K.O. Uplisova

ACOUSTICAL CHARACTERISTICS OF TALKING-BIRDS' VOWEL-LIKE SOUNDS

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

Recently an interest in the mechanisms of production and perception of human speech sounds by talking birds increase in the connection with studying acoustical characteristics of vowels in different conditions of production. The goal of this work was determination the characteristics of talking birds' vowel those might help in their separation. The vowel-like sounds were isolated from words of canary, mynah and gray parrot. The analysis of waveform showed that vowel-like sounds of talking birds are not harmonic as a rule. In the spectrum of sounds it was estimated the amplitude and the frequency of the spectral maxima. The means of two most pronounced spectral maxima for several categories of vowels were similar. The amplitudes of spectral maxima relative to the amplitudes of the first spectral maximum were counted within each critical band. The disposition of the frequencies and the amplitudes of spectral maxima ratio significantly differed for different categories of vowels (o, a, i). Consequently, the possibility of separation bird-talking vowels exists on the strength of the differences of arrangement the frequency positions and amplitude ratios of the most pronounced spectral maxima.

INTRODUCTION. Recently an interest in the mechanisms of production and perception of human speech sounds by talking birds increase in the connection with studying acoustical characteristics of vowels in different conditions of production. It may suppose that humans can recognize the different categories of talking birds' vowels, they have some specific features common with vowels of human speech. And these features were searched in some works, where acoustical characteristics of talking birds vowel were investigated [1, 2, 3]. Not so long ago the detailed research acoustical characteristics of birds imitations [4] demonstrated that vowel-like sounds of Gray and Amazon parrots have the fundamental frequency near 300 Hz and cannot be distinguished based on the means of the first and second formants (these characteristics are important for recognition of long vowels in adult speech). Using the other methods of analysis hadn't given a definite result. The question how can humans recognized the speech of talking birds in spite of their high fundamental frequency and different mechanisms of production arises afterwards. The purpose of this work is determination the characteristics of talking birds' vowel-like sounds, those might help in their recognition.

MATERIALS AND METHODS. The numerous avian speech utterances of mynah (Gracula religiosa), canary (Serinus canaria) and gray parrot (Psittacus erithacus) were analyzed to decide the task of this research.

Research workers of the department of vertebrate zoology Saint-Petersburg State University gave us mynah and canary records from the A. S. Malthchevski collection of animal sounds. Gray parrot vocalizations were recorded on a SONI WM – D6C (frequency response 40 – 15000 Hz ± 3 dB) at home with DAK UEM-83R microphone and EAG MD-14N. We used IBM Pentium II with 16 bit sound card Creative Labs AVE 64 for digitizing the records with the frequency 44100 Hz.

Full records of birds consisted of species vocalizations, speech and instrumental imitations. To estimate the temporal and spectral characteristics of sounds we selected clear human speech imitations. The vowels were isolated based on similarity in dynamic spectrum, waveform and sound. If selected sound wasn't classified as vowel, it wasn't analyzed.

Spectral analysis was performed based on fast Fourier transformation and the data were weighed using a Hamming window (1024 filter). We analyzed the amplitude and the frequency for all spectral maxima with frequency lower than 7000 Hz. The values of the formants and amplitudes of the corresponding spectral components formed the base of data for each category of vowels. The means for all sounds of each category were distributed among the critical band of human hearing [5]. The amplitudes of spectral maxima relative to the amplitudes of the first spectral maximum were calculated. The dependence the medians of amplitudes of spectral maxima ratio and the frequency was constructed with correction on the human's audiogram. The reliability of distinctions of getting dependences for each vowel was examined by Kolmogorov-Smirnov criterion.

RESULTS AND DISCUSSION It was selected 63 canary's sounds [i], 37 mynah sounds [o], 35 mynah sounds [a], 31 mynah sounds [i] and 100 parrot's sounds [a], 100 parrot's sounds [o], 25 parrot's sounds [i].

Analysis of vowel-like sounds waveform and spectrum showed that the vowels imitating by birds in many cases aren't harmonic signals because the periods of waveform aren't repeated in strictly identical periods of time. The spectrum of these sounds haven't clear expressed harmonics divide by the equally frequency mean coincide with the frequency of the first harmonic (fundamental frequency). In this connection we didn't use the terms "fundamental frequency" and "formant" in this work. In the work about parrots' imitation signals [4], the authors didn't find the sample of waveform different from humans.

In this work there were vowel-like sounds, the acoustical characteristics of which were similar with harmonic (Fig. 1a, b). But a lot of sounds had the signs of amplitude modulation (Fig. 1c, d, e, f). There are some researches of imitations signals of budgerigars [6], mynah [1], treachers and cardinals [7], which give evidence of the possibility of this way of sound generation.

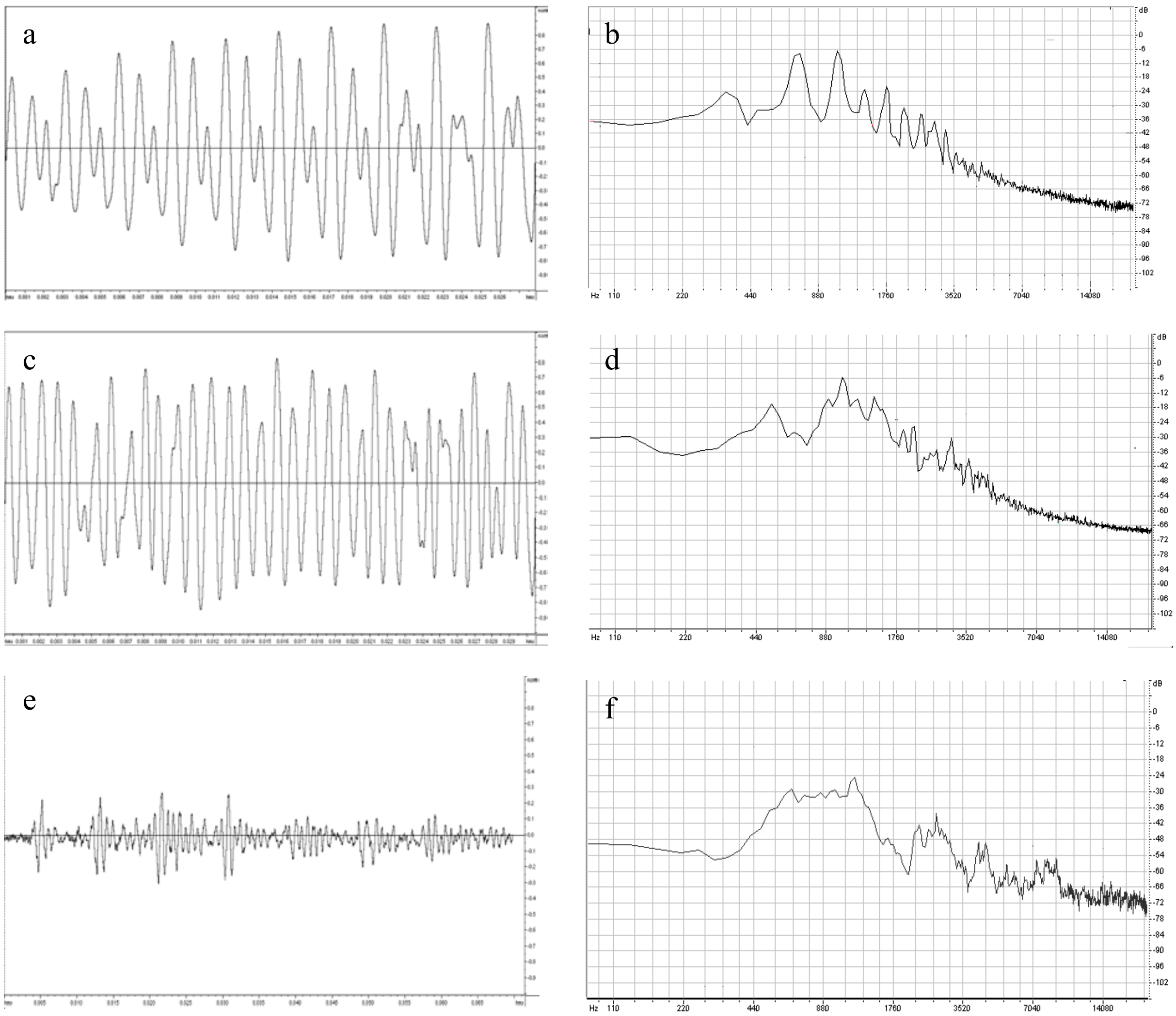

Fig. 1. The waveform and the spectrum of talking birds vowel-like sounds.
The waveform of vowel [a] for mynah (a, c) and parrot (e). The axes are time, ms and amplitude. The b, d, f are the spectrum of the vowel [a] for the same birds. The axes are frequency, Hz and amplitude, dB.

Analyze of spectrum showed that parrot's first spectral maxima means range across vowels from 260-396 Hz, compared to 270-308 Hz for the reported sample of gray parrot in the work of Pepperberg

[4]. It F1, F2 means across vowels were 855 Hz and 1386 Hz for [a], 599 Hz and 890 Hz for [o], 2253 Hz and 2901 Hz for [i]. In my work the means of two most expressed spectral maxima are 872 Hz and 1433 Hz for [a], 684 Hz and 1360 Hz for [o], 900 Hz and 2330 Hz for [i]. Mynah's means F1, F2 range across vowels from 645-899 Hz and 929-1960 Hz, compared to 380-840 Hz and 877-2660 Hz for the reported sample of mynah in the work of Klatt, Stefanski [1]. But the variability in our work is smaller.

Similarities and differences exist in absolute formant values for bird's vowels and human one. We did not find significant differences between the positions the first and second spectral maxima in two-formant space for [o], [a], [i]-like sounds of birds. And they were located in the regions of two-formant space that did not correspond to similar sounds of humans.

Recently, they were conduct series of works about the characteristics of vowel-like sounds in infants and the vowel of song speech [8, 9]. The authors had demonstrated that owing to high fundamental frequency of vocalizations the vowel-like sounds cannot be distinguished based on the characteristics of distribution of the first and second formants in two-formant space, but their amplitude ratios and frequency position of the spectral maximum differed from one another. Therefore the amplitudes relative to the amplitude of the spectral components corresponding to the first spectral maximum were compared within each critical band [5] for birds. The graph of dependence the medians of amplitudes of spectral maxima ratio and the frequency was constructed with correction on the human's audiogram. We compared the varied sounds for each bird and, one vowel within different birds. The results for the parrot's and mynah's vowel are shown in the Fig. 2.

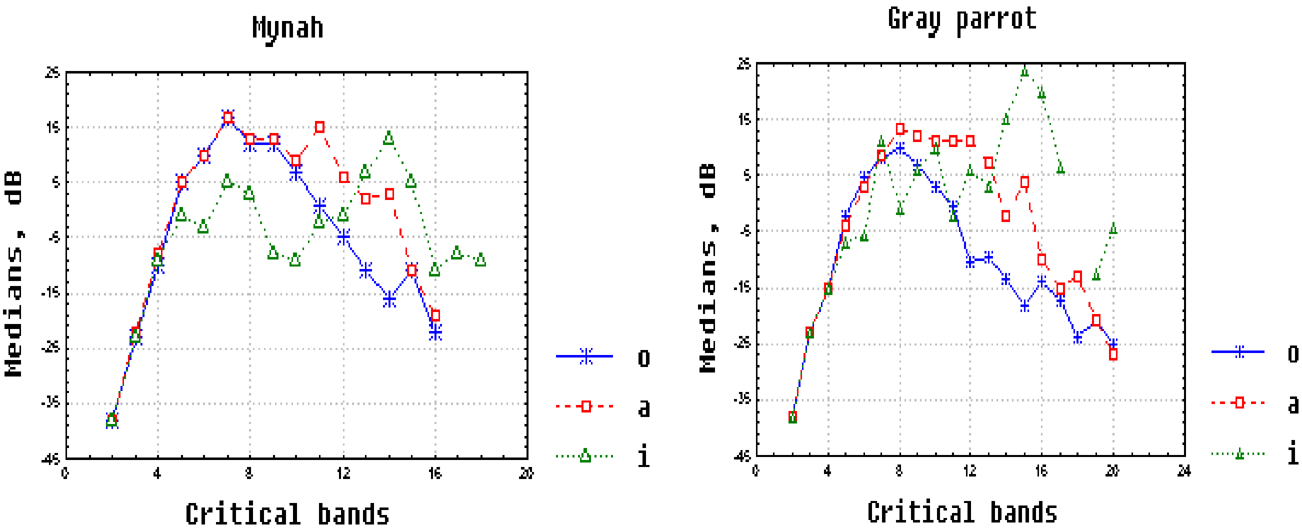

Fig. 2. The distributions of the relative amplitides of spectral maxima of vowel-like sounds by critical bands for mynah and parrot.

The ordinate shows the ratios of the amplitudes of spectral maxima to the amplitides of the Fo spectral maximums, dB. The abscissa shows the number of critical bands.

These figers demonstrate that the characters of distributions are different for each category of vowels. The sound [o] has one pronounced maximum range from 630 to 912 Hz. The sound [a] has the first maximum in the same range and the second maximim range from 1260 to 1720 Hz. And the sound [i] has two or three pronounced maxima: the first – in the range same within [a] and [o], the second – in the range1990- 2690 Hz. The sound's [i] maxima are in the 7 and 14 bands for mynah, 7 and 15 bands for parrot and 9 and 17 bands for canary. The reliability of distinctions of getting dependences for each vowel was examined by Kolmogorov-Smirnov criterion.

In this way the data of literature and the results of this work demonstrate the great variety of mean values spectral maxima for different birds. They show that it isn't possible to compare the frequency means of birds imitation's vowels with human'sounds characteristics in all cases and their places on two-formant space don't coincide with human's one [4]. Neither formant's means nor distance between them aren't the characteristic for distinguishing different categories of vowels imitating by birds, because their acoustical characteristics vary very widely. The amplitudes relative to the fundamental frequency significantly differed one category of vowel imitating by birds from another one. Because the similar distributions were received for vowel-like vocalizations of infants, the vowels of adult speech and the vowels of vocal speech [8, 9] it may assume that this criterion is common for different categories of vowels independently of the way of generation.

The work was supported by RFBR 02-06-80275

REFERENCES

1. Klatt D. H., Stefanski R. K. How does a mynah bird imitate human speech // J. Acoust. Soc. Amer. 1974. Vol. 55, № 4. pp. 822-832.
2. Hienz R. D., Sachs M. B. Sinnott J. M. Discrimination of steady-state vowels by blackbirds and pigeons // J. Acoust. Soc. Amer. 1981, Vol. 70, P. 699-706.
3. Dooling R. J., Brown S. D. Speech perception by budgerigars: spoken vowels // Percept. Psychophys. 1990, Vol. 47, № 6, pp. 568-574.
4. Patterson D. K., Pepperberg I. M. A comparative study of human and parrot phonation: acoustic and articulatory correlates of vowels // J. Acoust. Soc. Amer. 1994. Vol. 96, № 2. pp. 634-648.
5. Zwicker, E., Feldtkeller R. Das ohr als nachrichtenempfanger // S. Hirzel, Stuttgart, 1967.
6. Banta Lavenex P. Vocal production mechanisms in the budgerigar: The presence and implications of amplitude modulation // J. Acoust. Soc. Amer. 1999, Vol. 106, №1, pp. 491-505.
7. Goller F., Larsen O. N. A new mechanism of sound generation in songbirds // PNAS, 1997, Vol. 94, pp. 14787-14791.
8. 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, Vol. 368, pp. 471–473. (In Russian).
9. Andreeva N.G., Kulikov G.A., Samokishchuk A.P. Common features of the amplitude-frequency characteristics of vowels in different forms of speech // Acoustical Physics., Moscow, 2002, Vol. 48, № 5, pp. 620-622