The 2008 MedSLT System

Manny Rayner[1], Pierrette Bouillon[1], Jane Brotanek[2], Glenn Flores[2]
Sonia Halimi[1], Beth Ann Hockey[3], Hitoshi Isahara[4], Kyoko Kanzaki[4]
Elisabeth Kron[5], Yukie Nakao[6], Marianne Santaholma[1]
Marianne Starlander[1], Nikos Tsourakis[1]

[1] University of Geneva, TIM/ISSCO, 40 bvd du Pont-d'Arve, CH-1211 Geneva 4, Switzerland
{Emmanuel.Rayner,Pierrette.Bouillon,Nikolaos.Tsourakis}@issco.unige.ch
{Sonia.Halimi,Marianne.Santaholma,Marianne.Starlander}@eti.unige.ch

[2] UT Southwestern Medical Center, Children's Medical Center of Dallas
{Glenn.Flores,Jane.Brotanek}@utsouthwestern.edu

[3] Mail Stop 19-26, UCSC UARC, NASA Ames Research Center, Moffett Field, CA 94035–1000
bahockey@ucsc.edu

[4] NICT, 3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto, Japan 619-0289
{isahara,kanzaki}@nict.go.jp

[5] 3 St Margarets Road, Cambridge CB3 0LT, England
elisabethkron@yahoo.co.uk

[6] University of Nantes, LINA, 2, rue de la Houssinière, BP 92208 44322 Nantes Cedex 03
yukie.nakao@univ-nantes.fr

Abstract

MedSLT is a grammar-based medical speech translation system intended for use in doctor-patient diagnosis dialogues, which provides coverage of several different subdomains and multiple language pairs. Vocabulary ranges from about 350 to 1000 surface words, depending on the language and subdomain. We will demo three different versions of the system: an any-to-any multilingual version involving the languages Japanese, English, French and Arabic, a bidirectional English ↔ Spanish version, and a mobile version running on a hand-held PDA. We will also demo the Regulus development environment, focussing on features which support rapid prototyping of grammar-based speech translation systems.

1 Introduction

MedSLT is a medium-vocabulary grammar-based medical speech translation system built on top of the Regulus platform (Rayner et al., 2006). It is intended for use in doctor-patient diagnosis dialogues, and provides coverage of several subdomains and a large number of different language-pairs. Coverage is based on standard examination questions obtained from physicians, and focusses primarily on yes/no questions, though there is also support for WH-questions and elliptical utterances.

Detailed descriptions of MedSLT can be found in earlier papers (Bouillon et al., 2005; Bouillon et al., 2008)[1]. In the rest of this note, we will briefly sketch several versions of the system that we intend to demo at the workshop, each of which displays new features developed over the last year. Section 2 describes an any-language-to-any-language multilingual version of the system; Section 3, a bidirectional English ↔ Spanish version; Section 4, a version running on a mobile PDA

[1] All MedSLT publications are available on-line at `http://www.issco.unige.ch/projects/medslt/publications.shtml`.

Coling 2008: Proceedings of the workshop on Speech Processing for Safety Critical Translation and Pervasive Applications, pages 32–35
Manchester, August 2008

platform; and Section 5, the Regulus development environment.

2 A multilingual version

During the last few months, we have reorganised the MedSLT translation model in several ways[2]. In particular, we give a much more central role to the interlingua; we now treat this as a language in its own right, defined by a normal Regulus grammar, and using a syntax which essentially amounts to a greatly simplified form of English. Making the interlingua into another language has made it easy to enforce tight constraints on well-formedness of interlingual semantic expressions, since checking well-formedness now just amounts to performing generation using the interlingua grammar.

Another major advantage of the scheme is that it is also possible to systematise multilingual development, and only work with translation from source language to interlingua, and from interlingua to target language; here, the important point is that the human-readable interlingua surface syntax makes it feasible in practice to evaluate translation between normal languages and the interlingua. Development of rules for translation *to* interlingua is based on appropriate corpora for each source language. Development of rules for translating *from* interlingua uses a corpus which is formed by merging together the results of translating each of the individual source-language corpora into interlingua.

We will demonstrate our new capabilities in interlingua-based translation, using a version of the system which translates doctor questions in the headache domain from any language to any language in the set {English, French, Japanese, Arabic}. Table 1 gives examples of the coverage of the English-input headache-domain version, and Table 2 summarises recognition performance in this domain for the three input languages where we have so far performed serious evaluations. Differences in the sizes of the recognition vocabularies are primarily due to differences in use of inflection.

3 A bidirectional version

The system from the preceding section is unidirectional; all communication is in the doctor-to-patient direction, the expectation being that the patient will respond non-verbally. Our second demo, an early version of which is described in (Bouillon et al., 2007), supports bidirectional translation for the sore throat domain, in the English ↔ Spanish pair. Here, the English-speaking doctor typically asks WH-questions, and the Spanish-speaking patient responds with elliptical utterances, which are translated as full sentence responses. A short example dialogue is shown in Table 3.

[2] The ideas in the section are described at greater length in (Bouillon et al., 2008).

Language	Vocab	WER	SemER
English	447	6%	11%
French	1025	8%	10%
Japanese	422	3%	4%

Table 2: Recognition performance for English, French and Japanese headache-domain recognisers. "Vocab" = number of surface words in source language recogniser vocabulary; "WER" = Word Error Rate for source language recogniser, on in-coverage material; "SemER" = semantic error rate for source language recogniser, on in-coverage material.

Doctor:	Where is the pain?
	¿Dónde le duele?
Patient:	En la garganta.
	I experience the pain in my throat.
Doctor:	How long have you had a pain in your throat?
	¿Desde cuándo le duele la garganta?
Patient:	Más de tres días.
	I have experienced the pain in my throat for more than three days.

Table 3: Short dialogue with bidirectional English ↔ Spanish version. System translations are in italics.

4 A mobile platform version

When we have shown MedSLT to medical professionals, one of the most common complaints has been that a laptop is not an ideal platform for use in emergency medical situations. Our third demo shows an experimental version of the system using a client/server architecture. The client, which contains the user interface, runs on a Nokia Linux N800 Internet Tablet; most of the heavy processing, including in particular speech recognition, is hosted on the remote server, with the nodes communicating over a wireless network. A picture of

Where?	Is the pain above your eye?
When?	Have you had the pain for more than a month?
How long?	Does the pain typically last a few minutes?
How often?	Do you get headaches several times a week?
How?	Is it a stabbing pain?
Associated symptoms?	Do you vomit when you get the headaches?
Why?	Does bright light make the pain worse?
What helps?	Does sleep make the pain better?
Background?	Do you have a history of sinus disease?

Table 1: Examples of English MedSLT coverage

the tablet, showing the user interface, is presented in Figure 1. The sentences appearing under the back-translation at the top are produced by an online help component, and are intended to guide the user into the grammar's coverage (Chatzichrisafis et al., 2006).

The architecture is described further in (Tsourakis et al., 2008), which also gives performance results for another Regulus applications. These strongly suggest that recognition performance in the client/server environment is no worse than on a laptop, as long as a comparable microphone is used.

5 The development environment

Our final demo highlights the new Regulus development environment (Kron et al., 2007), which has over the last few months acquired a large amount of new functionality designed to facilitate rapid prototyping of spoken language applications[3]. The developer initially constructs and debugs her components (grammar, translation rules etc) in a text view. As soon as they are consistent, she is able to compile the source-language grammar into a recogniser, and combine this with other components to run a complete speech translation system within the development environment. Connections between components are defined by a simple config file. Figure 2 shows an example.

References

Bouillon, P., M. Rayner, N. Chatzichrisafis, B.A. Hockey, M. Santaholma, M. Starlander, Y. Nakao, K. Kanzaki, and H. Isahara. 2005. A generic multilingual open source platform for limited-domain medical speech translation. In *Proceedings of the 10th Conference of the European Association for Machine Translation (EAMT)*, pages 50–58, Budapest, Hungary.

Bouillon, P., G. Flores, M. Starlander, N. Chatzichrisafis, M. Santaholma, N. Tsourakis, M. Rayner, and B.A. Hockey. 2007. A bidirectional grammar-based medical speech translator. In *Proceedings of the ACL Workshop on Grammar-based Approaches to Spoken Language Processing*, pages 41–48, Prague, Czech Republic.

Bouillon, P., S. Halimi, Y. Nakao, K. Kanzaki, H. Isahara, N. Tsourakis, M. Starlander, B.A. Hockey, and M. Rayner. 2008. Developing non-european translation pairs in a medium-vocabulary medical speech translation system. In *Proceedings of LREC 2008*, Marrakesh, Morocco.

Chatzichrisafis, N., P. Bouillon, M. Rayner, M. Santaholma, M. Starlander, and B.A. Hockey. 2006. Evaluating task performance for a unidirectional controlled language medical speech translation system. In *Proceedings of the HLT-NAACL International Workshop on Medical Speech Translation*, pages 9–16, New York.

Kron, E., M. Rayner, P. Bouillon, and M. Santaholma. 2007. A development environment for building grammar-based speech-enabled applications. In *Proceedings of the ACL Workshop on Grammar-based Approaches to Spoken Language Processing*, pages 49–52, Prague, Czech Republic.

Rayner, M., B.A. Hockey, and P. Bouillon. 2006. *Putting Linguistics into Speech Recognition: The Regulus Grammar Compiler*. CSLI Press, Chicago.

Tsourakis, N., M. Georghescul, P. Bouillon, and M. Rayner. 2008. Building mobile spoken dialogue applications using regulus. In *Proceedings of LREC 2008*, Marrakesh, Morocco.

[3]This work is presented in a paper currently under review.

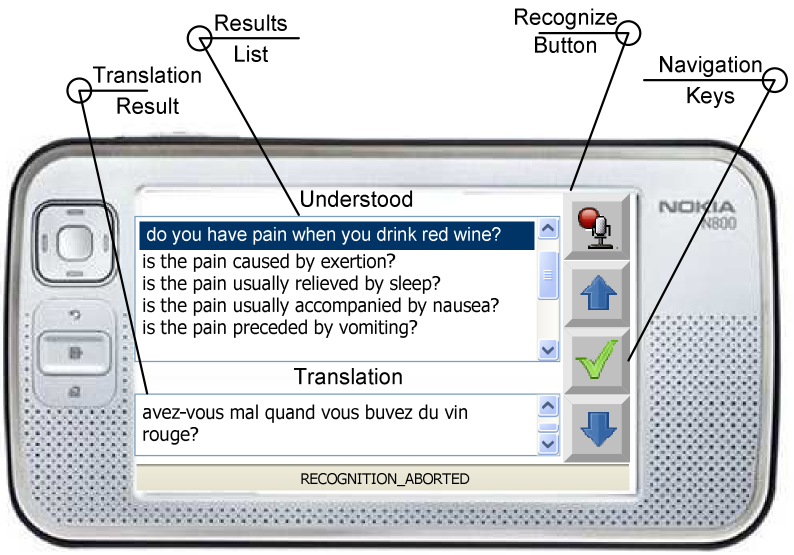

Figure 1: Mobile version of the MedSLT system, running on a Nokia tablet.

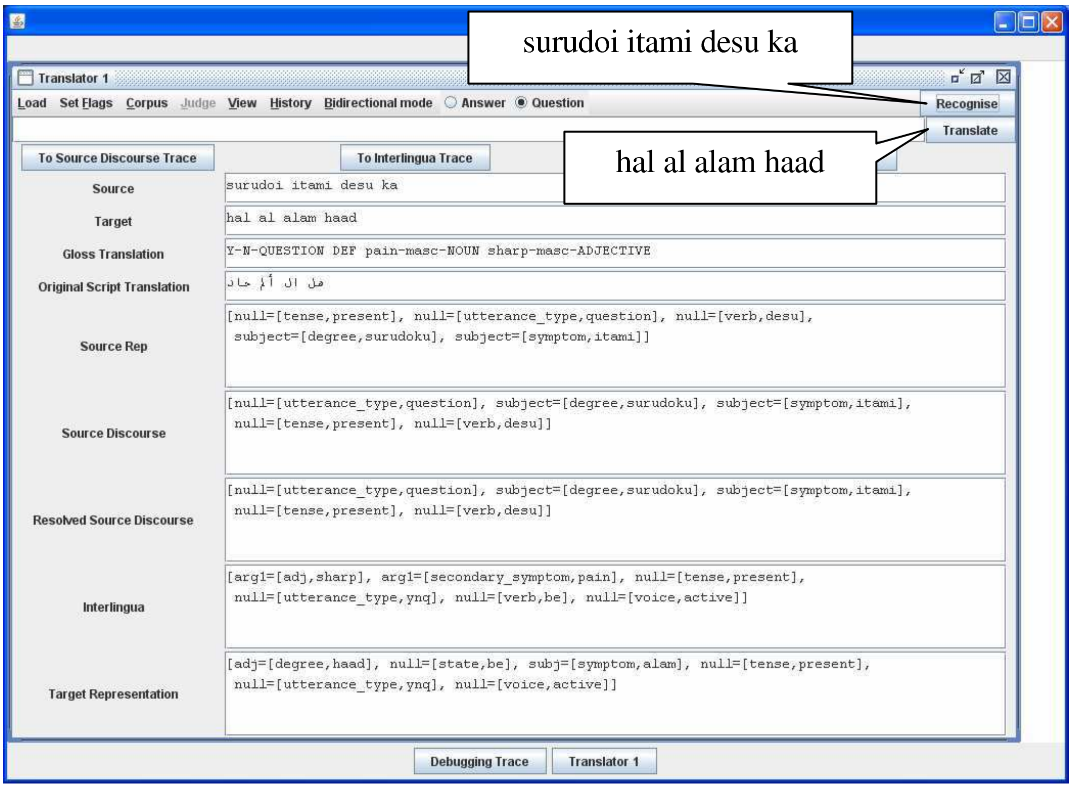

Figure 2: Speech to speech translation from the development environment, using a Japanese to Arabic translator built from MedSLT components. The user presses the Recognise button (top right), speaks in Japanese, and receives a spoken translation in Arabic together with screen display of various processing results. The application is defined by a config file which combines a Japanese recogniser and analysis grammar, Japanese to Interlingua and Interlingua to Arabic translation rules, an Arabic generation grammar, and recorded Arabic wavfiles used to construct a spoken result.