A History of Secure Voice Coding: Insights Drawn from the Career of One of the Earliest Practitioners of the Art of Speech Coding

JOSEPH P. CAMPBELL, JR., and RICHARD A. DEAN

The history of speech coding is closely tied to the career of Tom Tremain. He joined the National Security Agency in 1959 as an Air Force lieutenant assigned to duty at the Agency. Little did he know then that this assignment would shape his career as well as the future of speech coding.

Thomas E. Tremain was the U.S. government's senior speech scientist. He was a recognized leader and an expert in speech science. Tom's work spanned five decades of state-of-the-art modem and speech coding innovations that are the basis of virtually every U.S. and NATO modem and speech coding standard. His efforts have been critical to U.S. and NATO tactical and strategic secure communications programs.
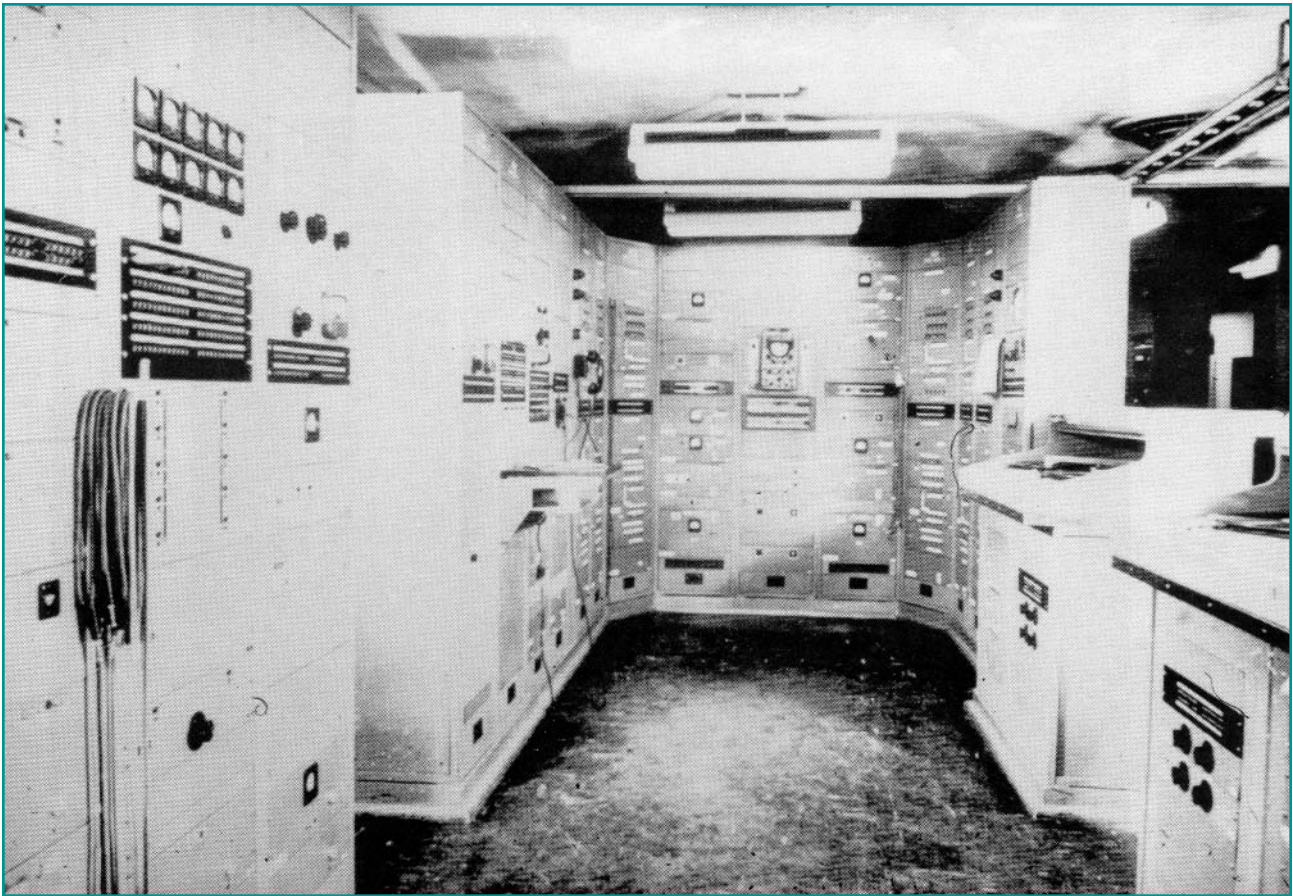
[Editor's Note: Since the original article was published, Mr. Tremain passed away on October 5, 1995.]

NSA inherited the responsibilities, traditions, and expertise of voice coding for enciphered speech applications from the Army Security Agency in 1952. The engineers and technicians who had participated in the famous SIGSALY vocoder used by Roosevelt and Churchill for planning the "D-Day" invasion [1, 2] were still developing their craft at NSA. SIGSALY, shown in Figure 1, was a vocoder-based system related to the "Talking Machine" first introduced by Homer Dudley of Bell Labs at the 1939 World's Fair. Developed with Bell Labs, it consisted of a bank of ten bandpass filters spaced approximately at the bands of equal articulation for speech, from baseband up to 3000 Hz. Each filter could be excited by one of six logarithmically spaced amplitudes developed by Harry Nyquist in the first applica-

tion of PCM. A "Buzz"/"Hiss" generator was used as an exciter for the vocoder corresponding to the voiced/unvoiced attribute of each 20-ms speech segment. Balance of the "Buzz"/"Hiss" generator, or voicing, represented a major factor in the quality of the speech. Early practitioners of speech coders, like Tom, can still be found today speaking "Aaahhh"/"Sshhhhh" into voice coders to test this balance.

From the time of SIGSALY until Tom arrived at NSA, several generations of voice coders had been developed in conjunction with Bell Labs. The KO-6 voice coder, developed in 1949 and deployed in limited quantity, was a close approximation to the 1200-bps SIGSALY voice coder. This was followed in 1953 by the 1650-bps KY-9 shown in Figure 2. Using a 12-channel vocoder and hand-made transistors, it was one of the earliest applications of solid-state technology. This reduced the weight of SIGSALY's vacuum tube technology from 55 tons to a mere 565 pounds! In 1961, Tom's first project was the development of the HY-2 vocoder, the last generation of U.S. channel vocoder technology. The HY-2, shown in Figure 3, was a 16-channel 2400-bps system using "Flyball" color-coded modular logic to reduce the weight to 100 pounds. Between 1962 and 1964, Tom created the first simulation of a channel formant vocoder in a mainframe digital computer and, between 1966 and 1968, he helped develop the first digital channel vocoder.

Even the best of U.S. vocoder technology at that time was limited by the analog technology that was the basis of its implementation. As the analog filters and amplifiers vary with age and
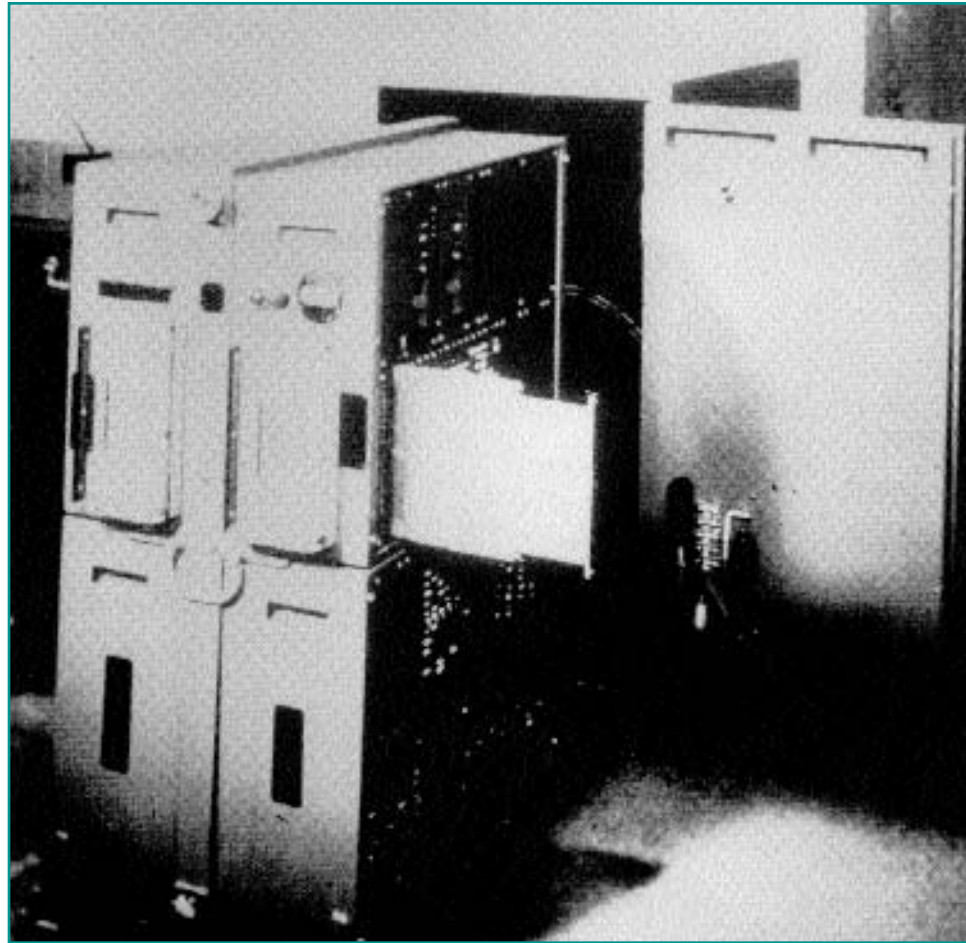
Fig. 1. A view of SIGSALY through the rear doors of a tractor trailer. The first real-time secure voice system, circa 1944

temperature, so does the sensitive tracking required between the speech analyzer and speech synthesizer. Performance in the field never approached laboratory performance; users, starting with Churchill and Roosevelt [2], were reluctant to use systems that had a synthetic "Donald Duck" quality. President Johnson refused to use the HY-2 because of its poor quality and, as a result, deployment was limited.

Tom's most significant contribution to voice coding was in seeing the possibilities of digital signal processing for voice. In an era when the state of the art was analog tuned circuits, Tom imagined a change to computer-based processing of speech – a radical shift in thinking. Tom pioneered this new approach, again collaborating with Bell Labs, to develop the Linear Predictive

Coding (LPC) generation of voice coders. Tom's predecessors and contemporaries never quite understood how the deck of punched cards he carried down the hall to a Honeywell computer was ever going to result in a vocoder. Tom developed the subtle techniques necessary to load and invert a matrix in real time in the fixed-point arithmetic necessary for such an operation. A former director of research at NSA ridiculed Tom for suggesting such an outlandish proposal; that very method is common practice today. So that systems could run faster on the computers of the day, he developed new constructs, such as the Average Magnitude Difference Function (AMDF) as a replacement for the multiply-intensive autocorrelation function. Finally, in 1974 Tom demonstrated the first real-time computer simulation of LPC-10 on the CSP-30 computer, a mile-

Fig. 2.
The KY-9
transistor-
ized secure
voice sys-
tem,
circa 1953

Tom Tremain at work on a prototype 1960s era
Channel Vocoder

stone in signal processing history. This led to a whole new family of NSA vocoder products, the STU products, built around the first generation of AMD2901/TRW multiplier-based high-speed bit-slice signal processors that forever changed the way voice coding was accomplished. Today's STU-III, shown in Figure 4, is the third-generation desktop telephone that uses an enhanced LPC-10 and supports secure voice users throughout the government. Voice coders, long associated only with exotic encryption schemes, are finding numerous applications today for wireless communications, voice mail, and synthetic voice applications. Today's voice coders used in satellite communications and hand-held digital cellular telephones are direct successors of Tom's early work.
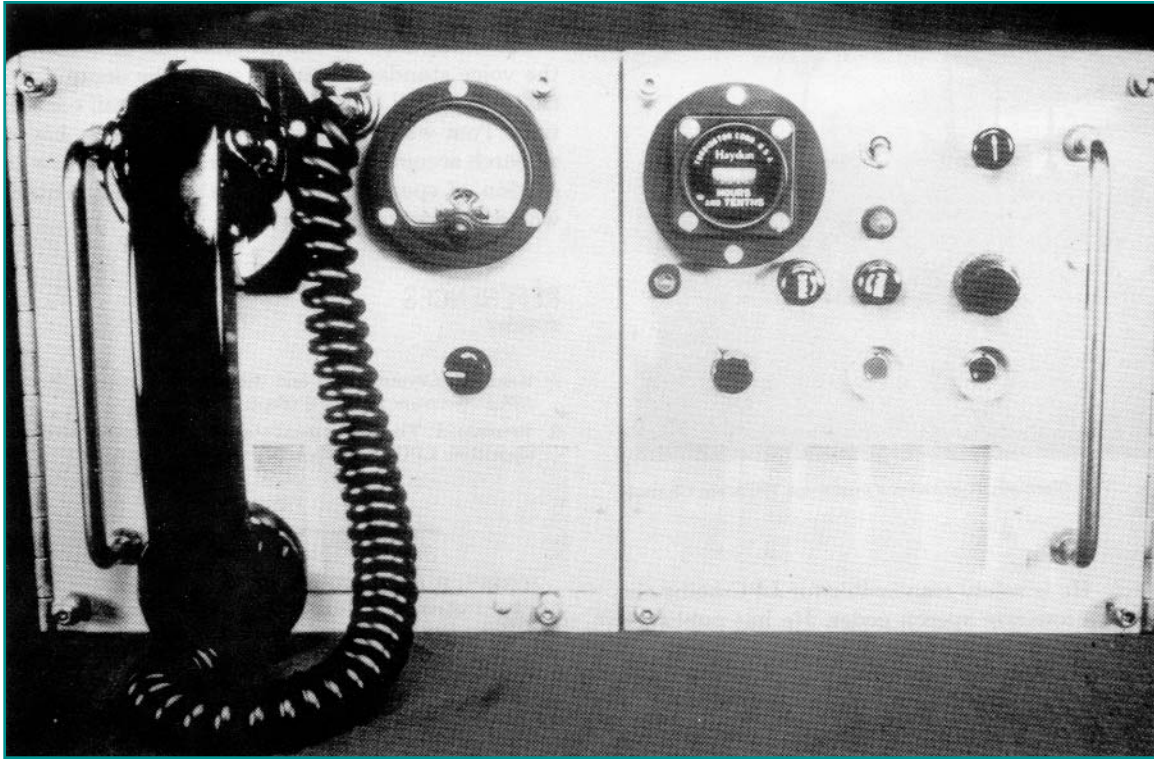
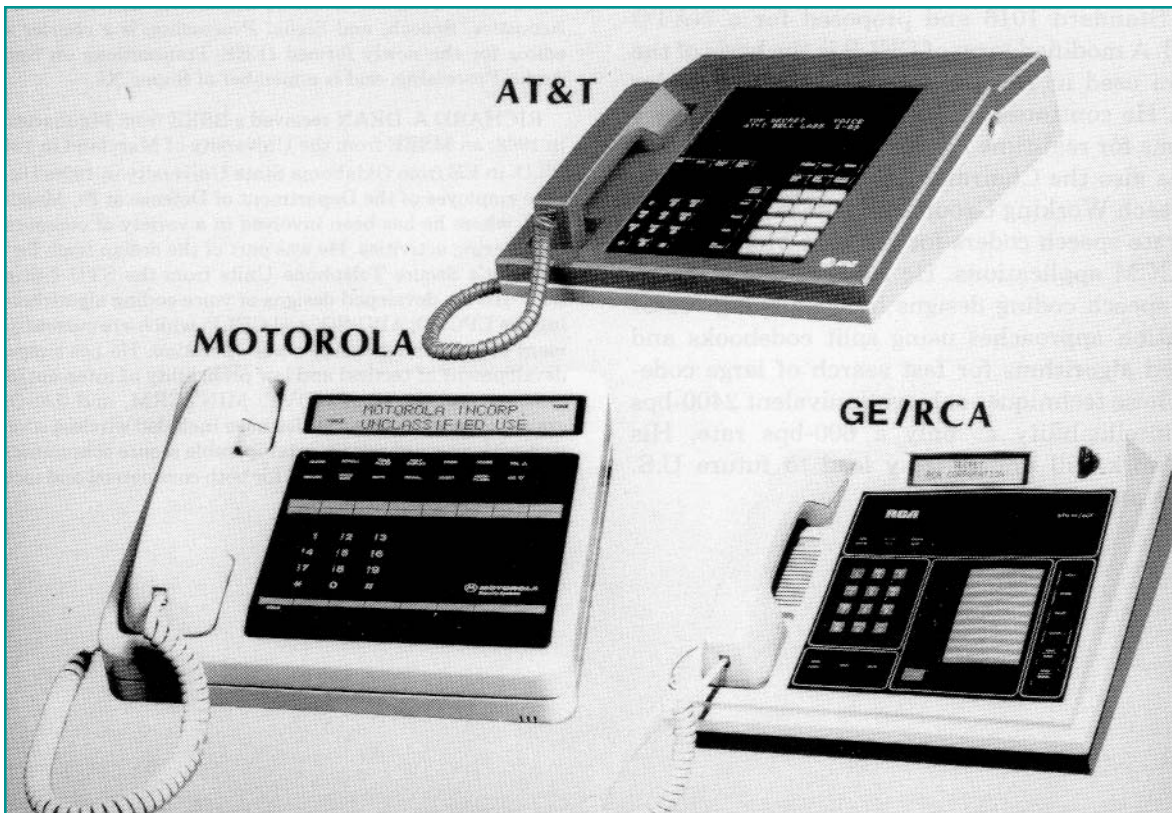Fig. 3. The HY-2 channel vocoder, circa 1961

Fig. 4. The STU-III secure voice terminal family, circa 1986

Tom was the chairman of the U.S. government's Digital Voice Processing Consortium for over twenty years. He reported directly to the under secretary of defense for command, control, communications, and intelligence. The Consortium is the recognized U.S. government forum for voice algorithm developments. It promotes innovative research, removes duplication, evaluates new algorithms, and promotes technology insertion. Under Tom's leadership, the Consortium became a prestigious U.S. forum for advancing state-of-the-art speech technology research. The Consortium activities and Tom's personal designs have had a far reaching impact on operational and planned communications systems totaling billions of dollars of national assets. His voice compression work facilitated high-quality secure communications worldwide. Senior U.S. government officials, national communications leaders, and other senior scientists frequently sought his opinions. He instituted speech testing comparison standards, promoted the establishment of a world-recognized independent test facility, initiated innovative voice compression development programs, and started and monitored voice research programs at universities. His original work on continuously variable slope delta modulation, linear predictive coders, vocoders, adaptive predictive coders, code excited linear predictors, modem technology, channel simulation, and intelligibility and quality test methodology forms the basis of most voice communications used today. For these efforts, the secretary of defense awarded Tom with U.S. DoD Meritorious Civilian Service Medals in 1985 and 1992.

Tom chaired many of the IEEE ICASSP and Speech Tech sessions and was a primary organizer of early ICASSPs and Speech Coding Workshops. He personally disseminated throughout the world the LPC source code that enabled much of today's research. He was solely responsible for LPC being the ubiquitous low-rate speech coder of its day. He published numerous extensively referenced papers. One of his papers

[3] on LPC is perhaps the most widely referenced paper in speech coding.

Tom's later achievements were as numerous and important as his long-term successes. His team's Code Excited Linear Prediction (CELP) speech coder, used in the STU-III, was endorsed as Federal Standard 1016 and proposed for a

Tom Tremain (left) was recognized by the IEEE in 1993 with a special award "For his major contributions and sustained leadership in digital speech coding, modem technology, and U.S. and NATO standards for speech coding and communication." [4]

NATO standard. A modified form of CELP is the basis of the algorithm used in North American digital cellular systems.

Tom was also the chairman of the NATO Narrowband Speech Working Group where he introduced low bit rate speech coders for HF ECCM and VHF/UHF ECCM applications. He led research in low bit rate speech coding designs by developing vector quantization approaches using split codebooks and tree-based algorithms for fast search of large codebooks. These techniques achieve equivalent 2400-bps speech intelligibility at only a 600-bps rate.

In looking back on Tom's career, one might point to the voice standards that he set or the organizations that he chaired as a measure of his overall contribution. Tom would more likely have pointed not to his own research accomplishments, but rather to the new generation of speech researchers that he mentored who still carry on his work.

The authors and their colleagues. Photo was taken in front of "SIGSALY" exhibit at its October 13, 2000, unveiling ceremony at the National Cryptologic Museum. Those pictured span four generations of voice coding and information security researchers. From left: Mary "Tina" A. Kohler, who worked on code excited linear predictive coding (CELP) and led the evaluation of the current generation of MELP vocoders; Joseph P. Campbell, Jr., who worked on LPC and led the CELP team; Mitchell R. Brown, who dismantled SIGSALYs and developed analog vocoders; Mahlon E. Doyle, who operated a SIGSALY and became a leading cryptographic designer; and Richard A. Dean, who worked on digital vocoders (including LPC).

References

[1] Kahn, D. "Cryptography and the origins of spread spectrum." *IEEE Spectrum* 21, No. 9 (September 1984), 70-80.

[2] Doyle, M. "Private communication." October 13, 2000. The extent of actual use by Roosevelt and Churchill is still unclear and needs further research to clarify; however, there is evidence of a call between Truman and Churchill on VE-Day (Victory in Europe).

[3] Tremain, T. "The government standard linear predictive coding algorithm: LPC-10." *Speech Technology Magazine* (April 1982), 40-49.

[4] Campbell, J. "In Memory of Thomas E. Tremain 1934-1995." *IEEE Transactions on Speech and Audio Processing 4*, No. 1 (January 1996), 1.