



AUDIOTEX: Computer Input and Output by Voice Telephony

R I Anderson

Audiotex means the output of computer data vocally, frequently remotely by telephone. The use of the name audiotex is to allow people to distinguish this offering from the other speech technologies like voice messaging even though the hardware (and even software) may be exactly the same.

Many of the customers for Ferranti voice messaging systems had a requirement for recorded announcements of changing information eg the name and phone number of the on call engineer at a plant, the plat du jour in a hotel etc. The 'billboard' mailbox is a standard feature on Ferranti voice messaging systems and as you can see from the examples has many applications.

A measure of the acceptance of recorded information against the resistance to recording a message is shown by the example of the on call engineer. Our system is quite capable of recording a message in the on-call engineers mailbox and then ringing the appropriate phone number or pager number as allocated by the engineer. The management of the plant however, decided that there was a risk of some people not leaving a message.

Recorded messages for retrieval over the phone have a long history, the time signal being the first in the UK. Though the time signal was recorded on drums most of the others were tape systems. Endless tapes meant the message was constantly replayed and you joined the message partway through, pairs of rewinding tapes gave a message which started at the beginning but still required separate phone lines for each message and two tapes per line. The multi-threading capabilities of digital computers means that the same message can be accessed by many people at the same time with each person hearing the message starting at the beginning - but only one message need exist.

The applications of the DTMF control utilised by voice messaging allows information to be scanned, skipped and replayed in an arbitrary way. Until you start to use the DTMF in a fully interactive voice response system you see only small improvements in efficiency or productivity. The fully computerised solution allows access to other computer held data and gives very powerful information handling.

Applications using some of this power have started to appear in the UK this year. One example is the horoscope service. This is a fairly trivial example since the DTMF input of birthdate merely determines the astrological sign. I would have twelve different telephone numbers for the different signs and dispense with DTMF. The home penetration of DTMF in the UK is low and I am sure Comms Managers are busy putting blocks on numbers beginning 007.

A second example is the stock exchange listing. This is currently being offered by two different companies and uses the full mainframe access and DTMF request of information capability of the systems. Use is made of the computer power to provide extra facilities such as portfolio listing where a client has his share holding prices latest prices read to him. Ultimately those systems will no doubt be extended to accept orders to sell and buy shares over the phone.

These information systems barely scrape the surface of what is possible. True audiotex services go much beyond the mere 'reading' of computer data or replaying of recordings. Home banking is a much better example of true interaction and structural dialogue between man and computer.

Though the security employed and therefore the man-machine interface varies widely among the different systems the principles are common. A user identifies himself by a variety of methods ranging from use of his ATM card in a reader to account number and a security number pressed on a DTMF phone or tone sender. The range of services available varies considerably and is linked with the security. A system with low security (DTMF tones only) normally restricts the ability to transfer money to outside accounts. A system requiring high security (ATM card plus encrypted PIN number) may allow transfer of money to any bank account.

These systems which started operation earlier this year are starting to give new improved services to the general public. The moves to automated banking started by 'hole in the wall' cash dispensers has moved on to sophisticated ATMs, home telephone banking and EFTPOS. The fact that people queue in the rain to use an ATM during normal banking hours shows that people not only accept but prefer automated services. Voice in the home is even more convenient and user friendly though without the facility for cash dispensing.

When users of voice output computers become accustomed to the structure and patterns of input required they rapidly become irritated with the lengthy verbal prompts required for naive users.

Two main strategies to overcome this have been adopted by suppliers. The first of these was multi-tier prompting where a user is introduced with one level of initial prompts and later has reduced or curtailed prompts played with access to the full prompt by pressing a help key. Less proactive but more satisfying is the second method of providing pre-emptible prompts. Input can be continued to completion of a transaction or to a situation where unfamiliarity requires listening to the prompts. With foreign language versions of Ferranti's Voice Manager product I have no need to listen to the prompts. The European prompts are relatively simple to follow but Mandarin Chinese I find is more likely to confuse me than help.

Since much of business is international or in countries where more than one language is in common use eg Belgium or Switzerland, a multi-lingual capability is required. Due to the sensitivity of people being greeted in the wrong language first, I feel the best solution is language selected by phone number. We have supplied many different implementations using a variety of techniques to different customers around the world.

User perception of the voice can create dramatic changes in the pattern of usage of a voice system even where no multi-lingualism is required. Our first speaker was an employee with a local Manchester accent. When we changed to the professional accentless BBC English announcer usage dropped. On surveying our user base the reason was 'the new girl isn't as friendly'. However when we shipped our first production system to the South-East of England one of the survey responses was 'we don't like the Northern accent'. What they would have thought of the original Manchester accent I hesitate to conjecture.

The current arguments between the competing technologies of voice, text and image as to which is the best in terms of communications are specious. Voice mail, electronic mail, and fax all have their part to play - my own company like many others uses all three. I use all three but to achieve different objectives. I also use post couriers and two way telephony. All these methods of communication have their part to play in business. The trade-offs between speed accuracy cost and flexibility will always dictate the use of different media.

With the coming of ISDN and digital transmission end to end the economics of the various media will cause changes in the emphasis of use. One of the obvious developments will be the synergy achieved from using multi-media transmission. Already speech annotation of text messages is deliverable. It will obviously shortly be speech annotation of any PC file so that it may be used not only for dictation and proof reading of typists work but for verbal commentary on spreadsheets business plans etc.

The use of ISDN will allow voice annotation of all communication such as fax as well as the current forms. Videotelephony and videoconferencing using multi-media have already been announced this year and many more combinations and synergy between communications styles and media will occur. The speed of convergence is increasing all the time.