

Title : A COMPUTERIZED SPEECH RECOGNITION SYSTEM FOR ANESTHESIA RECORD KEEPING

Authors : Andrew J. Sarnat M.D., Michael L. Quinn Ph.B., N. Ty Smith M.D.

Affiliation: Department of Anesthesia, University of California and
VA Medical Center, San Diego, California 92103

Introduction. The anesthesia record, though covering at most a few hours, is dense with information. Every anesthesiologist knows the problem of keeping accurate and complete records by hand while other tasks demand attention. Physiologic monitored data can be recorded by a computer directly and easily, but the development of a fully computerized record has been limited by available methods for entry of such data as drug injections or blood transfusions. Currently available methods of data entry include alphanumeric keyboards, specialized keyboards, bar code readers, and digitizing pads, all of which share the disadvantage of diverting the user's eyes and hands. At UCSD we are developing a system of data entry using human speech, where the user wears a microphone and a computer "understands" spoken words.

Description. The system is called EARS (Entry to the Anesthesia Record by Speech) and comprises a general-purpose micro-computer, a speech recognition computer (SRC), a speech synthesizer, and a close-talking, noise-cancelling microphone. We developed a vocabulary of approximately 350 words to describe drugs, inhalation agents, fluids (including blood products), monitoring, airway, equipment, positioning, intraoperative events, blood gases, and chest auscultation. The vocabulary can be easily modified or expanded.

Words are spoken one at a time, separated by short pauses. Before first using EARS for data entry, each user pronounces every vocabulary word while the SRC produces compressed digital representations of the words in that user's voice (the templates), which are stored in memory. During data entry, each utterance is similarly digitized and compared against the templates. The best-match template identifies the recognized word. Since each user has his own templates, consistency of pronunciation is more important than pronunciation itself, and any dialect or accent is accommodated.

Correct word recognition is more likely when an utterance is compared against fewer templates. Therefore the vocabulary is subdivided, and recognition is restricted to template subsets. A set of rules (the syntax) determines which subsets follow which, and as words are spoken by the user, the SRC moves from subset to subset, always anticipating the next word. This allows access to a

virtually unlimited vocabulary without overburdening the SRC.

Data entry is organized into sentences. A sentence is a sequence of from two to nine words describing a single item for the record, beginning with a subject word, e.g., "Event tourniquet down." Shortened sentences are often acceptable due to implied meaning, e.g., "Drug morphine four milligrams IV" and "Drug morphine four" are equally acceptable. The user signals the end of a sentence by a long pause. After a complete sentence is entered, EARS verifies its recognition by reading back the sentence (using the speech synthesizer) through a loudspeaker or the anesthesiologist's earpiece. Upon hearing the machine's version of the sentence, the user may use voice commands to (1) accept and record the sentence as is; (2) erase and re-enter it; or (3) change or add single words until correct. Once a sentence is verified and accepted, recognition can begin anew on the next sentence.

Ambiguous utterances, noise, and other spurious sounds are rejected by the SRC. By voice commands the user may render EARS unresponsive to speech, hold outside conversation with others, then "wake up" the system for further data entry. As an option, each active subset may be listed, in turn, on a TV screen to help new users learn the vocabulary and syntax.

Discussion. The system described above is not a stand-alone record keeper. It is intended to serve as an input device into some other computer which would maintain an actual hard-copy anesthesia record. EARS has been tested under real conditions in the operating room. After adjusting its noise rejection characteristics for this environment, the recognition accuracy appeared quite acceptable. It was not disturbed by the use of electrocautery.

In the operation of EARS, two-way verbal communication is established between man and machine. Although its development is far from complete, we feel there is great potential for highly natural and efficient automated anesthesia record keeping.