A reconstruction of Farkas Kempelen’s speaking machine

Nikléczy P.  Olaszy G.

Kempelen Farkas Speech Research Laboratory, Hungarian Academy of Sciences, Budapest, Hungary
n inkleczy,olaszy@nytud.hu

Abstract

The first “speaking machine” of the world was created by the Hungarian polyhistor Farkas Kempelen. He can also be referred to as the first phonetician in the world. He went on improving his speaking machine for twenty-two years, and described the final version in a book published in 1791 in Vienna. The reconstruction was made based on this book. What we wanted to make was not just an exhibition piece but a machine that actually worked. Thus we can go back more than 200 years and study the working of one of the most precious instruments of the Baroque period. We can try out the ways of producing sounds that Kempelen wrote so many pages about in his book. The acoustic patterns of the machine’s speech can be studied by today’s sophisticated signal processing methods and prove or disprove Kempelen’s claims by measurement data. Besides these we took it to be an important task in terms of the history of science to contribute to our knowledge of the beginnings of phonetic research.

Introduction

Farkas Kempelen (1734–1804), who created the first “speaking machine” of the world, was a famous scientist and inventor of the late 18th century [1]. He lived and worked in a world of mechanical instruments of the Baroque period and, with his novel ideas, he made outstanding achievements and stretched the possibilities of the day to the utmost. He can also be referred to as the first phonetician in the world as he created the first mechanical instrument ever that was able to produce sound effects that were rather similar to human speech [2]. He studied animals’ cries, non-speech human sound production (whistling, hawking, snoring, kissing), and even speech defects, in order to better understand the mechanism of human speech production. On the basis of his results, he also made some proposals concerning the treatment of speech defects. His descriptions and the copperplate engravings in his book provided his readers with otherwise unavailable information about the speech organs, their operation, and participation in the articulation of the individual speech sounds, as well as their “imitations” in the actual speaking machine.

1. The speaking machine

Kempelen started building his first speaking machine in 1769 and he worked relentlessly on the implementation of his grand plan: “I continued my experiments with undescribable patience and I am amazed to this very day at how I was able to work on my plan for months on end without ever making a step forward. However, the certainty that speech must be imitable strengthened my perseverance...” [3]. After twenty-two years of research and development, he completed the final product that he then described in his book. During his initial experiments, he thought that in order to imitate the individual vowels he should operate separate devices like the pipes of an organ. In 1773, he was already able to produce four sound types resembling a vowel each with this method, but with respect to [1], he failed. It was then that he hit upon the principle “one voice – one mouth”, that is, the idea that the same basic mechanisms (of voice production and vocal tract) are to be used for the production of the various speech sounds. The scientist let four years’ results be lost without hesitation to try producing speech with this new approach. He wrote about this in his book as follows: “...I had to start everything anew, but I did not mind the effort or the costs because I thought that the six letters that I gained by them and that were to shed light on my new dim path in what followed, amply rewarded me. But the matter did not remain at that; with the tiresome work of my own hands, I constructed a lot more and threw them away again. ... If I wrote down all my failures in as much detail as the above, my book would easily run into another extra volume... Let it suffice to say that, all in all, I threw away as much machinery with ease as would be hard to carry off even with the help of a strong horse.” After many years of experimentation, the idea of a serial machine fitted with a bellow (see Figure 2, based on Kempelen’s original engravings) and following the principle of human sound production finally took shape.

From the bellow, air flowed into the “voice box”, an air-tight wooden element (11 x 7 x 6 cm). Within the box were to be found a construction of a vibrating reed for voice production, as well as two valves (in fact 2 small wooden boxes with a cover) on the right side one and on the left side one, the cover of which could be opened by the help of keys on top of the “voice box” (the valves were shut in their position of rest). As
a continuation of the vibrating reed, Kempelen installed a tube corresponding to the pharynx, from which he opened two nostrils (these protruded upwards as two additional little tubes). That tube disembogued into a rubber funnel (articulation channel) that stood for the oral cavity and through which voice left the machine. It was by altering the shape of that funnel by hand that sounds resembling the various vowel qualities could be produced.

![Figure 2](image1)

Figure 2
A drawing of the speaking machine (upper part), and the “voice box” with the various keys on it (lower part). The ‘c’ part represented the rubber funnel

The two valves inside the “voice box”, when open, let the air flow towards the right and left side of the box, respectively. Kempelen attached a resonator made from a flute to the left side of the box and tuned it so that it produced a fricative like [ʃ]. To the right side, he attached another resonator that produced a low intensity noise resembling [f] as air flew through it. In order to produce a trilled [r], he devised a special procedure. He made use of the vibration of the reed that was originally to produce voice. He fitted a very thin vertical metal pin right above the reed (0.5 mm away). On top of the box (where the keys to the two valves were also located) he made a third key to produce [r]. By pressing the key, the metal pin was allowed to touch the reed, whereby it started moving up and down and repeatedly tapped at a little piece of wood that the key ended in. That was how a trill was produced, even if it resembled a uvular [R] more than it did the intended [r].

The speaking machine was to be operated by both hands. The right elbow was used to press the bellow, whereas the fingers of the right hand had to stop or release the two nostrils and operate the three keys at the same time. The five fingers had the following duties: the thumb had to operate the [r] key; the index finger was responsible for the [ʃ] key; the middle and ring fingers served to stop and let go the nostrils; and the little finger to press the [f] key. The left hand served to adjust the shape of the rubber funnel or to stop it completely (for instance, when an [m] or a voiceless fricative was to be produced). In the book, Kempelen gives detailed descriptions for each sound specifying what movements and fine motions had to be performed in order to produce the given sound. As can be seen from this brief summary, it must have taken complex and well-rehearsed sets of movements to produce intelligible syllables or words with this machine. Consequently, long practice was needed before one’s fingers, hands and elbow could work with the required precision and simultaneity. Kempelen wrote, “In three weeks’ time, amazing skill can be gained in playing the instrument…” [3]. Kempelen’s machine talked in a child’s voice, and it was possible to produce words in various languages with it: “I can make it repeat any French or Italian word that is pronounced for me; on the other hand, a longish German word always takes me a lot of effort and only rarely do I succeed in making it quite comprehensible”[3].

2. The reconstruction

Our reconstruction of Kempelen’s machine was based on the Hungarian version of his 1791 book and on our visits to the Deutsches Museum, Munich, and to the Hochschule für ange wandte Kunst, Vienna. (Kempelen had built a number of machines during his experiments. One such earlier machine of his is kept in Munich, in the Deutsches Museum.) The replica was designed and built by the authors of the present paper. Our aim was to build a copy of Kempelen’s last and final version of the machine, the one he described in his book in detail. What we wanted to make was not just an exhibition piece but a machine that actually worked. This was important for us as a number of reasons. If the machine works, we can produce sounds (or even sound sequences) with it. We can try out the ways of producing sounds that he wrote so many pages about in his book. Thus we can go back 210 years and study the working of one of the most precious instruments of the Baroque period. The acoustic patterns of the machine’s speech can be studied by today’s sophisticated signal processing methods and prove or disprove Kempelen’s claims by measurement data. It was another important aspect of our decision that we were to present researchers of speech and of Kempelen with an instrument that had been unavailable so far. We took it to be an important task in terms of the history of science to contribute to our knowledge of the beginnings of phonetic research. And finally, we wanted to erect a monument, as it were, to the memory of Farkas Kempelen and his work in phonetics.

![Figure 3](image2)

Figure 3
The working replica of the speaking machine built in 2001
The machine was reconstructed in its original dimensions (100 cm x 40 cm x 40 cm). During the reconstruction work, we had to pass through dead ends similar to the ones Kempelen had to experience. This was because what we had to replicate was a 210-year-old mechanical instrument, which did not physically exist, and even though we knew its basic features from the contemporary description, certain details of how to build it were missing. Here are a few examples. The air-tight attachment of the bellow or the design of its valve was not described by Kempelen. He also failed to mention what kind of wood he used for the voice-box. The connection between the keys that operate the valves and the valve-caps had to be established such that the voice-box remains as close to air-tight as possible. The hermetic sealing of the voice-box is important because the air flowing in from the bellow has to result in an increase of the pressure of the air that either makes the reed vibrate or vigorously passes through one of the side valves (if they are opened). If the voice-box leaks and some of the air escapes, the machine does not produce the appropriate sounds. Another problem was caused by the determination of the size of the supplementary bellow that serves to amplify the release of voiceless stops and the force of the backlash spring. This is because, with unsuitable size and spring pressure, the apparatus does not get filled with air or does not give appropriate supplementary pressure for stop release. The reconstructed “voice-box” is shown in Figure 4.

A similarly grave problem was to find the appropriate solution to how to shape the vibrating reed that produces voice. We have, again, followed Kempelen’s detailed description and drawings, but we also used today’s technology (for having voice spectrograms displayed, determining frequency components, etc.). We made and tried several reeds. (We had to build each into the machine separately, itself not a simple task given that the reed is located in the most central interior part of the machine.) The size, thickness, and way of fixing of the reed all bear upon the quality of the voice produced. Kempelen admitted that he himself had had difficulties with making a reed producing the appropriate vibrations. He could only guess that, in order to produce speech sounds, he needed a fundamental vibration that resembled the air vibrations produced by the movement of vocal cords.

Today we would say that the vibration had to exhibit an intensive structure of harmonics. Although he did not have an instrument by which he could have investigated this, he had good enough sense to apply the vibrating reed of a common tavern bagpipe as a substitute for vocal cords. Instrumental measurements also helped our work in producing sounds and sound sequences. Today, a lot more is known about both the articulation mechanisms and the acoustic patterns of the individual speech sounds. For instance, in producing vowels, we determined the size adjustments of the rubber bell standing for the mouth, respectively the hand movements needed to achieve them, with feedback from the resulting formant structures taken into consideration. This facilitated our work greatly, especially since Kempelen’s own descriptions gave but rough guidance: “All vowels are formed with only our left hand, that is, they are determined by a larger or smaller distance of our hand from the edge of the funnel. I cannot exactly give how much our hand has to withdraw for each, since a lot depends on whether the elastic funnel has a larger or smaller diameter; on the other hand, the distances that our hand has to keep will be easily found out by practice and by the help of our hearing. All I wish to give is roughly at which parts do we have to look for each vowel and in what sequence these follow one another by way of the increasingly smaller openings.

For the vowel A that I take to be the basic sound of speech, our hand is quite far from the opening of the mouth so that the voice can leave freely.

Next comes E. For this, we bend our hand and put it on the lower edge of the mouth so that it is roughly an inch away from the upper edge.

For O, more than a half of the hollow of our hand is placed close to the upper edge of the mouth.

For U, we put our palm quite close to the whole mouth opening but in a way that it is not fully covered, so that the voice is still heard.

For I, we press our palm hard against the full rim of the mouth and keep only our index finger as far as to make a small opening at its lower joint through which voice has to rush forth, with somewhat more force than for the other vowels”.

3. Working with the machine

The formant structures of the vowels we produced are shown in Figure 5. With the machine, after some practice, we produced short sound sequences in Hungarian (mama ‘Mom’, papa ‘Dad’, sás ‘edge’, etc.), German (es war ‘it was’), French (je t’aime ‘I love you’), and English (I go). The oscillograms and spectrograms of Hungarian mama and papa are shown in Figure 6.

Our results have confirmed that Kempelen must have been able to produce sounds and sound sequences on the basis of the guidance he gave in his book. The “speech” we were able to produce was ragged and slow since to produce each sound we had to perform the appropriate movements with both hands in a split second (today’s average speech rate is 13–16 sounds/s for Hungarian). Kempelen himself mentioned this physical limitation of the method. Hence, the machine is only suitable for producing short utterances. Sound quality was
furthermore influenced by the actual speech sounds that the sequence to be produced contains.

![Figure 5](image1)

Figure 5
Oscillograms and spectrograms of vowels produced by the reconstructed machine (á=[a], a=[ə], o=[o], ü=[y], e=[ɛ], ö=[œ], i=[i])

![Figure 6](image2)

Figure 6
The structure of the sequences mama ‘mam’ and papa ‘dad’ produced by the reconstructed machine (oscillograms and spectrograms)

Of the stops, the most adequate imitations can be achieved for [b], [p]. Of the fricatives, the easiest to produce are [ʃ], [f], since these have separate keys. Their intensity compared to that of the vowels, however, falls short of the required value due to the structure of the machine, a fact that makes their inclusion in sound sequences rather awkward. Hence, in sequences containing fricatives, the intensity of the latter is a lot lower than would be desirable. The sound [v] can be produced in a quality comparable to that of vowels, provided the hand movement is sufficiently swift. Affricates are difficult to produce, since their component phases require changes executed in such a short time that the necessary skill needs a lot of practice to acquire. The sound [l] can also be produced but in sequences the position of the hand has to be changed extremely fast. Of the trills, it is only the uvular type that can be produced (this is quite advantageous for German, though), given that the metal pin has to touch the vibrating reed and, while it takes over its vibration, it also hampers the movement of the reed to some extent, due to its own weight.

4. Conclusions

The reconstruction of Farkas Kempelen’s speaking machine is primarily of great importance for the history of the science of phonetics. The machine is in working order, therefore it makes it possible for us to produce speech sounds and sound sequences by eighteenth-century technology, the way Kempelen himself produced them in his time. Practicing the use of the machine has pinpointed difficulties that Kempelen also must have had to face. It has been demonstrated that Kempelen was only able to produce isolated sounds, syllables, words, and possibly very short sentences with his machine. Kempelen’s idea of mechanical speech production seems to be revived in the twenty-first century. Experiments are conducted by Hideyuki Sawada at the Kawaga university in Japan to produce mechanical voice production [4]. A small (roughly man-size) mechanical apparatuses are meant to produce speech. The application will serve as the voice of robots. The instrument imitates the human vocal tract with a suitably sized tube. The air needed for voice production is supplied by a small but high-performance pump, and the alteration of the cross-section of the vocal tract is done by stepper motors. Controlling is done by a microprocessor. The first experimental results show that vowels can already be produced in an appropriate quality.

Acknowledgements

The authors thank all the help that has been given to the reconstruction by the Deutsches Museum, Munich, and the Hochschule für angewandte Kunst, Vienna. This reconstruction was sponsored by Millenáris Kht., and the reconstructed machine was subsequently exhibited in Budapest for more than a year as part of the exhibition “Dreamers of Dreams; World Famous Hungarians”.

5. References