Speech Recognition Reassessed, July 15, 2015

A. Michael Noll
Columns · 2015
Original: http://quello.msu.edu/speech-recognition-reassessed-by-a-michael-noll/
Archived copy, Internet Archive snapshot 2015-08-10

Speech Recognition Reassessed by A. Michael Noll

by William Dutton

July 15th, 2015

Speech Recognition Reassessed

A. Michael Noll

© 2015 AMN

My new Garmin GPS navigation unit has made me reassess my previously negative opinion of automatic speech recognition. I am now impressed. But it has taken many decades for me to change my mind.

Back in the 1960s, when I was working in speech research at Bell Labs, speech recognition was in its infancy. Not only was the performance not very good, but also the applications were challenging to identify. A keyboard and knobs were far easier to use. Speech recognition a half-century ago required the largest computers that were then available – and they did not recognize speech in real time. Today's speech recognition is much better and produces results in real time – and on devices we have in our cars or carry in our pockets. The technology has progressed significantly.

John R. Pierce (the famed father of Telstar) had written a paper "Whither Speech Recognition?" in the Journal of the Acoustical Society of America in 1969. He predicted a dismal fate for automatic speech recognition. I followed with my own paper taking a similarly skeptical view of automatic speech production.* I believed that graphical display of information was better than machines that spoke to us. But I did acknowledge that speech recognition might help a "driver to keep eyes on the road."

We thought that imperfect automatic speech production would be more acceptable than imperfect speech recognition. That is because we believed that humans were better at understanding automatic synthesized speech than computers were at understanding human speech.

My Garmin represents the state of the art in both automatic speech recognition and production, as it not only recognizes speech but also creates synthetic-speech directions when navigating me along a route. Neither is perfect. Some pronunciations are comically wrong – and it will not recognize the

names of some restaurants. Most of the time, it is great – but at other times, it is frustrating. But speech is much better – and safer — than attempting to touch the screen to enter data while driving.

Since my Garmin GPS unit sometimes will not recognize the correct pronunciation, I have to pronounce words incorrectly but in a way that it does recognize. My Garmin is making me conform to it, and I wonder whether we will over time have people with a Garmin accent!

I am told that the speech recognition by Google and Apple are very good. These systems send the speech to a remote cloud-based computer that has considerable processing power and speech-recognition software. But when using a computer, I still find it is easier to just type my request for information. Speaking to a computer, for me, just seems like more energy and effort. But I guess that if I used a smart phone, then speaking might be easier than typing on a small screen. However, at my old age, I am not smart enough for a smart phone!

*Noll, A. Michael, "Whither Speech Production?" The Journal of the Acoustical Society of America, Vol. 47, No. 6 (Part 2), June 1970, pp. 1614-1616.

Tags: automatic speech, Garmin, GPS, John Pierce, speech, speech recognition

One thought on "Speech Recognition Reassessed by A. Michael Noll"

William Dutton says: July 15, 2015 at 2:14 pm This reinforces my view that the key issue in thinking about alternative futures is timing. Reply
This reinforces my view that the key issue in thinking about alternative futures is timing.

Leave a Reply

Archival reading copy rendered from the Internet Archive snapshot of the page above, for the A. Michael Noll collection at bre.art, with the author's permission. Text is the author's; this layout is not the original publication's.