SOME NOTES ON TALKING BIRDS

by   Hendrik Mol

INTRODUCTION

" A little bird told me

.... "

One of the topics of phonetics, if not its main topic , is the
study of the mechanism of speech and hearing. On the one hand , the
phonetician tries to develop an acoustic theory of speech production
explaining how articulatory and vocal activities are being trans-
formed into acoustic vibrations of the air particles. On the other
hand, he tries to discover how the acoustic vibrations reaching the
ear of the listener set up patterns of nervous activity in the
fibres of the acoustic nerve and the rest of the nervous system. The
mechanism of speech and hearing in man is based on the fact that a
speaker is able to produce articulatory cues the acoustic translat-
ions of which can be detected as acoustic cues by the nervous sys-
tem of the listener. The latter is conditioned to detecting these
acoustic cues. The task of the phonetician is to discover and study
the cues mentioned above; he should not allow himself to be carried
away by the strong tendency of the linguist to bundle these cues for
the purposes of writing or phonemic abstraction.

No doubt talking birds base their imitations on the acoustic
cues they receive and of which they reproduce fair caricatures. They
cannot be expected to bother with the concept of the phoneme. Neither
can they be expected to treat speech in a way that differs from the
manner in which they process imitable sounds like whistling, barking,
coughing etc. We think we might learn something from the birds in
our endeavours to break the code of speech transmission because the
sound producing apparatus of the bird differs so fundamentally from
that of Man, whereas its ear seems to be built in a less complicated
manner than the human auditory system.

It is the purpose of this paper to focus the attention on some
fundamental differences between talking birds and talking men.

DIFFERENCES BETWEEN THE ORGANS OF HEARING

The human ear displays the well-known microphonic construction of an auricle, an external meatus , an eardrum , three ossicles, two aural muscles and a 33 mm long coiled cochlea containing both inner and outer hair cells.

The parrot misses the auricle, has only one ossicle and only one muscle, whereas its 3 mm long rudimentary cochlea , too short to be called coiled, shows no distinction between inner and outer hair cells.

In spite of this greater simplicity the ear of the talking bird is able to process speech waves so effectively that the bird is in a position to render fair imitations of human talkers.

This proves that the complexity of the mammalian ear is not a conditio sine qua non for the development of speech. In the avian ear a mechanical frequency analysis seems less probable. Therefore we may expect that the ear of the bird works in the time domain. As a matter of fact we only recently discovered that the ear models [1] described by us in 1959 were models of the ear of the bird rather than models of the mammalian ear. Nevertheless, we still believe that part of the human ear, to wit the 'parrot' zone near the stapes, really behaves like the avian ear [2]. The spatial development of the mammalian ear should not be seen as a typical development towards the acquisition of speech.

[1] H. Mol and E.M. Uhlenbeck, Hearing and the concept of the phoneme , Lingua VII, 2 (1959) , p. 161 - 185 .

[2] H. Mol, Fundamentals of Phonetics, I : The organ of hearing, (1963) , The Hague, Mouton .

DIFFERENCES BETWEEN THE ORGANS OF SPEECH

During vowel production the human vocal tract may be considered as a slender tube open at the mouth side and closed at the throat side. The vibrating vocal folds alternatively open and close the end of the wind pipe, in that way cutting the expelled breath stream into a continuous series of air puffs. In other words: the larynx is analogous to a machine gun shooting gaseous bullets. Under normal conditions an air puff starts off gradually whereas it is cut off rather sharply at its end; it is this closing snap that sets up a collection of powerful damped oscillations in the vocal tract. Each damped oscillation , defined as a formant here, is characterized by its own initial amplitude, its own rate of decay and its own frequency. Formants are normally ranked according to their frequencies: F_1 , F_2 , F_3 , F_4 etc, F_1 pertaining to the formant with the lowest frequency. When a closing snap of the vocal folds hits the vocal tract all formants F_1 , F_2 , F_3 etc. pop up and are superimposed one on another. Quite often these superimposed damped oscillations are clearly visible in the oscillogram, at least to the experienced and willing eye.

In the special case where the cross-area of the vocal tract is constant, or practically so, the formants are approximately:

$$F_1 \approx \frac{c}{4\,l} \quad , \quad F_2 = 3\,\frac{c}{4\,l} \quad , \quad F_3 = 5\,\frac{c}{4\,l} \quad \text{etc.}$$

In these formulas c represents the velocity of sound of the warm air in the vocal tract, amounting to some 35000 m/s , whereas l stands for the length of the tract as measured between the mouth opening and the vocal folds.

The above formulas describe the typical resonances of the organ pipe closed at one end. For the human vocal tract this is the neutral position called the ' schwa ' .

The vocal tract, however, can do more than merely produce the schwa : its ' selling point ' is , that it is able to introduce a rather abrupt step in its cross-area and to place this step at different positions between the mouth opening and the vocal folds .

The so-called twin-tube model [1] is the most simple multiple-tube
model providing such a step : it is thought to consist of two tubes
in cascade, to wit the mouth tube in front (length l_1 , cross-
area S_1) and the throat tube in back (length l_2 , cross-area S_2),
meeting each other at the step $S_2 - S_2$ or formulated as a
quotient , $\dfrac{S_1}{S_2}$.

The corresponding formant formula of the twin-tube model may
be written as follows :

$$\cos 2\pi F \frac{l_1 + l_2}{c} = \frac{1 - \dfrac{S_1}{S_2}}{1 + \dfrac{S_1}{S_2}} \cos 2\pi F \frac{l_1 - l_2}{c}$$

Simple though the twin-tube model may be, it illustrates how , through
adjustment of its three parameters l_1 , l_2 and $\dfrac{S_1}{S_2}$ that are

inspired by the vocal tract, its formants cover the same wide range
as the formants of human talkers do for the different vowels of their
inventory.

For the covering of this range the talking bird is in a less
favourable position. Because its larynx does not contain the suitable
vibrating parts it cannot use this lung-protecting device for driving
its vocal tract: for this purpose it resorts to its syrinx, a sort of
whistle located at the spot where the trachea bifurcates to produce
the bronchi. In the parrot the lowest ring of the trachea has acquired
two oblong lateral apertures, each of which has been covered by a
membrane. Three pairs of muscles control the action of the syrinx of
the parrot. The vibrating membranes rhythmically change the cross-area
of the trachea in that way modulating the breath stream.

Our tape recordings of parrots and budgerigars indicate, that
these birds are able to produce practically sinusoidal oscillations as

[1] H. Mol, Fundamentals of phonetics, II: Acoustical models generating
the formants of the vowel phonemes, (1969) Mouton, The Hague -
Paris.

well as damped oscillations. The latter type of sound requires a
resonator, in this case the trachea, driven by sharp-edged air puffs.
Oscillograms reveal that, quite often, the puffs are symmetrical show-
ing an opening snap as well as a closing snap, so that the vocal tract
is excited twice during each vocal period, a condition sometimes met
in pathological human voices. Moreover, it strikes the eye that the
duration of the air puff may change continuously from period to period,
in that way impressing an irregular character on the oscillogram.

Essentially, the trachea of the bird behaves like a tube with con-
stant cross-area closed at one end. It is a degenerated twin-tube with
$S_1 = S_2$. As mentioned before, the formants of such a tube are odd
multiples of $\dfrac{c}{4.1}$. For the bird it is a severe handicap that ,
though it is able to adjust the length 1 of its trachea within wide
limits by muscular activity , the frequency relation between these for-
mants is fixed. Consequently , we may expect the parrot to take refuge
in tricks in order to produce p e r i o d i c i t i e s   that are fair
caricatures of the human formants. We studied the oscillograms of a
specimen of Psittacus erithacus erithacus , the well-known grey parrot
with a short red tail, and found, for instance, this swindler imitated
F_2 of a human [ə] as a damped oscillation with a frequency correspon-
ding to the first formant of its trachea . F_1 of human [ə] was com-
pletely omitted; as it were the bird ' speculated ' that the human
listener would take its F_0 , that is the frequency of the air puffs
of its syrinx, for the missing F_1 . In the Dutch word ' koekoek '
(phonetic transcription [kukuk]) , meaning cuckoo , the same bird
produced excellent [k]'s , but realized the [u]'s as practically
sinusoidal oscillations at two different frequencies corresponding with
the appropriate intonation.

Another subject , our hand-reared budgerigar 'Peter' , echoed
human [i] by delivering F_2 but omitting F_1 .

Though the larynx of the bird plays no role in vowel production
we suspect it to be active in the production of fricatives and stops.

ABSTRACTION AND IMITATION

The traditional loose formulation of how a child learns to talk
is to say that it 'imitates ' the grown-ups in its environment.
Because its vocal tract is too short, however, it is not able to
reproduce the absolute ' adult ' formant positions by far. Also ,
its larynx does not permit it to imitate the much lower pitch of
adults. In the long run the child discovers that the grown-ups use
a systematic set of contrasts between the vowels they produce. The
child gradually learns to reproduce these contrasts within the phys-
ical constraints of its youthful, short vocal tract. As time goes by
the axial dimensions of its vocal tract increase by growth, in that
way forcing down all formant frequencies by the same factor. The
child does not and indeed cannot arrest this shift by changing the
muscular commands it issues to its articulators. On the other hand ,
there is no ' phonemic ' reason for trying to resist the formant
shift because the latter does not incapacitate the system of con-
trasts that ' carries ' speech communication . We can say that the
child learns to talk in its own voice and with his own formants. In
learning to talk it abstracts rather than imitates.

On the other hand, the talking bird with its versatile syrinx is
a master in reproducing the frequency of the glottal air puffs of its
example , be it female of male.

Also, the bird imitates the formants of its human example, as
accurate as possible in so far they are within its possibilities and
are not stealthily omitted. It gives a real, though partial specific-
ation of a particular human talker. It reacts like a grammophone
record; it imitates a squeaking door with equal zest as human speech.
It cannot be expected to use the phonemes, those products of human
abstraction, as functional units in this process. We might as well
doubt whether Man uses the phonemes as functional units in the mechan-
ism of speech and hearing. Without in the least pretending that we
have definitely exposed the phoneme as a ' mere ' abstraction we
think the research on talking birds throws reasonable doubt on the
uniqueness of the phonemic (= alphabetic) approach.

It stands to reason that the talking bird (and also the human talker) can only reproduce acoustic cues it can notice via its auditory system and, moreover, is able to produce with its organs of speech.

Oscillograms show the bird is a typical second formant generator. This raises interesting questions such as: does the parrot ' notice ' the first formant (F_1) as well as the second formant (F_2) but limits itself to reproducing F_2 because it has no cavities to reproduce the low F_1 anyhow ? Or, is this statement of the problem too anthropomorphic and is the auditory system of the bird a typical F_2 detector that differentiates the sound signal with respect to time , in that way automatically ' drowning ' F_1 ?

Is it a lucky coincidence that the bird can only ' notice ' the formants its slide-tube vocal tract is able to produce ?

WHY DO BIRDS TALK ?

Leaving in mid-air why man talks , we can ask ourselves why birds talk. It is open to doubt [*] whether talking birds in their natural environment do imitate other animals. There must be some other reason for talking birds to echo human speech.

In order to improve our insight we bred budgerigars in our institute, separating the young from their mother as soon as they could be handreared. The little silent birds attached themselves to the person who fed and nursed them. In these conditions they started to imitate the words spoken to them.

I remember sticking my finger into the cage in order to caress a budgerigar. To my surprise he started ' feeding ' from his crop the nail of my index-finger, apparently taking me for a budgerigar. When flying around freely he used to land on my head and start pulling hairs through his beak as a token of intimacy . These phenomena led me to the following hypothesis:

Talking birds are born without wired-in programs for controlling their vocal organs. They start life with the inborn instruction to imitate the sounds they hear from those who nurse them, usually their parents in the nest. When this role is taken over by man the small birds automatically switch over to the (speech)sounds of man. In other

[*] Sebeok , private communication

words, birds talk by mistake.

Quite interestingly, we never succeeded in teaching a female budgerigar to talk. Male budgerigars may be recognized by the blue caps around their nostrils.

SPIN-OFF OF THE STUDY OF TALKING BIRDS.

Study of the oscillograms of talking birds is essential to the study of speaker specification and identification, because the bird produces a partial specification of the talker that is accepted by human listeners.

It is worthwhile to consider the talking birds as a pathological human talker because his organs of speech differ so fundamentally from those of its human colleague. The experience gained in studying the oscillograms of talking birds may be applied with advantage to the interpretation of the oscillograms pertaining to pathological human voices. In that way the study of the voice of a queer bird , the parrot, may be beneficial to the study of the voice of another queer bird : Man .

Hendrik Mol
Institute of Phonetic Sciences
of the University of Amsterdam
Herengracht 338
A M S T E R D A M