A Cybernetic Study of Speaking and Singing

A Cybernetic Study of Speaking and Singing

By

Ion Piso

Translated by Ligia Tomoiagă

Cambridge
Scholars
Publishing

A Cybernetic Study of Speaking and Singing

By Ion Piso
Translated by Ligia Tomoiagă

This book first published 2017

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-4438-8923-7
ISBN (13): 978-1-4438-8923-0

CONTENTS

ACKNOWLEDGEMENTS

I would like to express my very special thanks to Dr. med. Rodica Mureşanu, from the Cluj laboratories, where I conducted my experiments. She was extremely helpful and proved to be both a valuable professional, and an art lover.

I would also like to thank Mrs. Karen Richards, who was our team's English Language Advisor, and who helped with proofreading, editing and gave advice on the readability of our translation. I am most grateful for the work she has done.

INTRODUCTION

MOTTO: ARS SINE SCIENTIA NIHIL
(A DICTUM BELONGING TO OLD OPERATIVE MASONS)

This study arises from my wish to elucidate further, the differences between speaking and singing, as a cybernetic phenomenon (cf. Wiener). That is, the phenomena of circularity (cf. St. Odobleja); also known as reversed connection or feedback. The differences between singing and speaking lie in the intimate mechanisms that produce speaking on the one hand, and singing, on the other; they are two physiological functions that are different, but also quite similar, in many respects.

They are *similar* in that they are both the *product* of the same phonating apparatus; they are also similar because, in many languages and dialects, speech has a melodic intonation and because there are only a few cases in which vocal music does not use words.

They are *different* because not all languages are spoken with a melodic vocal intonation; moreover, one can sing without words, as in the case of cadenzas (or pacing), of colouraturas (or grace-notes), and when someone vocalizes.

A more detailed analysis of the physiology of speaking and singing shows other defining characteristics, which help us better understand that they are not one and the same thing. For example, with some notable exceptions – as in languages such as Welsh and German, where some words are constructed entirely of a string of consonants - the spoken word is articulated as a blend of both consonants and vowels. Singing, on the other hand relies almost entirely on the ‘pure’ articulation of the vowel, to convey meaning.

Further, in speech, when we pronounce vowels, the sonority and colour of the voice is limited to the harmonics of those vowels. This is called *vocal harmonics* – that is, the frequency of the sound used in speech, which is lower-pitched with men, and higher-pitched with women. Thus, the production of each vowel is the result of a certain frequency that characterizes it (see Helmholtz’s analysis below), which also accounts for the different place in the oral cavity where it is produced, and where these frequencies are born; for instance, for [a] it is backwards, for [e] it is frontal according to the volume of the resonance cavity specific to each vowel.

In singing, though, and I am referring to *professional, performance singing* in which the sonority of the voice is not limited to the vowel harmonics, the main weight is placed on the *extra-vowel harmonics* of the fundamental sound, which takes place in the resonance cavities: the pharynx and the oral cavity. A special vocal technique harnesses and develops the sonorous signals in these resonance cavities. They are independent of the specificity and colour of the particular vowels. Consequently, besides the vowel harmonics, in singing, the extra-vowel harmonics appear and prevail. These extra-vowel harmonics contribute greatly in obtaining a remarkable sonority (without relying on electronic amplification through the microphone), both in terms of intensity and richness (amplitude) of the sound, and in terms of the quality of the sonority. As we will see during this analysis, the obtaining of extra-vowel harmonics depends on the achievement of a certain muscle-abdominal-diaphragm tension, by using a special technique, which helps us to maintain better control, and a reduction of the quantity of air during the sonorous exhalation. Moreover, the technique entails a conscious but not *direct* opposition between the antagonistic muscles of the larynx, the inner cricothyroids and the thyroartenoids (see the Traissac tables – see note below).

In conclusion, we might say that the phonating apparatus allows two usages, in two variants, according to the proportion between the two types of harmonics – vowel and extra vowel ones. Insomuch as the number of the extra vowel ones increase, speaking becomes singing, and when they decrease, we get closer and closer to speaking, when almost only the vowel harmonics remain.

The fact that the apparatus is *bi-functional* is not obvious at first glance; the differences appear in the specific manner in which each of the two functions: for speaking in one way, and for professional singing in an entirely different way[1]. This bi-functionality will be demonstrated in this study, with the aid of the most advanced investigating methods that can be applied to sonorous frequencies, and presented graphically by digital means, that is, by the analysis of the spectrum of the sonorous vibrations of vowels, as well as that of consonants – the oscillograms of these sonorous signals.

*

[1] See I. Piso, *Criza Operei*, Bucuresti 2015, or I. Piso, *The Crisis of the Opera?*, CSP, 2013.

The study of the various harmonic frequencies – some specific to vowels, other to consonants – also opens up, in a surprising manner, the possibility of an excursion into the world of languages, idioms and dialects.

By analysing the inherent mechanism of producing sonorous signals by the phonating apparatus and the reception of these harmonic vibrations by the auditory apparatus - called the phenomenon of *circularity and reversed connection* (Odobleja), or *feedback* (Wiener), on the general basis of reversibility, of reversed connection between the two apparatuses, also currently called *Cybernetics*[2], I reached the conclusion that even if the genesis of dialects and idioms could not be explained by it, we could at least justify or clarify their actual being. From the point of view of the frequency of certain sonorous signals, and, implicitly their penetration force, a dialect appears as an *alternative* to that specific *official* language, as it strives to enhance the *efficiency* of communication.

The degree of *penetration* of sonorous signals, in speech, clarifies to a certain extent, the existence of idioms, as well as of dialects, and justifies the way the language is spoken. In other words, in transmitting and receiving articulate acoustic signals, the strictly pragmatic point of view (of efficiency) that certain idioms and dialects adopted, always played the most important part, becoming a sort of *critique*, thus leading to a bettering (I am referring to dialects) of that specific *official* language. Ultimately, the purpose of dialects does not consist in anything but the communicator's initial intention to convey meaning, as efficiently, clearly and as penetratingly as possible, using minimal effort. Idioms also follow

[2] See Jon Piso, *Cibernetica fonatiei in canto/ The Cybernetics of Phonation in Canto*, Bucharest, 2000 (a Doctoral Thesis that was awarded the appreciation *magna cum laude*), about the cybernetic phenomena that are due to *feed-back* (according to Norbert Wiener – 1948), or to *circularity by reversed connection* (according to Stefan Odobleja, published also in Paris, like Wiener's, but 10 years earlier, in 1938). Here is a short explanation of the essence of the neuro-circularity or neuro-cybernetic mechanisms (from the Greek *Kybernesis*). When I emit through the phonating apparatus a sonorous signal (a sound), as a result of a command from of the brain, the signal reaches instantaneously to the auditory apparatus, which verifies it. Almost at the very same time, the brain centre of the auditory apparatus decides whether the sound corresponds to the exigencies that it (the brain) requires. If the appreciation is positive, the phonating centre receives the *All Right* on the spot, and the action continues without modifications. If not, the brain will send a new command to the phonating apparatus, which contains the modification required by the auditory apparatus. Thus, this is the so-called *reversed connection,* or *feedback* – which can be reduced to the formula command and control of command.

the same route, in the preferences with which they use sonorous signals of certain frequencies. The phenomenon is mostly obvious in the case of languages that use an etymological spelling. This is why orthoepy has a correcting function as to sonorous efficiency; for instance, in English, the word *information* is pronounced differently to the French, /ɪnfəˈmeɪʃn/, from which it comes. The word in French ends with the sonorous signal /on/ whose harmonics have a frequency around 110 Hz, while the harmonics of the consonant /ʃ/, preceded by /i/ at the end of /ɪnfəˈmeɪʃn/ are spread between 1500 and 6000 Hz; consequently, the power of penetration of the word, and its chance to be better received, is much enhanced in English orthoepy.

At the same time, in this word that is spelled *information*, the vowel /o/, when pronounced in English, is replaced with the vowel /ə/, for the following pragmatic reasons. After the fricative /f/ (see the table with consonants in the Romanian version, below), which is pronounced in the front of the mouth (around the lips) by the expulsion of a large quantity of air (during exhalation), the cybernetic mechanism does not command a return from the front articulation of the fricative /f/ to the vowel /o/ (as in French); this happens because, in order to pronounce /o/ the formation of a *tube* is necessary, which starts at the back of the oral cavity, behind the hard palate[3], where formant I ends[4]; this tube is prolonged through the alveolar region (the teeth) in the oral cavity, where the lips are pushed to take the round form for O, by the activation of the oral sphincter; thus, in English orthoepy, the phoneme /o/ is replaced by the phoneme /ə/, which is formed immediately behind the phoneme /e/[5], so it is much closer to the articulation point of the fricative /f/. We will also note that the articulation of /ə/ takes place by means of a slight retraction and lowering of the tongue (similar to the manner in which initial /i/ is pronounced). This place where /ə/ is produced benefits from this vicinity, which has enhanced resonance, as the sound is directly transmitted through the hard palate to the jaw sinuses, thus realizing a spontaneous amplification of that particular sonorous signal. By the simple replacement of /o/ with /ə/, an important economy in cybernetic circuits is realized, and commands that would have been followed by many muscle modifications, in the frontal part of the phonating apparatus (the oral cavity and the oral sphincter), also commanded by cybernetic circuits, are no longer necessary.

The same mechanism takes place in the transformation of the syllable *ma* in the English orthoepic system, to pronounce /ma/ as in the French

[3] See the first figure at page
[4] See Guth's investigations at page
[5] See the first figure at page

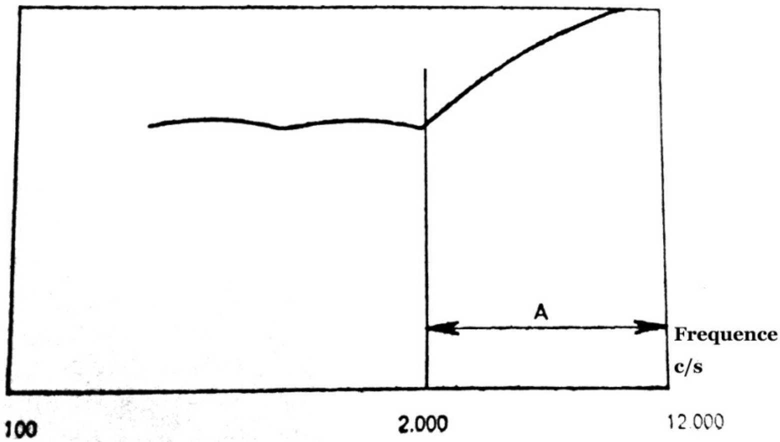
pronunciation, for instance, after the obstruction that is formed by the phonic nasal /m/, there appears a maximum opening of the oral pavilion, by a lowering of the mandible, so that the phoneme /a/ can be produced; a vowel with a penetrating sound of only 220 Hz. The English pronunciation, though, manages to by-pass the phonetic *dead end* that is caused by the low penetration of /a/ – which also requires many cybernetic commands, using the diphthong /ei/; this is an ingenious solution, because the result is surprisingly penetrating. The frequencies of the vocalic harmonic, of the two, are as follows: 880 Hz for /e/, and 2347 for /i/. Thus, the diphthong possesses body, sonority and penetration, due to the articulation of /e/, which is placed around the hard palate, frontally (thus amplified by the sinuses, as we have shown above); this sonority is ended by /i/, the vowel which is the most penetrating, due to its frequency of 2347 Hz, and to its harmonic peaks that go over 90 dB at 2800 Hz, as well as between 3400-3700 Hz (see Chart no. 1).

As for consonant /t/, which follows after the syllable *ma* in the word *information*, it is transformed in the dental fricative /s/ in French, which lacks sonorous relief (see Chart no. 8), and has harmonics between 200 Hz and 4000 Hz, with an intensity much lower than 60 dB (an intensity that is not sufficient for the human ear to hear). English orthoepy replaces the dental fricative /s/, though, with the sibilant /ʃ/, with harmonics between 1500 Hz, to 6000 Hz, which even has a harmonic with a frequency of 2200 Hz of an intensity of over 90 dB (see chart 9). Due to the increased level of penetration of this sibilant, the human auditory apparatus receives it in an optimum manner.

At this point, it is important to note, that when the vowels /a/ and /o/ are sounded, within a word and the speaker wishes to make himself/herself better understood, the tendency is to speak more loudly, thus emitting them in forte (the tenor Alfredo Kraus, maintained that vowel /a/ is a muckworm of harmonics). This solution proves to be counterproductive, however, as the intensity of these vowels, when emitted more powerfully, *shadows* the neighbouring consonants, thus the clarity of communication decreases, instead of increasing. I surmise that this is the reason why English reduces the frequency of these vowels as much as possible, as a precautionary measure, as we will see later on in this study.

Regarding vowel /a/, we must see that its use within the English orthoepy, especially when it precedes vowel /i/, leads to an increase in the efficiency of sonority, that is, of communication. In order that we may explain this, I will take as an example two words: the first is the personal pronoun *I*, which is frequently used; the other is the frequently pronounced word, *China*. In both cases, English orthoepy places an /a/ before the vowel /i/, on the intention of increasing the sonorous efficiency of this vowel. If we look at the oscillogram of /i/, (chart no. 1), we notice that between 800 Hz and 2200 Hz, sonorous signals do not appear on the oscillogram, due to their insignificant intensity. It is known that in normal conversation, sonorous signals that have a frequency between 1000 Hz and 2000 Hz are optimal for human hearing; still, the vowel /i/ is not perceived clearly, between 1000 and 2000 Hz. To facilitate the clear reception of a word containing the vowel /i/, the individual who articulates the word, should increase the intensity of the sonorous signal, and raise his/her voice, which is neither appropriate, nor socially acceptable. Imagine a conversation, in which each interlocutor emits, from time to time, an aggressive exclamation, with an intensity above the norm, and with a high shrieking frequency, in order to make sure that the other receives the message correctly. Therefore, by placing vowel /a/ before /i/, the English speaker, finds the optimum solution. In this way, the diphthong /ai/ also has a body, and maintains the necessary pithiness, which leads to good reception, because the spoken vowel /a/ has harmonics between 0 and 3500, and intensities well above 60 dB (see chart no. 5).

If we come back to the word *information*, as it is pronounced in English /ɪn.fəˈmeɪ.ʃən/, instead of getting *drowned* at its end with the non-penetrating, almost deafening sound /on/ (110 Hz) as in French, it increases in frequency and resonates with a particularly penetrating force. As A. Tomatis shows, in his study *L'oreille et la vie*, in the *ethnogram*, native English speakers emit sonorous sounds of the harmonics with high frequency, which can move from 2000 Hz to over 12000 Hz (116). It is known that on the basis of cybernetic circuit control and command, only those sonorous signals are emitted, which the auditory apparatus of the speaker can hear.

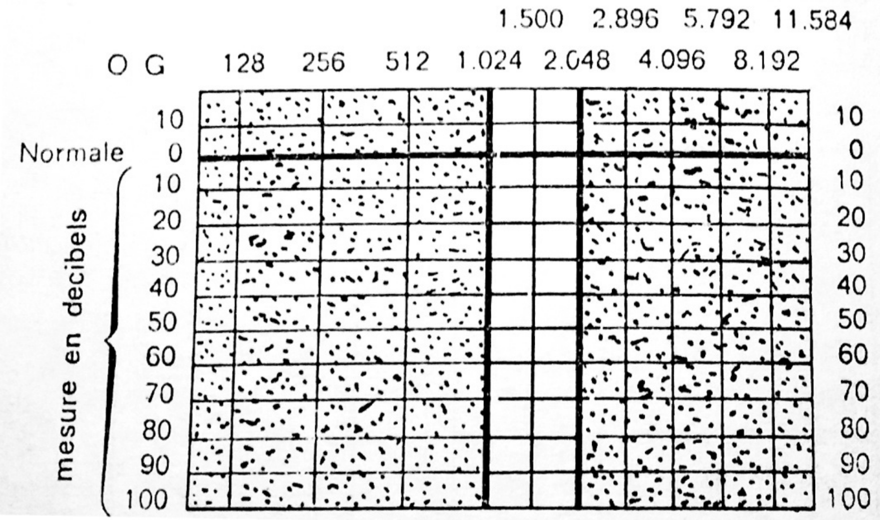


The impressive curve of audibility, from 2000 Hz to above 12000Hz, of native English Speakers
(According to A. Tomatis)

In his own study, Dr. Tomatis quotes the great tenor, Benjamino Gigili, when commenting on why English circumvents the French orthoepy: "J'ai remarqué combien vous avez de belles voix chez vous en France mais je n'ai jamais compris pourquoi les Français chantent en donnant l'impression d'être à l'octave au-dessous." / I remarked what beautiful voices you have in France, though I have never understood why, when a French singer sings, I always get the feeling that he is one octave lower"[6]. Obviously, Gigli did not say that a French mezzo-soprano has the ambit of a baritone, and sings with a voice which is one octave lower, rather he is suggesting that the harmonics of the French language, which are cybernetically dependent on the selectivity of the French ear, can rarely go above 2000 Hz[7].

---

[6] A. Tomatis, *L'oreille et le vie*, p. 118
[7] See I. Piso, *Cibernetica fonaţiei în canto /The Cybernetics of Phonation in Canto/*, Bucharest: Editura Muzicala, 2000, p.83



The band of selectivity of the ear of a French type, which is limited between 1000 and 2000 Hz
(According to A. Tomatis)

As for the consonant ‘r’, in the middle of the word *information*, in the English orthoepy [ɪn.fəˈmeɪ.ʃən] it completely disappears in the way people in some regions pronounce the word, due to its harshness, and the fact that it lacks in penetration (instead of being rolled from the tip of the tongue just behind the incisors); it is not easily distinguished and if too much accentuated, has an unpleasant repercussion on the vocal chords. In the French orthoepy, the solution is to shift its place from the alveolar region (from the teeth) in the velar region, where [r] is produced due to the exhaled air, which meets part of the palatal veil (the uvula) that touches the back part of the tongue, thus producing this specific sound, This does not bring an increase in either the sonority, or penetration of the sound. Consequently, as [r] is *inefficient* in terms of its penetration and not easy to articulate, English orthoepy gave it up, altogether. In fact, many consider the consonant [r] a relic of primitivism, even a sign of vulgarity or of lack of refinement[8].

In conclusion, and as an explanation for the fact that the orthoepy of English is so far from its spelling (in which the etymology of many words is found), the pragmatism of English orthoepy should be emphasized, that

[8] See, I. Piso, *Cibernetica fonaţiei în canto /The Cybernetics of Phonation in Canto/*, p.71

is, the use of any means that make communication as clear and efficient as possible, without involving too many moves of the inferior mandible, and of the oral muscles (the orbicular or the oral sphincter, etc.)

Also, for practical reasons, to preserve energy, the article - which in German (related to English) is rather difficult, from a grammatical point of view, but also because of the guttural, phonic consonant [d] (*der, die, das* and all their complicated forms in declination), employs the larynx, maximizing the effective use of its articulation mechanism; thus, the [ðə] – a sibilant followed by [ə] – comes as a more general form, exactly because it does not necessitate an accentuation in articulation or sonority, which is seen as superfluous.

*

This modality of phonetic analysis, of English orthoepy, is based on the study of the oscillograms of phonemes, vowels and consonants (see charts 1-16 at the end of the volume) used in the words of this language and could become a practical guide for foreigners who wish to learn English; thus, clarifying the difference between its spelling and pronunciation, giving additional insights into the language's etymology. Spoken English, as compared to written English, will appear much more transparent, to the learner.

I have also developed my ideas by referring to the words of the world-famous musician, Sir Yehudi Menuhin, who was also a brilliant teacher, and who possessed not only a vast, noteworthy and profound knowledge, but also a very comprehensive one: "Ganz bewusstes verstehen ist jedem eher instinctiven Begriffen gewiss überlegen/ A complete, conscious comprehension is certainly superior to an instinctive understanding, which is not explained analytically/." I therefore hope to add clarity to the student's 'instinctive understanding' and prior learning, of the subject, in this study.

Although this is a comparatively short study, it is a first step on a road full of promise for those who are attracted to the study of phonetics, as well as the myriad aspects of speech, but especially of singing, from the point of view of the cybernetic phenomena, that are active while we sing.

N.B. In order to highlight, in a visual-graphic manner, the real, scientific objectivity of the results and conclusions I have reached, I considered it necessary to make use of a series of computer analyses of the specters of sonorous (acoustic) vibrations, obtained in the Phoniatric Laboratory of the University Clinic of the Medical School in Cluj-Napoca, in collaboration with Dr. med. Rodica Mureşanu, a phoniatric specialist; I want to mention here how grateful I am for this collaboration, and to congratulate her passion and commitment to this study.

BASIC ELEMENTS

Propaedeutics

MOTTO: EINE NEUE THEORIE SIEGT NICHT DADURCH, DASS IHRE GEGNER
SICH ÜBERZEUGEN LASSEN, SONDERN, DASS SIE AUSSTERBEN. /
A NEW THEORY CONQUERS NOT BECAUSE THE ADVERSARIES
LET THEMSELVES BE CONVINCED, BUT BECAUSE THEY DIE/
(MAX PLANCK, NOBLE PRIZE WINNER)

WESTERN THOUGHT CONSIDERS, AS FOREIGN FROM THE WORLD OF VALUES,
ANYTHING THAT CAN BE EXPLAINED BY THE LAWS OF NATURE.
TO BE INTELLIGIBLE FROM A SCIENTIFIC POINT OF VIEW
EQUALS A KIND OF DE-VALORIZATION.
(KONRAD LORENZ, NOBLE PRIZE WINNER)

A. Speech transmits messages through acoustic signals formed by consonants and vowels, which can be received by the auditory apparatus. In order that reception is optimum, messages have to fulfill certain conditions imposed by the laws of acoustic physics, as well as those of the physiology of the phonating and the auditory apparatuses, that is, the human ear. Between the two apparatuses a cybernetic link is established, of *reversed connection*, or *feedback*; consequently such a link has permanent circularity.

I will analyse vowels and consonants one by one, with regard to the frequencies that characterise these acoustic signals, with the intention of discovering if there is any difference between speaking and singing in the way they are used in both; moreover, I will try to show the practical implications that are subsequent to such analysis.

1) In the case of normal conversation, when the intensity of sonorous messages does not exceed 20-30 dB, acoustic signals (sounds), whose frequency is between 1000 Hz (Si4) and 2000 Hz (Si 5) – see the ivories of the piano at the end of the study, and the frequency of the acoustic signals from 32.69 Hz (Do0) to 4184.32 Hz (Do7) – are perceived by the

ear in an optimum manner, as compared to those that have a higher frequency (higher sounds), or a lower one (lower sounds).

2) When acoustic signals contain harmonics between 2500 Hz and 3500 Hz, they can be perceived clearer and are more distinctive, at longer distances, due to the fact that they have a smaller opening and are more penetrating.

For us to understand the difference between speaking and singing from the perspective of the frequency of the sonorous signals, it is necessary that we clarify from the very beginning that they are composed of the frequency of the base sound, over which harmonics are superposed. The harmonics of the human voice are of two types: *vocalic and extra-vocalic*. In contrast with vocalic harmonics, which are characteristic for each vowel (see Hermann von Helmholtz's classification below), the extra-vocalic harmonics are those that constitute the specific richness of the timbre or specific sound of a voice, irrespective of the vowel that is articulated). The more numerous and intense extra-vocalic harmonics are, the more professional the technique of that specific singer becomes. A deficient vocal technique is characterized by a reduced number of extra-vocalic harmonics. Thus, the closer to simple speech it becomes; such deficiency will be felt, then, in the quality of the voice – its richness, brightness and sonority –ultimately, it will affect the beauty of its sound.

Vowels have the following frequencies (according to Helmholtz, Koenig, Rocard[1]):

/an/, /on/, /in/ – nasals = at la1 (110 Hz)
/a/ = la2 (220 Hz), the most open vowel
/o/ = la3 (440 Hz)
/e/ = la 4 (880 Hz)
/i/ = re6 (2347 Hz), the most closed vowel

If we take these frequencies into consideration, and the fact that the speaking voice has a base frequency of approximately 300-500 Hz for women, and 130-280 Hz for men, we will understand why in some idioms, spoken by populations that have been used to urban life for many centuries, the speakers tend to *sharpen* certain words, preferring to use vowels that are more penetrating; thus, speakers are *seasoning* many words with those sonorous signals that have a higher frequency, especially /e/ and /i/, in the case of English. Why populations in urban areas? Because, contrary to people in rural areas, they did not have to speak so

[1] See Y. Rocard, *Dynamique generale des vibrations*, Paris, 1943

loudly, to emphasize open vowels like /a/ and /o/ frequently or excessively, so they could hear each other with clarity; nor did they have to use extremely powerful acoustic signals. In this way, urban populations, *adjusted to* the low, *extra-vocalic* frequencies of the speaking voice, and the *vocalic ones* of /a/ and /o/.

Consequently, in orthoepy, /e/ and /i/ will often *substitute* the vowels 'a' and 'o' that appear in the spelling of some words, for etymological reasons, as they are vowels whose power of penetration is limited. The following examples will clarify this process. When such a substitution happens, the sonorous effect will have a rather metallic tone, which is sometimes even strident, due to the predominating high frequencies. Due to the fact that both the base frequency of the human voice, and the frequency of certain phonemes – vowels /a/ and /o/, as well as that of certain consonants – does not exceed 500-600 Hz, which makes the reception of such signals less than optimal, even deficient, at times, some languages have resorted to a replacement of these phonemes with others that are more penetrating, characterized by a higher frequency; they have done so in order to eliminate any inconvenience in communication, and make sonorous messages heard with greater ease and clarity.

B. The procedure is straightforward in, for instance, English, which is one of the European idioms that people have *worked on* the most, from the point of view of speech penetration. A few examples can illustrate the phenomenon. The word 'association' is pronounced /əˌsəʊʃiˈeɪʃn/, in which we can notice the replacement of 'a' (220 Hz¬) and 'on' (110 Hz) from French, with much closer vowels: /e/ (880 Hz) and /i/ (2347 Hz); the word 'pigeon' is pronounced /ˈpɪdʒɪn/ - the same procedure appears when more *guttural* (phonic) consonants are replaced with front ones (a-phonic), or when dental fricatives are replaced with more sonorous (labial) fricatives - e.g. the fricative consonant /s/ with the sonorous fricative /ʃ/ (see also figures 8, 10 and 16, in which we analyse consonants /s/, /ʃ/, /tʃ/ and /z/). We will come back along the way and explain their function. The word 'gingival' is pronounced /ˌdʒɪndʒɪˈvəl/, 'tea' is pronounced /ti:/; 'cream' is pronounced /cri:m/ (both with a long /i:/), while 'infection' is pronounced /ɪnˈfekʃn/, as the final part of the word, due to the nasal /n/ (110 Hz) which has a much reduced penetration, similar to the words 'association' or 'pigeon'. The group /ʃn/ raises the frequency towards 2400 Hz, so the phonetic penetration is much increased (see chart 9). Before we go any further, here are the correspondents between a-phonic and phonic consonants (in Romanian, but not only).

Strong a-phonic consonants	Phonic consonants	
P	B	*Plosives, labial occlusive*
T	D	*Plosives, dental occlusive*
C (palatal)	G (velar)	*Plosives, occlusive*
F	V	*Oral fricatives*
S	Z	*Dental fricatives*
Ş /ʃ/	J	*Sibilant fricatives*

The examples above prove the existence of a mechanism, which is conducted and coordinated by the principle of circularity (a cybernetic principle) and which has become installed in speech under the control of the auditory apparatus; its purpose is to enhance the penetration of phonation in English, without the speaker having to speak more loudly or without an increased level of Decibels (dB) of these acoustic signals, which would come as a consequence of an increased effort, both at the level of the larynx and that of the lungs (the exhaling jet while speaking). Thus, with only moderate effort, the speaker can obtain a more efficient result, making his/her speech clearer and more penetrating[2].

The frequent presence in speech, of these penetrating vowels, is part of a reflex mechanism, that is, part of the circularity, or the cybernetics; it is generalised, its purpose being to bring the sonority to the front part of the oral cavity (a cavity which has a reduced space[3]). This space also has the advantage of being very easily *modeled* by the tip of the tongue (this

---

[2] A similar tendency regarding consonants that are formed in the middle of the palate, that is, to replace some of the plosives, can be noticed in some regions in Romania. Around Braşov (the Bârsa County) we can often hear /ghine/ instead of /bine/, /ghie/ instead of /vie/, /ghiţăl/ instead of /viţel/, etc. The consonant /v/, even though a phonic one, being a plosive consumes so much air; /b/, solicits the larynx much more than the group *ghi*, which is easy to articulate, and, with the same effort of the phonating apparatus is emitted as a more sonorous sound. The larynx, in the case of the plosives, almost painfully feels the expulsion of a greater quantity of air, after stopping it, when we intend to speak louder, in comparison, the group *ghi* takes away this pressure. Making the articulation easier, with a greater sonorous result, can also be accomplished by replacing /p/ a bi-labial plosive in the word /piept/, with /t/, that is /tiept/ (as Hasdeu remarks in his work *Cuvente den bătrâni / Words of yore*, reproducing from Anton Pann a dialogue that refers to the Braşov *dialect* in an ironic manner); they also pronounce it /chiept/, as we can hear it in the Bran or Zărneşti region.

[3] This usually occurs when the exhaled air makes the sound slide too much to the front, provoking a turbulence in the oral cavity, that is, when they speak or sing using a lot of air.

organ's most mobile part), thus forming in this part of the oral cavity an adequate space, that is, a much-reduced one. Consequently, this smaller space is convenient for the formation of acoustic signals, with an increased frequency, which are more penetrating; this is the case in vowel /e/ (880 Hz), but especially the case in vowel /i/ (2347 Hz).

C) Nevertheless, a shift of the sonority from the *back* to the *front* of the cavity, creates impairment to the richness of the sound (not to be mistaken for power), which will be felt in the singing voice. The articulation of these closed vowels, /i/ and /e/ takes place in the front part of the oral cavity, so the opportunity to be loaded with a rich cascade of extra vowel harmonics – so that the vowels may thus acquire a rich sonorous body – is limited. These sounds will be penetrating, but not only poorer in their harmonics, but also with blank spaces, that is, with portions in which there are no harmonics over the intensity of 90 dB (also see the charts that analyse the spoken vowels /i/ and /e/).

In the case of vowel /i/ in speech (chart 1), a vowel that has an intensity of 58.20 dB, the harmonics between 800 Hz and 2200 Hz – which should form the sonorous body of this vowel – are practically nonexistent, as we can see in the diagram in section C (which shows the distribution of the harmonics of this sound according to frequency – in Hz – on the ordinate axis, and the intensity – in dB – on the abscissa). They start to appear only between 2200 Hz and 4000 Hz, that is, beyond the spectrum of normal audibility, presenting intensities less than 90 dB.

The vowel /e/ in speech (chart 2), with an intensity of 66.02 dB, proves to have a slightly better situation between 1000 Hz and 2000 Hz (as we may see in diagram in the section C); nevertheless, above 2000 Hz, the intensity of the frequency of the harmonics is inferior as compared to vowel /i/. This explains why, when used in singing, the vowel /i/, as it is formed in speech, has a sonorous capacity that has a rather strident effect, which is, consequently, unsatisfying from the point of view of the richness and the roundness of the sound.

When a male voice pronounces the word *tea* /ti:/, the spectrum of sonorous frequencies contains – besides the fundamental sound between 130 Hz and 280 Hz, that is the height that the word was uttered, and the specific height of the vowel /i/ (2347 Hz) – too few intermediate harmonics; these represent the extra-vocalic sound, which gives body to the sound that is emitted with the utterance of this particular word, /ti:/.

D) As a result of this, when a singer uses the above-mentioned articulation (the cybernetic circuit functions according to normal usage), in a reflexive manner, and sings a text in which this *sharp* word appears, in such a way that the closed vowel /I/ will sound penetrating, it will lack richness in harmonics, being a meagre sonorous substance, lacking consistency and will be perceived, by the trained musical ear, as aggressive and strident.

Moreover, in singing, when the vowel /a/ follows after the vowel /i/ in a word (for instance in the word *viaţă/ life in Romanian/*), or the other way round as in the word *aicea/ here* in Romanian, it will be very difficult, if not impossible, for the singer to obtain a vocal line made up of homogeneous sounds, that are close to each other in that they have resembling harmonics, because the sonorous spectrum built from the vocalic harmonics of these two vowels is very different: for /i/ it is over 2300 Hz, and for /a/ only 220 Hz[4].

E) In the first stages of studying professional singing, when he/she has not yet developed the vocal technique necessary for accomplished performance, the singer will sing as he/she speaks, as there is a natural tendency to hold to the known phonetic patterns of our mother tongue; thus, when a singer who is used to the sonority of English meets a vowel configuration that resembles the one he/she recognises from his/her mother tongue - even when singing in a foreign language – he/she will produce penetrating sounds by virtue of habit, but which lack sonorous richness. In the case of the word *tea* /ti:/ there will be greater anteriority, which will also cause a projection of the sound forward; the sound /t/ therefore, being front and plosive, will have a reduced chance of acquiring extra-vocalic harmonics, when sung, because of the sound /i/. The result is that the sonorous body is even poorer. Comparing the vocal performance of this singer with that of a professional singer, let's say, a Romanian or an Italian one, the vocal efficiency of the first will appear much different, and not to the better, unfortunately.

F) If we leave the Albion and go southwards to the south of the Italian Peninsula, in Mezzogiorno, we meet with a contrary tendency. Here, the articulation of vowels, are produced closer to the centre of the mouth, in the area of hard palate. The singer in Naples or Calabria prevents the vowels from slipping forwards – especially /e/ and /i/ - by replacing the aphonic consonants (see the table of consonants above) /s/, /p/, /t/, /c/, /f/, /ʃ/, which accompany those vowels in certain words, with their phonic, or throaty[5] counterparts; these - /b/, /d/, /g/, /v/, /z/, /j/ - bring sonority to the voice because they narrow (and reduce) the turbulence of air. In these

---

[4] In English, those who speak it have an ear that is used to such sonority, and so everything *seems* to be all right for them

[5] See *Cibernetica fonaţiei în canto /The Cybernetics of Phonation in Canto/,* Bucharest: Editura Muzicala, 200, pp. 65-73

consonant substitutions[6], they also add the substitution of the vowels /o/, /a/, and /e/ with the vowel /ə/; for instance, instead of /santo/ we will hear in the Naples dialect /zandə/; instead of /spinto/, /zbində/; instead of /sempre/, /zembrə/; instead of /turbamento/, /durbəmendə/; instead of /fronte/, /vrondə/; instead of /particolarmente/, /bardigolarmemdə/; instead of /fratello/, /vradelə/, etc. etc.

G) I demonstrated in another study[7] how the illustrious Neapolitan, Caruso - who benefited instinctively from a rich sonority in speech, which he also redounded upon his singing in a reflexive and delightful manner – modified consonants and vowels in a word in order to obtain a greater sonority; this is the case in the song *La Danza*[8] (a Naples tarantella), by Rossini. In the refrain of this canzone, there appears the pronounced word *frinche* /frinke/, which projects the voice to the front, at the teeth, due to

---

[6] The throaty consonants influence phonemes in their vicinity, even when least expected, as proven by the following example: the DEX [the Romanian Explanatory Dictionary (translator's note)} on page 529, explains the formation of the word *înfiera / burning brand/* from the particle *în / in/* + the word *fier / iron/*, while for the word *îmbolnăvi / to get sick/* they show that it comes from the particle *în /in/* + the word *bolnav / sick* (p. 514) In the constitution of both words the same particle *în* participates, but, while in the word *înfiera* this particle remains unchanged, in the word *îmbolnăvi,* /n/ is transformed into /m/, under the influence of the glottal /b/, which is a bi-labial occlusive plosive. When we split this word into syllables, though, that is, we separate /n/ from /b/, /n/ thus getting free from the influence of /b/, it is no longer transformed into /m/; therefore, we will split the word into the syllables *în-bol-nă-vi,* and never *îm-bol-nă-vi*. As the consonant /f/, which is a sibilant, does not have the same effect on /n/, the latter remains unchanged in the word *înfiera.* This is because in this mechanism: /m/ is a consonant, whose obstruction by the closing of the oral orifice is more forceful, and takes place more to the back of the mouth as compared to /n/; that is, /m/ is closer to where /b/ is formed, and therefore /b/ takes /n/ slightly backwards and transforms it into /m/. In the word *înfiera,* /n/ will remain unchanged, as it is articulated in the vicinity of /f/, which is a front sibilant par excellence. In fact, /b/ and /m/ are pronounced with the mouth initially open, being plosives, while /n/ and /f/ require the lips to be parted.

[7] See the booklet and the CD annexed to the *Cibernetica fonaţiei în canto / The Cybernetics of Phonation in Canto/*, as well as pages 69 and 144 in the same study, in which I remark how dialects in the south of Italy, especially the ones in Naples, manifest a marked preference for the transformation of a-phonic consonants (/s/, /p/, /c/, /t/, /f/, /s/) into phonic ones (/b/, /g/, /d/, /v/, /z/, /j/), thus forgetting about classical orthoepy of Italian in a spontaneous manner, that is instinctively, in order to give the voice additional sonority.

[8] See his recording in February 13, 1912 (the collection that contains the integral of Caruso recordings, *Bayer Dacapo*, vol. 8, nr. 10)

the swishing sound /f/, an a-phonic fricative (see chart 16), and to the vowel /i/, followed by the plosive, occlusive /k/, which, in turn, is followed by another vowel /e/. All of these have the tendency to be articulated in the front of the mouth. Caruso, however, pronounced the word /frinke/ as /vrengə/. Why? He did so due to the phenomenon of cybernetic circularity by feedback; Caruso's ear asked for the replacement of the fricative /f/ with its phonic counterpart /v/[9], which is less anterior,

---

[9] A similar tendency can be met with in old Romanian. The phenomenon is highlighted by B. P. Hasdeu [one of Romania's first linguists (translator's note)], in *Cuvente den bătrâni / Words of yore/*. Studying the most important collection of old Romanian texts, of the 16th century (*Textele Măhăcene / The MăhăceneTexts)*, that is, the beginning of the Apocalypse of "Saint Paul the Apostle", pp. 167-171), Hasdeu remarks that in several different places, instead of *sfântu* /sfɨntu/, with an /f/, there appears /svəntu/ with a /v/; and instead of *te sfieşti* /sfieʃti/, there appears the form *te svieşti* /svieʃti/. The authors of these texts were writing according to "local and momentum needs", as Hasdeu states – that is, as they spoke; thus, the author reaches the conclusion that those particular texts represent the language Romanians spoke between 1550 and 1600. The author, who in some cases was the translator, of these *Măhăcene* texts, was the priest Grigore, from the village of Măhaciu. During the sung liturgy, the word *sfânt* /sfɨnt/ appeared frequently, which may explain why, in order to produce it more easily and more sonorously, the priest turned to the procedure used by Caruso; thus, he replaced the a-phonic sibilant /f/ with the phonic oral fricative /v/, which is pronounced with closed vocal chords; a singing position. Because of this, the orthoepic form made its way in the spelling one. My explanation strengthens Hasdeu's explanation; those texts are the mirror of the language spoken in those times, and also confirms my other theory, according to which the tendency in singing is to replace an a-phonic consonant - in this case the a-phonic fricative /f/ - with the phonic fricative /v/, which is much more sonorous, as it employs the larynx; /v/ is articulated with closed chords, in a phonating position (see chart 4, according to Cornut).
At a distance of more than 1000 km, and almost 500 years, the exigency of the musical ear of two *singers* – of the priest Grigore and of the famous singer Caruso – found, independently of each other, *the same* phonetic solution. Both idioms, Romanian and Italian, have made their sonorous bread from the same dough, but, due to geo-political reasons, they *have risen* differently. We can also find, in the centuries that followed, examples that show a special preference for the replacement of the a-phonics with phonics; *throaty* ones, for instance: the inscription Brâncoveanu [a Romanian Voivode] placed on the belvedere of the Mogoşoaia palace says: "This palace was built and adorned by the Luminous Highness Voivode, I, *Costandin* Basarab" [in the original Romanian, where the name Constantine is spelled normally Constantin, not Costandin, as it appears in this inscription (translator's note)]. This is also true of the founder of the tower in Broșteni, who appears in the murals of the church in that town, and who is called

and consumes less air. Instead of /i/ (2347 Hz), which is the most anterior vowel, whose stridency is accentuated by the fricative /f/, we can hear something more like its neighbour, /e/ (880 Hz), which is less anterior. The palatal occlusive /c/ is also replaced by the velar occlusive /g/, which is typically *throaty*, and which is the most sonorous consonant in singing, the final /e/ becoming /ə/. We will concentrate largely on the particular virtues of this vowel in this study. From the *original* word, only /r/ and /n/ are preserved. All these modifications were done only in view of making the word more sonorous, giving it body; making it *vibrant*.

I will describe the inner mechanism that explains the Caruso phenomenon: the vocal chords behave differently in the case of an a-phonic consonant, in comparison to the phonic ones. The production of a consonant in the group of the a-phonic /p/, /c/, /t/, /f/, /s/, /ʃ/ comes about through the exhalation of a great quantity of air, which passes through the vocal chords, which are *apart* from each other (see Cornut's chart); therefore, they will not vibrate, while other consonants, from the group of the phonic ones /b/, /d/, /g/, /v/, /z/, /j/, produce a coming together of the vocal chords, so their vibration starts right after the air is accumulated in the back of the occlusion (with the occlusive ones), or after the moment of tension, in the case of fricatives. By the use of this technique, the genius singer[10] proves, once again, his infallible vocal instinct. He can coordinate and order this manner of singing, which uses the cybernetic circularity between the larynx and the ear; thus, he creates the conditions that allow him to sing with a vibrant, sonorous, round voice, which lacks any kind of

---

"Master *Constandin* Cuţui". In both cases, Constantin was replaced with *Costandin*.

[10] Caruso applies, consistently, the technique of *voice inhaling*, of which he is conscious. This is born out in the words he frequently used: "Don't forget to sing within yourself! („Auch denk' daran in Dir selbst zu singen!" – „Pense aussi à chanter en toi!" in *How to Sing,* by Enrico Caruso, London, Cincinati, New York: The John Church Company, no year), which means that one should not push the voice outside of oneself, before it is well loaded with harmonics, that is, before it acquires a rich sonorous *body*.

In fact, even in the 19th century, the Escudier brothers, musical critics, were talking about the shows of the Italian Opera in Paris, noting: "Si l'on veut savoir pourquoi ces artistes se font entendre dans les grands salles, sans effort, c'est qu'ils ont la voix **dedans**, et que c'est a force de l'avoir en dedans, que les sons finissent par sortir clairs, ronds, vibrants " / If we want to know, to understand why the voices of these artists can be heard so well in big halls, without effort on their part, the secret lies in the fact that they do not project the voice out on purpose, but they sing it *within themselves*; therefore, the voice comes out clear, round, and vibrant/ (see Jean Laurens, *Problèmes du chant français*, Paris: Subervie-Rodez, 1977, p.20

stridency or asperity; the chords will only vibrate at the vowel /e/ in the word /vrengə/. In this manner, he modified the sonority of this word substantially. The alert tempo does not provide the singer with a means of avoiding stridencies in the singing of the word, this being the only way of maintaining control over the sonority of the voice. Its sound, which is under the threat of being modified (even of a total change in sound) as a consequence of the imposed anterior articulation – especially by the vowel /i/ and the fricative consonant /f/, does not allow a reduction of exhaled air during the phonation; on the contrary, it would mean a loss of non-sonorous air, which is lacking in vibration.

The word /frinke/ is repeated, in the way described, above, 24 times during this song. The fact that Caruso replaces it 24 times with his own version cannot be considered accidental. However, his modification does not alter the meaning of the text; rather it acts as an exclamation. In this way, Caruso avoided all obstacles that prevent the round, ample sonority of the voice, which was so important to him. The collaboration between the larynx and the ear is obvious; the producer of the sound (the larynx) and the control of the sound (the ear) collaborate due to the phenomenon of circularity (according to Odobleja), by reversed connection or feedback, called Cybernetics (according to Wiener), to accomplish an optimum result in singing.

*

This helps us to understand how modifications can be made to the way words are spoken, to produce a greater quality of sound, when the word is sung. This said, there are many mastersingers, of worldwide fame, who expound the theory that the singer should be influenced by the spoken word, when singing. Such masters base their teaching of singing technique on the premise, ' sing as you speak' and this advice is taken in good faith and practised mot-à-mot, by many students. Consequently, in my view, their voices fail to reach their true potential of sonority and richness.

H) Experimental studies in the domain of phonetics, have brought us to the following conclusions: the resonance cavities that favour the forming of vowels, are regions in the oral and oral-pharyngeal spaces[11], which have various volumes, specific to each vowel (the volume being reduced to the minimum in the case of the vowel /i/); their form is of no consequence - see the studies belonging to Hermann von Helmholtz, with the apparatus he invented, which is called the Helmholtz Resonator – there are many

[11] For details about the resonators, see *Cibernetica fonaţiei în canto / The Cybernetics of Phonation in Canto/*, pp. 58-62, and 114-115

articles about this invention, together with the remarks made by Doctor Alfred Tomatis[12], to be found on Google).

This is one element of Helmholtz Resonator (found on Google)

Logically, if there existed any possibility for the formation of a space identical in volume (or at least as close as possible) for the pronunciation of the vowel /i/ in the middle of the oral cavity, this vowel would no longer be the *natural* vowel – the front vowel we use in speech. This vowel could be formed in that *new* space. Thus, there would exist the possibility that the spoken pronunciation of the vowels would not be the

[12] *L'oreille et la vie*, Paris: Laffont, 1977; *L'oreille et la voix*, Paris: Lafont, 1987

only way of articulating them; that is, the place where vowels are formed in speech could be shifted when singing, by using a new type of articulation. This is especially a desideratum, as the singer becomes aware, that the way sounds are pronounced in speaking rarely favour an optimum sonority in singing, especially in the upper part of the range, with acute sounds. But, before we see if there exists an effective and practical way of accomplishing a better sonority in singing by *shifting* the place where vowels are articulated, let us see what is the difference between the two manners of articulation.

I) In order to understand the difference between the mechanism of articulation in speech and in professional performing singing, by the technique of *inhaling the voice* (see the details and the explanation below), that is, of the different manners in which vowels are formed in the two cases, we have to corroborate Hermann von Helmholtz's theory of classification of vowels according to their specific frequency (above) with the results of Gérard Guth's[13] investigations. This researcher, from the Institute of Phonetics in Strasbourg, noticed how sounds are spaced out, according to frequency, along the main resonators; that is, the two cavities, the oral and the oral-laryngeal. Regarded through the viewpoint of these investigations, the theories of Helmholtz, Koenig, Rocard et al, can reveal the process of the formation of vowels, both in speech and in professional singing as two rather different mechanisms.

By placing a tiny microphone, on a micro-probe, G. Guth analysed the trajectory of sounds of various frequencies along the oral-pharyngeal and oral resonators. Here are his results in summary:

Each portion of this trajectory favours a certain spectra of sonorous waves. If at the level of the vocal chords the sound is a-vocalic[14] - which means it will have the same sound for all vowels – only representing the fundamental frequency of that sound, the sonorous vibrations (the sounds) that are developed in the pharynx are between 250 Hz and 700 Hz, which constitutes the first forming stage. The oral cavity, per se, favours the production of *harmonics with frequencies between 700 Hz and 2000 Hz,*

---

[13] See *Cibernetica fonaţiei în canto / The Cybernetics of Phonation in Canto/*, p. 116

[14] According to R. Husson (1962) and G. Cornut (1998), the fundamental frequency of the laryngeal sound has the following characteristics: a determined height, which is the result of the *frequency* of the sonorous vibrations of the vocal chords; a weak sonorous intensity, as a result of the *amplitude of these* sonorous vibrations; and a timbre, or sound, conditioned by the form of the sonorous impulse (see *Cibernetica fonaţiei în canto / The Cybernetics of Phonation in Canto/*, p.111).

even higher than 2000 Hz, which correspond to the second forming stage. Consequently, starting with the upper part of the pharynx, approximately at the level of the palatal veil, the lower frequencies lose amplitude and start favouring the higher ones, little by little, as we get closer to the alveolar region (the teeth), where the sonorous spectrum contains high frequencies. (Some sounds[15] also appear here). Obviously, the forming of vowels in speech is due only to mobile resonators, of which the oral cavity, the tongue and the oral orifice[16] play the most important part in the modelling of this resonating space.

In *speech*, vowels /o/ (440 Hz) and /a/ (220 Hz) are formed at the level of the first forming stage (250 Hz – 700 Hz), in the back of the oral cavity (the oral-pharynx), which is the resonating cavity that has the biggest volume. (The sinusoid of vowel /a/ as shown in section C – which contains the analysis of harmonics – looks like the reversed same section C of the consonant /s/, see charts 5 and 8 comparatively. Thereabouts start the harmonics in the case of /s/ and end those that vowel /a/ contains). In speech, there is also a reversed relation between the harmonics of vowel /a/ and vowel /i/. In the part where the harmonics of vowel /i/ have less intensity, vowel /a/, on the contrary, has harmonics of high intensity (see the respective charts).

The headquarters of the forming stage 2 (700 Hz – over 2000 Hz) starts where the forming stage 1 ends. Vowel /e/ (880 Hz) is articulated in that place. In the front part of the forming stage 1, in the alveolar region (the teeth), where the volume of the oral cavity can be thoroughly reduced, vowel /i/ (2347 Hz) is produced, as it is the closest vowel.

Consequently, vowels in speech are arranged in a fan that starts in the back and ends in front, with the closest one, where noises are also

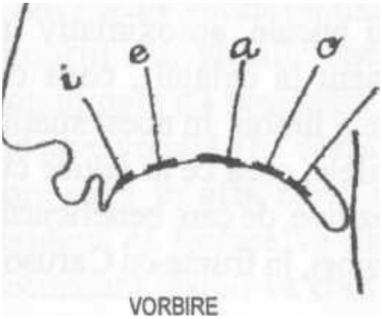
---

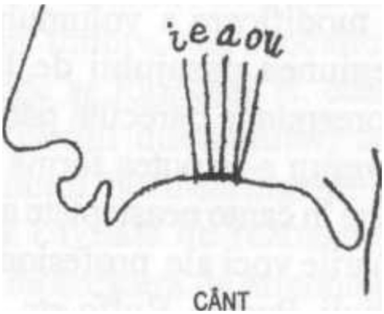
[15] See the charts, at the end of this chapter, illustrating the oscillograms of some consonants, which, in fact are noises. If we compare the register of their harmonics, section C (the low register) of the vowel /i/ (chart 1), with that of the consonant /ʃ/, section C (the same register, chart 9), we will notice a certain surprising resemblance between 1000 Hz and 3000 Hz, where the harmonics that are best perceived by the human ear appear, as they are more penetrating (see above).

[16] Brunings hears the laryngeal sound in 1938 for the first time, by studying the phonation of subjects with tracheotomies. In 1954, Portmann and Beckmann analyse sound at the level of the larynx, by means of a spectrometer, thus noticing that the laryngeal sound has an a-vocalic character. The timbre of this sound remains unchanged, even if the subject intends to pronounce different vowels. The mechanism of producing the fundamental (laryngeal) sound, that is of the way in which its height is determined, is analysed by Ion Piso in *Cibernetica fonaţiei în canto / The Cybernetics of Phonation in Canto/*, p. 111-112

formed[17], which is not of little importance for vocal professionals, who want to sing with a voice that is round, bright and flexible.

In short, we may say that the volume of each resonating cavity generates the vowel that corresponds to it, that is, the forming of vowels depends on the dimension of the resonating cavity, which takes over and amplifies the fundamental sound, which is produced by the vocal chords. The two drawings below suggest the area in which vowels are formed, on the palate in speech, and as compared to their formation in professional, performance singing (the *inhaling* of the voice).





Speaking Singing

© Piso

J) As the volume of these resonating cavities can be modified, particularly by the mobility of the tongue – at the command of certain cybernetic circuits[18] - the forming of each vowel is not necessarily linked to its *natural* place (the one that it corresponds to in the articulation of spoken sounds), as it solely depends on the volume of the resonating cavity, which can be shaped, or modelled.

This means that back vowels - /a/[19] and /o/ - can also be articulated more to the front, and the front ones more to the back. Thus, we can obtain a brighter quality of the vowels /a/ and /o/ - due to the use of secondary

[17] In many male and female singers, especially in tenors, but in sopranos, too, this is the reason why the vowel /i/ emerges with a stridency of sound

[18] See *Cibernetica fonaţiei în canto / The Cybernetics of Phonation in Canto/*, especially Chapter III, pp. 118 -154

[19] The tenor Alfredo Kraus said about vowel /a/ that it is a muckworm in terms of harmonious frequencies (see Ion Piso *Cibernetica fonaţiei în canto / The Cybernetics of Phonation in Canto/*, p. 79)