

A "POLYGLOTTAL" SPEECH SYNTHESIS – MODIFICATIONS FOR A REPLICA OF KEMPELEN'S SPEAKING MACHINE

Fabian Brackhane

*Leibniz Institute for the German Language, Mannheim (Germany)
brackhane@ids-mannheim.de*

Abstract: The speaking machine of WOLFGANG VON KEMPELEN (around 1770) is one of the most popular objects from the early days of phonetic sciences. But although there is a detailed design description by KEMPELEN himself, many details of his synthesis device are puzzling until today due to insufficient description, unusual materials and processing. In particular, this applies for the *reed pipe* and its *reed* as the representation of the human glottis. Therefore various constructional variants and materials have been used for that place in replicas, leading to the question whether these differences have a notable impact on the sound characteristics of the speaking machine. In this study, the influence of different materials and processing of the *reed* on the sound characteristics of a self-made replica of KEMPELEN'S speaking machine is presented.

Index Terms: Speech synthesis, history of phonetics, Wolfgang von Kempelen

1 Introduction

In the 18th century there have been several approaches to reproducing human anatomy in terms of speech production and, thereby, to synthesis of speech in this way. For instance, only slightly earlier than WOLFGANG VON KEMPELEN (1734–1804), ERASMUS DARWIN (1731–1802) and CHRISTIAN GOTTLIEB KRATZENSTEIN (1723–1795) have worked on this idea. But the speaking machine of WOLFGANG VON KEMPELEN is the most mature development dating from that time as well as the best documented one.

Regardless of its indisputable limitations, KEMPELEN'S speaking machine is one of the most famous and iconic objects of the early history of phonetic sciences. For his attempt to recreate the human anatomy, KEMPELEN used several components borrowed from contemporary organ-building. Aided by the fact that there is an extensive description of the machine and its structure by KEMPELEN himself [1], a large number of replicas and further developments was made since the early 19th century. Since the 1960s, over a dozen replicas have been made [3, p. 103]. Most of them tried to follow KEMPELEN'S instructions as best possible. But, for various reasons, sometimes there had to be compromises.

KEMPELEN'S idea of using technology from organ-building for the purpose of synthesizing human speech dates back at least until the late 17th century. In 1699, the French mathematician and philosopher BERNARD LAMY published his idea of generating human speech sounds by *organ pipes* [4, pp. 153–159]. During the 18th century, the *organ stop* called "vox humana" – made from *reed pipes* – more and more moved into the focus of consideration [5]. So it is not surprising that aspects of this discussion can also be found in KEMPELEN'S work.

At half of the modern replicas of KEMPELEN'S speaking machine a rather common compromise made concerns at this reed pipe, which is to simulate the glottis. These are operated



Figure 1 – A (left): Manufacturing steps for a reed made from ivory (raw piano key covering, thinned out, leathered). B (right): Reeds 1–6 (without leathering). Scaled figure.

with *reeds* that were not made of ivory as described by KEMPELEN but were made of brass [3, p. 103]. Reeds of brass have been common in organ building since the middle ages. So it is not known why he chose ivory instead. His method of further processing of the reed is also not known. In this paper we want to investigate the question of whether there is a significant influence of that reed's material and processing on the sound quality of the speaking machine.

2 Motivation

The reed pipe is a very crucial part of the speaking machine. This component is key for the assessment of the synthesis. But the information provided by KEMPELEN on the manufacturing of the reed pipe is insufficient: There are no exact measurements and the ostensibly scaled drawing on tab. XVIII in [1] proves to be problematic in perspective. But especially choice and processing of the reed's material pose a puzzle. Neither is ivory a common material for reeds, nor is it its further processing. Not only KEMPELEN equipped the upper side of the *shallot* with a thin layer of leather to damp the noisy parts of the sound (which is common in organ-building). He also leathered the bottom side of his ivory reed which is completely unusual. In addition, the thickness of the reed (that influences the sound perceptible) was described by him only vaguely ("like a playing card"), which is not least due to the then insufficient methods of measurement [1, p. 411].

For several modern replicas of KEMPELEN's machine it turned out that reeds made of ivory are difficult to put into operation (they work most reliably with a thickness of 0.2–0.4 mm).¹ Today, ivory is a material that is very difficult to obtain. The ethically most acceptable option is to use key coverings from old pianos (Fig. 1a). But such ivory has often been significantly affected by aging processes, rendering the ivory somewhat brittle and therefore even harder to put into operation. As a second problem, the abnormal leathering of the reed's bottom turned out to inhibit the already borderline operability of ivory reeds.

For those reasons, the use of ivory in the speaking machine's reed turns out to be a central problem. However, it is questionable, whether the (normally significant) influence of the reed's material on the sound of the reed pipe remains despite the double leathering of the reed pipe. It is conceivable, on the contrary, that that processing neutralizes the characteristics of the material to a certain extent (cf. [5, p. 111, footnote]).

During the years 2007–2017 the author had an opportunity for constructing four replicas of KEMPELEN's speaking machine. Each time, producing good (i. e. working) reeds from ivory has been a question of luck and chance. The fourth replica has been made in 2017 for the *Leibniz Institute for the German Language* (Mannheim, Germany). Besides reproducing KEMPELEN's approach to speech synthesis, this replica had been intended for exploring newly developed, innovative components, which had not been not possible with nearly all other existing replicas .

¹The thinner a reed is, the easier it will vibrate but the more vulnerable it is at the same time.

Here, again, arose the question whether it is even necessary to make the reed from ivory. Presumably, the leathering of the reed would minimize the influence of its material on the sound of the reed pipe anyway. So the sound quality of ivory reeds with and without leathering needed to be tested. In addition, conventional reeds made of brass and their influence on the sounding of the speaking machine should also be tested. Finally, TAKAYUKI ARAI'S innovative design of reed pipes was also to be tested as a possible alternative [6].

3 The speaking machine

3.1 Design

With his speaking machine, KEMPELEN tried to implement his insights into human anatomy and articulation processes. For synthesising human speech he tried to model the human vocal tract as best as possible by developing mechanical representations for the lungs, the thorax, the glottis, the mouth and the nasal cavity. This approach was limited, not least by technological restrictions of KEMPELEN'S time. For instance he failed to model a representation of the human tongue, which, of course, is crucial for speech production.² Because of the missing tongue and lips only a few speech sounds as /a, o, m, b/ are actually "articulatable" with the speaking machine. Other sounds have to be simulated by means of "substitutional articulation".

Another part of the speaking machine is just an approximated implementation of the human anatomy: The representation of the glottis, that consists of a reed pipe taken from classical pipe organ building (the functional principle can also be found in wind instruments like the clarinet or the saxophone) [1, p. 410]. In contrast to the human vocal folds, in reed pipes there is only one single vibrating element (the reed), not two. Therefore a *double reed* as used in wind instruments like the oboe or the bassoon would have been a closer analogon. Effectively, KEMPELEN tested double reeds with his speaking machine but he abandoned that without providing reasons [1, pp. 390–396, 399].³

The speaking machine is operated by a player in the manner of a musical instrument: While the left hand manipulates the oral cavity, the right hand operates the nasal cavity and fricative generators and the right elbows controls the airflow from the bellows. KEMPELEN'S own description in [1] gives a good first impression of its basic functioning but it turns out not to be as detailed and precise as it appears at first glance.⁴ Therefore, making replicas comes down to constructive decisions by trial and error even in important details.

3.2 The replica of the IDS

The replica used for the study presented here was made by the author in 2017 for the Leibniz Institute for the German Language (IDS) at Mannheim, Germany. Its construction of course is mainly based on [1], but is specially designed for experiments with alternative and additional components diverging from KEMPELEN'S specifications.

The reed pipe of this replica was basically designed according to the full-scaled illustration on tab. XVIII in [1] (which unfortunately is not well drawn in perspective, c.f. Fig. 2). Due to experiences with earlier replicas it had been modified in its shape (round but square) and not made from one single piece of wood. Instead, the shallot was made as a separate component. Its dimensions again were taken from KEMPELEN'S book. From his drawing it appears that shallot

²This circumstance was in fact discussed by KEMPELEN himself but trivialized at the same time [1, p. 432].

³In a previous study it could be shown that besides the insufficient robustness one main reason for this is that double reeds need a much higher air pressure than organ reed pipes [5, pp. 180–182].

⁴C. f. [2, pp. xlvii–xlix] for a discussion of KEMPELEN'S motives.

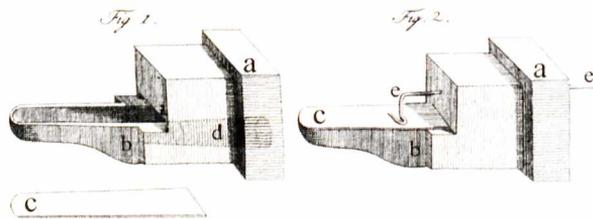


Figure 2 – Reed pipe made by KEMPELEN, according to [1, Tab. XVIII].

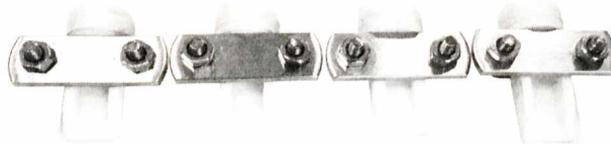


Figure 3 – Reed pipes nos. 7–10 made by TAKAYUKI ARAI (from left to right). The metal strips are for fixating the reed only, so their dimensions only influence the fundamental frequency of the pipe (the reeds are transparent and therefore difficult to see.).

and reed have to have a relevant length⁵ of 20–25 mm and a width of 10–15 mm. For the replica, dimensions of 22 x 12 mm have proven to be the best choice. In addition to this "primary" reed pipe, several experimental sound devices have meanwhile been constructed for this replica, of which some will be investigated below in this paper.

3.3 Organ reed pipes

Reed pipes as they are used in organ building, at least consist of three components (c.f. Fig. 2): The *block* (a), the *shallot* (b) and the *reed* itself (c). Shallot and reed together are fixed by a wedge in a drilling (d) that goes throughout the block. The *tuning spring* (e) allows to adjust the tuning of the reed to some extend. The reed itself is a thin piece of springy material (mostly brass), the shallot is a hollow tube segment with an open and flattened upper side. This opening is covered by the reed in its resting position.

When air flows, it raises the reed from the opening of the shallot so that the air can escape through the shallot. Because of the elastic recoil forces of the reed it begins to vibrate and the air column is stimulated. The degree of air pressure does not only influence the loudness of the sound but also its frequency.

3.4 ARAI reeds

TAKAYUKI ARAI developed an extraordinary design variant of reed pipes for didactic purposes (Fig. 3). The reed itself is made from transparent, but sturdy plastic film (polyester). The shallot is produced as a 3D print from nylon. Deviating from normal reed pipes, its front end is round bevelled, which probably has an effect comparable to the curvature of common reeds. These sound devices have a strong, surprisingly voice-like sound. Consequently, their suitability with KEMPELEN'S speaking machine should be tested.

⁵The reed's overall length has to be longer in order to be able to fasten it in the *block*.

4 Experimental design

4.1 Reeds

4.1.1 Ivory and brass reeds

For the measurements ten different reeds in three different designs were used: Firstly three specimens each made from ivory (reeds 1–3) and brass (reeds 4–6) with a vibrating length of 25 mm (Fig. 1b). For the brass reeds 4–6, it was possible to get them with three different, controlled thicknesses. They were made available by an organ builder and are of identical material composition and configuration as reeds of regular organ reed pipes.⁶ For the ivory reeds 1–3 some degree of randomness in terms of thickness had to be accepted for manufacturing reasons: Thinning out ivory evenly was hardly possible due to its composition and the tools available, so the thickness of the ivory reeds is not completely uniform over their entire surface.

Reeds 1–6 first were recorded without a leathering (Table 1, "a"). Subsequently, one ivory reed (3) and all brass reeds 4–6 were equipped with a layer of leather ("b"). For the brass reeds the leathering can be removed in case of poorer functionality. For the ivory reeds this was not possible because of the vulnerability of the material. So it was not reasonable to experimentally leather all of them. For the leathering split cow leather (thickness 0.4 mm⁷) was used.

4.1.2 ARAI reeds

Four different configurations of the reed pipe variant developed by TAKAYUKI ARAI (Fig. 3) were tested (reeds 7–10, tab. 2). Their configurations were adjusted to their fundamental frequency, not to comparable measurements with the other reeds. While the width of all four reeds is 10 mm, slightly different combinations of four constructional parameters were made:

- Length of the shallot/reed
- Thickness of the reed
- Radius of the shallot's curvature
- Inner dimension of the shallots

4.2 Air supply

The endeavour to do comparable measurements led to a construction-related problem: The reed pipe of the speaking machine is stimulated by air from a bellows. In normal use this is operated by the right elbow of the player. Operated with a constant weight, this so called wedge-bellows generates for physical reasons a decreasing air pressure towards the end (of a phrase/utterance). Thereby, fundamental frequency and intensity of the produced sound decrease as well. In normal use, based on auditory feedback, the player intuitively readjusts the pressure he exerts on the bellows to maintain an even sound level (or to create the intended prosodic contour), so this characteristic is not noticeable.

For our measurements this manually operating procedure was not reasonable because the exerted pressure would not have been measurable. So a constant weight was used knowing that only the first few seconds of a recording would be analyzable (Fig. 4). For our recordings we used a weight of 4,000 g (with a counterweight of 1,000 g).

⁶This is not the case for their dimensions, because according to KEMPELEN's instructions they are significantly broader (in proportion to their length) than conventional reeds.

⁷Sic! KEMPELEN mentions "glove leather" that, of course, must have had a certain thickness [1, p. 411].

Table 1 – Reeds made of ivory and brass (measurements in mm).

No.	Length	Thickness	Material	Leather
1	25	~0.25	Ivory	no
2	25	~0.4	Ivory	no
3a	25	~0.4	Ivory	no
3b	25	~0.8	Ivory	yes
4a	25	0.2	Brass	no
4b	25	0.6	Brass	yes
5a	25	0.22	Brass	no
5b	25	0.62	Brass	yes
6a	25	0.25	Brass	no
6b	25	0.65	Brass	yes

Table 2 – ARAI reeds (measurements in mm).

No.	Length	Thickness	Radius	Depth
7	10	0.1	15	3
8	12	0.1	15	4
9	12	0.1	20	3
10	14	0.2	20	4

4.3 Measurements

Recordings were made using a Sony PCM-D50 recorder and a Sennheiser ME2-II microphone (cardioid characteristic). The microphone was hung above the mouth funnel and directly over the top of the opened box of the speaking machine. For each reed and condition (with or without leather) three repetitions were made. The values reported in Tab. 3 are mean values of the three recorded repetitions. In case of fundamental frequency and intensity values these only refer to the stable phase of the excitation (cf. Fig. 4).

5 Results

5.1 Fundamental frequency and intensity

At the same length, the three different ivory reeds 1–3 show considerably different fundamental frequencies (Tab. 3.). While reeds 2 and 3a (that have nearly the same thickness) show similar fundamental frequencies, the thinner reed 1 shows a much lower fundamental frequency. For the reeds made of brass in not leathered condition (4a, 5a, 6a), fundamental frequency and thickness do not seem to be directly related. While 5a and 6a show a similar frequency, the fundamental frequency of 4a is 30 Hz higher. This is changing with the leathering. With this, all three brass reeds (4b, 5b, 6b) have the same, lower fundamental frequency. For the ivory reed 3, the fundamental frequency change is much stronger. The fundamental frequency frequencies of the ARAI-reeds 7–10 are comparable to those of reeds 4b, 5b and 6b. Interestingly, this reed construction is less sensitive to decreasing air pressure from the bellows than conventional reeds (fig. 4). For all reeds in question the intensity increases when equipped with a leathering.

5.2 Vibration characteristics and spectral analysis

Fig. 5 shows two periods for reeds 1–6 in all 10 conditions and exemplarily for reeds 9–10. Reed 2 shows a less complex sound characteristic compared to all other *reeds*. For reeds 3–6 the leathering influences not only fundamental frequency and intensity but also the vibration

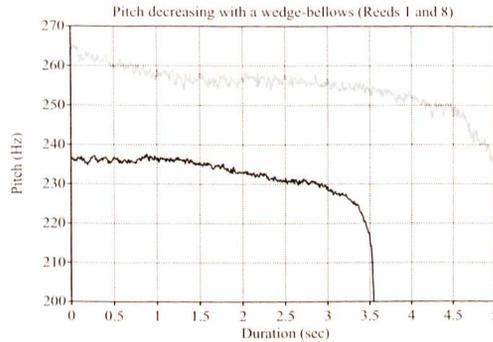


Figure 4 – Decreasing of fundamental frequency when blowing on with a wedge-bellows and a constant weight of 4,000 g for reed 1 (black line) and reed 8 (grey line). Figure shows measurements of single recordings.

characteristics substantially. The signal of reed 3b is much more intensive than 3a, also the spectral distribution of energy has changed (Fig. 6). But at the same time the strongest harmonic changes: While for reed 3a the 4th harmonic is the next strongest to the fundamental tone, it is the 6th harmonic for reed 3b. But the frequency of both harmonics is nearly the same. Something similar applies to the brass reeds 4–6. In not leathered condition the 5th (6th respectively for 6a) harmonics are the strongest ones next to the fundamental tone. For 4b it is the 7th instead, for 5b and 6b the 6th harmonic. Reeds 7–10 show a characteristic significantly different from the ivory reeds as well as from the ones from brass (figs. 5 and 6). But some similarity to reed 1 is also obvious.

5.3 Air supply

For reeds 1–3a (no leather), reeds 2 and 3a spend less air than the thinner reed 1. For reed 3, the leathering reduces the vibration time significantly. This is all the more true for reeds 4–6, while the values of these reeds within the two conditions are relatively comparable with each other.

Reeds 7–9 show a slightly decreasing vibration time, that is lower than for the not leathered reeds but higher than for the leathered ones. While wind pressure used was just about for reeds 1–9, reed 10 needed a much higher pressure, so that had to be helped manually. This is also reflected in the much shorter duration and the higher noise content.

The effect of the leathering on the stable phase of the fundamental frequency in vibrating time shows clear differences: For not leathered ivory reeds 1–3a the proportion of the stable phase is just under 40 %, for reed 3b it increases to 56 %. This increase can also be observed in the reeds 4 and 6, where it remains nearly constant for reed 5.

6 Discussion

From today's perspective, KEMPELEN'S *speaking machine* may seem like an amusing footnote in the history of phonetic sciences. But this was not the case in KEMPELEN'S lifetime, although not many understood the true meaning of this invention. Despite all its shortcomings, it was one of the very first serious approaches to speech synthesis based on empirically based findings. Many of its limitations are caused by contemporary technological limits.

It is not known which hypothesis caused KEMPELEN to choose ivory as material for his speaking machine's reed. Maybe some speculations about special characteristics of this material were decisive. This assumption could be supported by the fact that KEMPELEN'S "competitor"

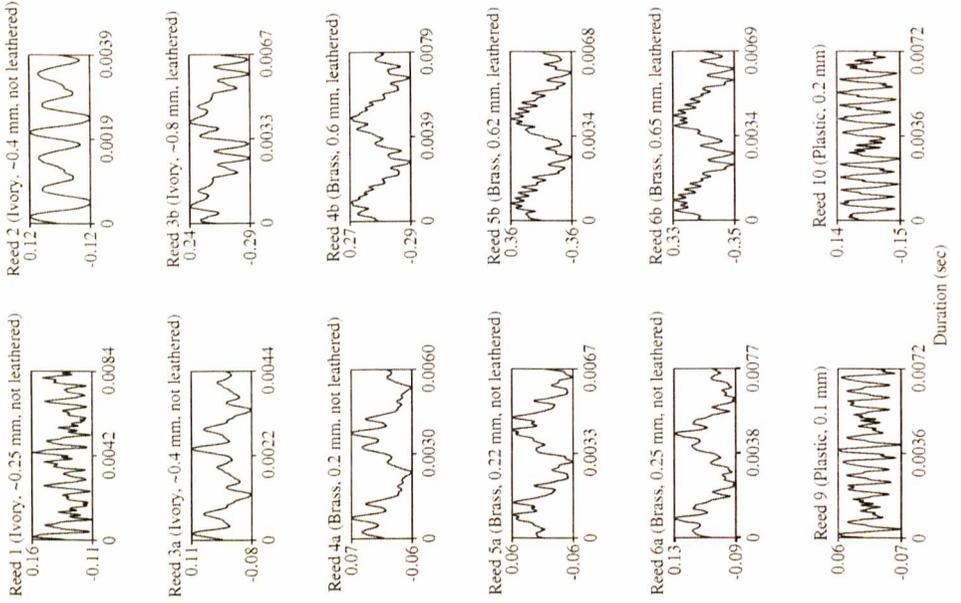


Figure 5 – Oscillograms of two periods each for reeds 1–6, 9 and 10.

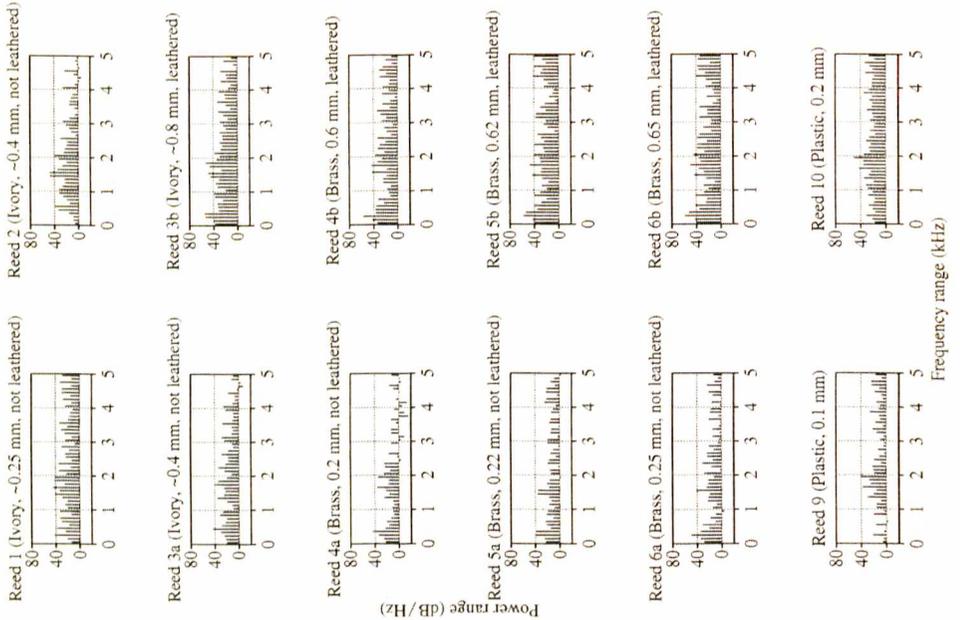


Figure 6 – Long Time Average Spectra (100 Hz) for for reeds 1–6, 9 and 10.

Table 3 – Synopsis of the reeds used in this study. (Tildes mark approximative measures. Values for fundamental frequency ("F0") and intensity are mean values referring to the stable phase of the sounds. Duration values (in seconds) are means over three recordings.)

No.	Material (mm)	Leather (Hz)	Thickness (dB)	F0 (Hz)	Intensity (dB)	Duration (ms)	Total Duration (ms)
1	Ivory	no	~0.25	237	71	1.5	3.9
2	Ivory	no	~0.4	511	70	2.4	6.7
3a	Ivory	no	~0.4	446	68	2.7	7.1
3b	Ivory	yes	~0.8	299	78	2.1	3.8
4a	Brass	no	0.2	329	65	3.4	6.7
4b	Brass	yes	0.6	285	77	2.3	3.1
5a	Brass	no	0.22	299	65	3.9	6.4
5b	Brass	yes	0.62	287	80	2.0	3.5
6a	Brass	no	0.25	298	70	4.0	5.8
6b	Brass	yes	0.65	287	79	2.3	3.1
7	plastic	no	0.1	280	80	3.3	5.7
8	plastic	no	0.1	261	81	3.6	5.3
9	plastic	no	0.1	275	83	3.0	5.2
10	plastic	no	0.2	276	92	0.5	1.7

CHRISTIAN GOTTLIEB KRATZENSTEIN (1723–1795) also used ivory for the reed pipes of his "vowel organ". As KEMPELEN, KRATZENSTEIN intended to soften the normally "metallic" sound of reed pipes and make it sound more human-like [7, p. 80].

Caused by the limited precision in thinning out the ivory, a comparability of ivory reeds with each other and with "standard" reeds made from brass is only possible to a limited extend. Nonetheless some basic observations can be made. As Fig. 5 shows, the vibration characteristics of ivory reeds do not necessarily have to be in a way other than that one of brass reeds. In the study presented here, reeds 3a, 4a and 6a show comparable behaviour. The conspicuously deviating behavior of reed 1 presumably is due to the considerable uneven thickness of this reed. Although all ivory reeds are more or less uneven, it is particularly noticeable with this one because the unevenness is more important with the small thickness.

Besides, Fig. 5 confirms the auditive impression, that reed 2 is the most succesful variant of the ivory reeds used in this study: Its oscillation behaviour is less complex than that of reeds 1 and 3 in particular, resulting in a softer sound.

But one way or the other, KEMPELEN'S design of the reed pipe ensured (unintentionally) that the choice of material of the reed is almost negligible. The (proportional massive) layer of leather reduces the vibration characteristics of the reed to a minimum.

At the same time, the influence of the reed's thickness on the height of the fundamental frequency is almost eliminated. So, as long as we know the aproximate length of KEMPELEN'S reed we can model the sound impact of his speaking machine very well.

Comparing the ARAI reed pipes, some fundamental differences regarding their vibration behaviour can be observed. The proportions of the harmonics is substantially different from reeds 1-6. In addition, the influence of the reed's thickness can be observed very well: Reed 10 was not operable with the air pressure that was sufficient for the remaining nine reeds.

However, this study also has shown that operating the bellows with a constant weight is not optimal. Although this procedure allows it to operate the speaking machine with a controlled air pressure it severely restricts the possibilities of investigation because of the bellows' characteristics. So a blower instead of the bellows would be more suitable at least for experiments.

7 Acknowledgements

Many thanks to TAKAYUKI ARAI for making his reed pipes available for experiments as well as to DENIS ARNOLD, ANJA GEUMANN and two anonymous reviewers for helpful comments.

References

- [1] W. Kempelen, *Wolfgang von Kempelen k. k. wirklichen Hofraths Mechanismus der menschlichen Sprache nebst Beschreibung seiner sprechenden Maschine*. Vienna: Degen, 1791
- [2] W. Kempelen, *Der Mechanismus der menschlichen Sprache / The Mechanism of Human Speech. Kommentierte Transliteration und Übersetzung ins Englische / Commented Transliteration and Translation into English*. Herausgegeben von / Edited by Fabian Brackhane, Richard Sproat & Jürgen Trouvain. Studentexte zur Sprachkommunikation, Vol. 87+88. Dresden: TUDpress, 2017.
- [3] F. Brackhane, "Die Sprechmaschine Wolfgang von Kempelens – Von den Originalen bis zu den Nachbauten," *Phonus (Reports in Phonetics, University of the Saarland)*, vol. 16, pp. 49–148, 2011.
- [4] B. Lamy, *La rhetorique ou l'art de parler*. Amsterdam: Marret, 1699.
- [5] F. Brackhane, "Kann was natürlicher, als Vox humana, klingen?" – *Ein Beitrag zur Geschichte der mechanischen Sprachsynthese*. *Phonus (Reports in Phonetics, University of the Saarland)*, vol. 18. Saarbrücken: IPUS, 2015.
- [6] T. Arai, *Vocal-tract models and their applications in education for intuitive understanding of speech production*. *Acoust. Sci. Tech.*, Vol.37, No.4, pp.148-156.
- [7] Chr. G. Kratzenstein, *Tentamen resolvendi problema. Petersburg 1781*. Übersetzt und kommentiert von Christian Korpiun. Dresden: TUDpress, 2016.