

## SPEECH RECOGNITION

Yeskov I. E., *student*

Speech recognition technologies allow computers equipped with microphones to interpret human speech, e.g. for transcription or as a control method.

Such systems can be classified as to whether they require the user to "train" the system to recognize their own particular speech patterns or not, whether the system can recognize continuous speech or requires users to break up their speech into discrete words, and whether the vocabulary the system recognizes is small (in the order of tens or at most hundreds of words), or large (thousands of words).

Systems requiring a short amount of training can (as of 2001) capture continuous speech with a large vocabulary at normal pace with an accuracy of about 98% (getting two words in one hundred wrong), and different systems that require no training can recognize a small number of words (for instance, the ten digits of the decimal system) as spoken by most English speakers. Such systems are popular for routing incoming phone calls to their destinations in large organizations.

Commercial systems for speech recognition have been available off-the-shelf since the 1990s. However, it is interesting to note that despite the apparent success of the technology, few people use such speech recognition systems.

It appears that most computer users can create and edit documents more quickly with a conventional keyboard, despite the fact that most people are able to speak considerably faster than they can type. Additionally, heavy use of the speech organs results in vocal loading.

Some of the key technical problems in speech recognition are that:

- Inter-speaker differences are often large and difficult to account for. It is not clear which characteristics of speech are speaker-independent.
- The interpretation of many phonemes, words and phrases are context sensitive. For example, phonemes are often shorter in long words than in short words. Words have different meanings in different sentences, e.g. "Philip lies" could be interpreted either as Philip being a liar, or that Philip is lying on a bed.

Chuchilina L. M. *EL adviser*