A SYLLABLE APPROACH TO THE SPEECH INFORMATICS

A.KNIPPER
Institute for the Problems of Information Transmission Academy of Sciences
101447, Moscow, GSP-4, USSR

ABSTRACT
A problem devoted to trends of syllable development for the usage in speech informatics systems is under consideration. It is noted that simple open CV syllables (C is a consonant, V is a vowel) are the most stable discrete phonetic units of continuous speech with respect to the context, speaker variability and noise. Problems of that type syllable classification and statistics for the Russian speech and their possible application in the design of the speech information are discussed. Some experiments on compilation syllable synthesis of the Russian speech of free contents and on analysis of the speech signal using CV fragments are briefed.

INTRODUCTION
The main problem of the speech informatics is development of man-machine communication systems on the base continuous speech. In that case speech communication between a user and a system is ensured with the best conditions. In continuous speech recognition / understanding the most promising approach is representation of the speech flow with the help of symbol sequences similar to the speech transcript with afterwards decoding at the word or phrase level /1, 2, 3/. The main requirement of that approach is transformation of a continuous signal into a discrete sequence of speech elements, phonetically stable to speaker variability, context, noise and other factors which influence the speech signal features. That case in the process of speech recognition / understanding the system operates with its new vocabulary training the efficiency for users is insured /4/.

In the process of synthesis of any piece of speech information an inverse problem is solved, i.e. a letter sequence is transcribed by phonetic symbols and then is transformed into the corresponding acoustic signal, and besides for comfortable usage it is desirable to synthesise any voice and any speaker at any type according to a user choice.

The choice of a phoneme as a phonetical symbol for a speech communication system is the most reasonable and convincing, as it permits relatively easily to pass to the conventional letter representation of any data, accessible and intelligible by broad circles of users. However, numerous researches on phonetics and speech informatics show that there is no direct relationship between speech segments and phonemes. The same sounds match speech segments with essentially different spectral and temporal characteristics, that is determined by context, positional and speaker variability of the speech. Simple open CV syllables have more stable characteristic especially those that are not cut off from left and right from a transition line, named CV fragments /5, 6/. It is considered that such syllables match base speech elements for Russian, Italian, Japanese and other languages /7, 8, 9, 10, 11, 12/ and is more widely used in different speech informatics. In the following sections it is shown that usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech. Usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech.

BASE ELEMENTS OF THE RUSSIAN SPEECH
It is considered that open syllable is a speech universal unit for the majority of languages /1, 2, 3/. In the speech informatics open syllables may be also preferred, due to the fact that there is a distinct transition from the corresponding consonant to the vowel in the interval of that open syllable that makes easier to label the continuous speech visually and with the help of technical means /5, 6, 7/. The number of open syllables for Russian speech is great, i.e. about 2500 /14/.

In the process of synthesis of any piece of speech information an inverse problem is solved, i.e. a letter sequence is transcribed by phonetic symbols and then is transformed into the corresponding acoustic signal, and besides for comfortable usage it is desirable to synthesise any voice and any speaker at any type according to a user choice.

The choice of a phoneme as a phonetical symbol for a speech communication system is the most reasonable and convincing, as it permits relatively easily to pass to the conventional letter representation of any data, accessible and intelligible by broad circles of users. However, numerous researches on phonetics and speech informatics show that there is no direct relationship between speech segments and phonemes. The same sounds match speech segments with essentially different spectral and temporal characteristics, that is determined by context, positional and speaker variability of the speech. Simple open CV syllables have more stable characteristic especially those that are not cut off from left and right from a transition line, named CV fragments /5, 6/. It is considered that such syllables match base speech elements for Russian, Italian, Japanese and other languages /7, 8, 9, 10, 11, 12/ and is more widely used in different speech informatics. In the following sections it is shown that usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech. Usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech.

BASE ELEMENTS OF THE RUSSIAN SPEECH
It is considered that open syllable is a speech universal unit for the majority of languages /1, 2, 3/. In the speech informatics open syllables may be also preferred, due to the fact that there is a distinct transition from the corresponding consonant to the vowel in the interval of that open syllable that makes easier to label the continuous speech visually and with the help of technical means /5, 6, 7/. The number of open syllables for Russian speech is great, i.e. about 2500 /14/.

In the process of synthesis of any piece of speech information an inverse problem is solved, i.e. a letter sequence is transcribed by phonetic symbols and then is transformed into the corresponding acoustic signal, and besides for comfortable usage it is desirable to synthesise any voice and any speaker at any type according to a user choice.

The choice of a phoneme as a phonetical symbol for a speech communication system is the most reasonable and convincing, as it permits relatively easily to pass to the conventional letter representation of any data, accessible and intelligible by broad circles of users. However, numerous researches on phonetics and speech informatics show that there is no direct relationship between speech segments and phonemes. The same sounds match speech segments with essentially different spectral and temporal characteristics, that is determined by context, positional and speaker variability of the speech. Simple open CV syllables have more stable characteristic especially those that are not cut off from left and right from a transition line, named CV fragments /5, 6/. It is considered that such syllables match base speech elements for Russian, Italian, Japanese and other languages /7, 8, 9, 10, 11, 12/ and is more widely used in different speech informatics. In the following sections it is shown that usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech. Usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech.

BASE ELEMENTS OF THE RUSSIAN SPEECH
It is considered that open syllable is a speech universal unit for the majority of languages /1, 2, 3/. In the speech informatics open syllables may be also preferred, due to the fact that there is a distinct transition from the corresponding consonant to the vowel in the interval of that open syllable that makes easier to label the continuous speech visually and with the help of technical means /5, 6, 7/. The number of open syllables for Russian speech is great, i.e. about 2500 /14/.

In the process of synthesis of any piece of speech information an inverse problem is solved, i.e. a letter sequence is transcribed by phonetic symbols and then is transformed into the corresponding acoustic signal, and besides for comfortable usage it is desirable to synthesise any voice and any speaker at any type according to a user choice.

The choice of a phoneme as a phonetical symbol for a speech communication system is the most reasonable and convincing, as it permits relatively easily to pass to the conventional letter representation of any data, accessible and intelligible by broad circles of users. However, numerous researches on phonetics and speech informatics show that there is no direct relationship between speech segments and phonemes. The same sounds match speech segments with essentially different spectral and temporal characteristics, that is determined by context, positional and speaker variability of the speech. Simple open CV syllables have more stable characteristic especially those that are not cut off from left and right from a transition line, named CV fragments /5, 6/. It is considered that such syllables match base speech elements for Russian, Italian, Japanese and other languages /7, 8, 9, 10, 11, 12/ and is more widely used in different speech informatics. In the following sections it is shown that usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech. Usage of CV syllables as base units of Russian permits to perform a rather distinct classification of context depended pairs of speech elements for Russian speech.
Consonants in the base CV syllables, Diphthongs A, belong to the soft vowel as well. Vowels pair A, V. It is similar properties in ranking according to their typical duration and for the articulation, but they essentially differ in spectral and temporal characteristics of a segment of sounds. At the same time the place of articulation for the consonants in the base CV syllables is achieved due to the maximum place of articulation of any syllable. The difference in transition of sounds is that effect is more associated with the place of articulation than the manner of production of consonants. Thus, the characteristics of consonants and vowels in the context of a speech signal are determined than the problem of the speech synthesis is much more complicated than the problem of the speech synthesis is much more complicated. Hence, the place of articulation of a consonant is the context of the syllable.

The syllable analysis has good prospects for usage in speech synthesis, since it establishes sufficiently adequate description of the physical and phonetic properties of a speech signal. However, that and higher levels of sound are provided by each national language and therefore should be thoroughly studied for any language.

CV ANALYSIS OF CONTINUOUS SPEECH

The syllable analysis of continuous speech pursuing an aim of automatic transcription of a speech signal is much more complicated than the speech synthesis. Difficulties of the speech analysis can be mainly caused by the variety of a speech signal and were briefly introduced in Introduction. However, the choice of an analysis unit is a very important step in the speech analysis. In addition to the number and the type of the base CV syllables included in the speech signal, the spectral and temporal characteristics become more important as well. The choice of a base CV syllables is an important step in the speech synthesis. Hence, the speech signal is extracted from the analyzed segment of sounds in the base CV syllables.

REFERENCES