Are you wasting your time & money?

By Tom Houwing, director, voiceandvision

Lots of technology, tools, software, middleware, licences, speech platforms, speech engines and the technical staff to get it all into place and under control. Pre-packaged solutions, templates and out-of-the-box-applications and the decision whether to build the application and its infra-structure on sight at the business customer or to have it hosted by platform-providers. Your organisation around the technical staff, which needs to get it all into place and under control, combined with trainings in order to enable your staff and project-leaders to handle and overview specific and new techniques.

All these investments can play a significant part in developing a speech recognition-based system and almost everyone at SpeechTek east 2006, can tell you as it is their line of work.

In New York they met once again: Companies from all over the world that deal with voice-automation-technology, in one way or another.
Because researchers and business analysts predicted voice as a very interesting future business, many ideas have been worked out and transformed into concrete results.

Since 1999 we've seen a clear increase of the number of companies that come to this conference. SpeechTek east, which takes place in New York during late summer, has its 12th anniversary. Since three years ago SpeechTek west takes place in San Francisco, during early spring. However, this west-variant takes place without AVIOS (Applied Voice Input/Output Society) and therefore it's somewhat smaller.

During the last couple of years the voice-industry has been preparing in order to get ready to 'attack' the expected explosion of demand that still has not taken place.

In the US a lot of companies are going from touchtone to speech, with the same IVR that was used with the touchtone system. This implies restricted flexibility as far as extended voice-recognition is concerned: dual -slot-filling or support of varying caller input require more complex grammar development.
The standard example in this context is 'say one' next to 'press one', which unfortunately doesn't show any respect for speech-specific dynamics in comparison to touchtone-based applications. The result is a voice recognition-based system without any added value.
Furthermore, both in the US and in Europe still the attention is almost entirely focussed upon the technique and infra-structure and at the same time more complex processes covered by call-centre agents, marketing-driven content, or even content represented on a website, shall now be covered by voice.

The awareness of the average SpeechTek attendee is definitely there: VUI is important.

The vendor will agree: That's definitely true, but first let me show you our product.

The overall reaction of the business customer is clear: vendors promise you heaven on earth. Then, once the project started, we were confronted with one problem after the other. It showed that their product wasn't that amazing after all. And now, way behind schedule, we still do not have our speech-application and we've spent considerably more money then we planned.

What seems to be very important is a dedicated professional team that listens to the business customer and supports the project.
End-users and business customer need to experience the added value of a speech-recognition-based system. The business customer wants to invest in a system that saves money and time. The end -user wants to experience an efficient and effective system that is easy and maybe fun to use, but also saves money and time just the same. Caller acceptance, that is the perceived public opinion, will drive and stimulate the use of speech-recognition-based systems.

As far as speech-recognition-based systems are concerned the speech recognition and VUI are responsible for making that happen.

Almost every SpeechTek attendee seems to agree, but there are only a few who really understand.
VUI stands for Voice User Interface. That is, the system uses the caller's voice as means of input. The system initiates callers to speak or give answers. The caller's input needs to be recognized in order to be able to go on with the dialog.

Within the man-machine interaction, this pseudo-dialog, aspects of human communication need to be respected and professionally taken care of within the framework and possibilities of today's speech recognition. This requires a specialized skill -set.

The recognition-machine must work properly to work with voice applications I the first place and for the first time this year, the speakers at SpeechTek openly communicate that in this context there is still a lot that needs to be done.

Then it's the integration and positioning of the VUI consultant/designer within existing organisation-structures and projects, together with solid grammar design and specialized recognition tuning that in the end assures any caller-acceptance.