International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:12, No:8, 2018

Online Multilingual Dictionary Using Hamburg Notation for Avatar-Based Indian Sign Language Generation System

Sugandhi, Parteek Kumar, Sanmeet Kaur

***Abstract*—**Sign Language (SL) is used by deaf and other people who cannot speak but can hear or have a problem with spoken languages due to some disability. It is a visual gesture language that makes use of either one hand or both hands, arms, face, body to convey meanings and thoughts. SL automation system is an effective way which provides an interface to communicate with normal people using a computer. In this paper, an avatar based dictionary has been proposed for text to Indian Sign Language (ISL) generation system. This research work will also depict a literature review on SL corpus available for various SL s over the years. For ISL generation system, a written form of SL is required and there are certain techniques available for writing the SL. The system uses Hamburg sign language Notation System (HamNoSys) and Signing Gesture Mark-up Language (SiGML) for ISL generation. It is developed in PHP using Web Graphics Library (WebGL) technology for 3D avatar animation. A multilingual ISL dictionary is developed using HamNoSys for both English and Hindi Language. This dictionary will be used as a database to associate signs with words or phrases of a spoken language. It provides an interface for admin panel to manage the dictionary, i.e., modification, addition, or deletion of a word. Through this interface, HamNoSys can be developed and stored in a database and these notations can be converted into its corresponding SiGML file manually. The system takes natural language input sentence in English and Hindi language and generate 3D sign animation using an avatar. SL generation systems have potential applications in many domains such as healthcare sector, media, educational institutes, commercial sectors, transportation services etc. This research work will help the researchers to understand various techniques used for writing SL and generation of Sign Language systems.

***Keywords*—**Avatar, dictionary, HamNoSys, hearing-impaired, Indian Sign Language, sign language.

I. Introduction

SL is a gestural language used by hearing-impaired people for their communication and education. It is not a unique language as it varies from region to region. SL includes facial expressions, head or body gestures, and eye gaze as non-manual signs whereas manual signs include hand orientation, shape, location and movement. Each word is represented as a different sign.

There are various types of SL as the spoken languages; they also differ from country to country. It is a complete natural language. Each of these SL has different signs to represent a word or sentence. According to the Ethnologue Languages of the World, there are more than 200 SL which are used by deaf communities all over the world [1]. These are Chinese Sign Language (CSL), Japanese Sign Language (JSL), Spanish Sign Language (LSE), British Sign Language (BSL), Australian Sign Language (Auslan), ISL, German Sign Language (DGS) etc. Fig. 1 depicts some of various types of SL.

Fig. 1 Various Types of SL

A. ISL

There are 12.3 million people who use ISL as a medium of communication [2]. Only 2% of deaf people get any education. ISL in India varies to some extend in different parts of country like Hindi language spoken in Delhi and Mumbai. Research on ISL indicates that it is very similar all over India [4]. Some educational institutes have started teaching through ISL like deaf schools at Indore, Patiala Deaf School, Bhagat Puran Singh School for Deaf at Amritsar, Ali Yavar Jung National Institute for the Hearing Handicapped (NIHH) at Mumbai etc. [3]. Moreover, National Council for Education, Research and Training (NCERT) came forward for giving education to students through ISL [4].

B. Features of ISL

There are various features of ISL described as follows.

- ISL uses non-manual component to indicate question, negation and suggestion.
- It uses spatial locations for representing a directional verb and anaphora.
- It represents the sign of stemmed word i.e. root word.
- For each alphabet and number, it represents its finger-spelling. For example, 10 in ISL is represented by finger-spelling of 1 followed by finger-spelling of 0.
- In ISL, name of a person is represented by his/her feature

Sugandhi, Parteek Kumar and Sanmeet Kaur are with the Department of Computer Science and Engineering, Thapar Institute of Engineering and Technology, Patiala, PB, 147002 India (e-mail: ssugandhi phd16@thapar.edu).

or finger-spelled.

- Past, Present and Future tenses are represented by sign at specific locations at the end of ISL sentence.
- Interrogative sentences like What, When, Where etc. are also represented by sign at the end of sentence.

C. Outline

The paper is divided into 8 sections. Section II describes the literature survey with the need of proposed system and Section III provides the detail of data collection for processing of ISL dictionary. Section IV explains the techniques used for creation of ISL dictionary whereas Section V includes the methodlogy of the system. Section VI describes the components of ISL dictionary and Section VII depicts the outcome of the system. Conclusion and Future scope of this research work is presented in Section VIII.

II. Literature Survey

There are various online multilingual dictionaries available for different SL. These are briefly discussed as follows.

Fuertes et al. (2006) had developed Spanish Sign Language-Spanish (DILSE) dictionary. It is a bilingual corpus of LSE, which is available electronically for the deaf communities. The interface two options for search, *i.e.*, Spanish and sign [5]. An online multilingual dictionary was created by Suzuki et al. (2006) for various sign languages which includes American Sign Language (ASL), Japense Sign Language (JSL) and Korean Sign Language (KSL). It consists 218 entries from JSL and 71 corresponding words from KSL [6]. A Danish Sign Language (DTS) dictionary was developed Kristoffersen and Troelsgrd (2010). The dictionary interface shows headword, video of sign, their synonyms and references [7], [8].

A BSL SignBank was developed for BSL by Cormier et al. (2011). It consists of 2,528 video clips of Deaf people using BSL. They have also collected the information about signers and the description of signs using annotation tool EUDICO Linguistic Annotator (ELAN) [9]. It can be publicly accessible and online available [10]. An ISL Dictionary was proposed by Martin et al. (2013) for disaster domain. The videos are traced to convert it into avatar animation. This dictionary provides the information of disaster to deaf. It consists of 600 sentences and 2000 words [11]. A DGS Corpus was developed by Thomas Hanke *et. al* (2016) using gloss and Hamburg Notation System (HamNoSys). It uses iLex annotation tool for both manual and non-manual signs. It includes about 6000 sign entries [12].

There are various corpus available for different sign languages but there is no such corpus exists for general domain in ISL, which can be accessible online and provides interface for avatar animation in ISL for each word or sentence. As there is lack of ISL interpreters so there is a need for a interface for learning and communicating in ISL. The propose system provides the online access to public so that they can learn ISL or communicate with others through this system.

III. Data Collection

The data for ISL Corpus has been collected from various sources. It has been taken from the books and videos provided by Bhagat Puran Singh School for Deaf. Each video is verified manually by the SL expert and any modification in the sign has been recorded. Further, the SL experts from Patiala School for Deaf provides the book and video for each sign and it is verified by them.

After the verification of data for every sign in ISL, the data is processed for creation of online multilingual dictionary for both English and Hindi language. For this purpose, HamNoSys is used which is explained in the next section.

IV. Techniques Used for Generation of ISL Dictionary

There are various techniques available for representing signs like Stokoe Notation [13], Sign Writing [14], and HamNoSys. The proposed system uses HamNoSys, SiGML, and avatar animation for ISL signs. These techniques are described as follows.

A. HamNoSys

An online multilingual ISL dictionary has been developed using HamNoSys notation. HamNoSys is developed by Prill-witz et al. in 1984. It has approximately 200 characters [15]. HamNoSys is a standard transcription and used internationally for all SL. It includes manual and non-manual components. It is a phonetic notation which consists of differ-ent parameters like configuration of hand, movement, location, head and body gesture described as follows.

1) *Handhsape:* For representing a sign, there are various handshapes symbols available in HamNoSys. Fig. 2 shows some of various types of handshapes and finger positions used in HamNoSys.

Fig. 2 Types of Handshapes

2) *Hand orientation:* The hand orientation includes the direction of fingers and palm. It can be left, right, upward, inward, downward or outward depending upon the position of signer/body. Fig. 3 shows the symbol used for hand orientation.

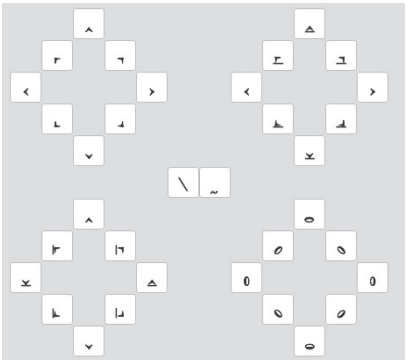

Fig. 3 Symbols used for Hand orientation

3) *Location:* The location of a sign can be defined using the location symbols like above head, head, shoulders, forehead, nose, mouth, lips, cheek, ears, neck, chest, stomach or below stomach. The position like left or right can also be defined. Further, the hands are touched or close to touch can also be shown using these symbols. Fig. 4 depicts the location symbols.

Fig. 4 Location symbols used in HamNoSys

4) *Hand Movement:* There are different types of directional and rotational movements present in HamNoSys. It defines the movement of hand from the initial location to the destination. The movement can be repeated one or more times. Fig. 5 shows the symbol used for movements.

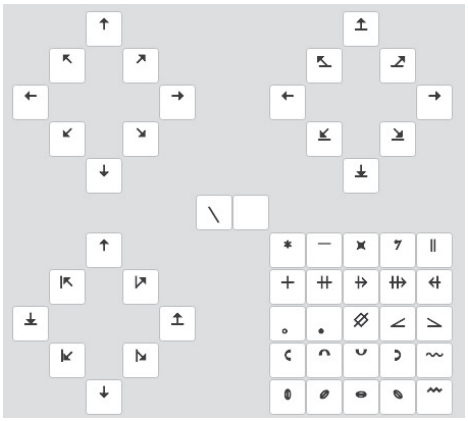

Fig. 5 Symbols used for movements

Fig. 6 depicts the finger and rotational hand movements used in HamNoSys.

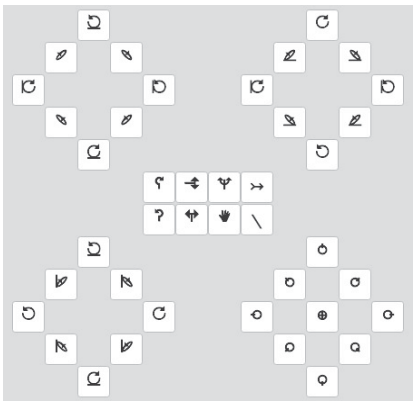

Fig. 6 Rotational movements

5) *Others:* It includes the symbols used for representing two-handed signs. Non-manual features like facial, body and head gestures are also included. Fig. 7 presents the symbol used for two-handed signs.

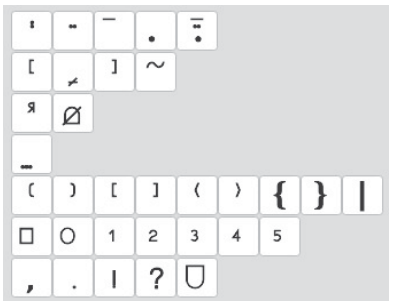

Fig. 7 Symbols used for two-handed signs

B. SiGML

SiGML was devel-oped at the University of East Anglia (UEA) for specifying signing sequences. It is like XML and uses different tags for representing their corresponding HamNoSys symbol. It is given as input to animation server for generating frames which defines the pose of avatar. Fig. 8 shows the SiGML of word Alone.

```
<sigml>
<hns_sign gloss="alone">
        <hamnosys_nonmanual>
        </hamnosys_nonmanual>
        <hamnosys_manual>
                <hamfinger2/>
                <hamthumbacrossmod/>
                <hamextfingeru/>
                <hampalmu/>
                <hamshouldertop/>
                <hamlrat/>
                <hamcircleu/>
                <hammoveo/>
                <hamsmallmod/>
        </hamnosys_manual>
</hns_sign>
</sigml>
```

Fig. 8 SiGML of *Alone*

C. Avatar

Avatar also known as virtual human was developed by Glauert et al. (2010) at University of East Anglia. It takes SiGML as input which is generated from HamNoSys notation and processes the animation frames. These frames generate 3D avatar animation. Fig. 9 shows the avatar animation of word *Alone.*

Fig. 9 Avatar animation of word *Alone*

V. Methodology

The workflow of proposed system is depicted in Fig. 10. The system accepts user input in two languages namely Hindi and English. The input tool uses Google Application Programming Interface (API) for Hindi transliteration. The input is processed through a parser and translated into an ISL sentence. For the translation process, the morphological information for each word of the sentence will be retrieved from the parsed output. Thereafter, stemming is applied on the basis of POS tag information to convert it into ISL sentence. After fetching the root words of the input sentence, its HamNoSys are extracted from the Database. Thereafter, its corresponding SiGML is looked up in the cache, where it is saved after generation. This SiGML is taken as input by animation server which uses Web Graphics Library (WebGL) for generating 3D avatar animation. The system can be accessed at URL *www.islfromtext.in* and and mobile application is also available at Google Play Store [16].

VI. Components of Dictionary

Different words along with their corresponding HamNoSys notation are stored in the database. To manage these words various functionalities are developed which are discussed as follows.

A. Database

The database statistics shows the number of words present in the dictionary for both English and Hindi language. It has different filters which can be used to find the words present in dictionary without its HamNoSys, corresponding Hindi word, or which are not verified. It consists of 2000 English and 3286 Hindi words. It also have 110 example sentences which can be accessed through the system interface.

B. Word Manager

Word Manager in the proposed system shows all the words present in the dictionary at a glance. The user can easily search the number, alphabet or word by using its initials or select the word from there. Fig. 11 shows the Word Manager of the proposed system.

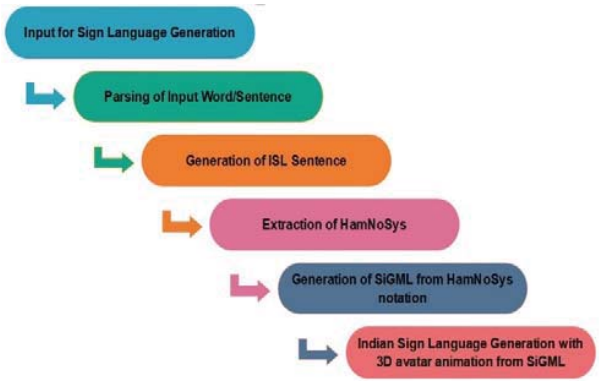

Fig. 10 Workflow of System

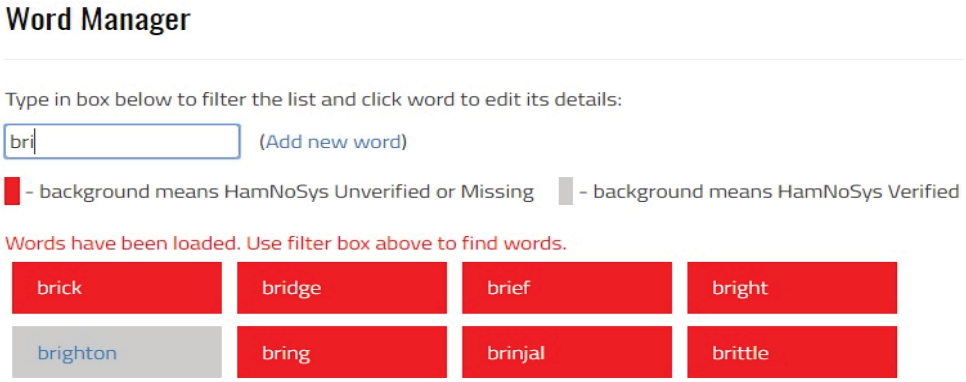

Fig. 11 Word Manager of Dictionary

C. Edit Panel

When a new word is created, its HamNoSys and corresponding Hindi word can be added by Edit panel. Each word stored in the database can be easily modified through Edit Panel. Fig. 12 depicts the Edit Panel for modifying word *Alone* in the dictionary.

D. HamNoSys Writing Tool

There is a interface for writing HamNoSys for each word. It provides the virtual keyboard for inserting symbols for representing ISL signs. Fig. 13 shows the HamNoSys virtual Keyboard.

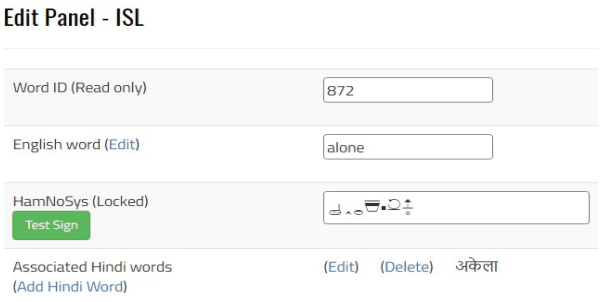

Fig. 12 Edit panel for modifying word *Alone*

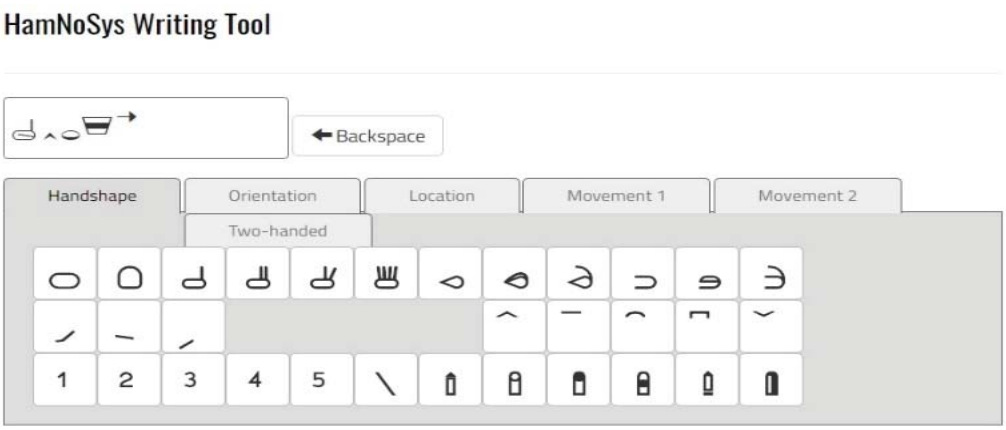

Fig. 13 HamNoSys Writing Tool

E. HamNoSys to SiGML Conversion Tool

The proposed system provides the feature of automatic conversion of HamNoSys to its corresponding SiGML. The user needs to define HamNoSys and its gloss, the output SiGML is generated automatically. Fig. 14 describes the HamNoSys to SiGMl conversion tool for word Alone.

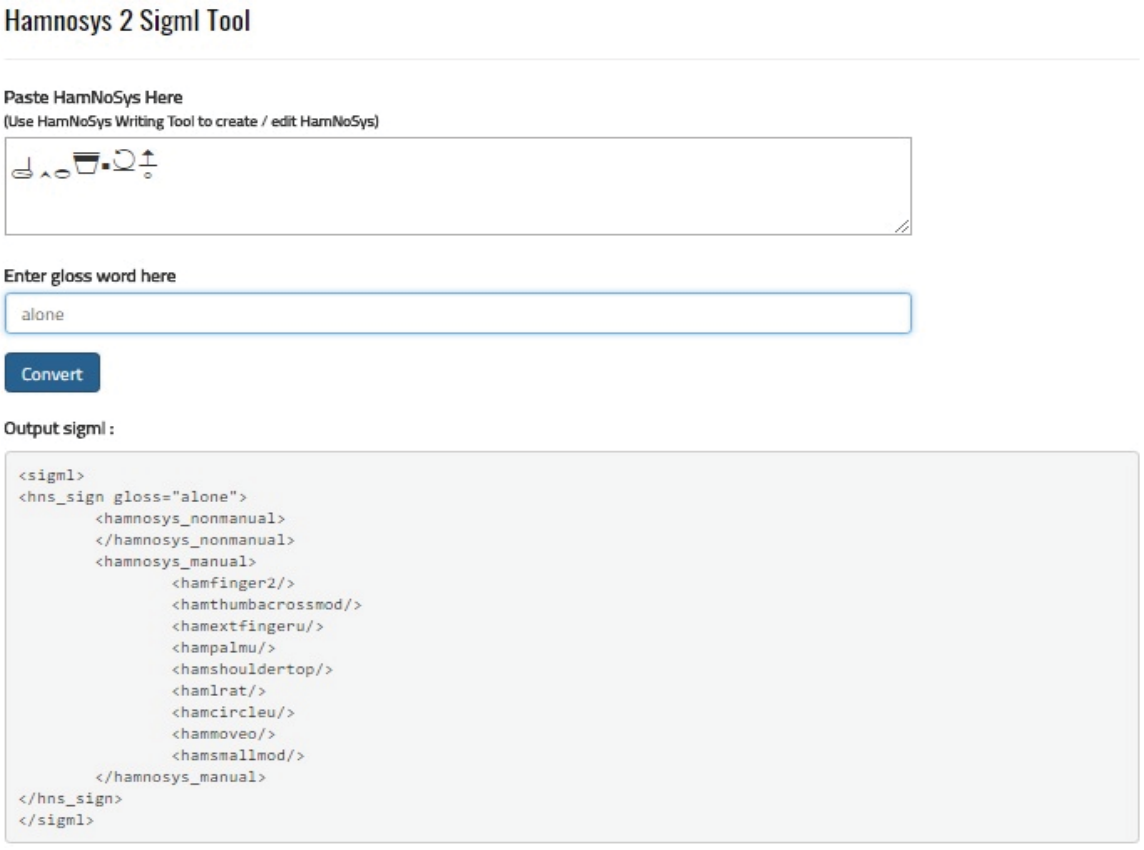

Fig. 14 HamNoSys to SiGML Conversion Tool

VII. Outcome

The system provides the online multilingual dictionary for ISL. It uses the techniques like HamNoSys for creation of dictionary and SiGML and Avatar for its processing. Table I depicts the processing of word Time for its ISL animation. The proposed system generate 3D avatar animation for any number, alphabet, word or sentences. It has 2000 English and 3286 Hindi words with 110 example sentences.

The proposed system can be used by the people for communication at public places like railway stations, banks, airport etc. It can also be used for education and learning of ISL.

TABLE I
DESCRIPTION OF ISL ANIMATION FOR WORD *TIME*

Word	HamNoSys	SiGML	Avatar Animation
Time		<sigml> <hns_sign gloss="time"> <hamnosys_nonmanual> </hamnosys_nonmanual> <hamnosys_manual> <hamparbegin/> <hamfinger2/> <hamthumbacrossmod/> <hamfingerbendmod/> <hamplus/> <hamfist/> <hamparend/> <hamparbegin/> <hamextfingerl/> <hampalmd/> <hamplus/> <hamextfingero/> <hampalmd/> <hamparend/> <hamparbegin/> <hamshoulders/> <hamplus/> <hamchest/> <hamparend/> <hammoved/> <hamparbegin/> <hamwristback/> <hamhandback/> <hamplus/> <hamindexfinger/> <hamfingertip/> <hamparend/> <hamtouch/> <hamseqbegin/> <hammoveu/> <hamsmallmod/> <hammoved/> <hamsmallmod/> <hamseqend/> <hamrepeatfromstart/> </hamnosys_manual> </hns_sign> </sigml>	

Fig. 15 analyzes the input sentence and shows the avatar animation of its ISL sentence *These are books.*

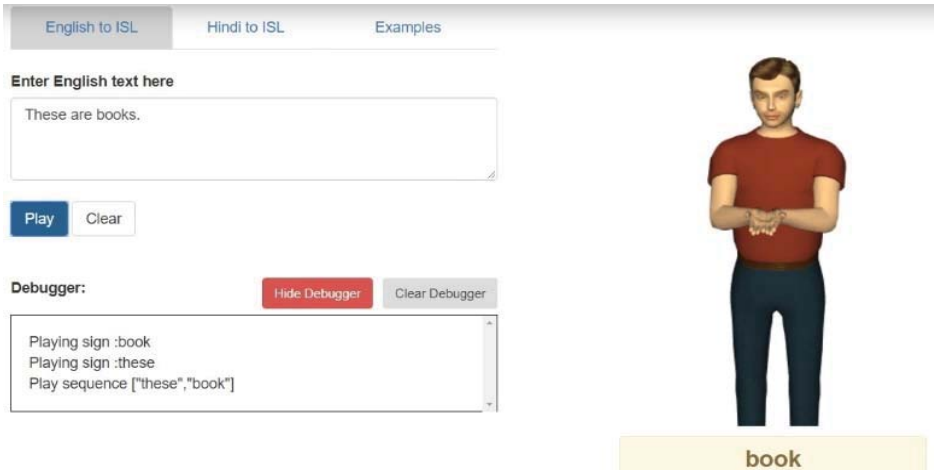

Fig. 15 ISL animation of sentence *These are books.*

VIII. CONCLUSION AND FUTURE SCOPE

An online multilingual ISL dictionary has been developed for Indian signs using different resources and books available for ISL. It is used as a database to associate signs with words or phrases of a spoken language. Thus, this dictionary consists of different words along with their corresponding HamNoSys notation. The dictionary is available at URL *www.islfromtext.in*. The system displays the 3D ISL avatar animation for each input sentence. The mobile application for ISL dictionary is available at Google Play Store [16].

In Future, this dictionary will be extended for ISL signs so the researchers can. The work to develop a ISL generation system from Hindi input language is in the process which will help the researchers and other in the field of SL generation.

ACKNOWLEDGMENT

This Publication is an outcome of the R&D work undertaken in the project under the Visvesvaraya PhD Scheme of Min-istry of Electronics and Information Technology,Government of India, being implemented by *Digital India Corporation* (formerly Media Lab Asia). We also acknowledge the support of *Department of Science and Technology* (DST), Government of India, under research project '*Automatic generation of sign language from Hindi Text for communication and education of*

hearing impaired people' supported by Science for Equity, Empowerment and Development (SEED) division with vide number 'SEED/TIDE/011/2015' to carry this work.

REFERENCES

[1] S. Chalk, "Do we know how many sign languages there are in the world?," 2014. (Online). Available: *http://www.clarion-uk.com/know-many-sign-languages-world.*
[2] D. C. I (Deaf Community of India), "The indian sign language community," 2012. (Online). Available: *https://joshuaproject.net/assets/media/profiles/text/t19007_in.pdf.*
[3] U. Zeshan, M. N. Vasishta, and M. Sethna, "Implementation of indian sign language in educational settings," Asia Pacific Disability Rehabilitation Journal, vol. 16, no. 1, pp. 16–40, 2005.
[4] S. Sinha, Indian Sign Language: An Analysis of Its Grammar. Gallaudet University Press, 2017.
[5] J. L. Fuertes, A. L. Gonzalez,´ G. Mariscal, and C. Ruiz, "Bilingual sign language dictionary," in International Conference on Computers for Handicapped Persons. Springer, 2006, pp. 599–606.
[6] E. Suzuki, T. Suzuki, and K. Kakihana, "On the web trilingual sign language dictionary to learn the foreign sign language without learning a target spoken language," in Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06), 2006.
[7] J. H. Kristoffersen and T. Troelsgard, "The danish sign language dictionary," in Proceedings of the XIV EURALEX International Congress. Leeuwarden: Fryske Akademy, 2010, pp. 1549–1554.
[8] J. H. Kristoffersen and T. TroelsgArd, The electronic lexicographical treatment of sign languages: The Danish Sign Language Dictionary. Oxford University Press, 2012.
[9] K. Cormier *et. al*, "British Sign Language (BSL) corpus project," 2011. (Online). Available: *http://bslcorpusproject.org.*
[10] J. Fenlon, K. Cormier, and A. Schembri, "Building bsl signbank: The lemma dilemma revisited," International Journal of Lexicography, vol. 28, no. 2, pp. 169–206, 2015.
[11] P. M. Martin, S. Belhe, S. Mudliar, M. Kulkarni, and S. Sahasrabudhe, "An indian sign language (isl) corpus of the domain disaster message using avatar," in Proceedings of the Third International Symposium in Sign Language Translations and Technology (SLTAT-2013), 2013, pp. 1–4.
[12] T. Hanke, "Dgs corpus project-development of a corpus based electronic dictionary German sign language/German," in Proceedings of the Corpus Linguistics Conference 2009 (CL2009), 2009, p. 355.
[13] W. C. Stokoe, D. C. Casterline, and C. G. Croneberg, A dictionary of American Sign Language on linguistic principles. Linstok Press, 1976.
[14] S. Parkhurst and D. Parkhurst, SignWriting: A Complete System for Writing and Reading Signed Languages. SignWriting Press, 2007.
[15] T. Hanke, "Hamnosys-representing sign language data in language resources and language processing contexts," in LREC, vol. 4, 2004, pp. 1–6.
[16] D. Singh *et. al*, "Sanket - Text to Indian Sign Language Translator," 2018. (Online). Available: *http://play.google.com/store/apps/details?id=in.dsingh.sanket.*

Sugandhi is a Research Scholar in the Department of Computer Science and Engineering at TIET, Patiala. She has earned her BCA degree in 2012 from Punjabi University, Patiala and MCA degree in 2015 from TIET, Patiala. She has 1 year of experience as Software Developer with Times Internet Ltd. She is developing a system for ISL (Indian Sign Language) Automation that can be used by hearing impaired communities for the sake of communication. Her areas of interest are Software Engineering and Natural Language Processing (NLP).

Parteek Kumar is Associate Professor in the Department of Computer Science and Engineering at TIET, Patiala. He has more than eighteen years of academic experience, including six years at D.A.V College Amritsar. He has earned his B.Tech from SLIET and MS from BITS Pilani. He has completed his Ph.D from TIET. He has published more than 75 research papers and articles in Journals, Conferences and Magazines of repute. He has research interest in area of NLP, Machine Learning and Human Computer Interface for Education of Hearing Impaired people.

Sanmeet Kaur is Assistant Professor in Computer Science and Engineering Department, TIET, Patiala, Punjab, India. She has earned her ME (Software Engineering) in 2008 and Ph.D in 2015 from TIET, Patiala. She has published more than 30 research papers and articles in Journals, Conferences and Magazines of repute. Her areas of interest are Network Security, Software Testing and Software Engineering.