

# Draw a Circle! – Spoken man-machine communication for design

**Sheri Hunnicutt, Alice Carlberger, Elisabet Rosengren, Nathalie Talbot**

Dept. of Speech, Music and Hearing, KTH  
Dr. Kristinas väg 31  
S-100 44 Stockholm, Sweden  
fax: +46 8 790 7854  
tel: +46 8 790 9268

**Corine Bickley**

Concentra Ltd.  
Viscount Centre II, Milburn Hill Rd.,  
University of Warwick, Science Park  
Coventry CV4 7HS, U.K.  
fax: +44 1203 419875  
tel: +44 1203 692323

ICAD, a product from Concentra Ltd., is a sophisticated generative design program for engineers, architects and other designers. It is used to produce product designs, product configurations and sales proposals and allows a user to substantially enhance productivity. This productivity gain results from the automatic application of a set of rules within the generative model, rules that can be quickly re-evaluated each time a user changes inputs to the model. Employing these rules to capture and retain knowledge in a product model, an ICAD user can make variations in a pre-existing procedure in a fraction of the time that it would take to produce a second model. Since minor variations of pre-existing practices and procedures have been estimated to encompass up to 80% of all engineering activity, this tool can give a user significant advantage. ICAD accepts input from the user keyboard and mouse action.

In a European Union project called ENABL (Bickley et al, 1998), the Dept. of Speech, Music and Hearing at KTH is a partner in the development of a different input modality for the ICAD program. This modality is speech. At KTH work is being done to provide speech recognition for a new interface for ICAD (Carlson and Hunnicutt, 1996; Ström, 1997; Carlberger, 1998; Talbot, 1999). The ENABL project is for the benefit of disabled users, designers who are not able to use a keyboard and mouse combination. The commercial benefit, however, is expected to be gained by combining the two modalities so that able-bodied designers have the option of using manual or speech input. It is expected that some functions will be more easily accessed by speech, and that others will lend themselves more to a manual input approach. A combination of the two would give a designer freer expression.

A major question being addressed in the ENABL project is what linguistic changes are desirable to provide for spoken input. For example, in spoken input, complete phrases or sentences may be more natural; in manual input, clicking in dialog boxes may be the preferred method of providing information and giving commands. The options need to be considered and evaluated.

Another issue to consider when implementing an input system to an engineering design program is the positioning of objects (or their parts) on the screen. In the regular ICAD program, this can be done by specifying coordinates through a keyboard or by pointing using the mouse. In the ENABL application, positioning is accomplished by voice. Voice also allows the position of an object to be specified relative to another object (or its parts) on the screen. An example is "Place Object1 20 millimeters below Object2." Another question which arises is how to refer to objects (or their parts). In the regular program, objects can simply be named. In the ENABL application, it has been decided to use a system of combining a name from the standardized naming alphabet (Adam, Bertil, Ceasar, etc. for Swedish) with a number (0-99). This simple compounding system eliminates the need to store the name of each object in the lexicon explicitly, and promotes generality across applications. Proper names were chosen over letters for ease of recognition, as the latter are notoriously difficult to recognize due to a high degree of inter-rhyming, and to word brevity with a consequent tendency for reduction and elimination in context.

Since persons with manual disabilities also often have a vocal disability, the digressive characteristics of the dysarthria present a challenge to this input modality. A part of the ENABL project is dedicated to analysing speech from persons with dysarthric voices and to evaluating

their performance using a speech recognition engine. The study focuses on 12 speakers with various abnormal voice characteristics. Recordings are being used to test the performance of the recognizer developed for the ENABL project. The recognition result will be calculated as the percentage of correctly recognized words and as the word error rate in which substitutions, deletions and insertions are taken into account. Factors which can affect the recognition performance are, for example, severity of dysarthria, intelligibility, voice intensity, frequency of pauses, length of pauses, and the occurrence of pauses within a word.

It is hoped that project conclusions regarding the linguistics and acoustic-phonetics of spoken design will contribute valuably to this emerging modality.

## REFERENCES

- Bickley, C., Carlson, R., Cudd, P., Hunnicutt, S. & Reimers, B. (1998) "TIDE-ENABL – The First Year," *Improving the Quality of Life for the European Citizen* (I.P. Porrero & E. Ballabio, eds.), pp. 258-261.
- Carlberger A. (1998) "Lexicons and Grammar for Speech Recognition in an Engineering Design Program (ICAD)," in *Proceedings of FONETIK 98* (eds. P. Branderud and H. Traunmüller), Dept. of Linguistics, Stockholm University, pp. 172-175.
- Carlson R. & Hunnicutt S. (1996) "Generic and domain-specific aspects of the Waxholm NLP and Dialog modules," in *Proceedings of ICSLP 96*, Philadelphia, USA, pp. 677-680.
- Ström, N. (1997) "Automatic continuous speech recognition with rapid speaker adaptation for human/machine interaction," doctoral dissertation, KTH, Stockholm.
- Talbot, N. (to be published) "Speech Recognition in the ENABL Project," *Talking to Computers II*, University of Sheffield, U.K., July 1999.