What's next

Dragon Professional

The History of Speech Recognition – Part 1

Dragon speech recognition has been around since the early 1980s. It has always been at the leading edge, pushing the available technology hard in pursuit of its goal of turning spoken words into editable text, and controlling computers with voice. Dragon is part of the history of speech recognition, and also part of its future. In Part 1 of this two-part blog, we look at the early history of speech recognition and how Dragon changed the paradigm. In Part 2, we will look at how Dragon has developed, embraced new technologies, improved accuracy, and enhanced productivity across a range of vertical sectors. Stay tuned!

Alixson Bell

Posted 6 June 2019

It was 1982 when Dragon Systems was founded by Dr Jim Baker and Dr Janet Baker. They produced voice recognition software that could turn spoken words into text that appeared on screen. It was a major achievement considering the limitations of computers at the time. For example, here are some things that indicate the state of computing in 1982:

- Microsoft Windows was still a few years away – it was launched in 1985.
- Laptop computers were rare, large and expensive – the Grid Compass 1100 cost thousands of dollars.
- The internet existed, but the World Wide Web was still years away – it was 1990 when Tim Berners-Lee devised HTML, which allowed the creation of web sites.

Early days

Even before computers there were examples of voice capture. For example in 1881 Alexander Graham Bell (inventor of the telephone) had a hand in developing a system for cutting grooves into a wax cylinder in response to the voice. Later, in the early 20th Century, came the Dictaphone recording onto wax, then plastic, then magnetic tape as technologies advanced.

But these were all just about recording speech to play it back. The big breakthrough – and what we would today understand as speech *recognition* came with computing systems. There were lots of parallel development strands going on in the 1950s and 1960s. For example:

- In 1952 Bell Labs came up with Audrey (Automatic Digit Recognition). This could only recognise the numbers 0 to 9. The speaker had to pause between each word, and Audrey had to be trained to each speaker's voice. But it worked. Audrey would 'recognise' a number and flash a corresponding light.
- In 1962 IBM revealed Shoebox at the World Fair. Shoebox could understand 16 English words. It would listen to the words and complete an instruction for example adding up numbers and providing the result.

The different developments going on at that time were based on matching spoken words up with voice patterns. They worked word-by-word, and were not able to produce sentences.

The next big breakthrough came in 1971 with Harpy. This was funded by Darpa (the US Department of Defense research agency), and was a joint effort that included Carnegie Mellon University, Stanford Research Institute and IBM. Harpy cold work with ordinary speech and pick out individual words, but it only had a vocabulary of around 1000 words.

Enter the Dragon

The biggest advance yet came in 1982 when Dr Jim Baker and Dr Janet Baker launched Dragon Systems and prototyped a voice recognition system that was based around mathematical models. The Bakers were mathematicians and the system they came up with was based a hidden Markov model – using statistics to predict words, phrases and sentences.

This allowed for much more than just identifying words. It also allowed for working with syntax and context. That's really important for efficient general-purpose speech recognition, which needs to be able to produce meaningful sentences. For example, to produce a grammatically accurate sentence it is important to know which word is meant out of several that might share the same pronunciation but have different meaning and/or spelling.

In 1990, Dragon Dictate was launched as the first general purpose large vocabulary speech to text dictation system. This was a groundbreaking product for Dragon, but it required users to pause between individual words. By 1997, that problem had been overcome. Dragon Naturally Speaking v1 launched that year. It allowed for *continuous* speech recognition – users could speak in their natural way without leaving pauses between spoken words.

In Part 2, we will look at how Dragon has developed, embraced new technologies, improved accuracy, and enhanced productivity across a range of vertical sectors.

Tags: accessibility, Dragon speech recognition, productivity, Speech Recognition

More Information

	Learn more Learn more about Dragon Learn more

About Alixson Bell

Alixson graduated from the University of Washington in 2008 and promptly started her career in the computer software industry at a company called Varolii Corporation in Seattle, WA. After Varolii was acquired by Nuance, she quickly became enamored by the many ways to improve customer satisfaction through omni-channel customer service solutions, all powered by AI. She is now a Senior marketing project manager for Nuance, but more than that, she is a firm believer in Nuance's technology and teams who are tirelessly focused on customer-centric outcomes and results.

View all posts by Alixson Bell