

ZUR HEUTIGEN BEDEUTUNG DER SPRECHMASCHINE VON WOLFGANG VON KEMPELEN

Jürgen Trouvain & Fabian Brackhane

*Institut für Phonetik, Universität des Saarlandes
trouvain@coli.uni-saarland.de*

Abstract: Die wissenschaftliche Beschäftigung mit der Kempelen'schen Sprechmaschine erfolgt zumeist aus wissenschaftshistorischen Motiven heraus. Der vorliegende Aufsatz widmet sich der Frage, welche Bedeutung der Sprechmaschine *heutzutage* zukommt. Neben möglichen Erklärungen, weswegen die Sprechmaschine auf Wissenschaftler wie Nicht-Wissenschaftler faszinierend wirkt, beschreiben wir den Einsatz von Nachbauten als Instrument zur Demonstration und auch zur Erforschung der Erzeugung von Sprachschall.

1 Einführung

Wolfgang von Kempelens "Mechanismus der menschlichen Sprache" von 1791 [1] und seine darin beschriebene Sprechmaschine haben ohne Zweifel eine große historische Bedeutung für die phonetischen Wissenschaften erlangt (vgl. z.B. [2,3,4]). Verschiedene Nachbauten der Sprechmaschine bezeugen eine gewisse Popularität und ihre einzigartige Stellung in der Forschung, die sich der Erzeugung von Sprachschall widmet. Die wissenschaftliche Beschäftigung mit der Sprechmaschine erfolgt zumeist aus wissenschaftshistorischen Motiven heraus. Der vorliegende Aufsatz widmet sich der Frage, welche Bedeutung der Sprechmaschine heutzutage zukommt.

2 Authentizität

Es gibt keine akustischen Zeugnisse vom Klang der Kempelen'schen Sprechmaschine. Wir können uns nur auf zwei Arten von Quellen verlassen, die beide ihre Nachteile haben. Zum einen gibt es Schilderungen von Zeitgenossen Wolfgang von Kempelens, die einer der Demonstrationen seiner Sprechmaschine beiwohnten. Zum anderen gibt es moderne Nachbauten der Sprechmaschine, welche – sofern sie funktionstüchtig sind – zum Ertönen gebracht werden können.

2.1 Zeitgenössische Beschreibungen

Wie haben nun Zeitgenossen den Klang beschrieben? Zwei anonyme Autoren von Zeitschriftenartikeln ähneln sich in ihren Beschreibungen vom ersten Eindruck der Maschine:

"Hierauf streckte er die rechte Hand durch das große Loch ins Kästgen, drückte mit dem Arme und Ellenbogen den Blasebalg nieder und sprach völlig mit der Stimme eines drey bis vier jährigen Kindes, sehr deutlich und vollkommen gut artikuliret in dem Kästgen aus: 'Oh Maman, Maman, on m'a fait mal!'" [5: 180]

"Das erste was wir hörten war: 'Mama, Papa, à ma chère Mama on m'a fait du mal.' Und nun konnte jeder in der Gesellschaft ein Wort fordern. Alle sprach die Maschine mit der größten Deutlichkeit aus. Auch die doppelten Vocalen und Konsonanten pronunziert sie sehr rein und richtig. Der Ton ist wie bei einem Kind von drei Jahren. Zuweilen kam das verlangte Wort nicht

gleich zum erstenmal richtig heraus, der Künstler mußte verschiedene Versuche machen. Er entschuldigte sich damit, daß einer, der die Violinen macht, sie darum nicht auch fertig spielen könne." [6: 483 f.]

Es lassen sich noch etliche weitere Zeugnisse finden, in denen sich ähnlich begeistert geäußert wird. Darunter sind aber auch skeptische Stimmen zu finden, welche die Sprechmaschine als Blendwerk ansehen, genauso wie Kempelens Schachtürken auch. Für eine ausführlichere Behandlung der zeitgenössischen Rezeption (u.a. von Goethe) siehe [7].

2.2 Klang von heute

Wir können nicht sicher sein wie weiland die "Original"-Sprechmaschine klang. War die von ihr produzierte Sprache tatsächlich so natürlich? Konnte man den erzeugten Schall wirklich als den einer Kinderstimme vernehmen? Die Berichte der Zeitgenossen stehen mit großer Sicherheit unter dem Einfluss mehrerer Faktoren. Zum einen zeigte Kempelen die Maschine bei seinen Demonstrationen stets unmittelbar nach seinem Schachtürken, wodurch das Publikum schon in eine gewisse Begeisterung versetzt war. Des Weiteren weist Köster [3] auf die wahrscheinlich eminente Rolle der Autosuggestion bei den beschriebenen Vorführungen hin: durch vorangehenden Wunsch der Äußerungen wusste das Auditorium bereits was es hören wollte und sollte.

Wie authentisch klingt heutzutage die Sprechmaschine? In einer anderen Studie [8] zeigen die Autoren, dass Zuhörer in einem Test mit offenem Antwortformat, in dem Alltagsgeräusche und auch die Wörter "Mama" bzw. "Papa", die mit einem Nachbau erzeugt wurden, dargeboten wurden, diese überwiegend als einer kindlichen Stimme zugehörig beschrieben. Diese Wirkung war mitunter so stark, dass einzelne Versuchspersonen auch nach dem Test nicht glauben wollten, dass es sich *nicht* um eine menschliche Kinderstimme handelte.

Auch wenn diese Wirkung des Authentischen bei anderen Wörtern und vor allem bei längeren Äußerungen nicht durchzuhalten ist, so bliebe die Beobachtung festzuhalten, dass die Synthesequalität der Sprechmaschine fesselnd wirkt.

3 Einfachheit

Ein genaueres Hinsehen auf und vor allem in die Sprechmaschine verrät, dass es sich um eine Konstruktion handelt, zu der das Attribut "(relativ) einfach, aber genial" passt.

3.1 Die Bestandteile

In Tabelle 1 sehen wir, aus welchen Komponenten die Sprechmaschine besteht. Als Material finden Holz, Leder, Gummi und Metall sowie Elfenbein und Leim Verwendung, um die verschiedenen Bestandteile zu konstruieren (vgl. auch Abbildung 1). Jede Komponente hat ein Analogon zur Anatomie bzw. Sprechphysiologie beim Menschen, wobei eine anatomische Korrektheit nicht immer erzielt wird, so beispielsweise die Nasenlöcher, die sich unmittelbar der Glottis anschließen. Gleiches gilt für die Frikativ-Generatoren, welche direkt der "Luftröhre" angegliedert sind.

Analogon	Bestandteil	Material
Lunge	Balg + Galgen	Holz, Leder, Blei (Gewicht)
Luftröhre/Brustkorb	Windlade	Holz, Schrauben, Kupfer (Rohr)
Kehlkopf	Nuss (Zylinder)	Holz
Stimmlippe/Glottis	Zungenblatt + Zungenkehle	Holz, Leder, Leim, Elfenbein
Nasenraum	Rundscheibe + Nasenröhren	Holz + Messing
Vokaltrakt	Schalltrichter	Gummi
Friktion	"Blockflöten"	Holz, Messing (Hebel)

Tabelle 1 – Die Bestandteile der Sprechmaschine. Nuss, Zungenblatt und Zungenkehle werden im Orgelbau zusammen Zungenpfeife genannt.

3.2 Das Innenleben als die eigentliche Sprechmaschine

Obwohl die Sprechmaschine üblicherweise mit ihrem Resonanzkasten gezeigt wird, in den hinein der Balg geführt wird, so hat dieser Hohlkörper nur eine zu vernachlässigende akustische Wirkung. Die eigentliche Erzeugung des Sprachschalls findet im Inneren dieses Kastens statt. In Abbildung 1 kann man die zentralen Bestandteile, wie sie in Tabelle 1 aufgeführt sind, in Kempelens Sicht und von einer Replik vergleichen.

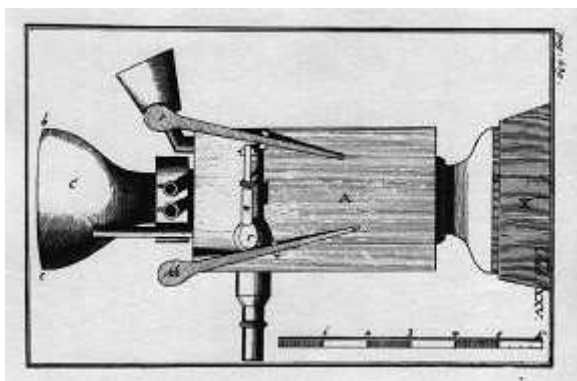


Abbildung 1 – Links: Innenleben der Sprechmaschine als Originalzeichnung aus Kempelen [1]. Rechts: Das Innenleben aus einem der in Saarbrücken entstandenen Nachbauten.

3.3 Die Sprechmaschine als "Musikinstrument"

Genau betrachtet kann man die Sprechmaschine auch als Musikinstrument bezeichnen. Dies nicht zuletzt deswegen, weil als Prinzip der Anregung der Luft eine im Orgelbau gebräuchliche Zungenpfeife Verwendung findet. Ebenso bedarf es wie bei jedem Musikinstrument einer gewissen Übung, während der man den motorischen Ablauf und dessen Koordinierung trainiert.

Daher lassen sich leicht Vergleiche zu artikulatorischer Sprachsynthese (z.B. [9]) und der ihr zu Grunde liegenden Analyse von Sprechgesten anstellen. Der in Abbildung 2 dargestellte Artikulationsplan kann als eine Art Partitur angesehen werden, bei der – im Gegensatz zu einem natürlichen Sprecher – der Sprechmaschinenspieler Arm- und Handgesten als Sprechgesten ausführt. Wir beschränken uns an dieser Stelle auf drei Parameter:

1. zur Steuerung des "sub-glottalen" Luftdrucks (durch rechten Ellenbogen "E"),
2. zum Anheben des "Velums" (Zeige- u. Ringfinger der rechten Hand "R" decken "Nasnlöcher" ab),
3. zum Bilden eines Verschlusses und gradueller Simulation der "Zungenkonstriktion" (durch den linken Handteller "L", der den Schalltrichter abdeckt bzw. graduell freigibt).

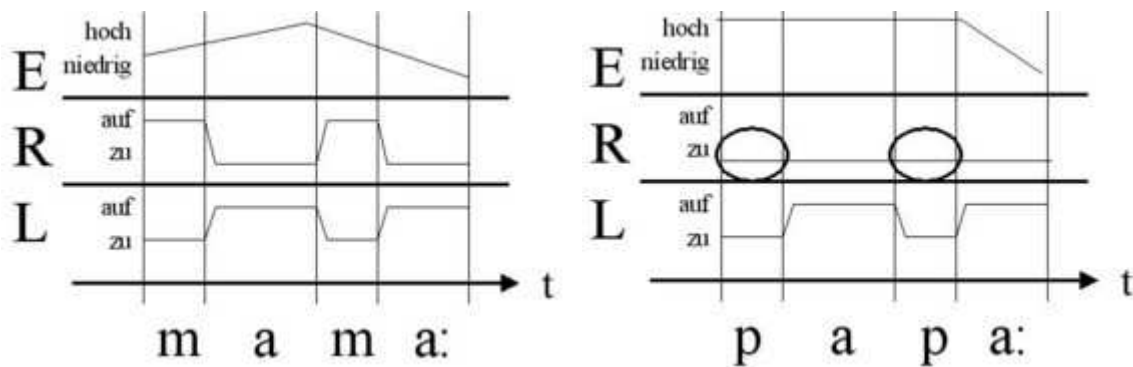


Abbildung 2 – Partitur von "Mama" und "Papa". E = Druck durch rechten Ellenbogen; R = Rechte Hand: Zeige- u. Ringfinger decken "Nasnlöcher" ab; L = Linker Handteller simuliert Grad der "Zungenkonstriktion" durch Abdecken bzw. durch Entfernen des Schalltrichters.

Der hauptsächliche Unterschied zwischen "Mama" und "Papa" besteht im Verschließen bzw. Öffnen der "Nasnlöcher" während der konsonantischen Artikulation im Wort. In Analogie zu den Sprechbewegungen für [m] und [p] beim Menschen, bei dem für [p] aktiv das Velum angehoben werden muss, wird bei der Maschine der nasale Ausgang durch zwei Finger der rechten Hand versperrt (siehe eingekreiste Stellen bei [p] in "Papa" in Abbildung 2). Durch das Stoppen der Luft an der oralen wie auch an der nasalen Stelle wird der Anblasedruck, der das so genannte Zungenblatt (aus Elfenbein) als "Stimmlipe" zum Schwingen bringt, unterdrückt, so dass kein Stimmton entstehen kann.

3.4 Nachbaubarkeit

Das Erkennen der prinzipiellen Einfachheit ihrer Bestandteile lässt den ein oder anderen Begeisterten wohl auch zu dem Gedanken gelangen "Ich könnte das auch bauen bzw. nachbauen".

Es besteht wohl für viele ein Unterschied in der *mentalen* Umsetzbarkeit zwischen Software- und Hardware-Lösungen, denn technologische Produkte lassen üblicherweise nicht an eine eigene Umsetzbarkeit denken. "Low Tech" (im besten Sinne) scheint durch ihre Nachbaubarkeit zusätzlich zu fesseln. Dabei spielt sicherlich auch der Aspekt des Haptischen eine Rolle.

Allerdings gibt es auch schwierige Komponenten wie zum Beispiel den Balg – dieser könnte jedoch auch gekauft werden. Die modifizierten Blockflöten sind eher optional, da sie lediglich für Zischlaute gebraucht werden. Tatsächlich problematisch ist die Herstellung der Zungenpfeife ("Kehlkopf") bestehend aus der Zungenkehle, Zungenblatt ("Stimmlipe") und der sie umgebenden Nuss aus Holz. Abschließend kann festgehalten werden, dass sich bei der Herstellung mehrerer Nachbauten durch den zweiten Autor der Sinnspruch "der Teufel steckt im Detail" bewahrheitete.

4 Originale und Repliken

4.1 Original(e)

Auch wenn immer von *der* Sprechmaschine die Rede ist, so sollte man sich im Klaren darüber sein, dass Kempelen selbst mit mehreren Versionen seiner Sprechmaschine parallel zueinander experimentierte. Da es demnach *die* Sprechmaschine gar nicht gegeben hat, kann es sich bei dem historischen Exemplar im Deutschen Museum (vgl. Abbildung 3) nicht um *das Original* handeln [10], als das es aber häufig bezeichnet wird.

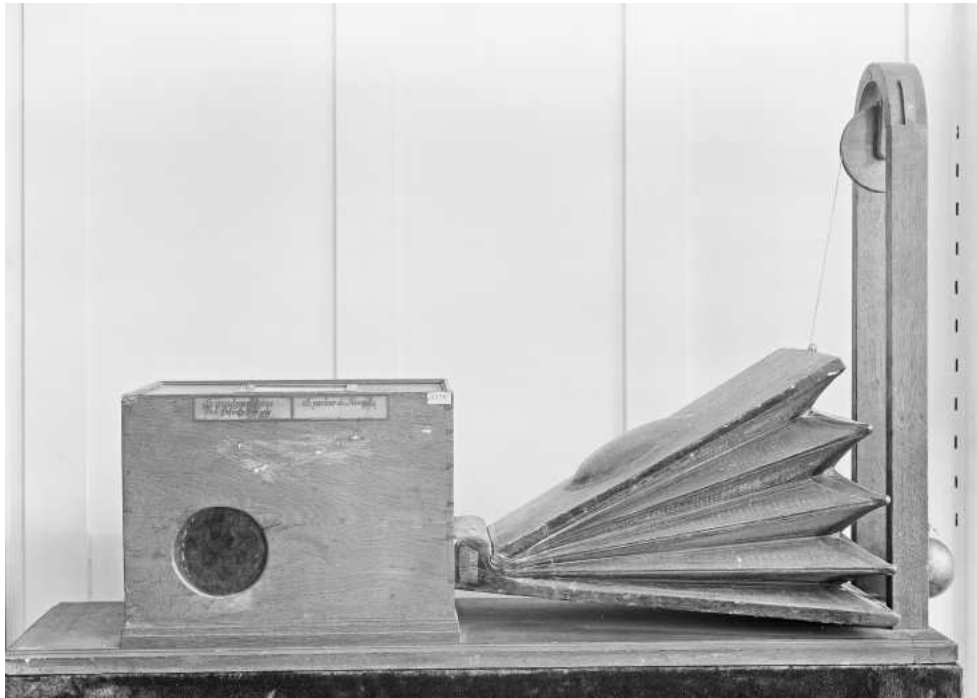


Abbildung 3 – Das Exemplar im Deutschen Museum in München.

4.2 Nachbauten

Im Laufe der zweiten Hälfte des 20. Jahrhunderts sowie zu Anfang des jetzigen Jahrhunderts hat es mehrere Versuche gegeben, die Sprechmaschine nach der Vorlage im *Mechanismus* nachzubauen. Dabei haben nicht alle Nachbauten Wert darauf gelegt sich in allen Konstruktionsdetails am historischen Vorbild zu orientieren.

So verzichteten beispielsweise die Nachbauten aus Wien (Universität für angewandte Kunst) und England (Universität York, Lehrstuhl Musiktechnologie) auf Resonanzkasten und Galgen, die sonst bei der Zurschaustellung üblich sind. Bei diesen Repliken wird Wert auf die Spielbarkeit als Instrument z.B. bei Aufführungen gelegt.

Im Gegensatz dazu handelt es sich bei den beiden Repliken aus Budapest (Akademie der Wissenschaften) um Arbeiten, die versuchten sowohl von der baulichen als auch von der phonetisch-akustischen Perspektive die größtmögliche Originaltreue zu erreichen. Diese Ziele gelten auch für die deutschen Nachbauten, die 2007-2009 in Saarbrücken (Universität des Saarlandes, Lehrstuhl Phonetik) entstanden sind. Zwei dieser Exemplare befinden sich nun in Dresden (TU Dresden, Lehrstuhl Kommunikationsakustik) und in Paderborn (Heinz Nixdorf MuseumsForum).

Die obige Auflistung dient lediglich der Veranschaulichung der Variabilität der Nachbauten. Es handelt sich dabei nicht um eine vollständige Auflistung, man kann

davon ausgehen, dass derzeit ein gutes Dutzend (teils) funktionsfähiger Repliken in Europa existiert (siehe [7]).

5 Die Sprechmaschine als Demonstrationsinstrument

Als Demonstrationsinstrument ist die Sprechmaschine zu allen Zeiten als außergewöhnlich und überzeugend zu bezeichnen. Dies trifft auf die europäischen Höfe zu Kempelens Zeiten genauso zu wie auf heutigen Unterricht vor Schülern und Studenten. Ein Instrument, das nur aus Holz, Metall, Leder, Gummi und einem Plättchen Elfenbein besteht, fasziniert trotz – oder auch wegen – elektronischer Sprachgenerierungsmethoden, die wir seit mehreren Jahrzehnten kennen.

Auch wenn in den vorhergehenden Abschnitten versucht worden ist das Faszinosum Sprechmaschine durch ihre Authentizität, Einfachheit sowie ihre Nachbaubarkeit zu erklären, so ist es doch wichtig zu erfassen, wer sich auf welche Weise mit welcher wissenschaftlichen Frage angesprochen fühlt.

Nach etlichen Vorführungen der Saarbrücker Nachbauten vor ganz unterschiedlichem Publikum, auch wenn diese in aller Regel nur "Mama" und "Papa" enthielten, können wir berichten, dass eigentlich niemand unbeeindruckt geblieben ist. Dies trifft auf Studenten genauso zu wie auf Professoren, auf Kinder ebenso wie auf Technik- und Lebenserfahrenere, auf solche mit technischer Ausbildung wie Ingenieure und solche mit technischer Ausbildung wie Logopäden. Auch Vertreter der phonetischen Wissenschaften scheinen sich der Faszination der Sprechmaschine nicht vollständig entziehen zu können.

Man kann durchaus behaupten, dass Repliken der Sprechmaschine sehr gut als Objekt zur Illustration der Erzeugung von Sprachschall dienen können – und dies multi-modal. Denn ein weiterer wichtiger pädagogischer Aspekt dieser informellen Beobachtungen ist die Tatsache, dass man die Maschine anfassen und damit nicht nur akustisch und visuell, sondern auch haptisch erfassen kann. Das Erfassen im wortwörtlichen Sinne in Verbindung mit anderen Modalitäten bietet eine hervorragende und eher ungewöhnliche Möglichkeit, den sonst in vielen Punkten unsichtbaren Ablauf der Erzeugung und Übertragung von Sprachschall zu demonstrieren. Dazu gehört auch, dass der Bediener des Instruments sich stetig gezwungen fühlt beim Spielen still mitzuartikulieren. Offenbar erleichtert es dem Spieler die Tätigkeit der Sprech-Erzeugung *durch die Hände*, wenn die kognitive Steuerung der Sprech-Erzeugung *durch den Sprechapparat* unhörbar simuliert wird. Möglicherweise kann der Spieler dieses "innere Sprechen" nur unter bewusster Anstrengung unterdrücken. Der Betrachter gewinnt dann den Eindruck, dass der Spieler mit der Stimme der Maschine artikuliert.

Das "Befassen" (mit) der Sprechmaschine sensibilisiert zu guter Letzt für die Frage "Wie kann es sein, dass dieses Holz-Leder-Metall-Konstrukt wie ein Mensch sprechen kann?" Diese Frage führt unweigerlich zum Kern des Strebens der phonetischen Wissenschaften: "Wie kommt es eigentlich dazu, dass der Mensch sprechen kann?" Auch Kempelen selbst hat seinerzeit im 18. Jahrhundert mit dieser Fragestellung begonnen. Er war sich aber über die Grenzen seiner praktischen Umsetzung seiner Forschungsarbeit voll bewusst. Er schließt sein Hauptwerk [1] unter anderem mit dem Wunsch, dass seine Leser "dieser neuen, noch in ihrer Kindheit befindlichen Erfindung einige Aufmerksamkeit schenken, und sie durch ihr Nachdenken und Bemühen weiter vorfrücken werden."

6 Die Sprechmaschine als Forschungsinstrument

Unseres Erachtens geht die Bedeutung der Sprechmaschine über diejenige eines einzigartigen Instruments zur Demonstration und Sensibilisierung für Sprachschallerzeugung jedoch hinaus. Für die überzeugende Stimmqualität verantwortlich ist der richtige Anblasedruck auf ein geeignetes Material, das zum Schwingen gebracht wird. Dabei handelt es sich hier eher um ein Korrelat eines passiven Kehlkopfes mit einseitig gelähmter Stimmlippe, das weder eine Steuerung der Adduktion und Abduktion noch eine Änderung der Form der Stimmlippen erlaubt. In dieser Hinsicht ähneln solche Experimente zur Phonation eher solchen mit post mortem entfernten Kehlköpfen (z.B. [11]) als der Modellierung des Phonationsprozesses in vivo.

6.1 Rolle des sub-glottalen Resonanzraumes

Durch Tests mit verschiedenen großen Windladen, welche die Verbindung der Zungenpfeife als Stimmgeber und dem Blasebalg als "Lunge" darstellt, hat sich gezeigt, dass die Größe dieses "sub-glottalen" Resonanzraumes einen entscheidenden Einfluss auf den Grad der Authentizität der künstlichen Kinderstimme hat [8]. Je nach Größe und Holzart bei der Anfertigung der Windlade erhält man bessere oder schlechtere Ergebnisse bei Authentizitätstests.

Es stellt sich hiermit die Frage, welche Rolle der sub-glottale Raum und der darin erzeugte Luftdruck für die Erzeugung von Sprachschall sowie für die Individualität einer Stimme spielt. Hinweise auf sub-glottale Resonanzeigenschaften kommen in phonetischen Lehrbüchern häufig (noch) nicht vor. Allerdings ist in dieser Angelegenheit noch mehr Grundlagenforschung notwendig (wie z.B. [12]).

Diese Fragestellung kann auch auf natürlich wirkende Stimmen bei artikulatorischen Synthesen ausgeweitet werden. Ein Vorteil bei einem Nachbau ist hierbei, dass sich verschiedene "sub-glottale" Hohlkörper innerhalb weniger Sekunden austauschen lassen.

6.2 Nachgiebigkeit bei stimmlosen Plosiven

Eine der Voraussetzungen für die Phonation ist ein ausreichend großer Abfall des transglottalen Luftdrucks. Durch das Anstauen der Luft wird bei Plosiven (und anderen Obstruenten) der orale Luftdruck erhöht, was zu einer Reduktion des transglottalen Luftdruckabfalls führt – und damit zu Entstimmlichung. Bei üblichen Geschwindigkeiten des Luftstroms wäre das Nachlassen der Phonation nach spätestens 15 ms erreicht [13], die Verschlussphasen (voll) stimmhafter Plosive sind aber üblicherweise länger als dieser Schwellwert. Man nimmt daher eine Ausweitung des Vokaltrakts an, um mehr Platz für die sich weiter anstauende Luft zu haben und den Zeitpunkt des Druckausgleichs und damit der Entstimmlichung hinauszuzögern.

Eine aktiv gesteuerte Vokaltrakterweiterung kann beispielsweise durch Anheben des Larynx oder Absenken des Zungenkörpers geschehen. Eine passive Erweiterung des Trakts ergibt sich durch Nachgiebigkeit des Gewebes ("compliance") [13]. Genau dieser Effekt wird bei der Sprechmaschine durch einen Plosivbalg erzielt. Dieser Balg befindet sich direkt unterhalb des "Nasenraums" und beide Hohlkörper sind mit einem Röhrrchen miteinander verbunden (Abbildung 4).

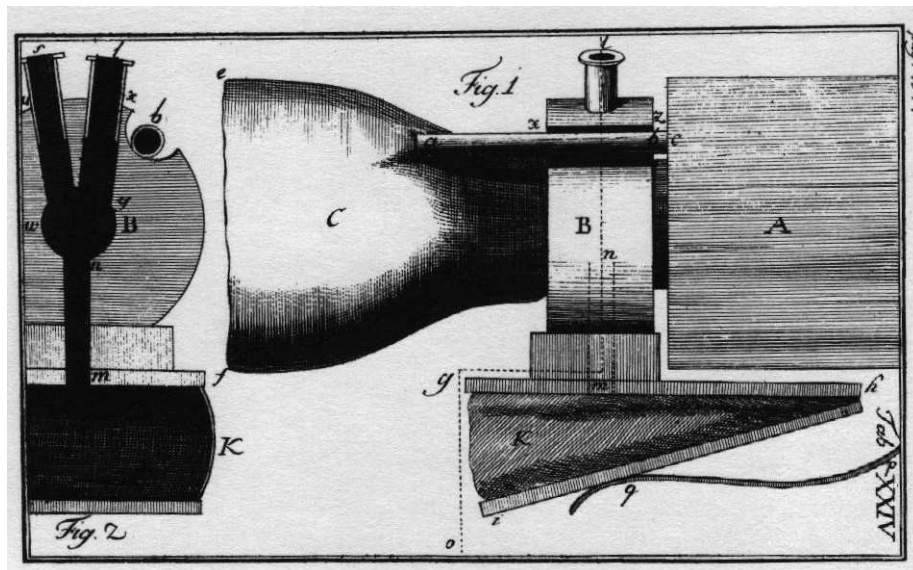


Abbildung 4 – Plosivbalg "K" im Querschnitt von vorne (linke Seite) und in der Seitenansicht des Innenlebens der Sprechmaschine (vom Originalstich aus [1]). Der "Nasenraum" ("B") ist durch ein Röhren ("n") mit dem darunter befindlichen Plosivbalg verbunden.

Dieser zweite, viel kleinere Balg hat also den Effekt die Plosive länger stimmhaft zu halten, was sich auch an verschiedenen Nachbauten nachvollziehen lässt. Repliken *ohne* diesen Balg eignen sich viel besser zur Erzeugung stimmloser Plosive mit anschließender Aspiration, bevor die Stimmhaftigkeit des Vokals beginnt. Bei Repliken *mit* Balg ist dieser Effekt nicht möglich, dafür klingt "Papa" dann eher wie "Baba". Ein dynamisches Ab- bzw. Hinzuschalten des zweiten Balgs wäre eine sehr gute Lösung.

Es ist hier allerdings anzumerken, dass Kempelen mit der Installation des kleinen Zusatzbalges den exakt gegenteiligen Effekt hervorzurufen versuchte: "Um die Explosion bey den stummen Mitlautern zu verstärken, habe ich noch einen anderen, nicht minder wichtigen Zusatz gemacht. Ich habe nämlich [...] einen kleinen Blasebalg angebracht [...]" [1: 437]

Um überhaupt einen Unterschied zwischen [b] und [p] bei der Sprechmaschine zu erwirken, wird bei [p] unmittelbar vor und während des Verschlusses der beiden Öffnungen ein größerer Druck aufgebaut als bei [b] (auch in Abbildung 2 erkennbar). Allerdings darf nach der Verschlusslösung der "sub-glottale" Anblasedruck nicht zu groß sein, da es sonst nicht zur Schwingung der "Stimmlippe" kommt. (Diesen Effekt haben wir möglicherweise auch beim Weinen bzw. Schreien beim Menschen, dann wahrscheinlich in Verbindung mit einer Anspannung der Stimmlippen selbst.)

6.3 Die Preisfragen der Petersburger Akademie

Weitere Überlegungen betreffen die Verbindung zwischen Erzeugung menschenähnlicher Stimmen und Orgelbau, so wie sie auch in einer Preisfrage der Petersburger Akademie von 1780 unter Vorsitz von Leonhard Euler angeregt wurde:

1. Qualis sit natura et character sonorum litterarum vocalium a, e, i, o, u tam insigniter inter se diversorum.
2. Annon construere queant instrumenta ordini tuborum organicorum, sub termino vocis humanae noto, similia, quae litterarum vocalium a, e, i, o, u sonos expriment. (vgl. [14])

1. Wie ist die Natur und der Charakter der Vokal-Buchstaben a, e, i, o, u, die sich so deutlich voneinander unterscheiden.
2. Ist es nicht möglich, Instrumente nach Art derjenigen Orgelpfeifen zu bauen, die als *Vox humana* bekannt sind und die den Klang der Vokal-Buchstaben a, e, i, o, u hervorbringen. (eigene Übersetzung)

Den Preis für die Beantwortung dieser Fragen erhielt Johann Gottlieb Kratzenstein [14]. Er lieferte wohl Pfeifen, welche die genannten Vokale erzeugen – diese Pfeifen zeigen in ihrer Form allerdings keine Analogie zu der Vokalerzeugung in einem menschlichen Vokaltrakt und scheinen auch sonst nicht in ihrer Form akustisch motiviert. Zudem erzeugen sie lediglich *statische* oder *isolierte* Vokale. Auf einer Art "Orgel" sollte für jeden Einzelvokal eine eigene Taste eine separate Pfeife bedienen. Kempelen hingegen hat einen wichtigen nächsten Schritt unternommen. Mit seiner Sprechmaschine hat er die zentrale Bedeutung der Koartikulation erkannt und hat diese auch in seine Maschine integriert:

"Izt fieng ich an einzusehen, daß sich die einzelnen Buchstaben zwar erfinden, aber auf die Art, wie ich es angriff, nimmermehr in Sylben zusammenbinden ließen, und daß ich schlechterdings der Natur folgen müßte, die nur eine Stimmritze, und nur einen Mund hat, zudem alle Laute herausgehen, und eben nur darum sich miteinander verbinden." [1: 407]

Nach wie vor ungeklärt ist die Beantwortung der zweiten Frage nach der *Vox humana*. Damit ist entgegen anderer Übersetzungen (z.B. [15]) nicht die menschliche Stimme gemeint, sondern das *Orgel*-Register, welches im Orgelbau bereits seit Jahrhunderten existiert (vgl. auch [16]). Dort wird dieses Register in Kombination mit einem so genannten "Tremulanten" verwendet, um einen Klang ähnlich der Vibrato-Singstimme des Menschen zu erzeugen. Bei einem Tremulanten wird der zur Pfeife (als Instrument der Anregung) gelenkte Luftstrom periodisch unterbrochen. Dabei handelt es sich um eine der Pfeife *vorgeschaltete* Konstruktion. Eventuell böten sich hiermit neue Möglichkeiten der Modellierung der maschinellen Singstimme, auch wenn sowohl professionelle Sänger als auch Singsynthesen glottale Parameter verändern (z.B. [17,18]), um Vibrato zu erzeugen.

Kempelen griff im Laufe seiner Forschungsarbeit den Gedanken, das Orgelregister *Vox humana* als Grundlage einer Sprachsynthese zu nehmen, ebenfalls auf. Nicht zuletzt deswegen verwendete er auch schließlich eine dem Orgelbau entlehnte Zungenpfeife (mit nur *einem* schwingenden Element, ähnlich der Funktionsweise eines Klarinettenmundstücks) als Repräsentation der Stimmlippen. Dabei verwarf er seine zunächst auf Basis der menschlichen Phonation entwickelte Konstruktion auf Grundlage eines Oboenmundstücks (mit zwei gegeneinander schwingenden Elementen, sehr ähnlich den menschlichen Stimmlippen). Trotz seiner Kenntnis der Kratzenstein'schen Arbeit griff er dessen (bei grundsätzlich falscher Auffassung des Stimmgebungsprozesses) in einigen Aspekten bessere Konstruktion nicht auf. Stattdessen suchte er sich mit teils höchst ungewöhnlichen Modifizierungen der Orgelpfeifen zu behelfen, um einen der menschlichen Stimme ähnlichen Klang zu erreichen. Eine Kombination der Kratzenstein'schen Zungenpfeife mit der Kempelen'schen Sprechmaschine würde ein hochinteressantes Forschungsobjekt ergeben.

7 Sprachsynthese im 18. vs. 21. Jahrhundert

Das 18. Jahrhundert könnte man auch das Jahrhundert der Automaten bezeichnen, was natürlich auch Sprech-Automaten miteinbezieht [3]. Allerdings geht es bei diesen Maschinen um die *Wiedergabe* von Sprachschall, bei der Erfindung Kempelens aber um die *Erzeugung* von Sprachschall. Kempelens Sprechmaschine könnte man als die erste funktionierende Hardware-Sprachsynthese bezeichnen. Es ist erstaunlich und bewundernswert, dass die historische Sprechmaschine den Vergleich zu Hardware-synthesen des 21. Jahrhunderts (z.B. [19]) nicht zu scheuen braucht, weist sie doch einen bedeutend besseren Stimmklang auf, der zu Kempelens Zeiten genauso wie heute ausreicht, um Zweisilbler mit einer Kinderstimme für unvoreingenommene Zuhörer täuschend nachzubilden wie bereits in Sektion 2 besprochen.

Eigentlich war die Sprechmaschine als Hilfsmittel für Gehörlose gedacht – eine Motivation, die einer anderen Erfindung Kempelens ähnelt, als er für einen blinden Menschen eine Schreib-Maschine entwickelt und erbaut hat (an dieser Stelle sei auf seine Biografie verwiesen [20]). Hier zeigt sich der starke Zusammenhang von Sprechkompetenz und sozialer Akzeptanz: es zählt nur, wer sprechen kann.

Aber nicht nur der Wunsch, dass es Sprachsynthese an sich geben soll, wird im 18. Jahrhundert laut gedacht, sondern auch darüber in welcher Weise die Ersatzstimmen klingen sollen. Leonhard Euler schreibt 1761 in seinen populärwissenschaftlichen "Briefen an eine Prinzessin" [21]:

"Ce seroit sans doute une plus importantes découvertes , que des construire une machine qui fut proper à exprimer tous les sons de nos paroles avec toutes les articulations. [...] Les prédicateurs & les orateurs , dont la voix n'est pas assés forte ou agreeable , pourroient alors jouer leur sermons & discours fut une telle machine , tout de meme que les organists jouent de musique. La chose ne me paroît pas impossible."

"Es wäre ohne Zweifel eine der wichtigsten Entdeckungen eine Maschine zu konstruieren, die sauber alle Töne unserer Sprache mit all ihren Artikulationen ausdrücken kann. [...] Die Prediger und Redner, deren Stimmen nicht stark oder angenehm genug wäre, könnten alsdann ihre Predigten und Reden auf einer solchen Maschine spielen, so wie jetzt die Organisten Musikstücke spielen." (eigene Übersetzung)

Wir müssen uns am Anfang des 21. Jahrhunderts und ca. 230 Jahre nach Kempelens Erfindung der Sprechmaschine die Frage gefallen lassen, ob denn mit den heutigen Sprachsynthesen Predigten und Reden im Sinne Eulers erzeugbar sind. Bedenkt man, dass sich die Sprachsynthese-Forschung erst innerhalb der letzten 10 Jahre Themen wie Emotionen, Affekt und andere Formen des sprecherischen Ausdrucks verstärkt annimmt (vgl. [22]), so kann man natürlich etliche Fortschritte verzeichnen. Für eine überzeugende Beantwortung der obigen Frage mit "ja" bleibt indes noch vieles zu tun. Ein Schritt in die richtige Richtung ist das Erkennen, dass "unsere Sprache mit all ihren Artikulationen ausdrücken" *weitaus* mehr bedeutet als textuelle Information verständlich zu übertragen.

8 Fazit

Auch wenn sich die Bedeutung der Kempelenschen Sprechmaschine vornehmlich in historischen Dimensionen bewegt, so bleibt dennoch festzuhalten, dass diese Maschine

für die Forscher von heute eines deutlich repräsentiert: eine über die Disziplinen hinweg ausstrahlende Faszination des wissenschaftlichen Gegenstandes der menschlichen Stimme sowie Anregungsquelle für weitere Forschung und Entwicklung von Sprach- und Singsynthesen.

Danksagung

Die Autoren möchten sich herzlich bei Herrn Orgelbaumeister Stephan Mayer (Heusweiler/Saar) und sein Team für die Unterstützung bei der Anfertigung der Nachbauten bedanken.

Literatur

- [1] W. von Kempelen. 1791. Über den Mechanismus der menschlichen Sprache. nebst Beschreibung einer sprechenden Maschine. Wien: J.V. Degen. Faksimile-Nachdruck von 1970, Stuttgart: Frommann-Holzboog.
- [2] H. Dudley & T.H. Tarnoczy 1950. The speaking machine of Wolfgang von Kempelen. *Journal of the Acoustical Society of America* 22, 151-166.
- [3] J.-P. Köster 1972. Historische Entwicklung von Syntheseapparaten zur Erzeugung statischer und vokalartiger Signale nebst Untersuchungen zur Synthese deutscher Vokale. Hamburg: Buske.
- [4] B. Pompino-Marschall 2004. Von Kempelen's contribution to the theory of acoustic articulation. *Grazer Linguistische Studien* 62, 137-147.
- [5] Anonymus. 1784a. Ueber Herrn von Kempelens Schach=Spieler und Sprach=Maschine. Zweeter Brief. *Teutscher Merkur*, November 1784.
- [6] Anonymus. 1784b. Schreiben über die Kempelische Schachspiel- und Redemaschine. *Hessische Beyträge zur Gelehrsamkeit und Kunst*. Bd. 1, Nr. 3, Frankfurt.
- [7] F. Brackhane. 2009. Die Sprechmaschine Wolfgang von Kempelens – eine phonetische Analyse. Magisterarbeit an der Universität des Saarlandes.
- [8] F. Brackhane & J. Trouvain 2008. What makes "Mama" and "Papa" acceptable? – Experiments with a replica of von Kempelen's speaking machine. *Proceedings of the 8th International Speech Production Seminar (ISSP)*, Straßburg, 333-336.
- [9] B. Kröger & P. Birkholz. 2009. Artikulatorische Sprachsynthese. *Tagungsbände der 20. ESSV (Band 1)*, Dresden, 182-189.
- [10] Information aus einem persönlichen Gespräch mit der Kempelen-Biographin Alice Reininger, September 2009.
- [11] F. Alipour & S. Jaiswal 2008. Phonatory characteristics of excised pig, sheep and cow larynges. *Journal of the Acoustical Society of America* 123 (6), 4572-4581.
- [12] W. Wokurek & A. Madsack. 2008. Messung subglottaler Resonanzen mit Beschleunigungssensoren. *Fortschritte der Akustik – DAGA 2008*, Dresden, 125-126.
- [13] J.J. Ohala & C.J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. In: J. J. Wolf & D. H. Klatt (eds.), *Speech communication papers*. New York: Acoust. Soc. of Am., 89 - 92.
- [14] Chr. G. Kratzenstein 1781. *Tentamen Resolvendi Problema ab Akademia Scientiarum Imperiali Petropolitana ad Annum 1780 Publicae Propositum*.
- [15] K.J. Kohler. 2000. The future of phonetics. *Journal of the International Phonetic Association* 30, 1-24.
- [16] F. Brackhane. 2009. Die linguistische "Sprachorgel" – Über die Verbindung von Orgelbau und Sprachsynthese. *organ – Journal für die Orgel* 4/09, 42-46.
- [17] I. Titze. 1994. *Principles of Voice Production*. Englewood Cliffs: Prentice-Hall.
- [18] P. Birkholz. 2007. Articulatory synthesis of singing. *Special Session on "Synthesis*

- of Singing", Proceedings Interspeech, Antwerpen.
- [19] K. Fukui, Y. Ishikawa, E. Shintaku, K. Ohno, N. Sakakibara, A. Takanishi & M. Honda. 2008. Vocal cord model to control various voices for anthropomorphic talking robot. Proceedings of the 8th International Speech Production Seminar (ISSP), Straßburg, 341-344.
- [20] A. Reininger. 2007. Wolfgang von Kempelen. Eine Biografie. Praesens Verlag. Wien.
- [21] L. Euler. 1761. Lettres à une princesse d'Allemagne sur divers sujets de physique & de philosophie. 137. Brief vom 16. Juni 1761.
- [22] F. Burkhardt & J. Stegmann. 2009. Emotional speech synthesis: Applications, history and possible future. Tagungsbände der 20. ESSV (Band 1), Dresden, 190-199.