A Framework for Rapid Development of Limited-Domain Speech-to-Sign Phrasal Translators

Manny Rayner[1], Pierrette Bouillon[1], Sarah Ebling[2], Irene Strasly[1], Nikos Tsourakis[1]

(1) University of Geneva, FTI/TIM, Switzerland
(2) University of Zurich, Institute of Computational Linguistics, Switzerland

We present a web platform which permits rapid development of limited-domain speech-to-sign translators; the work leverages a line of research on limited-domain speech applications that has been pursued at Geneva University (Rayner et al 2006, 2014b, Bouillon et al 2008), extending it to sign language output. The type of application supported typically uses a vocabulary of a few hundred to a couple of thousand words/signs, and grammatical coverage of a few dozen to a few hundred syntactic patterns. Although limited in scope, experience shows that translation applications of this kind are able to support goal-directed exchanges in a specific situation, for example buying a ticket or being admitted to a hospital. Applications can be packaged either as plain speech translators or as interactive spoken questionnaires (Armando et al 2014), where the system automatically advances through the sequence of questions and allows the Deaf partner to respond by choosing an answer from a push-button menu (Cox et al 2002, Tryggvason 2004). The framework is entirely web-based. Applications are accessed through a normal browser, with audio handled using the methods from (Fuchs et al 2012). Development is also carried out over the web, with content uploaded to the server through an FTP client using the model of (Rayner et al 2015a). Signed language generation is performed using the JASigning avatar (Jennings et al. 2010). Initial development work has focused on a train passenger announcement domain (Ebling 2016).

The rule formalism used is a version of Regulus Lite (Rayner et al 2014a), and is designed to be accessible to sign language linguists who have only modest skills in computer science. Translation is performed in two stages. In the first stage, parametrized phrase translation rules are written as groups of lines in Excel spreadsheet tables. “Table-rules” are based on the diagrams common in sign language linguistics (Neidle et al 2000, Johnston and Schembri 2007); one line is a sequence of sign glosses, while others contain non-manual information (eye gaze, head tilt, etc). These rules are compiled into a speech recognition package that can be run on the commercial Nuance platform; the runtime output is a bundle of synchronized parallel output streams (Rayner et al 2015b). The second translation stage uses rules, also written in the Regulus Lite notation, which map sign glosses and tokens representing non-manual elements to HamNoSys sequences (Prillwitz et al. 1989) and SiGML elements (Elliott et al. 2000). The output is finally realized in concrete form by the avatar.

While the goals of the project are practical, the framework we are developing has interesting implications for theoretical work in sign language linguistics: it makes it easy for non-signing experimenters to generate parametrized sets of signed utterances using an avatar, where one part of the utterance is systematically varied, and by obtaining judgments on them from native speakers get data on possible phonological, syntactic and semantic constraints on signed language.

References

Armando, Alejandro, et al. (2014). A Tool for Building Multilingual Voice Questionnaires. Proc Translation and the Computer. London, England.

Bouillon, Pierrette, et al. (2008). Many-to-many multilingual medical speech translation on a PDA. Proc AMTA. Waikiki Beach, Hawaii.

Cox S. et al. (2002). Tessa, a system to aid communication with deaf people. In Proceedings of the Fifth International ACM Conference on Assistive Technologies, Assets '02, 205–212. New York, NY.

Ebling, S. (2016). Building a parallel corpus of German/Swiss German Sign Language train announcements. To appear in *International Journal of Corpus Linguistics*.

Elliott, R. et al. (2000). The development of language processing support for the ViSiCAST project. Proc 4th International ACM Conference in Assistive Technologies, Arlington, VA.

Fuchs, Matthew, Nikolaos Tsourakis, and Manny Rayner. (2012). A scalable architecture for web deployment of spoken dialogue systems. Proc LREC. Istanbul, Turkey.

Jennings, V. et al. (2010). Requirements for a signing avatar. In 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies,

Johnston, Trevor, and Adam Schembri. (2007). *Australian Sign Language (Auslan): An introduction to sign language linguistics*. Cambridge University Press.

Hanke, T. (2004). HamNoSys - Representing Sign Language Data in Language Resources and Language Processing Contexts. Proc LREC Workshop on Representation and Processing of Sign Languages. Paris, France.

Neidle, Carol, et al. (2000). *The Syntax of American Sign Language: Functional categories and hierarchical structure*. Cambridge, MA: MIT Press

Prillwitz, S. et al. (1989). HamNoSys: Version 2.0: Hamburger Notationsystem für Gebärdensprachen: eine Einführung. Hamburg: Signum.

Rayner, M., B.A. Hockey and P. Bouillon. (2006). *Putting Linguistics into Speech Recognition: the Regulus Grammar Compiler*. University of Chicago Press.

Rayner, M., C. Baur, and N. Tsourakis. (2014a). CALL-SLT Lite: A Minimal Framework for Building Interactive Speech-Enabled CALL Applications. Proc. WOCCI. Singapore.

Rayner, M., et al. (2014b). CALL-SLT: A Spoken CALL System based on grammar and speech recognition. Linguistic Issues in Language Technology,**10.2**

Rayner, M., et al. (2015a). Helping non-expert users develop online spoken CALL courses. Proc SLaTE, Leipzig, Germany

Rayner, M., et al. (2015b). Helping domain experts build speech translation systems. Proc FETLT, Sevilla, Spain.

Tryggvason, J. (2004). Vanessa: A system for council information centre assistants to communicate using sign language. http://www.visicast.cmp.uea.ac.uk/eSIGN/Images/VANESSA.pdf.