Access to Lexical Knowledge in Interpersonal Communication Aids

by Constantine Stephanidis and Margherita Antona

The TIDE-ACCESS Project 1001 produces innovative methodologies and tools for access to computer-based applications by potentially all users, including people with disabilities. It focuses on the unified creation of User Interfaces adaptable to different user needs, characteristics and preferences, and at the same time aims at the design of innovative applications offering functionalities suitable for the potential target user groups.

One of TIDE-ACCESS project's application domains is Interpersonal Communication Aids, ie devices enabling people with speaking difficulties to communicate. Communication aids provide users with a vocabulary display (of either a natural language or a symbolic language, eg Bliss) and functions for composing and sending messages. In the framework of ACCESS, an innovative modular architecture is elaborated facilitating the design of highly tailorable and adaptable communication aids.

An important aspect of adaptability in communication aids concerns the selection of an appropriate language and vocabulary set for each user without precluding communication with users of different languages.

The User Vocabulary and Meaning Mapping Module, designed and implemented at ICS-FORTH in the context of this project, is a knowledge base containing lexical information related to symbolic and orthographic languages in Interpersonal Communication Aids. Its main purpose is to make multilingual knowledge available both during the configuration of devices and their use for communication. During configuration, usually performed by a therapist, the Module makes it possible to consult the vocabularies of already defined symbolic or natural languages and define new user languages along with their vocabularies. During communication, it ensures translation between whatever user language at a lexical level.

The Module is language independent, and is designed in such a way to allow encoding of any type of linguistic knowledge. Information is represented by means of a typed feature structure representation language (the ALE, Attribute Logic Engine, developed at CMU and available for research purposes), while retrieval, language definition and vocabulary definition functions are implemented in Prolog.

Currently encoded information includes:

- defined languages and their characteristics (symbolic or orthographic nature, subset relationship to other defined languages, etc)
- part of speech categories
- a simple domain model for common communication topics like food, family and people in general, house equipment, weather, etc
- vocabularies for two orthographic and two symbolic languages (English, Finnish, Bliss and Pictograms).

Lexical entries in each vocabulary contain features representing parts of speech and domain model concepts. Retrieval of lexical items takes place by unification of features structure descriptions. Lexical translational equivalence between items of all languages is defined as feature value identity on domain model concepts.

The Module can be further expanded by:

- included languages
- syntactic information
- domain model information and size
- vocabulary information and size.

Future work in the direction of representing and exploiting linguistic information in communication aids will build on the existing knowledge base to include parsing, message-to-message translation and other techniques such as word prediction and other rate-enhancement techniques.

Please contact:
Constantine Stephanidis - ICS-FORTH
Tel: +30 81 391741
E-mail: cs@ics.forth.gr