
**Abstract**

A system for the recognition of spoken phrases is described. The recognizer assumes that the input utterance contains one of a known set of allowable phrases, which may be spoken within a longer carrier sentence. Analysis is performed on a syllable-by-syllable basis with only the strong syllables considered in the recognition process. Each strong syllable is represented in terms of a set of distinguishing acoustic measurements taken at time points in and around the syllable nucleus. Phrases are represented as sequences of strong syllables. All parameters used in recognition are derived from LPC coefficients. Input speech is limited to 3.3 kHz upper frequency. Recognition is completed within 1-3 s after the utterance is spoken. An interactive training facility allows flexible composition of key phrase sets.

Testing was performed for a number of phrase sets each containing ten or fewer phrases, and included equal numbers of talkers used in training and talkers not used in training. Average phrase recognition accuracy was 95 percent when parameters were derived from unquantized (i.e., 16-bit) LPC coefficients and 90 percent when the LPC coefficients were transmitted to the recognizer across the ARPA network at 3500 b/s. The recognizer has been incorporated into a user interface system where the parameters required to set up a point-to-point ARPANET voice connection can be established remotely by voice.