Historic Phonetic Devices in the Education in Electrical Engineering and Information Technology

Rüdiger Hoffmann, Dieter Mehnert

Technische Universität Dresden, Chair for System Theory and Speech Technology, 01062 Dresden, Germany

Abstract — **This paper deals with some experiences with the application of historic phonetic devices in the education in engineering. The exhibits are part of the historic acoustic-phonetic collection (HAPS) of the Dresden University. This collection demonstrates the development of experimental phonetics and speech technology in Germany from the beginning at the end of the 19th century until the introduction of the computer in speech processing. Selected application examples are demonstrated for sound production, sound recording, sound analysis, and sound synthesis.**

Index Terms — **Acoustic devices, acoustic signal processing, electromechanical devices, engineering education, history, speech processing.**

I. Introduction

Academic education in acoustics and speech technology is frequently performed at the departments of electrical engineering. Therefore, numerous objects which are demonstrating the history of this scientific field can be found in historic collections of electrotechnology. At the respective department of the Dresden University of Technology, acoustic problems have been considered since the appointment of Heinrich Barkhausen (1881–1956) in 1911, who is mainly known as a pioneer of radio engineering. After the 1950th, speech technology was included into to the classical acoustic disciplines in Dresden, starting with the development of a vocoder [1] and continuing with research in speech recognition and speech synthesis which is still going on. Clearly, a number of objects are available from the past, illustrating the way of speech processing devices from electronic valves, transistors, and integrated circuits up to embedded systems. These relics formed the starting point for our historic acoustic-phonetic collection (HAPS).

Although speech technology is basing on electronics and, in its later times, computer engineering, it has strong roots which range back into the pre-electronic era. Since the end of the 19th century, the rapid development of experimental phonetics required the invention of numerous sophisticated devices. They were applied to create the foundations not only of the modern speech sciences but also of the speech technology which we find in our most recent communication devices. Therefore it is recommended to expand a historic collection in speech communication by objects coming from the pre-electronic times of early experimental phonetics.

In our Dresden collection, this extension was enabled by the close cooperation with some phonetic institutes in Germany. Traditionally, there were close connections between speech technology research in Dresden and phonetics in Berlin [2]. The fusion with the former phonetic collection of the Hamburg University in 2005 was the most important step to form a collection of high completeness in its field, which is unique at least in Europe. The collection is able to demonstrate the development of speech technology from the end of the 19th century until the introduction of the computer in the speech laboratories. There are some short descriptions of the collection (e. g., [3]) and a catalogue [4].

Historic collections at universities have to support the research as well as the teaching activities. This paper deals with the inclusion of the historic acoustic-phonetic collection in the education at the department of electrical engineering and information technology. For this purpose, we continuously pursue a project which develops teaching units where historic mechanical objects are presented along with their modern electronic counterparts.

In the following, we want to demonstrate some examples, which were selected to cover the following partial fields of different lectures:

- Sound production,
- sound recording,
- sound analysis,
- sound synthesis.

II. Sound Production

The collection includes a big amount of chimes, tuning forks and other sound generators which aimed as standard devices in the past. Usually, mechanic and electric systems can be mathematically described in a comparable way. This approach has growing importance in the modern field of mechatronics [5]. Therefore, the historic objects are valuable demonstrators for the unified theory of acoustical systems. Remarkably, they show impressive values of the resonance quality which is very hard to be reached by electronic means.

Fig. 1. Set of tuning forks for audiometric purposes. Manufactured by the company H. Pfau, Berlin.

An interesting application of tuning forks builds a bridge to psychoacoustics. One of our most interesting exhibits is a set of six tuning forks (Figure 1) with a controlled excitation mechanism, which was used for estimating the threshold of hearing. A listener is hearing normally if he perceives the sound of a fork longer than a defined time span.

E. g., if the tuning fork for 2,000 Hz is excited, a sound pressure level of 97 dB is produced. The time span, which is indicated at the device, is 90 s. After that time, we measured a level of 17 dB. Comparing the requirements of modern audiometry, this level is in the "normal-hearing" range. A complete set of measurements was published in [6].

Historic sound generators have also been used to study the mechanism of voicing. The collection includes a number of originals of the so-called cushion pipes from the German phonetician Franz Wethlo (1877–1960) which were pioneering mechanical models of the human larynx. Again, we investigated the performance of these historic instruments using state-of-the-art technology [7].

III. Sound Recording

It must be remembered that all the sound samples which served to develop our basic knowledge of the acoustic nature of speech were recorded by mechanical means (kymographs, wax cylinders, etc.). We are able to demonstrate the development of sound recording from the kymograph up to digital electronic recording by museum objects.

The transducer, which converts the acoustic vibrations into the movement of the pointer of the kymograph, was the most crucial element in producing a kymographic recording. The approach which was frequently applied dates back to E. J. Marey (1830–1904) who used it for recordings of different physiologic motions. Later, it was widely applied in experimental phonetics by P.-J. Rousselot (1846–1924) and his successors. We have demonstrated in [8] that the different constructions showed very specific transfer functions. The experimenters knew the restrictions and selected the capsule type which produced the largest deflection for a given voice (Figure 2). Nevertheless, the magnitudes which are achievable were rather small, and sophisticated technologies had to be developed for measuring the speech parameters (mainly pitch [9]) with sufficient accuracy.

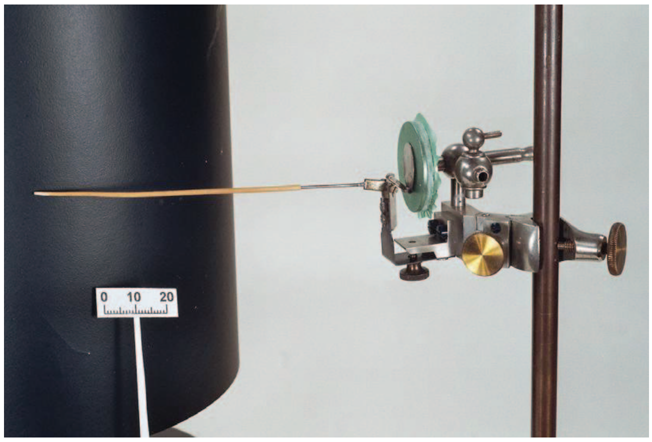


Fig. 2. One version of Marey's capsules writing at the revolving drum of a kymograph. Manufactured by E. Zimmermann, Leipzig, 1913.

Kymographs as well as the tools for interpreting the kymographic recordings have been continuously improved and commercially produced until the 1950th.

IV. Sound Analysis

The analysis of sounds aims predominantly to the estimation of their spectral composition. As already mentioned, a number of mechanical resonators were available for a qualitative analysis with restricted accuracy [8]. For a quantitative analysis, Fourier analysis had to be applied at the kymographic (or other) recordings of the sounds.

Spectral analysis as the fundamental method for describing sound signals was basically available with the works from Joseph Fourier (1768–1830). Due to the complexity of the Fourier analysis, computing aids were developed very early. Among them, the harmonic analyzers from Yule and Mader are the most important ones. George Udny Yule (1871–1951)

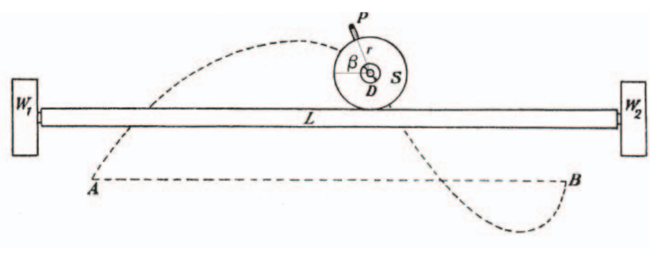


Fig. 3. Principle of the harmonic analyser from Yule, adapted from the explanation of the device by F. A. Willers [11].

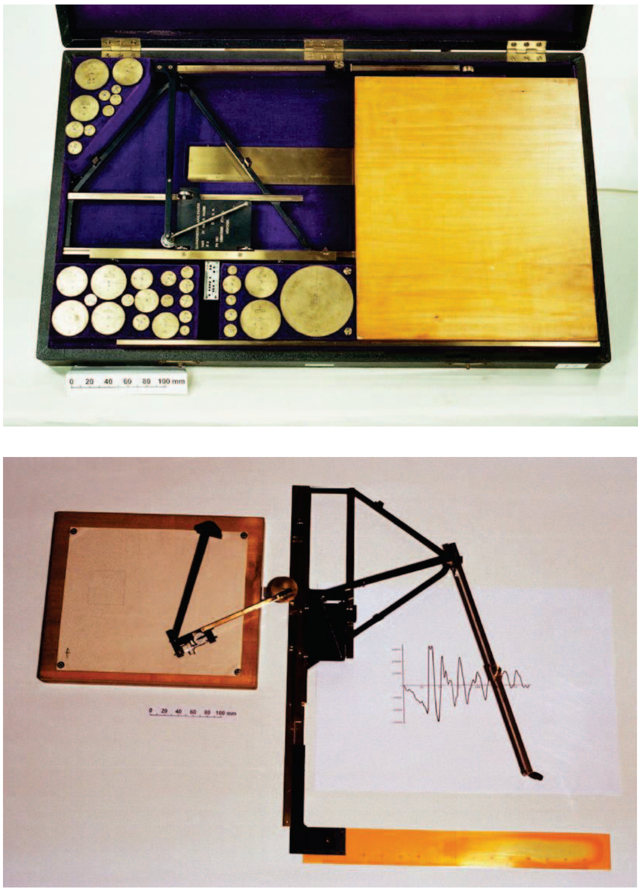


Fig. 4. Harmonic analyser from Mader, manufactured by the company Gebr. Störzl, Munich, 1910. – Above: Device in the transport box. The different cogwheels are required for calculating the Fourier coefficients of different order. – Below: The analyser in action, combined with a planimeter.

described in 1895 his idea of a mechanical analog computer, which transforms one period of a given waveform into that function which must be simply integrated by a planimeter to get a Fourier coefficient [10]. The planimeter is attached to a point *P* at the periphery of a cogwheel, the movement of which performs the required harmonic transformation (see Figure 3).

The drawback of the device lies in the fixed length *AB* of the signal period to be analysed. The harmonic analyser from Mader [11] combined the approach from Yule with a set of levers which allowed the device to be applied to drawings of signal periods of arbitrary length *AB*. This device was frequently applied in experimental phonetics and produced until the 1950th (see Figure 4).

The explanation of the harmonic analysis with the Mader analyser should be subdivided into three items:

- how to calculate integrals by a planimeter,
- how the Yule analyser works,
- how the Yule analyser is improved the Mader analyser.

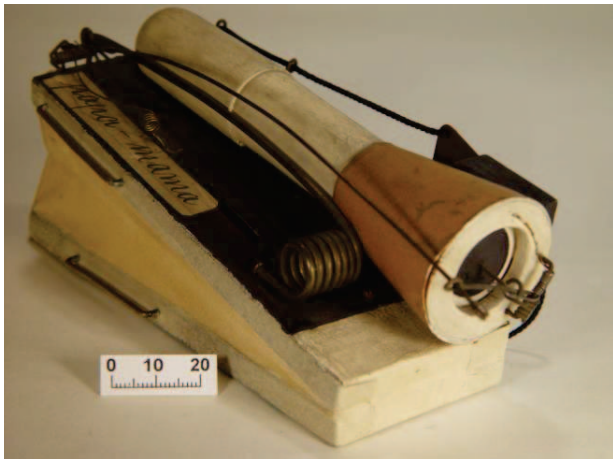


Fig. 5. Mechanical voice for “mama” and “papa”, probably produced by Hugo Hölbe, Sonneberg, Germany, at the end of the 19th century. The bellow is lifted for the photograph and will move downward if the device is speaking.

The core of the explanation, of course, should be focused on the Yule analyser which is described very well, e. g., in [12]. We do this regularly in our lectures on signal processing. It is very impressive to demonstrate the students, how a harmonic analyzer works, and to compare it with the FFT algorithms which are preferably applied since their invention in the 1960th.

V. Sound Synthesis

The development of devices which are able to produce human-like sounds is very old, but the scientifically based speech synthesis started in the age of enlightening with the work from several authors. The most influential work was the “speaking machine” from Wolfgang von Kempelen (1734–1804) who coined the famous sentence: „It is possible to make a machine which speaks everything." [13]

It is less known that the German mechanic and musician Johann Nepomuk Mälzel (1772–1838) bought and marketed the automata of Kempelen after his dead and received a patent in Paris in 1824 on a puppet voice which was able to pronounce the words “mama” and “papa” [14]. This was remarkably the first commercial spin-off in the history of speech technology! Later on, the manufacturers of toys and puppets produced such “voices” in large extent.

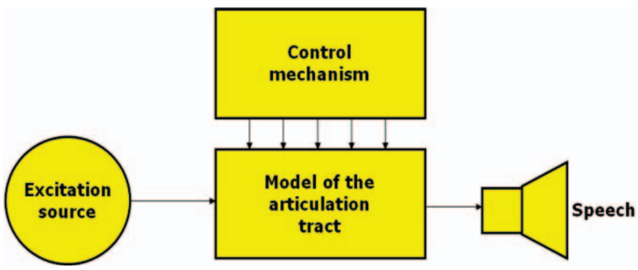


Fig. 6..General structure of a parametric speech synthesis system.

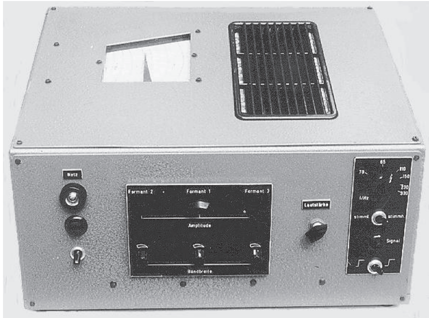
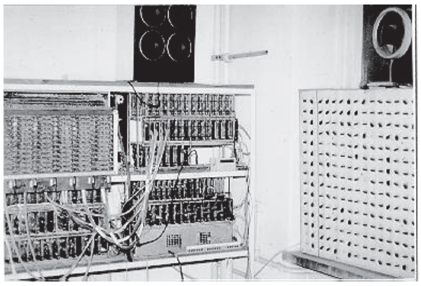
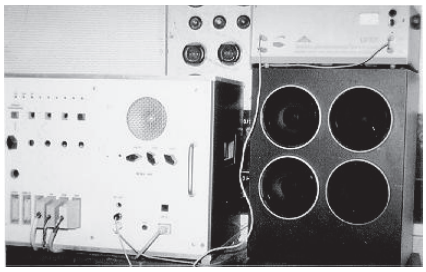
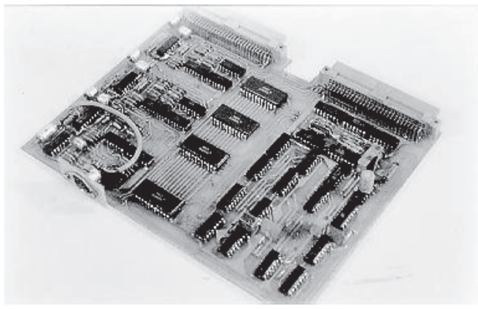
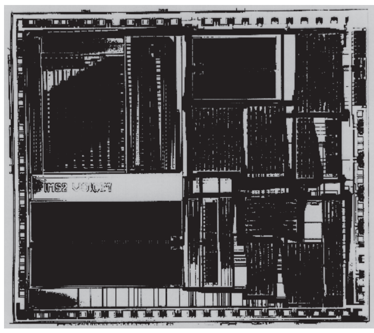
Device	Year	Technical Base	Control	Photograph
Vowel synthesizer	1962	Electronic valves	Manually	
SYNI	1975	Germanium transistors	Paper tape	
ROSY	1977	Gatter circuits (SSI)	Minicomputer	
TUSY	1987	Integrated circuits (MSI, LSI)	Microprocessor	
VOICE 1	1993	VLSI circuit (ASIC)	Embedding device	

Fig. 7: Historic development of speech synthesis devices following the paradigm of Figure 6, demonstrated using the developments of the TU Dresden guided by Walter Tscheschner (1927–2004). From [17].

The HAPS owns several examples which were manufactured at the end of the 19th century (Figure 5). They came to the phonetic laboratory because the otologist Johannes Kessel (1839–1907) proposed to apply them in the rehabilitation of hard-hearing patients [15].

The mechanical voices are suited to demonstrate very fundamental aspects of modeling and analysis-by-synthesis which we have utilized, e. g., in [16]. Analysis-by-synthesis is very natural in speech technology because everybody will agree that building a speech based system means to design and implement a model of that what humans do if they are speaking or listening. Kempelen started his development with the statement [13]: "I thought that a speaking machine needs nothing than a lung, a larynx, and a mouth." Indeed, if we look at the device in Figure 5, the lung is a bellow, the larynx is a metallic tongue, and the mouth (better the articulation tract) is a pipe.

Kempelen controlled his "speaking machine" manually. Therefore he forgot to mention a very essential part: a control unit. In the case of the mechanical voices, the control is performed by a mechanical arrangement which is driven by the movement of the bellow. With this supplement, the structure of a speech synthesis system is completed, as it is illustrated in Figure 6.

The step from the mechanical to the electronic technology appears to be very natural also in this case. The excitation source is implemented by a combination of a noise generator (for unvoiced sounds) and a periodic signal generator (for voiced sounds). The articulation tract can be modeled in different ways by a combination of electronic filters, which follow a theoretic fundament which was mainly developed by Gunnar Fant (1919–2009) and Manfred Schröder (1926–2009).

The subsequent development of electronic speech synthesizers followed the progress of electronics as can be seen in Figure 7 [17], starting with electronic valves and finally arriving at VLSI circuits. In parallel, the control units made use of the evolving computer engineering, which enabled the incorporation of text-processing components. Recent speech synthesis systems are therefore called text-to-speech (TTS) systems.

Although synthetic speech is alternatively generated since the 1990-th by the concatenation of segments of recorded natural speech, the parametric approach of Figure 6 remains the most important paradigm [16].

VI. Conclusion

These examples show that our modern speech communication equipment has two roots in the 19th century: early electric communication devices like the first telephones, and early mechanical equipment for research in acoustics and phonetics. Both lines converged later on, mainly after the invention of the electronic valve. This historic line can be demonstrated by our museum items.

Acknowledgement

The authors wish to thank Dr. Rolf Dietzel for kind support, especially for taking the photographs.

References

[1] R. Hoffmann, "On the development of early vocoders," *Proc. 2nd IEEE Histelcon, Madrid.,* pp. 359–364, November 2010.

[2] R. Hoffmann, and D. Mehnert, "Berlin–Dresden traditions in experimental phonetics and speech communication", *Un siècle de phonétique expérimentale (ed. by L.-J. Boe and C.-E. Vilain),* Lyon: ENS Éditions 2010, pp. 191–208.

[3] R. Hoffmann, and D. Mehnert, "Early experimental phonetics in Germany – Historic traces in the collection of TU Dresden", *Proc. 16th Int. Congress of Phonetic Sciences, Saarbrücken,* pp. 881–884, August 2007.

[4] D. Mehnert, *Historische phonetische Geräte (Catalogue),* Dresden: TUDpress 2012 (Studientexte zur Sprachkommunikation, vol. 62).

[5] A. Lenk, R. G. Ballas, R. Werthschützky, and G. Pfeifer, *Electromechanical systems in microtechnology and mechatronics,* Berlin etc.: Springer 2011.

[6] R. Dietzel, R. Hoffmann, D. Mehnert, and D. Richter, „Mechanische Resonatoren als Schallquellen und Messmittel in der experimentellen Akustik", *DAGA 2012, Fortschritte der Akustik,* pp. 125–126, March 2012.

[7] R. Hoffmann, D. Mehnert, R. Dietzel, and U. Kordon, "Acoustic experiments with Wethlo's larynx model", *Grazer Linguistische Studien,* vol. 62, pp. 51–59, 2004.

[8] R. Hoffmann, D. Mehnert, and R. Dietzel, "Measuring the accuracy of historic phonetic instruments", *Proc. 17th Int. Congress of Phonetic Sciences, HongKong,* pp. 176–179, August 2011.

[9] D. Mehnert, and R. Hoffmann, "Measuring pitch with historic phonetic devices", *Proc. 3rd Int. Conf. on Speech Prosody, Dresden,* vol. 2, pp. 927–931, May 2006.

[10] G. U. Yule, "On a simple form of harmonic analyser", *The London, Edinburgh, and Dublin philosophical magazine and journal of science,* ser. 5, vol. 39, pp. 367–374, 1895.

[11] O. Mader, "Ein einfacher harmonischer Analysator", *ETZ*, vol. 30, pp. 847–849, 1909.

[12] F. A. Willers, *Mathematische Instrumente,* Berlin and Leipzig: de Gruyter 1926 (Sammlung Göschen, vol. 922).

[13] W. v. Kempelen, *Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine,* Wien: J. Degen 1791.

[14] J. Maelzel, *Mecanique ou automate dite poupée parlante,* French Patent No. 1600, 31. Jan. 1824.

[15] J. Kessel, "Demonstration von Apparaten zur Erzeugung künstlicher Laute", *Verhandlungen der Versammlung Deutscher Ohrenärzte und Taubstummenlehrer in München am 16. Sept. 1899,* Berlin: Staude 1900, pp. 28-29.

[16] R. Hoffmann, "Analysis-by-synthesis in prosody research", *Keynote lecture, Proc. 6th Int. Conf. on Speech Prosdy, Shanghai,* vol. 1, pp. 1-6, May 2012.

[17] R. Hoffmann, "Speech synthesis on the way to embedded systems", *Keynote lecture, Proc. 9th Int. Conf. on Speech and Computer (SPECOM), St. Petersburg,* pp. 17–26, June 2006.