A Personalized American Sign Language Game to Improve Short-Term Memory for Deaf Children

Language Recognition

Pranay Agrawal

College of Computing

Georgia Institute of Technology

Supervisor

Prof. Thad Starner and Prof. Thomas Ploetz

In partial fulfillment of the requirements for the degree of

Bachelor of Science in Computer Science

May 5, 2022

Contents

1 Introduction

Short-term phonological memory is the capacity for holding speech-related information for a short period of time. The formation of this short-term memory in the phonological store is impacted drastically by a person's language acquisition at a young age [5]. Even though American Sign Language (ASL) is the primary mode of communication for most children today who are born deaf in the United States of America, sign language acquisition has shown be to as effective as spoken language in developing short-term memory [15].

Deaf children born to deaf parents learn ASL from their parents and hearing children learn their first language from their parents. However, 90-95% of deaf and hard of hearing children are born to hearing parents that do not know enough ASL to teach their children. This lack of adequate language exposure can have detrimental effects on working memory in these kids especially between the critical language learning ages of two and five [10, 3]. Studies have shown that these pre-lingually deafened children with weaker working memory are able to repeat only one or two signs compared to the average four to six signs [3]. This deficiency of short-term memory combined with the weakening of other language skills can lead to Language Deprivation Syndrome (LDS), a condition with poor lifelong outcomes, including a 2-7x increase in mental health problems [5], 50% unemployment rate [11], and a 3-30x increase in suicide rates [16]. As a result, it is crucial to have proper development of short-term memory through language acquisition. Studies have demonstrated that children capable of expressing themselves in ASL at an early age quickly increase their short-term memory, thereby mitigating the risk of suffering from LDS [14].

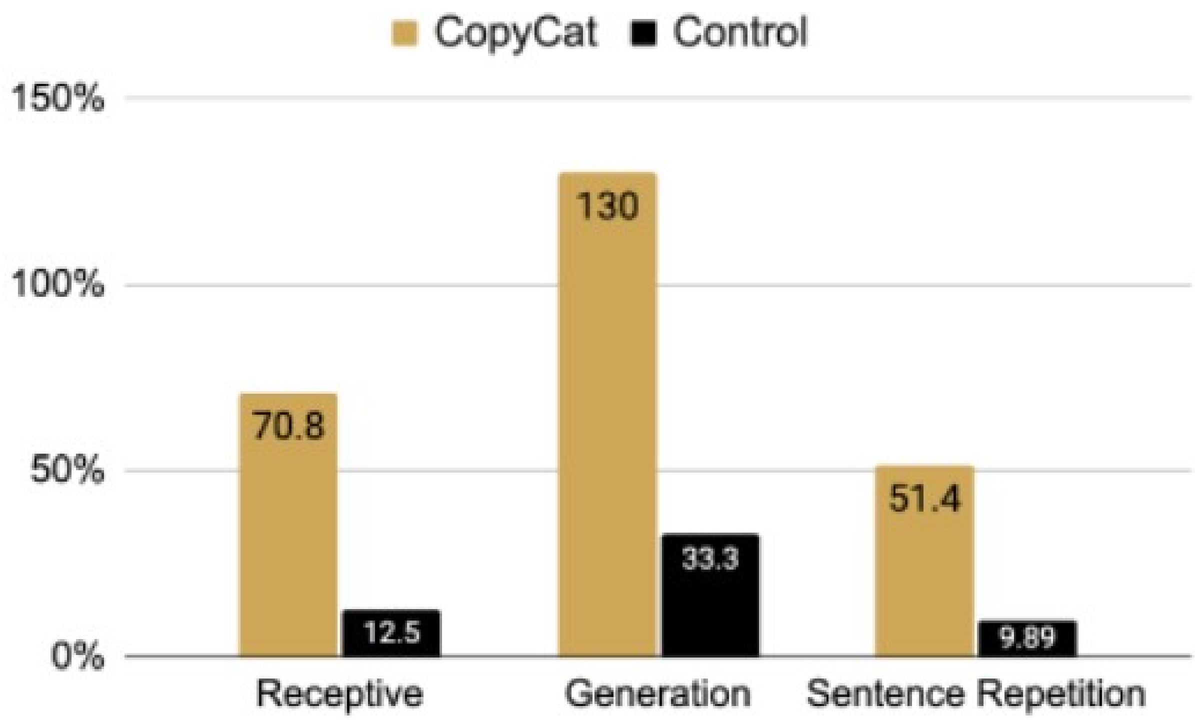

Figure 1: Observed Working Memory Increase from Old CopyCat versions. All improvements are statistically significant at $p < 0.05$

Fortunately, previous research has demonstrated that interactive games can significantly improve short-term memory skills. A Wizard of Oz experiment conducted by Weaver et al. showcased that children who play CopyCat alongside the class curriculum increase their scores on Receptive, Generative, and Sentence Repetition metrics more than children who do not play CopyCat (Figure 1) [14]. As these metrics provide a good estimate of a child's short term memory, one can expect similar results as long as the recognition accuracy is similar to a human sign linguist's accuracy of 92%.

As technology becomes more advanced and recognition models become more accurate, applications of gesture technology have become more widespread. One major application involves the recognition and verification of sign language for practical communication as well as basic learning. However, the majority of results published involve controlled lighting conditions and sensor-based gloves, leading to minimal practically in the real world [9]. Furthermore, many systems do not recognize the importance of interaction between the human and the computer, leading to an absence in the full scope of benefits that technology such as this could provide. This limited environment and learning that often involves the expensive financial cost and hassle of a custom setup usually leads to a cumbersome experience for both the child as well as his or her parents [4].

As a result, the exploration and development of an inexpensive and accurate system in a practical setting that focuses on the benefit of language acquisition with key user interactions is much needed for the future generations of deaf children. Based on this evidence, we are developing an educational machine-learning backed game called CopyCat. While previous attempts at educational ASL games typically focus on language reception rather than phrase generation [4, 3, 18], CopyCat is an interactive and entertaining game that helps deaf children hone their ASL generation skills and facilitates short-term language memory acquisition. The game presents the user with a scene, asks the user to describe it using sign language, and then uses natural language processing to determine the correctness of the signed phrase. As seen in Figure 2, the game presents the child with a scene on the right-hand side (lion above gray wall).

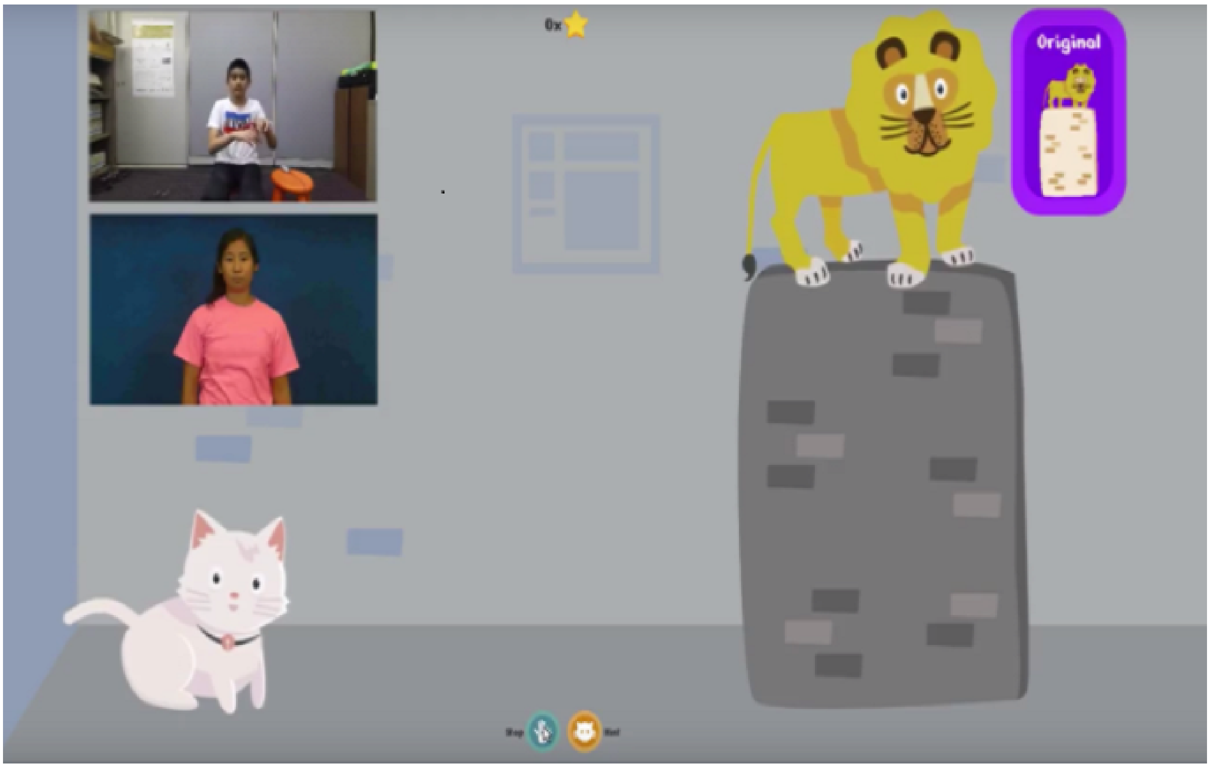

Figure 2: CopyCat Game UI

The practical scale and impact of gesture technology on the deaf community currently remains limited. Understanding how to further the CopyCat system will significantly widen the impact of gesture technology on deaf children and their quality of life. Through the use of our system at an early age, we expect to counteract the deficiency in short-term memory that deaf children of hearing parents often face. This will help these children avoid continuous language deprivation and therefore life-threatening consequences such as Language Deprivation Syndrome. Furthermore, overcoming these challenges will allow deaf children to participate more confidently in academics and limit the socioeconomic risk of inadequate education.

2 Literature Review

Language acquisition is an important contributor in forming short-term memory [5]. Unfortunately, 90-95% of deaf and hard of hearing children are raised in linguistically impoverished homes with minimal language interaction [8]. This minimal language interaction leads to problems with communication from an early age and decreases the child's short-term working memory. As a result, deaf children from hearing households face immediate issues in school compared to their peers such as lower academic performance and poorer reading ability [11]. Long term, delayed language acquisition will lead to Language Deprivation Syndrome.

As research has linked early development of sign language as an effective method to improve short-term memory, most of the past studies in interactive American Sign Language (ASL) software focuses on the ability of students to comprehend language [13]. However, due to a lack of repetition with feedback and the importance of the user generating language, benefits remain marginal. Furthermore, the limited size of ASL datasets make the task challenging. As a result, despite the advancements made in this field over the last few decades, there are many difficult challenges yet to be addressed. Bragg et al. outlined the shortcomings that public sign language datasets have that limit the power and generalizability of systems trained on them [2]. Unlike speech recognition, which has been trained on corpora containing millions of words, sign language corpora typically contain fewer than 100,000 articulated signs. Larger datasets for sign language recognition are needed for generalization to unseen scenarios and users, but these can be time-consuming and expensive to assimilate. Most sign language datasets contain only individual signs though most real-world use cases of sign language processing involve natural conversational with complete sentences and longer utterances. Our dataset tries to address this challenge by collecting sentence data.

Yin et al. showed translation of German Sign Language with a state-of-the-art STMC-Transformer [17]. By doing so, they obtained a word error rate of 21% on the PHOENIX14T dataset consisting of signed weather reports. However, Transformers often require larger datasets than HMMs as they have more parameters to train and must learn the grammar from scratch. By restricting the signed vocabulary to a limited set of signed phrases, Bansal et al. on CopyCat showed that HMMs outperform Transformers on average by 17% on ASL recognition of phrases with a predefined grammatical structure [1].

Furthermore, receptive language, expressive language, and sentence repetition are three crucial measures of working memory. CopyCat is a game that has shown an increase of approximately 70% in receptive language, 130% in expressive language, and 50% in sentence repetition in two weeks [1]. Although both the recognition error and verification error of this system were negligible, the system involved the use of expensive custom-built colored gloves with embedded accelerometers in a kiosk to control position and lighting [4]. Furthermore, the proper equipment required a custom setup that cost over $3,000, significantly limiting the accessibility of the project and hence the benefits.

A future iteration of CopyCat then explored the potential of the Azure Kinect 3D depth camera body tracking as a solution to minimize the equipment and the majority of the cost required [7][18]. However, the results on the Kinect system averaged only 74% word and 36% sentence accuracy with Hidden Markov Models (HMMs), creating a negative outcome on the overall user experience [18]. Therefore, the scale and the impact of gesture technology on the deaf community remained restricted.

3 Materials and Methods

3.1 CopyCat Game

Subject	Adjective	Preposition	Object
Monkey	Black	In	Bed
Lion	Grey	Above	Chair
Alligator	Orange	Below	Wall
Snake	Blue		Wagon
	White		Flowers
			Box

Table 1: CopyCat game vocabulary

The Adaptive CopyCat game consists of an 18 sign vocabulary set (Table 1) with a total of 58 different phrases. Sophisticated linguistic features in ASL like facial gestures are not considered in the game to reduce overall complexity. Three-sign, four-sign, and five-sign phrases each have their own level representing the difficulty in the game. Each phrase follows a specific sign-based grammar:

[adjective1] subject preposition [adjective2] object

All levels include the subject, preposition, and object signs. Neither adjectives are used in the three-sign level. One adjective is used in the four-sign level and both adjectives are used in the five-sign level.

3.2 Feature Extraction

Using a Microsoft Azure Kinect depth camera, each frame of the 4K RGBD video is captured. A human pose estimator is then used to extract the features from each frame. While we previously considered Google MediaPipe, Microsoft Azure Kinect SDK, and AlphaPose, as seen in Figure 3, based on results from [1] that all three pose estimators perform with relatively the same accuracy, we have decided to only consider Microsoft Azure Kinect SDK moving forward. The Azure Kinect SDK v1.3.0 combines its depth readings with its 4K RGB data to calculate the (x,y,z) coordinates of 32 joint positions [6]. Since the Azure Kinect SDK returns three coordinates per joint position, this results in 96 features in total. In addition to these raw features, we post-process pose recognition to generate additional features such as delta features, relative features, and normalized features. Throughout the past years, the CopyCat team has assembled a 4K RGBD dataset by collecting videos recorded using Azure Kinect from 12 adults at Georgia Tech with little ASL experience, contributing a public pose estimator feature dataset of appropriately 7000 videos.

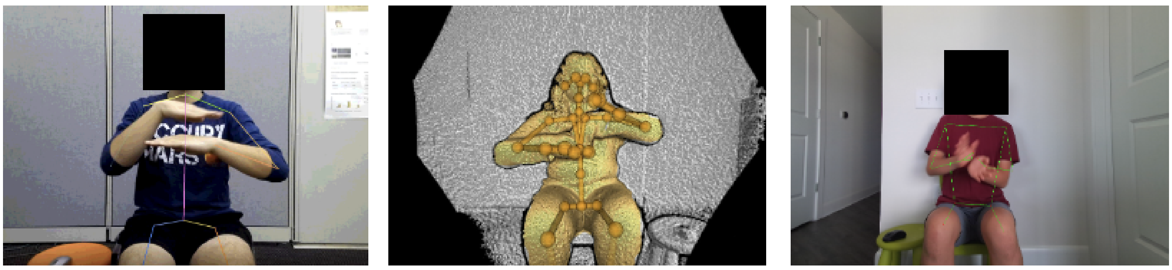

Figure 3: AlphaPose (left), Azure Kinect (middle), and MediaPipe (right)

3.3 Hidden Markov Model

Selected features are then integrated with the appropriate recognition model to enable training. We implement HMMs using the Hidden Markov Model Toolkit (HTK) with Kaldi to enable HMM training on sequential 2D and 3D data.

$X = Features, Y = Label, T = TotalTimesteps$
Initialize HMM Parameters $\theta = (A, B, \pi)$ randomly
$\alpha(X_0) = P[Y_0, X_0] = P[Y_0|X_0]P[X_0]$
$\beta(X_T) = 1$
while $i \leq iterations$ **do**
  for $k = 0 \rightarrow T$ **do**
    $\alpha(X_k) = \sum_{X_{k-1}} \alpha(X_{k-1})P(X_k|X_{k-1})P(Y_k|X_k)$
  end for
  for $k = N \rightarrow 0$ **do**
    $\beta(X_k) = \sum_{X_{k+1}} \beta(X_{k+1})P(X_{k+1}|X_k)P(Y_{k+1}|X_{k+1})$
  end for
  $\eta(X_k) = \frac{\alpha(X_k)\beta(X_k)}{\sum_{X_k} \alpha(X_k)\beta(X_k)}$
  $\epsilon(X_k, X_{k+1}) = \frac{\alpha(X_k)\beta(X_{k+1})P[X_{k+1}|X_k]P[Y_{k+1}|X_{k+1}]}{\sum_{X_k} \alpha(X_k)\beta(X_{k+1})P[X_{k+1}|X_k]P[Y_{k+1}|X_{k+1}]}$
  $\pi_0^* = \eta(X_0)$
  $A_{ij}^* = \frac{\sum_k \epsilon(X_k=j, X_{k-1}=i)}{\sum_k \eta(X_{k-1}=i)}$
  $B_{ij}^* = \frac{\sum_k \eta(X_k=i)1_{Y_{k=j}}}{\sum_k \eta(X_k=i)}$
end while

Algorithm 1: Baum-Welch Re-estimation

create path matrix $viterbi[N, T]$
for $s = 1 \rightarrow N$ **do**
  $viterbi[s, 1] = \pi_s * b_s(\varnothing_1)$
  $backp[s, 1] = 0$
end for
for $= 2 \rightarrow T$ **do**
  for $s = 1 \rightarrow N$ **do**
    $viterbi[s, t] = \max viterbi[s', t-1]a_{s',s}b_s(\varnothing_t)$
    $backp[s, t] =$ argmax $viterbi[s', t-1]a_{s',s}b_s(o_t)$
  end for
end for
$bestpathprop = \max viterbi[s, T]$
$bestpathpointer =$ argmax $viterbi[s, T]$
$bestpath =$ path starting at $bestpathpointer$, follows backp[] to states back in time

Algorithm 2: Viterbi Decoding

HMMs are probabilistic models that try to understand Markov processes – events where the current state depends only upon the previous state. HMMs model such processes by learning transition probabilities between states and emission probabilities of different observations at each state. The states are hidden (unobservable), hence the reason that HMMs are "hidden". Using a specialized case of the EM algorithm, the Baum-Welch re-estimation algorithm (Algorithm 1) is used to train an HMM for each sign for roughly 200 iterations. Every 25 iterations, the number of mixtures in the Gaussian mixture distribution is increased, allowing for more complex observations to be made. Finally, the Viterbi Decoding algorithm (Algorithm 2) is used to find the most likely sequence of hidden states and HMM models based on the trained selected features of the pose estimator.

3.4 Adaptive CopyCat

The study improves the accuracy and recognition of the machine-learning backed system. In order to achieve this improvement, we introduced a Progressive User Adaptive (PUA) model (Algorithm 3). This adaptive model progressively personalizes the model on the user as he or she plays the CopyCat game that advances in complexity from 3-sign phrases to 5-sign phrases. PUA modeling adds the new user's signed phrases that the system detected as correct in the previous levels. When the user advances from the 3-sign level to the 4-sign level, the model will retrain on the original dataset plus repetitions of all of the user's correctly signed 3-sign phrases, as determined by the user independent recognizer. When advancing to the 5-sign level, repetitions of the user's correctly signed 4-sign phrases are also added. Since the amount of new data collected is significantly smaller than the amount of existing training data, data repetition is used to increase the influence of new data on the training model. Experimental results have indicated that a 20x repetition of each new record most effectively incorporates new data into PUA models.

X = Original Features Data-set
$R \leftarrow$ Number of phrase repetitions (20)
Y_n = Labels of n-sign data in X, T = Total Time steps
S_n = System-detected correct n-sign words
$level$ = next level the user will enter (3, 4, 5, etc...)
for $n = 4 \rightarrow level$ **do**
    Perform Baum-Welch with parameters X, Y_{n-1}
    for $i = 1 \rightarrow R$ **do**
        $X \leftarrow (X, S_{n-1}^{orig})$
    end for
end for

Algorithm 3: Progressive User Adaptive Model

To test the accuracy of PUA, we perform leave-one-signer-out testing. We emulate a previously unseen user with the left-out user and perform the PUA procedure as described above (starting with the user independent recognizer trained using the 11 other users' data).

As a final test to avoid overfitting, we introduce a new participant six months after video data from the first 12 users was collected. This user signed each phrase from our dataset four times correctly and once incorrectly. This 13th user served as the novel user for our recognition system and helped us evaluate the model with purposely incorrect signs. Even if the actual signed phrase did not match the expected phrase, if the model predicted a match, the model would still incorporate the data into the training for the future levels.

4 Results and Discussions

4.1 HMMs vs Transformers

Recognition Model	HMMs Average	Transformers Average
AlphaPose	**90.6 (71.9)**	67.2 (51.1)
Kinect	**90.5 (71.6)**	75.7 (62.9)
MediaPipe	**90.4 (71.7)**	77.5 (65.9)

Table 2: User independent word (sentence) percent accuracy using HMMs and Transformers

Previous work from Bansal et al. has shown that HMMs are better suited for ASL recognition for the CopyCat game compared to deep learning alternatives such as LSTMs and Transformers [1]. HMMs perform better on time series and pattern recognition problems along with low training data requirements. Due to the limited data available for ASL and the grammatical phrase structure of CopyCat, HMMs achieve over 17% better word accuracy than Transformers in the User Independent setting on 8 users (Table 2).

4.2 User Independent vs Adaptive HMMs

Recognition Model	User Independent	Progressive User Adaptive
3-sign	95.0 (87.0)	NA
4-sign	95.3 (84.2)	**96.7 (88.9)**
5-sign	96.0 (84.3)	**97.9 (92.0)**

Table 3: User independent word (sentence) percent accuracy using User Independent and Progressive User Adaptive model with HMMs

After expanding the dataset from 8 users to 12 users, randomizing the training data order, and improving the list of features selected, HMMs trained on user independent data for Kinect achieve on average roughly 95% word accuracy and 85% sentence accuracy (Table 3). However, these models do not achieve the 92% sentence accuracy of a human sign linguist, a threshold CopyCat recognition aims to achieve in order to achieve the Wizard of Oz experiment benefits (Figure 1).

Using HMMs trained on progressive user adaptive data for Kinect has shown significant improvements (Table 3). For 4-sign phrases, the average sentence accuracy increased by 5.7% absolute to 88.9%. Average sentence accuracy increased for the 5-sign phrases by 7.7% absolute to 92.0%. While we are clearly not achieving 92% sentence accuracy with user independent training, PUA eventually gets there over time, where we are able to achieve human sign linguistic accuracy for 5-sign phrases. Note that PUA 3-sign uses the user independent model as the user has not completed any level of CopyCat.

In the group's original studies, When testing on our target population of 6-8 year-old Deaf children of hearing parents, we were impressed by how persistent they were at playing the game. When the game would obviously misclassify a phrase, one child signed "STUPID CAT. LOOK AT ME!" and then repeated the phrase. We designed the game to declare a level unfinished after four unsuccessful attempts at signing when given the correct phrase. Such an occurrence would cause the game to reduce the length of the required phrase for the next scene, while repeated successes would increase the length of the required phrase. Thus, while there are implications for the recognizer to mis-verify a child's signing, given the recognition rates seen here, there is a good chance that the recognizer would accept a second or third correct attempt by the child. This may avoid the child from becoming frustrated as the system will eventually recognize their attempt as correct.

3-sign	UI Incorrectly Signed	UI Correctly Signed
attempt 1	97.2 (105/108)	94.6 (105/111)
attempt 2	n/a	99.1 (110/111)
attempt 3	n/a	99.1 (110/111)
attempt 4	n/a	100 (111/111)
4-sign		
attempt 1	92.2 (107/116)	86.0 (92/107)
attempt 2	n/a	96.3 (103/107)
attempt 3	n/a	99.1 (106/107)
attempt 4	n/a	100 (107/107)
5-sign		
attempt 1	100 (108/108)	87.8 (97/108)
attempt 2	n/a	98.1 (106/108)
attempt 3	n/a	100 (108/108)
attempt 4	n/a	100 (108/108)

Table 4: User Independent (UI) percentage phrases correctly recognized as correct or incorrect from the previously unseen 13th participant

Table 4 shows the results from our 13th participant, recruited six months after our initial 12. This participant signed the same phrase correctly four times in a row to emulate a player making repeated attempts at having the recognizer approving their phrase. The user's user independent accuracies are quite good, suggesting that our recognizer is not overfitting due to tuning of hyperparameters during cross validation. As can be seen on the rightmost column of the table, the recognizer is highly likely to accept a phrase within two attempts.

Before providing the four repeated phrases, the 13th participant was first prompted to sign the phrase incorrectly, with only one sign in the phrase being incorrect. The system recognized that this “incorrect” phrase was, in face, incorrect an average of 96.4% of the time. Both the false positive rate and overall accuracy observed with the 13th participant are better than the operational glove system deployed with the Georgia School for the Deaf. In future studies, we plan for this participant to play Adaptive CopyCat live with the progressive user adaptive algorithm.

5 Future Work

There are many opportunities for building upon the current work.

5.1 Model Analysis

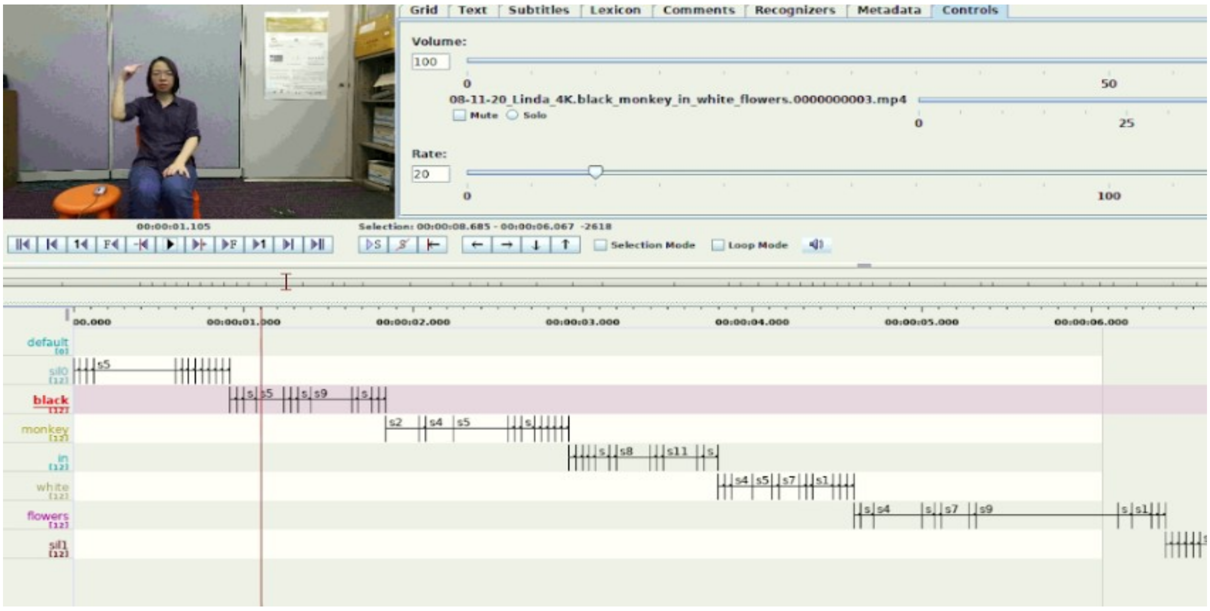

Figure 4: ELAN annotation tool for sign language

We have been working on analyzing our models and the features we are pulling from pose estimators to improve performance. We have built upon ELAN, an annotation tool used widely in the sign language community, to look into state and word boundaries, as well as features such as log likelihoods (Figures 4 and 5). One use of this data is to understand which words and phrases are confused the most. For example, the prepositions "in" and "above" are often confused, so using the ELAN tool, we can figure out which of our model states are getting confused. We can then train specialized classifiers for those confused words to drastically improve performance. Furthermore, we are looking into ways that we can utilize information theory to order phrases such that we can extract the most information about a user in a small amount of phrases for progressive user adaptive.

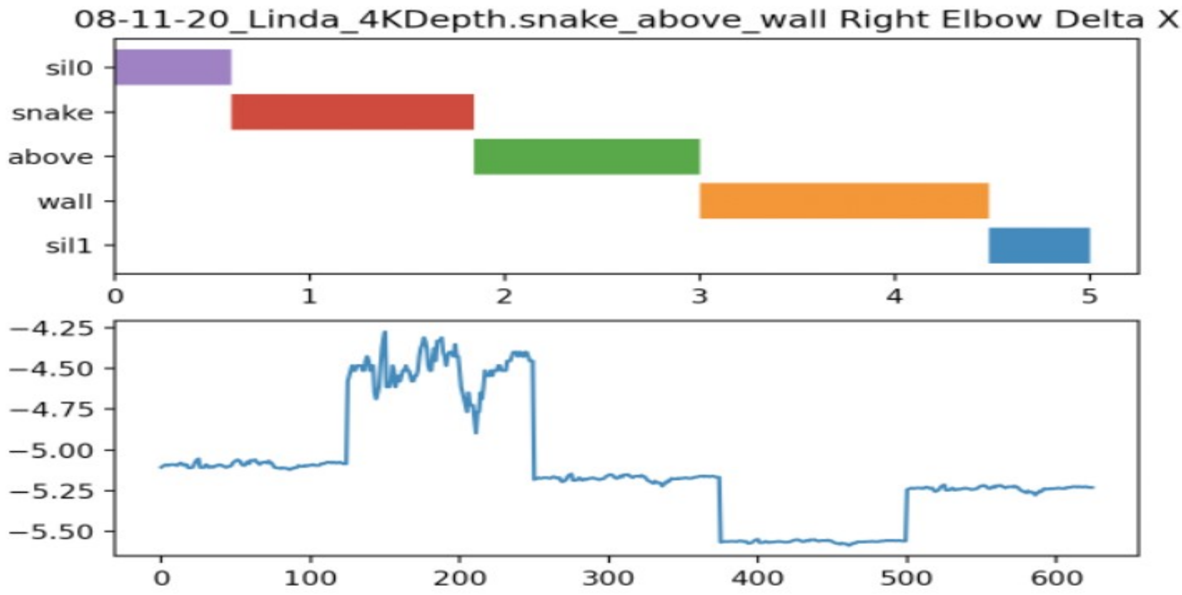

Figure 5: Custom tool for investigating log likelihood and boundaries

5.2 User Studies

We will also conduct a large-scale user study on deaf children to verify the educational benefits of the new CopyCat game. The study will be designed to confirm education benefits, ease of use of the system, and recognition accuracies in real life scenarios. We will be connecting with deaf children at the Atlanta School of the Deaf for conducting this user study. For conducting effectiveness studies on deaf children, we will integrate our final machine learning model with our full-fledged PC game in a setting similar to the one used by Weaver et al [14].

5.3 Live Adaptive CopyCat

The 13th participant has played the complete CopyCat game using a user independent model from the 12 existing users in the dataset. In doing so, we were able to test the accuracy of the model in a simulated setting with especially incorrectly signed data as well. The next step would be to have a similar participant play the complete CopyCat game using the progressive user adaptive model. By incorporating the data of the user as he or she plays the game, we can investigate the possible concerns of tuning the model towards the user's signing. Furthermore, we can investigate the effects of tuning the model on a phrase by phrase basis and incorporating phrases that are better recognized by the model earlier on in the game.

5.4 CopyCat Game and Dataset

An obvious improvement that will most likely improve accuracy is increasing the size of the dataset, expanding the training dataset. While this can be performed by recruiting additional signers, expanding the breadth of the current dataset from 12 users, CopyCat can also be extended past 5-sign phrases. Not only will this expansion allow for increased complexity in the game, but one can also expect further sentence accuracy improvements using the progressive user adaptive algorithm. Using insights from these results, we will move on to our target population of deaf children. As the current system is trained on adults instead of 6-8 year olds, we need to not only test the system with children for the education effect but also for the recognition and verification accuracies. Expanding the dataset to include the target population is essential for model generalization.

6 Conclusion

The results we present above suggest that a game based on sign language recognition is feasible. We show that Adaptive CopyCat successfully adapts to the addition of a new user and has the ability to perform at human sign linguistic sentence accuracy over time. Furthermore, we show that the user independent model does not suffer from overfitting on the data. Results on the CopyCat children dataset from the older system using custom gloves and accelerometers revealed that verification yielded higher accuracies compared to recognition. We observe that with our new system on adult users using pose estimators, we are able to achieve an accuracy over the 82% threshold obtained from verification on the older system. This result is not only promising, but it also opens an avenue of incorporating verification into our current game version to evaluate potential increase in the accuracy. In the near future, we hope to evaluate Adaptive CopyCat with the 13th participant and test our current system with deaf children by incorporating their videos into our current dataset.

References

[1] Bansal, D., Prerna, R., Matthew, S., Agrawal, P., et al. “CopyCat: Using Sign Language Recognition to Help Deaf Children Acquire Language Skills.” Association for Computing Machinery, (2021). https://doi.org/10.1145/3411763.3451523

[2] Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., et al. “Sign language recognition, generation, and translation: An interdisciplinary perspective.” ACM-SIGACCESS, (2019): 16-31.

[3] Brashear, H., Henderson, V., Park, K., Hamilton, H., Lee, S., and Starner, T. “American sign language recognition in game development for deaf children.” ACM-SIGACCESS, (2006): 79-86. https://doi.org/10.1145/1168987.1169002

[4] Brashear, H., Starner, T., Lukowicz, P., & Junker, H. “Using multiple sensors for mobile sign language recognition.” IEEE, (2003). https://doi.org/10.1109/ISWC.2003.1241392

[5] Brown, Patricia & Cornes, Andrew. “Mental Health of Deaf and Hard-of-Hearing Adolescents: What the Students Say.” Journal of deaf studies and deaf education, (2014). https://doi.org/10.1093/deafed/enu031

[6] Coulter, D., Wang, Y., & Meadows, P. “Azure Kinect body tracking joints.” https://docs.microsoft.com/en-us/azure/kinect-dk/body-joints

[7] Kratimenos, A., Pavlakos, G., & Maragos, P. “3D Hands, Face and Body Extraction for Sign Language Recognition.” ECCV SLRTP, (2020). Retrieved from https://slrtp.com/papers/extended_abstracts/SLRTP.EA.15.022.paper.pdf

[8] Mitchell, Ross & Karchmer, Michael. “Chasing the Mythical Ten Percent: Parental Hearing Status of Deaf and Hard of Hearing Students in the United States.” Sign Language Studies, no. 2 (2004): 138-163. https://doi.org/10.1353/sls.2004.0005

[9] Neiva, D. H. & Zanchettin, C. “Gesture recognition: A review focusing on sign language in a mobile context.” Elsevier, (2018). https://doi.org/10.1016/j.eswa.2018.01.051

[10] Newport, Elissa L. “Maturational constraints on language learning.” Cognitive science, no. 1 (1990): 11-28. https://doi.org/10.1207/s15516709cog1401_2

[11] Perkins-Dock, R. E., Battle, T. R., Edgerton, J. M., & McNeill, J. N. “A Survey of Barriers to Employment for Individuals who are Deaf.” JADARA, no. 2 (2015). Retrieved from https://repository.wcsu.edu/jadara/vol49/iss2/3/

[12] Sullivan, Patricia & Knutson, John. “Maltreatment and Disabilities: A Population-Based Epidemiological Study.” Child abuse & neglect, (2000): 1257-73. https://doi.org/10.1016/S0145-2134(00)00190-3

[13] Henderson, V., Lee, S., Brashear, H., Hamilton, H, Starner, S., & Hamilton, S. “Development of an American Sign Language Game for Deaf Children.” Interaction design and children, (2005). https://dl.acm.org/doi/10.1145/1109540.1109550

[14] Weaver, A., Hamilton, H., Zafrulla, Z., Brashear, H., Starner, T., Presti, P., & Bruckman, A. “Improving the language ability of deaf signing children through an interactive American Sign Language-based video game.” ISLS, (2010). https://doi.dx.org/10.22318/icls2010.2.306

[15] Wilson, M., & Emmorey, K. “Comparing Sign Language and Speech Reveals a Universal Limit on Short-Term Memory Capacity.” Psychological Science, (2006): p. 682-683. https://doi.org/10.1111/j.1467-9280.2006.01766.x

[16] Windfuhr, KL., Turner, O., Windfuhr, K., & Kapur, N. “Suicide in deaf population: A literature review.” Annals of general psychiatry, (2007). https://doi.org/10.1186/1744-859X-6-26

[17] Yin, K., & Read J. “Better Sign Language Translation with STMC-Transformer.” arXiv, (2020). arXiv:2004.00588v2

[18] Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., & Presti, P. “American Sign Language Recognition with the Kinect.” ICMI, (2011). https://doi.org/10.1145/2070481.2070532