SignON: Bridging the gap between Sign and Spoken Languages

SignON: cerrando la brecha entre las lenguas de signos y las lenguas orales

H. Saggion
Universitat Pompeu Fabra
horacio.saggion@upf.edu

D. Shterionov
Tilburg University
D.Shterionov@tilburguniversity.edu

G. Labaka
Universidad del País Vasco
gorka.labaka@ehu.eus

T. Van de Cruys
KU Leuven
tim.vandecruys@kuleuven.be

V. Vandeghinste
Instituut voor de Nederlandse Taal
and KU Leuven
vincent@ccl.kuleuven.be

J. Blat
Universitat Pompeu Fabra
josep.blat@upf.edu

Abstract: This article presents an overview of the SignON European project which aims to develop technology for automatic translation between sign and oral languages (and vice-versa). In order to achieve this objective, the project takes a multi-disciplinary approach by involving the deaf community, sign language linguistics, research in sign language recognition, speech recognition, natural language processing and machine translation (MT), 3D animation and avatar technology, and application development. The project follows a user-centered, community driven approach to the development of technology.
Keywords: Neural Machine Translation; Sign Language; Automatic Sign Language Recognition; Text Simplification; Avatar.

Resumen: Este artículo describe el proyecto europeo SignON, que tiene como objetivo desarrollar tecnología de traducción automática entre lenguas de signos y lenguas orales. Para lograr este objetivo, el proyecto adopta un enfoque multidisciplinario al involucrar signantes de lenguas de signos, lingüístas de las lenguas de signos, tecnología de reconocimiento automático de lengua de signos, reconocimiento automático de voz, procesamiento del lenguaje natural y traducción automática, animación 3D y la tecnología de avatar y desarrollo de aplicaciones. El proyecto sigue un enfoque centrado en el usuario e impulsado por la comunidad sorda para el desarrollo de una tecnología apropiada.
Palabras clave: Traducción automática neuronal; lenguas de signos; reconocimiento automático de lengua de signos; simplificación de textos; avatar.

1 Introduction

Access to information is a human right. In the modern, globalised world this implies access to multilingual content and cross-lingual communication with others. Crossing language barriers is essential for global information exchange and unobstructed, fair communication. The World Health Organisation (WHO) reports that there are some 466 million people in the world today with disabling hearing loss[1]; it is estimated that this number will double by 2050. According to the World Federation of the Deaf (WFD), over 70 million people are deaf and communicate primarily via a sign language (SL).

Machine translation (MT) (Koehn, 2009) is a core technique for reducing language barriers that has advanced, and seen many breakthroughs since it began in the 1950s (Johnson et al., 2017), to reach quality lev-

[1] https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

els comparable to humans (Hassan et al., 2018). Despite the significant advances of MT for spoken languages in the recent couple of decades, MT is in its infancy when it comes to SLs. The complexity of the problem, automatically translating between SLs or SL and spoken languages, requires a multidisciplinary approach (Bragg et al., 2019).

In this paper we present the SignON project which focuses on the research and development of a *sign language translation mobile application and an open communications framework*. SignON aims to rectify the lack of technology and services for automatic translation between sign and spoken languages, through an inclusive, human-centric digital transformation solution facilitating communication between deaf and hard of hearing (DHH) and hearing individuals.

2 Project overview

SignON is a Horizon 2020 project which aims to develop a communication service that translates between sign and spoken (in both text and audio modalities) languages and caters for the communication needs between DHH and hearing individuals. Currently, human interpreters are the main medium for sign-to-spoken, spoken-to-sign and sign-to-sign language translation. The availability and cost of these professionals is often a limiting factor in communication between signers and non-signers. The SignON communication service will translate between sign and spoken languages, bridging language gaps when professional interpretation is unavailable.

2.1 Objectives

A primary objective of the SignON project is to create a service that translates between sign and verbal languages. This high-level objective is broken down to the following 6 lower-level objectives:

1. **Co-creation workflow and community.** We aim to bring researchers and developers in a close collaboration with the main stakeholder groups to drive the research and development in SignON.
2. Development of the **SignON Framework and Mobile application** which will deliver the SignON service to the user.
3. **Automated recognition and understanding of SL and verbal language input** through advanced sign language recognition (SLR), automatic speech recognition (ASR), and natural language understanding (NLU).
4. Research and development of a novel **Language Independent Meaning Representation** for interlingua MT. It will be based on current vector representations (Lample et al., 2018), symbolic components (Baker, 2014) or hybrid representations of the input/output message.
5. **Sign, speech and text synthesis.** SignON will convert an SL specific syntactic-semantic representation in the target SL via a customizable 3D virtual signer (i.e. avatar). It will also produce text output in the different oral languages adapted to the user by, for example, simplifying the text.
6. **Wide-range of supported languages and extensibility of the framework.** During the project we will provide support for the following SLs: Irish SL (ISL), British SL (BSL), Flemish SL (VGT), Dutch SL (NGT) and Spanish SL (LSE) as well as English, Irish, Dutch and Spanish verbal languages. However, we design the SignON application and framework to be extensible to new sign and spoken/written languages.

2.2 Challenges

Achieving the aforementioned objectives, and thus the envisaged service is not a simple endeavour.

First, the difference between sign and verbal languages as well as between the different SLs (Vermeerbergen and Van Herreweghe, 2010) makes it difficult to adopt translation processes that have been developed for verbal languages for the use-cases of sign-to-verbal (and vice-versa) or sign-to-sign translation.

Second, there has been a lack of sufficiently advanced technology. In order to recognise and understand SL, tools need to be able to process digital representations of signers (e.g. 2D or 3D video) and to compress the underlying information into a meaningful (for both humans and machines) representation (De Coster, Van Herreweghe, and Dambre, 2020). Often, attempts at sign language recognition (SLR) require cumbersome additional hardware, such as gloves and bracelets, which incorrectly assume that an SL is simply articulated only on the hands.

A **third** issue is the scarce amounts of anno-

tated SL materials or parallel data in which signs are linked to text making it difficult to train state-of-the-art models.

Forth, the gap between DHH and hearing communities is big and it is often expressed as a lack of demand for and of willingness to adopt technological solutions. While technology could be of enormous benefit for each of these communities, it has not yet reached the expectations of its potential users.

To address these challenges SignON is exploring: (i) a multilingual representation common for both sign and verbal languages (InterL); (ii) sophisticated deep learning methods for recognition; (iii) efficient on-the-fly synthesis of detailed 3D avatars; (iv) an adaptive pipeline to allow the updating of the models based on user input; and (v) a co-creation methodology bringing together SignON researchers and the DHH community.

2.3 Current Developments

MT for SLs has been addressed with different approaches from rule-based methods (Porta et al., 2014), through statistical (Morrissey, 2008) and to neural machine translation (NMT) (Yin and Read, 2020). Given the objective of developing a multilingual, extensible, language independent framework for representing and translating language, the work we conducted in the first 6 months of this project focused on (i) the adaptation of the state of the art mBART model (Liu et al., 2020) to translating between English, Spanish and Dutch in both bilingual (the model is fine-tuned on two languages) and multilingual (the model is fine-tuned on all languages sequentially) settings as well as to experiment with text simplification (Saggion, 2017); (ii) infrastructure and framework development to support all interleaved components; and (iii) analysis of the current stakeholders' attitude, perception and vision related to SL translation.

3 Consortium

The following organizations participate in the SignON consortium:

1. Dublin City University (DCU) (coordinator), Ireland
2. Fincons Group (FINC), Switzerland
3. Instituut voor de Nederlandse Taal (INT), The Netherlands
4. University of the Basque Country (UPV/EHU), Spain
5. The National Microelectronics Applications Centre Ltd (MAC), Ireland
6. Pompeu Fabra University (UPF), Spain
7. Technological University Dublin (TUDublin), Ireland
8. Trinity College Dublin (TCD), Ireland
9. VRT, Belgium
10. Ghent University (UGent), Belgium
11. Vlaams GebarentaalCentrum (Flemish Sign Language Centre – VGTC), Belgium
12. University College Dublin (UCD), Ireland
13. Stichting Katholieke Universiteit (RU), The Netherlands
14. Nederlandse TaalUnie (NTU), The Netherlands
15. Katholieke Universiteit Leuven (KULeuven), Belgium
16. European Union of the Deaf (EUD), Belgium
17. Tilburg University (TiU) (scientific lead), The Netherlands

4 A Co-creation Approach

SignON aims to reduce the gap between the stakeholder communities through a user-centred and community-driven research and development approach, involving stakeholder-led user profiles from its inception. Our co-creation strategy relies on a continuous communication and collaboration with DHH communities to iteratively (re)define use-cases, co-design and co-develop the SignON service and application, assess the quality and validate their acceptance.

Through co-creation we will ensure that the developed solution is (i) accepted by the users; and (ii) that it will continue to evolve beyond the lifetime of the SignON project.

5 SignON App and framework

This project will develop a free, open-source service and framework for conversion between video (capturing and understanding sign language), audio (for speech, including atypical speech) and text, translating between sign and spoken languages, delivered to its users via an easy to use mobile application. The operational workflow of the SignON application and framework is illustrated in Figure 1.

The SignON communication and translation mobile application, each user's interface to the overall cloud platform and SignON framework, will run on standard modern smartphone and tablet devices without the need for special equipment. Further details on the early-stage development

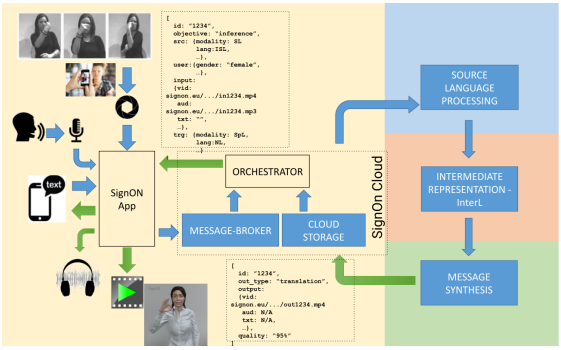

Figure 1: The SignON application and framework workflow.

of the SignON application can be found in (O'Flaherty et al., 2021).

6 Conclusions and Future work

This paper presents an overview of the SignON project. This project interleaves state-of-the-art research with continuous communication and verification with the user communities, a process that we refer to as *co-creation*. In its lifetime, SignON focuses on English, Irish, Dutch and Spanish verbal languages and the following sign languages: ISL, BSL, VGT, NGT, LSE. However, the design of the SignON framework allows for easy integration of new languages.

Acknowledgments

This work is supported by the European Commission under the Horizon 2020 program ICT-57-2020 - "An empowering, inclusive Next Generation Internet" with Grant Agreement number 101017255. We thank all members of the SignON consortium.

References

Baker, C. F. 2014. FrameNet: A knowledge base for natural language processing. In *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929-2014)*, pages 1–5, June.

Bragg, D., O. Koller, M. Bellard, L. Berke, P. Boudreault, A. Braffort, N. Caselli, M. Huenerfauth, H. Kacorri, T. Verhoef, C. Vogler, and M. Ringel Morris. 2019. Sign language recognition, generation, and translation: An interdisciplinary perspective. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, page 16–31.

De Coster, M., M. Van Herreweghe, and J. Dambre. 2020. Sign language recognition with transformer networks. In *Proceedings of the 12th Language Resources and Evaluation Conference*, May.

Hassan, H., A. Aue, C. Chen, V. Chowdhary, J. Clark, C. Federmann, X. Huang, M. Junczys-Dowmunt, W. Lewis, M. Li, S. Liu, T. Liu, R. Luo, A. Menezes, T. Qin, F. Seide, X. Tan, F. Tian, L. Wu, S. Wu, Y. Xia, D. Zhang, Z. Zhang, and M. Zhou. 2018. Achieving human parity on automatic chinese to english news translation. *ArXiv*, abs/1803.05567.

Johnson, M., M. Schuster, Q. V. Le, M. Krikun, Y. Wu, Z. Chen, N. Thorat, F. Viégas, M. Wattenberg, G. Corrado, M. Hughes, and J. Dean. 2017. Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics*, 5:339–351.

Koehn, P. 2009. *Statistical Machine Translation*. Cambridge University Press.

Lample, G., A. Conneau, L. Denoyer, and M. Ranzato. 2018. Unsupervised machine translation using monolingual corpora only. In *International Conference on Learning Representations*.

Liu, Y., J. Gu, N. Goyal, X. Li, S. Edunov, M. Ghazvininejad, M. Lewis, and L. Zettlemoyer. 2020. Multilingual denoising pre-training for neural machine translation.

Morrissey, S. 2008. *Data-Driven Machine Translation for Sign Languages*. Ph.D. thesis, Dublin City University.

O'Flaherty, J., M. P. Scipioni, M. Villa, E. Keane, and M. Giovanelli. 2021. Early-stage development of the signon application and open framework – challenges and opportunities. In *Proceedings of Machine Translation Summit XVIII Volume 2: User Track*, virtual, August. European Association for Machine Translation.

Porta, J., F. López-Colino, J. Tejedor, and J. Colás. 2014. A rule-based translation from written spanish to spanish sign language glosses. *Computer Speech & Language*, 28:788–811, 05.

Saggion, H. 2017. *Automatic Text Simplification*. Morgan & Claypool Publishers.

Vermeerbergen, M. and M. Van Herreweghe. 2010. Sign languages and sign language research. In *The handbook of psycholinguistic and cognitive processes: perspectives in communication disorders*. Psychology Press, pages 709–729.

Yin, K. and J. Read. 2020. Better sign language translation with STMC-transformer. In *Proceedings of COLING*, pages 5975–5989, December.