

Vowel-Like Sound Structure in an African Grey Parrot (*Psittacus erithacus*) Vocal Production

L. Bottoni¹, S. Masin¹ and D. Lenti-Boero^{*2}

¹Dipartimento di Scienze dell'Ambiente e del Territorio, Università degli Studi Milano Bicocca, Piazza della Scienza, 1 20126, Milano, Italy

²Corso di Laurea in Psicologia, Università della Val D'Aosta, Strada Cappuccini 2a, Aosta, Italy

Abstract: In our study we taught a female African Grey 11 Italian words: vowel-like sounds were extracted from comprehensible words after critical listening, and pitch frequency (P_{kf}) was measured for the first three formants of each vowel. Similarly, formants from human vowels were isolated and measured. The analysis run on formant frequencies mean values of both samples revealed that human vowels could be separated on the basis of the first three formants. Comparison between each human vowel and its parrot counterpart revealed that *four* out of *five* parrot vowels could be considered statistically different from human ones regarding the first two formants, but comparison between F₂/F₁ and F₃/F₂ are not significant. Our results suggest that formant spaces do exist in the vocalic production of a *talking bird*. This leads to interesting conclusions about generalization skills involved in speech recognition, vowel parsing patterns and label production.

INTRODUCTION

Vocal learning in Vertebrates is evolutionarily rare: among mammals only cetaceans, elephants, some primates, including humans and a bat species have vocalization variable enough that learning can be experimentally established as contributing to the repertoire of adults [1-5]. Among birds, vocal learning is studied in few species of songbirds, hummingbirds and parrots [6-10].

The so-called talking-birds produce human speech and use it in a semantic fashion by means of different cognitive skills ranging from comprehension, sound discrimination and phonological awareness, as in humans [11-21].

We define imitation as an advanced behavior whereby an individual observes and replicates another's with an understanding of its context and meaning; while in mimicry whereby a given behavior is copied without any sign of context understanding [20, 22]. In mimicry, the bird memorizes a growing number of labels, more or less related to a semantic background, but experimental evidence on solitary sound plays and babble-like activities exhibited by African Greys seems to shed light upon a more complex scenario: parrots like to play with their vocal repertoire and show a great plasticity in label acquisition [18, 23]. The segmentation process is a form of combinatory behavior, in which the subject produces new words by parsing and recombining phonemes of extant words into new words [24, 25]. The segmentation of a communicative code by means of intentional recombination of existent labels is considered one of the milestones for the acquisition of a phonological awareness [26], that is, the consciousness that words are made up of a finite number of sounds that can be recombined into an almost infinite

number of patterns (limited only by the constraints of a given language). This process is part of the speech learning practices in pre-school children and is used as a discriminant feature to tell mimicry apart from imitation as far as vocal learning is concerned [16, 23-25]. Till now, there were no evidences that segmentation arises in any Vertebrate species, except humans [27]. Most "non talker" animals are not expected to show their own representation of phonemes since they are not stimulated by any vocal input. According to Pepperberg [25], however, at least one of her experimental parrots, Alex, engaged in intentional label parsing and recombination of subunits of words to assemble new labels. Intentional "label parsing" was separated by Pepperberg from what she calls "babble-luck" pattern [25]: in the first pattern, the animal seemed to intentionally choose phonemes in his repertoire to parse and glue together with the precise purpose to assembling a desired label. Quite on the contrary the "babble-luck" patterns involve a free recombination of phonemes, often as a part of solitary sound play, with the occasional and fortuitous creation of a new label as in first year old humans [12, 28, 29].

While the interest in parrot cognitive skills has risen with the research on the functional use of speech, evidence on differences between human and *talking birds* vocalic utterances has somehow remained obscure and ignored. Part of the reason could be found in the difficulties to anatomically describe bird sound production patterns in a satisfactory way, by means of a systematic functional research, since this means a multilevel research focused on parrots' phonetic and articulatory structures [19, 30-32].

The interest in mechanisms of production and comprehension of vowel-like sounds uttered by talking birds has recently increased because of the opportunities provided by comparative studies of *phonetic spaces* in human vowels and bird sounds. These shed light upon problems of speech production and perception by means of characteristics of the

*Address correspondence to this author at the Corso di Laurea in Psicologia, Università della Val D'Aosta, Strada Cappuccini 2a, Aosta, Italy;
E-mail: dlentiboero@univda.it

vocal string and on similarities and differences between bird and human *vocal products*, rather than on vocal anatomical structures: an original approach to the problem.

One of the most important and puzzling questions in this field has focused on the presence and position of formant regions in talking bird's vocal products. Lieberman [33] suggested that avian sounds are perceived as speech because they have energy at the same formant frequency of human vocalizations. Patterson and Pepperberg [34], however, hypothesized that parrots were unlikely to match human formant values, but could manipulate some acoustic feature such as the pattern of spacing between the first and second formant, in order to separate vowel categories. Some reasons exist for believing that such a manipulation would adequately separate vowels. The main reason is that the second formant is such an important correlate of human production that it is used to categorize vowels. Parrots and other talking birds might adopt a different strategy for perception and production, indeed African Grey Parrot vowels show differences between absolute formant values and those of humans [34]. Another comparative study on acoustical features of vowel sounds in different species of birds evidenced that in human vowels the most intense of the spectral maxima (called *global maxima*, GM) was usually the first, less frequently the second one, but in bird samples (cuckoo, *Cuculus canorus*; rook, *Corvus frugilegus*; raven, *Corvus corax*; jay, *Garrulus glandarius*; jackdaw, *Corvus monedula*; mallard, *Anser platyrhynchos*) the scattering of maximum intensity areas was much more frequent and an unstable GM resulted [35]. On the other hand, recent study suggests that the presence of formant regions in talking birds acoustic correlates of vowels could not be proven useful in vowel discrimination [36].

The aim of our work was to study Italian vowel-like sounds recorded from an African Grey Parrot (*Psittacus erithacus*) in order to understand if something similar to formant spaces existed in its utterances and if the phonetic space of each vowel-like sound differed from human ones, otherwise showing some inner homogeneity. In addition putting on an Italian-speaking parrot could be useful to clarify some talking bird acoustics because of the fewer and not overlapping Italian vocalic-sounds in comparison with English. Finally, we tried to investigate, at an empirical level, if some kind of label parsing and segmentation occurred in the process of label acquisition by our parrot.

METHODS

Subject and Housing

The subject of this study is a female adult African Grey Parrot named Teo, captive born in 1998 and hand-raised. The parrot was kept in a standard cage (90x50x80 cm) equipped with toys for psychophysical welfare, and exposed to human contact and speech. Teo was allowed out of the cage for free exploration and free interaction with its trainer, a fundamental issue in parrot training, except during sessions [12]. During her life Teo is always sharing a room with another captive parrot, for social enrichment. The first Teo's companion was a Jardine's Parrot (*Poicephalus gulielmi*), subsequently, it was substituted by a young male African Grey Parrot. Both companion parrots were chosen among naive (non trained) animals, in order to avoid any influence on Teo speech learning patterns. Teo was previously trained

in music discrimination [37], but never formally trained in speech production.

Stimuli and General Procedure

Eleven Italian words were chosen among the 100 most frequent words in the vocabulary of Italian children aged 8-17 months [38] particularly those referring to meaningful things for a parrot and to social games (i.e. cucù when the experimenter played hide-and-seek emerging from an hide) and were mostly disyllabic (Table 1). Words listed in Table 1 include 24 vowel sounds: representing that five out of seven Italian vocalic sound.

Table 1. Words Chosen for Training Sessions, Together with Situational Context

Italian Word	Meaning	Social Context
Teo	African Grey's name	As a greeting
Vito	Another Parrot in the lab	As a greeting
Pasta	Pieces of raw pasta	To obtain pasta to chew
Pane	Pieces of bread	To obtain food
Banana	Pieces of ripe banana	To obtain food
Uva	Berries of grape	To obtain food
Acqua	water	To obtain fruit juice
Cucù!	Exclamation used in children hide-and-seek play	Play.
Pronto?	Italian for "Hallo?"	Play: a telephone rings
Ciao	Typical Italian greeting	As a greeting
Basta!	Shut up!	To scold Teo

Words were introduced in groups of 2-3 at a time; test sessions were performed one per day at the most and occurred three times per week from 9 a.m. to 11 a.m., (i.e. the earlier daily peak activity) from January 2003 to June 2004. According to the Model/Rival approach [12, 18, 25, 28], we trained Teo to referential use of Italian words by means of three-way interactions between two human speakers and the parrot itself. While the bird watched, two female experimenters handled an object in which the bird has demonstrated interest, or simulated a social game like *hide-and-seek*. Then the trainer questioned the other experimenter (the parrot's *model* and *rival* for the trainer's attention) about the object, or invited her to join the game. The parrot was requested to name the object, or repeat the social label for the game: the trainer rewards correct responses with the object, only if the word was judged identifiable by both trainers, otherwise, the parrot was encouraged to "talk better". Trainers introduced pronunciation errors too to demonstrate the consequences of making mistakes.

New words were introduced only after the previous group was properly learnt as described below.

Construction of Database

Two hundred and fifty-five sessions (intended as the whole human and parrot vocal interactions) were audio recorded with a Power Macintosh G4 and an *Mbox* Digidesign

(Inc.), for a total of 9600 min. Sounds were recorded by means of a Sennheiser K6 microphone with a directional module Sennheiser ME 67. Sounds were sampled at 48000 Hz and the sample size was 16 bit. Audio files were processed with *ProTools* sound analysis software.

In order to isolate and select vowels to be analyzed we only used those words, which had been clearly understood by both two judges among the whole recorded database, as described by Patel and Daniele [39]. By means of this procedure we isolated, 82 utterances composed by the 11 words taught to the parrot. Among those 82 words all vowel-like sounds were isolated and arranged in sound files for further selection: a different couple of judges [39] listened to them and were able to clearly tell apart recognizable vowel-like sounds. At the end of this second screening we obtained a total amount of 212 vowel-like sounds clearly recognizable.

Data Analysis

Among the 212 vowel-like sounds selected by judges we measured and analyzed only those lasting 100ms or more (195 vowel-like sounds), which is the minimum length required for a vocal bout to be correctly understood [40]. This procedure resulted in a total of the following five Italian vowels: 65 [A]=A almost like in the English word *such*, 25 [æ]=E like in *elephant*, 25 [I]=I like in *fish*, 40 [o]=O as in *old* and 40 [:u]=U almost like in *moon*. Sound comparison between English-trained parrot (Alex) and Italian-trained parrot (Teo) is not always possible because of phonetic differences in some vowels. Similarly, vowels from trainers' words were recorded and we selected 70 samples for human A[A], 25 samples for human E[æ] and I[I] vowels, 35 samples for O[o] vowel and 40 samples for human U[:u] vowel.

Spectrograms for each vowel were obtained from waveforms, by means of Canary 1.2.4 [41] software: a Hamming window, with a FFT size of 2048 points was used. We measured frequencies of first, second and third formant and mean frequency values for each formant were obtained for the five vowel. Then we calculated the ratio between second and first formant (R_1) and between third and second formant (R_2).

Statistical analysis was performed by means of SPSS.

RESULTS

Clues of Segmentation in Phonemes Shaping Patterns

We recorded several attempts of producing words by means of segmentation processes of Teo. Three main cases are described below.

Basta! (*shut up!*) Teo autonomously attempted to produce this word (which experimenters used to scold the parrot when she was too noisy) only two months after the beginning of the experiment (March 2003). *Basta!* Was first uttered as an alternate version of a previously learnt similar word: *Pasta*, but soon the parrot adjusted the prosodic contour of the new word (stress on the first vowel) which sounded more similar to an exclamation.

Acqua (*water*). This word was introduced after pasta, and referred to fruit juice administered to the parrot with a metal spoon. During the training for *Acqua*, Teo started pronouncing *-sta* (the second syllable of *pasta*), isolated from the first

syllable *pa-*, then shifted to *-qua*. Afterwards she tried to adjust the prosodic contour of her version of *acqua*, in two different ways: she prolonged the final vowel, thus producing a first version, which sounded like *-quaa*, after she introduced the first syllable and filled the gap between the first and the second syllable with a pause: the result sounded like *a/pause/quaa*.

Banana. This word was one of the most difficult for Teo to learn, being trisyllabic. She used the first syllable of *Basta* as a starting nucleus and elongated the final vowel, to render the prosodic rhythm. The first version of *Banana* sounded like *-baaa*, and was followed in a couple of month by two more sophisticated versions: the first one resulted with a duplication of the first syllable, together with the final vowel elongation and sounded like *-babaaa*.

The second version was somehow more complex, being composed by the previously learnt syllable *-ba* and by a new dyad *-na*, which Teo elongated in the same way she did with the syllable *-ba*, the final result sounding like *banaaa*. These two versions persisted for more than six months and were never completely replaced by the correct word, coexisting with *banana* in every session.

Human and Parrot Vowels Comparison

For all isolated vowels, we computed narrow band spectrograms to look for F_0 well evidenced in Figs. (1-5).

Then we measured pitch frequencies (P_{kf}) for the main three formant regions on wide band spectrograms (Figs. 1-5).

As a first step, analysis of variance on first three formant frequency values was run separately on both samples (human and parrot vowels).

ANOVA on human samples revealed that highly significant differences did exist among all three formants of the five vowels; in bird's vowel samples ANOVA was highly significant for each formant; but Tukey's test (confidence interval 95%) showed that only F_2 and F_3 separate samples into 5 homogeneous subsets (results are reported in Appendix).

Then, we run an ANOVA test for each vowel on both human and avian samples (Table 2).

This analysis showed that four vowels out of five can be distinguished on the basis of the first two formant mean values between human and avian samples (the only non significant sample was I[I]). Two opposite figures are portrayed in E[æ] and O[o] vowel human and parrot comparison charts: as for E[æ] vowel (Fig. 6) the differences are significant for all three formant mean values, and the parrot frequencies are higher than human ones.

For the O[o] vowel (Fig. 7) the differences are highly significant for the first and the second formant, but parrot frequencies are lower than human ones. In Table 3 are reported human and parrot formant ratios, comparison between R_1 and R_2 parrot values vs R_1 and R_2 human values run separately by means of a T test were non significant.

Finally, a dispersion chart of African Grey Parrot vowel formant spaces was drawn by means of F_2 and F_3 formant values (Fig. 8) because of their diagnostic importance in vowels discrimination.

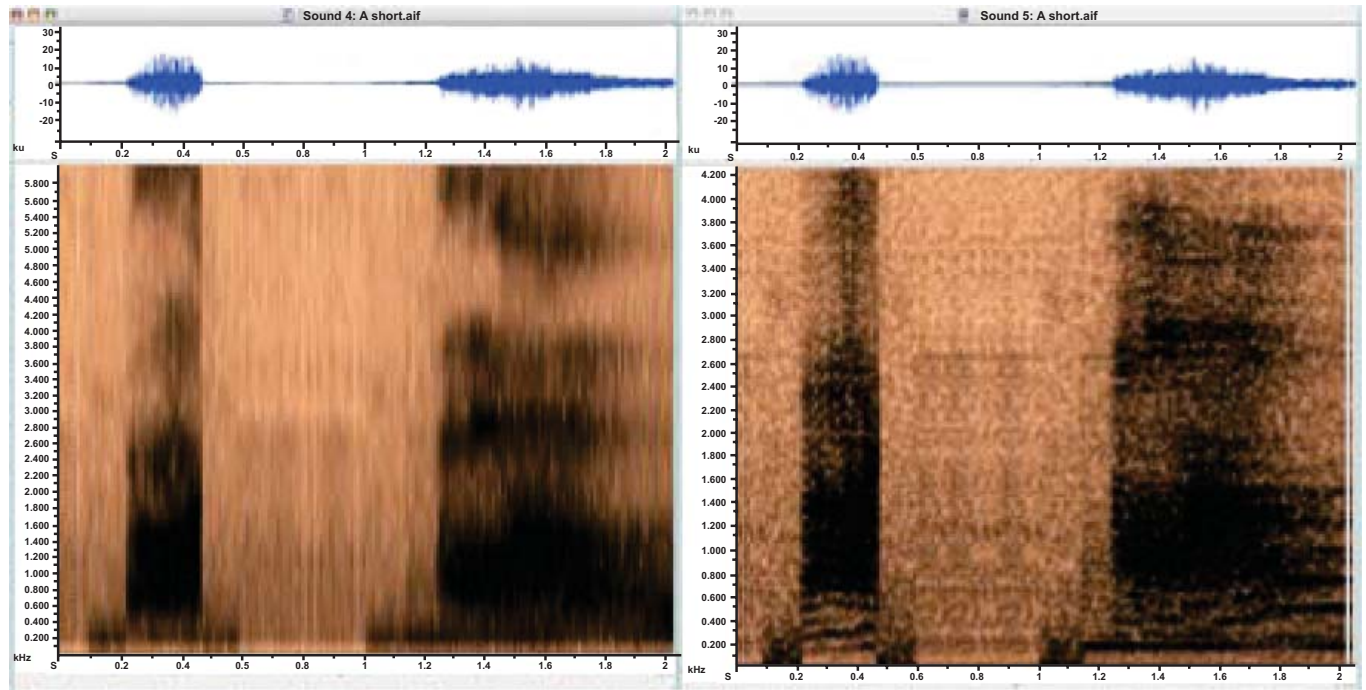


Fig. (1). Waveform and spectrogram of parrot A [ʌ] vowel. Spectrograms were traced by means Raven Pro 1.3 software. Narrow band parameters: window type Hamming, window size 1500 samples, 3 dB Filter Bandwidth 46 Hz, DFT size 16384. Wide band parameters: window type Hamming, window size 230 samples, 3 dB Filter Bandwidth 300 Hz, DFT size 1024.

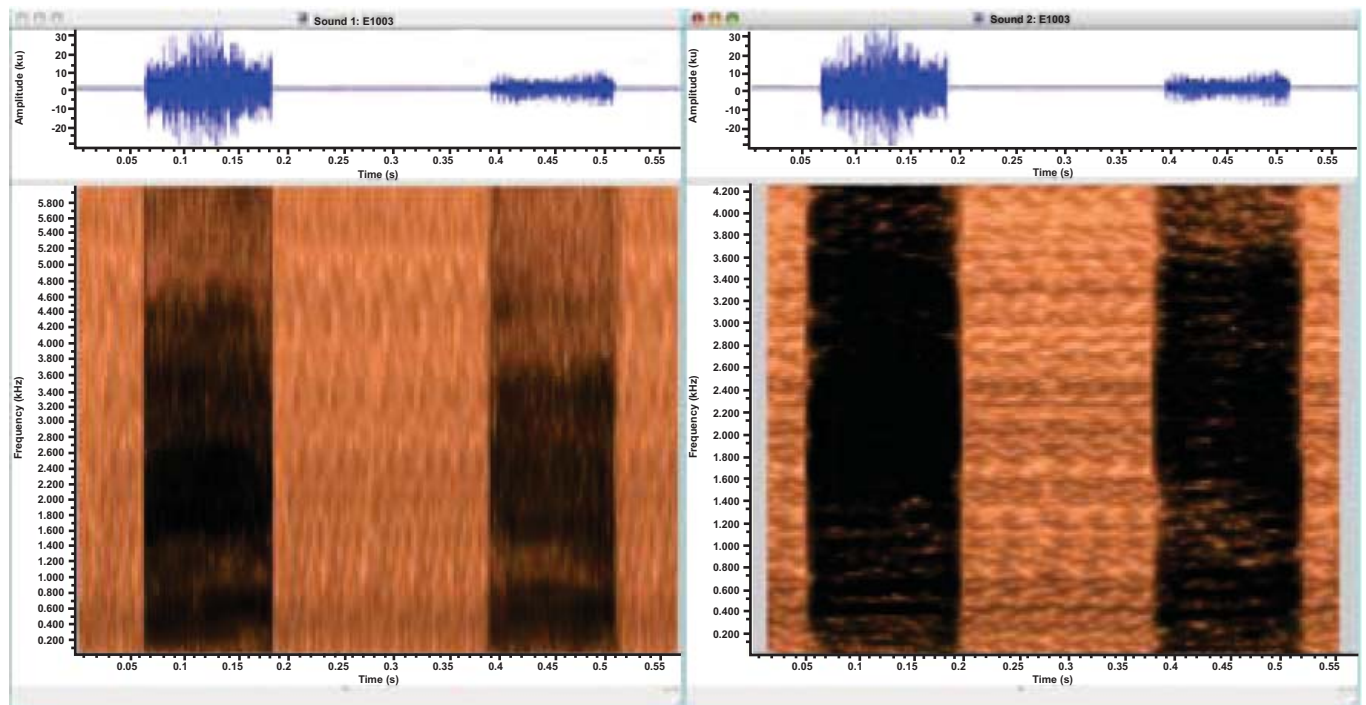


Fig. (2). Waveform and spectrogram of parrot E [æ] vowel. Parameters as in Fig. (1).

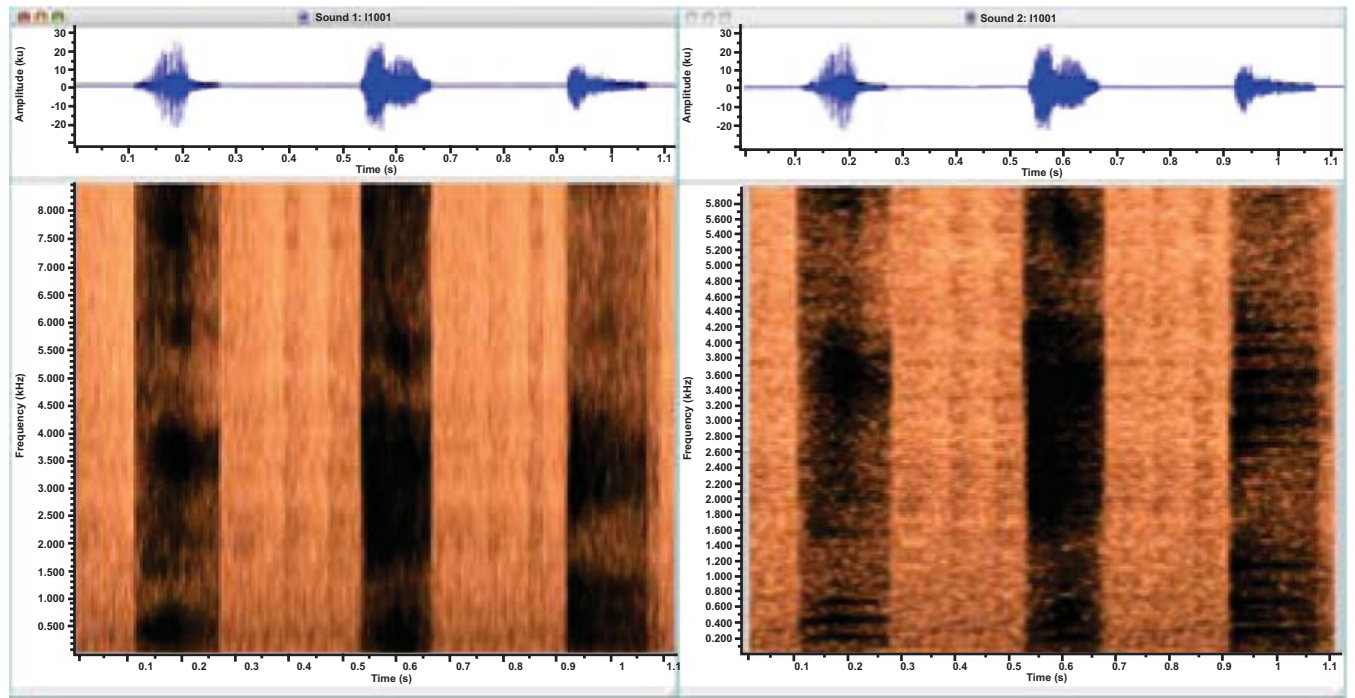


Fig. (3). Waveform and spectrogram of parrot I [I] vowel. Parameters as in Fig. (1).

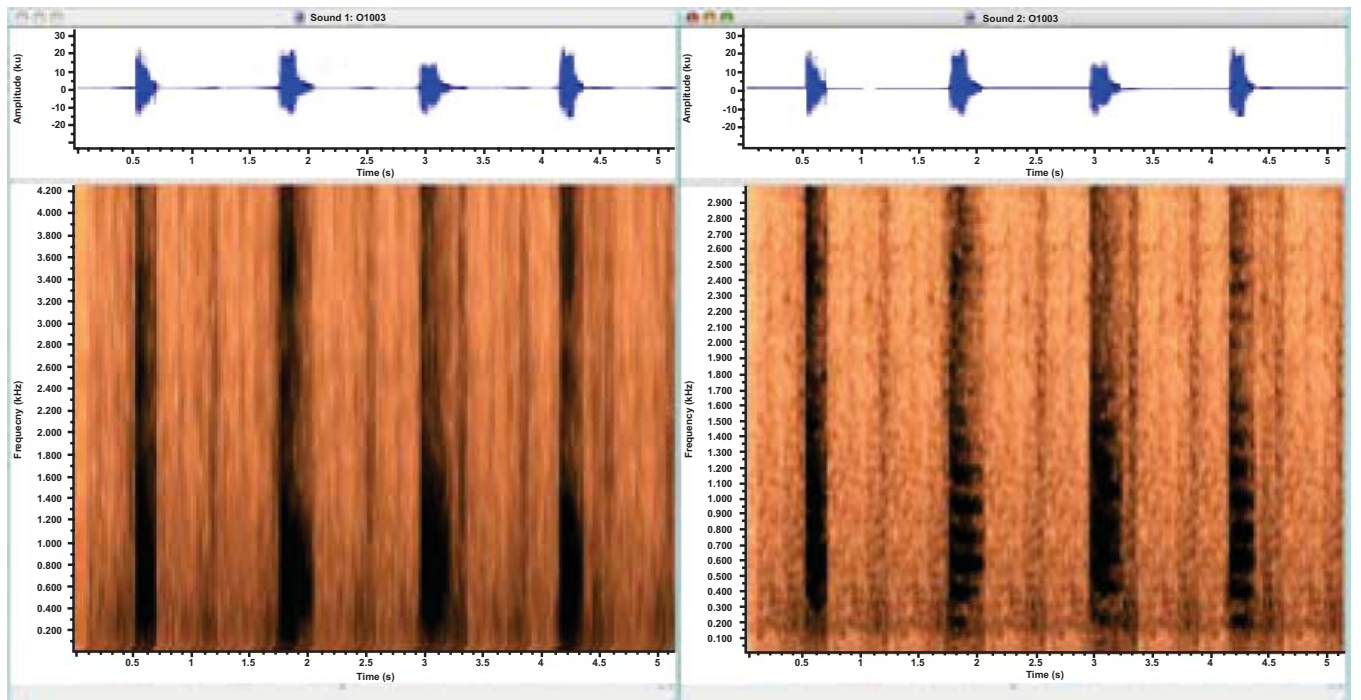


Fig. (4). Waveform and spectrogram of parrot O [o] vowel. Parameters as in Fig. (1).

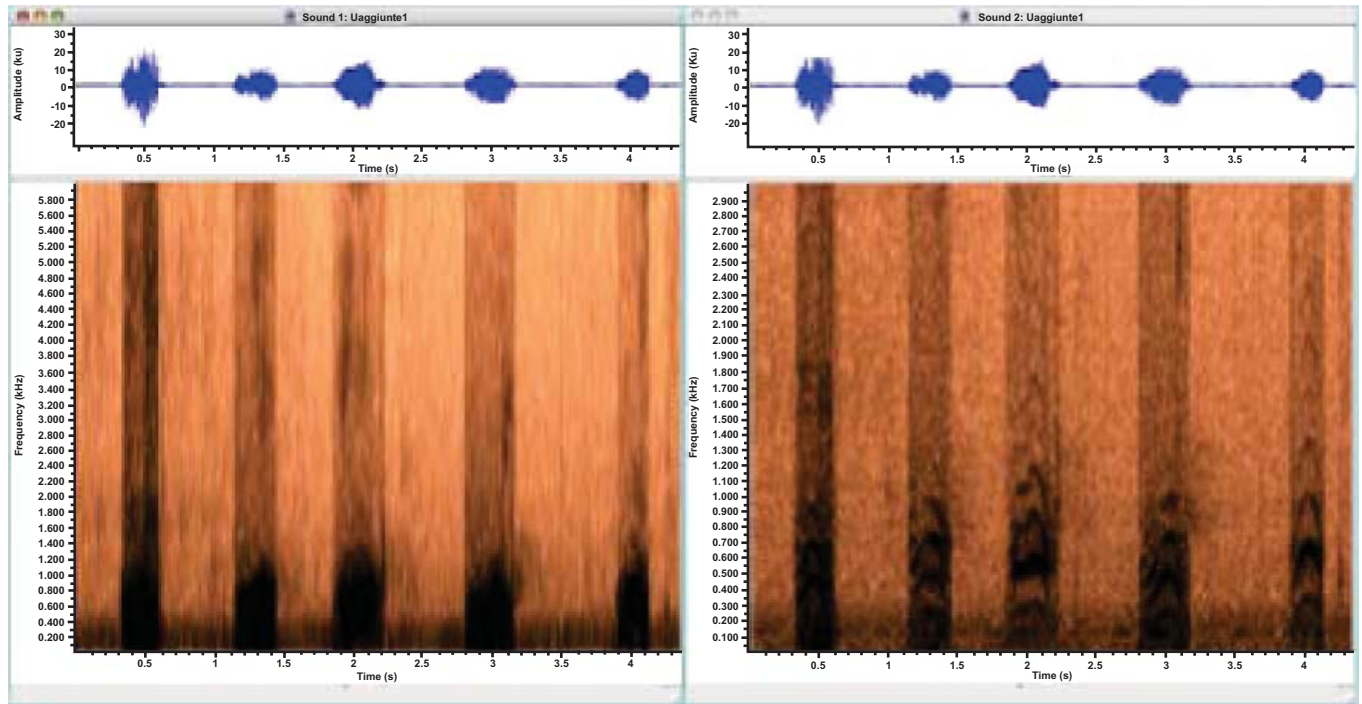


Fig. (5). Waveform and spectrogram of parrot U [:u] vowel. Parameters as in Fig. (1).

Table 2. Mean ± SD and p Values of ANOVA Test on the First Three Formants of African Grey* vs Human Vowels°

Vowel	F ₁	F ₂	F ₃
A	*X=637.5+-194.0 °X=806.9+-146.3 F=33.12; p<0.001	*X=1.052.6+283.8 °X=1246.5+-304.6 F=14.58; p<0.001	*X=2.018.1+642.5 °X=1974.8+-504.3 F=0.19; p=0.66
E	*X=671.2+119.9 °X=305.6+242.7 F=28.52; p<0.001	*X=2105.6+-451.2 °X=686.3+-343.7 F=156.55; p<0.001	*X=3122.8+-761.3 °X=1087.5+-595.0 F=110.93; p<0.001
I	*X=463.1+170.9 °X=496.3+221.5 F=0.35; p=0.55	*X=2341.9+-994.6 °X=944.0+-310.4 F=44.99; p<0.001	*X=3734.1+1060.5 °X=1929.1+-710.3 F=49.99; p<0.001
O	*X=334.4+96.1 °X=507.8+-129.0 F=44.97; p<0.001	*X=557.9+-141.6 °X=806.2+-192.2 F=41.90; p<0.001	*X=1512.0+48.7 °X=1648.5+702.3 F=1.00; p=0.321
U	*X=424.8+108.5 °X=483.3+-127.7 F=4.92; p=0.03	*X=738.3+198.2 °X=838.6+-211.2 F=4.85; p=0.03	*X=+1057.1+266.6 °X=1412.7+489.7 F=16.36; p<0.001

As the chart shows, a certain degree of overlapping is therefore present in parrot formant areas.

The groups of O[o] and E[æ] vowels seem to be quite separate on the basis of second and third formants values, while A[A] and U[:u] vowels are considerably overlapping.

A group of 5 out of 25 samples of the vowel I[i] segregate above 2550 Hz.

Table 3. Formant Mean Ratio: R₁ = F₂/F₁, R₂ = F₃/F₂ of Both Parrot and Human

Vowel	R ₁ parrot	R ₁ human	R ₂ parrot	R ₂ human
A	1,5	1,5	2,8	3,3
E	3,1	3,5	4,5	4,7
I	6,8	8,5	14,4	11,1
O	1,9	1,5	4,5	4,2
U	2,9	2,3	4,9	5

DISCUSSION

Existing data on the production of vowel-like sounds by a number of so-called *talking birds* paint a complex reality. No evidences of formant existence were found by Uplisova [36] in o, a and i vowel-like African Grey Parrot sounds. Quite on the contrary, Patterson and Pepperberg [34], evidenced both F₁ and F₂ frequencies, although the bird range of values was not as great as the human range and was especially lacking in low frequency values compared to humans. Other researches demonstrating that humans use dynamic information to separate vowels, suggest that formant data is only one of several cues used for vowel perception [42]; Gentilucci and Cattaneo [43] advocated the position of lips and mouth as visual cues for comprehension in human communication patterns. Nevertheless, in the present study, listeners performing compre-

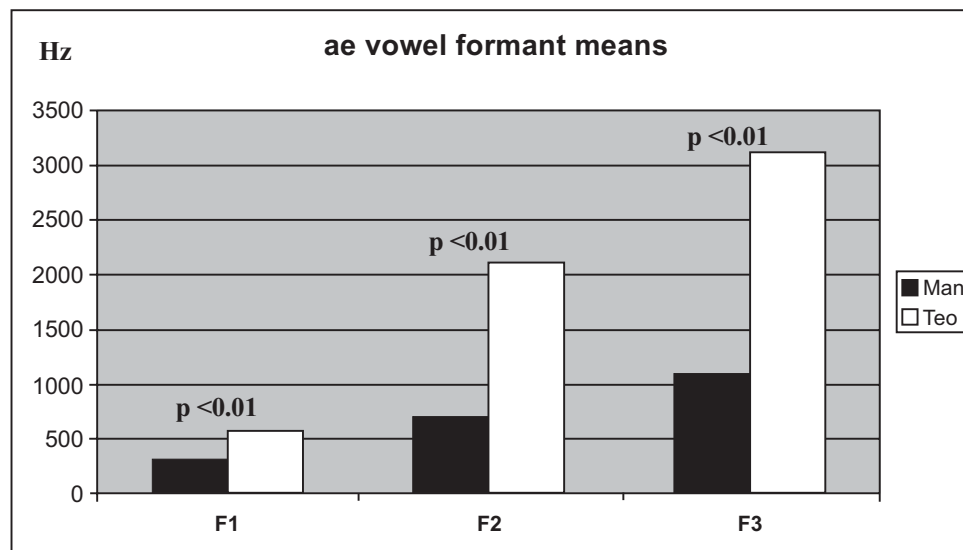


Fig. (6). Comparison histogram for ANOVA analysis on F₁, F₂, F₃ of the human and parrot E[æ] vowel.

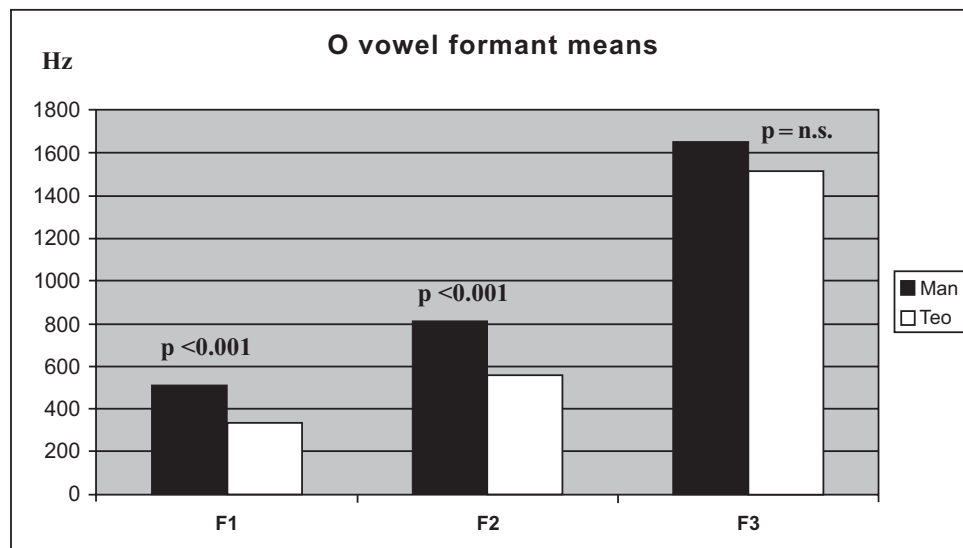


Fig. (7). Comparison histogram for ANOVA analysis on F₁, F₂, F₃ of the human and parrot O[o] vowel.

hension tests perceived bird's vocalizations as analogous to human speech. Peterson and Barney [44] found some formant frequency variability from one human speaker to another and in their dispersion charts a substantial degree of overlapping in the formant frequency patterns among adjacent vowels was evident, and latter studies confirmed it [45, 46].

Our data highlight the differences in F₁, F₂ and F₃, and formant spaces could be appreciated in spectrogram of our parrot vowel-like sounds. Variance analysis of parrot and human samples stressed that for bird samples F₂ and F₃ seem more important for vowel discrimination.

Comparison of human and parrot samples showed that four vowels out of five can be distinguished on the basis of the first two formants mean values, the only controversial vowel being the I[I]. This because 25% of the samples of the word "Vito", the only word containing the I[I] sonorous sound, was substituted by an avian whistle (in the Italian language I[I] in "ciao" is unexpressed). The difference in articulatory and phonetic apparatuses can impair the ability

of a parrot to produce certain sounds: thus the need to operate a substitution of certain sound categories with something else. Trained parrots seem to be able to overcome this problem if properly corrected, but at the beginning of training we were not aware of this problem, so Teo was never scolded when she produced whistled I[I] sounds instead of true vowels; nevertheless in 75% of the cases Teo pronounced a proper sonorous Italian I[I].

The present study, performed in Italian, confirms previous evidence found by Patterson and Pepperberg [34] showing that significant differences exist in parrot English vowel-like sounds, in comparison with human ones. The differences in F₁ values between human and parrot vocalic sounds (but not among parrot's sole samples) strengthen the hypothesis of a preferential use of F₂ and F₃ by parrots to differentiate vowels. An interesting add that can be reported from Patterson and Pepperberg [34] is the difference between F₁ and F₂ mean values of our African Grey U vowel and the only English phonetically comparable vowel recorded by Pepperberg's parrot which shows higher values.

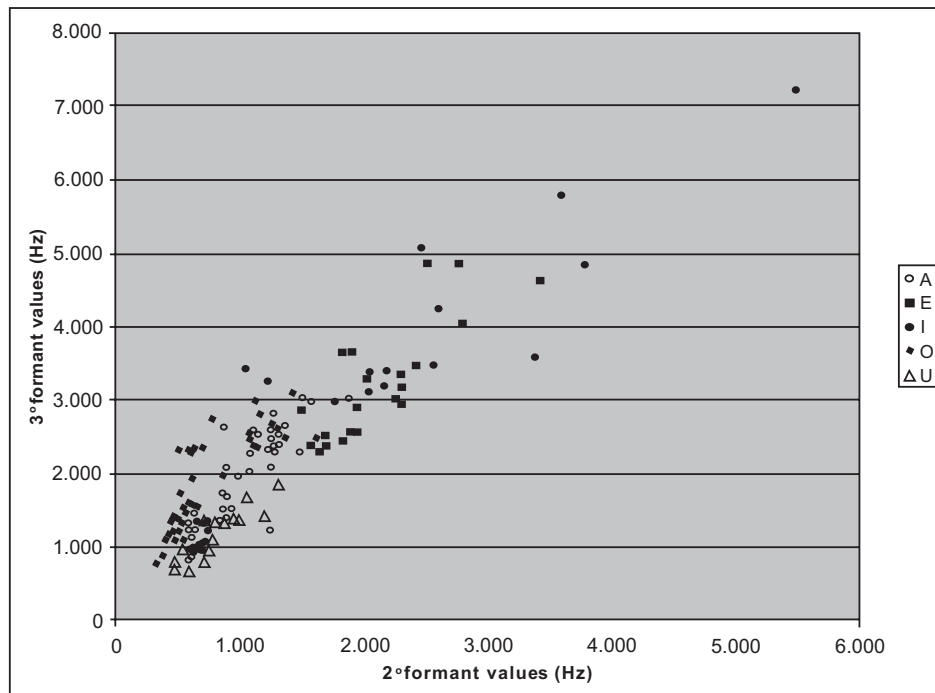


Fig. (8). Dispersion chart of the second and third formant of the five parrot vowels.

Definitely, parrot speech seems to be different enough in acoustical features to be separated by variance analysis but, at the same time, similar enough to be understood by a human listener, both as complete words and as isolated vowels. That is confirmed by the comparison of parrot and human R_1 and R_2 values, which doesn't show any significant difference. This evidence sheds some light upon the complexity of the process of generalization required to understand the speech of a talking bird. Strange [42] stressed that in human speech there is sufficient information within single syllables to allow the listener to identify vowels, even in different consonant context, this could be true even in the case of bird's speech.

Like speech, birdsong acquisition has distinct sensory and sensorimotor phases: in some birds species as in humans, auditory feedback is necessary for mature vocal performance and this should be true even in the case of a parrot producing human speech. The principal sensory input guiding learning is auditory, but in birds as well as in humans, visual input and social interactions can greatly influence vocal learning [47]. The process of self-listening is a well-known pattern in birds' vocal learning, as well as social context [11, 48, 49]. On the basis of these considerations, we could suppose that Teo must somehow *chisel* her utterances by means of an auditory feedback process, which could allow her to compare her *version* of a word with the *correct* one heard from human trainers and introduce adjustments, if needed. In the latter case, however, the process of generalization required to evaluate the best rendering of human speech seems more complex. As a matter of fact, in the allo-specific free imitation by birds, the imitated sound is often deeply modified to adjust it to the general setting of the whole song, thus personalizing the single individual performance, a process well studied in Cowbirds (*Molothrus ater*) and Starlings (*Sturnus vulgaris*) [50, 51] and known in African Greys [9].

In our case indeed, a strong pressure on a *correct rendering* was applied and *innovation* was neither allowed nor rewarded. This process of *correct* word production could be difficult for a parrot, even more challenging than sound production/imitation itself, an easier process, lacking the restraints that drive the animal to produce one and only one understandable version of a word.

Despite those training constraints, qualitative analysis of Teo's learning process suggests multidimensional cognitive abilities: 1) discrimination between "p/b" for "pasta/basta" suggests the existence of categorical perception of phones, similarly to Pepperberg's Grey parrot [34, 32] and rhesus macaques, dogs, chinchillas, quails and, of course, humans [21, 52-55]; 2) the ability to reuse learnt strings of phones in producing new words (ba-sta/pa-sta; ba-nana); 3) the ability to isolate, then shift and adjust the prosodic contour of syllables in a kind of "acoustic morphing" (sta/quaa) [56].

One final consideration on the pressure for speech correct rendering may be added: at the end of the experiment the animal was not subjected to further training and was housed with another African Grey (a young male, naive to speech, for breeding purposes). This led to interesting consequences in the evolution of Teo repertoire: one year after from the end of training, the animal continued to utter all the words learned during the experiment, using them to interact with the other parrot, but most of the words were deeply modified and lots of new versions of the words arose. As an example, words like *Banana* experienced a process of decline, boosted by the proliferation of several new combinations like *Banail/Banee/Banasta*, arisen from a continuous process of parsing-and-gluing of the extant repertoire. Interestingly, Ugo, the new male African Grey, took immediate part in this sound play, producing some brand-new combinations. We could hypothesize that - without any further input by trainers - all that was learnt could be used by the animal to arrange a

duet or a courtship parade, in order to establish a pair bond with the other African Grey, according to its evolutionary context [9]. In absence of a semantic pressure, which selects one version of a word, a strong tendency towards innovation soon arose.

An analogous pattern was observed about the acquisition of a musical code by the same parrot during a previous experiment: trained to use the Temperate Scale to answer to couple of musical notes played by a keyboard with appropriate sequences, at the end of the experiment she continued to

use learnt notes by assembling original sequences [37].

In conclusion, our contribution might be a first step to evaluate how vocal communication (and music too) are at least in part shaped by evolutionary selective pressures [57].

ACKNOWLEDGEMENTS

This research has been supported by grants from the 2001 Support Program for scientific research of National relevance - MIUR (Ministero dell’Istruzione, dell’Università e della Ricerca).

APPENDIX

Human F1 Vowel Comparison

ANOVA

F1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6226055,138	4	1556513,784	56,257	,000
Within Groups	5284526,427	191	27667,678		
Total	11510581,564	195			

Multiple Comparisons

Dependent Variable: F1

Tukey HSD

(I) FACTOR	(J) FACTOR	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	501,28	38,76	,000	395,57	607,00
	3	310,60	38,76	,000	204,88	416,31
	4	299,08	34,43	,000	205,14	393,01
	5	323,63	32,71	,000	234,40	412,86
2	1	-501,28	38,76	,000	-607,00	-395,57
	3	-190,69	47,05	,000	-319,02	-62,35
	4	-202,21	43,56	,000	-321,02	-83,39
	5	-177,65	42,21	,000	-292,79	-62,52
3	1	-310,60	38,76	,000	-416,31	-204,88
	2	190,69	47,05	,000	62,35	319,02
	4	-11,52	43,56	,999	-130,33	107,29
	5	13,03	42,21	,998	-102,10	128,17
4	1	-299,08	34,43	,000	-393,01	-205,14
	2	202,21	43,56	,000	83,39	321,02
	3	11,52	43,56	,999	-107,29	130,33
	5	24,55	38,28	,968	-79,87	128,97
5	1	-323,63	32,71	,000	-412,86	-234,40
	2	177,65	42,21	,000	62,52	292,79
	3	-13,03	42,21	,998	-128,17	102,10
	4	-24,55	38,28	,968	-128,97	79,87

The mean difference is significant at the .05 level.

F1
Tukey HSD

	N	Subset for alpha = .05		
FACTOR		1	2	3
2	25	305,63		
5	41		483,29	
3	25		496,32	
4	35		507,84	
1	70			806,92
Sig.		1,000	,974	1,000

Means for groups in homogeneous subsets are displayed.

Human F2 Vowel Comparison

ANOVA
F2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	9027177,136	4	2256794,284	29,542	,000
Within Groups	14591244,919	191	76393,952		
Total	23618422,055	195			

Multiple Comparisons
Dependent Variable: F2
Tukey HSD

(I) FACTOR	(J) FACTOR	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	560,24	64,40	,000	384,58	735,91
	3	302,50	64,40	,000	126,84	478,17
	4	440,30	57,22	,000	284,22	596,38
	5	407,92	54,36	,000	259,64	556,19
2	1	-560,24	64,40	,000	-735,91	-384,58
	3	-257,74	78,18	,009	-470,99	-44,49
	4	-119,95	72,38	,461	-317,37	77,48
	5	-152,33	70,14	,190	-343,64	38,99
3	1	-302,50	64,40	,000	-478,17	-126,84
	2	257,74	78,18	,009	44,49	470,99
	4	137,79	72,38	,315	-59,63	335,22
	5	105,41	70,14	,560	-85,90	296,73
4	1	-440,30	57,22	,000	-596,38	-284,22
	2	119,95	72,38	,461	-77,48	317,37
	3	-137,79	72,38	,315	-335,22	59,63
	5	-32,38	63,61	,987	-205,89	141,13
5	1	-407,92	54,36	,000	-556,19	-259,64
	2	152,33	70,14	,190	-38,99	343,64
	3	-105,41	70,14	,560	-296,73	85,90
	4	32,38	63,61	,987	-141,13	205,89

The mean difference is significant at the .05 level.

F2

Tukey HSD

	N	Subset for alpha = .05		
FACTOR		1	2	3
2	25	686,26		
4	35	806,21	806,21	
5	41	838,59	838,59	
3	25		944,00	
1	70			1246,50
Sig.		,154	,240	1,000

Means for groups in homogeneous subsets are displayed.

Human F3 Vowel Comparison

ANOVA

F3

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	19374666,879	4	4843666,720	14,341	,000
Within Groups	64511636,717	191	337757,260		
Total	83886303,596	195			

Multiple Comparisons

Dependent Variable: F3

Tukey HSD

(I) FACTOR	(J) FACTOR	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	887,32	135,41	,000	517,95	1256,68
	3	45,72	135,41	,997	-323,64	415,08
	4	326,30	120,31	,052	-1,89	654,49
	5	562,06	114,29	,000	250,29	873,83
2	1	-887,32	135,41	,000	-1256,68	-517,95
	3	-841,60	164,38	,000	-1289,99	-393,21
	4	-561,02	152,19	,002	-976,15	-145,89
	5	-325,26	147,47	,178	-727,53	77,01
3	1	-45,72	135,41	,997	-415,08	323,64
	2	841,60	164,38	,000	393,21	1289,99
	4	280,58	152,19	,348	-134,55	695,71
	5	516,34	147,47	,004	114,06	918,61
4	1	-326,30	120,31	,052	-654,49	1,89
	2	561,02	152,19	,002	145,89	976,15
	3	-280,58	152,19	,348	-695,71	134,55
	5	235,76	133,75	,396	-129,07	600,59
5	1	-562,06	114,29	,000	-873,83	-250,29
	2	325,26	147,47	,178	-77,01	727,53
	3	-516,34	147,47	,004	-918,61	-114,06
	4	-235,76	133,75	,396	-600,59	129,07

The mean difference is significant at the .05 level.

F3

Tukey HSD

FACTOR	N	Subset for Alpha = .05		
		1	2	3
2	25	1087,48		
5	41	1412,74	1412,74	
4	35		1648,50	1648,50
3	25			1929,08
1	70			1974,80
Sig.		,143	,452	,141

Means for groups in homogeneous subsets are displayed.

Parrot F1 Vowel Comparison

ANOVA

F1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2719793,568	4	679948,392	32,938	,000
Within Groups	3942854,024	191	20643,215		
Total	6662647,592	195			

Multiple Comparisons

Dependent Variable: F1

Tukey HSD

(I) FACTOR	(J) FACTOR	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	72,18760	33,81296	,205	-20,04668	164,42188
	3	174,37996	33,81296	,000	82,14568	266,61424
	4	303,08691	28,65451	,000	224,92374	381,25007
	5	212,69540	28,87331	,000	133,93539	291,45541
2	1	-72,18760	33,81296	,205	-164,42188	20,04668
	3	102,19236	40,63812	,087	-8,65945	213,04417
	4	230,89931	36,45849	,000	131,44861	330,35001
	5	140,50781	36,63071	,001	40,58733	240,42828
3	1	-174,37996	33,81296	,000	-266,61424	-82,14568
	2	-102,19236	40,63812	,087	-213,04417	8,65945
	4	128,70695	36,45849	,004	29,25625	228,15765
	5	38,31545	36,63071	,834	-61,60503	138,23592
4	1	-303,08691	28,65451	,000	-381,25007	-224,92374
	2	-230,89931	36,45849	,000	-330,35001	-131,44861
	3	-128,70695	36,45849	,004	-228,15765	-29,25625
	5	-90,39150	31,93076	,037	-177,49155	-3,29146
5	1	-212,69540	28,87331	,000	-291,45541	-133,93539
	2	-140,50781	36,63071	,001	-240,42828	-40,58733
	3	-38,31545	36,63071	,834	-138,23592	61,60503
	4	90,39150	31,93076	,037	3,29146	177,49155

The mean difference is significant at the .05 level.

F1
Tukey HSD

FACTOR	N	Subset for alpha = .05		
		1	2	3
4	41	334,41317		
5	40	424,80468	424,80468	
3	25		463,12012	
2	25			565,31248
1	65			637,50008
Sig.		,068	,802	,225

Means for groups in homogeneous subsets are displayed.

Parrot F2 Vowel Comparison

ANOVA
F2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	79892074,871	4	19973018,718	105,625	,000
Within Groups	36116792,389	191	189093,154		
Total	116008867,260	195			

Multiple Comparisons
Dependent Variable: F2
Tukey HSD

(I) FATTORE	(J) FATTORE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-1053,02	102,34	,000	-1332,17	-773,87
	3	-1289,26	102,34	,000	-1568,41	-1010,11
	4	494,68	86,72	,000	258,12	731,25
	5	314,36	87,39	,003	75,99	552,73
2	1	1053,02	102,34	,000	773,87	1332,17
	3	-236,24	122,99	,306	-571,74	99,26
	4	1547,70	110,34	,000	1246,71	1848,70
	5	1367,38	110,87	,000	1064,96	1669,79
3	1	1289,26	102,34	,000	1010,11	1568,41
	2	236,24	122,99	,306	-99,26	571,74
	4	1783,94	110,34	,000	1482,95	2084,94
	5	1603,62	110,87	,000	1301,20	1906,03
4	1	-494,68	86,72	,000	-731,25	-258,12
	2	-1547,70	110,34	,000	-1848,70	-1246,71
	3	-1783,94	110,34	,000	-2084,94	-1482,95
	5	-180,32	96,64	,336	-443,94	83,29
5	1	-314,36	87,39	,003	-552,73	-75,99
	2	-1367,38	110,87	,000	-1669,79	-1064,96
	3	-1603,62	110,87	,000	-1906,03	-1301,20
	4	180,32	96,64	,336	-83,29	443,94

The mean difference is significant at the .05 level.

F2
Tukey HSD

FATTORE	N	Subset for alpha = .05		
		1	2	3
4	41	557,94		
5	40	738,26		
1	65		1052,62	
2	25			2105,64
3	25			2341,88
Sig.		,420	1,000	,159

Means for groups in homogeneous subsets are displayed.

Parrot F3 Vowel Comparison

ANOVA
F3

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	150920259,247	4	37730064,812	90,702	,000
Within Groups	79452097,082	191	415979,566		
Total	230372356,329	195			

Multiple Comparisons
Dependent Variable: F3
Tukey HSD

(I) FATTORE	(J) FATTORE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-1104,74	151,79	,000	-1518,78	-690,71
	3	-1716,02	151,79	,000	-2130,06	-1301,99
	4	506,12	128,63	,001	155,25	856,99
	5	961,02	129,61	,000	607,47	1314,58
2	1	1104,74	151,79	,000	690,71	1518,78
	3	-611,28	182,42	,007	-1108,89	-113,67
	4	1610,86	163,66	,000	1164,43	2057,29
	5	2065,77	164,43	,000	1617,23	2514,31
3	1	1716,02	151,79	,000	1301,99	2130,06
	2	611,28	182,42	,007	113,67	1108,89
	4	2222,14	163,66	,000	1775,71	2668,57
	5	2677,05	164,43	,000	2228,51	3125,59
4	1	-506,12	128,63	,001	-856,99	-155,25
	2	-1610,86	163,66	,000	-2057,29	-1164,43
	3	-2222,14	163,66	,000	-2668,57	-1775,71
	5	454,90	143,34	,013	63,91	845,89
5	1	-961,02	129,61	,000	-1314,58	-607,47
	2	-2065,77	164,43	,000	-2514,31	-1617,23
	3	-2677,05	164,43	,000	-3125,59	-2228,51
	4	-454,90	143,34	,013	-845,89	-63,91

The mean difference is significant at the .05 level.

F3
Tukey HSD

	N	Subset for alpha = .05				
FATTORE		1	2	3	4	5
5	40	1057,07				
4	41		1511,98			
1	65			2018,10		
2	25				3122,84	
3	25					3734,12
Sig.		1,000	1,000	1,000	1,000	1,000

Means for groups in homogeneous subsets are displayed.

REFERENCES

- [1] Nottebohm F. The origins of vocal learning. *Am Nat* 1972; 106: 116
- [2] Poole JH, Tyack PL, Stoeger-Horwath AS, Watwood S. Animal behaviour: elephants are capable of vocal learning. *Nature* 2005; 434: 455-6.
- [3] Janik VM, Sayigh LS, Wells RS. Signature whistle shape conveys identity information to bottlenose dolphins. *Proc Natl Acad Sci USA* 2006; 103: 8293-7.
- [4] Esser KM. Modeling aspects of speech processing in bats-behavioral and neurophysiological studies. *Speech Commun* 2002; 41: 179-88.
- [5] Fitch TW, Hauser MD, Chomsky N. The nature of the language faculty and its implications for evolution of language. *Cognition* 2005; 97: 211-25.
- [6] Spencer KL, Pepperberg IM. Culture: in the beak of the beholder? *Behav Brain Sci* 2001; 24: 341-2.
- [7] Jarvis ED, Ribeiro S, Silva ML, Ventura D, Viellard JM, Mello CV. Behaviourally driven gene expression reveals ong nuclei in hummingbird brain. *Nature* 2000; 406: 628-32
- [8] Geberzahn N, Hultsch H. Rules of song development and their use in vocal interaction in birds with large repertoires. *An Acad Bras Cienc* 2004; 76: 209-18.
- [9] Cruickshank AJ, Gautier JP, Chappuis C. Vocal mimicry in wild African Grey Parrots, *Psittacus erithacus*. *Ibis* 1993; 135: 293-9.
- [10] Snowdon CT, Hausberger M. Social influences on vocal development. Cambridge University Press, Cambridge MA, 1997.
- [11] Marler P. Origins of Music and Speech: Insights from Animals. In: Wallin NL, Merker S, Brown B, Eds. *The Origins of Music*. MIT Press 2001 Cambridge, pp. 345-70.
- [12] Pepperberg IM. Functional vocalizations by an African Grey Parrot (*Psittacus erithacus*). *Zeitsch Tierpsychol* 1981; 55: 139-60.
- [13] Pepperberg IM. Cognition in the African Grey Parrot: Preliminary evidence for auditory/vocal comprehension of the class concept. *Anim Learn Behav* 1983; 11: 179-85.
- [14] Pepperberg IM. The value of the Piagetian framework for comparative cognitive studies. *Anim Cogn* 2002; 5: 177-82.
- [15] Pepperberg IM. "Insightful" string-pulling in Grey Parrot (*Psittacus erithacus*) is affected by vocal competence. *Anim Cogn* 2004; 7: 263-266.
- [16] Pepperberg IM. Cognitive and communicative abilities of Grey parrots, *Applied Anim Behav Sci* 2006; 100: 77-86.
- [17] Pepperberg IM, Lynn SK. Perceptual consciousness in Grey Parrots. *Am Zool* 2000; 40: 393-401.
- [18] Pepperberg IM, Brese K, Harris B. Solitary sound play during acquisition of English vocalizations by an African Grey Parrot (*Psittacus erithacus*): possible parallels with children's monologue speech. *Appl Psycholinguistics* 1991; 12: 151-78.
- [19] Jarvis ED, Gunturkun O, Bruce L, et al. Avian brains and a new understanding of vertebrate brain evolution. *Nat Rev Neurosci*, 2005; 6: 151-9.
- [20] Todt D. Social learning of vocal patterns and mode of their application in Grey parrots. *Zeit fur Tierpsychol* 1975; 39: 178-88.
- [21] Barsalou LW. *Cognitive Psychology*. An overview for cognitive scientists. Lawrence Erlbaum Associates Hillsdale, New Jersey, 1992.
- [22] Armstrong EA. The nature and function of animal mimesis. *Bull Anl Behav* 1951; 9: 46-8.
- [23] Hurley S, Chater N. *Perspective on Imitation: from Neurosciences to Social Science*. MIT Press Cambridge, MA 2005.
- [24] Anthony JL, Francis D. Development of Phonological Awareness. *Curr Direct Psychol Sci* 2005; 14, 225-259.
- [25] Pepperberg IM. Grey Parrots do not always "parrot": the roles of imitation and phonological awareness in the creation of new labels from existing vocalizations. *Lang Sci* 2007; 29: 1-13.
- [26] Safran JR. Musical learning and language development. In: Avanzini GC, Faienza C, Minciacchi D, Lopez L, Majno L, Eds. *The Neuroscience and Music*. Annals of the New York Academy of Sciences, The New York Academy of Sciences, 2003; pp. 397-401.
- [27] Leonard LB. Fillers across languages and language abilities. *J Child Lang* 2001; 28: 257-61.
- [28] de Boysson-Bardie, B. *How language comes to children*. From birth to two years. MIT Press. Cambridge, Massachusetts 1999.
- [29] Lenti Boero D, Bottoni L. From crying to words: unique or multi-level selective pressures? *Brain Behav Sci* 2006; 29: 292-3.
- [30] Banta Lavenex P. Vocal production mechanisms in the Budgerigar (*Melopsittacus undulatus*). The presence and implications of amplitude modulation. *J Acoust Soc Am* 1999; 10: 491-505.
- [31] Beckers G, Nelson B, Suthers R. Vocal-Tract Filtering by Lingual Articulation in a Parrot. *Curr Biol* 2004; 14: 1592-7.
- [32] Patterson DK, Pepperberg IM. Acoustic and articulatory correlates of stop consonants in a parrot and a human subject. *J Acoust Soc Am* 1998; 103: 2197-215.
- [33] Lieberman P. *The biology and evolution of language*. Harvard University Press, Boston 1984.
- [34] Patterson DK, Pepperberg IM. A comparative study of human and parrot phonation: acoustic and articulatory correlates of vowels. *J Acoust Soc Am* 1994; 96: 634-48.
- [35] Il'ichev VD, Silaeva OL. Comparison of Species-Specific Bird Signals and Human Speech. *Biol Bull*, 2002; 29: 56-61.
- [36] Uplisova KO. Acoustical characteristics of talkig-birds' vowel-like sounds. *Acta of the XIII Session of the Russian Acoustical Society Moscow*, 2003.
- [37] Bottoni L, Massa R, Lenti-Boero D. The Grey Parrot (*Psittacus erithacus*) as a musician: an experiment with temperate scale. *Ethol Ecol Evol* 2003; 15: 133-41.
- [38] Caselli MC, Casadio P. *Il primo vocabolario del bambino*. Franco Angeli, Milano, 1995.
- [39] Patel A, Daniele JR. An empirical comparison of rhythm in language and music. *Cognition* 2003; 87: 35-45.
- [40] Frova A. *Fisica della musica*. Bologna, Zanichelli, 1999.
- [41] Charif RA, Mitchell S, Clark CW. *Canary 1.2 User's Manual*. Cornell Laboratory of Ornithology Press Ithaca, NY 1995.
- [42] Strange W. Evolving theories of vowel perception. *The J Acoust Soc Am* 1989; 85: 2081-7.

- [43] Gentilucci M, Cattaneo L. Automatic audiovisual integration in speech perception. *Exp Brain Res* 2005; 167: 66-75.
- [44] Peterson GE, Barney HL. Control methods used in the study of vowels. *J Acoust Soc Am* 1952; 24: 175-84.
- [45] Gonzalez J. Formant frequencies and body size of speaker: a weak relationship in adult humans. *J Phonol* 2004; 32: 277-87.
- [46] Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. *J Acoust Soc Am* 1995; 97: 3099-111.
- [47] Margoliash D. Offline learning and the role of autogenous speech: new suggestions from birdsong research. *Speech Commun* 2003; 41: 165-78.
- [48] Baptista LF, Petrinovich L. Social interaction, sensitive phases and the song template hypothesis in the white-crowned sparrow. *Anim Behav* 1984; 32: 172-81.
- [49] Geberzahn N, Hultsch H. Rules of song development and their use in vocal interaction in birds with large repertoires. *An Acad Bras Cienc* 2004; 76: 209-18.
- [50] Freeberg TM, West MJ, King AP, Duncan SD, Sengelaub DR. Cultures, genes and neurons in the development of song and singing in brown-headed cowbirds (*Molothrus ater*). *J Comp Physiol [B]* 2002; 188: 993-1002.
- [51] Hausberger M, Richard-Yris A, Henry L, Lepage L, Schmidt, I. 1995 Song Sharing Reflects the Social Organization in a Captive Group of European Starlings (*Sturnus vulgaris*). *J Comp Physiol [B]* 1995; 109: 222-41.
- [52] Kuhl PK, Miller JD. Speech perception by chinchilla. *J Acoust Soc Am* 1975; 57: S49-S50.
- [53] Kluender KR, Diehl RL, Killeen PR. Japanese quail can learn phonetic categories. *Science* 1987; 237: 1195-7.
- [54] Miller JD. Perception of speech sounds in animals: evidence for speech processing by mammalian auditory mechanisms. In: Bullock T, Ed., Dahlem workshop on recognition of complex acoustic signals. Life Sciences report, Abakon, Berlin, 1977; pp. 49-58.
- [55] Morse PA, Snowdon CT. An investigation of categorical speech discrimination by rhesus monkeys. *Percept Psychophys* 1975; 17: 9-16.
- [56] Todt D. Spontaneous recombinations of vocal patterns in parrots. *Naturwissenschaften* 1975; 62: 399-400.
- [57] Lenti Boero D, Bottoni L. Why we experience musical emotions: intrinsic musicality in an evolutionary perspective. *Brain Behav Sci*; in press.

Received: January 26, 2009

Revised: February 20, 2009

Accepted: March 12, 2009

© Bottoni et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.