International Journal of Scientific and Technical Advancements

ISSN: 2454-1532

Investigating the Quality of Speech of Birds Using Linear Predictive Coding

Randhir Singh[1#], Ajay Kumar[2], Parveen Kumar Lehana[3]
[1]I. K. Gujral Punjab Technical University, Kapurthala, Punjab, India
[2]Beant College of Engineering and Technology, Gurdaspur, Punjab, India
[3]Department of Electronics, University of Jammu, Jammu, India
Email address: [#]errandhir81@gmail.com

Abstract— For both human beings and birds, sound is the most affective mode for communication over long distances. The aim of this research is to investigate the quality of the speech of parrots and crows in comparison to human beings using Linear Predictive Coding (LPC) as an analysis-synthesis platform. Analysis using Linear Predictive Coding (PESQ), Mean opinion score (MOS), and visual inspection of the spectrograms shows that the LPC model is able to synthesize the phrases uttered by the parrots and crows efficiently.

Keywords— Bird calls; Bird songs; LPC; MOS; PESQ; Speech production; Speech quality.

I. Introduction

Human beings, some animals, and few birds are have developed a certain level of intelligence. They have developed different methods of communication, including speech and singing. Among these, speech signals stands out the most natural, efficient, and the affective mode of communication among humans, animals, and machines. Speech signal consists of known small units of sound called phonemes. It conveys information about age, sex, emotion, and even the state of health of the speaker [1]. The most important organs of the human beings involved in speech production are lungs, larynx, vocal tract, tongue, lips, muscles that move these organs, nerves, and finally the brain where are the signals to be transmitted to the muscles are conceptualized. Speech is a time-varying signal. It depends on known physical movements of the articulators (jaw, tongue, velum, lips, and mouth). It is different for every speaker and may be slow, fast, or varying in speed. It may also have low pitch, high pitch, or be whispered. It has widely varying kinds of environmental noise. Speech has an unlimited number of words.

The frequency is one of the most important parameters of the sound. The sounds are distinguished from each other with the help of their frequencies. When the frequency of a sound increases, the sound gets high-pitched and irritating. When the frequency of a sound decreases, the sound gets deepen. The sound waves are waves that occur from the vibration of the materials. The spectral content of human speech falls in the range 70 Hz to 10 kHz. This range may vary from person to person. A normal human speech has a frequency interval of 100 Hz - 3200 Hz and its magnitude is in the range of 30 dB - 90 dB. A human ear can perceive sounds in the frequency range between 16 Hz and 20 kHz. Due to the variances in vocal tract length, children, female, and male speech are different. Changes in regional accents lead to the changes in resonant frequencies, durations, and pitch. Individuals have resonant frequency patterns and duration patterns that are unique. This allows us to identify the speaker. Human speech is band limited in the range 300-3400 Hz. Digitizing speech requires a minimal sampling rate of 8 kHz. Speech is often quantized with 8 bits/sample resulting in a data rate of 64 kbps.

Similar to human beings, birds also use sound signals for communicating their emotions and is an ideal method of communication over long distances or for poor light [2-5]. The sound signals of the birds may be classified as calls and songs. The calls are generally of short duration, unmusical signal, less complex than songs and produced by both males and females. It is used for immediate requirement of contact, alarming threat, keeping in touch while flying, announcing their location, and about food sources. The songs are musical, complex as compared to calls, and sung usually only by males. The purpose of the song may be an announcement of their territory, to attract females, and to compete with other males. The organs used by the birds are lungs, bronchi, syrinx, trachea, larynx, mouth, and beak [6, 7]. Sound produces in the syrinx, located between bronchi [8] and the tracheas or sometimes it may be completely in the bronchi or tracheas, depending upon the specie [9]. The role of syrinx in birds is similar to that of vocal cord in human beings. Its shape varies with the size of the species. The vocal tract of birds modulates the sound produced by syrinx similar to the vocal tract of human beings. The similarity of the sound production mechanism of birds and human beings has been confirmed and described by several researchers [10-12].

The animals such as whales, dolphins, songbirds, bats, hummingbirds, parrots, crow, and primates also have vocal capabilities as like human beings [13-15]. Parrots are mostly found in tropical and subtropical regions of the globe like central and South America, Caribbean, Africa, Asia, India, New Zealand, Australia, etc. They are brightly colored birds with short curved beaks, biologically known as *Psittacines* birds. There are about 350 different types of parrot species and about 100 of these are declared as endangered. The size and weight of parrots vary with the species. The biggest parrot is the Hyacinth Macaw and the smallest is a pygmy parrot. The heaviest parrot is the Kakapo. These birds have first and fourth toes pointing backward; and the second and third pointing forward. They eat seeds, nuts, fruit, buds, and other plant material. Parrots are the only birds that use foot to bring food to their mouth. Parrots mostly nest in the hollows of old trees

Randhir Singh, Ajay Kumar, and Parveen Kumar Lehana, "Investigating the quality of speech of birds using linear predictive coding," *International Journal of Scientific and Technical Advancements*, Volume 3, Issue 3, pp. 1-12, 2017.

ISSN: 2454-1532

or cliffs. They lay 2 to7 eggs and incubate them for 22 to 30 days. When the babies hatch, the mother usually feeds them and the father brings food for them.

Only few species of parrots are known to imitate. For example, Alexandrine parrots (*Psittacula eupatria*) are found to have the ability of mimicking the human speech [4, 16]. They have a clear voice and enjoy communicating with human. They are start mimicking words as early as 4 months old. The Alexandrine parrot species includes both Indian and African Ringneck parrots. Alexandrine parrots have the dark green bodies, long tails, red beaks, and yellow eyes. They have only major visible difference from Indian Ringneck parrots that their maroon patches on the shoulders and larger beaks. These species are sexually dimorphic which means males and females look different form each other. Alexandrine male parrots have a dark black ring and pink collars, which are absent in female (Fig. 1). The male Alexandrine parrots start to develop their ring around after18 months and matures as long as three years old. The tails of female are shorter than those of the male. The female is usually lighter than the male. Detailed information of Alexandrine parrot is listed in table I [4], [16-19].

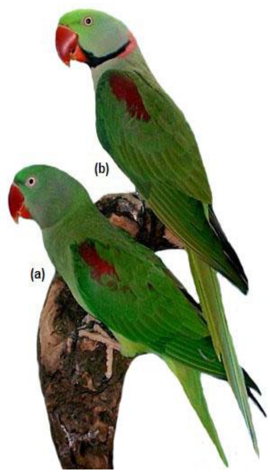

Fig. 1. Alexandrine parrot: a female, b male [4].

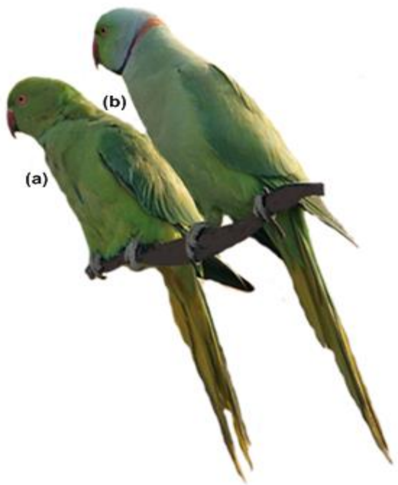

Fig. 2. Indian Ringneck parrot: a female, b male.

Another species of parrots, Indian Ringneck parrots (*Psittacula krameri*) or Indian Ringnecked parakeets, also known as Rose-ringed parakeets, also have imitating capabilities. They are found in India and Pakistan. Both sexes are known for their intelligent and have the ability to mimic human speech [5, 16]. They are classified as smaller parrots with long tails, which generally refer to as "parakeet”. They have a hooked red-beak, green body, a long tail (6-7inch), about 120 - 140 g of weight, and yellow feathers underling on the wings. The Indian Ringneck parrots are also sexually dimorphic (Fig. 2), mature males have a black marking on the face and rose-coloured ring around their neck, which gives the species their name. Females have a very faint ring around the neck. Detailed information of Indian Ringneck parrot is listed in table II [5, 16, 17, 20, 21].

Some of the crows also have imitating capabilities. Crows belong to order *Passeriformes* and the family *Corvidae* of birds [22], which also includes ravens, magpies, and jays. They are found all over the world except Antarctica. There are about 40 - 43 species of the *Corvidae* depend on the investigation by the different researcher [23]. As per the fourth edition of National Geographic’s Field Guide to the Birds of North America (2002), the body length of some of the Corvidae family are – Blue Jay Cyanocitta cristata 28 cm, American Crow Corvus brachyrhynchos 45 cm, Black-billed Magpie Pica hudsonia 48 cm, Common Raven Corvus corax 61 cm. They are known to be the most adaptable, boldness and extremely intelligent birds [16, 24, 25]. They are also known for their problem-solving skills, tools making, and amazing communication skills.

TABLE I. Detailed information of Alexandrine parrot.

	Alexandrine parrot
Common Names	Alexandrian Parrot, Alexandrine Parakeet
Scientific Name	*Psittacula eupatria*
Original Homeland	India, Afghanistan, Pakistan, Sri Lanka, Burma, Thailand, and China
Habitat	Highland forests, lowland forests, cultivated areas, urban areas, parks, and gardens
Appearance	Male: Green body, red beak, maroon patches on wings, and pink and black rings on neck Female: Green body, red beak, and a faint ring around the neck
Size	Length: 55-60 cm (including tail feathers) Wing length: 18.9-21.5 cm Tail: 21.5-35.5 cm
Weight	200 - 300 g
Average Life Span	35 - 40 years
Diet	Seeds: sunflower, safflower, mixed canary Grains: mixed millet, maize, wheat, rice Fruit: apple, mango, guava, grape, peach, grapes, berries Vegetables: peas, celery, cabbage, carrot etc.
Sexual Maturity	3 - 4 years
Number of eggs lay	2 to 4
Incubation Period	21 - 26 days
Number of Young	2 to 4
Fledging age	7 weeks
Mimicking Capability	Yes

TABLE II. Detailed information of Indian Ringneck parrot

	Indian Ringneck parrot
Common Names	Indian Ringneck Parrot, Indian Ringnecked Parakeet, Rose-Ringed Parakeet
Scientific Name	*Psittacula krameri*
Original Homeland	India, Pakistan, Burma, Nepal, Sri Lanka, Bhutan, Tibet and China
Habitat	Semi-deserts, woodlands, cultivated areas, urban areas, parks, and gardens
Appearance	Male: Green body, red beak, and pink and black rings on neck Female: Green body, Red beak, a faint ring around the neck
Size	Length: 37-43 cm (including tail feathers) Wing length: 15-17.5 cm Tail: 15-18 cm
Weight	120 - 140 g
Average Life Span	20 - 30 years
Diet	Seeds: sunflower, safflower, mixed canary Grains: mixed millet, maize, wheat, rice Fruit: apple, mango, guava, grape, peach, grapes, berries Vegetables: peas, celery, cabbage, carrot etc.
Sexual Maturity	3 - 4 years
Number of eggs lay	3 to 6
Incubation Period	21 - 24 days
Number of Young	2 to 4
Fledging age	7 weeks
Mimicking Capability	Yes

Out of the 43 species of crow, the Indian house crow is easily found in rural and urban area of India. The Indian house crow (*Corvus splendens*), also called Indian Grey-necked crow, Ceylon crow, or Colombo crow is an intelligent bird [26]. The Indian house crow originated in the Indian subcontinent and spread to other parts of the world by shipping, where it is considered as one of the world's most invasive bird species [27]. Indian house crows are 42 - 44 cm long (including body and tail) and weigh 300 - 400 g. Their feathers are glossy black all over its body, except its neck and breast, which is smoky-grey in color and not glossy (Fig. 3). It has black beak, legs, and feet. Both males and females are lookalike except the size, which slightly larger for males [27]. Juveniles have little or no gloss on their feathers [28]. Indian house crows lives close to people (in villages, towns and cities), where they can find food scraps and other rubbish. Indian House Crows are omnivorous i.e. feeding on both plants and animals. It consumes various insects, lizards, frogs, fish, snakes, mice, fruits and cereal seeds, even eggs and nestlings of other birds. Indian house Crows are very social and non-migratory birds. They have an integrated family, which protects them from horned-owls, red-tailed hawks, and raccoons. They mate for life. Indian house crows live for approximately 6 years in the wild [28]. Detailed information of Indian house crow is listed in table III [16, 24, 25, 29].

Fig. 3. Indian House Crow.

Due to its invasive nature, Indian house crows can create negative effects in around 25 or more countries throughout Africa, the Middle East and South East Asia [30, 31, 32, 33]. It is now considered as one of the world's most invasive bird species [27] and has negative effects on agriculture, tourism, human health, traffic, transportation, and biodiversity [31]. Indian house crows are serious agricultural pest feeding on a variety of crops, mainly maize [34, 35] and damage orchards [36, 37]. The Indian house crows are also responsible for the decline of small reptiles and amphibians, insects, fish, birds and mammals, and domestic animals [27]. It disturbs tourists and local people with their loud calls, and aggressive attacks while attempting to steal food [36]. It also poses a bird strike risk to aero planes due to their flocking nature [31]. It is also known to transmit pathogens, which affect people and domestic animals [38, 39]. These species also have been reported for carrier and transmitter of diseases, such as cholera, dysentery, West Nile Virus [38, 40], bird flu [41] as a fecal contaminator of human environments and water sources [42].

There are several models for speech analysis and synthesis. For example, LPC has widely been used for the analysis and synthesis of human speech [43]. The objective of this paper is to investigate the capabilities of LPC for the analysis and synthesis of bird calls particularly parrots and crows. LPC analysis and synthesis procedure is presented in the following section. Section III is devoted to methodology of the research work. Results and discussions are presented in Section IV followed by the conclusions.

II. LINEAR PREDICTIVE CODING (LPC)

In the year 1967, B.S. Atal and M.R. Schroeder introduced first time linear prediction for the processing of speech signals [44]. In the year 1982, this model became a federal government standard. Linear Predictive Coding (LPC) is an all pole model for speech analysis and synthesis. It has the capability to provide accurate approximation of speech parameters and efficient for computation purposes. LPC relates the present sample of speech as a linear combination of the past p samples, where p is called the order of LPC. The method of eliminating the formants from the speech signal is known as inverse filtering and the remaining signal is known

Randhir Singh, Ajay Kumar, and Parveen Kumar Lehana, "Investigating the quality of speech of birds using linear predictive coding," *International Journal of Scientific and Technical Advancements*, Volume 3, Issue 3, pp. 1-12, 2017.

as residual signal. LPC synthesizes the speech signal using an inverse mechanism; it employs the residual signal to generate a source speech signal. In the synthesis procedure, the formants are used to generate an all-pole filter and the filter processes the source signal to generate synthesized output speech [45]. LPC techniques are the most commonly used in speech coding, speech recognition, speech synthesis, speaker recognition and verification and for speech storage.

TABLE III. Detailed information of Indian house crow.

	Indian House crow
Common Names	Indian House crow, House crow, Indian crow, Grey-necked crow, Ceylon crow, Colombo crow
Scientific Name	*Corvus splendens*
Family	Corvidae
Original Homeland	India, Pakistan, Nepal, Bangladesh, Sri Lanka, Maldives and Laccadive Islands, South West Thailand and coastal southern Iran
Habitat	The house crow is generally a lowland species, found in tropical and subtropical areas. Lives close to people, found in villages, towns and cities (usually in areas where they can find food scraps and other rubbish). It needs some trees too. No populations of house crows are known to live independently of people
Appearance	Forehead, crown, throat and upper breast are a richly glossed black, whilst the neck and breast are grayish in color. Eyes are brown. The bill, wings, tail, legs, and feet are also black
Size	Length: 41-43 cm Wingspan: 76-85 cm
Weight	300 – 400 g
Average Life Span	6 years
Diet	Omnivorous. The diet includes seeds, fruit, grain, nectar, berries, bird's eggs, nestlings, mammals, reptiles, amphibians, fish, insects, carrion and food scraps
Sexing	Sexes alike
Sexual Maturity	2-4 years
Number of eggs laid	3 to 5
Incubation Period	Female incubates the eggs, the incubation period is 16 – 17 days
Number of Young	2 to 4
Fledging age	21–28 days
Mimicking Capability	Mimicking some phrases

Linear Predictive Coding (LPC) [43-49] is an analysis-synthesis method of lossy speech compression that approximate to model the human production of sound. LPC is originally designed for speech coding systems, but may be also used in speech synthesis. This technique approximates the basic speech parameters (pitch, formants, spectra, vocal tract area functions, etc.) and achieves a low bit rate. It can also reduce calculations and accuracy in speech recognition process. The basic LPC is based on the source-filter-model of speech. Linear prediction models are also referred to as autoregressive modeling or AR-modeling. From a frame of natural speech the digital filter coefficients are estimated automatically

The basis LPC model predicted the current speech sample from a finite number of previous samples by a linear combination with the small error term called residual signal [50]. The covariance method and the autocorrelation method are normally used to calculate these coefficients. Only the autocorrelation method guaranteed the stable of the filter.

In synthesis stage the used excitation is estimated for voiced sounds by a train of impulses and for unvoiced by random noise. After the excitation, signal is gained and filtered with a digital filter for each coefficient. The filter order is generally between 10 and 12 at 8 KHz sampling rate and for higher quality at 22 kHz sampling rate, the filter order required is between 20 and 24. The coefficients are regularly updated after every 5-10 ms [51]. A linear predictive coding model of speech is diagrammatically shown in Fig.4, where G represents Gain. LPC techniques are the most commonly used in speech coding, speech recognition, speech synthesis, speaker recognition and verification and for speech storage.

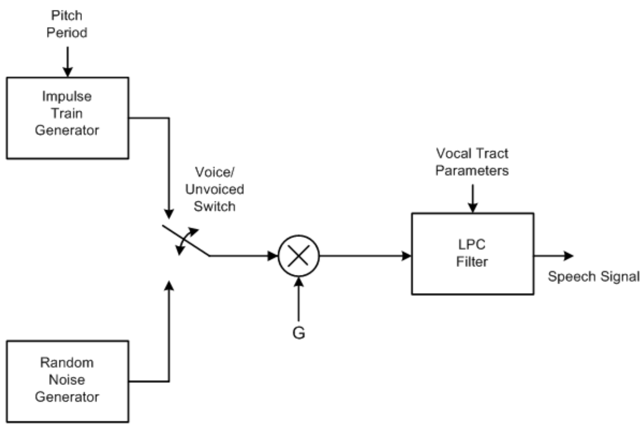

Fig. 4. LPC model of speech.

LPC is a short-term estimation method because speech signals are suitable only for a short interval of time. There are two techniques for processing, short-term analysis methods. In the first technique, the windowed speech frame $s_w(n)$ is obtained by multiplying each speech frame with the window function $w(n)$. The LPC coefficients for each frame are estimated from the autocorrelation vector using a Levinson or a Durbin recursion method. In the second technique, the covariance method of analysis is used to get the LPC coefficients.

The LPC method predicted the present speech sample $s(n)$ at time n as a linear combination of the past p speech samples in the following way:

$$s(n) \approx a_1 s(n-1) + a_2 s(n-2) + ... + a_p s(n-p) \quad (1)$$

where $a_1, ..., a_p$ are prediction coefficients. By including an excitation term $Gu(n)$, the equation (1) can be transformed

$$s(n) = \sum_{i=1}^{p} a_i s(n-i) + Gu(n) \quad (2)$$

where G is the gain parameter and $u(n)$ is the normalized excitation. Transforming equation (2) to z-domain, we get:

Randhir Singh, Ajay Kumar, and Parveen Kumar Lehana, "Investigating the quality of speech of birds using linear predictive coding," *International Journal of Scientific and Technical Advancements*, Volume 3, Issue 3, pp. 1-12, 2017.

$$S(z)=\sum_{i=1}^{p}a_i z^{-i}S(z)+GU(z) \quad (3)$$

and therefore, the transfer function $H(z)$ becomes

$$H(z)=\frac{S(z)}{GU(z)}=\frac{1}{1-\sum_{i=1}^{p}a_i z^{-1}}=\frac{1}{A(z)} \quad (4)$$

This resembles to the transfer function of a digital time varying filter. The LPC model is designed to obtain the speech parameters i.e. classification of voiced/unvoiced, the gain, the pitch period, and the prediction coefficients. Higher the order of the model better the all-pole model can estimate the spoken sounds. A linear predictor with coefficients is defined as follows:

$$P(z)=\sum_{k=1}^{p}\alpha_k z^{-k} \quad (5)$$

and their output is

$$\tilde{s}(n)=\sum_{k=1}^{p}\alpha_k s(n-k) \quad (6)$$

The prediction error $e(n)$ is defined as:

$$e(n)=s(n)-\tilde{s}(n)=s(n)-\sum_{k=1}^{p}\alpha_k s(n-k) \quad (7)$$

which is the output of the system

$$A(z)=1-\sum_{k=1}^{p}\alpha_k z^{-k} \quad (8)$$

If $a_k=\alpha_k$, then $H(z)=\frac{G}{A(z)}$.

The main aim now is to compute the set of coefficients α_k that minimizes the square of the prediction error.

The mean short-time prediction error per frame can be represented as follows:

$$E_n=\sum_m e_n^2(m)=\left[s_n(m)-\sum_{k=1}^{p}\alpha_m s_n(m-k)\right]^2 \quad (9)$$

where $s_n(m)$ is a segment of speech selected in the neighborhood of sample n: $s_n(m)=s(m+n)$.

The values of the coefficients α_k which minimize the total prediction error E_n can be obtained from $\frac{\partial E_n}{\partial \alpha_i}=0$, $i=1,2,....,p$.

The differentiated expression can be resulted in the next equation:

$$\sum_m s_n(m-i)s_n(m-k)=\sum_{k=1}^{p}\alpha'_k\sum_m s_n(m-i)s_n(m-k) \quad 1\le i\le p \quad (10)$$

where α'_k are the values of α_k that minimize E_n.

Defining $\Phi_n(i,k)=\sum_m s_n(m-i)s_n(m-k)$ equation (10) can be written as:

$$\sum_{k=1}^{p}\alpha_k\Phi_n(i,k)=\Phi_n(i,0) \qquad i=1,2,......,p \quad (11)$$

This is a system of p equations with p variables that can be solved to find the α_k coefficients for the segment $s_n(m)$.

It can be demonstrated that

$$E_n=\sum s_n^2(m)-\sum_{k=1}^{p}\alpha_k\sum_m s_n(m)s_n(m-k) \quad (12)$$

and in compact form:

$$E_n=\Phi_n(0,0)-\sum_{k=1}^{p}\alpha_k\Phi_n(0,k) \quad (13)$$

Now, the values $\Phi_n(i,k)$ for $1\le i\le p$ and $0\le k\le p$, and the α_k coefficients are obtained by solving equation (11). By using the autocorrelation technique the system given by equation (13) can be solved.

The autocorrelation technique considers the segments $s_n(m)=0$ outside the interval $0\le m\le N-1$ and $s_n(m)=s(m+n)w(m)$ in the interval, where $w(m)$ is a finite-length window. If $s_n(m)$ differs from zero for $0\le m\le N-1$, the correspondent prediction error $e_n(m)$ for a linear predictor of order p will be different from zero in the interval $0\le m\le N-1+p$.

Therefore, $E_n=\sum_{m=0}^{N-1+p}e_n^2(m)$

Using this technique, the prediction error is large at the beginning and at the end of the interval due to the prediction of null samples in the extremes.

Due this reason, a windowing process (Hamming window) should be applied to every segment for reducing the border values. By taking the consideration that $s_n(m)$ is null outside the interval $0\le m\le N-1$.

It can be demonstrated that

$$\Phi_n(i,k)=\sum_{m=0}^{N-1+p}s_n(m-i)s_n(m-k) \quad 0\le k\le p \quad 1\le i\le p \quad (14)$$

which can be rewritten as

$$\Phi_n(i,k)=\sum_{m=0}^{N-1-(i-k)}s_n(m)s_n(m+i-k) \quad 1\le i\le p \quad 0\le k\le p \quad (15)$$

For this case, $\Phi_n(i,k)$ is related to the short-time autocorrelation function valued for $i-k$:

$$\Phi_n(i,k)=R_n(i,k)$$

here $R_n(k)=\sum_{m=0}^{N-1-k}s_n(m)s_n(m+k)$ is a pair function, so:

$$\Phi_n(i,k)=R_n(|i-k|) \quad i=1,2,...,p \quad k=0,1,...,p$$

Hence,

$$\sum_{k=1}^{p}\alpha_k R_n(|i-k|)=R_n(i) \qquad 1\le i\le p$$

In a similar way, the square prediction error is:

$$E_n=R_n(0)-\sum_{k=1}^{p}\alpha_k R_n(k) \quad (16)$$

In matrix-vector form the equation system can be expressed as:

$$\begin{bmatrix} R_n(0) & R_n(1) & R_n(2) & \dots & R_n(p-1) \\ R_n(1) & R_n(0) & R_n(1) & \dots & R_n(p-2) \\ R_n(2) & R_n(1) & R_n(0) & \dots & R_n(p-3) \\ \dots & \dots & \dots & \dots & \dots \\ R_n(p-1) & R_n(p-2) & R_n(p-3) & \dots & R_n(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \dots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R_n(1) \\ R_n(2) \\ R_n(3) \\ \dots \\ R_n(p) \end{bmatrix} \quad (17)$$

Levinson-Durbin algorithm can be used to solve the above matrix equation more efficiently.

The Levinson-Durbin's algorithm is employed to solve equation systems in which the elements across the diagonal are identical and coefficients of the matrix are symmetric (Teoplitz matrix). The complete Levinson-Durbin algorithm is:

$$E^{(0)} = R(0) \text{ for } 1 \le i \le p$$

$$k_i = \frac{\left[R(i) - \sum_{j=1}^{i-1} \alpha_j^{(i-1)} R(i-j) \right]}{E^{(i-1)}}$$

$$\alpha_i^{(i)} = k_i \text{ for } 1 \le j \le i-1$$

$$\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)}$$

$$E^{(i)} = (1-k_i^2) E^{(i-1)}$$

These equations are solved recursively for $i = 1, 2, ..., p$ and the final solution is given by:

$\alpha_j = \alpha_j^{(p)}$ where $1 \le j \le p$.

There are several advantages of LPC platform [52]. It reduces the bit rate of the speech signal resulting in ease of transmission as it requires less bandwidth. LPC uses encryption of the signal and hence, is more secure. LPC model is also associated with few disadvantages as well. It is an all-pole model, therefore, difficult to synthesize nasal sounds. Data gets faded when transmitting the data on the long distance because this is a lossy compression technique. The quality of speech is also poor with short plosives due to the time-scale events may be shorter than the frame size used for the analysis.

III. METHODOLOGY

The methodology employed for investigation related to the validity of LPC model for the analysis, synthesis, and comparison of speech generated by human beings (female), parrots (Indian Ringneck parrots), and crows (Indian house crows) may be divided into the following four sub-sections. Flow chart of the methodology is shown in Fig. 5.

Speech Recording: The phrases uttered by human beings and parrots were recorded using a high quality audio system in an acoustically treated room at sampling frequency of 16 kHz with 16 bit quantization. The recording of the calls of the crow was carrying out in the natural environment as the arrangements for recording could not be completed for this bird. The crows used in the investigation were not professionally trained; it was difficult to make them to imitate the phrases spoken by the human subjects. Hence their natural calls were recorded for the investigation. Some of the calls were also obtained from the Internet and other resources. A total of 21 phrases were used for processing, among them, 7 phrases were selected for each subject.

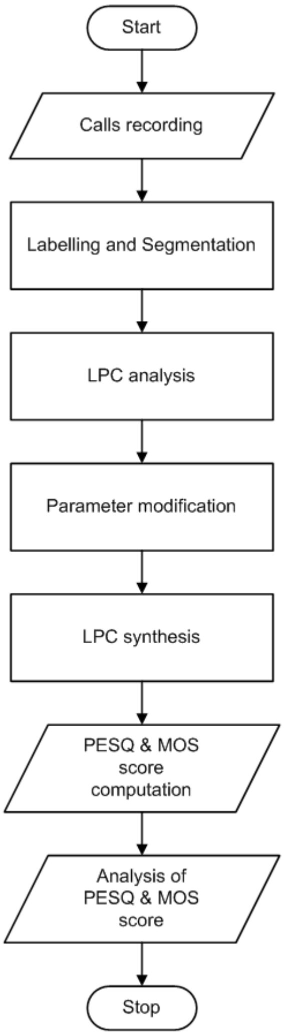

Fig. 5. Flow chart of the methodology.

Segmentation and Labelling: The recorded phrases were segmented and labelled in accordance with the content. The irrelevant information from the beginning and end section was removed using PRAAT software.

Analysis and Synthesis: The segmented phrases were analysed and synthesized using LPC based model. The optimum value of the analysis window and LPC order were estimated based on the PESQ score of the synthesised phrases. For investigating the effect of residue and other parameters, re-synthesis was carried out by modifying the residue.

Speech Evaluation: The evaluation of the synthesized speech has been carried out using mean opinion score (MOS) [53, 54, 55] and Perceptual Evaluation of Speech Quality (PESQ) [56, 57]. In MOS test, the subjects respond the quality of the speech presented to him on 1-5 scale (1: bad, 2: poor, 3: fair, 4: good, 5: excellent). MOS is more accurate, but time-

Randhir Singh, Ajay Kumar, and Parveen Kumar Lehana, "Investigating the quality of speech of birds using linear predictive coding," *International Journal of Scientific and Technical Advancements*, Volume 3, Issue 3, pp. 1-12, 2017.

consuming and non-repeatable. Hence, PESQ is mostly preferred for preliminary evaluation. The final results are evaluated using MOS.

IV. Results and Discussions

The spectral frequency content of human beings, parrots, and crows is different as the sound production mechanisms are not identical. Wrong selection of analysis window may deteriorate the output parameters. Hence, the optimum value of analysis window is obtained by investigating the PESQ score for different value of analysis window varying from 2 ms to 30 ms. It was observed that the PESQ score is satisfactory for the analysis window length around 10 ms for human beings, 12 ms for parrots and 10 ms for crows. Although, the quality of the synthesized output using original LPC residue was better than that of the synthesized using random noise as the residual, only random residual was preferred to reduce the number of parameter required after LPC based analysis of the phrases.

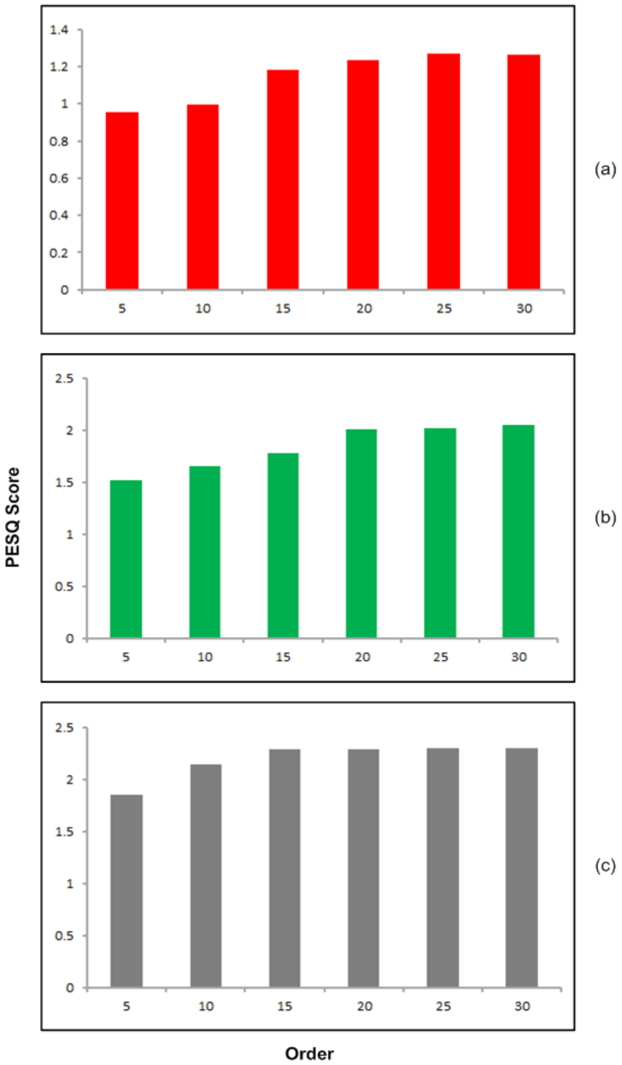

Fig. 6. PESQ score of the synthesized phrases uttered by: a human beings (female), b parrots (Indian Ringneck parrots), and c crows (Indian House crows) for different values of order synthesized with random LPC residual.

Second important parameter for LPC analysis-synthesis is the order. For estimating the optimal value of LPC order the investigations were conducted by varying the order from 5 to 30. The results are shown in Fig. 6 for human beings, parrots, and crows. The analysis shows that the optimum value of the order may be taken around 25 for human beings, 20 for parrots, and 15 for crows for satisfactory quality of the synthesized LPC output.

The spectrograms for the original and synthesized phrases uttered by human beings, parrots and crows are shown in Fig. 7, Fig. 8, and Fig. 9, respectively. The visual inspection of the spectrograms shows that the LPC model is able to synthesize the phrases uttered by human beings along with the birds efficiently as the formant structure and evaluation with respect to time is maintained in all phrases.

The results of the investigation carried out using subjective evaluation (MOS) are shown in Table IV and plotted as histograms in Fig. 10. It has been observed that the MOS of the original recorded speech of human beings is relatively more as compare to crows and the parrots. When speech is synthesized using random residual, the score of all the three, i.e., human beings , parrots, and crows decreases. The maximum reduction was observed for the human beings and minimum for the parrots. The results with original LPC residual were comparable to the original recorded speech for all the three subjects. Further, comparatively low value of the standard deviation confirms repeatability of the results.

V. Conclusions

Investigations have been carried out to study the quality of the speech uttered by parrots and crows in comparison to human beings by using LPC as the analysis-synthesis platform. It was observed that the quality of the synthesized output using original LPC residue was better than that of the synthesized using random noise as the residual taking the order 25 for human beings, 20 for parrots, and 15 for crows. The visual inspection of the spectrograms shows that the LPC model is able to synthesize the phrases uttered by the parrots and crows efficiently. MOS of the original recorded human speech is relatively more as compared to that of crows and the parrots.

Acknowledgment

The authors are thankful to I. K. Gujral Punjab Technical University, Kapurthala, Punjab, Sri Sai College of Engineering & Technology, Badhani, Beant College of Engineering & Technology, Gurdaspur, and DSP Lab, University of Jammu for providing resources for the research work.

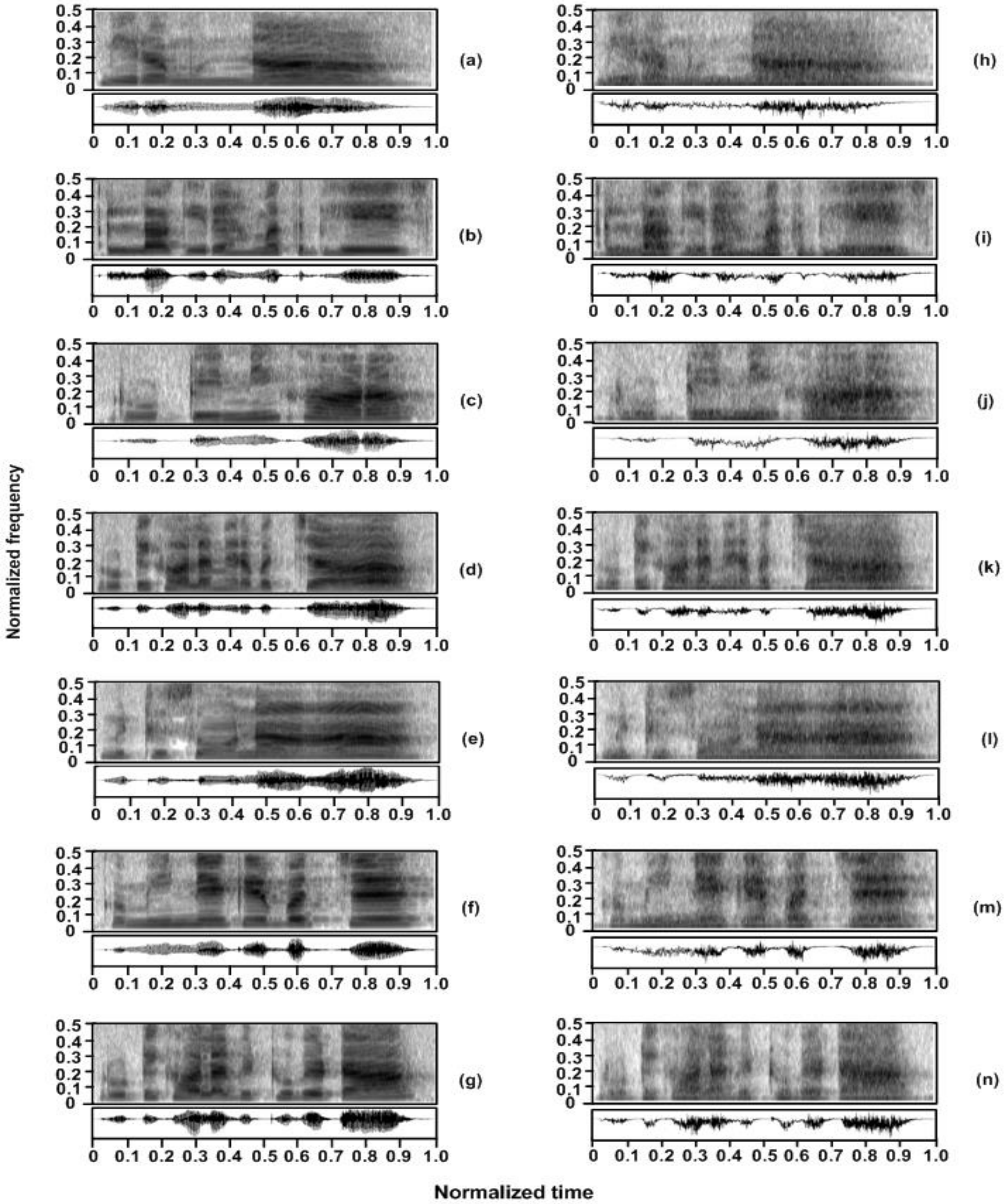

Fig. 7. Spectrograms of the original recorded phrases uttered by human (female) speaker (Column 1) and the same phrases synthesized by LPC with random residual (Column 2). The phrases in the spectrograms are (**a, h**) /mera munna/, (**b, i**) /əlla ki: rəɦəmət $b^{ɦ}$eʤ/, (**c, j**) /roʈi: ni: $k^{ɦ}$a rəɦa/, (**d, k**) /roʈi: $k^{ɦ}$a le mera bəʈʃʈʃa/, (**e, l**) /miʈ^{ɦ}ʈ^{ɦ}u: soɦəna/, (**f, m**) /kəmi:ne ke bəʈʃʈʃe/, and (**g, n**) /roʈi: $k^{ɦ}$ a le miʈ^{ɦ}ʈ^{ɦ}u: beʈa/.

ISSN: 2454-1532

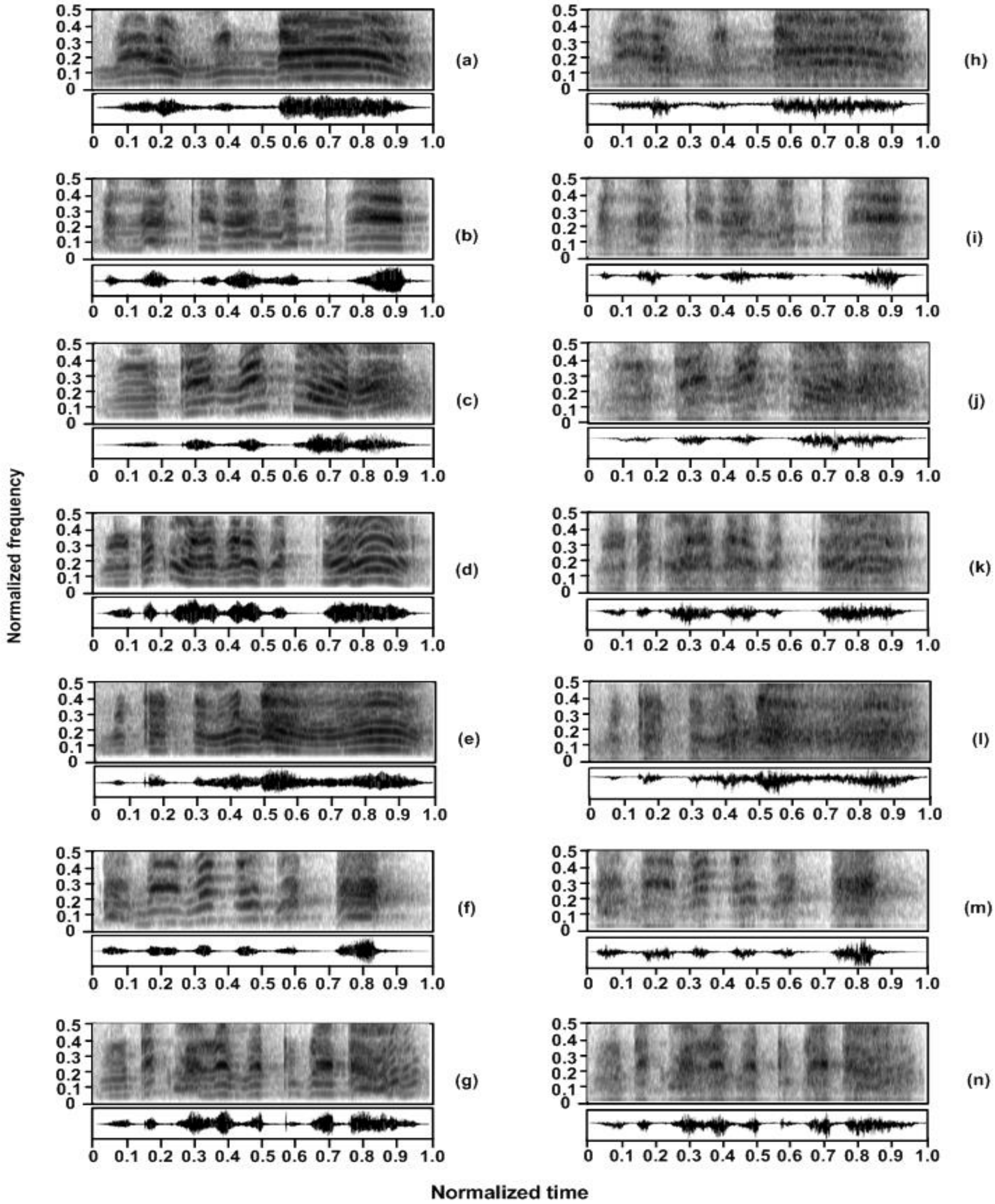

Fig. 8. Spectrograms of the original recorded phrases uttered by the parrot (Column 1) and the same phrases synthesized by LPC with random residual (Column 2). The phrases in the spectrograms are (**a, h**) /mera munna/, (**b, i**) /əlla ki: rəɦəmət bʱedʒ/, (**c, j**) /roʈi: ni: kʱa rəɦa/, (**d, k**) /roʈi: kʱa le mera bətʃtʃa/, (**e, l**) /miʈʱʈʱu: soɦəna/, (**f, m**) /kəmi:ne ke bətʃtʃe/, and (**g, n**) /roʈi: kʱ a le miʈʱʈʱu: beʈa/.

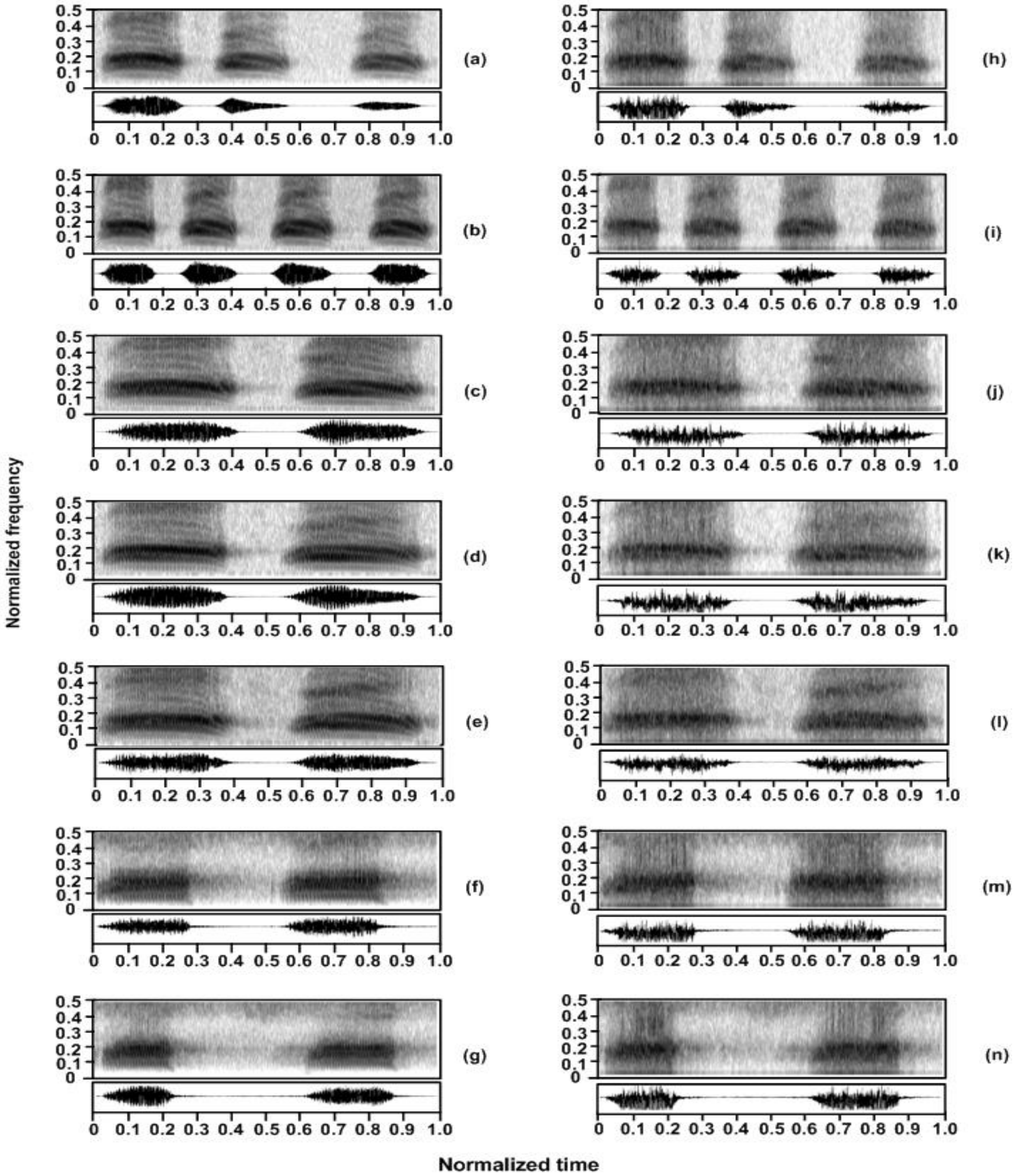

Fig. 9. Spectrograms of the original recorded phrases uttered by crow (Column 1) and the same phrases synthesized by LPC with random residual (Column 2).

TABLE IV. MOS for the original recorded signal, synthesized using random residue, and the synthesized output using original residue for human beings, parrots, and crows.

	Original			Random residue			Original residue	
	MOS	S.D.		MOS	S.D.		MOS	S.D.
Human	4.3	0.5		2.4	1.0		4.2	0.6
Parrot	2.9	0.9		1.9	1.1		3.1	0.9
Crow	3.7	0.8		2.0	0.8		3.7	0.8
S.D. means standard deviation								

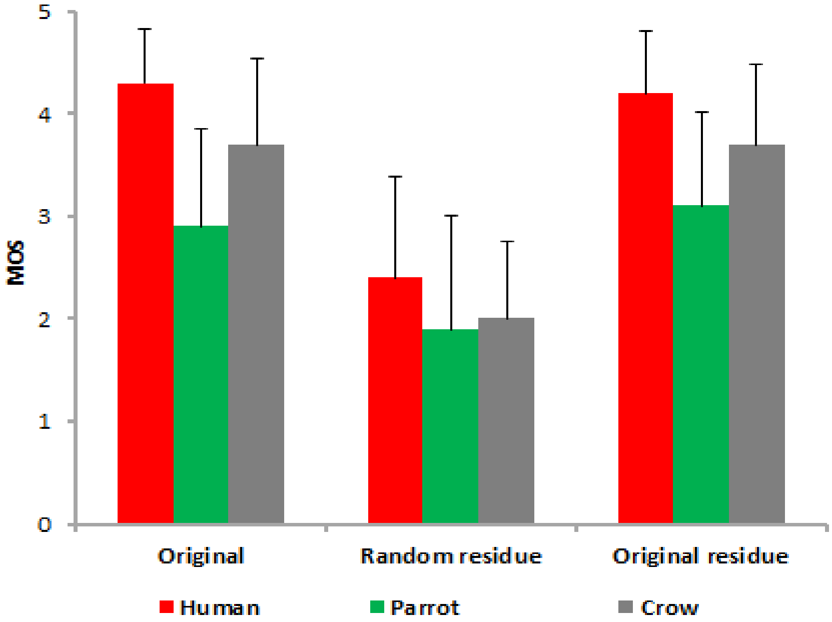

Fig. 10. Plot of histograms for Table 4. The height of the histograms is proportional to the MOS and the SD is also shown on the top of the corresponding histograms.

References

[1] D. O'Saughnessy, *Speech Communication - Human and Machine*, New York: Addison-Wesley, 1987.

[2] C. K. Catchpole, P. J. B Slater, *Bird Song: Biological Themes and Variations*, Cambridge University Press, Cambridge, UK, 1995.

[3] A. Kumar, "Acoustic communication in birds: Differences in songs and calls, their production and biological significance," *Resonance*, vol. 8, pp. 44-55, 2003.

[4] R. Singh, A. Kumar, P. K. Lehana, "Investigations of the Quality of Speech Imitated by Alexandrine Parrot (Psittacula eupatria)," *Circuits Syst. Signal Process.*, vol.36, pp. 2292–2314, 2017.

[5] R. Singh, A. Kumar, P. K. Lehana, "Effect of bandwidth modifications on the quality of speech imitated by Alexandrine and Indian Ringneck parrots," *Int. J. Speech Technol.*, 2017, doi: 10.1007/s10772-017-9437-x.

[6] S. Nowicki, "Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere," *Nature*, vol. 325, pp. 53-55, 1987.

[7] B. S. Nelson, G. J. L. Beckers, and R. A. Suthers, "Vocal tract filtering and sound radiation in a songbird," *J. Exp. Biol*., vol. 208, pp. 297-308, 2005.

[8] R. A. Suthers and S. A. Zollinger, "Producing song: the vocal apparatus," *Ann. NY Acad. Sci.*, vol. 1016, pp. 109-129, 2004.

[9] A.S. King, *Functional anatomy of the syrinx, in Form and Function in Birds*. ed. by A.S. King, J. McLelland. London: Academic Press, 1989.

[10] G. J. L. Beckers, "Bird speech perception and vocal production: a comparison with humans," *Hum. Biol*., vol. 83, pp. 191-212, 2011.

[11] A. J. Doupe and P. K. Kuhl, "Birdsong and human speech: common themes and mechanisms," *Annu. Rev. Neurosci*., vol. 22, pp. 567-631, 1999.

[12] V. R. Ohms, P. Escudero, K. Lammers, and C. ten Cate, "Zebra finches and Dutch adults exhibit the same cue weighting bias in vowel perception," *Anim. Cog*. vol. 15, pp. 155-161, 2011.

[13] W. T. Fitch, "The evolution of language: a comparative review," *Biology and Philosophy*, vol. 20, pp. 193–230, 2005.

[14] D. R. Chamberlain and G. W. Cornwell, "Selected vocalizations of the Common Crow," *The Auk*, vol. 88, pp. 613–634, 1971.

[15] A. O. Gross, Eastern Crow, in Life Histories of North American Jays, Crows, and Titmice. in AC Bent (partial author and ed) United States National Museum, Smithsonian Institution, Bulletin 191, Washington, District of Columbia, pp. 226–259, 1946.

[16] S. Ali, *The Book of Indian Birds*, Bombay: The Bombay Natural History Society, 1943.

[17] J.M. Forshaw, *Parrots of the World: An Identification Guide*, Princeton: Princeton University Press, 2006.

[18] S. Sindel, "The alexandrine parakeet, Psittacula eupatria (linee)," *the Journal of the Avicultural Society of New South Wales.*

[19] *Wildlife Crime Control Bureau, A Manual on Wildlife Species in Trade*, Ministry of Environment and Forests, Government of India, New Delhi, 2011.

[20] S Ali and S. D. Ripley, *Handbook of the Birds of India and Pakistan, together with those of Bangladesh, Nepal, Bhutan and Sri Lanka*, Bombay, London: Oxford University Press, 1969

[21] P. B. Michael and W. Michael, "Nestling Development of Ring-necked Parakeets (*Psittacula krameri*) in a Nest Box Population," *The Open Ornithology Journal*, vol. 6, pp. 9-24, 2013.

[22] G. Patra and S. Chakrabarti, Avian Diversity in and around Digha, District—East Midnapore (West Bengal, India), *Advances in Bioscience and Biotechnology*, vol. 5, pp. 596-602, 2014.

[23] T. Reaume, *The American Crow, Naturally*, An illustrated ebook, August, 2013.

[24] S. Ali and S. D. Ripely, *Compact Handbook of Birds of India and Pakistan*, Bombay: Oxford University Press, 1987.

[25] D. Goodwin, Crows of the world, London: British Museum, Natural History, 1976.

[26] "House crow: animal pest alert". agric.wa.gov. Government of Western Australia: Department of Agriculture and Food. Retrieved 30 October 2015.

[27] GISD, Global Invasive Species Database of the IUCN/ISSG (Invasive Species Specialist Group of the World Conservation Union). Fact sheet on Corvus splendens. [Online]. Available: http://www.issg/database.

[28] Department of Agriculture and Food, Animal Pest Alert: House Crow. National Animal Pest Alert Information Sheet 2. Department of Agriculture and Food, Perth, 2008.

[29] R. Singh, A. Kumar, G. Singh, and P. K. Lehana, "Estimation of phonemes in the calls of crows using vector quantization," *International Journal of Computational Linguistics and Natural Language Processing*, vol. 3, pp. 565-571, 2014.

[30] C. Ryall, "Predation and harassment of native bird species by the Indian house crow corvus splendens in Mombasa, Kenya," *Scopus*, vol. 16, no. 1, pp. 1-8, 1992.

Randhir Singh, Ajay Kumar, and Parveen Kumar Lehana, "Investigating the quality of speech of birds using linear predictive coding," *International Journal of Scientific and Technical Advancements*, Volume 3, Issue 3, pp. 1-12, 2017.

[31] C. Ryall, "The pest status of the Indian house crow Corvus splendens in Mombasa and a survey of its expansion of range in coastal Kenya," in *Proceedings 7th Pan African Ornithological Congress,* L. Bennun, Ed., Nairobi, pp. 303-310, Aug. 1988.

[32] C. Ryall, "Further records of range extension in the house crow corvus splendens," *BOC Bulletin*, vol. 122, no. 3, pp. 231-240, 2002.

[33] C. Ryall, "Further records and updates of range extension in house crow corvus splendens," *BOC Bullentin*, vol.136(1), pp. 39-45, 2016

[34] H. A. Khan, "Damage patterns of house crow (Corvus splendens) on some food crops in Faisalabad, Pakistan," *Int. J. Agric. Biol.* vol. 4, no 4, pp. 500–502. 2002.

[35] A. Nyari, C. Ryall, and A. T. Peterson, "Global invasive potential of the house crow Corvus splendens based on ecological niche modelling," *Journal of Avian Biology,* vol. 37, pp. 306–311, 2006.

[36] M. S. Dhindsa, P. S. Sandhu, H. K. Saini, and H. S. Toor, "House crow damage to sprouting sunflower," *Tropical Pest Management*, vol. 37, no. 2, pp. 179-181, 1991.

[37] C. J. Feare and Y. Mungroo, "The status and management of the house crow corvus splendens (Vieillot) in mauritius," *Biological Conservation,* vol. 51 pp. 63-70, 1990.

[38] J. E. Cooper, "Health studies on the Indian house crow (Corvus splendens), Avian Pathology," vol. 25, no. 2, pp. 381-386, 1996.

[39] P. Roy, A.T. Venugopalan, and R. Manvell, "Isolation of Newcastle disease virus from an Indian house crow," *Tropical Animal Health and Production*, vol. 30, no. 3, pp. 177-178, 1998.

[40] A. M. Kamel, "Potential Impacts of Invasive House Crows (Corvus Splendens) Bird Species in Ismailia Governorate, Egypt: Ecology, Control and Risk Management," *Journal of Life Sciences and Technologies,* vol. 2, no. 2, 2014

[41] Alien Crows Targeted for Total Extermination in South Africa. Availableonline:http://news.nationalgeographic.com/news/2006/05/060509_crows.html.

[42] D. L. Fraser, G. Aguilar, W. Nagle, M. Galbraith, and C. Ryall, The House Crow (Corvus splendens): A Threat to New Zealand?, *ISPRS Int. J. Geo-Inf.*, vol. 4, pp. 725-740, 2015.

[43] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall Signal Processing Series. Prentice-Hall Inc., Engelwood Cliffs, New Jersey, 1978.

[44] B. S. Atal and M. R. Schroeder. "Predictive coding of speech signals", *in proceedings of 1967 IEEE Conf, Communication Processing*, pp. 360–361.

[45] S.D Gunjal and R.D Raut," Advance source coding techniques for audio speech signal: A survey," *International journal of computer technology and applications*, vol. 3, no. 4, pp. 1335-1342, 2012.

[46] M. M. Sondhi, "New Methods of Pitch Extraction," *IEEE Trans. Audio and Electroacoustics*, vol. AU-16, no. 2, pp. 262-266, 1968.

[47] B. Atal and L. Hanauer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," *J. Acoust. Soc. Am*., vol. 50, no. 2, pp. 637-655, 1971.

[48] J. Hai, M. Joo, Improved Linear Predictive Coding Method for Speech Recognition, *in proceedings of ICICS-PCM*, Singapore, pp. 15-18, December 2003.

[49] J. Bradbury "Linear predictive coding" December 5, 2000.

[50] K. Paliwal and W. Kleijn, "Quantization of LPC parameters," *Speech Coding and Synthesis*, pp. 433-466, 1995.

[51] L. Rabiner, and B. H. Juang, *Fundamentals of speech recognition*, New Jersy: Prentice Hall, 1993.

[52] M. N. Raja, P. R. Jangid, S. M. Gulhane, "Linear Predictive Coding," *International journal of engineering sciences & research technology,* vol. 4, no. 4, pp. 373-379, 2015.

[53] T. Ganchev, A. Lazaridis, I. Mporas, and N. Fakotakis, *Performance evaluation for voice conversion systems*, Berlin: Springer, 2008.

[54] ITU-T, Methods for subjective determination of transmission quality, Tech. Rep. ITU-T Recommendation P.800, ITU, 1996.

[55] Z. Shuang, F. Meng, and Y. Qin, "Voice conversion by combining frequency warping with unit selection," *in Proceedings IEEE Int. Conf. Acoust., Speech, Signal Process*., Las Vegas, NV, pp. 4661-4664, 2008.

[56] ITU-T Rec. P. 862, Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow band telephone networks and speech codecs, 2001.

[57] A. W. Rix, M. P. Hollier, The perceptual analysis measurement system for robust end-to-end speech quality assessment, *in proceedings IEEE Int. Conf. Acoust., Speech, Signal Process*., Istanbul, Turkey, pp. 1515-1518, 2000.

Randhir Singh, Ajay Kumar, and Parveen Kumar Lehana, "Investigating the quality of speech of birds using linear predictive coding," *International Journal of Scientific and Technical Advancements*, Volume 3, Issue 3, pp. 1-12, 2017.