MyVoice goes Spanish. Cross-lingual adaptation of a voice controlled PC tool for handicapped people *

Zoraida Callejas
Univ. Granada
Granada Spain
zoraida@ugr.es

Jan Nouza
Tech. Univ. Liberec
Liberec Czech
jan.nouza@tul.cz

Petr Cerva
Tech. Univ. Liberec
Liberec Czech
petr.cerva@tul.cz

Ramón López-Cózar
Univ. Granada
Granada Spain
rlopezc@ugr.es

Resumen: En este artículo presentamos la adaptación del sistema MyVoice del idioma checo al español. MyVoice se desarrolló con la idea de permitir a usuarios con discapacidad motora controlar sus ordenadores y aplicaciones informáticas de forma oral. Nuestro objetivo era adaptarlo de forma rápida al español empleando únicamente los recursos disponibles para el idioma checo. Los resultados experimentales muestran que se puede conseguir hasta un 96.73% de precisión en el reconocimiento del habla española empleando el motor de reconocimiento del habla checo del sistema MyVoice.
Palabras clave: cross-linguistic, reconocimiento del habla, aplicaciones para discapacitados

Abstract: In this paper, we present the cross-lingual adaptation of the MyVoice system from the Czech to the Spanish language. MyVoice was developed to allow motor-handicapped people to voice control their PCs and applications. Our objective was to cost-efficiently adapt it to the Spanish language using uniquely the resources available for Czech. Experimental results show that up to 96.73% recognition accuracy can be achieved for Spanish using MyVoice's Czech speech recognition environment.
Keywords: cross-linguistic, speech recognition, applications for handicapped

1 The MyVoice system

MyVoice is a software tool to control the PC and its programs orally. It recognizes voice commands and interprets them into one or more basic actions which include virtual managing of keyboard, moving mouse, clicking mouse buttons, printing strings and executing programs. MyVoice was developed with the purpose of facilitating Czech motor-handicapped people the access to new technologies, and has been succesfully used by them since 2005 (Nouza, Nouza, and Cerva, 2005).

MyVoice is structured in several command groups, each of them dealing with an specific task, this way for example the group that controls the mouse is different from the one that deals with keyboard but they can be accessed easily from each other by a voice command. The grouping of commands makes interaction easier as the user is aware of the valid words he can utter at each time and can easily navigate between groups. Furthermore, as a specific vocabulary was defined for each task, better recognition results are achieved.

The system was designed to be user friendly and customizable and it can be easily adapted to user preferences employing its configuration window. From there, phonetics of the words can be changed, commands can be added, edited and deleted, and new command groups can be easily introduced without the need of having any expert knowledge about computers.

MyVoice system was carefully designed and implemented and has been warmly welcomed by the Czech handicapped community. Our aim was to make it available also

for the Spanish users withoug building a new system from the scratch, but rather using the already developed resources for the Czech language. In order to reach this objective, we carried out a cross-lingual adaptation of the system so that recognition of Spanish commands could be done over the Czech speech recognition environment (i.e. acoustic, lexical an linguistic models), as explained in the next section.

2 MyVoice cross-lingual adaptation to Spanish

MyVoice commands were translated to Spanish and a cross-lingual adaptation procedure of the Czech recognizer was carried out. The Czech recognizer's decoding module works with a lexicon of alphabetically ordered words, each of them represented by its text and phonetic form. For the cross-language application we used Spanish text along with an automatically generated Czech phonetic representation. The phonemes built for the Czech recognizer could be then applied to the new task of recognizing Spanish words, using the Czech phonetic form to construct the acoustic models of the words by concatenating the corresponding phoneme models.

To automatically generate the Czech phonetic representation of the Spanish commands, a correspondence between Spanish and Czech phonemes was carried out by one Spanish native speaker and supervised by several Czech native speakers. The accuracy of such correspondences depends on the number of phonemes present in each language and the similarity between them. However, Czech and Spanish languages are very different in their origin, as Czech belongs to the family of Slavic languages like Russian, and Spanish is an Italic language like Italian or French. Thus, one of the challenges of our work was to obtain satisfactory mapping for such different languages; especially when previous researches had obtained poor results in cross-language tasks between Slavic and Italic languages, for example in (Žgank et al., 2004) with Slovenian and Spanish.

3 Experimental results

Our first experiments were carried out with a female Spanish native speaker employing the MyVoice software for carrying out her daily activities with the PC. For speech recognition

a gender dependent model was used obtaining a 93,92% accuracy rate. We carried out speaker adaptation to try to further improve this result. After adaptation to our female speaker, 96,73% accuracy was obtained. It is important to note that these results are for real interaction with MyVoice, in which vocabulary is restricted at each step to the list of commands in the current group, the size of the group ranges between 5 and 137 commands. To obtain meaningful results for the different speaker models with independence of the groups visited during the interaction, we carried out an offline speech recognition process in which we used the whole MyVoice vocabulary, which is composed of 432 commands. With a gender dependent user model we obtained 91.03% accuracy, which is improved by speaker adaptation reaching a 96.58% accuracy.

4 Conclusions

In this paper we have presented the adaptation of the MyVoice system for orally controlling PC, from Czech to Spanish language. We have empirically demonstrated that cross-lingual adaptation of the speech recognition environment can done in a short time carrying out an expert-driven correspondence between both languages' phonetic alphabets. Experimental results using the Spanish version of MyVoice showed that a 96.58% offline and 96.73% online performance can be obtained. Thus, these are very promising results as they show that portability of speech recognizers can be ensured in a straightforward way and that this approach can achieve good results even with very phonetically different languages as Czech (Slavic) and Spanish (Italic).

References

Žgank, A., Z. Kacic, F. Diehl, K. Vicsi, G. Szaszak, J. Juhar, and S. Lihan. 2004. The cost278 masper initiative - crosslingual speech recognition with large telephone databases. In *Proceedings of LREC 2004*, Lisbon, Portugal, May.

Nouza, J., T. Nouza, and P. Cerva. 2005. A multi-functional voice-control aid for disabled persons. In *Proceedings of International Conference on Speech and Computer (SPECOM 2005)*, pages 715–718, Patras, Greece, October.