K.O. Uplisova

TALKING BIRDS: THE FEATURES OF MODIFICATED VOWEL SOUNDS

Saint-Petersburg State University
Russia, 199034, Saint-Petersburg, University emb., 7/9
Tel.: (812) 328-9706
E-mail: sehmet@fromru.com

Ability of some species of birds to imitate human speech represents particular interest because of the structure of their vocal apparatus essentially differs from human's one. It leads to that resonant characteristics of a syrinx of birds cannot be same as at a larynx of the human. As a result in most cases values of formants of vowel sounds of talking birds differ from human vowel. Nevertheless, vowels are identified with high probability. With the purpose of revealing of the acoustic keys, allowing to distinguish imitating sounds, modification of a series of vowels has been lead. The modification was spent by suppression of amplitude of spectral components, or shift of all frequencies of a spectrum. It has been revealed, that if in a spectrum of a sound there are expressed spectral components at a range of 1300-1700 Hz, this sound with high probability will be identified as "a", whether irrespective of there are expressed spectral components on low and high frequencies. If frequencies of the basic energetically expressed components are below 1000 Hz the sound will be identified as "o", irrespective of presence of the expressed spectral components above 1700 Hz, but not in case of their presence in a range of 1300-1700 Hz. In case of presence of energetically expressed spectral components with frequency below 500 Hz and at absence of the expressed spectral components on higher frequencies the sound will be identified as "u". If the expressed spectral components will be present since 1700 Hz a sound will be identified as "i".

INTRODUCTION The significance of a problem of recognition and perception of speech by the human is reflected in many works, these questions devoted to development. However, despite of the long time research, till now many aspects of perception and recognition of speech remain obscure, for example, what acoustic characteristics of sounds are perceptual significant for identification of phonetic categories of vowels. According to classical representations [1, 2] key attributes of vowels are frequencies of the first, the second and, sometimes, third of the formants reflecting resonant frequencies of a voice tract of the human. For each category of vowels frequencies of the first and second formants are stable, and their values are not crossed with values of formants of sounds of other category [3]. However, for example, formants of sounds with high fundamental frequency (above 300 Hz), babies present at vowel-like vocalizations and sounds of vocal speech, do not get in the borders certain for sounds of speech of the adult speaker. Nevertheless, similar vowels are identified by auditors with high probability [4, 5].

In this connection, the special interest is the ability of some species of birds to imitate human speech. For this purpose it is necessary for a bird to allocate those acoustic keys on the basis of which it will create a copy of a sound, words or phrases despite of individual distinctions of voices of humans surrounding it and characteristics of the acoustical system which has been adjusted on feature of species-specific signals. The anatomy of a vocal apparatus of birds also essentially differs from human [6, 7] and, hence, its resonant characteristics cannot be same as at the human. Moreover, some species of birds use two sources of sound vibration at a sound production. As the sounds of speech imitated by birds, are adequately recognized by the human, it is natural to assume, that there are the acoustic features, allowing to carry a sound to this or that category. [8]. In previous work [9, 10] the features of a spectrum of vowel sounds, stable for categories of vowel sounds "a", "o", "u" and "i", imitated by different birds have been specify. The account of values of frequency and amplitude of all spectral components of phonemes has allowed to reveal distinctions between the sounds belong to categories "a", "o", "u", "i", namely presence of energetically expressed spectral components at sounds of a category "o" in a range of 630-1260 Hz, "a" - 1260-1990 Hz, "u" - below 630 Hz, "i" - above 1990 Hz. However for check of the assumption that the maxima which are being on these frequencies, whether are critical for identification of this or that category it was necessary to modify sounds and to reveal the change of they phonetic identification.

Thus the given work is devoted to research of influence of change of acoustic characteristics of spectral maxima on phonetic identification of vowel sounds.

METHODS Vowel sounds of grey parrot and budgerigar with high probability identified as "a" and "o" have been selected for modification of spectral components. By means of fast Fourier transform (FFT-filter, Cool Edit Pro) were made changes of amplitude of spectral components. If at a spectrum of a sound there were expressed spectral maxima which could be characterized a unique value of frequency and amplitude reduction their amplitude was spent so that this maximum could not be allocated (the first series of modifications). If the spectrum of a sound was characterized by the expressed strip of spectral components, the change of its width by suppression of amplitude of spectral components in its certain part (the second series of modifications) was spent. For shifting frequencies (the third series) of vowels function Constant Stretch of program Cool Edit Pro was used. Shift aside lowering of frequency was carried out by steps with the relation of 0.79 all frequencies of the previous sound in relation to frequencies of the subsequent. For example if the frequencies of initial sound formants are 600 Hz, 1270 Hz, 2150 Hz as a result of a first step of shift aside lowering, frequencies of formants became 474 Hz, 1003 Hz, 1698 Hz, as a result of the second step - 374 Hz, 792 Hz and 1341 Hz, etc. Increase of frequencies was carried out with the relation 1.33. In this case frequencies of formants as a result of a first step are 798 Hz, 1689 Hz, and 2859 Hz. For each sound some steps of shifting of frequency aside its lowering and increase have been spent. The modificated sounds were listened by group of the seven auditors constantly participating in similar experiments. Whether before them the task was put to designate has changed the modificated sound a category in comparison with a category of not alter sound and how.

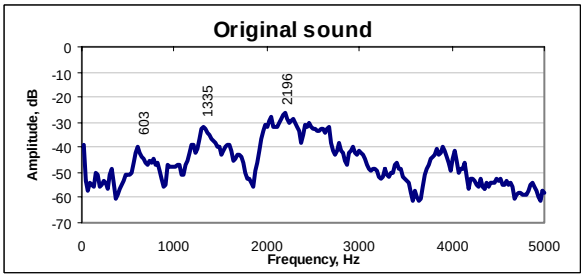
RESULTS And DISCUSSION The first modification was suppression of amplitude of spectral components below 750 Hz and above 1700 Hz. In this case no more than three auditors heard change of a category for all sounds. It has allowed to assume, that critical features of categories "a" and "o" are in a range of 750-1700 Hz.

In the first series it has been modified four sounds of the budgerigar (three "a" and one "o") and three sounds of a grey parrot (two "a" and one "o"). In three two formants sounds of the budgerigar suppression of the first maximum, and in one three formants sound, suppression of the first and second maximum, has led to essential lowering of identification the initial sounds as "a" and "o". The modificated sounds "a" and "o" with two formants began to be identified as "i" by all auditors. The identification of a three formants "a" also was changed at six auditors, but is less uniform (fig. 1, 1). However at preservation of first two maxima at three formants "a" (fig. 1, 2) essential change of identification of a category it has not been revealed. Suppression of the first formant (470-860 Hz) has not affected an estimation of a category (fig. 1, 3) while suppression of the second (1000-1700 Hz) has led to change of an estimation of a category practically at all auditors (fig. 1, 4). Suppression of the second and third formant has changed an estimation of a category "a" to a category "o" at the majority of auditors (fig. 1, 5). More than five auditors identified two modified three formants "a" by grey parrot as "o" if the second formant (1000-1700 Hz) was suppressed. Suppression of the first (400-900 Hz) has not affected identification of a sound as "a". Thus for sounds of categories "a" suppression of amplitude of a formant in the range of 900-1700 Hz leads to change of identification of a category of an initial sound "a". In this case more often the sound starts to be identified as "o".

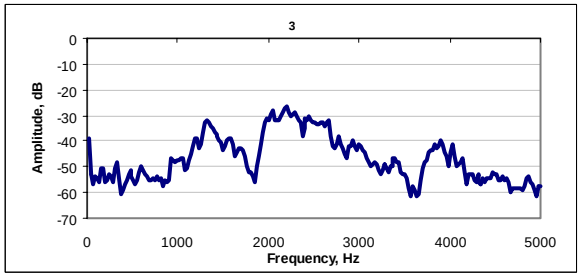
In the second series four sounds "a" by grey parrot has been modified. At a sound "a" with energetically expressed band of 700-1600 Hz suppression of amplitude of spectral components in a frequency range of 1300-2000 Hz has not caused changes identification of a category. All auditors identified the sound as "o" if the suppression of amplitude of spectral components in a range of 1000-2000 Hz was led; suppression of amplitude in a range of 800-2000 Hz has caused occurrence of estimations of a sound as "u" at a part of auditors. The similar picture was observed for other sounds. It has allowed to come to conclusion that at presence of energetically expressed components not above 1000-1100 Hz the sound will be identified as "o", not above 600-800 Hz - as "u".

In the third series the frequencies of two "a" of budgerigar and one grey parrot's "a" have been shifted. The formant frequencies of two initial sounds "a" of budgerigar were: 600, 1270, 2150 Hz (three formant sound) and 1380, 2600 Hz (two formant sound). As a result of the first step of shifting aside lowering the frequency of formants have made 450, 990, 1636 and 1033, 1980 Hz. Already the first step of shifting has led to change of identification of a sound by a part of auditors. The sound began to be identified as "o" a part of auditors. Only to the fourth step for three formant sound

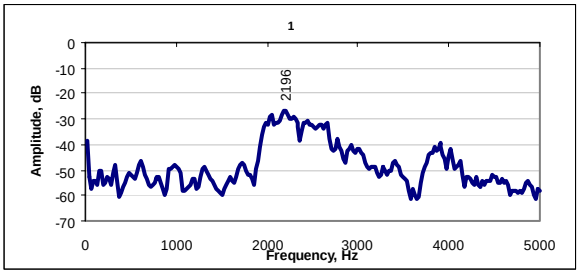
(frequency of formants 193, 409, 689 Hz) all auditors have fixed change of a category "a" on "o" or "u". At the two formant "a" from the third step (frequency of formants 580, 1119 Hz) identification remained constant: three auditors have determined a sound as "o" and "u", two as "a" and «e». Increase of frequencies has been carried out in six steps for three formant sound and five steps for two formant sound. Already as a result of a first step modification of three formant "a" two auditors identified a sound as "a" (frequencies of formants 750, 1636, 2700 Hz). All auditors have determined a category "a" at the fourth step (frequencies of formants 1530, 3337, 5770 Hz). Two formant "a" has been identified as "i" all auditors at the third step of transposing aside increases of frequency.



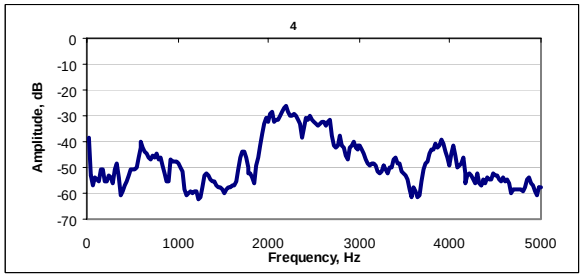
Original sound
Identification: 7 – «a»



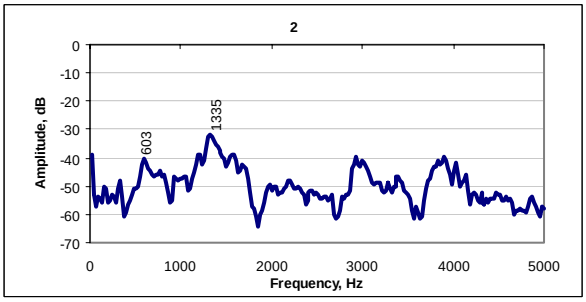
Amplitude suppression 470-860 Гц
Identification: 7 – «a».



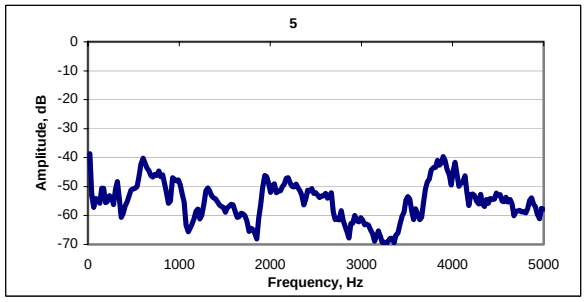
Amplitude suppression 470-1765 Гц
Identification: 3 – «и», 1 – «a», 1 – «o», 1 – «e», 1 – «э».



Amplitude suppression 1000-1700 Гц
Identification: 5 – «э», 1 – «e», 1 – «и».



Amplitude suppression 1722-2900 Гц
Identification: 4 – «a», 2 – «o», 1 – «y».



Amplitude suppression 1000-3500 Гц
Identification: 5 – «o», 2 – «y».

Fig. 1. Spectra of budgerigar sound “a” and it modifications
Under spectra the range of frequencies in which lowering amplitude of spectral components was spent, and amount of auditors stated this or that estimation is specified.

Already at a first step of shifting frequencies aside lowerings the sound "a" by grey parrot with energetically expressed strips of frequencies of 750-1800 Hz and 2200-2800 Hz identified a sound as "o" by three auditors. To the second step two auditors identified a sound as "u". To the fourth step of shifting three auditors identified a sound as "u", one as "o". For one auditor the estimation of a category of an initial sound has not changed no one of the modificated sounds. At transposing frequencies aside increases only to the sixth step two auditors identified a sound as "i", for the others the estimation of a category "a" has not changed.

Thus, transposing of frequencies of sounds leads to change of an estimation of a category "a" on "o" and "u" at lowering frequency, and on a category "i" at increase of frequency.

CONCLUSION If in a spectrum of a sound are present the expressed spectral components at a range of 1300-1700 Hz, this sound with high probability will be identified as "a", whether irrespective of there are expressed spectral components on low and high frequencies. If frequencies of the basic energetically expressed components are below 1000 Hz the sound will be identified as "o", irrespective of presence of the expressed spectral components above 1700 Hz, but not in case of their presence in a range of 1300-1700 Hz. At presence of spectral components in a range of 1100-1300 Hz, a sound it can be identified both as "a" and as "o". In case of presence of energetically expressed spectral components with frequency below 500 Hz and at absence of the expressed spectral components on higher frequencies the sound will be identified as "u". In case of simultaneous presence of the expressed spectral components with frequency below 500 Hz and above 1700 Hz, a sound it will be identified as "i", also it will be identified, if the expressed spectral components will be present only since 1700 Hz.

The work was supported by RFH №07-06-00546a

R E F E R E N C E S

1. Fant G. Analiz i sintez rechi. – Novosibirsk: Nauka, 1970. – 166 s. (In Russian)
2. Flanagan Dzh. Analiz, sintez i vospriyatie rechi. – M.: Svyaz', 1968. – 396 s. (In Russian)
3. Peterson G.E., Barney H.L. Control methods used in a study of the vowels // Journal of the Acoustical Society of America. – 1952. – V. 24. – P. 175-184.
4. Kulikov G. A., Andreeva N. G., Pavlikova M. I., Samokishuk A. P. Characteristics of vowel-like sounds in infants under six months of age // Doklady Biological Sciences. – 1999. - V.368. - P. 471–473. (In Russian)
5. Andreeva N.G., Kulikov G.A. Sung vowels' characteristics under different fundamental frequency // Sensory systems – 2004. - V.18. - №2. - P. 172-179.
6. Goller F., Larsen O.N. New perspectives on mechanisms of sound generation in songbirds // Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology. – 2002. – Vol. 188. – № 11-12. – P. 841-850.
7. Gaban-Lima R., Höfling E. Comparative anatomy of the syrinx in the tribe Arini (Aves: Psittacidae) // Braz. J. Morphol. Sci. – 2006. – V. 23. - № 3-4. – P. 501-512.
8. Patterson D.K., Pepperberg I.M. A comparative study of human and parrot phonation: Acoustic and articulatory correlates of vowels // JASA – 1994. – Vol. 96. – № 2. – P. 634-648.
9. Uplisova K.O. Acoustical characteristics of vowel in talking-birds' imitations // Sensory systems – 2004. - V. 18. - №3. - P. 199-205. (In Russian)
10. Uplisova K.O. The acoustic and auditor analysis of vowel-like sounds of gray parrot (Psittacus erithacus) and budgerigar (Melopsittacus undulatus) // Sensory systems – 2006. - V.20. - №3. - P. 229-237. (In Russian)