



## THE CONFLICT BETWEEN SPEECH AND TEXT IN THE AUTOMATED OFFICE

M F Lacey\*

### ABSTRACT

The application of speech based technologies in the office, apart from the humble telephone, has received only limited penetration. Where text based facilities are already available active resistance to the introduction of speech based systems has been observed. It is believed that in their attempts to introduce the technology suppliers have taken too little care to understand how naive users will approach and use the new facilities.

### INTRODUCTION

Change affects us in all areas of life. In the office the introduction of technology has progressed only slowly and then only in particular areas. Fear of change is a barrier to all technological (and other) introductions which require people to operate in a new way. However, this has often been overcome when the new way presents obvious benefits -even when the human-computer interface (hci) has been poorly designed.

But in the area of speech applications in the office, few such clear benefits have been apparent and resistance rather than adoption has been common. ITL's Information Management Processor (IMP) was available for several years and included the ability to annotate word processed documents with spoken comments. Despite making the speech controls similar to a tape recorder, the facility remained little used. In an American installation of a similar text and speech system, less than ten percent of users accounted for half the voice traffic; and less than thirty percent for four-fifths. Voice annotation, in part or whole, was used in less than thirty percent of all messages (ref 8).

Even the telephone is not immune. A recent UK survey (ref 1) found that a third of male and a quarter of female callers who encounter a telephone answering machine hang up immediately. Some research at Loughborough University indicates that the greeting message may engender a lower 'slamdown' rate - with formal greetings most successful in a limited trial (ref 7).

\* Principal Consultant, Eosys Limited, Clove House,  
The Broadway, Farnham Common, Bucks SL2 3PQ

## NEEDS

Too often facilities such as voice annotation, recognition, and store and forward have been introduced -especially in commercial environments on a "try it and see basis" with the result that few people used them successfully. As Culnan observed (ref 3), physical access is not, by itself, sufficient for promoting usage -especially in voice related applications where the information reception threshold is already inherently low. For professional people, more information is retained from reading a text than from listening to speech (ref 2).

With technologies such as telephone answering machines and voice mail systems there are no established norms for greetings or for structuring messages and this in itself gives rise to uncertainty on the part of users. So, not surprisingly, effective use of such systems has been found to relate to length of experience, rather than any assumed prior level of telephone usage (ref 4). That same research emphasised that it is essential to understand the nature of the user community's communications patterns and to recognise that prior experiences with similar technologies will significantly affect the community members' use. It is therefore doubly disappointing that suppliers of speech technology products continue to seek to iron out their faults in use in a real community, thus colouring individuals' future expectations against success and ease of use.

Licata (ref 6) isolated three key points which positively influence users in favour of a new technology. These were:

- their pre-implementation beliefs about the technology
- their current attitude structures over what is going to happen and what they will get out of it
- the reason for using the technology.

Thus it is important to engender a positive attitude to the technology amongst the user community prior to its introduction.

## PROBLEMS

There are three further problems (at least) associated with the use of voice in the office - apart from the major one of trying to introduce alternative media into situations which, after much effort and heartache, are now successfully using text based office systems.

The first is a human perception problem, the other two more functions of the current stage of technological developments.

It is important in any application that the user is able to visualise in some way what he believes is happening at any one point in a dialogue (textual or verbal). Thus the imposition of voice - or even key tone (DTMF) - control on what was initially designed as a text based application seldom works effectively unless the transactional information flows are minimal. Thus, for example, in an audiotex system, there is a need for a clear mapping of the database structure onto a comprehensible user understanding (ref 9).

A technological problem arises in voice recognition applications - that of mismatches. When these occur they dramatically undermine the users visualisation of the affect of his commands and lead to confusion and rejection. Further, such systems need testing in a live user environment for extraneous noise, false commands and, not least, to realise that "er" and "um" are part of a users vocabulary in either discrete or continuous word environments (ref 10).

Finally, in telephone based applications, there is currently a problem caused by the low penetration of DTMF phones in the UK and the limited success of speaker independent recognition which will only be solved with the passing of time (and the efforts of BT and Mercury to introduce DTMF and switchable phones).

#### RULES

In conclusion we state (after Jones et al, ref 5) some rules to be borne in mind when designing speech based applications for the automated office.

1. Engender a positive attitude amongst the user community prior to its introduction.
2. Only consider speech for relative infrequent use.
3. Beware of unnatural impositions when using speech technologies (eg discrete word recognisers).
4. Design a special command vocabulary for speech independently of any alternative text command vocabulary.
5. Ensure that the user receives feedback in an appropriate form at the right time when using the application.
6. Take special care over error detection and correction.
7. Don't forget that the user always has a choice whether or not to use the system.

## REFERENCES

1. Audience Selection (for British Telecom), Getting the Message (1986)
2. C. Brett, Voice Processing (On-line, London, 1985) pp 11-21
3. M J Culnan, 1984 (Quoted in Grantham and Vaske)
4. E G Grantham & J J Vaske, Behaviour and Information Technology 4, pp 327-335.
5. D Jones et al, Voice Processing (1985) pp 23-30.
6. B J Licata, 1982 (Quoted in Grantham and Vaske).
7. H S Maskery & B G Pearce, HUSAT Research Group (undated).
8. R T Nicholson, ACM Transactions on Office Information Systems, 3, pp 307-314.
9. M B Rosson & N M Mellem, American Voice 3/0 Systems Applications Conference, 1985.
10. K Swan, AVIOS Applications Conference, 1985.

Copyright (c) 1987 Eosys Limited. Used by permission.  
All rights reserved.