

Emerging trends: A tribute to Charles Wayne

KENNETH WARD CHURCH

IBM, Yorktown Heights, NY, USA

e-mail: kwchurch@us.ibm.com

*(Received 12 September 2017; accepted 12 September 2017;
first published online 9 October 2017)*

Abstract

Charles Wayne restarted funding in speech and language in the mid-1980s after a funding winter brought on by Pierce's glamour-and-deceit criticisms in the ALPAC report and 'Whither Speech Recognition'. Wayne introduced a new glamour-and-deceit-proof idea, an emphasis on evaluation. No other sort of program could have been funded at the time, at least in America. One could argue that Wayne has been so successful that the program no longer needs him to continue on. These days, shared tasks and leaderboards have become common place in speech and language (and vision and machine learning) research. That said, I am concerned that the community may not appreciate what it has got until it's gone. Wayne has been doing much more than merely running competitions, but he did what he did in such a subtle Columbo-like way. Going forward, government funding is being eclipsed by consumer markets. Those of us with research to sell need to find more and more ways to be relevant to potential sponsors given this new world order.

I just returned from the 2017 Annual Jelinek Memorial Workshop on Speech and Language Technology.¹ This year was the first year that the workshop was completely funded by industry, with no government support.

Speech and Language have been amazingly well funded since the mid-1980s, largely thanks to Charles Wayne. Many government projects are intended to run for five years or so by construction but somehow those of us working on speech and language have enjoyed nearly continuous funding for three decades. However, the music may not continue for much longer now that Charles Wayne is unlikely to continue to be as active as he has been. In addition, now that the technology has progressed to the point where industry is prepared to take the lead, it is harder to make the case that the government needs to jump start the speech-and-language business. Even though defense applications are just as relevant as they ever were to the government sponsors, it is becoming clear that industry is now prepared to invest in speech and language products and services at levels that go way beyond what can be done with typical government grants. That said, industrial investments tend to be more product focused, with less discretion for more speculative long-term research. It isn't clear if academic research will continue to enjoy the level of support that it has become accustomed to over the past three decades.

¹ <https://www.clsp.jhu.edu/workshops/>

Mark Liberman has given a number of talks on somewhat related topics.² He starts by describing a funding winter from 1975 to 1986, where there was no US research funding for machine translation or speech recognition. Liberman describes some of the events that gave rise to the winter. Pierce, a highly respected researcher (and Vice President) at Bell Labs, responsible for the transistor (among other things), is a no-nonsense engineer with little patience for Artificial Intelligence³:

- Artificial intelligence is real stupidity.
- Funding artificial intelligence is real stupidity.
- I thought of it the first time I saw it.
- After growing wildly for years, the field of computing appears to be reaching its infancy.
- I resent artificial intelligence because I feel that it is unfair to computers. But then, artificial intelligence people did devise LISP, which is pretty good.

Liberman attributes the funding winter to Pierce's criticisms in the ALPAC report⁴ (Pierce *et al.* 1966) and 'Whither Speech Recognition' (Pierce 1969). The two criticisms are similar in many ways though they discuss different subjects (machine translation and speech recognition, respectively). More substantive differences are length and tone. The ALPAC report is a committee effort. Documents produced by committees tend to be longer and more diplomatic.

The Committee cannot judge what the total annual expenditure for research and development toward improving translation should be. However, it should be spent hardheadedly toward important, realistic, and relatively short-range goals.⁵

'Whither Speech Recognition' was written a few years later, with Pierce as the sole author. Liberman selected these quotes from 'Whither Speech Recognition' to illustrate the difference in tone:

- A general phonetic typewriter is simply impossible unless the typewriter has an intelligence and a knowledge of language comparable to those of a native speaker of English.
- Most recognizers behave, not like scientists, but like mad inventors or untrustworthy engineers. The typical recognizer gets it into his head that he can solve 'the problem'. The basis for this is either individual inspiration (the 'mad inventor' source of knowledge) or acceptance of untested rules, schemes, or information (the untrustworthy engineer approach).
- The typical recognizer ... builds or programs an elaborate system that either does very little or flops in an obscure way. A lot of money and time are spent.

² See <https://www.simonsfoundation.org/event/reproducible-research-and-the-common-task-method> for a video, and <https://web.stanford.edu/~vcs/AAAS2011/AAAS2011Liberman.pdf> for slides.

³ See page 339 of Nilsson (2010) as well as https://en.wikipedia.org/wiki/John_R._Pierce

⁴ <http://www.mt-archive.info/ALPAC-1966.pdf>

⁵ See page 33 of Pierce *et al.* (1966).

No simple, clear, sure knowledge is gained. The work has been an experience, not an experiment.

Lieberman ends the discussion of the events leading up to the funding winter with a slide titled, ‘Tell us what you really think, John’.

- We are safe in asserting that speech recognition is attractive to money. The attraction is perhaps similar to the attraction of schemes for turning water into gasoline, extracting gold from the sea, curing cancer, or going to the moon. One doesn’t attract thoughtlessly given dollars by means of schemes for cutting the cost of soap by 10%. To sell suckers, one uses deceit and offers glamour.
- It is clear that glamour and any deceit in the field of speech recognition blind the takers of funds as much as they blind the givers of funds. Thus, we may pity workers whom we cannot respect.

There is little debate over the events that led to the funding winter. The more interesting part of Liberman’s argument is what came after the winter. Liberman attributes the 1986 funding restart to a particular DARPA program manager, Charles Wayne, and a new idea to protect against ‘glamour and deceit’.

Wayne’s new idea was to emphasize evaluation. There would be a well-defined objective evaluation, applied by a neutral agent (NIST) on shared datasets (many of which were distributed by the Linguistic Data Consortium).⁶ According to Liberman, no other sort of program could have been funded at the time, at least in America.

Wayne’s idea wasn’t an easy sell, especially at first. Liberman points out that ‘not everyone liked it’.

Many Piercian engineers were skeptical: you can’t turn water into gasoline, no matter what you measure.

And speech researchers, according to Liberman were ‘disgruntled’:

It’s like being in first grade again – you’re told exactly what to do, and then you’re tested over and over.

But Wayne’s idea eventually succeeded because ‘it worked’.

Why did it work? Liberman starts out with the obvious. It enabled funding to start because the project was glamour-and-deceit-proof, and to continue because funders could measure progress over time. Wayne’s idea makes it easy to produce plots such as Figure 1,⁷ which help sell the research program to potential sponsors.

A less obvious benefit of Wayne’s idea is that it enabled hill climbing. Researchers who had initially objected to being tested twice a year began to evaluate themselves every hour. An even less obvious benefit, according to Liberman, was the culture. Participation in the culture became so valuable that many groups joined without funding. As obvious as Wayne’s idea may appear to us today, Liberman reminds us that back in 1986, ‘This obvious way of working was a new idea to many!’

⁶ <https://www ldc.upenn.edu>

⁷ <https://www.nist.gov/itl/iad/mig/rich-transcription-evaluation>

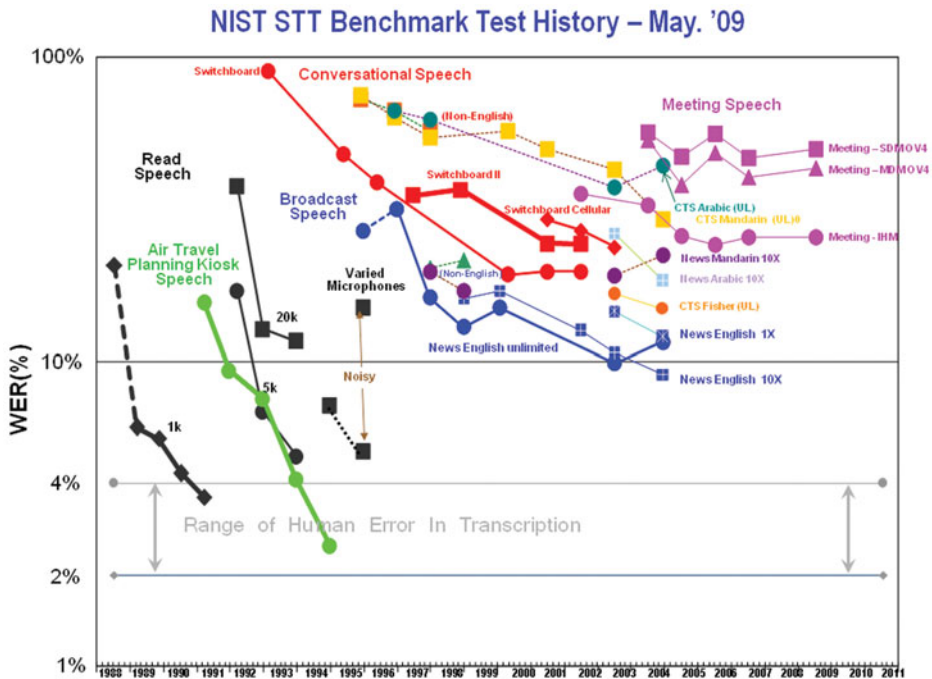


Fig. 1. (Colour online) Thirty years of progress in STT (Speech-to-Text).

Given Liberman's compelling description of the events leading up to the funding winter, as well as Wayne's funding restart in 1986 that led to three decades of prosperity, the obvious question is what happens to the field after Wayne?

One could argue that Wayne has been so successful that the program no longer needs him to continue on. These days, shared tasks and leaderboards have become common place in speech and language (and vision and machine learning) research. Papers describing shared tasks often receive massive citations (Godfrey *et al.* 1992; Marcus *et al.* 1993; Canavan *et al.* 1997; Cierri *et al.* 2004; Lewis *et al.* 2004; Bennett and Lanning 2007; Lin *et al.* 2014). I have no doubt that shared tasks and leaderboards (like Kaggle)⁸ are here to stay, but I am concerned that Wayne has been doing so much more than merely running competitions. He did what he did in such a subtle way that the community may not appreciate what's its got until its gone.⁹

Wayne used an unassuming Columbo-like approach to find out where everyone stood on every question, so he knew (before everyone else) what would fly and what wouldn't. The research community tends to think of funding agencies as the source of funding, but in fact, funding agencies are often middle men, somewhat like real estate agents. Real estate agents don't own the house (either before or after the transaction). Their job is merely to be market makers. They bring the relevant parties together.

⁸ <https://www.kaggle.com/competitions>

⁹ https://www.youtube.com/watch?v=Nym1P-B0_ws

In real estate, there are sellers' markets and buyers' markets. Sometimes the sellers are in a stronger position and sometimes the buyers are in a stronger position. Our business tends to be a buyers' market. It is relatively easy for program managers to find researchers who have excellent research to sell. The hard part of the job is to find buyers within the government who need to buy our research. Some program managers have focused on the sellers (research community), but the more successful ones tend to be the ones like Wayne that find arguments that work with the buyers.

Going forward, there are macro trends that suggest there will be fewer and fewer buyers in the government (with less and less impact on the market). I used to say, when I was a graduate student in the late 1970s, that I wouldn't use a computer that I could afford. There wasn't much of a consumer market in those days. Computers were so expensive that the market was limited to potential buyers that could afford them (mostly large enterprises and government). But that's completely changed these days. At some point, my wife came home with a PC that she bought to write her thesis on, and I had to admit that it compared favourably to the million-dollar PDP-10 that I wrote my thesis on. An iPhone today is more powerful (and much lighter) than a multi-million-dollar Cray Super-Computer back in the day.¹⁰

The government used to have considerable clout with computer companies. When they asked Cray for certain features, Cray was happy to listen. But these days, government and enterprise markets have given way to consumer markets. If a government agency wants a feature, they have to stand in line behind consumers (and no one will be all that interested in what the government has to say because there's more money in phones than super-computers).

What does this mean for us? The buyers of what we have to sell (academic research) used to be in government agencies, but going forward, they will be in consumer businesses. Consumer businesses have different priorities. When I worked at Microsoft, one of my colleagues was asked at a conference why he was working on machine translation *out of* English when much of the field was working on machine translation *into* English. He put his hand up to his ear and suggested that the US government was interested in listening to what the rest of the world was saying, whereas Microsoft was hoping to sell products and services to a larger market (beyond English). As a practical matter, the sponsor has considerable leverage to guide our research in so many ways, including whether we translate out of English, or into English. More generally, those of us with research to sell need to find more and more ways to be more and more relevant to more and more potential sponsors in a new world order where government and enterprise markets have been eclipsed by consumer markets.

References

- Bennett, J., and Lanning, S. 2007. The netflix prize. In *Proceedings of KDD Cup and Workshop*. California, USA: San Jose.
- Canavan, A., Graff, D., and Zipperlen, G. 1997. Callhome American English Speech. In *Linguistic Data Consortium*. Pennsylvania, PA, USA: University of Pennsylvania.

¹⁰ https://en.wikipedia.org/wiki/Cray_X-MP

- Cieri, C., Graff, D., Kimball, O., Miller, D., and Walker, K. 2004. Fisher English Training Speech. In *Linguistic Data Consortium*. Pennsylvania, PA, USA: University of Pennsylvania.
- Godfrey, J., Holliman, E., and McDaniel, J. 1992. SWITCHBOARD: Telephone speech corpus for research and development. In *ICASSP*, Washington, DC, USA: IEEE Computer Society, pp. 517–520.
- Lewis, D., Yang, Y., Rose, T., and Li, F. 2004. RCV1: A new benchmark collection for text categorization research. *JMLR* **5**: 361–397.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, L. 2014. Microsoft COCO: Common objects in context. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) *Computer Vision ECCV 2014*. Springer, Cham: Lecture Notes in Computer Science, vol 8693.
- Marcus, M., Marcinkiewicz, M. A., and Santorini, B. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics* **19**(2): 313–330, MIT Press.
- Nilsson, N. 2010. *The Quest for Artificial Intelligence*. 32 Avenue of the Americas, New York, NY: Cambridge University Press.
- Pierce, J. *et al.* 1966. *Machines: Computers in Translation and Linguistics*. Report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences. National Academy of Sciences, National Research Council. see <http://www.mt-archive.info/ALPAC-1966.pdf>.
- Pierce, J. 1969. Whither speech recognition? *JASA* **46**(4B): 1049–1051.