Engineering Applications of Artificial Intelligence 32 (2014) 258–269

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Translating bus information into sign language for deaf people

V. López-Ludeña [a], C. González-Morcillo [b], J.C. López [b], R. Barra-Chicote [a], R. Cordoba [a], R. San-Segundo [a,*]

[a] Speech Technology Group, E.T.S.I. Telecomunicación, UPM, Spain
[b] Grupo de Sistemas Inteligentes Aplicados, Dpto. de Tecnologías y Sistemas de Información, UCLM, Spain

ARTICLE INFO

Article history:
Received 22 March 2013
Received in revised form
7 October 2013
Accepted 11 February 2014
Available online 4 March 2014

Keywords:
Bus information
Translation into sign language
Deaf people
LSE
Lengua de Signos Española

ABSTRACT

This paper describes the application of language translation technologies for generating bus information in Spanish Sign Language (LSE: Lengua de Signos Española). In this work, two main systems have been developed: the first for translating text messages from information panels and the second for translating spoken Spanish into natural conversations at the information point of the bus company. Both systems are made up of a natural language translator (for converting a word sentence into a sequence of LSE signs), and a 3D avatar animation module (for playing back the signs). For the natural language translator, two technological approaches have been analyzed and integrated: an example-based strategy and a statistical translator. When translating spoken utterances, it is also necessary to incorporate a speech recognizer for decoding the spoken utterance into a word sequence, prior to the language translation module. This paper includes a detailed description of the field evaluation carried out in this domain. This evaluation has been carried out at the customer information office in Madrid involving both real bus company employees and deaf people. The evaluation includes objective measurements from the system and information from questionnaires. In the field evaluation, the whole translation presents an SER (Sign Error Rate) of less than 10% and a BLEU greater than 90%.

1. Introduction

According to wfdeaf.org (2013), there are more than 70 million deaf people in the world. This disability has serious implications for education and social inclusion. In Spain, there are 1,064,000 deaf people according to the INE (Spanish Institute of Statistics). Forty-seven percent of the deaf population do not have basic studies or are illiterate, and only between 1% and 3% have finished their university studies (as opposed to 21% of Spanish hearing people). Deaf people (especially those that became deaf before language acquisition) have serious problems when expressing themselves or understanding written texts. They have problems with verb tenses, concordances of gender and number, etc., and they have difficulties when creating a mental image of abstract concepts. These deficiencies have become apparent because of the lack of feedback in speak-listen procedures.

However, the Deaf[1] use a sign language (their mother tongue) for communicating. Sign languages are fully fledged languages that have a grammar and lexicon just like any spoken language, contrary to what most people think. Traditionally, deafness has been associated with people with learning problems but this is not the case. The use of sign languages defines the Deaf as a linguistic minority, with learning skills, cultural and group rights similar to other minority language communities. In 2007, the Spanish Government accepted Spanish Sign Language (LSE: Lengua de Signos Española) as one of the official languages in Spain, defining a long-term plan to invest in new resources for developing, disseminating and increasing the standardization of this language. LSE is a natural language with the same linguistic levels as other languages such as Spanish. Thanks to associations such as the Fundación CNSE, LSE is becoming the natural language for the Deaf to communicate.

This paper describes the efforts made to translate transport information into LSE, specifically bus information. The main target is to translate this information automatically (without human intervention).

This paper is organised as follows. Section 2 presents the state of the art. Section 3 describes the main language translation technologies considered in this work. The sign representation using an animated avatar is described in Section 4. Section 5 describes the system for translating panel information, and Section 6 the system for translating face-to-face conversations at the customer service office, including a field evaluation. Finally, Section 7 summaries the main conclusions of this work.

* Correspondence to: E.T.S.I. Telecomunicación. Ciudad Universitaria SN, 28040 Madrid, Spain.
E-mail address: lapiz@die.upm.es (R. San-Segundo).

[1] It is necessary to differentiate between "deaf" and "Deaf": the former refers to non-hearing people, and the latter refers to non-hearing people who use a sign language to communicate between themselves (their mother tongue), making them part of the "Deaf community".

http://dx.doi.org/10.1016/j.engappai.2014.02.006
0952-1976

2. State of the art

In the last 20 years, the European Commission and the USA Government have invested many resources into research into language translation. In Europe, there has been a large sequence of research projects: C-Star, ATR. Vermobil, Eutrans, LC-Star, PF-Star and, finally, TC-STAR, EuroMatrix, EuroMatrixPlus, FAUST, etc. Some of them focus on text translation and others on spoken language. The FAUST project focuses on computer-aided translation. In the USA, DARPA (Defense Advanced Research Projects Agency) is supporting the GALE program (http://www.darpa.mil/ipto/programs/gale/gale.asp). The goal of the DARPA GALE program has been to develop and apply computer software technologies to absorb, analyze and interpret huge volumes of speech and text in multiple languages. This program has also been promoted by the machine translation evaluation organised by the US Government, NIST (http://www.itl.nist.gov/iad/mig/tests/mt/).

The best performing translation systems are based on various types of statistical approaches (Och and Ney, 2002; Mariño et al., 2006), including example-based methods (Sumita et al., 2003), finite-state transducers (Casacuberta and Vidal, 2004) and other data-driven approaches. The progress made over the last 10 years is due to several factors such as efficient algorithms for training (Och and Ney, 2003), context dependent models (Zens et al., 2002), efficient algorithms for generation (Koehn et al., 2003), more powerful computers and bigger parallel corpora, and automatic error measurements (Papineli et al., 2002; Banerjee and Lavie, 2005; Agarwal and Lavie, 2008).

Another important effort in machine translation has been the organization of several Workshops on Statistical Machine Translation (SMT). On the webpage http://www.statmt.org/, it is possible to obtain all the information on these events. As a result of these workshops, there is a free machine translation system called Moses available from this web page (http://www.statmt.org/moses/). Moses is a phrase-based statistical machine translation system that allows you to build machine translation system models for any language pair, using a collection of translated texts (parallel corpus).

In recent years, several groups have shown interest in spoken language translation into sign languages, developing several prototypes: example-based (Morrissey and Way, 2005), rule-based (San-Segundo et al., 2008), grammar-based (Marshall and Sáfár, 2005), full sentence (Cox et al., 2002) or statistical (Bungeroth and Ney, 2004; SiSi system http://www-03.ibm.com/press/us/en/pressrelease/22316.wss; Morrissey et al., 2007) approaches. For LSE, it is important to highlight the authors' experience in developing speech into LSE translation systems in several domains (San-Segundo et al., 2008, 2011; López-Ludeña et al., 2012, 2013a). This kind of system can complement a Sign Language into Speech translation system, allowing a two-direction interaction (Cemil and Ming, 2011; Ibarguren et al., 2010).

As regards 3D avatars for representing signs, the VISICAST and eSIGN European Project (Essential Sign Language Information on Government Networks) (http://www.sign-lang.uni-hamburg.de/esign/) (Zwiterslood et al., 2004) has been one of the most significant research efforts into developing tools for the automatic generation of sign language contents. One of the partners in the VISICAST and eSIGN projects is the research group into Virtual Humans at the University of East Anglia (http://www.uea.ac.uk/cmp/research/graphicsvisionspeech/vh). This group has been involved in several projects as regards the generation of sign language using virtual humans: TESSA, SignTel, Visicast, eSIGN, SiSi, LinguaSign, etc.

This paper describes the effort in adapting translation technology for generating LSE content into the bus information domain. This technology has been used for translating both panel information and spoken Spanish into LSE in real interactions between a deaf person and a hearing person without an interpreter: deaf customers and bus company employees that provide bus information. The method and the system used in this research work have been developed during several years in previous research projects (San-Segundo et al., 2008, 2011; López-Ludeña et al., 2012, 2013a).

3. Language translation technology

In this work, several translation strategies (López-Ludeña et al., 2013a) have been adapted and evaluated in the bus information domain: example-based and statistical translation. The final translation module integrates all these technologies.

In order to use automatic translating technologies, it is essential to represent the Sign Language in a written form. In order to write down LSE, each sign of an LSE sentence is represented by a gloss, so a gloss sequence represents a sequence of signs. Glosses are words in capital letters with a similar meaning to the sign meaning. An example of glosses representing the sentence "¿a qué hora se abre? (what time do you open?)" would be "ABRIR HORA? (OPEN HOUR?)". There can be several signs represented by a gloss with '+', for example: "SABADO+DOMINGO (SATURDAY+SUNDAY)" to represent "fin de semana (weekend)". Also, there can be several words in Spanish that form only one gloss in LSE, this fact is marked with '-'. For example, "CAFE-CON-LECHE" for representing "café con leche (coffee with milk)". For more details about LSE and written LSE can be seen at López-Ludeña et al. (2012).

3.1. Example-based strategy

An example-based translation system uses a parallel corpus: set of sentences in the source language (from which one is translating) and its corresponding translations into the target language, and translates other similar source-language sentences. In order to determine whether one example (in the corpora) is similar enough to the text to be translated, the system computes a heuristic distance between both sentences. If the distance is less than a threshold, the translation output will be the same as the example translation. But if the distance is greater, the system cannot generate any output and it is necessary to consider other translation strategies.

In this case, the heuristic distance considered is the well-known Levenshtein distance (LD) (Levenshtein, 1966) divided by the number of words in the sentence to be translated (this distance is represented as a percentage). The Levenshtein Distance is a measurement of the similarity between two strings (or character sequences): source sequence (s) and target sequence (t). The distance is the number of deletions, insertions, or substitutions required to transform s into t. Because of this, it is also called the edit distance. Originally, this distance was used to measure the similarity between two strings (character sequences). But it was already used for defining a distance between word sequences (as has been used in this paper). The LD is computed using a dynamic programming algorithm that considers the following costs: 0 for identical words, 1 for insertions, 1 for deletions and 1 for substitutions.

In order to develop an example-based translation system, it is necessary a large amounts of pre-translated text to make a reasonable translator. But it is possible to generalize the examples in order to make them more effective: more than one string can match any given part of the example. Considering the following translation example for Spanish into LSE:

Spanish**:** "Veinte euros con diez céntimos" (Twenty Euros, ten cents)
LSE**:** "VEINTE COMA DIEZ EURO"

Now, if it is known that "veinte" and "diez" are numbers, it is possible to save this example in the corpus as