

Multilingual speech-based technology to talk about!

January 19 2005

Paving the way for much more intuitive, interactive, and user-friendly ‘spoken dialogue technology’, DUMAS developed a multilingual speech-based system that creates new ways to communicate.

DUMAS, a three-year IST-funded project, began by developing the Athos platform, a generic and modular framework for multilingual speech-based systems. A consortium of eight partners from Sweden, Finland, Germany and the UK, its researchers built on basic [speech technology](#), such as speech synthesis and recognition, and focused on dialogue level problems to develop systems that can process both spoken and text inputs in several languages, and provide appropriate verbal responses to the user.

DUMAS’ researchers produced 27 outputs on various levels, from research prototypes suitable for further exploration to fully commercialised and marketable products. Several results are in commercial use; five have commercial potential with a short lead time or are in the process of being commercialised at present; 12 are knowledge resources for exploitation in other research projects or in commercial products of various kinds; and five are advanced technology components for further research exploitation.

Today’s electronic voice message systems primarily provide routine service transactions, such as booking a train ticket or a hotel room. “It’s a fill-in-the-blank approach,” says Dr Björn Gambäck, DUMAS coordinator. “If the information you need is not programmed into the list of options, you may well be out of luck, especially when it is impossible

to reach a live person.”

The project explored and improved upon three main areas. “First,” explains Gambäck, “current technology has very limited capabilities when it comes to processing and understanding structured text, in particular text written in several different languages. Second, existing systems are designed for quite a limited set of conversational contexts and falls over when the user tries to do something outside of the systems’ scope or uses a language which isn’t very grammatical. Finally, the systems can’t remember how specific users behaved and thus can’t adjust to their needs.”

Supporting multilingual email interaction

AthosMail, the main demonstration prototype of the project, is an email application that deals with multilingual issues in several forms and environments, and whose functions can be adapted to different users, different situations and tasks. It has been used to explore adaptivity in dialogue, and to compare and test various adaptive methods in spoken dialogue, and supports multilingual interaction in English, Finnish and Swedish.

The Language Identifier, one of the building blocks of the research, is a text classification program that identifies the language of a given text. The language identifier developed for the DUMAS project was implemented for English, Finnish and Swedish, but the language selection can be extended quite easily. The language identification makes it possible to automatically choose the linguistic analysis tool of the correct language.

“The system can identify which sentences of an email are written in a particular language and switch the speech synthesiser to that language,” says Gambäck. “So if a sentence is in Swedish it’s read in a Swedish

voice, and if the next sentence is in English it's read by an English voice.”

Reading for the visually impaired

AthosNews, another output of the project, is a prototype telephone system that reads English and Finnish newspapers for the visually impaired or for people in situations where they are using their sight for other tasks. The database currently contains bulletins from the Finnish Federation for the Visually Impaired and the largest newspaper in Finland, Helsingin Sanomat, as well as the two most popular afternoon papers, Iltasanomat and Iltalehti.

The Finnish version of AthosNews is currently being tested with about 50 visually-impaired test users. The AthosNews system for English has been completed and successfully evaluated with a group of 10 users, divided evenly between blind and sighted.

“Feedback on this system has been encouraging,” says Gambäck. “It is not fully usable yet, but the missing functions, which are keypad navigation and control of the document reading process, are not technically difficult to put in place. We have had some discussions with external partners about the potential for developing a commercial or charitable industrial version.”

Giving desktop applications a voice

To date, three outputs of DUMAS' research are available.

Timehouse SpeechServer is a practical, commercially-available tool with some licenses already sold to customers. It gives voice to desktop applications and provides a powerful, easy-to-use interface to

Microsoft's Speech API (SAPI) for developers. It offers developers and informed users the capability to easily produce compact MP3 files for the Web.

Searcher is a system for indexing and retrieving text documents. Documents and queries are matched using the well-known and efficient Vector Space Model, which interprets documents and queries as vectors in a high-dimensional term space. Searcher is publicly released under the MIT license. The latest version can be retrieved online.

AthosCal, a prototype multimodal calendar application, runs on several platforms including PDAs and desktop computers. It is currently being tested by a number of Swedish users and plans are under way to extend its functionality for mobile usage to mobile phone clients.

Coming soon to a telephone or computer near you?

Take a conversation between a human and a computer where the computer informs the user about a doctor's appointment. The computer suggests a bus time and the user responds by asking for a taxi. The computer orders and tells the user where to wait. Gambäck says that such a conversation is technically feasible today. "This more elaborate type of scenario could easily be in place within a year or two," he says. "But nobody has done it yet. The problem is a lack of funding to develop the system and offer it on a commercial basis. Money, not technology, is holding it up."

"In the future, communication with electronic systems will have to be dynamic and adaptive," says the project's scientific coordinator Kristiina Jokinen. "People want systems that can learn through interaction and adapt their behaviour to different users and different situations, as opposed to today's systems, which require the user to adapt to the system." DUMAS' multilingual speech-based system sounds like part of

that future.

Source: [IST Results](#)

Citation: Multilingual speech-based technology to talk about! (2005, January 19) retrieved 19 April 2024 from <https://phys.org/news/2005-01-multilingual-speech-based-technology.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.