FROM SPEAKING MACHINES TO SPEECH SYNTHESIS

Jean-Sylvain Liénard

Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur

LIMSI - CNRS
BP 133 - 91403 ORSAY Cedex

ABSTRACT

Speech synthesis has a long history. The first scientific attempts date back to the latter part of the eighteenth century, with Kratzenstein's resonators, Mical's talking heads, and Kempelen's speaking machine. The functioning and results of these devices are presented, as well as those of two other famous machines, Faber's Euphonia and, more recently, Dudley's Voder. The second part of the paper introduces some elements of discussion concerning the magical aspect of speech, the modelling of the human speech production system, and the very concept of a speaking machine.

1 - INTRODUCTION

Kempelen was the eighteenth century inventor of the first speaking machine, the good functioning of which has been testified by numerous contemporaries. This achievement was not due to chance; it was part of the framework of a movement of interest in the topic of speech production, which itself was related to the wide curiosity of this time for the understanding and imitation of natural phenomena. Among the other attempts which appeared afterwards, the most accomplished was undoubtably Faber's machine, unfortunately ignored, probably because its author, unlike Kempelen, had not published a book describing it in detail. The last speaking machine which really deserved this name was that of Dudley, the Voder, which was also famous in its time, but which carried in itself the end of an era and the beginning of a new one. Today the term "speaking machine" and the dreams it denotes are abandoned to the benefit of that, seemingly more technical or rigourous, of "speech synthesis".

We shall try, in the following paper, to draw the main lines of this evolution, to show its ruptures and survivals and, perhaps, to extract from it some lessons for the future.

2 - HISTORICAL SURVEY

Kempelen's idea was not new. It corresponded to an old dream of humanity, which had been formulated as early as antiquity and Middle Ages. It was felt then that to give speech to inanimate objects was to give thought, a divine gift, specific to the human species. Making objects speak had a magical, if not sacrilegious, aspect. Even today, everything related to the mechanisation of speech retains some kind of mystery; the domain is intrinsically spectacular, and the researcher often has to restrain the untimely enthusiasm of his contemporaries.

2.1 - Myths and mystifications [1, 11, 14]

Several antique authors describe the talking statue of Memnon, the son of dawn in the mythology of ancient Egypt. This statue was said to emit intelligible speeches, and even seven-verse long oracles ! This extraordinary gift was due to the intervention of the egyptian priests; the statue emitted only one sound, evoking the breaking of a string. The phenomenon happened at sunrise; it may have been caused by the dilatation of some parts of the monument because, in the country in question, the temperature deviations between night and day may be very great.

Throughout history, there are many testimonies related to talking heads. At Lesbos a speaking head attributed to Orpheus was famous for having predicted the violent death of Cyrus the Great, during his expedition against the Scythians. Odin, the nordic magician, had a talking head which passed as wise Minos' head, and which gave divine answers...The mechanician Gerbert of Aurillac, who became pope at the turn of the first millenary under the name of Sylvester the Second, was supposed to have built a talking head of brass, which said the words "yes" and "no". Monk Albert, who became for posterity Albert the Great (thirteenth century), was reported to have build a head of baked clay, which spoke and moved. This masterpiece had such a sacrilegious character that Thomas Aquinas broke it into pieces...

It is useless to lengthen the list of examples. Most of those heads are a matter for legends. If however some of them seem to have existed, they probably worked with concealed pipes, or through the ventriloquist talents of their authors.

2.2 - The beginning of a scientific approach

Since the seventeenth century it has been possible to observe the maturation of ideas concerning the mechanism of speech production. The preoccupations then were of a philosophical and anatomical order. An alchemist, Van Helmont, imagined in 1668 a theory according to which the letter shapes of the Hebrew alphabet would represent the positions of the tongue in the mouth...The same year, a less fanciful study was performed by Wilkins, who defined for each speech sound a corresponding arrangement of the vocal organs. Debates on the nature of the voice producing organ arrived in the eighteenth century at the notion of vocal cords, due to Ferrein. Finally, in 1779, the Academy of Sciences of St Petersburg proposes as the topic of its annual contest the following questions:

a) What are the nature and character of the sounds of the vowels A E I O U (probably /a e i o u/), so different one from another ?

b) Can an instrument be constructed like the *vox humana* pipes of the organ, which shall accurately express the sounds of the vowels ?

At that time three researchers - Kratzenstein, Mical and Kempelen - had already obtained some results. But they did not know each other and, apparently, only Kratzenstein presented a realization at the Academy. He was the one who won the prize.

Christian Gottlieb Kratzenstein was a professor of physiology at Copenhagen. He presented five resonators, of which the internal volumes and openings corresponded to those of the mouth during the emission of the requested vowels. They were adjusted on a windchest and modified the timbre furnished by a free reed, except for the resonator producing the sound I, which received the airflow directly.

In 1778 Abbot Mical, a Frenchman who had a passion for mechanics, had built a talking head which could articulate a long sentence. Forced into the public eye by a friend's indiscretion, he destroyed his machine. But this inventor, modest as well as quick-tempered, started again and at the French Academy of Sciences in 1783, he presented a machine made of two talking heads, which pronounced the two following sentences in the form of a dialog:

- first head: "Le Roi a donné la Paix à l'Europe" *("The King has given Peace to Europe")*,

- second head: "La Paix couronne le Roi de Gloire") *("Peace crowns the King with Glory")*.

The committee put in charge of examining the machine was composed of distinguished scholars [3]. It worked out a long report in which, unfortunately, was no diagram. Here is a part of it (translated from [3]):

"... These two sentences are not clearly pronounced in all of their parts, especially the last one. This is mainly due to the fact that the basis of the voice produced by this machine greatly differs from a human voice; that, since some syllables result from the combination of several sounds, their joining does not occur with all the possible precision; and also that the pronunciation of several consonants needs to be perfected... One can consider it as made of two different parts:

a) A wind chamber, in which a bellows brings the air and from which this fluid escapes when different valves are raised. The air is then directed by ducts towards the cavities, where it is modified, and where it becomes sonorous.

b) A cylinder which moves levers, and which gives them the necessary impetus, either to raise the valves at the appropriate moment or to give the cavities where the sound modifies itself the shapes required by its diverse modifications."

From the description given in the report it follows that each of the resonators is fitted with its own reed, which probably explains some of the difficulties encountered in the stringing together of successive speech sounds.

Despite the flattering terms of the prerecorded sentences, the King did not express interest in Mical's machine. The inventor became impoverished and died in 1789.

2.3 - Kempelen's machine [5, 10, 11, 12, 13, 14]

Wolfgang von Kempelen, a nobleman living at the court of Austria-Hungary, was born in 1734 in Presburg (now Bratislava) and died in Vienna in 1804. He was an organiser and inventor of great talent. He was the designer of the fountains of Schoenbrunn palace, as well as of the plans of the royal castle of Buda. He was the organiser of a wool factory in south Hungary. But above all he was the author of two renowned machines: the chessplaying automaton, which was immortalized by Edgar Allen Poe in a famous novel, and the Speaking Machine. Kempelen never claimed the chessplayer to be a real automaton, but the trickery was so perfect that nobody was ever able to fault it. As for the speaking machine, it resulted from two decades of scientific investigations.

In his book [13], Kempelen presents at length his theory of speech production. Only the last chapter is devoted to the machine: in his mind the two are intimately related. His first trials date back to 1769. With a variable volume resonator and a bagpipe reed he succeeds in imitating the sound of some vowels. Then he makes several resonators producing the vowels /a/, /o/, /u/, as well as others in two articulated parts, producing the consonants /p/, /l/, /m/. These elements, fixed on an ogan windchest and put into action by a set of keys, constitute his first machine which is presently at the Vienna museum. But it is a failure. The sounds do not connect with each other, and the emission of the vowels is preceded by a sort of explosion which does not resemble a speech sound. Thus he gives up the result of two years of work, which probably exhibit the same kind of defects as those observed in Mical's machine, and resolutely adopts an anthropomorphic design: since Nature has provided us with a single glottis and a single mouth, it must be the same in a speaking machine.

The final machine, as described in the book, is composed of a bellows, a free reed, a windbox, a rubber open resonator and two openings, which play the part of the lungs, the vocal cords, the pharynx, the mouth and the nostrils (Fig 1). The operator's right elbow rests on the bellows and produces the air pressure. The right hand is busy with the different levers and openings on the top of the windbox, while the left hand more or less closes up the "mouth", the whole constituting an adjustable resonator. The reed length is fixed using a piece of wire, in order to produce a high pitched voice, attributable to a child.

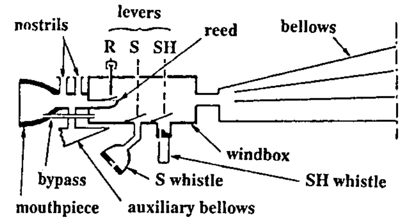

Fig 1 - functional diagram of Kempelen's machine

In order to pronounce /m/ and /n/ the mouth is closed up by the left hand and the sound escapes through one nostril (for /m/) or both (for /n/). This arrangement permits a smooth connexion between those consonants and the following sounds.

The phoneme /r/ is produced by creating some irregularities in the functioning of the reed: by depressing a key a brass needle is brought into contact with the reed during its vibration. The needle rebounds against a wooden stopper, which limits its course. This produces a scraping noise, the duration and intensity of which are adjusted by the operator according to the depression of the key.

The phoneme /l/ is produced by a quick movement of the left hand; two fingers are introduced into the mouth until the reed canal is partly closed up, so as to divide the airflow in two for a short time.

The phoneme /p/, and the other unvoiced plosives, are produced by closing up the mouth and nostrils, then by rapidly removing the left hand, which goes into the position necessary for the next vowel. In order to prevent any vibration of the reed during the occlusion Kempelen found it necessary to balance the pressures using a narrow pipe, which acts as a bypass of the reed canal. In addition, a small bellows located under the glottis increases the efficiency of the compression and contributes to a better restitution of the burst. The voiced plosives are produced in the same manner, with an extra airflow through one slightly uncovered nostril, so as to allow the reed to continue its vibration during the occlusion.

The fricatives /s/ and /ʃ/ are produced in accessory whistles located on both sides of the windbox, put into action by two valves controlled by levers. The phoneme /f/ is produced very simply by the unavoidable airflow losses which subsist when the mouth is closed up and the air pressure increased. The aspirated /h/ is obtained with the mouth open and a pressure low enough not to make the reed vibrate.

At the time of the demonstrations, the machine was covered with a small wooden box with two openings through which the operator could pass his hands. The top of the box was made of fabric. According to Kempelen the purpose of the box was to protect the machine from dust as well as to provide a passage for the sound. It might also have had the purpose of surrounding the machine with mystery, as was usual at that time, unless it was intended to prevent any imitator from copying its mechanism. The inventor describes his results in the following way (translated from [13]):

"... Although imperfect, it at least gives some good principles for designing a more perfect one. Finally I have brought it to the point where I can make it pronounce at the first trial and without any exception all of the Latin, French and Italian words that are proposed to me, some, it is true, better than the others, but at least several hundred words clearly and distinctly. For instance: Papa, Maman, Marianna, Roma, Maladie, Santé, Astronomie, Anatomie, Chapeau, Racine, Soupé, Charmante, Opéra, Comédie, Pantomime, as well as long and difficult words such as Constantinopolis, Monomotapa, Missisipi, Astrakan, Anastasius, etc... As for complete sentences, I can only pro-

nounce a few of them, for instance: Vous êtes mon ami - Je vous aime de tout mon coeur - Leopoldus Secundus - Romanorum Imperator, etc...".

These results were confirmed by numerous contemporaries. Grimm, the writer, who saw the machine in 1783, testified as follows (translated):

"... As it is today, the machine already clearly answers several questions; its voice is pleasant and soft; only the Rs are pronounced in a guttural way, with a tedious snoring noise. When one has not understood its answer, it repeats it again, but with a tone of infantile irritability and impatience... The pronunciation of Mr Abbot Mical's machine is far from being as distinct as that of Mr Kempelen's machine...".

Kempelen's machine was imitated several times. The copy which is exposed at the Deutsches Museum of Munich does not fit exactly with the description given in the book. It has two extra levers on the top of the windbox, one of which seems to control the length of the reed, that is the pitch of the voice. This kind of improvement might have been worked out by Kempelen himself after the publishing of the book. Another reconstitution was attempted by the physicist Sir Charles Wheatstone, some sixty years later [5, 14]. We ourselves made a reconstitution in order to check some points [10, 11]. In particular we could verify that the vowels were restituted only as crude approximations, except for /a/ and /o/, the device permitting only the creation and variation of a single wide formant in the 1000-2000 Hz region. Similar observations were made by Van den Broecke, who made a replica around the same time [12]. The consonants require some manual skills, which necessitate a long training period.

2.4 - Faber's machine [2, 6, 8, 11]

Kempelen claimed to have made his machine *"for the benefit of some master's hand, who would know how to raise it to the highest degree of perfection".* Actually his machine was missing a tongue and teeth, and he had envisioned the great improvements that variable pitch and keyboard control would have brought.

His real successor was Joseph Faber, a professor of mathematics born in Vienna around 1786, who could possibly have known Kempelen directly. He probably read Kempelen's book, because his machine comprised many similarities, as well as some of the improvements mentioned above. This machine, called "Euphonia", finished in 1835, was presented in several European main cities over a period of twenty years by Faber himself, then sold to Barnum, the famous show director. It appeared again in Paris around 1880, and probably lies today as a wreck in the basement of the Paris School of Medicine.

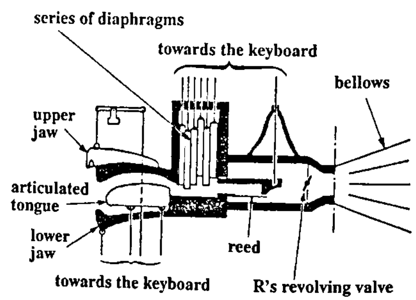

Fig 2 - Faber's machine diagram, after Du Moncel

This machine spoke, with normal or whispered voice, and sang "God save the Queen" ! It comprised a foot-manipulated bellows, a tongue, articulated jaws made of a flexible material, and a set of six diaphragms which modified for each sound the shape and section of the vocal tract (Fig 2). The controls were grouped on a 14-key keyboard; pressing down each in turn, one could obtain the following sounds : /a/, /o/, /u/, /i/, /e/, /l/, /r/, /v/, /f/, /s/, /ʃ/, /b/, /d/, /g/. Each key simultaneously controlled several parts, in fixed proportions. In other words, the control of Euphonia was practically phonemic in nature, even though some controls remained physiological or acoustical (nasality, voicing). The phoneme /r/, produced by a modulation of the airflow, was probably more plausible or pleasant than Kempelen's. The reed was articulated in such a way that the vibrating length could vary so as to control intonation.

One cannot help being impressed by the accuracy and modernity of the design of this machine. The part corresponding to the vocal tract was about 15 centimeters long. The rear portion of the tract was defined by a set of six sections of variable area and shape. The front part was defined by three parameters controlling the positions of the jaws and tongue, which evoke the now classical parameters of openness, aperture and place of articulation. Considering such a richness it is probable that this machine had the capability of working out realistic formants and transitions from one phoneme to the next.

Unfortunately nothing remains of Euphonia, except some descriptions which do not come from Faber himself and as such are necessarily superficial. It is certain that Euphonia was much more sophisticated than Kempelen's machine.

2.5 - Dudley: the rupture [4, 7]

A century passed before the reapparition of a speaking machine, which marked simultaneously a change of technology and a change of design. In 1937 Homer Dudley and his colleagues Riesz and Watkins, engineers with the Bell Telephone Company, finished out the VODER (VOice DEmonstratoR), which was exhibited to a large public in 1939, at the San Francisco exhibition and at the New York World's Fair. Externally the Voder looks somewhat like Faber's machine (an operator playing speech on a keyboard), but it differs from it in two respects. The first one deals with the physical nature of the vibrating phenomenon: sound is processed through its electrical analog, of which telephony has shown the equivalence as far as transmission is concerned. This analogy remains widely used nowadays, even though the signals are processed digitally. The other aspect concerns the parameters controlled by the operator, which are strictly related to the pitch and spectral envelope, without any reference to the vocal tract physiology or to phonetics.

The Voder is well enough known (Fig 3) for us not to spend too much time on its description. Let us just mention that the operator uses keys to control the signal amplitudes in ten spectral bands. Four extra keys and a pedal are used to control the occlusions and bursts of the plosives, the voiced/unvoiced feature and the pitch.

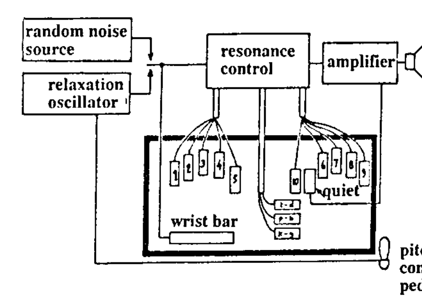

Fig 3 - Principe of the Voder, after Dudley

The Voder was so difficult to operate that the 24 telephone operators selected among 320 to demonstrate the machines had to receive a year of intensive training. The result was conclusive, however, and the operators could effectively play intelligible sentences on their machines, and even make them sing.

The Voder cannot be separated from the VOCODER (VOice CODER) presented by the same authors at the same time. In the Vocoder the control signals were automatically extracted from the analysis of real speech by means of a filterbank completed by a voicing and pitch analyser. The Vocoder is not a speaking machine, but a speech compression system. It is motivated by an economical stake, which will be the main driving force for speech research for forty years. Even nowadays the analysis/synthesis paradigm it illustrates remains prominent, despite a change of technology (from analog to digital) and some new methods of signal processing.

2.6 - Speech Synthesis today

We shall not present here an inventory which can be found elsewhere [7, 9], but only observe that the term of Speech Synthesis has substituted that of Speaking Machines. Speech Synthesis uses different types of knowledge according to the nature of the control parameters. At the lowest level these parameters are acoustic (they come or could come from a signal analysis very similar to the one implemented in the Vocoder) or articulatory (they represent the variables of a simulation of the vocal tract).

Controlling the synthesis process in terms of phonetic and prosodic parameters implies another step, in which the knowledge used, either explicitly (rules) or implicitly (segments), is related to the dynamic functioning of the vocal apparatus, partly guided by phonetic considerations. Let us recall here the large body of work that has been devoted to this aspect since the early fifties, in which Dennis Klatt took a major part [9].

Text-To-Speech Synthesis still requires a third step, mainly of linguistic nature (but not exclusively), which governs the interpretation of a text into oral terms. Finally, one should mention the Concept-To-Speech Synthesis, which could reveal itself to be extremely rich within the next few years in the context of Man-Machine Communication, but on which very little work has been done as yet.

Let us just observe that, after a twenty-year period during which the speech synthesis problem was wrongly considered as practically solved - a side effect of the Vocoder paradigm -, a powerful renewal of interest has now appeared, at all processing levels. This is due to many reasons that exceed the scope of the present paper. In the discussion which follows we shall get back to our initial topic concerning the Speaking Machines.

3 - WHAT COMES OUT OF THE SPEAKING MACHINES SAGA

We shall now try to distance ourselves from this historic evolution, in order to emphasize some of its aspects in relation to the contemporary views on speech processing.

3.1 - The magical aspect of speech

Throughout the early history of speaking machines the divine or magical aspect of speech was prominent. Even the eighteenth century scientific efforts were not definitely cleared of any mystification. Is this aspect really absent from the contemporary speech research ?

The fact of having a machine pronounce only words or sentences known in advance may look like a kind of mystification. However it reveals something which is basic in human understanding. In everyday life it is rare for a message to be totally unpredictable. Even when it is poorly articulated a message can be understood if it is partially predictable in the situation context. Conversely, when a speaker knows that his interlocutor can predict some of the message, he does not have to take the care of a perfect articulation. Kempelen had understood that point and made some use of it: *"... One is particularly misled when one knows in advance the word that the machine has to say, and when it pronounces it one imagines to have heard it..."*. Let us mention that this effect is extremely disturbing when one works out rules or patterns in speech synthesis, and that it necessitates the use objective listening tests, free from any uncontrolled previsibility, be it phonological, lexical, semantic or pragmatic. On the other hand, it indicates that speech synthesis could, in some cases, be thought as a predicting and interacting game between the machine accomplishing a task and the operator who supervises it.

In the same vein, Kempelen knew very well that some phonemes were not correctly pronounced by his machine (he used to replace /d/, /g/, /k/, /t/ by /p/ or /b/), and it was in full awareness that he gave it a child's voice: *"... the childish voice of the machine is always advantageous to it. One willingly forgives a child who sometimes stammers the mistake of using one letter in place of another, and one satisfies oneself of having understood what he meant..."*.

Hiding the active part of the machine in a box, as well as using some of the tricks mentioned above, brings conjuring tricks in mind. But one has to remember that at that time, today's scientific criteria were not strictly defined. Curiosity, ingenuousness, the capacity for amazement, were as essentiel to progress as the scientific method in the rigorous sense. The very idea transmitting speech at a distance or recording it seemed to be a dream (the poetic notion of "frozen speeches" had been formulated by François Rabelais in 1548).

Even today, in the latter part of the twentieth century, it is not certain that our research activities in the field of speech are perceived as being totally free of something magical. People are always surprised to hear a machine speak. When they realize that it is genuine, they tend to ascribe to it intelligence, language, and feelings like a human's; while it only superficially reproduces some of man's linguistic abilities. And it is highly significant that the potential users of TTS synthesis nowadays expect a more "natural" voice, although this is not absolutely necessary in most practical situations. Also, what makes speech such a fascinating domain of investigation to us, the rigorous speech scientists of 1990 ?

3.2 - Is it necessary to imitate human speech production ?

A great debate, opened in the earliest times of the speaking machines, is still going on today. On one side there is the idea that imitating nature as best as we can must improve speech synthesis. The degree of imitation is evidently a function of current knowledge and techniques. In the Middle Ages it was thought that it sufficed to materially imitate a human head for it to spontaneously produce speech, and if it did not one added to it some artifice... Faber illustrates best the success of this anthropomorphic view, which manifests itself today in the articulatory models. On the other side are the functional approaches, according to which it is not the conformity of the model to the original that counts the most, but the very result, the function, obtained by using different materials and techniques. The Voder is a perfect illustration of this view, in several respects: direct modification of the spectrum, ignoring the vocal tract functioning, and use of electronics to simulate acoustical phenomena.

We have no intention here of choosing between these views, each of which has its own merits and limitations. Obviously it is impossible to strictly imitate Nature in all respects; even if one succeeded in reconstituting a system presenting all the physical and physiological properties of the human apparatus, one would only have pushed the problem a little farther, because one would have then to build the equivalent of its nervous control, as well as the proprioceptive and auditory organs which allow it to learn and function. Conversely, any purely functional approach quickly encounters some limitations due to a lack of knowledge concerning the real vocal apparatus, which constitutes one of the possible realizations of the function that is investigated. It must be observed that, generally, the two views complement rather than oppose each other. In this spirit, Kempelen, after the failure of his first machine, succeeded by imitating the human speech production more closely. In a reciprocal spirit some notions which come from the physiology of speech production, such as the notion of a formant and the vocal source/vocal tract duality, are a great help in the functional systems as first approximations or sources of knowledge to be used in another form.

Let us observe that such a debate is extremely general and could concern speech analysis, visual perception, pattern recognition, artificial intelligence as well.

3.3 - What sense does it make to play a speaking machine ?

The notion of a speaking machine as a "speech instrument", in the sense of a "musical instrument" was abandoned after the Voder, to the benefit of systems using delayed controls. Interesting questions can be asked about the causes of this disappearance.

Firstly, the only use of a "speech instrument" is the demonstration that it is possible to play it. Mute people can be divided into two groups. For the first, the difficulty in speaking is due to an auditory deficiency; replacing their vocal apparatus by a manual device does not change anything, inasmuch as they cannot learn how to control it. For the second the problem comes from the poor functioning of the vocal cords; an artificial larynx is sufficient in that case.

Secondly all the realizations in the past have shown that it is extremely difficult to learn how to play such an instrument, which moreover delivers a result greatly inferior in quality to the normal production of any human being.

Does this mean that the notion of "speech instrument" is of no interest today ? Maybe not, because the interest of spontaneity, of real-time interaction, of the individual and expressive aspects of the voice have been forgotten a little too quickly. For a speaking machine to raise some practical interest its control should be simple to learn (i.e. it should be phonemic or syllabic in nature), and some expressive capabilities should be available. It should allow for the generation of several voice timbres, as well as of non-speech sounds, for musical or sound engineering purposes.

4 - CONCLUSION

Several speaking machines have been built in the past. For some of them it is beyond doubt that they worked satisfactorily for the listeners of their time. Their inventors had understood in an empirical way some of the structures of speech. Kempelen had understood the importance of the proper linking of successive phonemes, as well as the principle of the separation between the glottis and the vocal tract, not to mention some of the perceptual phenomena related to speech communication. Faber added to that a much more sophisticated modelling of the vocal tract, as well as the successful realization of a phonemic control using a keyboard. Dudley demonstrated with the Voder the possibility of reconstructing speech using electrical signals, without any reference to the physiology of the human apparatus, and without using a phonemic control.

The synthesizers which came later made it possible to experiment with pre-recorded controls; the concept was there in Mical's talking heads, but the technology and knowledge were too primitive to permit any serious investigation at that time. This evolution is now oriented towards Text-To-Speech synthesis, under the pressure of computer technology. During the progress in this direction some aspects of speech communication have been neglected. Spontaneity, expressivity, flexibility, interruptibility, underlying intelligence and sensitivity are expected by potential users of synthetic speech, because they want it to imitate human speech at a deep level.

Making a computer speak is not the same as playing a speaking

machine. Speaking machines, in the sense of "speech instruments", have disappeared, for they lacked the capability of being easily played and of being useful for something. With them disappeared the idea of a direct and instantaneous control by a human operator, as well as the idea of feedback from the listener to the operator, since everything has been frozen into the algorithms. Another aspect which has also disappeared, due to the success of the electrical analogies, is care for the real physical phenomenon of speech, which refers to fluid mechanics: what synthesizer, what articulatory model, takes into account the physical phenomena created by the airflow in the vocal tract, the role of the saliva, the directivity of the sound in the three dimensions of the space ? On those points as well as on a few others it may prove fruitful to adopt once more the naive attitude of the pionneers.

5 - REFERENCES

1 - BREWSTER D. : "Letters on Natural Magic", London, 1832.

2 - CHAPUIS A. et GELIS E. : "Le monde des automates", Paris 1928.

3 - DE MILLY, LE ROY, LAVOISIER, LAPLACE, FERRIEU, VICQ D'AZIR, Compte-rendu de l'Académie des Sciences, 3 septembre 1783.

4 - DUDLEY H., RIESZ R.R., WATKINS S.S.A. : "A Synthetic Speaker", J. Franklin Inst., **227**, 739-764, 1939.

5 - DUDLEY H. and TARNOCSY T. "The Speaking Machine of Wolfgang von Kempelen", J. Acoust. Soc. Amer. **22**, 151-166, 1950.

6 - DU MONCEL T.: "Le Téléphone", Paris, 1887.

7 - FLANAGAN J.L. : "Voices of Men and Machines", J. Acoust. Soc. Amer. **51**, 1375-1387, 1972.

8 - GARIEL : "La machine parlante de Mr Faber", Journal de Physique Théorique et Appliquée, T 8, 1879.

9 - KLATT D.H. : "Review of Text-To-Speech Conversion for English", J. Acoust. Soc. Amer. **82**, 737-793, 1987.

10 - LIENARD J.S. : "Reconstitution de la machine parlante de Kempelen", 4° Conférence Internationale d'Acoustique, Budapest, 1967.

11 - LIENARD J.S. : "La machine parlante de Kempelen", bulletin n°34 du Groupe d'Acoustique Musicale, Université Paris VI, mars 1968.

12 - VAN DEN BROECKE M. : "Wolfgang von Kempelen's Speaking Machine as a Performer", in Sound Structures, M. Van den Broecke ed., Foris publications, Dordrecht, 1983.

13 - VON KEMPELEN W. : "Le Mécanisme de la Parole, suivi d'une Description de la Machine Parlante", Vienne, 1791.

14 - WHEATSTONE C. : "The Scientific Papers", London, 1879.