STUDY IN APPLICATION OF NATURAL LANGUAGE PROCESSING IN MARITIME COMMUNICATIONS

Rosa de la Campa Portela[1], Benigno A. Rodríguez Gómez[2], Mª Carmen Meizoso López[3] and Begoña Rivero Martínez[1]

ABSTRACT

Recent research on maritime accidents done at the School of Nautical Studies at the University of A Coruña (Spain) shows that nearly 20 per cent of accidents in maritime settings have been due to, among other causes, communication problems derived from the lack or misuse of a common language.

Moreover, automatic speech translation, a technology that combines speech recognition and automatic translation, has been for more than a decade the focus of research as a tool for improving communication in different settings.

Similarly, this technology can be applied, to a greater or lesser extent, in communicative processes that take place in the maritime workplace so as to minimise problems stemming from multilingual environments, especially those in which communication takes place via the use of radio devices, in which the inclusion of an automatic translator could enable two people of different nationalities, for example, to communicate with each other in their own native languages. This article describes a system of this type whose feasibility is being studied at the Universidad de A Coruña, under the auspices of a university-financed project called "Language Industries Applied in the Maritime Sector".

What this project seeks, therefore, is to study the legal and technical possibilities, as well as the commercial suitability of developing an automatic translator for oral communications in the maritime sector, and in such a case, to

---

[1] Departamento de Ciencias de la Navegación y la Tierra. ETS Náutica y Máquinas. Universidad de A Coruña (rosamary@udc.es). [2] Departamento de Ingeniería Industrial. Escuela Universitaria Politécnica. Universidad de A Coruña, Ferrol (benigno@udc.es). [3] Departamento de Ingeniería Industrial. Escuela Universitaria Politécnica. Universidad de A Coruña, Ferrol (mmeizoso@cdf.udc.es)

establish the foundations for such a development, as well as to study the implications that the use of a device with such features would have on maritime safety. Likewise, this project seeks to develop the simulator for learning and using maritime phrases and for the creation of the corresponding curricular material.

Key words: maritime communications, human language technologies, automatic translation, comunicaciones marítimas, tecnologías del lenguaje, traducción automática.

INTRODUCTION

Recently, at the School of Nautical Studies at the University of A Coruña, a statistical study has been conducted on the subject of maritime accidents between 1994 and 2001 (de la Campa, 2003) with the aim of determining a possible relation between communication problems stemming from the lack of or misuse of a common language, and these accidents.

The study's results show that communication problems that stemmed from the language were a leading factor in approximately 20% of maritime accidents.

However, using different risk evaluation techniques and based on the data in the maritime accidents reports, it was determined that the risk of a maritime accident happening due to communication problems stemming from language is at a tolerable level. Hence it seems advisable to establish corrective means that lead to lowering this risk to the lowest possible level. Hence, the School of Nautical Studies at the University of A Coruña has proposed a series of corrective measures that mainly focus on the following points:

— The improvement of the quality of the process of teaching/learning maritime English;

— The use of new technologies, mainly the so-called language industries, in order to improve the communication process in the maritime sector;

— Heightening the awareness of the responsibility of maritime managers in relation to training, contracting and education of multilingual crews, as well as establishing and maintaining an adequate security policy.

This article focuses on the second point mentioned above and endeavours to offer an overall view of the various possibilities that the so-called language industries can offer the maritime industry with the aim of improving security in this sector.

LANGUAGE INDUSTRIES: GENERALALITIES

Computational linguistics is the study of computer systems that are useful for the understanding and generation of natural languages, which based on the language processes enable computers to undertake activities of a linguistic nature (such as translating), process textual data, or make personal access to stored data easier (Grishman, 1986).

Those products, techniques, services or activities which require an automatic treatment of natural language, that is, which use computational linguistics as a resource, but whose aim is not the description of the language, are what is known nowadays as language industries (Cabré, 1993). These language industries mainly proceed from the combination of two fields of study: linguistics and computer science, along with the contribution of other sciences such as engineering, psychology, logic and documentation.

These linguistic products contribute to perfecting the interaction and use of computer systems, to improving the effectiveness of the assimilation, analysis, selection, use and presentation of information, and to making means of generation and translation of natural language available (Oficina del Español en la Sociedad de la Información, 2000).

The products that are most directly useful for the maritime industry are those that are related to automatic translation, teaching, natural language interfaces and speech recognition and reproduction systems.

AUTOMATIC TRANSLATION

According to Hutchins (1986), there have been and are numerous reasons for attempting to achieve automation of the translation process. Perhaps the most important reason is the need to improve univocal communication between specialists from different nationalities, in order to decrease the amount of time it takes to translate technical documents and to increase the number of translations of this type into several languages.

The translation process can be done on any type of text, but automatic translation is more useful when it deals with technical texts, since they usually do not use rhetorical and literary tropes and are limited to well-defined spheres of language, and are characterised by a style that is limited to a series of determined structures.

However, and despite the wide range of methods, all of the automatic translation processes pose problems that are not associated with the technology, but rather with the language. Linguistic problems related to translation are those that deal with lexical ambiguity, syntactic complexity, the difference of vocabulary between languages, and elliptical or incorrect grammar constructions.

These problems can be reduced by implementing certain restrictions on the original text, such as the use of controlled languages by adapting the original texts with constructions and vocabulary that the programme can translate, or by restricting the types of texts by designing programmes for a specific and well defined variety of texts.

With regards to this point, and within the maritime sector, there is the possibility of studying the viability and usefulness of the automatic translation of communication processes based on the use of the Standard Marine Communication Phras-

es of the IMO, with the aim of avoiding misunderstandings due to, for example, differences in the pronunciation of the English language or to the limited knowledge of these phrases in said language.

NATURAL LANGUAGE PROCESSING

The technology of natural language processing "covers a broad range of activities with the eventual goal of enabling people to communicate with machines using natural communication skills" (Cole, 1997), and includes systems that are able to do speech synthesis processes (automatic generation of speech from a symbolic representation) and speech recognition (speech conversion into a symbolic representation). The latter process, that is, speech recognition, can take place on three different levels: recognition of the language that is being spoken, recognition of the content or meaning of what is being spoken, and recognition of the speaker (Mariani, 1991).

All of these processes could be applied in the medium term to improve activities in the maritime sector. Hence for example, many of the activities that require, during the operation of a ship, man-machine communication processes, could be controlled by voice recognition and/or speaker recognition devices so that they can assure that the orders that are given are understood and executed correctly. Likewise, a device that incorporates automatic translation to speech recognition and synthesis could act as the "interpreter" in operations that require very controlled language and take place in a multilingual environment.

THE PROJECT OF THE SCHOOL OF NAUTICAL STUDIES AT THE UNIVERSITY OF A CORUÑA

Introduction

As we stated before, all of these processes can be applied to a greater or lesser extent in maritime communications, with the goal of minimising the problems that arise from a multilingual environment, especially in those communications that take place via radio devices, in which the inclusion of an automatic translator would enable two people of different nationalities, for example, to communicate with each other in their native tongues. The viability of a system of this type is being studied at the School of Nautical Studies at the University of A Coruña in a project financed by this university, called "Language industries applied to the maritime sector".

This system could be of great value when the messages that are sent have been standardised. Consequently a quality translation and a perfect comprehension are guaranteed.

The development of a translator that can be used commercially requires a high level of "maturity" of the utilised technologies, as well as a state of legality that does not prevent its use. Even if these conditions are not fulfilled, the present state of

the technology enables, following the same steps, the development of simulator systems for the learning and use of the standard phrases.

What this project intends, therefore, is to study the legal and technical possibilities, as well as the commercial suitability of developing an automatic translator for oral communication in the maritime sector, and to establish the bases for this development. At the same time we seek to study the implications that the use of a device of such characteristics would have on maritime security. Likewise we wish to pursue the development of the simulator for learning and using maritime phrases and the preparation of the corresponding curricular material. In order to lay the foundation for this study, as well as for the development, it is necessary to make up a list in the form of a database of standard phrases that are commonly used in the marine sector.

The Project's Objectives

The aim of this project is to open up a line of research on the use of ICT [Information and Communication Technologies] in order to improve oral communication processes that use the standard phrases in the maritime sector.

These are the main objectives that have been established:
1. The creation of a digital sound database of standard phrases for maritime communication.
2. The development of a learning support system, as well as of a system of the usage of standard phrases.
3. The study of:
   a) the legal and technical possibilities,
   b) the commercial suitability of the development of an automatic translator for oral communication in the maritime sector,
   c) (if it is fitting) the establishment of the bases of said development, and the implications that the use of a device of such characteristics would have on maritime security.

All of these objectives share stages of development and interact among each other, and thus mutually reinforce each other. The attainment of these objectives would be the result of working on the following specific points:
1. Establish the univocal communication needs in the maritime sector and the implications of these communications on maritime security.
2. Study the International Maritime Organisation's provisions on the establishment of a common language for maritime communications as well as the use of standard vocabulary in such communications.
3. Study the suitability of the development of an automatic translator for maritime communications, taking into account the following:
   a) The legal implications of implementing this system.
   b) The implications on maritime security of the use of this system.

4. Establish the linguistic needs for the development of an automatic translator for maritime communications on the basis of:
   a) The attainment and analysis of genuine spoken samples of ship-to-ship and ship-to-land maritime communications.
   b) The study of genuine standard vocabulary in maritime communication on the basis of collected samples.
   c) The recognition and association between standard and non-standard phrases through the comparative study of collected samples and the standard glossaries issued by the International Maritime Organisation.
5. The development of a system of assessment (simulator) that enables the studying of the technological needs for the development of an automatic translator for maritime communications. This automatic translator will be readapted as a learning support system and as a system of the usage of standard phrases. An analysis of the features will focus on:
   a) The transformation of standard voice-to-text phrases and vice-versa, based on genuine samples, that should determine the rate of accuracy and reliability of the system.
   b) The study of disturbances and possibilities of elimination based on genuine samples.
6. Adaptation as a learning-assistance system and usage of the standard phrases. Preparation of the course
7. The establishment of the equipment needs in order to develop an automatic translator for maritime communications: physical characteristics of the equipment, composition and cost.
8. The study of the commercial prospects of developing such a translator.

Methodology

The current project bases a large part of its content and development on collecting and analysing genuine samples of communications in the maritime sector. These samples should be obtained on site, that is, on board the merchant ships in order to register the genuine ship-to-ship and ship-to-land communications, and in maritime traffic control towers or in other locations for genuine ship-to-land communications.

The presence on board of several members of the research team will be essential. They will be responsible for taking these samples and for beginning to process them.

Then these samples will be immediately analysed linguistically in the laboratory, and used to run different tests on language processing and to eliminate noise.

The conclusions of this analysis, and its comparison with the availability of speech processing technologies, translation, voice synthesis, all of them in real time will enable one to establish the equipment needs for the development of an automat-

ic translator for maritime communications: characteristics of the equipment, composition and cost.

The positive assessment of the technical and legal aspects for the development of the translator, leads to the study of commercial prospects, through the search for companies in the telecommunications sector that are interested in prototyping and undertaking the complete development of the system.

On the other hand, the negative assessment of the technical and/or legal aspects, does not permit a commercial translator in the short term, but it does not impede the development of the learning-assistance simulator system and of the system of usage of standard maritime phrases, developed on a PC and which consists of readapting the translator-assessment system that is complemented by the preparation of curricular materials.

The following is a summary of the work that has been done until now on each point:

1. Establish the univocal communication needs in the maritime sector and the implications of these communications on maritime safety.

The international character of maritime commerce entails getting people from very different nationalities, languages and cultures in contact, who will work efficiently in a complex setting.

The most common way people interact is, without a doubt, through oral communication, which in this environment becomes hindered by the aforementioned coexistence of different nationalities.

Communication on board a ship can take place in the crew members' native tongue, as long as they come from the same country or share a common mother tongue. However, multinational and multilingual crews are more and more usual each time, a factor that obliges ships to establish a common work language for the safe operation of the ship. Knowledge and the ability to use said language must be shared by the entire crew. It is with this aim in mind that English is the language that is chosen in the majority of the cases.

As for the interaction between ships or between these and shore services, English has become the means of communication at sea. As a result, the entire operation of the international maritime industry, safety at sea and the protection of the sea environment, depend in many aspects on the level of knowledge and use of this language. According to Loginovsky (2002), the globalisation of the maritime industry requires sailors to be more highly qualified and to have a higher level of training and certification than previously required. Communication is closely linked to safety nowadays, and communication problems have negative effects not only on the operability of the ship, but also on the social life on board.

We are aware, according to the bibliography consulted, that these communication problems can have several origins, among which the following stand out:

— The crew members' linguistic incompetence, as a result of a low level of English;
— Differences in the level of knowledge of English among the crew members;
— Difficulty to assimilate different accents and understand new and particular forms of English;
— Environmental difficulties that are inherent to this sector, such as voice distortions via VHF, or the high noise level that hinders communication in the engine room.

2. Study the International Maritime Organisation's provisions on the establishment of a common language for maritime communications as well as the use of standard vocabulary in such communications

Despite technological advances in means and procedures to facilitate the exchange of information between ships and between ships and shore services, such advances cannot have the desired positive effect if attention is not paid to the development of the crew members' linguistic skills: "linguistically under qualified officers, either on deck or in the engine room, create a danger to their ships, crews and passengers, to other vessels and to the marine environment" (Trenkner, 1996: 125)

The International Maritime Organisation is, obviously, aware of this situation, and recognising its responsibility with regards to maintaining safe navigation and maritime commerce, they have attempted for many years to improve verbal communication in this sector by championing the standardisation of use at the international level firstly with the approval of the Standard Marine Navigation Vocabulary (SMNV) in 1973, and currently with the renewal and expansion of this vocabulary as a result of the approval in 2001 of Standard Maritime Communication Phrases.

Moreover, with regards to the use of natural language recognition, it is necessary to keep in mind that the positive achievements of this technology are due to, among other things, the use of the so-called "controlled languages" among interlocutors.

These languages stand out mainly because of their simplicity based on the use of a limited vocabulary and the simplification of grammar rules.

In the maritime sector, Standard Maritime Communication Phrases, created by the International Maritime Organisation, fulfil the necessary conditions to be considered a controlled language, and are thus suitable for use in an automatic speech translation system.

3. Study the suitability of the development of an automatic translator for maritime communications.

One of the main objectives of this project is to study the feasibility of the application of new computer technologies and communication in maritime communications, with the aim of improving its effectiveness as well as to highlight the

aspects of maritime safety related to these technologies. In this sense we have summarised in this section the findings of a Delphi study undertaken to establish the legal, commercial and technical feasibility of the use of this type of resources, by consulting maritime experts who mainly work professionally in an international communication environment.

From among all the conclusions gathered in this study it is interesting to highlight the following points:

— A large majority of experts agree that oral communication between people of different nationalities can entail a safety problem in the maritime workplace; they also consider that the most problematic communications are those that take place with the exterior: ship-to-ship and ship-to-shore.

— Moreover, general knowledge of English is considered the most problematic area in exterior communications, as differences in pronunciation and accent and knowledge of standardised vocabulary are considered to be factors that are also very problematic in the development of such communications.

— To improve these exterior communications experts insist on the need, above all, to attach greater importance to the teaching of maritime English; they also indicate the need to involve maritime administrators and the appropriate authorities so that they insist upon a minimum level of English from those involved in such communications. The use of a standard international maritime language is highlighted as the third best solution to these communication problems.

— Once a possible design for equipment that would use speech recognition and automatic translation as an aid to exterior oral communications is proposed, almost 70% of the experts would recommend the development of equipment with these features, and 85% would offer to try out the equipment in case such equipment were developed and in a trial period.

— Almost 80% of the experts surveyed would select equipment that is independent of VHF, which in any case could be adapted to it whenever it were necessary.

— As for the more problematic areas that would have to be kept for the proper running of the equipment, the reliability of the voice recognition system and the time taken up for translations are underlined as the most problematic and difficult technical and linguistic impediments to overcome.

— Finally, according to the experts who were surveyed, user friendliness, price and reliability are, in this order, the most important features that are to be kept in mind when deciding on the possibility of acquiring equipment with such characteristics.
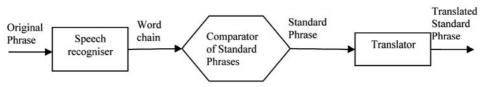
4. The attainment and analysis of genuine spoken samples of ship-to-ship and ship-to-land communications. The study of the genuine use of standard vocabulary in maritime communications based on the collected samples. The recognition and association between standard and non-standard phrases through a comparative study of the collected samples and the International Maritime Organisation's standard vocabulary.

This project bases a large part of its content and development on the gathering and analysis of real samples of communications in the maritime sector. These samples have been obtained in situ, that is to say on board merchant ships in order to register real ship-to-ship and ship-to-shore communications, and in traffic control towers or by means of other devices for real ship-to-land communications.

Up to now almost a hundred conversations have been obtained between ships, and between ships and land services. These have been transferred to a database for further analysis. Finally as well, a comparative analysis is being done of real vocabulary used in these conversations and vocabulary contained in the Standard Maritime Communication Phrases, as well as a study on the use of these phrases in real conversations.

5. The transformation of voice-to-text standard phrases, the study of disturbance and possibilities of elimination based on genuine samples. The transformation of text-to-voice standard phrases

We see our simulator as a system that links several sequential processes, in such a way that taking as the point of entry a phrase said by a speaker in his or her native tongue (Spanish in this case), it transforms it into a standardised phrase in meta(English), as shown in Graphic 1.

Graphic 1: sequential processes of simulator

Our aim is to assess current technologies in order to verify the reliability of the entire process. To this end we initially assessed each one of the stages in an independent manner:

   a) Speech recogniser; even though our problem is limited to a controlled language, we selected a commercial application for continuous language recognition with an extensive vocabulary, Dragon Naturally Speaking version 8. This processor requires the speaker to receive training. In order for the application to run under optimum conditions after the initial training

period, the speaker was given the specific vocabulary of standard phrases, and the group of phrases as a model of the language. Under these conditions the rate of accuracy in the recognition of phrases in continuous flow was approximately 97%. It must be noted that the processor shows a preference for a particular type of speaker during the training phase, and the required signal/noise ratio is very high.

b) Comparator with standard phrases; this module takes the word chain provided by the Recogniser, and establishes the similarity with the group of standard phrases, choosing from among these the one that is closest to the word chain. This module is in the development stage; until now the frequency of each word of the vocabulary has been established as well as the concordance between groups of words; these parameters allow for the reconstruction of a standard phrase starting from some of its elements, and thus they make it unnecessary for the word chain to be exactly the same as a particular standard phrase. What remains to be done for the complete development of the module is to add the work on the groups of real conversations, which are already at our disposal but have not yet been sufficiently processed.

c) Translator; this element is not difficult to develop since the translation is not real, but rather once the comparator identifies a standard phrase in any language, its translation into any other language is basically a matter of a consultation or a group of consultations (if the phrase contains variables) of a database.

The assessment of the state of the technology in order to undertake the proposed development is the following: The voice recognition system that is dependent upon the speaker in low noise conditions is sufficiently developed in order to undertake the recognition of standard maritime communication phrases.

It is now necessary to continue the study on the following points:
— Noise conditions
— The sturdiness of the recognition in the presence of noise
— The strategies of noise filtration (conventional or adapted)
— The similarity between used and standard phrases

We have assumed as significant sources of error the following: the speaker's use of non-standard phrases, and the errors made by the speech recogniser module, due either to surrounding noise that cannot be eliminated or to other causes.

The elimination of the aforementioned errors should take place in two ways: a technological solution involving the development of an efficient filtering system, and also the creation of a comparator module that would maximise the rate of accuracy in the selection of standard phrases; the second means is educational in nature as the officers who are involved in the communications should know and be able to use standard phrases correctly, according to the IMO's recommendations.

6. Adaptation as a learning-assistance system and usage of the standard phrases. Preparation of the course

With regards to the teaching of Standard Phrases, which is the area that interests us the most in terms of the objectives of this research, the IMO's own resolution (2001) indicates the following points of reference:

— The objective of Standard Phrases is not to provide a detailed study plan of maritime English.
— Part A should be an essential element of any type of curriculum designed to fulfil the requirements of training Agreement 78/95
— Standard Phrases are to be taught and learned in a selected way, according to the specific needs of the user, and not as a whole.
— Teaching should be based on common practices in the maritime sector and should be imparted in accordance with modern methods of language learning that are considered to be appropriate.

Moreover, keeping the second point in mind, in which Standard Phrases were considered to be a controlled language, suitable for use in the automatic translation process, it is important to indicate that with regards to the learning and teaching of controlled languages, as noted by Mitamura (1999), in order for the use of a controlled language to be successful, it is critical that the users be able to accept the notion of a controlled language, and that they be willing to receive training in the controlled language in question.

According to this author, it seems that the users who are trained and use the controlled language habitually attain better results and a higher level of communicative productivity.

Thus the need to establish appropriate training methodologies in the use of Standard Phrases becomes clear in order to comply with the International Maritime Organisation's regulations and also as a provision for the use of these phrases as a controlled language of maritime communications which is subject to being processed by recognition and speech translation equipment.

We have therefore dedicated this point to studying several of the methodologies that are used nowadays in the teaching of English in the maritime sector, which can be used, and in fact some of them are being used to teach the IMO's Standard Phrases, and which use, furthermore, the new information and communication technologies as a teaching tool in order to follow the Organisation's recommendation in relation to appropriate and modern teaching methods.

7. The establishment of equipment needs for the development of an automatic translator for maritime communications: equipment characteristics, composition and cost.

The type of communication that our research focuses on is done with VHF and VHF-DSC (Digital Selective Calling) equipment. The equipment that we propose as the final product must, necessarily, respect the specifications of these types of equipment, not impede their use and be activated upon the operator's request.

Given the current state of technological development, a simple solution is a PC built-in to the VHF equipment itself, while the external aspect of the system would include a feedback screen for the user so that he or she can be aware of the information that is being transmitted when using the automatic translation, or could even be informed that he or she should desist from using this mode.

This equipment will permit sending synthesised voice with the standard phrase and will also be able to emit a digital code that will enable receiving equipment of similar characteristics to recognise the phrase directly.

A more efficient solution could be the development of a processing card with a DSP as a nucleus of the same card.

Another one of the needs that has to be evaluated in terms of noise is the incorporation of a multi-microphone system that allows for an adaptable filtration in order to improve the signal/noise ratio.

8. The study of the commercial prospects of developing the translator.

During this first year of the development of this project, contact has been made with Arteixo Telecom, which after having shown interest in the development of the equipment is presently assessing the feasibility of its commercialisation. We trust their judgment, given the implementation of their equipment in the naval sector, even if until now they have focused on developments for the Fishing sector and the Navy.

At this point one might think of multiple scenarios, but we shall only consider two assuming that the rest are variations of these.

A technologically favourable scenario: If the development of this equipment is possible, this situation should be faced with the legal scenario: What bodies should approve the use of such type of equipment? We must not forget that its usage will take place in an international setting. This undoubtedly generates a significant delay between the technological availability and the possibility of using it.

A technologically unfavourable scenario: It is not possible with the current state of technology to create equipment with the appropriate characteristics. The eventual increase of features will make this scenario evolve towards the aforementioned situation. In this case the technological and legal delays can overlap, if we begin to work immediately on the latter aspect. In any case it is foreseeable that the delay due to legal restrictions will always be greater than due to technological reasons, because while the product is not available one cannot assess its effect on maritime security, which is the ultimate objective.

Either of the two scenarios leads to a similar situation, that is, the fact that the product will not be ready immediately; however given the fact that the possibility of creation has been proven, the development depends on the industrial's sector involvement. Thus, in the last part of the project and once the results of the prior stages have been evaluated, telecommunication companies will be contacted, especially those that are involved, or wish to be involved, in the maritime sector, to assess the possibility of collaboration and development.

REFERENCES

Cabré, M.T. (1993): *La terminología: teoría, metodología, aplicaciones*. Barcelona: Editorial Antártida/Empuries

Cole, R. (1997): "Introduction to survey of the state of the art in human language technology". En: Cole, R. (ed). *Survey of the state of the art in human language technology*. Cambridge: Cambridge University Press.

De la Campa Portela, R. (2003): "Communication breakdown". *Marine Engineers Review*. Dic/ene 2003, 34-35

Grishman, R. (1986): *Computational linguistics: an introduction*. Cambridge: Cambridge University Press.

Hutchins, W.J. (1986): *Machine translation: past, present, future*. Chichester: Ellis Horwood.

Mariani, J. (1991): "Las tecnologías del lenguaje". En: Vidal Beneyto, J. (ed). *Las industrias de la lengua*. Madrid: Pirámide, 326–369.

Oficina del Español en la Sociedad de la Información (ed) (2000): *Como aprovechar la fuerza del lenguaje*. Madrid: Instituto Cervantes.

Loginovski, V.A. (2002): "Verbal communication failures and safety at sea". *IAMU 3rd General Assembly*. Examinado en diciembre 2002. http://bell.mma.edu/~iamu2002/

Mitamura (1999): "Controlled language for multilingual machine translation". En: *Proceedings of Machine Translation Summit VII*. Singapur, 1999

Trenkner, P. (1996): "IMO–SMCP an attempt to meet increased communication requirements of ships' officers". *9th International Conference on Maritime Education and Training*. Japón: IMLA, 125–130.

OMI (2001): Res. A 918(22) "Stadard Marine Communication Phrases"

APLICACIÓN PARA PROCESAR EL LENGUAJE NATURAL EN LAS COMUNICACIONES MARÍTIMAS

RESUMEN

Un reciente estudio sobre accidentes marítimos llevado a cabo desde la Escuela Técnica Superior de Náutica y Máquinas de la Universidad de A Coruña, muestra que cerca de un 20% de los accidentes en el ámbito marítimo han sido debidos, entre otras causas, a problemas de comunicación derivados de la falta o mal uso de una lengua común.

Por otro lado, hace más de una década que la traducción automática del habla, tecnología que combina el reconocimiento del habla y la traducción automática, es objeto de investigación como herramienta para la mejora de las comunicaciones en diversos ámbitos.

Esta tecnología puede ser aplicada, de igual forma, en mayor o menor medida en los procesos comunicativos que tienen lugar en el ámbito marítimo con el fin de minimizar los problemas derivados de los ambientes multilingües, especialmente aquellos en que las comunicaciones tienen lugar mediante el uso de aparatos de radio, donde la inclusión de un traductor automático podría permitir a dos personas de diferente nacionalidad, por ejemplo, comunicarse entre ellos en sus lenguas nativas. El siguiente artículo nos describe un sistema de este tipo cuya viabilidad está siendo estudiada en la Universidad de A Coruña, al amparo de un proyecto financiado por dicha universidad denominado "Las industrias de la lengua aplicadas al ámbito marítimo".

Lo que el mencionado proyecto pretende, por lo tanto, es estudiar las posibilidades legales y técnicas, así como la conveniencia comercial del desarrollo de un traductor automático para las comunicaciones orales en el ámbito marítimo, y en su caso, establecer las bases de dicho desarrollo, a la vez que estudiar las implicaciones que sobre la seguridad marítima tiene el uso de un dispositivo de tales características. Igualmente se persigue el desarrollo del simulador para el aprendizaje y uso de las frases marítimas y la elaboración del material curricular correspondiente.

INTRODUCCIÓN

Recientemente, en la Escuela Técnica Superior de Náutica y Máquinas de la Universidad de A Coruña, se ha llevado a cabo un estudio estadístico sobre accidentes marítimos acaecidos entre los años 1994 y 2001 con el fin de determinar una posible relación entre los problemas comunicativos derivados de la falta o mal uso de una lengua común y dichos accidentes.

Los resultados obtenidos de tal estudio indican que, en aproximadamente un 20% de los accidentes marítimos, los problemas comunicativos derivados del lenguaje fueron un factor desencadenante.

Por otro lado, mediante el uso de diversas técnicas de evaluación de riesgos y basándose en los datos obtenidos de los informes sobre accidentes marítimos mencionados, se determinó que el riesgo de ocurrencia de un accidente marítimo debido a problemas comunicativos derivados de la lengua se encuentra en un nivel tolerable, por lo que parece recomendable el establecimiento de medidas correctoras que conlleven la disminución de dicho riesgo al mínimo nivel posible.

Así pues, y desde la propia Universidad de A Coruña, se han propuesto una serie de acciones correctoras centradas principalmente en los siguientes puntos:

— La mejora en la calidad del proceso enseñanza/aprendizaje de inglés marítimo.

— La utilización de las nuevas tecnologías, principalmente aquellas denominadas industrias de la lengua, para la mejora del proceso comunicativo en el ámbito marítimo.

— La exaltación y concienciación sobre la responsabilidad de los gestores marítimos en cuanto a la formación, contratación y educación de tripulaciones multilingües, así como sobre el establecimiento y mantenimiento de una adecuada política de seguridad.

El presente artículo se centra en el segundo de los puntos arriba indicados y pretende presentar una visión global de las varias posibilidades que las llamadas industrias de la lengua pueden ofrecer a la industria marítima con el fin de mejorar la seguridad en este ámbito.

OBJETIVOS

La tecnología del procesamiento del lenguaje natural cubre un amplio rango de actividades cuyo propósito es permitir a las personas comunicarse con las máquinas usando las habilidades propias de la comunicación natural, e incluye los sistemas capaces de realizar los procesos de síntesis del habla (generación automática del habla a partir de una representación simbólica) y reconocimiento del habla (conversión del habla en una representación simbólica). Este último proceso, el reconocimiento del habla, puede llevarse a cabo en tres niveles diferentes: reconocimiento de la lengua en la que se habla, reconocimiento del contenido o significado de lo que se habla, y reconocimiento del hablante.

Todos estos procesos podrían aplicarse, en un plazo de tiempo medio, en la mejora de las actividades en el ámbito marítimo. Así por ejemplo, muchas de las actividades que requieren, durante la operación del buque, procesos comunicativos hombre-máquina, podrían ser controladas mediante dispositivos de reconocimiento de la voz y/o reconocimiento del hablante, de manera que aseguren que las órdenes dictadas son correctamente entendidas y ejecutadas.

De la misma forma, un dispositivo que incorpore la traducción automática al reconocimiento y síntesis del habla podría realizar la función de "intérprete" en operaciones que requieran un lenguaje muy controlado y se desarrollen en un ambiente multilingüe.

Así mismo la aplicación de estos procesos a las comunicaciones marítimas permitirían minimizar los problemas surgidos por la multilingualidad del ámbito, sobre todo en aquellas comunicaciones que se realizan a través de aparatos de radio, en los que la inclusión de un traductor automático permitiría, por ejemplo, que dos personas de nacionalidades diferentes se comunicasen en sus lenguas nativas. La viabilidad de un sistema de este tipo está siendo estudiada desde la Escuela Técnica Superior de Náutica y Máquinas de la Universidad de A Coruña, a través de la realización de un proyecto financiado por dicha Universidad, denominado "Las industrias de la legua aplicadas al ámbito marítimo".

Este sistema podría ser de gran utilidad cuando los mensajes que se emiten se encuentran normalizados. De esta forma se garantiza una traducción de calidad y un entendimiento perfecto.

El desarrollo de un traductor utilizable comercialmente requiere un grado elevado de "madurez" de las tecnologías empleadas, así cómo un estado de legalidad que no imposibilite su uso. Aún cuando estas condiciones no se cumpliesen, el estado actual de la tecnología permite, siguiendo los mismos pasos, el desarrollo de sistemas simuladores para el aprendizaje y uso de las frases normalizadas

Lo que dicho proyecto pretende, por lo tanto, es estudiar las posibilidades legales y técnicas, así como la conveniencia comercial del desarrollo de un traductor automático para las comunicaciones orales en el ámbito marítimo, y en su caso, establecer las bases de dicho desarrollo, a la vez que estudiar las implicaciones que sobre la seguridad marítima tendría el uso de un dispositivo de tales características. Igualmente se persigue el desarrollo del simulador para el aprendizaje y uso de las frases marítimas y la elaboración del material curricular correspondiente. Para sentar la base tanto del estudio, como del desarrollo es necesario efectuar el registro, en una base de datos, de frases normalizadas en su forma habitual de uso en el medio marino.

METODOLOGÍA

El mencionado proyecto basa gran parte de su contenido y desarrollo en la toma y análisis de muestras reales sobre comunicaciones en el ámbito marítimo. Estas muestras deberán ser obtenidas in situ, es decir a bordo de buques mercantes para el registro de comunicaciones reales buque-buque y buque-tierra, y en torres de control de tráfico marítimo u otros dispositivos para las comunicaciones reales buque-tierra.

Será imprescindible la presencia a bordo de varios miembros del equipo investigador, que se encargarán de tomar dichas muestras y realizar un primer tratamiento informático de las mismas.

Seguidamente estas muestras serán analizadas lingüísticamente en el laboratorio, y utilizadas para realizar las diversas pruebas sobre procesamiento del lenguaje y supresión de perturbaciones.

Las conclusiones de este análisis, y su comparación con la disponibilidad de tecnologías de procesamiento del habla, traducción y síntesis de voz, todas ellas en tiempo real permitirán establecer las necesidades de equipamiento para el desarrollo de un traductor automático para las comunicaciones marítimas: características del equipo, composición y coste.

La evaluación positiva de los aspectos técnicos y legales para el desarrollo del traductor, conduce al estudio de las perspectivas comerciales, a través de la búsqueda de empresas del sector de las telecomunicaciones interesadas en prototipar y llevar a cabo el desarrollo completo del sistema.

Por el contrario la evaluación negativa de los aspectos técnicos y/o legales, no permite el desarrollo a corto plazo de un traductor comercial, pero no impide el desarrollo del sistema simulador de ayuda al aprendizaje y uso de las frases marítimas normalizadas, desarrollado sobre un entorno PC y consistente en la readaptación del sistema de evaluación del traductor que se complementa con la elaboración de materiales curriculares.

ESTADO ACTUAL DE PROYECTO

En el momento de redacción de este artículo el proyecto se encontraba en la siguiente fase:

Han sido completados un estudio sobre las particularidades de las comunicaciones externas en el ámbito marítimo y la realización de un cuestionario Delphi o panel de expertos, para determinar las opiniones y sugerencias de los destinatarios finales de esta tecnología, sobre la viabilidad de su desarrollo y posterior utilización, así como sobre las implicaciones en la seguridad marítima que el uso de tal equipo conllevaría.

Se ha realizado un estudio de las Frases Normalizadas de la OMI para las comunicaciones marítimas. Este estudio refleja las posibilidades de uso de estas frases como "lenguaje controlado" válido para la facilitación de la traducción automática. Así mismo se realiza un análisis probabilística exhaustivo de estas frases con ayuda de la herramienta WordSmith Tools, que nos ha permitido, entre otras cosas, conocer la probabilidad de uso de los términos contenidos en dichas frases.

Muestras reales están siendo obtenidas en distintos puntos y analizadas simultáneamente. En este punto se está realizando el entrenamiento del software de reconocimiento de voz "Dragon Naturally Speaking".

El desarrollo de un sistema de evaluación(simulador) que permita el estudio de las necesidades tecnológicas para el diseño de un traductor automático para las comunicaciones marítimas se encuentra igualmente en fase de desarrollo, para lo cual se está utilizando la herramienta de ayuda al aprendizaje "Lingus".