

DNA-based prediction of Nietzsche's voice

Flavia Montaggio¹, Patricia Montaggio², Imp Kerr³

Communicated by Margretta Covert, Yale University School of Medicine, New Haven, CT, February 28, 2015

Abstract

This paper presents a protocol for the accurate prediction of an individual's voice based on genotype data, specifically from single nucleotide polymorphisms (SNPs). We collected trace amounts of cellular material (Touch DNA) from books that belonged to the philosopher Friedrich Nietzsche (1844-1900). DNA was extracted and amplified using DOP-PCR technique. Five different genomic DNAs were generated. Nietzsche's genotype was singled out after comparison to genotype data from one living relative of the Nietzsche family. Nietzsche's genotype data was analyzed using a DNA-based phenotyping assay, termed VoiceRator, that incorporates the 24 most informative voice SNPs based on their association with genes related to the phenotypic expression of the vocal tract and larynx structure and function.

An SNP-based voice profile of Nietzsche was inferred. The profile data was converted into bio-measures that were used to 3D-print a vocal tract and larynx through which phonation was organically generated. A composite of seven Text-to-Speech simulations was made using a sound morphing software. The result is presented in audio format and illustrates the first attempt at simulating the voice of a deceased person.

1. Department of Forensic Genetics, University of Basel, Switzerland.

2. Department of Phoniatics, University of Basel, Switzerland.

3. Nietzsche-Archiv, Weimar, Germany.

Nietzsche's private library

German philosopher Friedrich Nietzsche (1844–1900) studied classical philology at the Universities of Bonn and Leipzig. In 1869, despite lacking his doctorate and being only 24 years old, he was appointed to the chair of philology at the University of Basel, Switzerland. In 1889, a mental breakdown left him an invalid under the care of his mother, and then his sister, Elisabeth Förster-Nietzsche. In 1897, Swiss feminist and Nietzsche's friend Meta von Salis purchased a large house in Weimar, the Villa Silberblick, to provide a home for Nietzsche and her sister. After Nietzsche's death, Elisabeth bought the villa and turned it into the *Nietzsche-Archiv*. While Nietzsche's original handwritten manuscripts and letters are stored at the Villa Silberblick (in a fireproof safe), his personal collection of books is kept at the Duchess Anna Amalia Library⁴, also located in Weimar.

Nietzsche's private library comprises more than 1,000 volumes with about 170 containing annotations left by Nietzsche. The books have been frequently handled by scholars and researchers for consultation, especially the annotated ones. With the assistance of Fräulein Imp Kerr of the *Nietzsche-Archiv*, we browsed Nietzsche's private library and retained 49 books bare of annotations (*lot AP9*), which we presumed would contain the least contaminated DNA sample from Nietzsche.

Among the 49 volumes selected (all showing minor discoloration) were sophisticated copies of Jean de La Bruyère's *Caractères*, Wilhelm Brambach's *Antike Rhythmik und Metrik*, Fyodor Dostoyevsky's *Notes from the Underground*, Arthur Schopenhauer's *Parerga*, and George Sand's *Lettres d'un Voyageur*.

4. In 2004, a fire broke out at the Duchess Anna Amalia Library. Twenty-five thousand books burned and over forty thousand were damaged beyond repair. However, the largest part of the collection—including Nietzsche's private library—was left untouched.

Collection of DNA Samples

Touch DNA sampling techniques are used at crime scenes to collect skin cells which are subsequently typed for DNA in the exact same manner as body fluids using standard laboratory procedures. Humans shed tens of thousands of skin cells each day and these cells may be transferred to surfaces our skin contacts. When you read a book, you may deposit a generous amount of skin cells on several pages. Touch DNA may be sampled from the touched pages. We tested Nietzsche's books for Touch DNA using Tape Lift method. We sampled the surface of 2,695 pages (55 pages per book) with pretreated 4" x 4" adhesive sheets to collect potential skin cells. From the samples collected, only 8 gave sufficient amounts of DNA to be considered exploitable.

PCR-based whole-genome amplification

Recent advances in sequencing technologies make it possible to routinely retrieve DNA sequences from century-old remains, even from small and damaged fragments.

DNA was extracted from samples by phenol-chloroform extraction method. DOP-PCR⁵ amplification was performed using Phi-29 DNA polymerase and random primers on 99 loci⁶ (2 loci on each autosomal pair, 2 loci on the X chromosome(s) and 1 locus on the Y chromosome) for each of the 8 samples. DNA extracts were amplified approximately 18,000-fold with a maximum

5. Degenerate oligonucleotide primed-polymerase chain reaction.

6. Locus (*plural loci*): specific location of a gene or DNA sequence on a chromosome.

bias of representation between the loci of 4-fold. Fragments were separated by pulsed-field gel electrophoresis (fig. 1-2).

Five different genomic DNAs (FuP1, FuP2, FuP3, FuP4, FuP5) were generated. Three from female individuals (lacking Y-chromosome specific genes) and two from male individuals (FuP2 and FuP5).

The Nietzsche family left no direct descendant but we were able to identify, locate and contact one indirect female descendant of the Nietzsche family, Aloysia Nag-Nietzsche (ALNN) from Stuttgart, Germany who agreed to take part in the study. DNA match was found between ALNN and FuP2, suggesting genotype FuP2 belonged to Nietzsche (fig. 3) (we also found a match between FuP1 and ALNN which encouraged us to think that genotype FuP1 belonged to Elisabeth Förster-Nietzsche).

SNP genotyping

The biomechanics and control of the human voice involve a multitude of tissues, membranes, muscles, glands, and bones. The voice is produced by vibration of the vocal folds (or vocal cords), which have a three-layer construction—epithelium, lamina propria, thyroarytenoid muscle—butterflying down cellular and extracellular matrices where collagen and elastin fibers, among other proteins, aggregate. The larynx—which size is affected by the production of testosterone—adjusts the length and tension of the vocal folds which tremors determine the frequency of the emitted sound (70-200 Hz for men, 140-400 Hz for women). The sound resonates in the nasal and mouth cavities where articulators like the tongue, palate, gums, velum, and uvula influence the general aerodynamics and acoustics. Furthermore, vocal folds physiology is impacted by humidity and viscosity factors, vocal resonance is

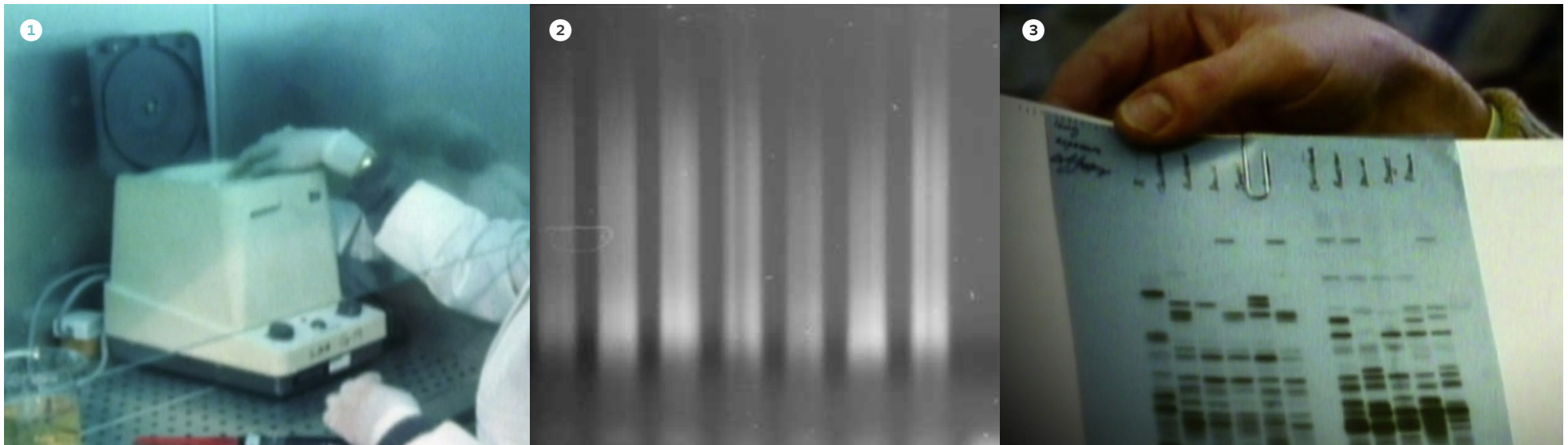


Figure 1. Fräulein Imp Kerr places PCR tubes into the thermal cycler. **Figure 2.** Lanes show gel analysis of DOP-PCR amplification products of DNA samples. No amplification is present in sample 8. **Figure 3.** Matching genomic DNAs of Friedrich Nietzsche (left) and indirect descendant Aloysia Nag-Nietzsche (right).

co-shaped by lung pressure and glottal flow, and brain circuits intervene in every movements underlying vocalizations.

One of the rapidly developing areas in forensic biology is the ability to predict characteristics of an individual's voice based on genotype data, specifically from single nucleotide polymorphisms (SNPs) (fig. 4). There have been several studies looking to identify the predominant SNP loci associated with human voice which might be used for accurate predictions. In 2012, Cooper *et al.* identified the 24 most informative voice SNPs (among more than 200 candidate "voice genes") based on their association with genes related to the phenotypic expression of the vocal tract structure and function (Table 1). The traits of focus indicated by these SNPs were found to be highly polymorphic and complex, involving several contributions from various gene-gene

interactions, suggesting intergenic complexity needed to be assessed through sets and subsets of SNPs (as opposed to single SNPs).

SNPs from Nietzsche's genotype were analyzed using a DNA-based phenotyping assay, termed VoiceRator, that incorporates the 24 voice predictor SNPs. The prediction model is based on a multinomial logistic regression method and uses phenotype and genotype data from 99 680 European individuals to which results are compared for parametrization. Based on these data the model has previously shown high prediction rates for vocal weight, texture, range, frequency and timbre in populations from Switzerland, France, and Germany.

Nietzsche's SNPs were benchmarked against 122 sets and subsets of predictor SNPs. Correlations between ensembles were derived for each SNP value.

| Gene | Name | Location | SNP | Association |
|---------|---|-------------|------------|--|
| MYH13 | Myosin-13 | 17p13.1 | rs16354291 | Laryngeal muscles |
| MYH8 | Myosin-8 | 17p13.1 | rs16789035 | Laryngeal muscles |
| KRT6A | Keratin 6A | 12q13.13 | rs24981908 | Epithelial tissues (tongue, oral mucosa) |
| KRT6B | Keratin 6B | 12q13.13 | rs17688854 | Epithelial tissues (tongue, oral mucosa) |
| KRT14 | Keratin 14 | 17q21.2 | rs74256550 | Epithelial cells |
| KRT16 | Keratin 16 | 17q21.2 | rs57829586 | Epithelial cells |
| KRT17 | Keratin 17 | 17q21.2 | rs82654506 | Cell differentiation in epithelia |
| COL4A1 | Collagen, Type IV, Alpha 1 | 13q34 | rs72498798 | Collagen (basal lamina, epithelium layer) |
| COL4A2 | Collagen, Type IV, Alpha 2 | 13q34 | rs52887413 | Collagen (basal lamina, epithelium layer) |
| COL4A3 | Collagen, Type IV, Alpha 3 | 2q36-q37 | rs49565648 | Collagen (basal lamina, epithelium layer) |
| COL4A4 | Collagen, Type IV, Alpha 4 | 2q35-q37 | rs73687386 | Collagen (basal lamina, epithelium layer) |
| COL4A5 | Collagen, Type IV, Alpha 5 | Xq22 | rs42736803 | Collagen (basal lamina, epithelium layer) |
| COL4A6 | Collagen, Type IV, Alpha 6 | Xq22 | rs27093543 | Collagen (basal lamina, epithelium layer) |
| MMP2 | Matrix Metalloproteinase 2 | 16q13-q21 | rs12846565 | Extracellular matrices of the vocal folds |
| ELN | Elastin | 7q11.23 | rs90887988 | Epithelial tissue |
| TBX22 | T-Box 22 | Xq21.1 | rs43562187 | Palate, nose |
| IRF6 | Interferon Regulatory Factor 6 | 1q32.3-q41 | rs19736143 | Palate (connective tissue) |
| AR | Androgen Receptor | Xq11-12 | rs22296857 | Testosterone |
| SHBG | Sex Hormone-Binding Globulin | 17p13.1 | rs89771654 | Testosterone |
| GDF6 | Growth Differentiation Factor 6 | 8q22.1 | rs96829456 | skull bones and joints |
| BMP2 | Bone Morphogenetic Protein 2 | 20p12 | rs90764367 | skull bones and cartilage |
| FOXP2 | Forkhead Box P2 | 7q31 | rs87241675 | Language regions of the brain, lungs, guts |
| CNTNAP2 | Contactin Associated Protein-Like 2 | 7q35 | rs88141243 | Cell adhesion molecules and receptors |
| SRPX2 | Sushi-Repeat Containing Protein, X-Linked 2 | Xq21.33-q23 | rs98724632 | Language centers of the brain |

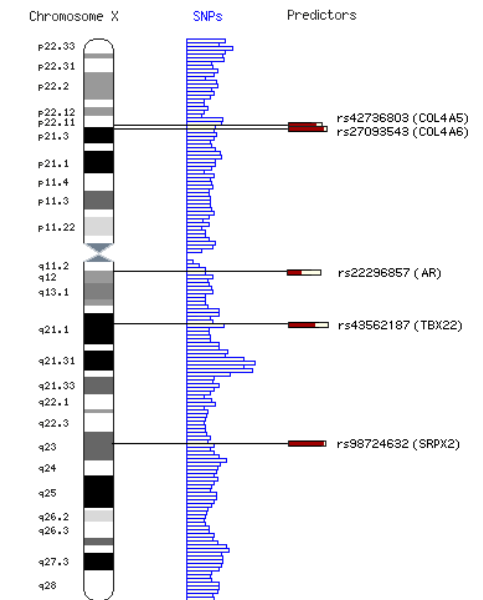


Figure 4. SNP loci on X chromosome.

Table 1. Top genetic predictors of human voice (Cooper *et al.*, 2012).

Contributions from the highest scoring sets were combined and re-parameterized, converted into bio-measures, and incorporated into the VoiceRator algorithm, leading to an exploitable SNP-based voice profile (fig. 5).

Biofabrication of Nietzsche's vocal tract and larynx

The application of 3D printing in tissue engineering has enabled new methods for the fabrication of organs and body parts using inkjet techniques. Synthetic biopolymers have been developed and combined with nanofibers and nano-structured particles to fabricate materials with selective bioactivity, as well as physical and chemical properties. Since 2014, the synthesis of biopolymers can be computer-tailored to spatially encode gene properties.

Bio-measures inferred from Nietzsche's voice profile data were used to build a 3D model of a vocal tract and larynx through which artificial voiced sounds were organically computed. The M-shaped model was 3D-printed with biopolymer-based composites (collagen, chondroitin sulphate, chitin) at subcellular resolution. Tissue-engineered constructs integrated phonatory aerodynamics, muscle contractions, viscoelastic properties, thermal agitation, glottal flow rate, phonatory fluid-structure interactions, frequency, vocal weight, texture, timbre, and prosody (fig. 6-7). Additional voice determinants (such as age⁷ and geography of accent) were manually input.

The model was coupled to an unsteady respiratory flow allowing variances

7. Human voice changes over time. In one longitudinal study by Endres *et al.* (1971), recordings of 7 speakers sampled over a time interval of up to 29 years were compared. The authors found a downward trend, as a function of increasing age, for fundamental frequency and formant frequencies.

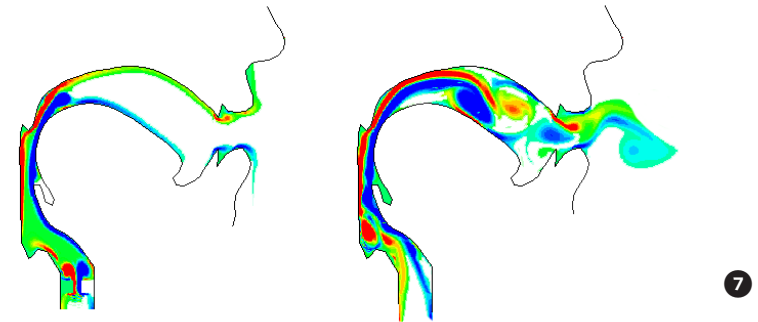
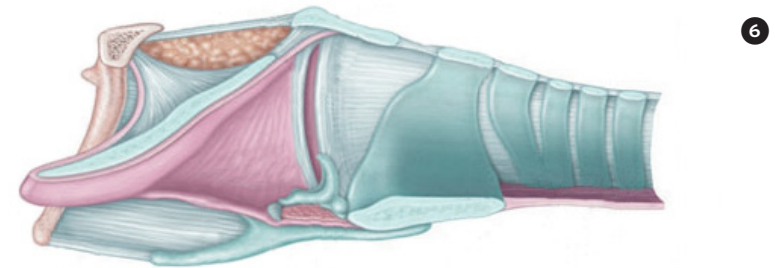
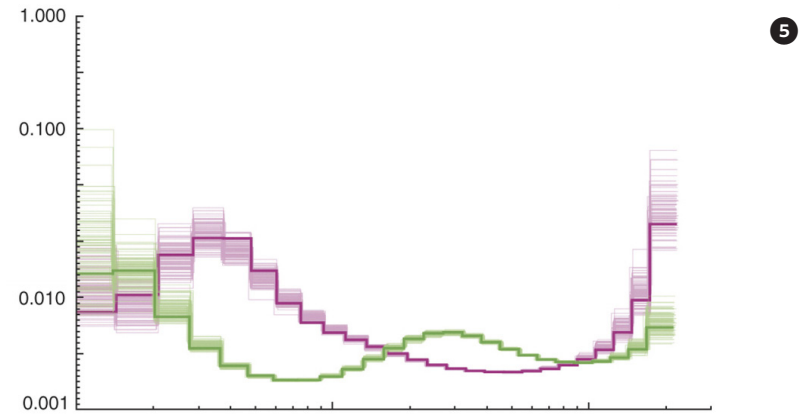


Figure 5. SNP-based voice profile of Nietzsche combining the two highest scoring sets of SNPs. **Figure 6.** 3D-printed larynx and vocal folds (sagittal section, detail). **Figure 7.** Computer-based modeling of glottal flow rate and thermal agitation.

from which lifelike phonation was obtained (previous phonation results compared well to subject-specific data—with a ~95 percent level of accuracy—demonstrating the precision of the VoiceRator modeling approach in addressing the complex interaction of biomechanics, fluid mechanics, and acoustics).

Simulation of Nietzsche's voice

We generated seven Text-to-Speech organic simulations compossible with Nietzsche's voice profile—one standard simulation, and six declensions incorporating tension asymmetry and mucosal wave-type variables. The voiced sounds generated by the model were recorded using a Neumann U87 microphone. We used a sound morphing software based on Gaussian Mixture Models to build a 7-fold monophonic composite out of the seven simulations—the audio clip is available here:

Nietzsche's voice at age 42, 7-fold composite (168KB)

▷ <http://www.impkerr.com/basel/nietzsche-composite.mp3>



Figure 8. Waveform of Nietzsche's voice at 42 (denoised 7-fold composite).

The input text was taken from a letter Nietzsche wrote to Georg Brandes on January 4, 1889: *Nachdem Du mich entdeckt hast, war es kein Kunststück mich zu finden: die Schwierigkeit ist jetzt die, mich zu verlieren (After you discovered me, it was no great feat to find me. The problem now is how to lose me).*

Phonaesthetics discussion

Nietzsche's vocalization exemplifies a flat, smooth, typical mild sounding voice: hued pitch, honeyed tone, low versatility in timbre, silvery hyponasal flow, disembodied texture with a touch of steel in vowels, and low head-nasal resonance (lower than expected in regards of Nietzsche's robust mandibles). The diction is unified and follows a three-wave prosodic structure with light elasticity. Range goes from low C to the B above middle C. Fundamental frequency ranges between 156.8 and 157.5 Hz.

Results

Nietzsche's voice simulation was tested against a control database of 4,287 voice recordings of living German individuals. Comparative tests based on 526 VoiceRator parameters were performed and unclocked 28 minor bias discrepancies between the control voice recordings and Nietzsche's voice simulation, resulting in a 94.68 percent level of accuracy consistent with previous results.

References

- [1] Diethel C (2003) *Nietzsche's Sister and The Will to Power: A Biography of Elisabeth Förster*. University of Illinois Press, Champaign, IL, pp. 92-93
- [2] Arneson N, Hughes S, Houlston R, Done S (2008) Whole-Genome Amplification by Degenerate Oligonucleotide Primed PCR (DOP-PCR). *Cold Spring Harbor Protocols* 01/2008. doi: 10.1101/pdb.prot4919
- [3] Williamson AL (2011) Touch DNA: Forensic Collection and Application to Investigations. *J Assoc Crime Scene Reconstr.* 2012;18(1);1-5
- [4] Romeika JM, Yan F (2013) Recent Advances in Forensic DNA Analysis. *Journal of Forensic Research* S12: 001. doi: 10.4172/2157-7145.S12-001
- [5] Cooper K, Harper S (2012) Testing of Voice Prediction Based on 24 SNPs. *Forensic Science International: Genetics*, 2012-6: 21–76. doi: 10.1016/j.forsciint.2012.06.008
- [6] Xue Q, Mittal R, Zheng X (2011) Computational Modeling of Phonatory Dynamics in a Tubular Three-dimensional Model of the Human Larynx. *Acoustical Society of America*, 2012 Sep; 132(3): 1602–1613. doi:10.1121/1.4740485
- [7] Li X, Cui R, Sun L, *et al.* (2014) 3D-Printed Biopolymers for Tissue Engineering Application. *International Journal of Polymer Science*, vol. 2014, Article ID 829145, 13 pages. doi: 10.1155/2014/829145
- [8] Service RF (2015) The Synthesis Machine Science. *Science*, Vol. 347 no. 6227 pp. 1190-1193. doi:10.1126/science.347.6227.1190
- [9] Kerr I (2009) Contesting Truth. *The Journal of Nietzsche Studies*, Issue 38 Autumn 2009, pp. 168-175

This study was funded by the Swiss National Science Foundation (201525-69812, MC-ME 7005). The computational resources were provided by Imp Kerr & Associates, NYC.