# Differences in the Stability of Russian and Greek Speech Signal towards Noise Effects

I.Oparin, A.Kozina, D.Kocharov

## 1 Goals of Work

The goal of our work is the study of some peculiarities of speech comprehension in noise, stability of speech signal towards noise effect and allocation of elements which influence correct recognition. The first aim was to find such a level of correlation signal/noise at which 100% recognition of the material visibly fails. Second we tried to analyze the effect of speaker`s individual speech characteristics on speech comprehension. And our third aim was to make an attempt to allocate those elements in speech signal which are the most importaint for a listener in the process of speech recognition in noise.

The speciality of our research is that it is held on the base of two languages: Russian and Greek. Part of the research connected with the Greek language was carried out with the help of Wire Communications Laboratory at University of Patras, Greece. Choosing of Greek language as a second language for the research can be explained by those facts that it is, like Russian, is a flective language and the place of a word stress is free. We were mostly interested in comparison of the results obtained for Russian and Greek languages. One of the main goals was to discover that features which are common for these two languages and those which present different tendencies.

## 2 Sound Material for the Research

The set of phrases consists of 6 sentences built from frequent lexics and 6 sentences of non-frequent words. In each of the groups there is one sentence made from one-syllable words, two sentences from two-syllable words and three three-syllable words (with different stress positions). Our set of phrases includes all phonemes of Russian language and take into account their natural frequency[1]. This material provides a researcher with an opportunity of taking into account such factors as word length, word frequency and rhythmic structure. Phrases for Greek language were composed according to the same principles as Russian ones. This fact provides best conditions for a comparison of results.

Russian phrases which are the material for the research were read by four non-professional speakers (teachers and students of philological faculty). They have got

[1] Bondarko L.V. Phonetics of Modern Russian Language. SPb, Russia. 1998. p. 155. (in Russian)

different pitches: male (low – 115Hz and high – 150Hz) and female: (low – 160Hz and high – 260Hz). Greek sentences were kindly recorded for us in muffled chamber at Patras University (Greece).

The basis of the experiment is the analysis of noised phrases by phonetically naive (non-professional) listeners who were asked to write down what they hear. Here we must admit that the number of Greek listeners is much more restricted than we would like to have (about 15) since it was impossible to find native speakers in St.Petersburg. In Russian part of the experiment more listeners were involved but at the same time we consider this our weak point.

### 2.1 Used Noises

Two different kinds of noise were used for the experiment.

1. Noise which was synthesized to cover evenly all phonetically important frequencies. It has got even decrease of intensity at high frequencies area. This noise was taken from the base of noises "NOISEX-92", #6. Below it goes as "synthesized noise".
2. Compiled noise – obtained by adding four speech signals (different phrases read by different speakers), then normalized with intensive segments deleted. Thus we obtained two different noises for both languages concerned.

## 3 The Results of the Experiments

### 3.1 Russian Language

As it was stated above one of the main aims of the research was to find such a noise level for each noise when 100% comprehension is lost and those changes which took place with the spectrum of the speech signal.

At present statistic analysis of the phrases revealed two interesting tendencies. First, it appeared that apprehension of different timbres depends on the type of noise. When synthesized noise was applied mistakes in correct recognition began to appear at signal-to-noise ratio about 6dB. There we can say that phrases read by high voices (male and female) are apprehended better by 10% than those which were read by low voices. At the same time correct recognition of high male and high female voices is approximately equal.

If compiled noise is applied "first-mistakes" level is also 6dB. But in this case no dependency between correct apprehension and timbre was revealed.

So we can say that the effect of compiled noise on speech comprehension does not depend on timbre while the effect of synthesized noise, on the contrary, depends on individual characteristics of a speaker. Moreover, when average rate of the pitch of male high and female low voices coincide, male voice is more stable towards noise

effect. That is why we can suppose that the stability of speech signal towards noise effect can be determined not only by spectral characteristics of a signal.

Within this investigation there was an attempt to determine what in speech signal is valuable for listener in noised speech perception. Due to it an analysis of sonograms of all phrases for every speaker was made to discover those syllables, sounds and their combinations which are in a signal-to-noise ratio of 6 dB noised so that can not be truly recognized for sure. That is we received a "real sounding" in a noise for every phrase – that could hear a listener during experiment.

The most general conclusion which can be made agrees with the well-known fact that in Russian the initial part of a word is more significant for a listener in speech perception. It explains the fact that percentage of recognition of words (without taking into account correctly recognized words) with the falling out in the middle is higher than of words with falling out in the end.

Infrequent phrases which have no more frequent quasihomonyms are being recognized best of all. In case of existence of such quasihomonyms, they are recognized instead of correct ones by listeners in majority of cases. If a word has less frequent homonyms they don't write them as possible variant.

### 3.2 Modern Greek

The results received for compiled noise with ratio 2dB and 3dB are very interesting. Despite of little difference of noise degree between each other (2,2 dB and 3 dB) these two auditions differ a lot in recognition degree. It is possible that it largely concerned with that phrases with 3 dB ratio have been given after phrases with 2,2 dB ratio to the same listeners (due to lack of Greek native speakers in St.Petersburg). But taking in account a high degree of the noise and that phrases have been given in arbitrary order, we can say that a great difference in recognition will hardly explained only by this factor. Probably, this signal/noise ratio is a kind of a limit beyond which a great change of true speech signal recognition in a noise happens.

Results obtained for signal-to-noise ratio 6db confirms the assumption that a noise ratio of 2,2/3 dB is some kind of threshold for a speech recognition in a noise. It appeared that the difference of answers between 3dB and 6dB of noise is not so considerable as between 2,2 dB and 3 dB. As a result we had that such noise decreasing as 0,8 dB causes much more changes in a correct recognition than decreasing of noise by 3 dB of original 3 dB noise.

It turned out for Modern Greek as for Russian that perception and recognition of the noised speech depend not only on its own features but also on features of the noise, it suffer from. Probably such data would not be proved in a case of larger number of listeners, but today it is impossible to attract more native speakers of Modern Greek.

On our opinion, considerable distinctions appear in a field of "frequent" and "infrequent" phrases. For frequent phrases in mode of 6 dB difference of recognition is irrelevant (recognition of the material in the synthesized noise is by 11% better, such discrepancy may be explained by inaccuracy of data because of small amount of

listeners). But a clear increasing of correct recognition degree when using the compiled noise in comparison with monotonous one may be seen for sentences made of infrequent words.

Such correlations may be seen as well for mode of 3 dB. There is almost no difference in correct recognition for phrases made of frequent words, but as in the previous mode the synthesized noise greatly worsens perception of "infrequent" phrases in comparison with compiled one (by 15%). It proves once more the fact that a correct recognition greatly depends not on features of speech signal itself but on features of a noise it being noised by.

## 4 A Comparison of the Results for both Languages

We can consider the ratio when true comprehension of the speech fails to be a common for both languages. Obviously, 6 dB correlation ratio may be considered as such one. At the same time a general percentage of recognition of all the Modern Greek phrases is commensurable with that of Russian sentences (nearly 80 %).

The main difference of Russian and Modern Greek phrases recognition is that the ratio of 2,2 / 3 dB is probably a kind of a "threshold" for perception of Greek and there is no such tendency for Russian.

Within investigating there were discovered some contrary tendencies for Russian and Modern Greek. The main of them, on our opinion, is that Russian phrases made of frequent words are rather better recognized in the compiled noise than and "infrequent" phrases are quite better in the synthesized one. And we can see a contrary situation in Greek: phrases, made of infrequent words are better recognized in compiled noise. Anyway we must admit that it is necessary to increase the number of listeners in the experiment to ensure the results.

## 5 Authors Affiliation

All authors are students of Philological Faculty, St.Petersburg State University. The work was done at the Laboratory of Experimental Phonetics, department of Phonetics.
Ilya Oparin (Department of Mathematical, Structural and Applied Linguistics), Ulysses@io8176.spb.edu
Anna Kozina (Department of Balkanic Languages), Kozina@ak8719.spb.edu
Daniil Kocharov (Department of Mathematical, Structural and Applied Linguistics), daniil_kocharov@mail.ru