FRANCESCO AVANZINI, PIERO COSI, ROLANDO FÜSTÖS, ANDREA SANDI

WHEN FANTASY MEETS SCIENCE: AN ATTEMPT TO RECREATE THE VOICE OF ÖTZI THE "ICEMAN"

Abstract: Ötzi the Iceman's, the mummy found some years ago on the top of the Similaun mountain, lived 5300 years ago, and obviously, we cannot say which language, which phonemes, or even which sounds could belong to him. In this work, still in its initial stage, a possible reconstruction of an "approximation" of Ötzi the Iceman's voice, recreating the timbre or colour of his stone-age possible vowels will be described.

Keywords: Ötzi, Vocal Tract, CT Scan.

1. Introduction

Ötzi (*the Iceman*) was found by two German hikers in 1991, frozen and mummified in Schnalstal glacier, Ötztal Alps, near Hauslabjoch in South Tyrol on the border between Austria and Italy.
Ötzi and his artefacts have been exhibited at the South Tyrol Museum of Archaeology in Bolzano, Italy since 1998 (see Figure 1).
Radiocarbon dating gives ages between 5300 and 5200 years, placing it in the Age of copper, the moment of transition between the Neolithic and Bronze Age. So Ötzi is an ancient mummified specimen of homo sapiens and is Europe's oldest known natural mummy, providing researchers with an interesting picture describing what life was like around 3300 BC.
There are particular difficulties working with a 5300-year-old, especially conserved body - by far the oldest mummified person ever found; those of ancient Egypt are at least 1,000 years younger. Nonetheless, during these last years, scientists have discovered many interesting things about the mummy. The discovery made by X-rays and CT scans that Ötzi had an arrowhead lodged in his left shoulder led researchers think that he died of blood loss from the wound (Jha 2007). Further research found bruises and cuts to the hands, wrists and chest, and cerebral trauma suggested a blow to the head. At present, researchers believe that death could have been also caused by a blow to the head but it is difficult to say if this was due to a fall, or from being struck with a rock by another person (Carrol, 2002).

Figure 1 - *Ötzi (the iceman), the mummy kept in the Archaeological Museum of Alto Adige of Bolzano*

Ötzi exhibits the oldest preserved tattoos in the world. These 61 tattoos are concentrated in areas where bones X-ray examination showed "age-conditioned or strain-induced degeneration", thus led to hypothesise that these tattoos may have been related to pain relief treatments similar to acupuncture (Singh, Ernst, 2008), and not to spiritual or magical meanings. If so, this is at least 2000 years before their previously known earliest use in China (c. 1000 BC).

Modern DNA sequence analysis (Palmer, 2012) revealed that Ötzi had brown eyes, blood type 'O', was lactose intolerant, and was likely to suffer heart disease. He was closely related to modern Corsicans and Sardinians and he was also the first known case of a person infected by the Lyme disease bacterium. He is middle-aged.

Summarizing in other words, thanks to him, experts were able to give us a quite complete picture of his past, and, by reconstructing his voice, we hope to gain more insight into what humans might have sounded like back then.

Obviously, we cannot say that we could reconstruct Ötzi's original voice because we miss too many crucial information from the mummy. However, with two reconstructed but true measurements, the length of both the vocal tract and the vocal cords, we have been able to recreate with a "high degree of fantasy", a hypothesized approximation of the mummy's vocal tract structure and consequently a hypothesized approximation of its voice.

2. Vocal Tract Reconstruction

We had to face several challenges as we worked to reconstruct the 5300-year-old mummy's vocal tract configuration.

We had to deal with Ötzi's position, which makes the CT scanning quite difficult. Ötzi's arm is covering his neck and throat, and, for our project, this is the worst position you can imagine. Moreover, the hyoid bone, or tongue-bone, was partially absorbed and dislocated.
Considering the scan of similar human body as a reference, and with the help of a special dedicated software, available at SINTAC Biomedical Engineering in Padova[1], we were able to virtually move Ötzi's arm from his position, reposition his skull in the erect position, reconstruct his vertebrae, from the first one (C1) closest to the skull to the first thoracic vertebra (T1), and reconstruct and reposition the hyoid bone.
At last, as illustrated in Figure 2, we ended up with a complete model of the vocal tract (*Vocal Tract Lenght=127.3mm, Vocal Chord Length=20.47)*, but still we underline again, that we were missing important data such as the tension and density of the vocal cords or the thickness and composition of the soft tissues that take an important role in the human voice.
MRI (magnetic resonance imaging) scans would have helped us getting more insights, but this technology could not be used because of the condition of Ötzi's mummified body. Thus, we had only to rely on mathematical models and a software that simulates the way the vocal tract works.
Taking into consideration that, despite its short slender body, Ötzi's had a rather large head, it was possible to hypothesize that his voice could probably had a fundamental frequency between 100 Hz and 150 Hz, in line with today's average male (voices).
With only these simple measurements and this hypothesized f0 finally, with the help of a specialized software developed by Peter Birkholz of the Institute of Acoustics and Speech Communication of the Technische Universität of Dresden (Birkholz, Lehnert, Neuschaefer-Rube, 2009), we let Ötzi speak with few vowel like sounds and few simple words/sentences.

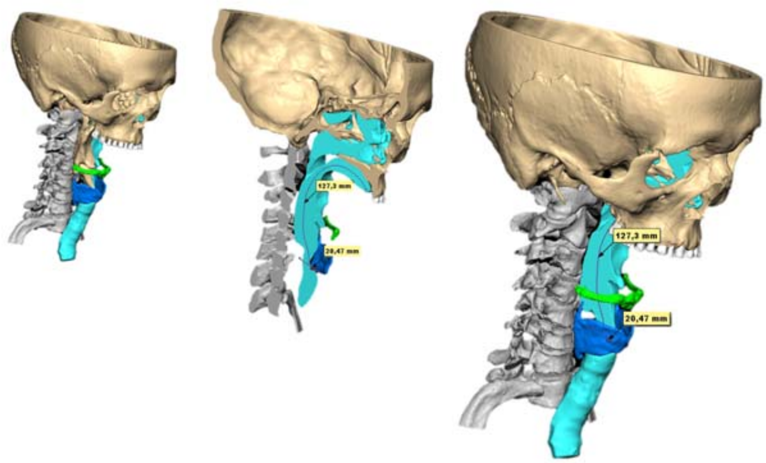
3. From Vocal Tract to Speech Sounds

Acoustic tube models of speech production have been heavily studied (Chiba and Kajiyama, 1941; Dunn, 1950; Stevens, Kasowski, Fant, 1953; Fant, 1959; Fant, 1960; Kelly and Lochbaum, 1962; Flanagan, Rabiner, 1973, Tousignant, Lefevre, Lecours, 1979). In these studies, it was shown that from a given tube shape, the resonance frequencies could be obtained, and the inverse problem of determining a unique tube shape from resonance characteristics has also received considerable attention in many other studies (Schroeder, Mermelstein 1965; Mermelstein, 1967; Schroeder, 1967; Heinz, 1967; Sondhi, Gopinath, 1971).

---

[1] SINTAC Biomedical Engineering, Padova
http://www.sintac.it/

Figure 2 - *Ötzi's vocal tract reconstructed model: Vocal Tract Lenght=127.3mm, Vocal Chord Length=20.47*



From Linear Prediction of Speech Model theory (Markel, Gray, 1976), it is well known that the vocal tract could be modeled as a succession of elementary cylindrical tubes, and the corresponding area function can be represented by reflection coefficients between adjacent sections of the tube. This LP model allows the parameters of these acoustic tube models to be estimated directly from the acoustical speech waveform.

The first attempt at directly computing an acoustic tube model of the vocal tract from the speech waveform was due to Atal (Atal, 1970) who demonstrated that the formant frequencies and bandwidths are sufficient to uniquely determine the areas of an acoustic tube having a specified number of sections. He also demonstrated that a transfer function with M poles is always realizable as the transfer function of an acoustic tube consisting of M cylindrical sections of equal length (Atal and Hanauer, 1971). Thus, a unique discrete tube shape can be reconstructed from a given-order polynomial transfer function. Wakita (Wakita, 1972) showed that the same acoustic tube model is equivalently represented from the inverse filter A(z) obtained by linear prediction of the acoustical speech waveforms. He also demonstrated the important experimental result that if the speech is properly pre-emphasized, and if boundary conditions of the acoustic tube are properly chosen, then very reasonable vocal tract shapes can be directly estimated using the autocorrelation method of linear prediction.

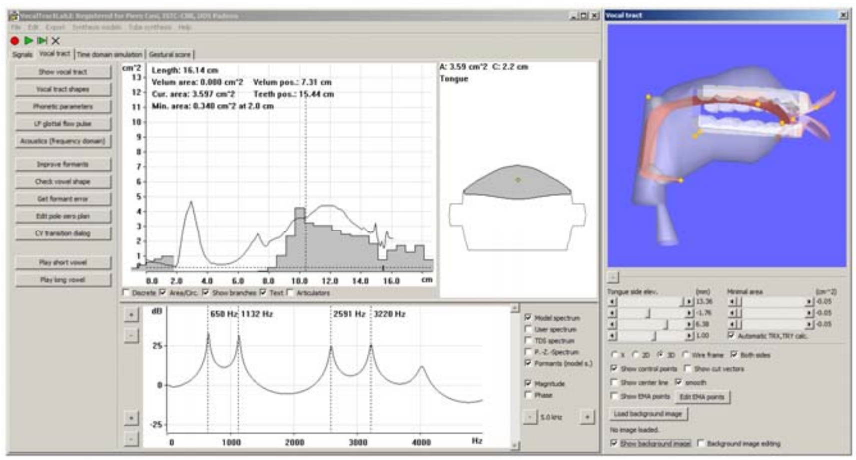
All these findings are the basis for the VocalTractLab (Birkholz, P., Lehnert, B., Neuschaefer-Rube, C., 2009) open source software[2].

---

[2] http://www.vocaltractlab.de/

The central element of VocalTractLab is a three-dimensional model of the human vocal tract, which represents the surfaces of the articulators and the vocal tract walls. An interactive visualization of the model is shown in Figure 3.

Figure 3 - V*isualization of a screenshot of the VocalTractLab software*



The shape and/or position of the articulators is defined by a number of vocal tract parameters, which can be changed interactively by dragging some control points (see the dots in the vocal tract picture in Figure 3). The pictures on the right side of Figure 3 display the cross section through the vocal tract at a selected position along the centerline (top), and the vocal tract area function (bottom). At the bottom of the screen is a graph of the volume velocity transfer function.

All possible articulators parameters are easily set by the user but, for Ötzi, we could have only two of them: the length of the vocal tract and the length of the vocal chords. The voiced sound corresponding to the adjusted vocal tract shape can be thus mathematically reconstructed, by using the above-described knowledge, and played back to the user.

4. Final Considerations

We would like again to repeat that, obviously, we could not say that we can reconstruct Ötzi's original voice because we miss too many crucial information from the mummy. However, with two true measurements, the length of both the vocal tract and the vocal cords, we have been able to recreate with a "high degree of fantasy", a hypothesized approximation of the mummy's vocal tract structure and consequently a hypothesized approximation of its voice.

Even knowing perfectly the Ötzi' respiratory capacity, his vocal chords dimensions and his vocal tract perfect configuration measurements, still, there will be

too many unknown factors, such as tissue characteristics to cite one, that make impossible a true reconstruction of its voice.
With all these limitations in mind, we wanted only to explore with the help of a lot of fantasy, but also “a certain degree” of science what could have been the old voice of Ötzi.

Acknowledgements

We would like to strongly acknowledge EURAC (the Institute for Mummies and the Iceman), and the South Tyrol Museum of Archaeology of Bolzano, for the courtesy to make us working on their Ötzi’s CT scannings, and Peter Birkholz of the Institute of Acoustics and Speech Communication of the Technische Universität of Dresden, for his kindness and patience, in explaining us the functioning of the VocalTractLab open source software, and for letting us use it, even for the complete speech synthesis stage.

Bibliography

ATAL, B.S. (1970). Determination ofthe Vocal Tract Shape Directly from the Speech Wave. In *Journal of Acoustical Society of America*, 47, 1979, 65(A).

ATAL, B.S., HANAUER, S.L. (1971). Speech Analysis and Synthesis by Linear Prediction of the Speech Wave. In *Journal of Acoustical Society of America*, 1971, 637-655.

BIRKHOLZ, P., LEHNERT, B., NEUSCHAEFER-RUBE, C. (2009). VocalTractLab – Ein neues Softwaretool für die artikulatorische Sprachsynthese in der Lehre. In *Proceedings of the 26th Jahrestagung der DGPP*, Leipzig, Germany, 2009, 209–211.

CARROLL, R. (2002), How Oetzi the Iceman was stabbed in the back and lost his fight for life, *The Guardian*, 2002.

CHIBA, T., KAJIYAMA, M. (1941). *The Vowel, Its Nature and Structure*, Tokyo Kaiseikan Publisher Company, Tokyo, 1941.

DUNN, H.K. (1950). The Calculation of Vowel Resonances, and an Electrical Vocal Tract. In *Journal of Acoustical Society of America*, 22, 1950, 740-753.

FANT, G. (1959). Acoustic analysis and synthesis of speech with applications to Swedish. Ericsson, Stockholm.

FANT, G. (1960). *Acoustic theory of speech production*. The Hague, Netherlands: Mouton.

FLANAGAN, L., RABINER, L.R. (1973). *Speech Synthesis*. Dowden. Hutchinson and Ross, Stroudsburg, Pennsylvania, 1973.

JHA, A. (2007). Iceman bled to death, scientists say, The Guardian (https://www.theguardian.com/science/2007/jun/07/archaeology.internationalnews) 2007.

HEINZ, M. (1967). Perturbation Functions for the Determination of Vocal Tract Area Functions from Vocal Tract Eigenvalues. In *Report Inst Tech., Stockholm Speech Trans. Lab. Q. Prog. and Status Rep.*, 1967, 1-14.

KELLY JR., L., LOCHBAUM, C. (1962)., Speech Synthesis. In *Proceedings of the Stockholm Speech Communication Seminar*. Reprinted in J. L. Flanagan and L. R. Rabiner, eds. (1973) *Speech Synthesis*. Dowden. Hutchinson and Ross, Stroudsburg, Pennsylvania, 1973, 127-130.

MARKEL, J.D., GRAY Jr., A.H. (1976). *Linear Prediction of Speech*. Springer-Verlag, 1976.

MERMELSTEIN, P. Determination of the Vocal-Tract Shape from Measured Formant Frequencies. In *Journal of Acoustical Society of America* 41, 1283-1294 (1967).

PALMER, J., (2012). *Oetzi the Iceman's nuclear genome gives new insights*. BBC News: Science & Environment, 2012 (http://www.bbc.com/news/science-environment-17191398).

SCHROEDER, M.R, MERMELSTEIN, P. (1965). Determination of Smoothed Cross-Sectional Area Function of the Vocal Tract from Formant Frequencies. In *Proceedings of the 5th lnternational Congress on Acoustics*, Liege, Belgium, D.E. Commins, editor (Imprimierie Georges Thone, Liege, 1965), paper A-24.

SCHROEDER, M.R (1967). Determination of the Geometry of the Human Vocal Tract by Acoustic Measurements. In *Journal of Acoustical Society of America* 41, 1967, 1002-1010.

SINGH, S., ERNST, E. (2008). *Aghi, pozioni e massaggi. La verità sulla medicina alternativa*, Rizzoli, 2008

SONDHI, M.M., GOPINATH, B. (1971). Determination of Vocal Tract Shape from Impulse Response at Lips. In *Journal of Acoustical Society of America* 49, 1971, 1867-1873.

STEVENS, K.N., KASOWSKI, S., FANT, C.G.M. (1953). An Electrical Analog of the Vocal Tract. In *Journal of Acoustical Society of America*, 25, 1953, 734-742.

TOUSIGNANT, B., LEFEVRE, J-P., LECOURS, M. (1979). Speech synthesis from vocal tract area function acoustical measurements. In Proceedings of ICASSP 1979, *IEEE International Conference on Acoustical Speech and Signal Processes*, Washington, D.C., April 2–4, 1979, 921–924.

WAKITA, H. (1972). Estimation of the Vocal Tract Shape by Optimal Inverse Filtering and Acoustic/Articulatory Conversion Methods. In *SCRL Monograph No.9, Speech Communications Research Laboratory*, Santa Barbara, California, 1972.

Author name, affiliation and email

Francesco Avanzini
ENT Department, General Hospital – Bozen
francesco.avanzini@asbz.it

Piero Cosi
ISTC-CNR, SSPD
Istituto di Scienze e Tecnologie della Cognizione - Consiglio Nazionale delle Ricerche
Sede Secondaria di Padova
Via Martiri della libertà, 2 - 35137 Padova
http://www.pd.istc.cnr.it/
piero.cosi@pd.istc.cnr.it

Rolando Füstös
ENT Department, General Hospital – Bozen
rolando.fustos@sabes.it

Andrea Sandi
SINTAC Biomedical Engineering, Padova
http://www.sintac.it/
andrea.sandi@sintac.it