

# NLP-enhanced self-study learning materials for quality healthcare in Europe

**Miriam Urbano Mendaña**  
University of Málaga  
miriamurbano@uma.es

**Gloria Corpas Pastor**  
University of Málaga  
gcorpas@uma.es

**Ruslan Mitkov**  
University of Wolverhampton  
r.mitkov@wlv.ac.uk

## Abstract

In this paper we present an overview of the TELL-ME project, which aims to develop innovative e-learning tools and self-study materials for teaching vocationally-specific languages to healthcare professionals, helping them to communicate at work. The TELL-ME e-learning platform incorporates a variety of NLP techniques to provide an array of diverse work-related exercises, self-assessment tools and an interactive dictionary of key vocabulary and concepts aimed at medics for Spanish, English and German. A prototype of the e-learning platform is currently under evaluation.

## 1 Introduction

A key priority of the European Union (EU) is to improve quality of life by enhancing its citizens' mobility. This includes the free movement of medical professionals and patients within other EU states. In fact, the access to quality healthcare is one of the European priorities in health and social welfare.

In line with this priority, we seek to improve access to quality healthcare in Europe through the training of health professionals. The TELL-ME Project (*Towards European Language Learning for Medical Professionals*)<sup>1</sup> aims to remove barriers to medical communication in multicultural and multilingual medical settings within the European Union. This project, involving institutions and hospitals in Spain, United Kingdom and Germany<sup>2</sup>, is part of the Lifelong

Learning Programme of the European Commission as one of the Leonardo da Vinci Multilateral Projects for Development of Innovation (ref. no. 517937-LLP-2011-UK-LEONARDO-LMP, January 2012 - December 2013).<sup>3</sup>

We aim to support those medical professionals already working outside their home EU countries and provide the tools for others to undertake such mobility for work. For this purpose, we have developed a comprehensive set of learning materials and support tools for teaching vocationally relevant English, Spanish and German to health professionals, helping them to communicate. TELL-ME incorporates a new language learning methodology into a coherent environment, providing users with access to personalised, interactive learning material in various forms including video, audio, corpora, and a dictionary, thanks to the interactive exercises which highlight gaps in the knowledge of the learner. The exercises are, in turn, generated in a semi-automatic manner which significantly increases productivity. The generation is based on pioneering methodology and cutting-edge NLP technology (Mitkov et al., 2006).

The TELL-ME e-learning platform could have a very positive impact on the current socio-demographic context. According to statistics provided by the OECD, 33% of the doctors in Britain and 12.6% of the doctors in Germany come from other countries. Particularly, in Germany, there is a huge political controversy about the shortage of health professionals, which is

---

<sup>1</sup>TELL-ME project has been funded with the European Commission support. This publication reflects only the authors' view. The Commission is not responsible for any use which may be made of the information in this paper.

<sup>2</sup>The consortium is made up of six partners from three European countries: the University of Wolverhampton,

---

United Kingdom, which is the project coordinator; Universität des Saarlandes, Germany; Universidad de Málaga, Spain; Universitätsmedizin Mannheim, Germany; NHS West Midlands Workforce Deanery, United Kingdom; and Hospital Pascual, Spain.

<sup>3</sup>For further information about the TELL-ME project, consult the European Union website (<http://www.adam-europe.eu/adam/project/view.htm?prj=8242>). See also Corpas and Mitkov (2013).

being addressed with the immigration of professionals from Eastern Europe. There is also a growing healthcare demand from European citizens who have moved from their home countries to others, in particular those citizens who have moved from Western Europe.

## 2 Related work

Nowadays, there is a lack of free or reasonably priced self-study courses/resources for medical language for professionals. No currently available courses bring together all necessary features in the way demonstrated by TELL-ME. The products available seem to be split into three broad areas: monolingual courses, meta-communication or communication aids.

*Monolingual courses* exist, for example those offered by the University of Bath or Tokyo Medical University. These are computer-based self-study courses. However, they do not contain all necessary features for real development of efficient medical communication skills, e.g. they may have medical vocabulary exercises but no pronunciation aids or exercises based on real doctor-patient scenarios.

Books and DVDs aimed at helping doctors improve their *meta-communication* skills focus on aspects such as active listening and non-verbal communication, and teaches learners how to handle a range of situations sensitively, e.g. breaking bad news, etc. (Glendinning and Holmström, 2005). While useful, these are monolingual and can only benefit learners with relatively high levels of communication skills (cf. Glendinning and Howard, 2007; Wilmanns and Schmitt, 2002; Karenberg, 2011).

Several *communication aids* exist to help minimise the language barriers between healthcare providers and patients. In most cases these take the form of either databases or computer-based translation tools. An example of this kind of resource is Lebab, which allows users to search for the relevant vocabulary using text and audio. Examples of computer-based translation tools include MedSLT and UniversalDoctor Speaker. With these tools doctors or patients can choose a topic or domain, find the relevant phrases or keywords, and translate them from their own language into the language in which they cannot communicate. As they aim to facilitate communication in typical consultations, phrases and terms focus on characteristic expressions used to describe symptoms, diagnosis and

treatment. These tools might be useful for patients with less well known languages, but are not very useful for vocational purposes.

There are also some platforms which include monolingual and bilingual material such as dictionaries, glossaries, and resources related to medical language. In this line, the Cosnautas<sup>4</sup> platform has been created by translators and health professionals. This platform includes, among other things, a bilingual dictionary, a directory of acronyms, abbreviations and symbols, and a collection of useful links for the medical writer and translator.

## 3 The e-learning platform

The TELL-ME project aims to develop a comprehensive set of learning materials and support tools for teaching vocationally relevant English, Spanish or German to health professionals, helping them in their communication. TELL-ME incorporates new language learning methodologies into an environment which provides users with access to personalised, interactive multi-modal learning materials supported by novel NLP technology and tools. The innovative NLP techniques single out the TELL-ME environment as original and different from all currently available resources for medical language for professionals.

In order to achieve the objective of creating an environment which allows learners to have access to a complete and personalised learning solution, we have designed an e-learning platform developed through Moodle. This e-learning platform consists of three courses, one for each of the target languages of the project. The courses, which can be accessed through the following URL: <http://www.tellme.uni-saarland.de/>, are structured around a set of 3D animation videos created using Xtranormal<sup>5</sup> from real medical interviews in first-contact situations, and take into account the ten most frequent specialties (emergency, internal medicine, ophthalmology, general surgery, vascular surgery, plastic surgery, traumatology, otorhinolaryngology, urology and gynaecology). There are a total of 900 dialogues, i.e. 300 interviews per language between the doctor and the patient who comes for the first time for consultation. Each dialogue covers the patient's family and medical history, symptoms, diagnosis and additional tests (if any)

<sup>4</sup> <<http://www.cosnautas.com/index.php>>.

<sup>5</sup> <<http://www.xtranormal.com/>>.

as well as the prescribed treatment, according to standard medical uses.

In addition, the courses include a section of general resources, such as medical image dictionaries and collaborative medical glossaries which are linked to exercises to improve vocabulary and develop memory, reasoning and understanding, as well as other skills.

Each unit is divided into three parts: a video clip, related exercises created from the project's medical corpus and related and complementary interactive NLP-based exercises also derived from the medical corpus. The last section of each unit is a self-learning and self-assessment tool that contains various types of exercises, e.g. fill-in-the-blank, click and drag exercises or multiple-choice questions. These automatic exercises will be generated according to users' demand, using NLP tools and techniques.

For example, the generation of multiple-choice questions is based on a ground-breaking computer-aided procedure which makes use of language resources such as electronic corpora to identify important concepts in text and generates questions about these concepts as well as multiple choice distractors semi-automatically (Mitkov et al., 2006). This procedure consists of three stages: (i) in the first stage the domain-specific terms are identified, (ii) in the second stage a stem is generated from a declarative finite clause using some transformational rules and, finally (iii) in the third stage distractors are selected using a strategy based on the semantic similarity.

The fact that this kind of exercise is automatically or semi-automatically generated in accordance with the users' demand and needs ensures individualised learning and brings added value to the TELL-ME e-learning platform.

Additionally, the TELL-ME corpora, consisting of over 20 million words, are a valuable teaching resource. They will be available on the platform through an open-structure corpus management system (cf. Orasan and Krishnamurthy, 2000). They are specialised medical comparable corpora, composed of monolingual untranslated documents, which are comparable to each other, because they have a similar layout and design. They have been compiled from electronic sources using BootCat, according to an *ad hoc* virtual corpus compilation protocol (Corpas Pastor, 2008).

The TELL-ME e-learning platform is currently being tested out on trainees by the medical partners and will be disseminated to key national medical policy makers. As a first part of the evaluation, an on-line trilingual questionnaire<sup>6</sup> has been designed to test out the platform by health professionals in order to receive users' feedback. For this, we have provided free access to a sample of resources which allows users to try the platform and give us feedback. In order to contact medics and other health professionals, the questionnaire has been distributed by medical partners to relevant e-mail lists and healthcare organisations in Spain, Britain and Germany. The medic partners have also organised in-house training activities for doctors willing to participate. In addition to practitioners' feedback, the questionnaire has been also distributed among students and University lecturers in Healthcare degrees, such as Medicine or Nursing, in the three target countries.

In addition, a blog has been created within the TELL-ME website that reflects the users' impressions after testing the prototype.

At present we are analysing the users' feedback received so far. Very preliminary results show that users find the platform quite useful, especially the *vocabulary* tab and the vocabulary exercises, and they also think that the course content is well organised by medical domains. Once the analysis of users' feedback is finished, we will make the necessary improvements for the final version of the platform, which will be presented in the final dissemination event in Brussels on October 31. In addition, more evaluation results will be included in the final version of the paper, if accepted.

#### 4 The TELL-ME website

In order to achieve all the objectives, the consortium established six work packages: Management, User Requirements, Implementation, Quality Assurance, Exploitation and Dissemination. Each partner is responsible for specific packages, but the consortium as a whole also participates in all the packages. The project website (<http://tellme-project.eu/>) assists the collaborative work of a geographically dispersed multidisciplinary team and provides public ac-

<sup>6</sup> These questionnaires are available on the following URLs: English (<http://encuestas.sci.uma.es/28429/lang-en>), Spanish (<http://encuestas.sci.uma.es/28429/lang-es>) and German (<http://encuestas.sci.uma.es/28429/lang-de>).

cess to the project outputs. More specifically, it consists of a free access public area and a private area restricted to the project researchers.

The public area contains general information about the project and a survey section and a blog, which reflect the users' feedback after testing the prototypes and using the didactic materials.

The private area is the most important part of the website and its main function is to provide support for the members of the project. This area is designed with a very specific set of objectives: firstly, to establish a unified, centralised repository of useful materials for researchers (the corpus, the doctor-patient dialogues, survey results and partial studies, etc.), and, secondly, to provide the principal investigator and the different partners with a way to plan, manage and monitor the activities of the consortium, organisational and management issues, timelines and calendars, summaries of meetings, presentations and the decisions reached.

The private area is linked to Google Analytics. Thus, it is possible to retrieve statistics on the visitors to the TELL-ME website and analyse their behaviour in real time.

The website is structured in accordance with the initially established dissemination plan for the project. While the public area is linked to social networks and gives visibility to the activities, publications and final products of the project, the private area allows a detailed analysis of the website's visits. This is complemented by brochures and leaflets (published in three languages) that have been developed within the project, distribution lists for medical users, newsletters, press releases, etc.

## 5 Conclusion

This paper has presented an overview of the TELL-ME project, with special attention to the resulting e-learning platform. The TELL-ME platform is designed to provide users with a rich and versatile tool which allows them to acquire adequate competence in speech and doctor-patient interaction in a new, technology-based learning environment, at their own pace and according to the needs marked by their own learning process. This e-learning platform is based on innovative NLP tools and techniques. Among other things, computer-assisted content generation, e.g. the semi-automatic generation of multiple-choice questions, singles out the TELL-ME e-learning platform as original and different from

all currently available resources for medical language for professionals.

## Acknowledgments

The work described in this paper is supported by the Lifelong Learning Programme of the European Commission (ref. no. 517937-LLP-2011-UK-LEONARDO-LMP) and also supported by Andalucía Tech's programme "Atracción de investigadores de reconocido prestigio".

## References

- Corpas Pastor, G. and Mitkov, R. 2013. "Nuevos entornos formativos para la comunicación médico-paciente en un contexto transnacional". *Uciencia*. 11. 44-46.
- Corpas Pastor, G. 2008. *Investigar con corpus en traducción: los retos de un nuevo paradigma*. (Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation, 49). Frankfurt, Berlin and New York: Peter Lang
- Glendinning, E. and Holmström, B. 2005. *English in Medicine: A Course in Communication Skills*. 3<sup>rd</sup>. ed. (with CD-ROM). Cambridge: Cambridge University Press.
- Glendinning, E. and Howard, R. 2007. *Professional English in Use Medicine*. Cambridge: Cambridge University Press.
- Karenberg, F. K. A. 2011. *Fachsprache Medizin im Schnellkurs: Für Studium und Berufspraxis*. 3.<sup>a</sup> ed. Stuttgart and New York: Schattauer,
- Lebab Systems. <http://lebab.es/>. As of 23 September 2013.
- MedSLT, 2007. <http://www.issco.unige.ch/en/research/projects/me/dslt/index.html>. As of 23 September 2013.
- Mitkov, R.; Ha, L.A.; Karamanis, N. 2006. "A Computer-aided Environment for Generating Multiple-choice Test Items". *Journal of Natural Language Engineering*. 12 (2). 177-194.
- Orasan, C. and Krishnamurthy, R. 2000. "An Open Architecture for the Construction and Administration of Corpora". *Proceedings of LREC'2000*. Athens, Greece. 793-800.
- UniversalDoctor and Tools, 2011. <http://www.universaldocor.com/>. As of 23 September 2013.
- Wilmanns, J. C. and Schmitt, G. 2002. *Die Medizin und ihre Sprache: Lehrbuch und Atlas der Medizinischen*. Landsberg: AG & Ko, KG.