ABSTRACT
With the increasing role of computers in teaching, there is little doubt that we will eventually want them to talk to us and allow us to speak to them. This review presents an experimental approach of voice I/O techniques to computer-based teaching English on the basis of Expert Type System. Advanced applications of the speech processing technology and some special linguistic/information problems are discussed.

INTRODUCTION
The (micro)computer in education has both traditional research on linguistic database and has provided a more precise experimental vehicle for controlling the presentation of instruction and measuring responses (1).

The basic hypothesis is that the computer offers an opportunity:
(i) to provide the teacher with a powerful resource to manage individual learning within the terminal room;
(ii) to enable the student to follow learning procedures which incorporate step-by-step feedback and to stimulate individual attention and to gain assistance discreetly;
(iii) to make it possible to the teacher to observe and monitor the progress of the student in detail.

The objectives of the preliminary study are, firstly, to investigate the feasibility of the approach and the appropriateness of the software facilities being employed.

AUTOMATED LEARNING SYSTEMS
Evidence indicates that it is most productive to teach grammatical and lexical basics in context. In order to carry out the exercise the student must thus draw upon reading (knowing) English letters, vocabulary, and connected utterances, spoken by ten Russian speakers. In computer-based English teaching the student receives all of his training from the computer; the material used in the initial study provides 20-40 minutes work for students. Testing and teaching may be done on the computer via keyboard input by students. Responses may be saved on diskette and at the end of testing the student's score can be displayed on the screen immediately, or a printed evaluation can be reproduced depending on the computer programme and the test construction.

Questions to be printed are chosen from a master list and their descriptions (up to 1000 items). Each Automated Learning System (ALS) comprises several learning volumes: Training volume with its priority scoring due to three levels of complexity, grammatical and lexical databases; Reference volume with its grammar and lexicon retrieval; Encouragement volume.

Potential linguistic problems are worked out before they are translated into hard- and software. Need goals, constraints are described first, thus helping to solve the complex problem to be divided into a hierarchy of smaller problems to understand/control the whole process of learning. Each training step is related to previous and subsequent stages, and misunderstandings among students are avoided. Questions/answers from the packet of exercises in Control volume should be typed with a 'minimum-energy' solution.

Finally, prediction models of recognizer performance can be used in determining optimal operating conditions for voice input (4). A major problem is the identification of those factors having significant influence on the performance of the speech recognizer. W.A. Lee (1982) has compiled an extensive list of more than 80 variables including language and task factors (number of training passes, rectification, size of the active vocabulary, inter-word confusability), human factors (age, native accent, speaking rate, channel and environmental factors). The performance of the recognizer = 98-99% recognition accuracy, thus the average number of words which are incorrectly recognized is counted at each stage of the task (sometimes referred to as the recognition vocabulary). DIS-332 includes a study of extensive database collections of both isolated and connected utterances, spoken by ten Russian speakers.

Vocabulary size = 500 words, performance of the recognizer = 99.99%, recognition time = real branching factor 10.

Above mentioned configuration provides a large scale of opportunities while using voice I/O techniques to computer-based teaching English on the basis of Expert Type System. Advanced applications of the speech processing technology and some special linguistic/information problems are discussed.
about 3 to 6 depends positively on vocabulary size, complexity and confusability. The results show that equally-positioned phonemes can be better distinguished from among one another when different articulation types are used (e.g. plosive—fricative). Such features as voicing, nasality, affrication, duration and place of articulation are the primary channels of the intelligibility-relevant information. Different vocals tend to be good features of words to be recognized if they do not belong simultaneously to the high vowels "e" and "i" and to the deep vowels "o" and "u". We believe that for each doubling of the vocabulary size, the recognition accuracy tends to decrease by a fixed amount, which is different for each talker.

Yet there can hardly be any more important task in speech recognition than determining now well algorithms or devices work. Thus the error rate as a performance measure conveys no information about performance except the relative number of errors made on a given task. It tells nothing about the distribution of errors and the costs of making particular errors; depends on vocabulary size and doesn't reflect large vocabulary difficulty, the inherent acoustic confusability, the difficulty of the speaker, or the environment. A new information-theoretic performance measure is based, in part, on the idea that automatic as well as human speech recognition systems can be modelled as communication channels. A more meaningful measure, called the Relative Information Loss (RIL) would normalize the amount of information lost in a recognition process with the amount transmitted/5/. Woodard and Nelson/6/ propose combining the 'Human Equivalent Noise Reference' (HENR) method with a RIL method. HENR is based on the confusions between speech sounds by humans listening in noise. The model predicts the percentage word recognition rate, and the confusions at any signal to noise ratio for any vocabulary which has been defined in phonetic terms. This combined approach may be used to relate device performance to task difficulty.

Grammatical constraints whether they will be stochastic or deterministic, have the effect of decreasing entropy, increasing redundancy and hence decreasing error rate (entropy is, of course, a statistical property). Each natural language requires that some assumption be made about the likelihood of occurrence of trained difficulty at a given point in sentence.

Two reasonable assumptions are that the difficulties are equiprobable or distributed to maximize entropy. Under the medium entropy assumption

$$H_{eq} = \frac{\log_2 L(G)}{E_1(G)\{w\}}$$

so that entropy in bits/word is the base-two logarithm of the size of the language divided by average sentence length. For ALS redundancy is to be increased up to 20% against existing Expert Systems.

CONCLUSIONS

The development of faster microprocessors, larger memories, better printers and storage devices, together with pricing competition, will play roles, too. But the factor likely to be judged most significant in the academic microcomputer revolution will probably be the rate at which these recognition systems have gained widespread acceptance by humans in serving their diverse educational needs on the basis of ALS.

REFERENCES

/3/Петров А.Н., Туркин В.Л. "Система речевого диалога", Автоматическое распознавание речевых образов, Ч.1, Наука, 1986, с.97-99