The 4th International Conference on Computing and Data Science (CONF-CDS 2022)
DOI: 10.54254/2755-2721/2/20220568

PakParse: Machine Translation from Text to Pakistan Sign Language for Effective Communication with Deaf People

Rana Arsal Mahmood[1], Haseeb Ahmad[1,*], Qurat-ul-Ann[1], and Muhammad Usman Ashraf[2]

[1] Department of Computer Science, National Textile University, Faisalabad, Pakistan
[2] Institute of Business Management Sciences University of Agriculture, Faisalabad, Pakistan

*Corresponding author's e-mail: haseeb_ad@hotmail.com

Abstract. The People around the world use different languages to communicate with each other. But there are many individuals who can't listen or speak. These people use gesture-based language for their communication which is called Sign Language. There is no universal Sign Language in world so it changes region wise. As Sign language can't be written and sign video approach is also expensive (in terms of create, store and manage), that's why sign animation through 3D avatar is best option. Recently Pakistan got its own Sign Language called Pakistan Sign Language (PSL). As a noteworthy population of deaf people live in country, there is no machine translator for PSL which can facilitate deaf in their communication and education. This work proposed a machine translator for Pakistan Sign Language called PakParse, which takes text (English & Urdu) as input, and animate corresponding signs by 3D avatar.

Keywords: Sign Language, Deaf, Pakistan Sign Language, Machine translator, HamNoSys, SiGML.

1. Introduction

Language is the most common and reliable source of communication between people. There are 6,906 different languages spoken around the world, which changes region to region [1]. But there is a large number of people who are hearing-impaired or speaking-impaired. So a big communication gap between healthy and deaf people exists. Deaf people use gesture-based language to communicate. This gesture-based language is called Sign Language. Sign language is composed of unique gestures and hand/arm movements. These gestures were supposed to explain the human feelings and thoughts as much as possible. Each particular sign represents a meaning or distinct letter of any language. But still, due to the absence of worldwide sign language, each country has its separate sign language. So it's still a problem to communicate with the deaf community of another country or region. American Sign Language (ASL), Chinese Sign Language (CSL), British Sign Language (BSL), Russian Sign Language (RSL), Swedish Sign Language (SSL), Pakistan Sign Language (PSL) and Indian Sign Language (ISL) are some famous Sign Languages practiced through the world [2]. These languages are different from each other in terms of signs.

© 2023 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2/20220568

1.1. Pakistan Sign Language

The deaf and mute community in Pakistan is over 0.1 million, which is 7.4% of the incapacitated population across the country, and only 5% of them have access to education. **Table 1** is showing a detailed disabled population of Pakistan. It is also notable that 55% of deaf people are in the 5-29 age group [3]. This vast number of people need the attention of authorities to reduce the intra-deaf and cross-deaf communication gap. This gap could only be reduced by using some gestures and signs which are recognized at least at the national level. Recently Pakistan got its sign language. Many nations have resolved this issue of their deaf community by launching different information technology-based projects to reduce the gap of communication [4]. In South and Central Asia, significant work has been done on Indian Sign Language and Chinese Sign Language but Pakistan Sign Language is still under development due to lack of resources and structured information related to grammar and syntax rules. Like all other SLs around the world, Pakistan Sign language also uses different gestures and signs to convey the meanings [5].

Table 1. Disabled population of Pakistan.

Sr.	Disability	Total population	As % of total disabled
1	Disability	265398	8.06
2	Blind	243683	7.40
3	Deaf & Mute	626785	19.00
4	Crippled	210854	6.40
5	Insane	250184	7.60
6	Mentally Retarded	270451	8.21
7	Having More than 1 disability	1426800	43.33

Pakistan sign language is gesture language having its syntax and vocabulary, which is majorly inspired by Urdu, which is the national language of Pakistan [6]. These signs are further classified into manual and non-manual signs. Manual signs consist of hand shapes, arm and body movements, and their locations [7]. Whereas non-manual signs include facial expressions, gazing, shoulder raising, and head movements. Every sign language is influenced by some spoken language as American Sign Language from English and French Sign Language from the French Language. Other regional languages such as Sindhi, Balochi, and Pashto also have their influence on Pakistan Sign Language. Another property of physical communication is fingerspelling. Which is about making of Urdu alphabets with the use of fingers. Urdu has 38 alphabets [8]. So there are 38 equivalent hand-shapes to represent them. These alphabet handshapes are used for short message communication. The deaf community across the country use these handshapes to communicate [9]. These alphabet handshapes are used where no specific sign exists. For this, the word is decomposed into characters of specified language and then each character is signed with its alphabet handshape.

2. Literature Review

People everywhere in the world have been researching on automatic SL translation system because other methods are expensive in many ways. These systems take input in the form of text or speech and produce animation through 3D avatars. Some of the systems are described in this subsection.

2.1. TESSA

This system was built on direct translation method. This system was conversion method from speech input into British Sign Language (BSL) which was supposed to minimize the deaf communication gap with a post office staff. The system takes input in English, lookup all words in English to Sign dictionary, integrate those morphemes together and create animation at end. In TESSA, the grammar which was

DOI: 10.54254/2755-2721/2/20220568

backbone of the system was formulaic grammar. In which multiple pre-defined phrases are saved for translation. The postal worker uses microphone set and language detector devices in which legit morphemes are saved. Speech recognizer match the phrase with legal phrase when staff speaks something. Post office staff's screen shows topics related to uttered phrase, i.e. "Bill Payments", "Postage" & "Passports". Staff select one phrase out of them according to the condition & relevant sign of that morpheme displays on display. TESSA is area specific project because it contains very minor amount of legit phrases to match and all related to post office work. It contains almost 370 phrases in it.

2.2. SignSynth

This is a sign-synthesis prototype under development at University of New Mexico. This system translates text into ASL. This system is almost identical to speech synthesis. Only the output modules are different. So backbone architecture is same. This system based on scripts written in Perl through the CGI (common gateway interface) for animating signs. The method consists of 3 main modules. Primary one is MENU CGI which displays list of options for signs through which operator can select phonological limits. Further options guide novel operators about ASCII-Stokoe. User can select handshapes, location & orientation for every pose. Next module, ASCII-Stokoe parser is for facilitation of experienced operators to set timings and non-manual features. User can type in Roman for finger spelling. This module gives output in the form of a tree which is input for conversion module. This module helps generating Web3D rotations which will use later. This Web3D rotation is sent to SignGen as input. SignGen binds them with Web3D avatar for creating file with animations. At the end, animation is played with help of plug-in. This system is free, and anybody can use it.

2.3. TEAM

TEAM project [10] was an English to ASL translator created at the University of Pennsylvania, which was based on synchronous tree adjoining rules of grammar, which was used to develop American Sign Language Syntactic Structure. A tree adjoining Grammar (TAG) is about rewriting the tree in which the elementary trees are fundamentals. This project not only focused on linguistic information but also the visual and three-dimensional information associated with ASL. The output of this section was American Sign Gloss with fixed parameters. The output was a virtual human as a signing avatar. This virtual avatar was a total of 80 joints and 135-degree freedom for rotation. But this project uses limited information, facial expressions and sentences to animate. This project gained particular success as it was one of the first of its kind. TEAM project only converted written text into American Sign Language and left the reverse portion (ASL to written text) for the future work. This will require American Sign Language recognizer and parser as in English; then, it will be possible. TEAM project was only specific for ASL but for other Sign Languages as well.

2.4. ViSiCAST

Elliott at el. [11] implemented a system in the University of East Anglia, which converts English into British sign language (BSL). As per authors, sign languages such as BSL built on its own syntax rules, morphology, and grammar. Also in Sign Languages, 'phoneme' are meaningful chunks of the text, which include hand orientation, movement & shape. At this phoneme level, sign languages are a multi-model in which, body postures and facial expressions mean a lot. Due to these problems, the ViSiCAST project was based on the partially-automatic type of sign order, in which natural language techniques are implemented in an interactive environment. Which, wherever needs, requires human interaction. To analyze the English input, they use CMU (Carnegie Mellon University) link parser. Then DRS (Discourse Representation Structures) semantic representation is generated. This DRS semantic representation then converted into meaningful phrases of BSL. Then Prolog grammar set of instructions were recycled to convert this into a virtual arrangement. While generating, the first portion of the development was based on head- driven phase structure semantics to generate symbolic Sign Language script. This symbolic SL script is known as 'Sign Gesture Markup Language (SiGML), which is written

DOI: 10.54254/2755-2721/2/20220568

lines of code required for the animation of a 3D avatar in Signing Space. This SiGML is simply comprehensible by render software, which animate avatar.

2.5. ZARDOZ

The ZARDOZ system [12] was a machine translator that translates the English language to SL consuming a chunk of manual representations as an Interlingua component. Some researchers of this project focused on AI representation and system architecture. So its design is heavy for reasoning and knowledge. In the first step, which is the analysis of input, the English text is decomposed into meaningful idiomatic phrases. Then these phrases went through syntactic parsing, which will fill out the spaces for situation/special event. The main benefit of this method was the easiness of reasoning components that could easily operate on the same syntax tree.

2.6. Boltay Hath

Boltay Hath is a system which identifies Pakistan Sign Language using special sensor-gloves [13]. Boltay Hath is based on a quite similar system which use sensor gloves to translate ASL into speech. This system was called Talking Hands. Gloves contain contact sensors and flex to sense the finger movements. Further ANN was diploid to get data from the gloves. These incoming streams of data are then classified into alphabets of English. This system was well-built with 92% precision rate. It first translates alphabets into text and then into speech.

3. Techniques Used in Proposed System

Just like spoken languages, SL can also be written for which multiple notations are available. We can use any notation system according to our requirements as there is no standard notation to follow specifically. These notations are helpful to represent the words of spoken language, which would be used later in the translation process. There are many techniques available to represent signs, but four of them are widely used. SignWriting [14], Stokoe Notation [15], Gloss and HamNoSys [16]. This system is based on Hamburg Notation System (HamNoSys).

3.1. Hamburg Notation system (HamNoSys)

HamNoSys (Hamburg Notation System) was developed in 1985 by the University of Hamburg in Germany. It has its SL representation for signs and gestures. It consists of almost 200 characters. Stokoe Notation System inspires HamNoSys.

HamNoSys is a transcript which is used globally for all kind of sign languages like phonetic alphabets for spoken words. All types of signs, including manual and non-manual gestures, are included in this transcript. This is a phonetic representation that includes different parameters, including hand, head, and body's location and movement. [17].

HamNoSys is decomposed into four main parts, including three subparts, as shown in

Figure.Modules given in solid boxes are required to represent a sign in HamNoSys, which includes initial configuration and arm/hand movements. The initial setup consists of Handshape, orientation, and hand location. On the other hand, modules given in dotted boxes are optional, which includes symmetry operator and non-manual features (NMF). The initial configuration module is used to represent manual signs (handshapes, hand movement, hand orientation, hand location), whereas non-manual signs include head tilting, facial expressions & shoulder raising. The symmetry module, which is on a very start position, is used to describe whether the sign is single-handed or double-handed. And the last module is used to specify whether the sign is static or dynamic. As HamNoSys was inspired by Stokoe notation, which only consists of basic hand shapes, HamNoSys has extended it to more than 200 unique symbol. Which includes:

- Handshapes of dominant hand.
- Location of both hands with respect to signer's body.
- Extended finger orientation of both hands.

DOI: 10.54254/2755-2721/2/20220568

- Palm orientation of dominant and non-dominant hand.
- Straight movement, curved movements and circular movements.

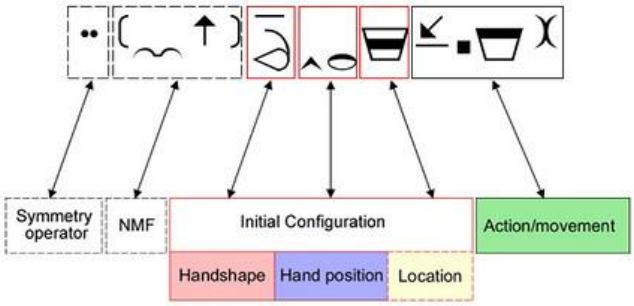

Figure 1. HamNoSys decomposition.

3.2. PSL HamNoSys Dataset

As for animation of signs by 3D avatar, each word is required to translate into HamNoSys code. As no previous work of this kind in Pakistan Sign Language is done, the dataset for PakParse is created. It includes 2000 regularly using words for translation.

3.3. SiGML

When the animation of the 3D avatar at run time is necessary, some framework is required for avatar's motion in such a way from which the related animation could generate automatically, instead of repeating pre-defined clips. This animation must be realistic in all directions, which ideally independent of specific Virtual Human's proportion. For this, a particular scripting language is required. SiGML (Signing Gesture Markup Language) is an XML, which was established by the University of East Anglia to represent sign sequences displayed in **Figure 2**. This language has different tags to notate relevant HamNoSys symbol. It builds on HamNoSys notations, and the output of this module will be lines of code having tags, which will be input for the avatar to animate the relevant sign.

```
<hamgestural_sign gloss="INFORMATIQUE">
  <sign_nonmanual>
    <head_tier> <head_movement
      movement="TL"/></head_tier>
  </sign_nonmanual>
  <sign_manual both_hands="true">
    <handconfig handshape="finger23spread"
      thumbpos="out"/>
    <handconfig extfidir="u"> </handconfig>
    <handconfig palmor="d"> </handconfig>
    <location_bodyarm location="shouldertop"/>
    <wristmotion motion="swinging"/>
  </sign_manual>
</hamgestural_sign>
```

Figure 2. SiGML of a word INFORMATIQUE.

3.4. 3D Avatars

Humanoids created digitally which can be rotate in 3 dimensions are called 3D avatars or virtual bodies. These avatars take input in SiGML format and generate corresponding animations. These animation

DOI: 10.54254/2755-2721/2/20220568

frames are inputted to 3D humanoids in sequence, which defines the standing posture of virtual body. These sequences also define the time stamps, i.e. at what period of time, avatar will be at what position.

When these animation frames are placed sequentially, then rendering software convert them into sign animation. This animation speed is also adjustable. Many avatars have been designed for this purpose as shown in **Figure 3**. "Anna", "Luna" and "Marc" are three basic avatars whereas many others also available. A comparison of gloss text, HamNoSys, corresponding SiGML and avatar animation is displayed below.

Figure 3. 3D avatars.

4. Methodology

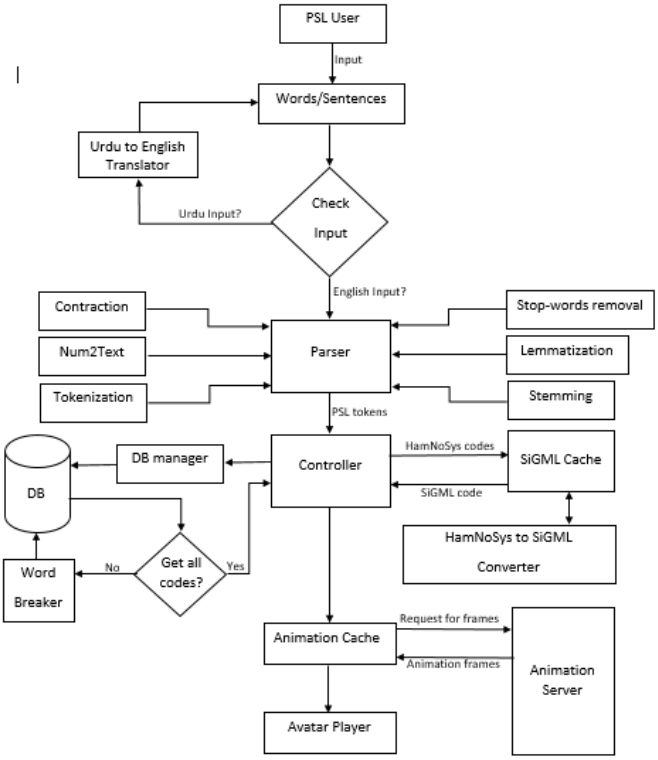

Figure 4. Methodology of PakParse.

This work is about machine translation of English/Urdu text into Pakistan Sign Language (PakParse) through 3D avatar sign animation It takes input in textual form (English/Urdu) up to 140 characters, translator is used in case of Urdu input. Then this English sentence is inputted to a parser which is

DOI: 10.54254/2755-2721/2/20220568

responsible to convert this English sentence into PSL sentence. For this conversion, preprocessing techniques are used which includes tokenization, stop words removal, stemming and lemmatization.

This preprocessing will generate related root words of sentence. Then its corresponding HamNoSys code from the database will be extracted. As database contains 2000 basic signs, the words having no specific sign in database will be break down into English characters so that they can finger spelled at the end. Then this HamNoSys code is converted into SiGML through a built in converter, which is helpful to generate animation frames for that sentence. These animation frames are played by 3D avatar at the end.

5. Result & Discussions

This system (PakParse) offers an online bilingual dictionary of Pakistan Sign Language (PSL). It is based on techniques such as HamNoSys for the generation of PSL signs, SiGML as XML for animation and JA SiGML URL APP to produce animation frames for 3D avatar. This proposed system creates sign animations for 2000 basic alphabet, number & words. The input limit is set to 140 characters to maintain the efficiency of the system. Though it contains 2000 basic sign dataset only, but those words having no sign in database will be breakdown into English characters and finger spelled by avatar at the end.

This system can be used by deaf or normal people at public places such as airports, railway stations, parks & banks. PakParse can also be used for education of deaf children and for learning Pakistan Sign Language.

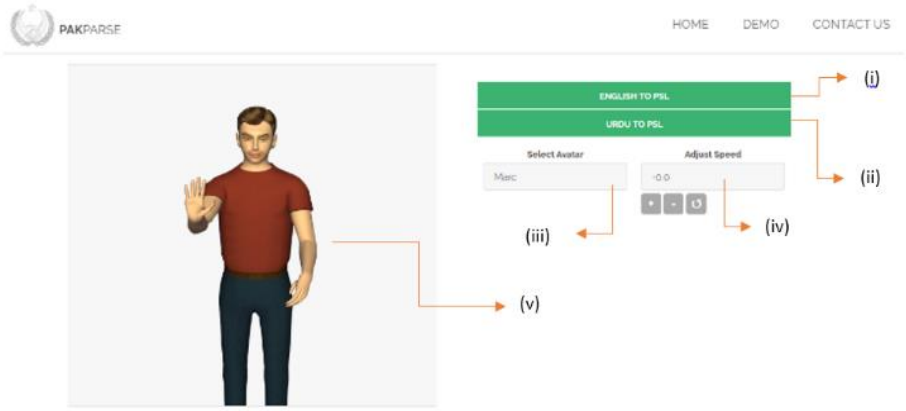

Figure 5. Screenshot of PakParse.

1. Textbox to enter text in English.
2. Textbox to enter text in Urdu.
3. Avatar selection menu
4. Adjust speed of avatar while signing.
5. Drag the avatar to change view of camera and zoom.

5.1. Evaluation

PakParse is evaluated by a team of deaf school in Faisalabad-Pakistan, consisting of five PSL experts and five students. Students were asked to get the meanings of signs performed by 3D avatar. And language experts were asked to check the accuracy level of avatar signs. 10 sentences were tested by participants. Accuracy of the system is 90.5% as per formula.

DOI: 10.54254/2755-2721/2/20220568

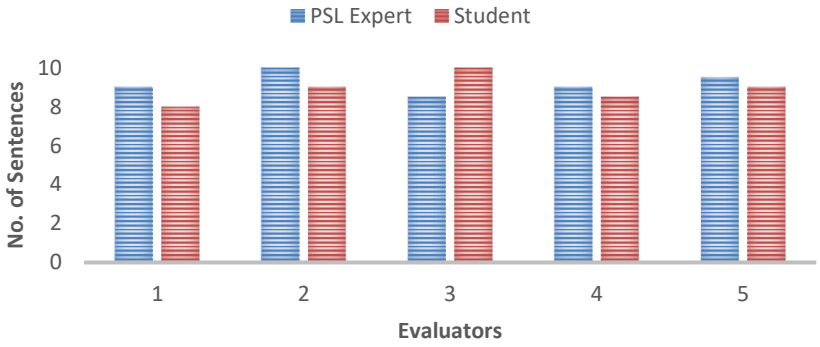

Figure 6. Evaluation of PakParse.

6. Conclusion

Deaf and mute people around the world, or people who take birth in such families have to learn sign language to communicate. They learn SL as their first language while other languages become their second choice. These people try to access information and knowledge in the form of SL only. There are multiple ways to facilitate them but many of them are expensive like video clips of signs. The recording, managing and storing of video clips is expensive in terms of cost, time and space. So sign animations through avatars is a promising approach. Pakistan recently got its separate Sign Language and as it's still in its development phase, there was no machine translator for it. That's why PakParse is proposed in this thesis which is a sign language generation system for PSL. It generates signs for English & Urdu languages according to PSL. To perform this, it takes sentence in any of above mentioned languages, and generate its root words with the help of preprocessing techniques. These root words are called PSL words which have their HamNoSys codes in database. This code is fetched and converted into XML form which is called SiGML. Then this SiGML is inputted to humanoid avatars available in JA SiGML URL APP which generate corresponding animations frames through a software and avatar animates according to it.

6.1. Future Scope

Database of PakParse has almost 2000 HamNoSys codes which can be extended to cover the whole PSL language. The system is fully scalable of it. HamNoSys generation system covers only two modules of HamNoSys code; handshape and hand orientation, which can be extended to cover all modules of HamNoSys. This system doesn't cover non-manual features which can also added in future to give a more realistic look.

References

[1] K. Nabeel, A. Sabir, A. Adnan, F. Kamran, F. Uzma, S. Muhammad and H. Jameel, "Speak Pakistan: Challenges in developing Pakistan sign language using information technology," South Asian Studies, vol. 2, p. 367, 2015.

[2] Marshall and I. Eva Safar, "The Architecture of an English Text-to Sign Languages Translation System," in Recent Advances in Natural Language Processing (RANLP), Tzigov Chark Bulgaria, 2001, pp. 223-228.

[3] Hassan, F. Bilal, A. Muhammad Shoaib, S. Adnan and Nabeel, "Pakistan Sign Language: Computer Vision Analysis & Recommendations," VFAST Transactions on Software Engineering, vol. 9, pp. 1-6, 2015.

[4] A. Aleem, K. Azhar, B. U. M Yousuf, M. Mehmood, R. Suleman, R. Sameer, A. Razi Ur and Israr, "Pakistan sign language recognition using statistical template matching," International

DOI: 10.54254/2755-2721/2/20220568

Journal of Information Technology, vol. 1, pp. 1-12, 2004.

[5] A. F. Sevilla, A. D. Esteban and J. M. Lahoz-Bengoechea, "Tools for the Use of SignWriting as a Language Resource," in Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives, 2020, pp. 203-208.

[6] Sulman, Z. Dr Nasir and Sadaf, "Pakistan sign language--a synopsis," Pakistan, June, 2000.

[7] Y. Rabia, K. Ahmet Alp and H. Kose, "Avatar-based Sign Language Training Interface for Primary School Education," in Workshop: Graphical and Robotic Embodied Agents for Therapeutic Systems, 2016.

[8] Dewani, B. Amirita, M. Sania, A. Mohsin, A. Wajiha, Z. Quratulain and B. Sayyid, "Sign Language e-Learning system for hearing-impaired community of Pakistan," International Journal of Information Technology, vol. 10, pp. 225-232, 2018.

[9] N. S. Khan, A. Abid and K. Abid, "A Novel Natural Language Processing (NLP)--Based Machine Translation Model for English to Pakistan Sign Language Translation," Cognitive Computation, pp. 1-18.

[10] Z. Liwei, Kipper, K. Schuler, W. Vogler, C. Badler, N. Palmer and Martha, "A machine translation system from English to American Sign Language," in Springer, 2000.

[11] B. Andrew, J. C. Stephen, R. Elliott, G. John, I. Marshall and S. Rankov, "Virtual signing: Capture, animation, storage and transmission-an overview of the visicast project," in IET, 2000.

[12] V. Tony, C. Alan and Collins, "The challenges of cross-modal translation: English-to-Sign-Language translation in the Zardoz system," Machine Translation, vol. 13, pp. 81-106, 1998.

[13] A. Aleem Khalid, B. A. M Yousuf, U. Mehmood, M. Suleman, R. Sameer, U. R. Razi and I. Ahmed, "Pakistan sign language recognition using statistical template matching," International Journal of Information Technology, vol. 1, pp. 1-12, 2004.

[14] Parkhurst, P. Steve and Dianne, SignWriting: A Complete System for Writing and Reading Signed Languages, SignWriting Press, 2007.

[15] Stokoe, C. William C, C. Dorothy C and C. G, A dictionary of American Sign Language on linguistic principles, Linstok Press, 1976.

[16] Thomas and Hanke, "HamNoSys-representing sign language data in language resources and language processing contexts," in LREC, 2004, pp. 1-6.

[17] K. Rupinder and K. Parteek, "HamNoSys generation system for sign language," in 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2014, pp. 2727-2734.