Speak Pakistan: Challenges in Developing Pakistan Sign Language using Information Technology

Nabeel Sabir Khan
University of Management and Technology, Lahore, Pakistan.
Adnan Abid
University of Management and Technology, Lahore, Pakistan.
Kamran Abid
University of the Punjab, Lahore, Pakistan.
Uzma Farooq
University of Management and Technology, Lahore, Pakistan.
Muhammad Shoaib Farooq
University of Management and Technology, Lahore, Pakistan.
Hamza Jameel
University of Management and Technology, Lahore, Pakistan.

Abstract

Gesture based communication called Sign Language (SL) is the fundamental communication channel between hard of hearing individuals. Communication through signing is a visual motion dialect. Hard of hearing individuals use gesture based communication as their primary medium for correspondence. Different countries have their own sign language as the United States of America has American Sign Language (ASL), China has Chinese Sign Language (CSL), India has Indian Sign Language (ISL), and similarly Pakistan has Pakistan Sign Language (PSL). Most of the developed nations have addressed the issues of their hearing impaired people by launching projects involving Information Technology to reduce this gap between a deaf and a normal person. In central and south Asia, a considerable work has been conducted on ISL and CSL. However, Pakistan Sign Language is a linguistically under-investigated in the absence of any structured information about the language contents, grammar, and tools and services for communication. Hence, the major contributions of this research are to highlight the challenges to bridge this communication gap for Pakistani deaf community by using the existing literature, and to propose an Information Technology based architectural framework to identify major components to build applications which may help bridging the gap between the deaf and normal people of the country.

Introduction

The most common and useful way of communication among human is speech but a large number of population in the world suffer from hearing/speech disability. So, there is a huge communication gap between these disabled and normal people. To bridge this gap, a language which is known as sign language exists. Sign language is comprised of gestures or visual representations of several different types.

Spoken languages vary region to region and about 6,909 (Linguistic Society of America, 2015) spoken languages exist in the world till now. Similarly, the languages of gestures (sign languages) vary from region to region, and about 138 (Pakistan Sign language, 2015) sign languages are known till today. Among them American Sign Language (ASL) and British Sign Language (BSL) are based on English language. Whereas, Indian Sign Language (ISL), and Chinese Sign Language (CSL) are also among the well-known sign languages. The grammars of these gesture based sign languages differ from grammars of spoken or written languages. The reason is that gesture based languages involve shapes and concepts, whereas spoken and written languages involve words and grammar rules, thus, both types of languages have significantly different grammatical structures (Debevc et al., 2014) (Marschark et al., 2004) .

The field of Information Technology (IT) is strongly influencing human life. Several different tools, technologies and devices have been built to help mankind resolve different problems. Similarly, people have worked on bridging this gap between the deaf and normal person by involving IT. The idea behind such IT tools and services is to enable the deaf to communicate with a normal person and vice versa. There can be numerous scenarios where such services can be useful to minimize or eliminate this communication gap.

Motivational Example: Consider a deaf person who wants to read an online newspaper written in normal English/Urdu language. He would not be able to do so, as he does not understand the grammatical structure of English/Urdu language. However, if the same is shown to him using the gestures in respective sign language, he will be able to understand that very easily. Creating an application that converts the written text to sign language and in turn this sign language to avatar performing the gestures can resolve this issue.

The rest of the article has been presented in the following manner: the next section explains the general concepts related to the sign language. This is followed by the challenges identified in the light of current state-of-the-art to enable Pakistani deaf community to interconnect with the normal people by realizing the scenarios like the one presented in the motivational example. The major components of an Information Technology infrastructure to bridge this gap have

been presented in the next section. Lastly, we present the conclusion and possible future directions for this research.

Concepts Involved in the Sign Language

A sign language uses manual communication and body language to convey meaning, as opposed to acoustically conveyed sound patterns (Sign Language, 2015) and to communicate with deaf people use signs. Each particular sign represents a distinct letter, word or phrase of the corresponding spoken language e.g. for the word "What" the sign in different Sign languages is shown in Figure 1.
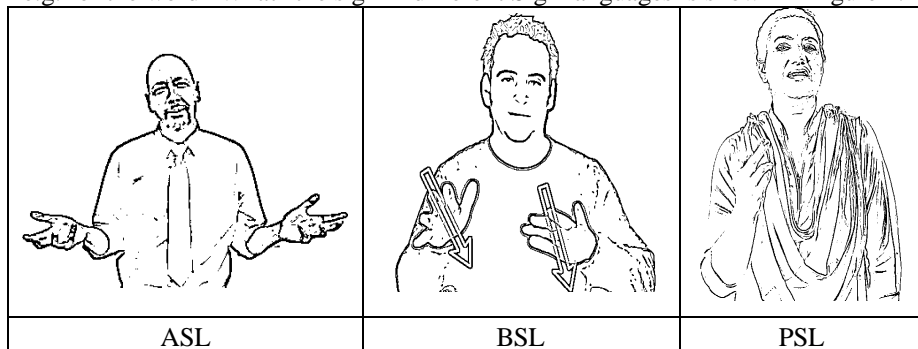
| ASL | BSL | PSL |

Figure 1: Sign of "What" in different sign languages.

Gestures

Sign languages uses gestures to make a sign for particular unit e.g. letter, word or phrase. These gestures are further decomposed into manual gestures and non-manual gestures. Manual gestures consist of hand shape, movement, location (Hall et al., 2015), (Al Qodri et al., 2012), and orientation as shown in Figure 2, whereas non-manual gestures consist of facial expression, head movement, posture and orientation of body (Al Qodri et al., 2012), shoulder raising, and mouthing, as shown in Figure 2. Mostly non-manual markers are used along with manual markers.
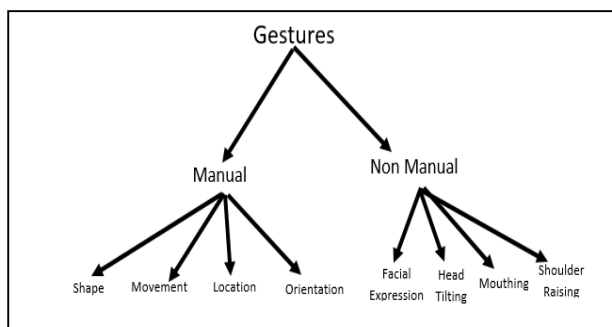
Figure 2: Components of Gestures

Manual gestures have two attributes *hands* and *dynamism* as shown in Figure 3. *Hands* involves the number of hands participating in performing the gesture, whereas dynamism has two possible instances namely, static and dynamic signs. Where, the type of signs which include constant movement of hands i.e. sign is performed in a flow is known as static sign. Whereas, the dynamic type of signs include variable movements of manual and non-manual markers i.e. sign is performed by combination of two or more signs. Therefore, a manual gesture can be single handed static, or single handed dynamic. Similarly, it can be double handed static or double handed dynamic in nature.
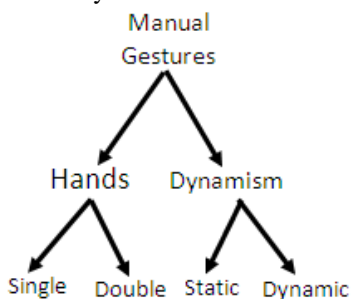

Figure 3: Attributes of Manual Features

Sign Writing Notations

Like spoken languages, Sign Languages can also be written down with the help of Sign Writing Notation Systems. Different notation systems are present for the representation of signs in Sign Language but no notation for sign languages is considered as standard till to date. The main advantages of using sign language notation systems are

- They are helpful in representing the words of the natural Language to a format that can be used later in the translation of text to sign language *animations*.
- They make the translation system scalable.
- Storage space

There are many notation systems used for Sign language writing among which the four most widely used Sign Writing Notation Systems are Stokoe, Gloss, SignWriting, and HamNoSys (Hutchinson, 2012). The basic representation of widely used sign writing notation symbols are shown in Figure 4.

aɒʒdɛʇⱦꞣↄꞩꞰ 𝑇𝑇 ∪Ꝺ ŋ ⊙ ꝺ ᴎ ɑ ⊤ ∧ ∨ ꞷ z ꙩ 𝟇)(_ ⱶ ⱦ [] □ ÷ 5 A B C G L R V X Y	(image of gloss notation text)
(a) Stokoe	**(b) Gloss**
(image of SignWriting symbols)	(image of HamNoSys symbols)
(c) SignWriting	**(d) HamNoSys**

Figure 4: Widely used sign writing notation symbols

The comparison between all these notations is provided in Table 2. The values in the table clearly reflect that among all these four notations HamNoSys is the most suitable choice because it provides the following advantages over all other notations.

It is not dependent on any sign language, so we can represent any sign language gesture using this notation. It can represent both manual and non-manual features of a particular sign. It is used for both *academic* and *research* purposes. Its representation is linear so instead of storing pictures we can store sign language gestures in textual format which helps us to minimize the space complexity. It can be represented in both ASCII and Unicode so it is easy to represent and store gestures in computer. So we take HamNoSys as a standard sign writing Notation in rest of the article.

Table 1: Comparison of widely used Sign Writing Notations

Sign Writing Notation System	Sign Language Dependant	Non-Manual Features Support	Objective	Arrangement	Computer Compatibility
Stokoe	Yes	No	Dictionary or Academic	Linear	Custom Font or ASCII codes
GLOSS	Yes	Yes	Academic	Linear	Custom Font or ASCII codes
SignWriting	No	Some	Public Use	Pictorial	ASCII or Unicode
HamNoSys	No	Yes	Academic	Linear	Custom Font or Unicode

HamNoSys Sign Writing Notation System

It is known as Hamburg Sign Language (HamNoSys) notation system introduced by the University of Hamburg in Germany in 1985 (Sign Language Phonology, 2015). It has its own predefined notations and phonetic transcriptions for the

definition of signs and gestures shown in Table1(c). It provides us a way to write signs in a computer understandable format which is easy to interpret and process. The origin of HamNoSys was basically Stokoe writing notation system and it gives us an alphabetic system to define different sign language parameters like hand-shape, hand-movement, hand-location and hand-orientation (Symbol Font for ASL, 2015).

HamNoSys has four basic components including three sub-components as shown in Figure 5. The components shown in solid boxes are mandatory components for the representation of Signs in HamNoSys which are Initial Configuration and Action/Movement. The initial Configuration component comprises of Handshape, location and orientation. The attributes shown in boxes with dotted border are optional that are Symmetry Operator and Non-Manual Features. From the Figure 5 we can easily conclude that HamNoSys notation has the capability to represent all components of gestures manual and non-manual as described in Fig2. The Initial configuration component in Fig5 can represent all manual gesture attributes including hand shape, movement, location and orientation. The non-manual feature component can represent facial expression, head tilting, mouthing and shoulder raising. The symmetry component is used to represent whether the gesture is single handed or double handed as explained in Figure3. The last component of HamNoSys is used to represent Dynamism of the gesture i.e. whether the gesture is static or dynamic.
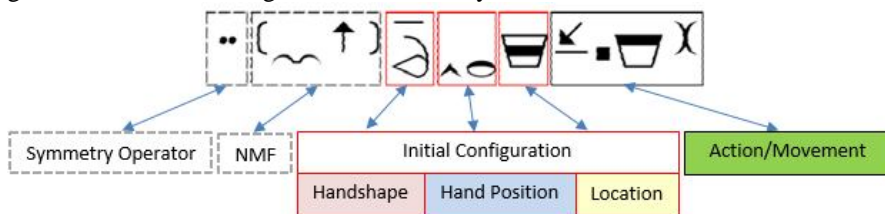

Figure 5: HamNoSys Components of Sign Gesture

Current State-of-the-art and Challenges

There are more than hundred sign languages in the world today. Generally, every country has its own sign language e.g. American, British, Japanese, Indian sign languages exist. Similarly, Pakistani sign language is called Pakistan Sign Language (PSL). According to an estimate by World Health Organization over 5% of the world's population which is more than 360 million people have disabling hearing loss, in which 328 million are adults and 32 million are children (World Health Organization, 2015). A significant part of the deaf population is young and sign language recognition system can turn them into useful human resources for certain positions. Whereas, data given by Population Census Organization of Pakistan more than 3.3 million people of the country are disabled, and among them 0.25 million suffer from hearing loss, that counts to 7.4% of the overall disabled population (Population Census Organization, 2015).

The research work related to Sign Language gestures detection has been done in many different ways. The approaches for gesture recognition are either hardware-based which use data gloves, Kinect, or other sensor based devices (Mohandes et al., 2007), or they are based on computer vision approaches which use digital camera and image processing algorithms (Rashid et al., 2009) (Khan et al., 2014). Some elementary work related to Pakistan sign language has also been conducted in both directions. There is a system named "Boltay Hath" that aims at recognizing Pakistan Sign Language using data gloves as its interface (Alvi et al., 2004). Similarly, a vision based approach to recognize Pakistan Sign Language alphabets has been presented by (Khan et al., 2014).

Machine translation has recently gained popularity and is being widely used to convert natural language (NL) text to a given sign language. An early work in this regard was conducted by Grieve-Smith (Grieve-Smith, 2002). Similarly, a linguistic analysis for the possible issues that may occur during machine translation have been discussed by (Speeers, 2002). A grammatical approach based on synchronous tree adjoining grammar has been proposed by (Zhoa et al., 2000), which has been further enhanced by (Huenerfauth, 2004). Whereas (Zahoa et al., 2000) presents English to ASL translation approach using tree adjoining grammar rules. Another dimension of machine translation involves statistical machine translation of sign languages, some work in German sign language using statistical machine translation has been presented in (Suszczanska et al., 2002). Likewise, example-based translation is another variant of translating sign language to natural language, (Bungeroth et al., 2004) presents such translation for Irish sign language. Similarly in South Africa a project South African Sign Language Machine Translation (SASL-MT) has been conducted to enable the deaf community of the country with the help of a machine translation system from English to SASL (Van Zijl et al., 2003).

It is clear from the literature review that people and governments of many different countries have worked in many different facets to enable their hearing impaired population communicate with the normal people. Translation from sign language to natural language and vice versa has been the core idea which has been implemented in many ways. Unfortunately, no significant work has been done for Pakistan sign language in this regard, and there is a great room for conducting research in various levels. Based on the approaches discussed earlier we have laid down the following challenges that should be addressed to help Pakistani deaf community communicate and use sources of information.

- Lack of availability of linguistic information. PSL has not been linguistically investigated properly.
- Absence of Standard Sign corpus based on different language granularity units.
- No standard grammar rules for sentence creation in PSL.
- No sign writing notation exists for PSL.

- Automating it all requires evaluation and no evaluation corpus to test the systems exist.

Proposed Framework

This research presenting a framework which will centralize all standardized gestures, their equivalent *HamNoSys*, grammar rules of PSL and then using these rules we will convert Urdu/English text to sign animations using an avatar. An architecture for the proposed framework has been presented in Figure 6. Here, we have presented all major components of the system and their interaction with each other. The diagram shows that there are different layers in the system including Storage Layer, Middleware Layer, and Application Layer.
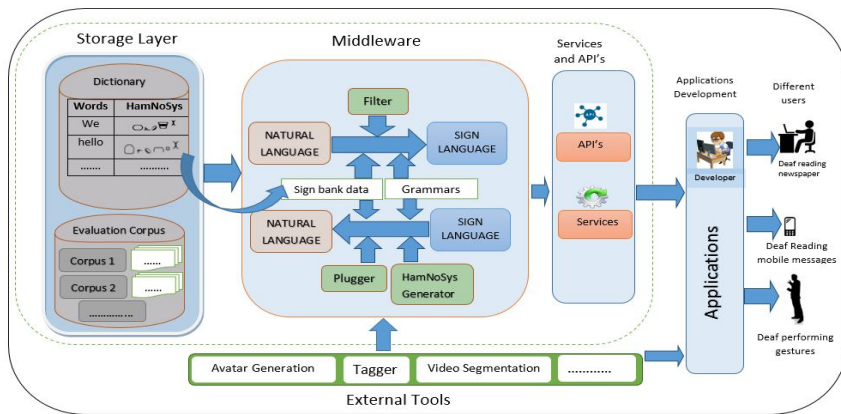

Figure 6: Architectural Framework

Components and services

The system is divided into following major components.
- Storage
- Middleware
- Services and API's

Storage Component involves the following two sub-components:
- *Standard Sign Bank*
- *Evaluation Corpus*

Standard Sign Bank: In order to make the translation possible from text to sign language or vice versa we need a corpus containing gestures of all the words along with their *HamNoSys* representation. Like natural languages the sign language also varies from region to region so same word has different gestures in PSL. We will store a word along with its all possible *HamNoSys* representations. We will make one of the *HamNoSys* as a standard of that particular word. For this

standardization purpose we will consult Pakistan sign language experts and interpreters so that this standard sign bank can be accepted globally among all PSL researchers and developers so that services and applications can be constructed that will be accepted by all deaf community of the country. We also ensure that while making this data bank the granularity of data units i.e. letters, words and phrases must be incorporated with the consent of PSL experts. This standard corpus also help us while translating from sign to text because if person use non-standard gesture during his communication still our system is capable to map that gesture to appropriate *HamNoSys*. It is pertinent to mention that we are not storing the images or any animation for the sign. But we are storing a digitized format of a gesture known as *HamNoSys*. This makes our system storage efficient, comprehensive. Furthermore, it also supports the cause of translation from sign language to natural language, and vice versa.

Evaluation Corpus: Research needs to be evaluated and such evaluation requires tests. This invites us to generate several gold standard corpuses for testing and evaluating all services/tools that we intend to develop. The Corpus contains sentences of all possible categories of the language along with their correct translations according to the rules of the grammar so that accuracy of provided services/tools can be measured and results can be improved.

Middleware Layer

This layer is the core of the whole framework. It consists of the following components: -

- Language Translator
    - o Natural Language to Pakistan Sign language Translation
    - o Pakistan Sign language to Natural Language Translation
- Grammar
    - o PSL Grammar
    - o Natural Language Grammar (Urdu, English) Plug-in based
- Sentence Manipulator
    - o Filter (Stop Word removal/Stemmer/Lemmatizer)
    - o Plugger (Add missing words)
- Video to *HamNoSys* Generator

The language translator module consists of two sub modules, first Natural Language (NL) → Sign Language (SL) converter, which converts text to sign language animation, and the other SL→NL which converts the video of SL to NL text. These sub components have been explained below.

Natural Language to Pakistan Sign language Translation

This component is responsible for translating the English/Urdu sentence to equivalent sign language sentence. This module takes sentence as input and using external service of tagger performs the morphological analysis of the sentence and

converts it into lexical units. Then it communicates with the NL grammar component to verify the sentence structure and generate parse trees. This generated parse tree is then fed to filter sub component of sentence manipulator which removes stop words like *a, an, the,* and other prepositions. This filtered tree becomes the input of PSL grammar, and then this module converts the NL filtered sentence into equivalent PSL sentence. This PSL sentence communicates with the Sign bank to generate *HamNoSys* regarding to each tagged lexical unit. In the end the generated *HamNoSys* will use the services of external service of Avatar to generate the Sign animations of the input sentence.

Pakistan Sign language to Natural Language Translation: To make two way communication possible this module will take video as input. The video is passed to external service of video processing which will preprocess the video and segment all gestures available in the video. The segmented video is passed to Video to *HamNoSys* generator which generates the corresponding *HamNoSys* of each segmented gesture. After this the corresponding words against each gesture are fed to *plugger* module which add missing words according to grammar of Natural language using certain algorithms and then SL→NL module generates appropriate sentence.

PSL Grammar: The grammar module is also subdivided into two sub modules. PSL grammar and NL grammar of English and Urdu. Grammar is the building block of any language's sentence structure. Every spoken language has some sort of grammatical structure for their sentence formation. Other than an important component, grammatical structure helps in verifying the syntax of the respective sentence.

Like all Sign Languages of the world PSL has its own grammatical rules for the construction of valid PSL sentence. This PSL grammar module is used by NL→PSL converter to convert the NL sentence into its equivalent PSL sentence.

NL grammar of English and Urdu: This sub component will be implemented as a plug in for Urdu and English languages. The major task of this module is to check the validity of Natural language sentence. As it will be a plug in we can replace it with any other language if the grammar of that language exists and it can also work for our regional languages like Punjabi, Pashtu etc.

In order to understand the differences between sentences structures of PSL and English consider the following examples shown in Table 2.

Table 2: Comparison of the structure of English sentence with PSL equivalents.

English sentence and structure	PSL sentence and structure
I am from Lahore	I from Lahore I from Lahore I From Lahore I
I am a teacher	I teacher I I teacher Teacher I

Variations in different PSL sentence formats makes it obvious that Grammar is the most important module for the accurate conversion from one language to another. To the best of our knowledge no such grammar exists for PSL.

Sentence Manipulator: This component is used to transform natural language sentence to Pakistan sign language sentence and vice versa. This in turn involves two sub-components namely Filter and Plugger. This reads the tree of natural language sentence and remove stop words and other un necessary details from the sentence that are not used by deaf people during their English/Urdu reading or writing. Whereas, the Plugger is used from Pakistan sign language sentence to natural language sentence. It will use certain algorithmic techniques to add the missing information in PSL sentence and transform in to equivalent NL sentence.

Video to HamNoSys Generator (VHG): The video processing service tracks and segments all the gestures in the sign language video. The gestures are then given to VHG that identifies the handshape, orientation, palm location and movements and maps these features to appropriate HamNoSys representation. This HamNoSys will then be used to generate words and NL grammar along with plugger converts those in valid NL sentences.

External Services: There are certain services that are external to the system:

Tagger**:** The Language Translator uses tagger service to break the sentence into its morphological structure, and helps tagging the input sentence to the parts of speech. This tagged result is further used in the grammar component to perform syntactic and semantic analysis on the input.

Video Segmentation**:** The Sign to NL module uses this service to segment the input video into different segments based on the gesture identification. That is, it will create a separate video segment for each identified gesture which will be processed further by VHG to generate HamNoSys.

Avatar Generation: This service shall be used while converting text/audio to sign language conversion. It would take the HamNoSys of tagged words from the sign bank, and then by using this HamNoSys it would generate avatar for each HamNoSys.

Services and APIs: The proposed middleware along with external services can be used by developers to develop certain applications for the deaf people, for instance, newspaper reading, mobile messaging reading, and writing an email etc. Similarly these applications can be used by deaf community to bridge the communication gap and get better job opportunities by minimizing the communication gap.

Conclusion and Future Work

In this research we have covered a literature review of the work done by different countries to enable their hearing impaired population communicate and to help them access the information in many different ways. Certainly, the usage of Information Technology cannot be denied in achieving such milestones. The

research analysis shows various different dimensions in which people have been working to resolve these issues. The unfortunate part is that no significant work exists for Pakistan sign language. We have highlighted this gap and possible challenges which should be addressed to help Pakistani deaf community. Apart from emphasizing upon the challenges, as another principal contribution of this research we have also proposed a general architectural framework which can help translating English or Urdu text/voice to animations of Pakistan sign language, and vice versa.

In future, we intend to address all these challenges by implementing all different components in the proposed architectural framework. To this end, we intend to start with the text to sign language translation, followed by defining and refining the grammatical structure for Pakistan sign language. We shall also develop a standard corpus for Pakistan sign language for all different granularity levels including letter, word, and phrase. We also plan to develop APIs and services for the developers and deaf community, respectively. Lastly, we shall develop evaluation corpus for the testing of all these services and tools for their effectiveness and accuracy.

References

Al Qodri Maarif, H., Akmeliawati, R., & Bilal, S. (2012, July). Malaysian Sign Language database for research. In Computer and Communication Engineering (ICCCE), 2012 International Conference on (pp. 798-801). IEEE.

Alvi, A. K., Azhar, M. Y. B., Usman, M., Mumtaz, S., Rafiq, S., Rehman, R. U., & Ahmed, I. (2004). Pakistan sign language recognition using statistical template matching. International Journal of Information Technology, 1(1), 1-12.

Bungeroth, J., & Ney, H. (2004, May). Statistical sign language translation. In Workshop on representation and processing of sign languages, LREC (Vol. 4).

Debevc, M., Milošević, D., & Kožuh, I. (2014). A Comparison of Comprehension Processes in Sign Language Interpreter Videos with or without Captions. PloS one, 10(5), e0127577-e0127577.

Grieve-Smith, A. B. (2002). SignSynth: A sign language synthesis application using Web3D and Perl. In Gesture and Sign Language in Human-Computer Interaction (pp. 134-145). Springer Berlin Heidelberg.

Hall, M. L., Ferreira, V. S., & Mayberry, R. I. (2015). Syntactic Priming in American Sign Language. PloS one, 10(3), e0119611.

Huenerfauth, M. (2004, May). A multi-path architecture for machine translation of English text into American Sign language animation. In Proceedings of the Student Research Workshop at HLT-NAACL 2004 (pp. 25-30). Association for Computational Linguistics.

Hutchinson, J. (2012). Literature Review: Analysis of Sign Language Notations for Parsing in Machine Translation of SASL.

Khan, N., Shahzada, A., Ata, S., Abid, A., Khan, Y., & ShoaibFarooq, M. (2014). A Vision Based Approach for Pakistan Sign Language Alphabets Recognition. Pensee, 76(3).

Linguistic Society of America, http://www.linguisticsociety.org/content/how-many-languages-are-there-world (Retrieved on 6-August-2015)

Marschark, M., Sapere, P., Convertino, C., Seewagen, R., & Maltzen, H. (2004). Comprehension of sign language interpreting: Deciphering a complex task situation. Sign Language Studies, 4(4), 345-368.

Mohandes, M., & Buraiky, S. (2007). Automation of the Arabic sign language recognition using the powerglove. AIML Journal, 7(1), 41-46.

Pakistan Sign Language , http://psl.org.pk/faq/ (Retrieved on 10-July-2015)

Population Census Orgnaization, Govt. Of Pakistanat http://www.census.gov.pk/index.php(Retrieved on 12-June-2015)

Rashid, O., Al-Hamadi, A., Panning, A., & Michaelis, B. (2009). Posture recognition using combined statistical and geometrical feature vectors based on SVM

Sign Language, https://en.wikipedia.org/wiki/Sign_language (Retrieved on 2-June-2015)

Sign Language Phonology, http://www.ling.fju.edu.tw/phono/sign.htm (Retrieved on 25-May-2015)

Speers, D. (2002). Representation of American sign language for machine translation.

Suszczańska, N. I. N. A., Szmal, P. R. Z. E. M. Y. S. Ł. A. W., & Francik, J. A. R. O. S. Ł. A. W. (2002, February). Translating Polish texts into sign language in the TGT system. In 20th IASTED International Multi-Conference. Applied Informatics AI (pp. 282-287)

Symbol Font for ASL, http://aslfont.github.io/Symbol-Font-For-ASL/ways-to-write.html (Retrieved on 20-June-2015)

Van Zijl, L., & Barker, D. (2003, February). South african sign language machine translation system. In Proceedings of the 2nd international conference on Computer graphics, virtual Reality, visualisation and interaction in Africa (pp. 49-52). ACM.

World Health Organization, http://www.who.int (Retrieved 15-March-2015)

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N., & Palmer, M. (2000). A machine translation system from English to American Sign Language. In Envisioning machine translation in the information future (pp. 54-67). Springer Berlin Heidelberg.

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N., & Palmer, M. (2000). A machine translation system from English to American Sign Language. InEnvisioning machine translation in the information future (pp. 54-67). Springer Berlin Heidelberg.

Biographical Note

Nabeel Sabir Khan University of Management and Technology, Lahore, Pakistan.

Adnan Abid University of Management and Technology, Lahore, Pakistan.

Kamran Abid University of the Punjab, Lahore, Pakistan.

Uzma Farooq University of Management and Technology, Lahore, Pakistan.

Muhammad Shoaib Farooq University of Management and Technology, Lahore, Pakistan.

Hamza Jameel University of Management and Technology, Lahore, Pakistan.

————————————————