Modeling and Predicting the Location of Pauses for the Generation of Animations of American Sign Language

Sedeeq Al-khazraji, Sushant Kafle, Matt Huenerfauth
Rochester Institute of Technology, Golisano College of Computing and Information Sciences
152 Lomb Memorial Drive, Rochester, NY 14623, 1-585-475-2459
{sha6709, sxk5664, matt.huenerfauth}@rit.edu

Abstract

Adding American Sign Language (ASL) animation to websites can improve information access for people who are deaf with low levels of English literacy. Given a script representing the sequence of ASL signs, we must generate an animation, but a challenge is selecting accurate speed and timing for the resulting animation. In this work, we analyzed motion-capture data recorded from human ASL signers to model the realistic timing of ASL movements, with a focus on where to insert prosodic breaks (pauses), based on the sentence syntax and other features. Our methodology includes extracting data from a pre-existing ASL corpus at our lab, selecting suitable features, and building machine learning models to predict where to insert pauses. We evaluated our model using cross-validation and compared various subsets of features. Our model had 80% accuracy at predicting pause locations, out-performing a baseline model on this task.

Keywords: American Sign Language, Animation Synthesis, Prosodic Breaks, Pauses, Modeling

1. Introduction

American Sign Language (ASL) is used as a primary means of communication for about one half million people (Mitchell et al., 2006). ASL is a natural language that consists of movements of the hands, head, body, and face to convey meaning, and it has its a syntax, word order, and lexicon which are distinct from spoken English. While there is great diversity in the English reading literacy skills among members of the Deaf community in the U.S., including some individuals with strong skills, there are also many individuals with lower English reading skills, due to reduced levels of language-exposure during childhood or other educational circumstances. In fact, standardized testing has shown that the majority of Deaf high school graduates in the U.S. (students who are completing secondary school, typically age 18) have English reading skills at the "fourth-grade" level (age 9 students in the U.S.) (Traxler, 2000). Because of the linguistic differences between ASL and English, there are people fluent in ASL but with difficulties reading English text.

Amid these literacy issues, some English text on websites may be too difficult to read. While adding videos of a human signer to websites may sound like a simple solution, this is impractical: Online information is often updated or generated automatically based on a query. A video would need to be recorded and uploaded, which would be costly and time-consuming. Professional animators can produce realistic animations of virtual humans, but the process is also slow. For these reasons, many researchers, e.g. (Adamo-Villani and Wilbur, 2015; Cox et al., 2002; Ebling and Glauert, 2016; Huenerfauth, 2004; Jennings et al, 2010; Kacorri, 2016; Kennaway et al., 2007; Lu, 2014; McDonald et al. 2016; Segout and Braffort, 2009), investigate the development of software that can generate understandable ASL animations of a virtual human signer automatically from an easy-to-update script. The challenge is that this software must configure the animation so that the movements are accurate and easily understood by ASL signers (Huenerfauth, 2008; Huenerfauth and Lu, 2010).

The primary focus of this paper is to investigate using motion-capture data that our lab has previously recorded from human signers to build predictive models for inserting pauses in ASL animations. Prior studies at our lab (Huenerfauth, 2009) have shown that adding linguistically motivated pauses and adjusting the duration of signs enhances the understandability of ASL animation (as measured on a comprehension task). Thus, our goal is to automate this aspect of animation synthesis and to create understandable ASL animation with better quality.

2. Literature Review

2.1 Linguistic Research on Pausing

Some prior psycholinguistic studies have focused on the timing and pausing in ASL and spoken English (Grosjean and Lane, 1977; Grosjean, 1977; Grosjean et al., 1979). For example, Grosjean et al., (1979) investigated the pause length and location of pauses in ASL, based on the sentence structure. Others studied the sign duration and sign speed, based on video observation, and they analyzed speaker and signer performances at different rates (Grosjean, 1979). Grosjean and Lane (1977) found that for spoken English, longer pauses take place at sentence boundaries (pause length longer than 445 ms); shorter pauses take place between noun phrases, verb phrases and conjoined sentences (pause length range between 245 and 445 ms), and the shortest pauses occurred within phrasal constituents (pause length less than 245 ms). For ASL, Grosjean and Lane (1977) analyzed videos to estimate an average pause length: between sentences (229 ms), between conjoined sentences (134 ms), pauses between noun or verb phrases (106 ms), within verb phrases (11 ms), and within noun phrases (6 ms). These findings suggested that pause length in ASL was related to the syntax structure of the sentence, and these findings have inspired the selection of features for our models, as discussed in section 4.2.

2.2 Rule-Based Pausing for Signing Animation

In some prior sign language animation systems, the speed and duration of signing are invariant or must be specified by a human authoring the message: The eSign project developed an animated avatar that performed sequences of signs from a lexicon. To specify the duration of each word, the authors examined the speed of human signers in videos performing each sign (Kennaway et al., 2007), but they did not vary the duration of signs based on where they appeared in sentences. Sign Smith Studio was a commercial product for ASL animation generation; it allowed the user to modify the temporal parameters of the signs in a sentence manually (by adjusting numerical values for the transition time between words, the hold/pause at the end of words, and a multiplier factor for hand speed during the sign). Therefore, a Sign Smith Studio user needed some skill in computer animation and instincts about how to set numerical values for ASL timing (Vcom3D, 2017).

Many projects have implemented sets of rules that govern the speed and timing of sign-language animations, e.g.:

- Huenerfauth (2009) built a model for the duration (length) of signs, the location of pauses, and the length of pauses in ASL. However, his model was based on rules he authored based on some published data in the psycholinguistics literature on ASL (summarized above in 2.1). He designed two algorithms: for calculating sign duration time and for calculating pause location and length. His sign-duration algorithm depended on whether specific signs had previously appeared in a passage and whether they were at the end of clauses, e.g. noun signs located at boundaries (sentence or clause) were lengthened in duration by a set percentage (12% and 8% respectively). For verb signs, subsequent occurrences were shortened in duration by 12%. The values used in these rules were based on averages reported in the linguistics literature, not on any data-driven machine-learning method.
- A more recent study by Villani and Wilbur (2015) also utilized a rule-based approach. Their system predicted how to add prosodic enhancements to ASL animations, including insertion of pauses and phrase-final lengthening of sign duration. To determine Pausing, Villani and Wilbur adopted linguistic values from (Pfau et al., 2005) to insert pauses between and within sentences. Regarding Phrase Final Lengthening, they increased the length of the last sign of a phrase based on a prior study by Wilbur (2009). Their initial user evaluation showed promising results from using this algorithmic approach to add prosodic features.
- Ebling and Glauert (2016) built a system for translating train announcements from German text to Swiss German Sign Language using the JASigning animation platform. The authors wrote a rule to insert a short pause after each item in lists, based on a suggestion from deaf users who viewed their system's animation output; however, they did not provide a general rule for when pauses should be inserted nor what the pause duration should be, in novel contexts.

2.3 Data-Driven Sign Language Research

While many advances in computational linguistics have come from data-driven methods based on machine-learning models, most prior work on sign language has been rule-based, because of the small quantity of training data, e.g. available audio/video recordings that have been linguistically-annotated. As additional signing corpora have recently become available, there has been a recent trend among sign-language researchers of applying data-driven approaches, as discussed in (Huenerfauth, 2014). For instance, various researchers have examined data-driven methods for sign language translation research:

- Bungeroth et al. (2006) created a corpus for German Sign Language and studied machine-translation and facial-expression issues, but not speed or timing.
- Morrissey and Way (2005) investigated example-based machine translation approaches for producing sign language from English text, using a corpus they annotated with manual and non-manual features. They generated word sequences for sign language, not any animation output, which would have required speed or timing information (Morrissey and Way, 2005). Most of these prior studies made use of small corpora containing texts on a special topic/domain, and none of them explicitly modeled speed and pausing of signs.
- Naert et al (2017) investigated automatically adjusting manual segmentation of sign language motion data.

Some researchers have used data-recordings from humans to generate animation output, for example:

- Segouat and Braffort (2009) used rotoscopy to create a French Sign Language corpus and built an animation system that combined different elements of human motion to create novel sentences. While they studied co-articulation (how the movements at the end of one sign are influenced by the beginning of the next), they did not model speed or timing issues directly.
- Cox et al. (2002) built and evaluated a system called "TESSA" for converting English speech to British Sign Language (BSL) animations, using some template-like phrases to build a limited set of sign language sentences. Since their system filled words into templates (rather than synthesizing complete phrases), they did not address timing and pausing issues, which is the focus of our current work.
- In prior work, our lab has used our ASL Motion-Capture Corpus (Lu and Huenerfauth, 2012; 2014) to investigate different aspects of ASL animation: inflecting verb movement (Lu, 2014), facial expression (Kacorri, 2016), and spatial reference point locations (Gohel, 2016).

3. Research Question

While there has been prior translation and animation-synthesis research that has utilized data-driven techniques, as described above, there has not been prior work that has utilized motion-capture corpora of ASL to directly train machine-learning models of speed, timing, or pause-insertion. Given the success of these prior projects (focusing on other aspects of animation) at using motion-capture recordings to build models of how human ASL signers behave, we therefore intend to use a similar method to investigate the following research question:

Research Question: Can we accurately predict where human signers insert pauses in their ASL signing, as evaluated via cross-validation on an annotated corpus of human ASL signing?

Table 1: Detailed information about selected features

Feature name	Explanation	Type
Sentence Boundary (SB), Clause Boundary (CB), Noun Phrase Boundary (NPB), Verb Phrase Boundary (VPB)	Is this inter-sign gap at the boundary of a sentence, clause, noun phrase or a verb phrase?	Categorical: {Yes, No}
Relative Proximity (RP)	How far is this inter-sign gap from midpoint of the current sentence? A detailed formula for calculating this value appears in Huenerfauth (2009).	Numerical
Complexity Index (CI)	The number of syntactic nodes that dominated this inter-sign gap. A detailed formula for calculating this value appears in Huenerfauth (2009).	Numerical
Sentence Length (SL), Noun Phrase Length (NPL), Verb Phrase Length (VPL)	Number of words in the current sentence, the current noun phrase (if applicable), or the current verb phrase (if applicable) at this inter-sign gap position in the corpus.	Numerical

4. Methodology

4.1 Data Preparation

To support our modeling work, we first needed to process and extract relevant information from our existing ASL Motion-Capture Corpus (Lu and Huenerfauth, 2014), which contains: *motion-capture movement data* available as .bvh files (Biovision hierarchical data, XML files representing human joint angles from a movement recording) and *linguistic annotation* (text files exported from the SignStream annotation tool (Neidle, 2002). A team of annotators that included Deaf native ASL signers and linguists, labeled the glosses and syntactic constituents (including sentence, clause, verb phrase, and noun phrase boundaries) in the ASL video and motion-capture corpus, using a process whereby two annotators independently annotated each file and met to discuss their annotations to arrive at a consensus annotation. The annotation included word labels, clause boundaries, and other syntactic information (Lu and Huenerfauth, 2014).

To process this data for our analysis, we wrote Python code to extract timing information and a subset of linguistic annotation properties to produce a comma-separated values (CSV) file with each row representing an inter-word "gap" location, after each of the 7138 words in the corpus, where a prosodic pause could potentially occur. One column was a "target" label that indicated whether this gap location in the corpus was where the human performed a "pause."

Since the original linguistic annotators did not specifically label which inter-sign gaps contained a prosodic pause and which did not, we needed to fill this value automatically: by identifying a threshold time duration to distinguish between regular end-of-sign "holds" (some signs end with the hands remaining in position for a moment) and longer prosodic-break "pauses" during signing. To calculate this threshold, we calculated mean hold time at the end of words, and we subtracted this value from the period of time when the hands were motionless at the end of words. After ranking these durations, we labeled the longest 25% as "pauses" and the remainder as "not pauses," following the typical ASL ratio of prosodic pauses in (Grosjean, 1979).

4.2 Feature Engineering

The remaining columns contain "predictor features," i.e. properties about this gap location (e.g., is this inter-sign gap a boundary between two sentences, what is the length of the current sentence, etc.) that may be relevant to predicting pauses. These features were calculated automatically from annotation present in the original corpus. In summary, our training data set consisted of nine predictor columns and one target column; there were 7138 rows representing gap locations after each ASL sign. Table 1 lists the predictor features implemented in this work and explains their meaning. The various boundary, relative proximity, and complexity index features listed in the table were included based on their use in determining pauses in prior linguistic work (Huenerfauth, 2009; Grosjean et al., 1979). As mentioned in Table 2, detailed formulas for some of these features are described in Huenerfauth (2009), as they were a key part of that prior rule-based model. All of the numerical features were scaled using unity-based normalization with the training minimum and maximum.

The reader may note that none of the features included in our model were lexically specific, i.e. they did not depend on the specific gloss/word labels for the individual signs that preceded or followed any inter-sign gap. This decision to avoid lexically-specific features was intentional, given the relatively small size of our training corpus. Furthermore, a small set of prompts had been used in the collection of this corpus (Huenerfauth and Lu, 2014); thus, we sought to avoid training a model of pause insertion that would be overly domain specific, given our limited data.

4.3 Selecting the Classification Models

Since our goal was to fit and test a model to predict pause locations in ASL animation and our target variable had values of ("there is a pause here" or "there is not a pause here"), we considered a traditional supervised classification approach to make an individualized prediction for the gap following each word in a sentence. Since we had both categorical and numerical predictor features (see the "Type" column in Table 1), we chose to investigate and compare several machine-learning algorithms that support mixed features, including: decision trees, support vector machines (SVM). In particular, we noted that prior work on pause prediction for English (Sarkar and Rao, 2015) or other modeling for ASL (Shibata et al., 2016) had successfully used decision-tree-based learning methods.

To select the optimal subset of features to use when building our model, we implemented code to exhaustively build and test versions of each model using all possible combinations of our predictor features, for a total of 511 different feature subsets. We trained a decision tree classifier (using a maximum of 100 branch nodes) and an SVM Linear classifier in MATLAB (MathWorks, 2017).

Aside from making independent predictions of the target variable ("pause" or "no pause") for each inter-sign gap location, we also investigated if there were dependencies between the values at subsequent gap locations. Specifically, we considered making predictions based on a +/-1 context window (i.e. the predictor features of the inter-sign gap immediately preceding and following the current inter-sign gap), thereby treating the problem as a sequence-tagging problem. For this purpose, we trained a Linear-Chain Conditional Random Field (CRF) model which operated on the context-features and looked for the most optimal path through all possible target sequences for a sequence of words in a sentence.

4.4 Cross-Validation Training and Evaluation

For the classifiers described above, we implemented a 5-fold cross-validation procedure, dividing our data into 80% training set and 20% testing set at each evaluation fold. We calculated the average accuracy and f-score across the 5 folds. To select the best working parameters for each of our models, we performed a grid-search to optimize the model performance. We compared our result with some baselines:

- **Baseline 1:** We inserted a pause at the end of every sentence (and nowhere else). The rationale is that if a human were to create an animation and manually chose to insert some pauses, the animator may likely put them at all of the sentence boundaries, as a simple approach.
- **Baseline 2:** We inserted a pause randomly at 25% of locations. To account for variation due to randomness, we ran it ten times (Table 2 presents the average).

5. Results Analysis

Table 2 shows the accuracy and f-score for each model with best the performing feature combinations. As shown in the table, Baseline 1 (which inserted a pause at all sentence boundaries) has good performance – which is expected as many pauses do occur at the end of sentences.

Table 2: Results of Each Pause-Prediction Classifier

Classifier	Accuracy	F-Score	Features
Linear-Chain CRF	**0.80**	0: 0.298 1: **0.880**	ALL
Decision Tree	0.76	0: 0.226 1: 0.858	CB, VPB
SVM (Linear)	0.76	0: 0.160 1: 0.868	CB, VPB
Baseline 1	0.77	0: **0.392** 1: 0.860	SB
Baseline 2	0.64	0: 0.227 1: 0.768	N/A

The SVM and Decision-tree models, which utilized features from the current inter-sign gap only, struggled to beat this baseline, in both accuracy and f-score. The linear-chain CRF model was our top performing model, with an accuracy of 80% and F-score comparable in performance to (and slightly exceeding) the Baseline 1.

6. Conclusions

In this work, we demonstrated our methodology for building models of one aspect of ASL animation timing, based on machine-learning modeling of a collection of motion-capture data. Specifically, our work has focused on building models of where people pause during signing, and we have successfully identified a set of features and a modeling approach that outperforms a commonly-used baseline for pause-placement (i.e. insert a pause at every sentence boundary). Notably, we have presented a model that utilizes a set of features related to the syntax structure of a sentence (rather than utilizing lexically specific features, such as word labels), which has enabled us to make use of a relatively small corpus to train our model.

We envision that this model could be incorporated as part of a system for automatically synthesizing animations of sign language, with the assumption that such a system is aware of the location of syntactic phrase boundaries during the generation of sentences (which is the basis for all features listed in Table 1), and thereby our model could utilize this information to automatically determine where to insert pauses in the resulting sign-language animation.

In future work, we plan to investigate additional predictive features and modeling techniques for this task, and to conduct a user-based study (with ASL signers evaluating the quality of animations resulting from this model). In subsequent work, we plan to investigate models of the duration (length) of both pauses and individual signs, with an ultimate goal of building software that can generate realistic and understandable animations of ASL, to make information more accessible for ASL signers who may prefer to receive information in the form of sign language or may have reduced reading literacy in written language.

7. Appendix

The Decision Tree and SVM classifiers were implemented in MATLAB using the *Classifier Learner Package*[1], while the Linear-Chain CRF classifier was implemented using the *sklearn-crfsuite*[2] package in Python. Table 3 displays the parameter settings used to build the respective models.

Table 3: Parameters for machine-learning models

Classifier	Function	Parameters
Linear-Chain CRF	*CRF*	algorithm: l2sgd c2: 0.0869 max_iterations: 100 all_possible_transitition: True
Decision Tree	*fitctree*	SplitCriterion: gdi MaxNumSplits: 100 Surrogate: off
SVM (Linear)	*fitcsvm*	KernelFunction: linear PolynomialOrder: [] KernelScale: auto BoxConstraint: 1

[1] https://www.mathworks.com/products/statistics/classification-learner.html

[2] https://sklearn-crfsuite.readthedocs.io/en/latest/

8. Bibliographical References

Adamo-Villani, N., & Wilbur, R. B. (2015, August). ASL-Pro: American sign language animation with prosodic elements. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 307-318). Springer, Cham.

Bungeroth, J., Stein, D., Dreuw, P., Zahedi, M., & Ney, H. (2006, May). A German sign language corpus of the domain weather report. In *Fifth International Conference on Language Resources and Evaluation* (pp. 2000-2003).

Cox, S., Lincoln, M., Tryggvason, J., Nakisa, M., Wells, M., Tutt, M., & Abbott, S. (2002, July). Tessa, a system to aid communication with deaf people. In *Proceedings of the fifth international ACM conference on Assistive technologies* (pp. 205-212). ACM.

Ebling, S., & Glauert, J. (2016). Building a Swiss German Sign Language avatar with JASigning and evaluating it among the Deaf community. *Universal Access in the Information Society*, *15*(4), 577-587.

Gohel, J. (2016). Modeling the locational distribution of spatial reference points established by ASL signers in a motion capture data corpus. *Master's capston report, R.I.T.*

Grosjean, F. (1977). The perception of rate in spoken and sign languages. *Attention, Perception, & Psychophysics*, *22*(4), 408-413.

Grosjean, F. (1979). A study of timing in a manual and a spoken language: American Sign Language and English. *Journal of Psycholinguistic Research*, *8*(4), 379-405.

Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive psychology*, *11*(1), 58-81.

Grosjean, F., & Lane, H. (1977). Pauses and syntax in American sign language. *Cognition*, *5*(2), 101-117.

Huenerfauth, M. (2004). Spatial and planning models of ASL classifier predicates for machine translation. In *Proceedings of the 10th International Conference on Theoretical and Methodological Issues in Machine Translation* (TMI 2004). Baltimore, MD, USA.

Huenerfauth, M. (2008). Generating American Sign Language Animation: Overcoming Misconceptions and Technical Challenges. *Universal Access in the Information Society*, 6(4). Springer-Verlag. DOI=https://doi.org/10.1007/s10209-007-0095-7

Huenerfauth, M. (2009). A linguistically motivated model for speed and pausing in animations of American Sign Language. *ACM Transactions on Accessible Computing (TACCESS)*, *2*(2), article 9, 31 pages. DOI=http://dx.doi.org/10.1145/1530064.1530067

Huenerfauth, M. (2014). Learning to generate understandable animations of American sign language. In *Proceedings of the 2nd Annual Effective Access Technologies Conference*, Rochester, NY, USA, June 2014. *R.I.T.*

Huenerfauth, M., & Kacorri, H. (2015). Augmenting EMBR virtual human animation system with mpeg-4 controls for producing ASL facial expressions. In *International Symposium on Sign Language Translation and Avatar Technology* (3).

Huenerfauth, M., & Lu, P. (2010). Modeling and synthesizing spatially inflected verbs for American Sign Language animations. In *Proceedings of the 12th international ACM SIGACCESS conference on computers and accessibility (ASSETS'10), 99-106. DOI=http://dx.doi.org/10.1145/1878803.1878823*

Jennings, V., Elliott, R., Kennaway, R., & Glauert, J. (2010, May). Requirements for a signing avatar. In *Proceedings of the 4th LREC Workshop on the Representation and Processing of Sign Languages* (pp. 133-136).

Kacorri, H. (2016). *Data-driven synthesis and evaluation of syntactic facial expressions in American Sign Language animation*. City University of New York.

Kennaway, J. R., Glauert, J. R., & Zwitserlood, I. (2007). Providing signed content on the Internet by synthesized animation. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *14*(3), 15.

Lu, P. (2014). *Data-driven synthesis of animations of spatially inflected American Sign Language verbs using human data*. City University of New York.

Lu, P., & Huenerfauth, M. (2012). CUNY American Sign Language motion-capture corpus: first release. In *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon, The 8th International Conference on Language Resources and Evaluation (LREC 2012), Istanbul, Turkey*.

Lu, P., & Huenerfauth, M. (2014). Collecting and evaluating the CUNY ASL corpus for research on American Sign Language animation. *Computer Speech & Language*, *28*(3), (May 2014), pp. 812-831. Elsevier. DOI=https://doi.org/10.1016/j.csl.2013.10.004

MathWorks Inc. 2017. Classification Learner-MATLAB. MATLAB and Statistics Toolbox Release 2017a. Retrieved April 13, 2017 from https://www.mathworks.com/products/statistics/classification-learner.html

McDonald, J., Wolfe, R., Schnepp, J., Hochgesang, J., Jamrozik, D. G., Stumbo, M., Berke, L., Bialek, M. & Thomas, F. (2016). An automated technique for real-time production of lifelike animations of American Sign Language. Universal Access in the Information Society 14(4):2016. pp. 551-566.

Mitchell, R. E., Young, T. A., Bachleda, B., & Karchmer, M. A. (2006). How many people use ASL in the United States? Why estimates need updating. *Sign Language Studies*, *6*(3), 306-335.

Morrissey, S., & Way, A. (2005). An example-based approach to translating sign language. *Proc. Workshop on Example-Based Machine Translation,* 109-116.

Naert, L., Larboulette, C., & Gibet S. (2017). Coarticulation Analysis for Sign Language Synthesis.

International Conference on Universal Access in Human-Computer Interaction, Vancouver, Canada.

Neidle, C. (2002). SignStream™: A Database Tool for Research on Visual-Gestural Language. In Brita Bergman, Penny Boyes-Braem, Thomas Hanke, and Elena Pizzuto, eds., Sign Transcription and Database Storage of Sign Information, a special issue of Sign Language and Linguistics 4 (2001):1/2, pp. 203-214.

Pfau, R., Steinbach, M., & Woll, B. (Eds.). (2012). *Sign language: An international handbook* (Vol. 37). Walter de Gruyter.

Sarkar, P., & Rao, K. S. (2015, July). Data-driven pause prediction for synthesis of storytelling style speech based on discourse modes. In Electronics, Computing and Communication Technologies (CONECCT), 2015 IEEE International Conference on (pp. 1-5). IEEE.

Segouat, J., & Braffort, A. (2009, February). Toward the study of sign language coarticulation: methodology proposal. In *Advances in Computer-Human Interactions, 2009. ACHI'09. Second International Conferences on* (pp. 369-374). IEEE.

Shibata, H., Nishimura, H., & Tanaka, H. (2016, July). Basic Investigation for Improvement of Sign Language Recognition Using Classification Scheme. In International Conference on Human Interface and the Management of Information (pp. 563-574). Springer International Publishing.

Traxler, C. B. (2000). The Stanford Achievement Test: National norming and performance standards for deaf and hard-of-hearing students. *Journal of deaf studies and deaf education*, *5*(4), 337-348.

Vcom3D, Inc, Sign Smith. Retrieved December 13, 2017 from http://www.vcom3d.com/signsmith.php

Wilbur, R. B. (2009). Effects of varying rate of signing on ASL manual signs and nonmanual markers. Language and speech, 52(2-3), 245-285.