Lernout & Hauspie:
Speech tech with a Flemish accent

Lernout & Hauspie has the biggest speech processing lab in Europe. "We're the only ones developing both speech digitizing, multilingual text-to-speech, and multilingual speech recognition."

Jo Lernout and Pol Hauspie are no greenhorns to the Belgian computer industry. The former was a sales director at Wang Belgium, the latter the founder of a successful accounting software house.

Perceiving in recent years the intense yet resolutely monoglot activities in speech processing by Japanese and American companies, they cast their lot together with overtly multilingual ambitions in mind.

Now, two short years hence, Lernout & Hauspie Speech Products is one of Flanders' technological prides, having achieved remarkable success in this field of speech processing. That is, speech recognition, speech synthesis (text-to-speech), and speech digitizing.

Homebase is Ieper, a part of Flanders best known for the stench of burning heretics in the 15th century, for the fields of red poppies which blossomed where the blood of many hundreds of thousands of young Europeans was spilled in the First World War, and for the hops which now delectify many of Belgium's noble ales.

Lernout & Hauspie Speech Products is quartered in the "t-zone" of Ieper, one of a number of pan-European EC-organized petrie dishes for the cultivation of Eurotech businesses. Growth-spurring medium comes in the form of ten years of tax-free dividends. Neighbors there include companies with ominous sounding names like RoboSoft.

Last May, Lernout & Hauspie took a text-to-speech system to Speech Tech 89 in New York City. Bart Verhaeghe relates, "It impressed a great many linguists and industry colleagues, despite the fact that it could only generate Dutch. They appreciated the system's skill in phrasing and prosody."

Like many US high-tech startups, a key element to the success of Lernout & Hauspie Speech Products has been its close ties to universities. Verhaeghe elucidates: "In addition to the text-to-speech system shown in New York which had been developed at the University of Ghent, a promising speech recognition project was underway at the University of Leuven. In collaborating to make these systems commercial successes, it's been a fortuitous combination of academic lab research and computer industry marketing savvy.

"Currently," says Marketing Director Verhaeghe, "we derive most of our income from our digitized speech storage system Digilog." Digilog is especially popular with large organizations which regularly need to verify the spoken word. A police department might need to trace a given incoming telephone call, or a stock exchange »

Speechtechies Pol Hauspie (left) and Jo Lernout (right): "Small has its advantages. We can make decisions in an instant. Decision-making within large companies can be extremely ponderous. Endless meetings. We call it TTM: Talking Too Much."

« might want to verify a spoken bid.

Previously, searching for a given conversation meant days of fast-forwarding through reels of tape. Digilog, however, can inventorize calls or conversations based on various criteria, such as incoming telephone line, time and date of call, length, telephone operator, etc., allowing extremely fast retrieval.

According to Verhaeghe, Lernout & Hauspie has the biggest speech processing lab in Europe. "We're the only ones developing both speech digitizing, multilingual text-to-speech, and multilingual speech recognition."

Lernout & Hauspie's audiotext systems, for home banking, tele-ordering, reservation systems, and other applications, are also selling like cold *geuze* beer on a warm day.

TALK TO ME

Like all good information-age prophets, they practice what they preach. Installed in their Brussels sales office is a voice mail system with true speech rec capabilities. Depending on access privileges, callers can instruct the system in natural language to record, replay, forward, delete, and further manipulate messages. The system has a vocabulary of about 40 words.

Lernout & Hauspie's speech synthesis system (DEPES) is based on a large database of diphones and triphones ("derived from a few hours of a native speaker reading the newspaper," adds Bart Verhaeghe).

First language was of course Dutch, or more accurately, Dutch with a Belgian accent. German followed, as did French, Spanish, and Italian. "DEPES is the only contiguous, prosodic speech synthesis system around," says Pol Hauspie proudly.

How directly related is their success to new hardware technology? "We are rather bit-hungry," Hauspie concedes with a smile. "But the timing is good. Optical storage is getting cheaper and cheaper per megabyte. Likewise, digital signal processors (DSP) are getting cheaper and more powerful. They're crucial in our text-to-speech and speech rec systems."

The company hopes to begin implementing optical disk "jukeboxes" within their systems shortly. This will mean being able to store up to 7,000 hours of speech (sampled at 16 kilobits per second) in their Digilog systems.

PHONETIC ENGINE

Current development activities are focused on what will be the most formidable armament in their langtech armory: the Phonetic Engine. Linguist/software engineer (and Leuven graduate) Roddi Coudron expects it to be ready within three years.

Traditionally, speech rec products are word-based. And for speaker-independent systems, it takes a fair bit of work to add a new word to the system. Lernout & Hauspie has a more ambitious approach, however.

Coudron: "For our phonetic engine, we will be entering all the allophones of a given language. We'll be recording 600 or so human speakers, sampling their voices at 40 KHz to obtain the generic parameters of each allophone."

Then, of course, they can discard the recordings. An inference engine will determine which sound corresponds to which allophone, and which word corresponds to which collection of allophones. "So, to enter new words, you'll simply have to key them in using phonetic nomenclature. Digital models will be automatically generated," says Coudron.

While Lernout & Hauspie expects simple continous speech rec within a year ("two–plus–two," for example), complex continuous speech rec ("ice–scream/I–scream") is at least three years off. Will the laurels go to the engineers or the linguists? Coudron: "Definitely the linguists. Semantic ambiguities can only be resolved by contextual analysis. Continuous speech rec will only be possible in combination with a hefty dose of natural language processing."

But how can a company of 36 people compete in a field where the likes of IBM, AT&T's Bell Labs, and NEC have speech processing research budgets the size of those of small governments?

"Rather than being in competition with them, we're actually rather attractive partners," explains Hauspie. "Because they tend to focus on American English or Japanese, they are very keen to collaborate with European partners for European languages. We are currently pursuing a number of possibilities in that regard," he confides conspiritorially.

"Another factor in our favor is that we are involved in all three fields: speech recognition, speech synthesis and speech digitizing." According to Hauspie, other companies tend to focus their research activities on just one of them.

"Besides, small has its advantages. We can respond quickly to opportunities. We can make decisions in an instant. Decision-making within large companies can be extremely ponderous. Endless meetings. We call it TTM: Talking Too Much."

Hauspie sums it up aphoristically by saying: "Your strengths are your weaknesses and your weaknesses are your strengths." In Lernout & Hauspie's case, he believes they have transcended their origins in linguistically obstreperous Belgium by capitalizing on their multilingual strengths.

Roddi Coudron echoes these sentiments: "Because Dutch-speaking Belgium has such a small linguistic base, we've been forced since we were young to be multilingual. As a result, there are a lot of translators and linguists in Belgium. It's our strength."

In consolidating these strengths, Lernout & Hauspie are the right paradigm for the European language industries: manage the multilingual challenge of polyglot Europe and the world will be at your doorstep.

— Colin Brace