



HOME BIOGRAPHY TOPICS RESOURCES NEWS CONTACT

AUTOMATED SPEECH RECOGNITION

Speech recognition technology has finally come of age. For banks, this means it is now possible to automate routine call centre transactions such as balance enquiries and funds transfers, thereby providing a more convenient, streamlined service, and dramatically reducing costs.

To make sense of the technology it is helpful to make two distinctions:

- Speaker dependent vs. speaker independent recognition.
- Basic Interactive Voice Response (IVR) vs. more sophisticated Automated Speech Recognition (ASR) and Natural Language (NL) systems.

SPEAKER DEPENDENCE VS. SPEAKER INDEPENDENCE

Speaker dependent systems only recognise a particular individual's voice, after fairly extensive training. The most familiar application of this technology are dictation products from suppliers such as Dragon and IBM. Dictation products are gradually becoming more acceptable but are unlikely to become widely used until it is possible to use continuous speech without prior training. Another application is [voice verification](#), where the unique physical characteristics of a particular voice are used to authenticate an individual. This is highly relevant to banks and is the subject of another "financial futures" web page.

Speaker independent speech recognition technology allows any person, with any accent or dialect, to communicate with a computer using continuous speech, a large vocabulary, and increasingly natural language patterns. This technology can be used with high quality microphones in PCs or kiosks, but its most exciting application for banks and other financial institutions is to enable routine transactions over the telephone.

IVR VS. ASR/NL

IVR has been around for some time and can be regarded as the "plumbing" of a modern speech recognition system. Vendors such as Periphonics, Syntellect and Intervoice supply industrial strength systems for call centres which handle advanced telephony and CTI (computer/telephony integration) features such as call direction, load balancing, and screen popping, as well as touchtone (or DTMF) selection from a menu of options, rudimentary recognition of a few words such as single digits and yes/no, and automated speech generation. IVR technology is used by many banks (eg NatWest's Actionline) and works reasonably well, but leaves a lot to be desired and will never be accepted by a large majority of the population.

ASR/NL technology from vendors such as Nuance, ALTech and Larnout & Hauspie sits on top of the IVR systems and is far more sophisticated. Because it operates at the level of phonemes rather than words, it is possible to recognise huge vocabularies, and sophisticated algorithms now allow customers to ask questions, make commands, and engage in a dialogue using increasingly natural, continuous speech, albeit in a limited subject domain.

THE SIGNIFICANCE OF SPEAKER INDEPENDENT AUTOMATED SPEECH RECOGNITION

Most banks, already saddled with costly branch networks, have built large call centres to handle telephone enquiries and transactions. These call centres are proving to be a victim of their own success - call volumes are growing rapidly and this means growing costs (on top of the branch costs) since good telephone operators are expensive. But up to 80% of call centre transactions are simple, routine transactions such as balance enquiries, funds transfers or pre-authorised payments which can easily be automated using the new generation of speech technology. Moreover, even complex transactions start with up to a minute's worth of routine identification and authentication of the customer, which can also be automated. The bottom line is massive potential cost savings for banks.

Of course this raises the question will customers accept talking to a computer rather than a human being? The answer seems to be definitely yes provided the dialogue is well designed. For example Charles Schwab's Voicebroker stock quotation system is accepted by over 90% of customers. In fact many customers actually prefer automated facilities - the service is available 24 hours a day, there is no waiting for a call to be answered, the transaction can be accomplished quickly, in a streamlined manner, with no potential embarrassment.

The important point about this technology is that it can be used by anyone (or at least anyone with access to a telephone), anywhere, without any prior training or special equipment. This is in sharp contrast to other new delivery methods such as Internet banking, and is the main reason why automated speech recognition over the telephone looks set to become a "killer app" over the next few years. At least three UK financial institutions are currently building speech applications and arguably most should follow.

A word of warning however. As with any new technology, it is the human factors surrounding the systems which at the end of the day determine success or failure. Good dialogue design is critical, and many other issues need to be carefully considered, such as customer authentication, interaction with other delivery channels, and links to host systems.

Interested? Please contact Nick Collin on nick@ncollin.demon.co.uk or **+44 (0)207 833 8765** with comments or questions.

Designed by - Blue Nostromo © 200505