Experts Debate Authenticity of "Shah" Tape

The CBS Evening News *broke the story, but it may prove to be a hoax*

ANNOUNCER: *This is the* CBS Evening News *with Walter Cronkite.*

CRONKITE: *The Shah of Iran was forced into exile 15 days ago. But about a week before he left Tehran, he met with his military staff, and according to a tape, reportedly made at that session, the Shah outlined a plan to put himself back in power. Experts who have studied the tape say it is the Shah's voice, and the tape was not edited. We have a report from Joseph Benti of KNXT in Los Angeles.*

It was the inauspicious start of the great Shah tape snafu. The *CBS Evening News* broke the story on 31 January 1979. Since that time, CBS New York, KNXT Los Angeles, the *New York Times*, several other news organizations, the White House, the Shah of Iran, and a flock of voice identification experts have been quietly battling over the authenticity of the tape. The outcome is still far from clear, which shows how a little science and a lot of media coverage can make a big deal out of a questionable story.

The tape was brought to KNXT, a CBS-owned station in Los Angeles, by an Iranian student who said it was a secret recording made as the Shah talked with his top military advisers shortly before he fled Iran. KNXT hired an interpreter and three audio experts to check the tape. After making a positive identification of the voice, which spoke in Farsi, the Persian language, KNXT fed the story to CBS New York. The gist of the tape, as aired on the *CBS Evening News*, was the Shah urging his advisers to create a prolonged civil war in order to give him a chance to return and regain power. "Through creating hostility and hatred between the army and the people," the voice said, "by ordering the soldiers to shoot freely and kill, you could throw these two weighty forces against each other. A long civil war thus created will gain us enough time during which we could devise counter measures, perhaps by introducing a government which would appear to some extent acceptable to the people."

Joseph Benti, reporting from Los Angeles, asked on the *CBS Evening News*: "How do we know that that is the voice of the Shah? For the past 2 weeks, we have spent hundreds of man-hours verifying that fact. Two of the leading voiceprint experts in this country, one at Michigan State, the other at UCLA, today stated that beyond any doubt, the voice you heard on that tape is, in fact, that of the Shah of Iran."

Others were not so sure.

On Friday, 2 February, the *New York Times* reported that their expert was having doubts about the identity of the voice. That same day, CBS New York sent a team out to Chicago to interview the *Times* expert, Anthony J. Pellicano. The CBS team arrived too late to report Pellicano's growing "doubts" about the tape, but that night's *CBS Evening News* did carry a story by Robert Pierpoint, filmed outside the White House, saying that "a government official in a position to know" had evidence that the tape had been tampered with.

It was an obvious oversight. KNXT news director Jay Feldman told *Science* that they had asked two of their voice experts to check the identity of the speaker, but had not asked them to look for signs of tampering. Their third expert had checked for tape tampering, but it was with a widget called a "voice stress analyzer," a controversial device that some call a hoax. The oversight was unfortunate. By rearranging the tape, or inserting a mimic's voice, or deleting key words (such as "not"), a trickster might totally change the original meaning of the tape. And it got worse. By the following Monday, 5 February, Pellicano had finished analyzing the tape for the *New York Times*. He called it an out-and-out fraud.

CBS New York was on the spot. What started as a Los Angeles–generated scoop had now turned into a thick tangle of conflicting stories. Cronkite broke the story on the *CBS Evening News*, saying that the Shah was telling his army to turn its rifles on the people. The *New York Times* now said it wasn't true. Informed but unnamed government sources said a real tape of the Shah had been rearranged and that a mimic had added words. To top things off, the Shah, through his Ambassador to the United States, Ardeshir Zahedi, declared the tape a complete fabrication. He also threatened to sue.

The CBS news empire shook. John Lane, vice president and deputy director of CBS Network News, told *Science* that the firm of Bolt, Beranek and Newman in Cambridge, Massachusetts, has now been asked to check out the tape. Richard Bolt is considered one of the top tape experts in the country, having most recently chaired the National Academy of Science's committee on the evaluation of sound spectrograms (see page 854). Asked why CBS had hired Bolt, Lane replied: "At this point we don't know who is right. We're just trying to find out what the truth is." According to Lane, it may take another week.

> What started as a scoop turned into a thick tangle of conflicting stories.

In the meantime, KNXT and the *New York Times* have turned to battling over the experience and reputations of their experts, KNXT saying they have the big names in the field, the *Times* touting Pellicano's rather off-beat track record. On the KNXT side, the credibility battle has degenerated into name calling, one KNXT staff person hitting Pellicano as "nothing but a small-time investigator out for some publicity."

The top card in KNXT's hand is Oscar L. Tosi, director of the Institute of Voice Identification at Michigan State, and one of the founders and now director of the board of the International Association of Voice Identification. It is the only organ-

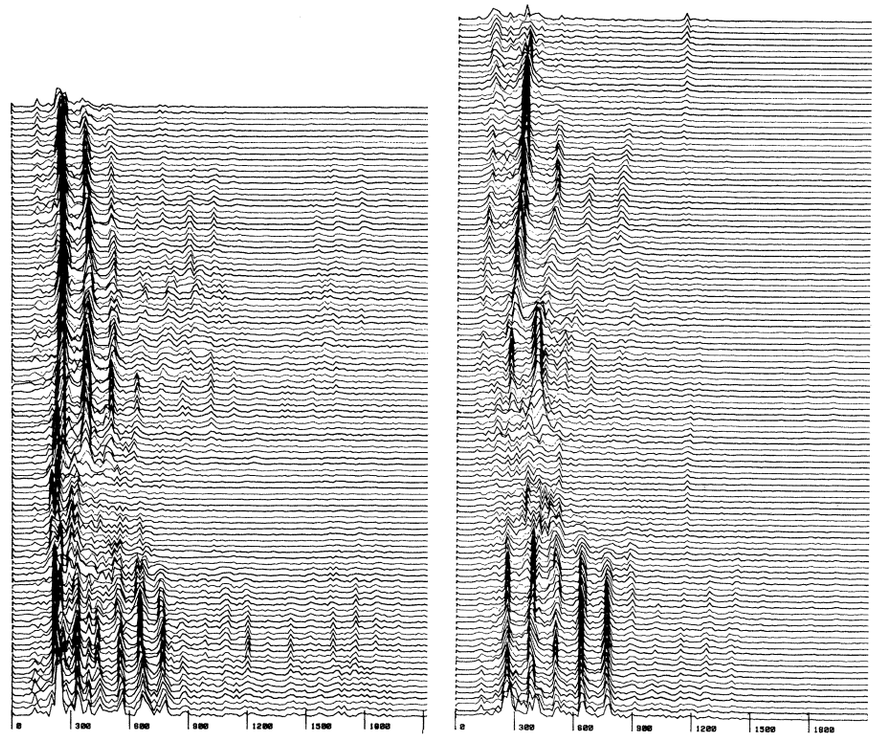
0036-8075/79/0302-0852$00.75/0

ization of its kind in the world (with a total of 18 members). Tosi also has studies, about 40 court appearances, a book, and long experience on his side. Next in the KNXT lineup is George Papcun, a doctoral candidate in Linguistics at UCLA who has done work in the UCLA Phonetics Laboratory and has testified in court about a dozen times on voice identification issues. Last on the KNXT roster is Robert Clark, a private investigator who has worked for the Los Angeles Police Department. Clark uses the "voice stress analyzer," though one observer had "no idea what Clark could possibly do with it."

The one and only expert of the *New York Times* is Pellicano, a self-styled maverick who is president of Voice Interpretation and Analysis, a one-man outfit on the outskirts of Chicago. Pellicano is a private investigator and computer buff who has no academic credentials and no court cases to his record but who has gained a reputation for bright ideas in audio analysis. He was an expert witness on the 18 ½-minute gap in one of President Nixon's tapes, showing how it could have been accidental. He also slighted the conspiracy findings of the House committee investigating the assassination of President John F. Kennedy, pointing out an obvious error in their analysis of a tape.

It is Pellicano, understandably, who touts techniques used in analyzing the Shah tape rather than flashing credentials. At first he worked with tapes provided by the *New York Times*, but after these proved to have too much hiss he was sent copies of the KNXT tapes by CBS New York. These included tapes that were known to be recordings of the Shah's voice and these were matched against the tape purported to be of the Shah's voice. Pellicano first asked an interpreter to pick out similar words from each tape. They were—in English—economy, people's, unfortunately, oil, mistakes, riots, and workers. He then made sound spectrograms, sometimes called voice prints, of the words. This process records the frequency and strength of a voice signal through time.

After these were compared by visual and computer methods, Pellicano gave his verdict: not the same person, because two words clearly contained frequencies not present in the matching spectrograms. Are the results believable? According to Ernst F. W. Alexanderson, president of Voice Identification, Inc., who supplies Pellicano and many others with their sound spectrographs, Pellicano's try was good, but it could have been better. He says that the minimum number of words for a visual comparison of spectrographs is ten, and they must be of exceptional clarity. Alexanderson told *Science*, moreover, that he had never heard of the computer technique that Pellicano used for further analysis of the tape.



Pictured are two of Pellicano's computer printouts. The horizontal axes show frequency (0 to 2100 hertz), and the vertical axes show time (0 to about 500 milliseconds). The left printout is a sample of the Shah saying, in Farsi, "mardum." In English it means "people's." The right printout is "mardum" from the tape under investigation. There is no way, says Pellicano, that they could have been made by the same person.

Papcun, working for CBS, took the standard approach, shying away from computer analysis and relying instead on the visual comparison of spectrograms. He told *Science* that he made more than 50 spectrograms from the tapes and used sounds and syllables, rather than whole words, to make his spectrograms. But, according to Alexanderson, "When you isolate such small signals, you increase the possibility of losing the unique features of a voice."

Side-stepping the whole problem of how much signal is needed for a good spectrogram, Tosi first asked a panel of five trained listeners, two of whom were native speakers of Farsi, to aurally compare the tapes in question with several tapes known to be recordings of the Shah. They came up with a positive identification. Tosi also analyzed the tape with a computer program that he had been working on. This program helps to identify a voice by electronically stacking signal upon signal until the individual variation between words is lost and only the unique features of a voice stand out. He calls it choral speech. But some critics slight it (and Pellicano's computer approach), saying that computer analysis is not yet an accepted method in any court of law. And they claim that even trained listeners, using Tosi's other method, can be fooled by a good mimic.

To help resolve the Shah tape riddle, CBS New York has asked Tosi to reanalyze the tape and explain the method of frequency filtering and amplification he used to improve the quality of the signal. According to Tosi, other news organizations, such as the *NBC Nightly News*, apparently asked their experts to look at the tape, but decided not to touch it because they found a restricted range of frequencies that might make a judgment difficult to reach.

And other news organizations were wary. On a radio talk show in Los Angeles, Robert W. Gibson, the foreign editor of the Los Angeles *Times*, recently debated the ethics of the Shah tape story with Jay Feldman, news director of KNXT. Gibson said the Los Angeles *Times* also received a copy of the tape, felt it was a rumor not worth investigating, and did not print a word on it. He also said that if the Los Angeles *Times* had done the story, they would have asked for opinions from at least 15 experts in voice identification before going to press.

Voice Analysis: It's Not So Easy

A controversy over the validity of voice analysis has ebbed and flowed ever since 1966, when the first voicegram was admitted as evidence in court. Since then, federal and state courts have ruled both for and against its use as evidence. To help clear the air, the Federal Bureau of Investigation in March 1976 asked the National Research Council of the National Academy of Sciences (NAS) to evaluate the technique. That report, "On the Theory and Practice of Voice Identification," has just been issued. It concludes that "the technical uncertainties concerning the present practice of voice identification are so great as to require that forensic [real-life] applications be approached with great caution."

To prevent voicegrams from being overvalued by judge or jury, the NAS report calls for several limitations on their use. In a jury trial, for instance, a hard-headed note of caution would be read aloud before any evidence was presented. Another recommendation: get both sides in a case to agree to abide by the evidence. Another: use voicegrams only if other evidence corroborates the identification. The final recommendation is to allow voicegrams only if opposing experts are scheduled to testify.

Caution is necessary, the NAS report says, because an individual's voicegram can change. In this sense, they are fundamentally different from finger prints (though they are sometimes mistakenly referred to as voice prints). The ridges on a finger never change, a case of duplication having never been found. Not so with speech. In most instances, there is some variability in one word repeated over and over by the same person—but this is less than the variability of that word repeated by other people. The NAS report warns, however, that no relationship has yet been found between the variation in an individual's voice and the variation among many.

Present practice, according to the report, involves aural and visual comparisons of one or more known voices with an unknown voice. With the aural method, the examiner listens to recordings of known and unknown voices in order to observe general similarities, to screen out less useful samples, and to index the recordings that are useful for further study. With the visual method, the examiner compares voice patterns, attempting to relate only the same phonemes, syllables, or words in each of the different voicegrams. Further, it is best to compare elements taken from the same phrase, or to isolate the element under analysis from other speech sounds.

Under ideal conditions in the laboratory, the report noted, aural-visual methods can be quite accurate, with error rates as low as 1 or 2 percent in controlled experiments. The largest study to date on voice identification was done in 1971 by Oscar L. Tosi at Michigan State University. For this study Tosi used a total of 34,992 trials, including 11,663 that mimicked the less-than-perfect conditions found outside the lab (rapid speech, background noise, telephone recordings). Using the visual inspection of the spectrograms, he found a mere 6.4 percent of these tests resulted in false identifications.

As a result of this study, many courts in the early 1970's admitted voicegrams as evidence. Up to 1975, they had been admitted by 14 of the 15 federal trial judges that had ruled on the issue, and by 35 of the 37 state courts that had ruled on it. But the mood did not last. Several later studies questioned Tosi's methodology and cast doubt on the validity of voicegrams that are used in less-than-lab-perfect conditions. As a result, they are currently accepted as evidence by courts in only 23 states, having recently been ruled against in the highest courts of California, Maryland, Michigan, and Pennsylvania.

The NAS study reflects the trend toward wariness. Considering that many of the nine committee members who wrote the report make some part of their living off voice identification and analysis, the report is refreshingly objective. The committee concluded that the analysis of sound spectrograms is more "empirical art" than science. There is today, they added, not enough information "to predict whether, and if so, when, the aural-visual process of voice identification will become a fully developed technology based solidly on science."—W.J.B.

Compounding the whole problem of the tape's validity is the possibility that widely scattered specialists are working with different tapes (though Tosi, Papcun, and Pellicano have now, thanks to CBS New York, all received the same versions). The *New York Times*, for instance, originally got a tape from Bahman Sholevar, an Iranian member of the Executive Council of the Iranian National Front in the United States, which opposes the Shah. KNXT got theirs from an unnamed member of the Moslem Student's Association in Los Angeles. And according to Robert Pierpoint at CBS, the government's tape was bought on the grounds of Tehran University, where hundreds of copies were being sold at discount prices.

The multitude of tapes and the conflicting analyses have led some anti-Shah Iranian students to see signs of a conspiracy. They claim that there is indeed a tape of the Shah inciting his troops to riot, but that SAVAK (the Iranian secret police), upon finding that such a tape existed, immediately counterfeited tapes, mixing the Shah and an imitator in an effort to discredit the whole affair.

Whatever the ultimate explanation, almost everyone now enmeshed in the Shah tape controversy says that more time is needed to pin down the truth—time to use a number of methods to check tape validity; time to compare tapes with those of other experts; time to analyze the results.

When the *CBS Evening News* broke the story on 31 January, the KNXT experts had been at work on the tape for only a few days, though the station, according to Feldman, received the tapes on 17 January. Tosi received his tapes on Monday, 29 January. Papcun received his first good samples of the Shah's voice in Farsi on that same day, though he had worked with the suspected Shah tape since 25 January. Tosi told *Science* that, in retrospect, he feels KNXT was too eager to make a positive identification, and that they should have asked him to check for tape tampering along with identity (though some have questioned why he didn't right at the start check it on his own).

Asked how long a complete check of the tape would take, Tosi replied, "Using the computer, the spectrograph, the oscilloscope, using trained listeners and moving by microseconds while trying to detect any spikes, any irregularities, I suppose it would take a good solid month, working full time, to analyze the validity of the tape." It seems a pity that someone didn't think of that sooner.

—WILLIAM J. BROAD