, but it poses some challenges of its own. Fingerspelling puts greater demands on avatar technology than those required for conventional game play. Using a 3D avatar for fingerspelling requires careful attention to simulating the flexible webbing between the thumb and index finger and mimicking the complex behavior of the base of the thumb (McDonald, et al., 2001).

Avatars suffer from a lack of physicality. Unless prevented, the thumb and fingers will pass through each other when transitioning between closed handshapes such as the ASL manual letters M, N, T, S and A. Figure 3 demonstrates the differences between a naive interpolation of the transition from N to A, and a human production of the same transition. In the naive interpolation, the index and middle fingers descend and the thumb cuts through the flesh of the two fingers on its way to the radial side of the hand. In contrast, a human signer will straighten the metacarpophalangeal joints of the index and middle fingers, lifting them upwards momentarily to allow the thumb to pass underneath the fingers unobstructed. Such motions are part of the movement envelope described by Akamatsu (Akamatsu, 1982)and are essential to a realistic display of fingerspelling. To simulate this physically-based transition via an avatar requires the system to prevent finger collisions and faithfully replicate the motion envelope produced by fluent signers.
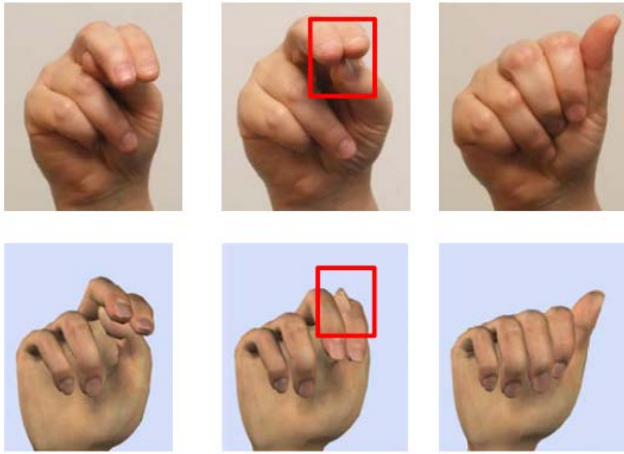
Figure 3. The contrast between human collision avoidance and naive interpolation in an avatar (Wolfe, McDonald, Toro, Baowidan, Moncrief, & Schnepp, 2015)

Accurately portraying this high level of realism in an avatar entails large computational requirements. For this reason, some previous efforts sacrificed realism to gain real-time speeds by using a simplified 3D avatar that did not accurately portray a human hand and/or did not prevent collisions (Su, 1998) (Dickson, 2013). Another early approach to collision avoidance was to move the hand into a neutral position between each letter (Adamo-Villani & Beni, 2004). However, the resulting motion of this approach does not follow the shape of the motion envelope and introduces handshapes not present in fingerspelling produced by fluent signers.

Other efforts (Wolfe, et al., 2006) (Toro, McDonald, & Wolfe, 2014) created a real-time fingerspelling avatar that required only modest computing resources and addressed the collision problem to produce a smoothly-flowing motion envelope. The approach involves a pre-rendering step that carefully organizes the animations into a series of small video clips, each containing a single letter-to-letter transition. These transition clips are very short: if the avatar is spelling at two letters a second, then there are fifteen frames in a transition; if the avatar is spelling at three letters a second, then there are ten frames in a transition. Since each clip had a transition between only two letters, the problem of collisions became more tractable. As part of the pre-rendering step, animation artists reviewed each video clip, and manually added animation keys to remove any collisions. For example, in the N to A transition mentioned earlier, an animator added keys to the index and middle metacarpophalangeal joints to cause the index and middle fingers to rotate up and out of the way of the thumb's path. Figure 4 shows selected frames from the video clip.

3. Results

To build a corpus that would satisfy the requirements of such a study, we first had to choose the medium for the recording. Two primary methods exist for capturing fingerspelling, video and motion capture. Motion capture can be quite invasive, requiring a glove or series of sensors applied to the fingers. We were concerned with capturing natural rapid fingerspelling, and since motion capture gloves or sensors have the potential to radically change or slow a signer's fingerspelling, we chose to work with video for this corpus.

3.1. Developing a Corpus

A primary challenge in building a video corpus for fingerspelling is the incredible rate at which letters are produced, particularly for rapid fingerspelling. For example, if fingerspelling is occurring at a rate of 5 fingerspelled letters per second, and video is recorded at only 30 frames per second (fps), then at most 6 video frames will be dedicated to each transition, and the result can be extremely blurry frames, especially when the fingers are moving quickly. Thus to capture transitions and co-articulators with high fidelity, we used a high-frame-rate video camera capable of 240 frames per second at standard definition 640x480 resolution. This allowed us to record clear frames even when recording the most rapid fingerspelling.

The corpus was designed to capture a range of different fingerspelling phenomena, both in the context of a larger signed discourse and in isolated examples. To accomplish this, we used two separate stimuli

- A script wherein a person is recounting the people invited to a wedding reception. This script made it natural to chain together lists of names with connecting phrases and thoughts. The names were chosen to exercise a range of letter combinations. The full script is included in Appendix A.
- A list of isolated words designed to elicit finger spelled letter combinations which involved transitions between open and closed handshapes. Thus this list contains some "worst-case" situations for extreme finger movement. The list of words is included in Appendix B.

Qualified ASL interpreters were hired to sign the scripts, and each script was recorded with the interpreter asked to sign with two different styles:

- A "teacher" style in which the intended audience had little fingerspelling recognition experience and needed fingerspelling that was as crisp and clear as possible. The result was similar to careful fingerspelling as described by Patrie ,
- A "fluent" style in the manner they would use to sign to a fluent or native signer. This corresponded to rapid fingerspelling.

Each style was recorded with the signer being asked to sign at an appropriate speed and then recorded again signing at a faster speed. Each script was projected as text directly in front of the signer, and recordings were taken from a front camera only.

The captured videos were cut into individual clips, each containing a single fingerspelled word. In the case of the

names from the wedding invitation discourse, transitions into the first letter were included in the clip to give context to the position and orientation of the first fingerspelled letter in the name. The result was a corpus of 527 fingerspellings of 80

unique words recorded in standard definition and at high frame rate, allowing us to clearly see the shape of the hand through the entire motion envelope.

Figure 4.   An animation of the fingerspelled word T-U-N-A, created from three clips depicting letter transitions (Wolfe, et al., 2006)

3.2. Annotating the Corpus

Each of the individual video clips were annotated in ELAN by a student familiar with fingerspelling and were then checked by a faculty member with ten years of experience in ASL handshape animation and fingerspelling. The following ELAN tiers were generated

- **Word**: containing a single annotation for the fingerspelled word spanning the entire motion.
- **Letter**: containing annotations for fully formed letters. The annotations span the length of time that the full handshape is held. This may include some movement in the orientation of the wrist that did not significantly affect the shape of the hand.
- **Coarticulation**: containing annotations of handshapes that have significant modifications, but where some aspect of the handshape was still recognizable. The annotations span the recognizable elements of the handshape. In addition, this tier includes all instances of coarticulation wherein two letters are signed within the same motion. In cases where there was ambiguity as to which tier a letter should be placed, the annotator favored including it in this "Coarticulation" tier.
- **Deletion**: contains annotations of letters that are deleted. The annotation length is not significant as the letter is completely unrecognizable anywhere in the sequence of frames. The annotation is placed between the annotations of surrounding present or coarticulated letters.

Figure 5 displays a screen capture of ELAN demonstrating our annotations for the fingerspelled word V-E-R-O-N-I-C-A. In this example, the letters R, N, C and A were fully formed but the O was deleted between the R and N. In addition, the E was altered so that it only involved the index and middle fingers as they were used subsequently to make the R. Interestingly, this example also contained a leading deletion because the V was subsumed by an initial enumeration sign involving the index finger.

3.3. Using the ELAN n-Gram Analytic Tool for Fingerspelling Analysis

In looking to improve the current fingerspelling display technology, we wished to determine the nature of the relationship between the speed of fingerspelling and the occurrences of coarticulation and deletion. In particular, we wished to determine how coarticulation and deletion could be incorporated into the tool's fingerspelling as the rate of fingerspelling increased. This would allow the tool to more faithfully reproduce rapid fingerspelling for more advanced students. To study this, we used the ELAN n-gram analytic tool (Berke, 2013) to analyze the following statistical patterns:

- Fingerspellings of words that involve any coarticulated or deleted letters,
- 3-grams that have either a coarticulation or a deletion as the second letter of the 3-gram,
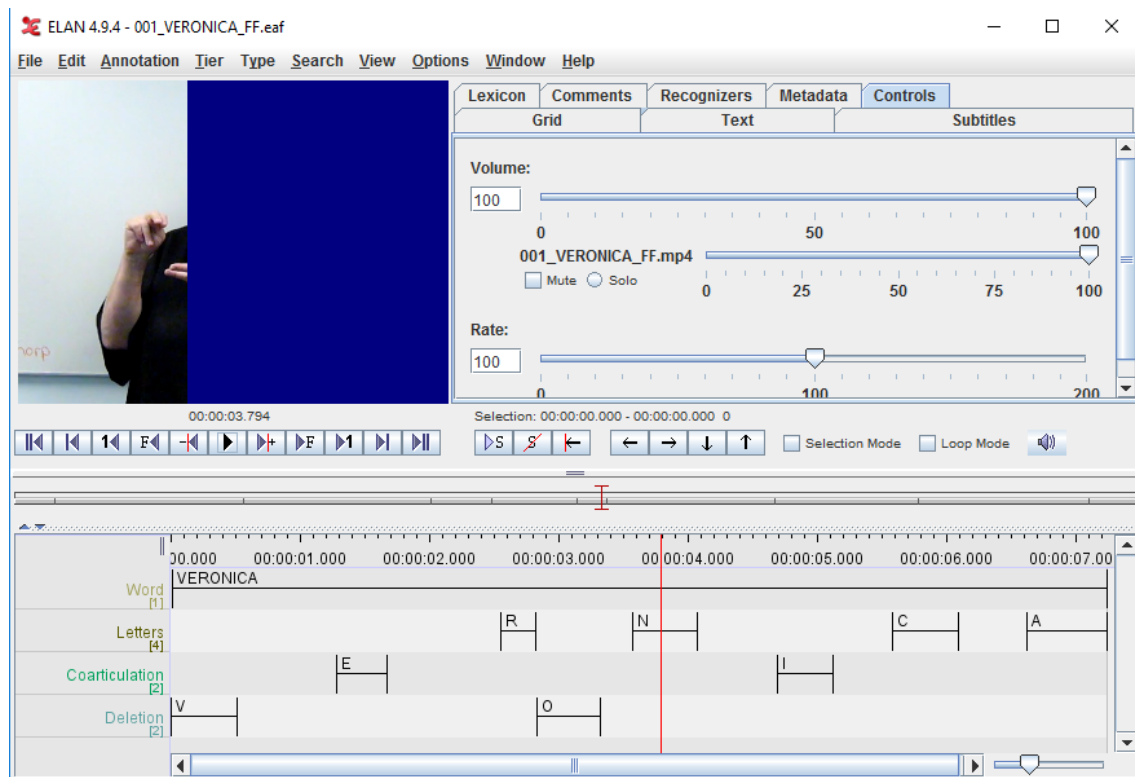- Letters that were most often coarticulated or deleted

Figure 5. Example of annotations in ELAN for the fingerspelled name "Veronica". The video portion is anonymized for publication. The timeline shows the entire word, and the current time is positioned within the letter 'N'.

For brevity, we will call any coarticulation or deletion generally a *modification* of a letter. Our first analysis was to get a bird's-eye-view of letter modification by looking at the overall frequency of occurrence in words. The results were consistent with past studies of fingerspelling coarticulation. Among the 527 fingerspellings of these words, coarticulation or deletion happened in 64% of the cases. Breaking this out between coarticulation and deletion yields the following rates over all speeds for both the careful and rapid fingerspelling:

- Coarticulation occurred in 39% of fingerspelled words.
- Deletion occurred in 12% of fingerspelled words.

To gauge the relationship between speed and coarticulation, we analyzed the full set 1390 individual 3-grams occurring in the corpus' 527 words. For example, in a fingerspelling of V-E-R-O-N-I-C-A, there would be the following six 3-grams

VER, ERO, RON, ONI, NIC, ICA

To avoid duplications, we chose to look only at coarticulation or deletion that occurred on the middle letter of each 3-gram. This necessarily excluded the initial and final letters of each word from the analysis, where coarticulation and deletion were both expected to be relatively rare. Thus the study only considered coarticulation and deletion of medial letters. Overall coarticulation and deletion of medial letters occurred in 21% of the 3-grams.

As a measure of the fingerspelling speed, we chose to look at the overall duration of each 3-gram, which allows for varying speed during a fingerspelling production. While certain letters do take more time to produce, J and Z for example, in the presence of two other letters in the 3-gram, these differences averages out over the duration of the 3-gram. A histogram of modification to 3-grams indexed by duration is shown in Figure 6, and reveals that modifications happen far more often with more rapid 3-grams.

To further quantify this, Figure 7 displays the percentage of 3-grams containing letter modifications by duration. The graph shows a clear decreasing linear relationship between the log-percentage of letters modified vs duration.

Running a regression analysis on this relationship yielded the following results:

Table 1. Regression results for modification percentage

Intercept	-.026
Slope	-2.19
Residual Standard Error	.166 (9 d.o.f)
R^2	.95

The computed slope is statistically significant ($p < .01$) and the overall F-test of the regression shows significance ($p < .01$). The resulting regression line is displayed in red in Figure 6.
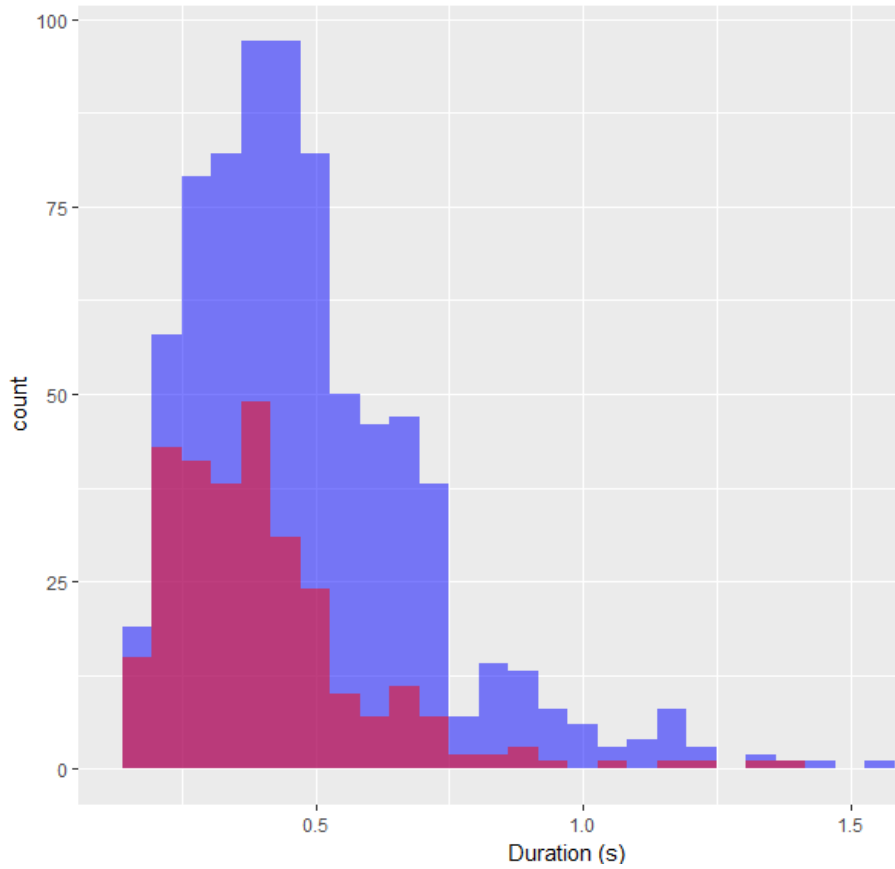
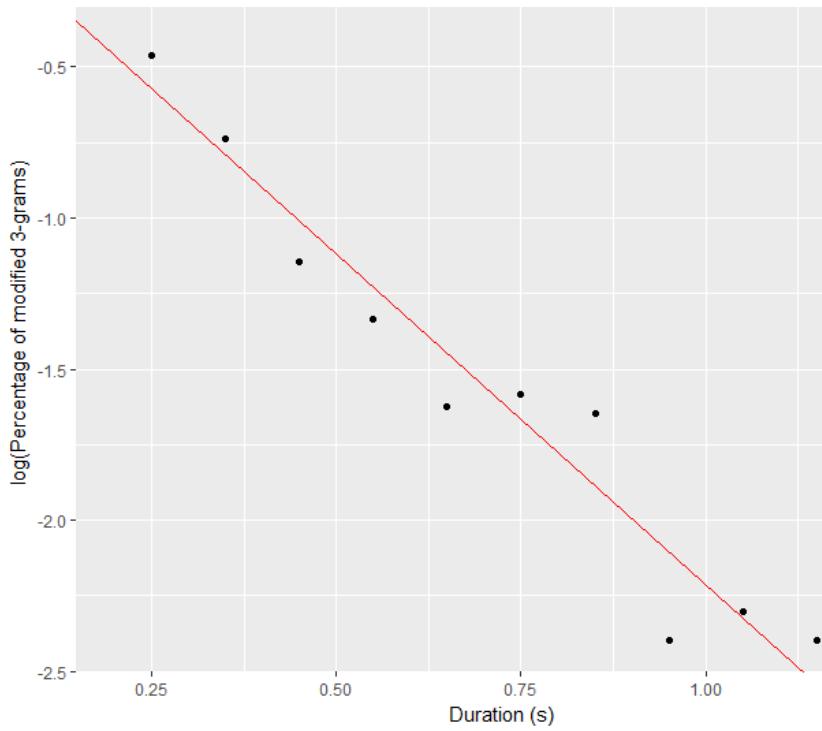**Figure 6**.   Histogram of 3-gram durations and modifications

Figure 7.   Relationship between percentage of modified letters and 3-gram duration
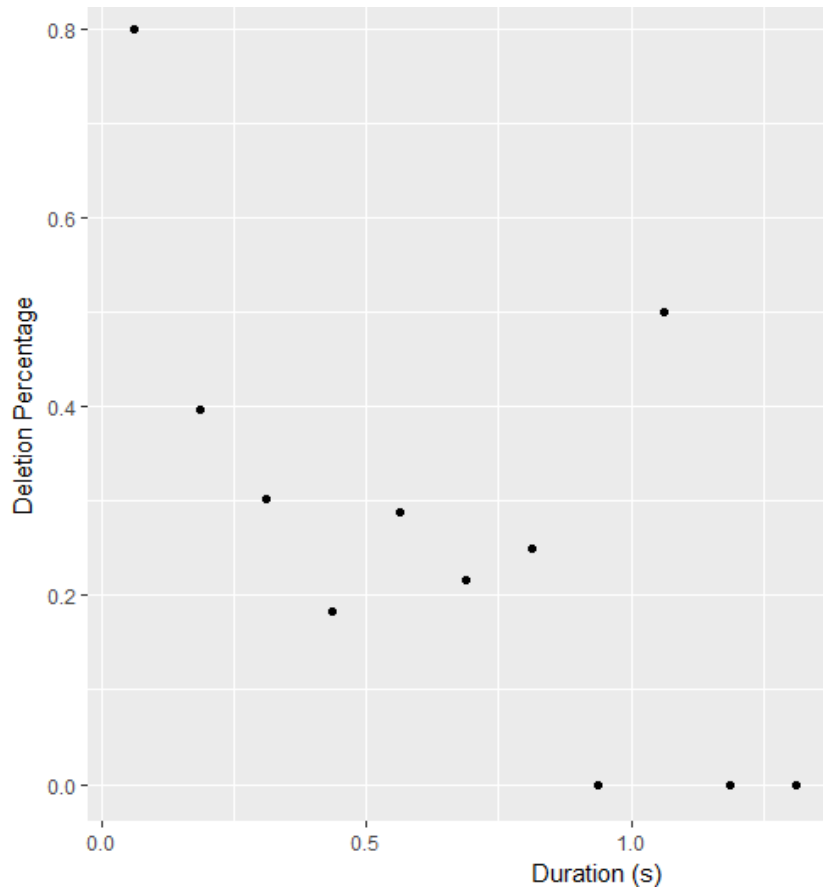
Figure 8. Deletion as percentage of total modifications vs duration

The resulting equation for probability based on 3-gram duration yields the following approximation for the percentage of letter modifications

$$p = .974e^{-2.19s} \qquad (1)$$

This indicates that at a duration of zero seconds, the percentage of modifications is essentially 100%, and each 0.1 second increase in duration decreases the percentage of modifications by a factor of 80%.

The graph in Figure 8 examines this relationship in more detail. Deletions are displayed as a percentage of total modifications indexed by duration. It shows that the deletions form an increasing percentage of the modifications at shorter durations.

Since these findings consist of a percentage or a probability that letters will be deleted or coarticulated, we can sharpen this analysis by looking at the overall frequencies of letter modifications to see which letters are more likely to be modified. Figure 9 shows the most frequently coarticulated letters whereas Figure 10 shows those that are most often deleted. These results will inform our proposed changes to the Fingerspelling technology.
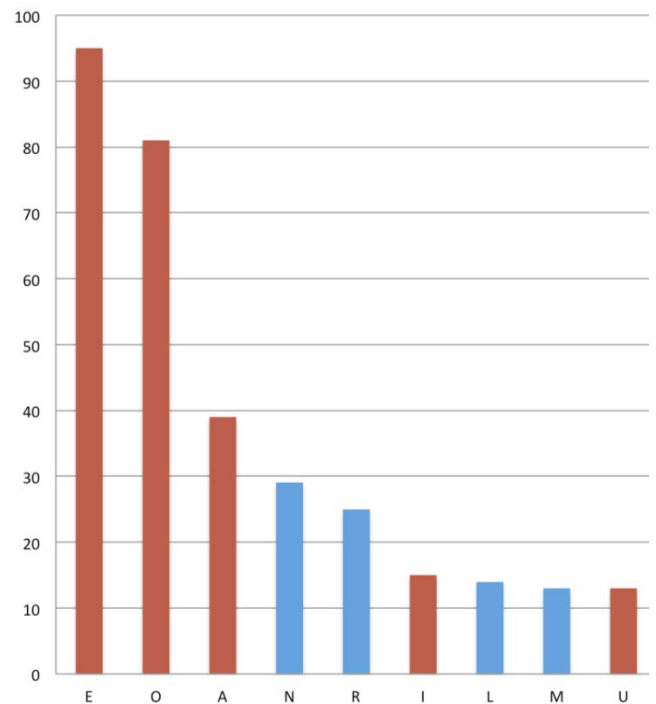

Figure 9. Coarticulated letter frequencies

4. Discussion and Conclusions

The results of the last section suggest the following modifications to the current fingerspelling technology:

- Rather than simply increasing the rate of playback, as the speed of fingerspelling increases, we should modify or delete letters. Further, as the rate of fingerspelling increases, coarticulations will give way to deletions.
- The first candidate letters for modification should be the vowels E, O, I and A, followed by such handshapes as M, N, L and R.
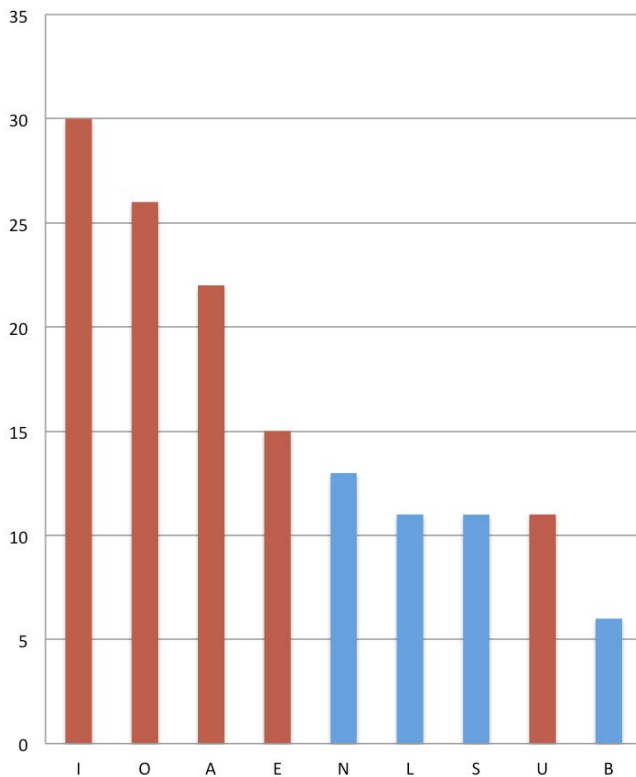

Figure 10.   Deleted letter frequencies

These findings suggest two modifications to the current fingerspelling technology to produce more a realistic rapid fingerspelling style than would be possible by simply increasing the frame rate. The first is to introduce deletions in medial letter positions. The second is to introduce coarticulation by careful editing of the transition clips before assembling and displaying the fingerspelled word.

Introducing deletions is simply implemented via a preprocessing step where the word to be spelled is edited to remove the letters affected by deletion. The first choices for deletion are medial vowels, followed by a preference for the medial letters M, N, L and R. For example, a deletion applied to the word M-A-R-Y would result in M-R-Y, which would utilize only the transitions M to R, and R to Y.

Introducing coarticulation is more involved and requires editing of the individual transition clips. As demonstrated in Figure 3, the initial frame of a transition clip depicts a letter in its canonical form. Instead of using the entire transition clip, a new software module shortens the clip to start with frame 2 or frame 3, depending on the speed of the fingerspelling. The letter undergoes coarticulatory effects, and the resulting animation is smoother and more accurately simulates the motion envelope described by Akamatsu (Akamatsu, 1982). The link

http://asl.cs.depaul.edu/video/LinguisticsAndLiteratureStudies/comparison.mp4

shows two versions of the word V-E-R-O-N-I-C-A. The version on the right demonstrates the effect of deletion and coarticulation, and more accurately simulates the motion envelope.

Future plans include additional analysis of the coarticulatory processes in rapid fingerspelling to make further modifications to the portrayal of automatic fingerspelling generation. As a future expansion, we envision having a native signer or qualified interpreter annotate the fingerspelling corpus with further detail for the types of modifications that the handshapes undergo. We further envision adding annotations for orientation changes.

Appendix A: Wedding Invitation Script

We're planning a family and friends reunion for our wedding anniversary celebration and need to make sure that we've invited everyone. To start there is the immediate family starting with my parents
  Priscilla & John
Then there are my wife's parents
  Doris and Peter
There are my two sisters
  Kate and Joanne
And my brothers
  Ian and Bob
Then there are my wife's siblings
  Jordan, Sue, Norma and Lori
Oh, I can't forget my
  Uncle Toni & Aunt Irma and also my uncle Quin and Aunt Harriet
Then there are my wife's uncles and aunts
  Xavier, Martha, Ron and Fiona
And her cousins
  Bob, Tammie, Walter and Zoie
That about does it for the family ... oh wait, I forgot about my cousins
  Cindy, Veronica and George
Ok, now for our friends. First there are our closest friends
  Dianna and Brandon
And also
  Natalie and Ian
Then there are my friends from work
  Lelia, Wilbur, Theo and Karen
There are my wife's colleagues
  Brenda, Violet and Ralph

Appendix B: Isolated Word Script

RACE	MACE	MAIN
MILE	MALE	MULE
BUS	MIDI	MINI
PROVIDE	SALE	SOLE
JOINT	JOAN	JOHN
JOB	AJOURN	OINTMENT
AMEND	MEND	MINT
MONEY	DIME	MISSION
VENT	OVEN	OVER
LICE	ACTION	MAIN
BILL	VETO	VICE
TAME	TEAM	TIME
TOMB	COVER	TUMBLE
MANSION	FILL	MILL
DIANA	SYLVIE	MIKE
NILES	TIMOTHY	TYLER
TYSON	FINN	DYNA
BILL	WILLIAM	CIARAN
KITTY	SYBILE	KYLE
KYANA	OFER	IAGO
OYA	FABIAN	IDELLE
YANN	JERED	JIMMY
DINNER	GIFT	SILLY
INNER	NITPICK	KITE
OFTEN	IFFY	MAYA
NICE	LICE	LYRE
BOB	FANNY	PETER
MORRIS	NOLAN	TILL
WALLACE	GROMMIT	SUSAN
MOLLY	NELLY	WALTER

REFERENCES

[1] Max Planck Institute for Psycholinguistics. (2013, September) The Language Archive. [Online]. http://tla.mpi.nl/tools/tla-tools/elan/

[2] Hennie Brugman and Albert Russel, "Annotating Multi-media/Multi-modal Resources with ELAN," in Proceedings of the 2nd Workshop on the Representation and Processing of Sign languages: Lexicographic matters and didactic scenarios, Paris, 2004, pp. 82-87.

[3] Onno Crasborn, Han Sloetjes, Eric Auer, and Peter Wittenburg, "Combining video and numeric data in the analysis of sign languages within the ELAN annotation software," in Proc. LREC 2006 workshop on representation & processing of sign language, Paris, 2006, pp. 82-87.

[4] Peter Wittenburg, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes, "Elan: a professional framework for multimodality research," in Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Paris, 2006, pp. 1556-1559.

[5] Philippe Dreuw and Hermann Ney, "Towards automatic sign language annotation for the ELAN tool," in LREC Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora, Paris, 2008.

[6] Eric Auer et al., "ELAN as Flexible Annotation Framework for Sound and Image Processing Detectors," in Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010), Paris, 2010, pp. 890-893.

[7] Han Sloetjes and Peter Wittenburg, "Annotation by category - Elan and ISO DCR," in Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008), Paris, 2008.

[8] Herman Stehouwer and Eric Auer, "Unlocking Language Archives Using Search," in Proceedings of the Workshop on Language Technologies for Digital Humanities and Cultural Heritage, 2011, pp. 19-26.

[9] Onno Crasborn, Micha Hulsbosch, Lari Lampen, and Han Sloetjes, "New multilayer concordance functions in ELAN and TROVA," in Proceedings of the Tilburg Gesture Research Meeting [TiGeR 2013], 2013.

[10] R. Sutton-Spence and B. Woll, "Fingerspelling as a resource for lexical innovation in British sign language," Sociolinguistics, language and society. Language and development, vol. 5, pp. 10-143., 1998.

[11] Robbin M. Battison, Lexical Borrowing in American Sign Language. Silver Spring, MD: Linstok Press, 1978.

[12] Clayton Valli and Ciel Lucas, Linguistics of American sign Language: An Introduction. Washington, DC: Gallaudet University Press, 1995.

[13] Michael Strong, Ed., Language Learning and Deafness. Cambridge: Cambridge University Press, 1988.

[14] Carol Padden, "The acquisition of fingerspelling by Deaf children," in Theoretical Issues in Sign Language Research, P Siple and S. Fischer, Eds. Chicago: University of Chicago, 1991, pp. 191-210.

[15] D. A. Grushkin, "Lexidactylophobia: The irrational fear of fingerspelling," American Annals of the Deaf, pp. 404-415, 1998.

[16] Carol Patrie and Marty Taylor. (2007) Outcomes for graduates of baccalaureate interpreter preparation programs specializing in interpreting in K-12th grade settings. [Online]. http://scholarworks.rit.edu/cgi/viewcontent.cgi?article=1067 &context=books

[17] Carol Patrie and Robert Johnson, Fingerspelled World Recognition through Rapid Serial Visual Presentation. San Diego, CA: Dawn Sign Press, 2011.

[18] Richard D. Zakia and Ralph Norman Haber, "Sequential letter and word recognition in deaf and hearing subjects," Attention, Perception & Psychophysics, vol. 9, no. 1, pp. 110-114, 1971.

[19] Carol Tane Akamatsu, "The acquisition of fingerspelling in pre-school children," Ann Arbor, 1982.

[20] Sherman Wilcox, The Phonetics of Fingerspelling. Amsterdam: John Benjamins Publishing, 1992.

[21] David F. Armstrong, William C. Stokoe, and Sherman E.

Wilcox, Gesture and the nature of language. Cambridge: Cambridge University Press, 1995.

[22] Thomas Jerde, John Soechting, and Martha Flanders, "Coarticulation in fluent fingerspelling," The Journal of Neuroscience, vol. 23, no. 6, pp. 2383-2393, March 2003. [Online]. http://www.jneurosci.org/content/23/6/2383.full.pdf+html

[23] Caitlin S. Channer, "Coarticulation in American Sign Language Fingerspelling," 2013.

[24] Deborah Stocks Wager, "Fingerspelling in American Sign Language: A case study of styles and reduction," 2012.

[25] Mary Thumann, "Fingerspelling in a word," Journal of Interpretation, pp. 104-126, 2009.

[26] Jonathan Keane, "Towards an articulatory model of handsape What fingerspelling tells us about the phonetics and phonology of handshape in American Sign Language," 2014.

[27] A Jaklic, D. Vodopivec, and V. Komac, "Learning sign language through multimedia," in International Conference on Multimedia Computing and Systems, Washington, DC, 1995, pp. 282-285.

[28] Bill Vicars. (2005) Dr. Bill Vicars' American Sign Language (ASL) Fingerspelling Practice Site. [Online]. http://asl.ms

[29] John McDonald et al., "An improved articulated model of the human hand," The Visual Computer, vol. 17, no. 3, pp. 158-166, May 2001.

[30] Rosalee Wolfe et al., "Promoting Better Deaf/Hearing Communication Through an Improved Interaction Design for Fingerspelling Practice," International Conference on Universal Access in Human-Computer Interaction, pp. 495-505, 2015.

[31] Augustine Su, "VRML-based Representations of ASL fingerspelling on the World-Wide Web," in The Third International ACM SIGCAPH Conference on Assistive Technologies, Marina del Rey, CA, 1998. [Online]. http://www.csdl.tamu.edu/~su/publications/assets98_.html

[32] Stewart Dickson, "Advanced Animation in Mathematica," The Mathematica Journal, vol. 15, no. 2, 2013.

[33] Nicoletta Adamo-Villani and Gerardo Beni, "Automated finger spelling by highly realistic 3D animation," British Journal of Educational Technology, vol. 35, no. 3, pp. 345-362, 2004.

[34] Rosalee Wolfe et al., "An Improved Tool for Fingerspelling Recognition," in Technology and Persons with Disabilities Conference, Los Angeles, 2006.

[35] Jorge Andres Toro, John C. McDonald, and Rosalee Wolfe, "Fostering better Deaf/hearing communication through a novel mobile app for fingerspelling," 2014, pp. 559-564.

[36] Larwan Berke, "Empowering ELAN with N-gram Analytics for Corpora," School of Computing, DePaul University, Chicago, 2013. [Online]. http://asl.cs.depaul.edu/software/nGramELANtechReport.pdf