At Haskins Laboratories the Veterans Administration has sponsored research with the goal of developing a high performance reading machine. The term “high performance” refers here to the combined behavior of both the reader and the machine. From the program's inception this requirement has been understood to mean that the machine must convert the printed text into clear intelligible English speech, so that the blind listener needs no recourse whatever to an intermediate code.

The first indications that speech, and speech alone, was the essential for high performance reading emerged from research undertaken by the Laboratories during the mid 1940's for the Committee on Sensory Devices. (This committee was established by Dr. Vannevar Bush as a part of the wartime Office of Scientific Research and Development to study potential applications of technology to alleviate the disabilities of war veterans.) Many reading devices which produced nonspeech acoustic outputs were examined and all were found to yield poor reading performances. This was despite the fact that some of the blind readers who were tested had practiced over long periods of time and were motivated to a degree that very few sighted people would be likely to match. Whether or not one subscribes to the view (as many of us at the Laboratories do) that speech is a unique code having no surrogate, the fact remains that reading machines with nonspeech output cannot provide easy, fluent access to the books that sighted people read, however useful such devices may be in more limited tasks. Learning is a chore and using them taxes the users' patience, so that ultimately they fail to meet the needs of more than a fraction of the blind.

It was for these reasons that the Veterans Administration program, revived by Dr. Eugene F. Murphy, Director of the Research Center for Prosthetics, included a project aimed at high performance. After several years of research in which many methods of generating speech were tried, a substantial portion of this objective has been reached. An operat-
ing prototype now exists in the Laboratories and various aspects of its performance are under evaluation. Results of these tests indicate that consideration should now be given to setting up a working-scale reading machine system in a large centrally located library. This pilot facility would supply subscribers with requested tape recordings of texts read by the machine, and would provide an opportunity for continued research and development to meet the problems that can be expected to arise in the early stages of any pioneering application.

The desirability of such a reading service made available countrywide is indisputable. Blinded veterans, in company with students and many other blind people who wish to take up some rewarding occupation, find themselves in need of rapid access to specialized printed information. Reading services provided by human readers, whether volunteer or paid, are invariably slow. It is not unusual for the voice recording of a new book to take several months to reach the hands of a blind subscriber. Machine spoken recordings—particularly if the publisher's typesetting tape is available—could be generated at high speed and be available to blind listeners within a few days of their request.

Here, briefly, is how we are now generating the synthetic speech that makes such a service possible. The existing laboratory prototype reader performs the conversion of typewritten text to audible speech in four main stages. Stage one is carried out by a small computer (and associated optical scanning equipment) which reads the typescript and temporarily stores the text in digital form on magnetic tape. The second stage uses a 150,000 word phonetic dictionary which is accommodated on a computer disc file. Having retrieved the text from magnetic tape, it is converted word by word from the alphabetic spelling to the corresponding phonetic spelling. The phonetic text is interspersed with small superscript marks which indicate the syllables that should receive stress and intonation symbols (\(\uparrow\) and \(\downarrow\)) which indicate the pitch contours that the computer should generate during speech synthesis. If the text contains words not found in the dictionary or lacks sufficient punctuation to achieve a satisfactory vocal output, then the third stage (manual editing) is invoked. Here a human editor views a phonetic display of the text and, with the aid of a keyboard, introduces the missing words or prosodic symbols as required. The machine then proceeds to the fourth and final stage, speech synthesis, in which the phonetic symbols are used to compute control data for a terminal analog speech synthesizer. Using a buzz source (filling the role of the human larynx), a noise source, and several dynamically controllable filters, the synthesizer is able to generate voicelike sequences of sound under computer program control which listeners hear as English speech. (Fig. 1).
This speech is found to be highly intelligible by most people (particularly after a little practice) and yet there is no doubt that it still sounds different from human speech. We expect that, with continued research, the computer synthesis programs and speech quality will continue to improve. Even now our experience in the Laboratories with student listeners (both blind and sighted) shows that the performance of the existing reading machine system could provide a significant service. Our
goal is to start this service on a pilot basis while improvements to the speech and plans for broader service are underway.

The technical description of the Prototype Reading Machine system which is given here was supplemented at the conference by the presentation of a 10-minute 16-mm. sound film. Figure 2 is a page of the phonetic text used in this film. The original alphabetic input follows:

Figure 2.—Phonetic text used in film presented at Chicago Conference.

The battle of Gettysburg was fought on the first, second and third days of July 1863. On November 19 of the same year, a portion of the battle field was dedicated as a final resting place for those who died there. The main address on that occasion was one of two hours in length, delivered by Edward Everett, the best known orator of the time. After his address, Lincoln delivered the short speech now so famous.