KEMPELEN’S SPEAKING MACHINE: EXPERIENCES WITH REPLICA

Fabian Brackhane\textsuperscript{1}, Mária Gósy\textsuperscript{2}

\textsuperscript{1} Institut für Deutsche Sprache (Mannheim), Germany,
\textsuperscript{2} Research Institute for Linguistics, Hungarian Academy of Sciences (Budapest), Hungary

brackhane@ids-mannheim.de

Abstract: There are a number of recent replicas of Wolfgang von Kempelen’s speaking machine. Although all of them are explicitly based on Kempelen’s own description nearly none of them are identical in construction and sound. In this paper we want to illustrate some of these differences and their reasons for five replicas built by ourselves.

1 Introduction

Wolfgang von Kempelen’s speaking machine presumably is the most famous, but certainly the most important early approach to create artificial speech. Apart from its strong reference to human anatomy this is also due to the fact that the construction had been extensively described in his also popular book “Mechanismus der menschlichen Sprache” (1791) [1].

Based on Kempelen’s description very early replicas have been made. The first one we know of dates around 1797 [2]. But the fascination with this device remains until today. Exemplarily from 1967 to the present day at least 11 replicas have been made by various researchers (see section 1.1.3). Interestingly, despite Kempelen’s description of building the machine in his book nearly no one is like the other. This can be explained by various reasons including the facts that there are some parts of the machine that (i) lack the detailed information and (ii) are not clearly described and/or illustrated in Kempelen’s own book, such as the missing information about the outer structure and material of the machine or the vague description of the reed pipe.

The goal of this paper is to discuss some constructional details of five 21\textsuperscript{st} century replicas compared to those of the original machine as described in [1]. We want to highlight the constructional consequences of some insufficient information of Kempelen’s own description. In addition, we will present some acoustic characteristics of the replicas’ sounding that could be a basis for further research of the replicas. Similarities and differences in the constructing processes could help to better understand the strengths and weaknesses of these replicas of a 226-year old mechanical speaking machine. For space reasons, we focus on the representations of four organs of articulation, such as lungs, glottis, nose and mouth and, additionally, on the creation of fricative and trill sounds. We will show that (i) all replicas have more similarities than differences, and (ii) their ‘child voice’, as was described in Kempelen’s era, can be proved by acoustic analysis.

1.1 Sources about the speaking machine

There are mainly two sources that give us reliable information about the construction of Kempelen’s approach to speech synthesis. Firstly, the reports of Carl Friedrich Hindenburg [3] and Johann Jacob Ebert [4], which are strongly related to each other and therefore could be counted as a single source. Secondly, of course, Kempelen’s own description in [1].
Hindenburg’s and Ebert’s publications are strongly related to each other.\footnote{Both were published by the same publisher within a year. Additionally, Ebert refers to Hindenburg’s book and replicates partly his illustration of the speaking machine.} They described an early version of the speaking machine presented to the public in 1784 during Kempelen’s journey through Europe presenting his “Mechanical chess player”. Both descriptions are rather short and partly hard to understand. Obviously neither Hindenburg nor Ebert had the opportunity to fully understand the operating principle of the machine. Nonetheless it is obvious that the device both authors had seen was very similar to the one described in [1] but seems to have been much more complex in its controlling (see Fig. 1).

Kempelen devoted the whole final fifth chapter of his book to his speaking machine and the history of its development. The description of the construction covers 42 pages with 9 additional scaled copper plates engraved by himself. At first glance this description seems to be very detailed and precise. However, if one wants to replicate the speaking machine based only on this information he becomes aware of many missing details. Some of these can be read off the drawings that do indeed show a high level of accuracy. Although Kempelen was known as a gifted illustrator and engraver some of his copper plates show shortcomings especially representing three-dimensional objects. Additionally, the “Mechanismus” seemingly describes no existing device but a more or less idealized (and reduced?) version.

From these sources it is not entirely clear what the speaking machine in its final form looked like. In the Deutsches Museum, Munich there is a speaking machine that at least seems to originate from Kempelen to a large extent. It has been copied conscientiously and shows more similarity to the descriptions of Hindenburg and Ebert in [3; 4] than to Kempelen’s own in [1].\footnote{For licensing reasons it is not possible to give a picture of this device here.}

A whole range of replicas has been made in the second half of the 20\textsuperscript{th} and early 21\textsuperscript{th} century (Table 1). Unfortunately only for some of them descriptions were published which furthermore mostly are not very detailed. For pictures of most of these replicas see [5].
Despite the fact that Kempelen’s device for creating artificial speech was not fully successful and at least is somewhat inconvenient, his speaking machine is a fascinating approach for understanding speech production by simulating speech by one’s own hands. Although nearly all of the speech organ representations show deviations from the human body, it was nonetheless the first approach to speech synthesis that takes the human anatomy seriously into account. Reading Kempelen’s description it is hard to imagine that this device could have anything in common with human speech.

On the other hand, the construction seems to be simple enough to just try it out of curiosity. Although one comes across many difficulties during the building process, the fascination grows side by side. In addition, the motivation of enhancing parts of the replica and/or trying different solutions leads to a more or less continuous work on this topic, and could shed light on further knowledge about some important nuances (odds and ends) of human articulation. Kempelen’s machine could be regarded as the predecessor of the present-day ‘articulatory synthesizers’ from several aspects.

All of the replicas basically rely on [1]. Nonetheless nearly none of them are identical but show changes in various details and various extend of their construction. In the following paragraphs we will explain the deviations in the construction of our five replicas compared to Kempelen’s description as well as the reasons of the deviations (Fig. 2). The speaking machine consists of three major parts which we will discuss separately: The outer structure (panel, box and bellows frame), windchest, and bellows.

## 2 Analyzed replicas

The five replicas we intended to compare are the followings: 1. replica developed in Budapest (called “replica 1” hereafter), 2. four replicas developed in Saarbrücken for several institutions (“replicas 2, 3\textsubscript{a}+\textsubscript{b} and 4”).

---

4 = faithful replica, 2 = with minor changes, that do not affect the functionality, 3 = with major changes, that partly affect the functionality and design, 4 = modern adaption

5 Personal message by David Howard.

6 Personal message by Jean Jeltsch
Figure 2. The replica models reported here: 1 (top left), 2 (top right), 3a+b (bottom left) and 4 (bottom right). The framework of 1 is intended to be very close to that of Kempelen’s original machine while replicas 2–4 show miscellaneous changes in constructional details.

The replica at Budapest (“replica 1”) was initiated by two phoneticians, Péter Nikléczy and Gábor Olaszy of the Phonetics Department of the Research Institute for Linguistics. The main reason for building the machine was to participate in the Exhibition for the Millennium of the Hungarian State in 2000. In addition, researchers wanted to get an impression of the acoustics of Kempelen’s original machine.

In 2007, the first author made a replica on the occasion of the 16th ICPHS congress in Saarbrücken (“replica 2”). This was on initiative of Jürgen Trouvain who has organized an accompanying exhibition on the history of phonetic sciences. Due to the complexity of the project the local organ building firm of Hugo Mayer (Heusweiler/Saar) was asked to support us with manpower and knowledge. The organ builder finally sponsored us by building the entire outer structure such as bellows, the foundation and the outer chest. Due to time constraints constructional compromises had to be made for the construction of some components such as the windchest and the reed pipe. These temporarily deviations will be illustrated too in the following sections.

In 2009, the first author was asked to build two identical replicas for the Technical University of Dresden and the Heinz Nixdorf MuseumsForum, Paderborn (“replicas 3a+b”). It was decided to continue cooperation with the organ builder for these replicas as well as for replica 4 made in 2017 for the Institut für Deutsche Sprache, Mannheim (“replica 4”). For these three replicas (“3a+b+4”) the organ builder was responsible for the outer parts, the bellows and the blanks of the windchests and the reed pipes.

Replica 1 was intended to be an exact implementation of Kempelen’s description. In contrast, replicas 2–4 were built according to pragmatic aspects. The overall constructional principle of [1] was retained, but we also wanted to be able to easily take the replicas apart for demonstration and (experimental) changes. So, some constructional aspects were adjusted to this aim. Replicas 2 and 4 were to be used for experiments and research and therefore conceptualized as constant work in progress (see also [10; 11]).

2.1 The outer structure
There are almost no details on the construction of the outer structures of these replicas. Such information is given by Kempelen neither about the material nor the dimensions. Only the proportions can be deduced from his scaled figures. As a consequence, the analyzed replicas show differences in this respect.
For replica 1, the outer structure, like all wooden parts, was made of teakwood. This kind of wood was judged to be the most appropriate one among the available hardwoods of the time. For replica 2, varnished spruce wood was used for cost reasons while for replicas 3\textsuperscript{a+b} and 4 these parts were made of waxed oak wood which is a very robust material and represents a high quality. The entire outer structure of replica 4 was defined in order to be able to dismantled and transported easily in a compact package.

The panel on which all parts of the machine (except the windchest) are fixed was made by Kempelen in such a way that it contained two separate boards connected by six wooden bases. This was likely to compensate the low height of the tables at Kempelen’s time (see replica in Fig. 2 top left, the speaking machine is generally operated by a person standing at the machine). For replicas 2 and 3, we omitted the (second) bottom panel but retained the six wooden bases in each case (see Fig. 2, top right and bottom left). For replica 4, these bases were omitted, as well (see Fig. 2 bottom right).

The wooden box that encloses the windchest has no functional purpose. It just covers the windchest for protection. In addition, Kempelen obviously did not want his audience to see how the machine was operated. The measurements of the boxes for all four replicas were taken from the specimen in the Deutsches Museum.

Nowadays, it is important that the audience should be able to see how speech is simulated. Although its cover is movable, the box turned out to be a hindrance presenting the replicas. For this reason, the box of replica 4 was made partially of acrylic glass (side parts in the front and on the right). In addition, the parts of the box were not glued but only put together with wooden dowels. So, a dismantling and space-saving transportation is possible.

2.2 Windchest

The windchest of replica 1 was made again of teakwood. The upper side is not glued but can be twisted off, so the internal parts can be shown with some effort.

For the first prototype of replica 2, laminated wood was chosen for cost reasons. In addition, this windchest was made with an inner volume one and a half times of Kempelen’s measures in order to have enough space in the windchest for making changes of some components. As it turned out this enlarged dimensions caused resonance vibrations and therefore damped the vibration of the reed pipe. So, another windchest with the original inner dimension was made after having found the right working constructions for all inner parts. This time we used nut wood following the specimen in the Deutsches Museum.

For replicas 3\textsuperscript{a+b} and 4, material and dimensions of the windchest were maintained. However, for the last one the wooden top was replaced by acrylic glass so that the audience could see what is happening inside the windchest during operation.

2.3 Levers for fricatives

If the windchest is made following Kempelen’s specifications it becomes difficult to arrange the controlling levers for the valves of the fricative generators. They easily get into conflict with the “nostrils” if these are arranged too straddled (if they are arranged too narrow, this results in operating problems, too). For the connection between the levers and the valves, Kempelen speaks of “thin wire”.

Therefore brass wire was used for replicas 2 and 3. But it turned out that wire (of course) is not flexible enough to compensate for even very small inaccuracies in the arrangement of levers, valves and the corresponding drilling in the top of the windchest. For this reason, the wire was replaced by nylon thread for replicas 2 and 4. The developers of replica 1 chose a specific textile thread that is usually used by shoemakers when sewing shoes.
2.4 Device for /r/-sounds

While replica 1 is equipped for simulating trill sounds following Kempelen’s description, replicas 2–4 do not contain a specific “r”-device. The reason for this difference is mainly language-specific meaning that replica 1 is used particularly for simulating Hungarian speech while the other replicas are used mainly for German utterances. Since in German the phoneme /r/, in contrast to Hungarian, can be represented in most cases by the vocalic allophone /ɐ/ or at least by a very weak fricative /ʁ/, there was no imperative to build this rather delicate part into the machine. Kempelen admits that the machine is best in producing utterances in French and Italian (which both require consonantal /ʁ/-sounds) while German speech is more difficult to produce than the ones mentioned previously [1: 455].

2.5 Reed pipe

2.5.1 Beating reeds

Kempelen’s representation of the human glottis is closely based on so called reed pipes with beating reeds as used in pipe organs (which have an operating principle similar to a clarinet, see Fig. 3). Kempelen’s reed pipe, however, differs from typical organ reed pipes in some constructive aspects. The reason for this is unknown but according to [1] Kempelen had no deeper knowledge of organ building.

Replica 1 firstly was equipped with a reed pipe made as close as possible to the requirements in [1]. Because of their unsatisfactory functionality after a year of trials a second one was made with slight modifications.

For replica 2, in the beginning a modern reed pipe was used that was not made for this purpose (and did not follow Kempelen’s specifications) but was made from lead and brass instead of wood and also had slightly different proportions. In 2008, this could be replaced by a wooden one close to Kempelen’s drawings. However, also in this case it was impossible to make the reed pipe following exactly the drawing on Kempelen’s Tab. XVIII: The conic drilling of the block as shown there is nearly impossible to make if the reed pipe is made from one single piece of wood as demanded by Kempelen. For this reason it was decided to make a cylindrical drilling that of course had to have the same inner diameter as the height of the shallot.

For replicas 3a+b and 4, the organ builder developed a completely new solution containing a loose wooden shallot that is clamped in the cylindrical drilling of the wooden block (Fig. 3). In every case we made the block of the reed pipe and the segment of the nostrils from one piece of wood, deviating from Kempelen’s specifications. Just putting the nose segment on the block resulted in a constant instability.

Kempelen chose ivory as material for the pipe’s reed. This is very unusual and has no model in organ building. Ivory is not very flexible but brittle and therefore hard to thin out.
Figure 4. The free reed pipe for replica 2. While the nostril segment and the block of the proper reed pipe are retained (mid), the shallot is replaced by an adapter that contains a free reed from a reed organ (right). On the left: The mouth funnel.

For every five replicas, very thin ivory coats of piano keys were polished and used. Kempelen’s information about the reed’s thickness (“like a playing card” [1: 411]) is not helpful. In fact, the thickness had to be explored empirically. Here, it is necessary not to thin out the reed too much to preserve some degree of tension. For replica 1 it is about 0.2 mm, for replicas 2–4 it ranges from 0.25 to 0.4 mm. Also, it is very helpful to get some curvature at the top of the reed. Otherwise it will work unreliably and only with very low air pressure. The leathering of the reed’s bottom (as demanded by Kempelen) significantly hampers its vibration. For this reason the reed of replica 1 is not leathered.

2.5.2 Free reed

Almost simultaneously to Kempelen’s work Christian Gottlieb Kratzenstein had worked on a “vowel organ” that did at least produce the vowels a, e, i, o, u [12]. In contrast to Kempelen, Kratzenstein’s theory on vowel articulation and speech production in general was very defective. Although he also used a reed pipe for simulating the human glottis, his supposed findings on human anatomy led him to the development of a variant of reed pipes, the free reed. With this construction the reed does not rest on the top of a shallot but it is mounted in a frame where it can freely oscillate.

These reeds are distinguished by a very smooth sound without noise components, but in contrast to beating reeds no variation of pitch by increasing or decreasing air pressure is possible. However, in order to experimentally combine Kempelen’s and Kratzenstein’s approach, the first author build a free reed for replica 2. Manufacturing of a free reed pipe has to be even more precise than for a beating reed. For this reason, we reused a free reed from a reed organ and combined it with a custom-made groove (Fig. 4). The sound of the speaking machine using this free reed was smooth as expected but because of its absolute monotony it was less authentic than using beating reeds.

2.6 Bellows

Kempelen gave no exact description concerning the construction, measurements or material of the bellows. It seems that it was obvious for him that the bellows should be made like traditional organ bellows, just smaller. Their size can be derived from proportions on Tab. XXII in [1] to some degree. The bellows of all five replicas are 40 cm in length and 26/11 cm in width (v-shaped), respectively.

However, we know about the bellows design and its number of direct folds. It is a classical wedge-bellows with four direct folds. For replica 1, following Kempelen’s depiction four folds were adjusted. For replica 2, the organ builder made only three folds for financial reasons which results in a very tight air reservoir. As a consequence replicas 3a+b and 4 got bellows with four folds.

There is also no exact definition of the construction and material of the valves. We made not-return valves as in traditional organ bellows.
2.7 Mouth funnel

For the funnel representing the mouth Kempelen again chose a very unusual material. He describes it as the upper part of a bottle made from natural rubber. Since these bottles were not produced any longer we had to think about another material and thought of casting it in rubber. So, for replica 1, a specific rubber object was made by a professional rubber worker according to Kempelen’s figures.

Due to time constraints in 2007, a mug made from rubber was used for replica 2 (Fig. 4, left). The bottom was cut off and the rubber was rolled into the right dimensions. Although this was considered as a temporary solution it had proven itself so well that replicas $3_{a+b}$ were also equipped with funnels like this. Apart from aesthetic aspects these rubber funnels have the danger of leakage in itself if they are not made very carefully. For replica 4, this sort of funnel was replaced by a tailor-made version made by 3d printing. The flexible plastic available is completely sufficient for our purposes.

The inner diameter of the mouth funnel seems to be around 45 mm on Kempelen’s drawings. This part was slightly enlarged to 55–60 mm being more comfortable to use. This variation seems to have no influence on the sound.

3 Fundamental frequency and spectra based on speech of replicas

Kempelen’s information on the dimensions of the reed pipe (that is absolutely crucial for the speaking machine) is very vague (cf. section 3.4). Therefore, it is not surprising that the voice qualities of all the replicas are somehow different. However, they all have in common that they sound as the voice of a child, caused by a relatively high fundamental frequency of the produced speech. This phenomenon was described by Kempelen himself as a useful feature to conceal the machine’s deficits, such as the lack of fricatives and articulatory disfluencies [1: 442].

The factors causing the voice quality characteristic of a child are primarily the proportions and thickness of the ivory reed (cf. section 3.4). The different dimensions of the diverse reeds cause differences in the mean fundamental frequency of the replicas reported here. Replica 1 has a pitch of ca. 350 Hz while that of replica 2 is ca. 250 Hz (for both reed pipes). The fundamental frequency of replicas $3_{a+b}$ and 4 is about 500 Hz and 400 Hz respectively.

Considerable differences can also be seen in the spectrograms and spectra when comparing them (Fig. 5). The faithful reed pipe of replica 2 shows a more complex spectrum than replica 1 but it contains more noise components and less differentiated distribution of energy at the same time. For replica 4 the reed’s very high fundamental frequency results in a spectrographic depiction that differs significantly from the others and resembles the spectrograms taken of very young children who have high fundamental frequency in their speech. All three replicas analyzed here show an energy maximum around 1,700 Hz. Replica 1 has an additional energy maximum around 1,100 Hz.

4 Discussion

Kempelen’s speaking machine unquestionably is an important keystone for the development of mechanical approaches to speech synthesis that has preserved its fascination until today. Motivated by his seemingly detailed description in [1] a considerable number of replicas have been made during the last 50 years. In this paper, we have described and partly compared five of them made by ourselves.

---

7 Also many contemporary correspondents reported a similar sound quality [2: 154; 3: 51; 4: 51].
8 The reed of one specimen of replica 3 shows a very similar spectrogram.

[32]
Although Kempelen’s own detailed description is a solid foundation for building replicas, information is missing for a large amount of details that are of vital importance. Their exact construction could only be investigated by trial and error. This leads to several constructional differences of the replicas reported here. Additional differences are due to different intentions from which the replicas were made: True copies following Kempelen’s description in every detail versus modified replications, which are adjusted to their intended usage for research and teaching. So, we can conclude that the replicas show great similarity in their general design but they have noticeable differences in constructional details, and as a consequence, they show differences in their acoustic results.

Our acoustic analysis proved that all of our replicas produce speech in a way that shows strong similarities in quality to the speech production of a child. But the construction of the reed pipe – the representation of the human glottis – is much more difficult than Kempelen’s description suggests. Different approaches to implement the requirements as well as to achieve a reliable functionality caused differences in the synthesis quality of the replicas. However, all of them share a similar high fundamental frequency when producing speech. Each replica is a close and high quality imitation of the machine built 226 years ago demonstrating various solutions for several constructional challenges. There is no question on a competition among the replicas. They provide a very useful opportunity to be able to look at a mechanic construction and functioning of the human speech organs with the 21st knowledge of human articulation.
Further acoustic and perceptual research of all the replicas’ speech will show various interrelations between constructions and sounding. For this, an examination of more different replicas than the five presented here is desirable for a more comprehensive impression.

5 Acknowledgements

Thanks to Annette Klepp for improving the English version of this paper and to M.S. for always having great ideas.

References


