SOUND, SURVEILLANCE AND SIGNAL ANALYSIS. The other "Lautarchiv"

Techno-productive "operational" surveillance: The MfS "Stimmenarchiv"
"Cold listening" to sensitive sound archives
"Cantometrics"
Is there a "sound of the archive"?
Audio-recordings and their media-archaeological understanding

<see LAUTARCHIVE>

<see W. E., SONISCHES GEDÄCHTNIS ALS FUNKTION TECHNISCHER SPEICHER" (in: Pfleiderer (ed.) 2011>

SOUND, SURVEILLANCE AND SIGNAL ANALYSIS. The other "Lautarchiv"

[Research outline, presented at conference *Listening to the Archive. Histories of Sound Data in the Humanities and Sciences*, February 11-13, 2016, Humboldt-Universität / Max-Planck-Institut für Wissenschaftsgeschichte, Berlin]

Techno-productive "operational" surveillance: The MfS "Stimmenarchiv"

Let us take the very name of the Berlin *Lautarchiv* in its literal meaning. German *Laut* is not the musical but the phonetic parameter of audio communication and therefore asks for signal analysis rather than for hermeneutics of cultural semantics. From a dialogue with the previous professor for linguistic forensics at the former section "Kriminalistik" of Humboldt University (Koristka), the hypothesis arose that recordings in the Sound Archive (the *Lautarchiv*) housed at my Institute might have been used for voice identification by the GDR State Security. It turned out from the files in the archives of the former MfS that Koristka's writings have been read as a central authority for voice recognition indeed,

and in 1985 an IT project called "Phonotek" (which refers both to the technologies of voice identification in the sense of „Phonotec“ and the audio archive in the sense of „Phonotheque“) started systematic analysis "operativ-relevanter Sprecherstimmen"[1].

1 File MfS JHS [Juristische Hochschule Potsdam] 22035: Gärtner, Andreas, Die Bedeutung der Sprechererkennung nach meßtechnischen, hörtechnischen und operativen Faktoren bei der Bearbeitung der gegnerischen Geheimdienste und anderer relevanter Bereiche durch die HA III. Die Anwendbarkeit der meßtechnischen Faktoren einer Stimme als mögliche Recherchekriterien im Informationsgewinnungsprozeß, typescript, 18.3.1989, GVS [Geheime Verschlußsache] o026-344/89, JHS [Juristische Hochschule Potsdam] =

<see LAUTARCHIV-MFS: sw "Dialekt"; further: Xerox copies from BStU files, esp. articles by Koristka>

This sounds like a classical topic for archival research by historians, reconstructing the discursive, administrative and technical context.[2] The media-archaeological research track, though, has a decisively different emphasis; it does not focus in the ideological implications of acoustic sourveillance, but on the surplus in acoustic knowledge which actually arose from the application of voice-identifying technologies. In the former MfS, it has been the so-called "Technisch-Operativer Sektor" which developed or applied such tools. In a Foucauldean understanding, any paranoia (especially in agencies of state power) actually generates new methods and technologies of increasing knowledge. Media archaeology takes the very term "operative" (which is the key expression throughout almost all MfS files) in a productive sense: information provided by machines. Knowledge which results from techno-operative research is "timeless" beyond its limited historical or ideological discourse, since the language of what Nick Montfort la belled „technical report“[3] here deals with signal evidence rather than with semantic heuristics.

[At the MfS, the term "operative" is not just an umbrella word for all kind of surveillance activities, but even technically tighly linked to the archive. A brochure on computing defines the Random Access Memory as such: "Arbeitsspeicher (auch: Hauptspeicher, Operativspeicher, Zentralspeicher) ist Bestandteil der Zentraleinheit einer Datenverarbeitungsanlge."[4]]

Beyond ideological barriers in terms of cultural analysis and political correctness, there is a techno-formal language from the files which "speaks to us" in a non-historical way once our attention switches from the historian's to the media archaeologist's mood. In a remarkable document which deals with

---

Diploma, p. 8

2 Karen Bijsterveld at Maastricht University) performs ongoing research on the Voice Archive (*Stimmenarchiv*) of the former Ministry of State Security (MfS) of ex-GDR

3 Nick Montfort, Beyond the Journal and the Blog. The Technical Report for Communication in the Humanities, im *online*-Journal Amodern 1, Themenausgabe „The Future of the Scholarly Journal“ (2016): http://amodern.net/article/beyond-the-journal-and-the-blog-the-technical-report-for-communication-in-the-humanities/ (Abruf 11. Februar 2016)

4 BStU, Akte MfS Abt. 26, Nr. 820: Broschüre (Druckschrift) *Technische Kommunikation. Überblick über wichtige Grundbegriffe der technischen Kommunikation (Computer, CAD/CAM, Telekommunkation)*, zusammengestellt von Hans Maschke, hg. v. Zentralinstitut für sozialist. Wirtschaftsführung beim ZK d. SED, 2. überarb. Aufl. Berlin 1985 ("Nur für den Dienstgebrauch"), 28

applications of psycho-acoustic science and signal acoustics (the polygraph) to speech identification, an epistemological rupture in the analysis of acoustic evidence is described: the human voice becomes subject of analysis in a double way: object of observation, but "subject" to machine listening. Understanding of human speech is no exclusive agency of human ears any more. The author, an MfS officer, makes use of the appropriate metapher of breaking through the "sonic wall".

["Eine wesentliche Seite der tschekistischen Theorie stellen all jene Erkenntnisse dar, die sich mit der Rolle des Menschen als Subjekt und Objekt der operativen Prozesse beschäftigen. Gleich, woher die Impulse aus diesem Teil der operativen Theorie kommen, ob aus der operativen Erfahrung <...> ob aus den Ergebnissen der Wissenschaften <...>."[5] And further: "<...> haben wir in der operativen Theorie vom Menschen eine 'Schallmauer' durchstoßen - so relativ das Bild auch für die Dimensionen unserer Arbeit sein mag."[6]]

Due to the abrupt ending of that German state, the administative files on voice recordings (*Stimmenarchiv*) in the archives of the former State Security of German Democratic Republic have survived and are immediately accessible in terms of academic textual research. But what about the technical accessability of the audio files themselves?

The challenge which arises from the actual telephone voice recordings by former GDR State Security is not the obsolete hardware to read magnetic data.

[At the archives of the former State Security of German Democratic Republic[7], there is a special department for recovering ("Erschließung") machine-readable data from obsolete magnetic storage discs, headed by Stephan Konopatzky who has asked the Signal Laboratory of Humboldt University Media Studies for help in retro-computing the data hidden on large antique magnetic recording hard drives. In its last decade, GRD State Security increasingly had changed from type-written to computational data processing indeed.]

The cassette tapes remaining from the so-called "voice archive" ("Stimmenarchiv") is analog audio signals which can still be accessed by any commerical tape deck. This is voices in rather

5 File MfS JHS No. 165 "Gutachten zu den Forschungsergebnissen <...> 'Die wissenschaftliche Bewertung des psychophysiologischen Verfahrens der Stimmanalyse, seine Einsatzmöglichkeiten - Grundsätze in der politisch-operativen Aufklärungsarbeit des MfS' (vorgelegt von Oberstleutnant Roitzsch und Hauptmann Lips)". Typescript, p. 3 (signed: "Scharbert, Oberst", Potsdam, 2nd October 1979, Juristische Hochschule Potsdam) <= BStU Bl. 32>

6 Typescript p. 5 <= BStU Bl. 34>

7 Der Bundesbeauftragte für die Stasi-Unterlagen (BStU) in Berlin

technical than traditionally archival latency which, as long as they require electro-magnetical transduction for re-play, can first of all (i. e. on the media-archaeological level) be "heard" by machines only.

[When the pick-up transduces phonograpic grooves from mechanical "inscription" of physical sound into electro-magnetic current, is this "listening"?

<see passage MEDMEM-OSLO>

Fig.: Tonabnehmer, Lautarchiv]

The GDR State Security's definition for "Speaker archives" (*Sprecherarchiv*) was storage and retrieval systems for audio tapes on the basis of what was called "operative data" on the one hand and "speaker-typical characteristics" on the other.

[<copy LAUTARCHIVE> "Sprecherarchive sind rechnergestütze Speicher- und Recherchesysteme, die auf der Grundlage operativer Daten und Sprechertypischer Merkmale arbeiten."[8]]

The core media-archaeological research question is this: What difference did and still *does* it make if the administration and retrieval of analog voice recordings is not only computer-aided (that is, on the level of metadata), but the audio recordings themselves transform *from signals to data* by digital sampling? Then *signal* recording transforms into *information* storage.

The term used in our conference title is "sound data"; let us rather make a difference between *collections of phonographic analog audio signals* and *digital storage of sonic data*, since this difference is an essential change of medium state.

At the moment of the break-down of the GDR State Security in autumn 1989, surveillance of telephone voices had reached the point of changing from human identification of recorded voices (supplied by computer-based meta-data retrieval) to the application of fully *algorithmic* software for automatic voice recognition - which would created a completely different kind of "archive".

<see Modulares Fourier-Analyse-System: COMP-SPRACHANALYSE-MFA-104>

["Und es ist bereits abzusehen, wann Sonogramme" <sic> auch zum computergesteuerten Sprachvergleich im öffentlichen Telefonnetz gespeichert werden können. Der ehemalige BKA-Präsident Horst Herold malte <...> aus, daß Computer in der Lage sein werden, Inforemationen wie mit den Sinnen eines Menschen zu erfassen.

8 File MfS BdL No. 273 , letter (typescript) from 16th November 1989, concept by the Operativ-Technischer Sektor [BStU archival page no. 3-5]: *Dienstliche Bestimmung zur künftigen Arbeit mit Sprecherarchiven im MfS* (November 1989)

Dann, gegen Ende des Jahrzehnts, könne es technisch möglich werden, die Fahndung nach gesuchten Straftätern unmittelbar auf Maschine zu übertragen."[9]]

<Suchwort "algorith" in LAUTARCHIVE>

A special task force investigated options of the application of computing algorithms for automatic voice identification; here computational intelligence and secret „intelligence service" converge. One report insists on the remaining, unbridgable gap between human hearing ("Höranalyse") for issues which require *semantic* understanding, and analysis by sonagraphic *audio-signal* processing. Such "Meßanalyse" enables *differential* realtime voice identification, by comparing the signal output to similar recording evidence - which is the functional "archive", the *Sprecherarchiv*.

[Once more the name Koristka turns up which has been the academic mastermind behind such reasoning on acoustics. Among the MfS files figures a copy of Koristika's article on forensic voice analysis.[10] Koristka himself wrote his habilitation with Horst Völz on the forensic use of magnetic tape recording; see ARCTIC-ARCHIVE: "Im Speicherzustand erreicht die 'aufgenommene' Information gewissermaßen statische Eigenschaften, die verschiedentlich auch durch die Bezeichnung 'eingefrorene Information' charakterisiert werden."[11] .Koristka refers to L. G. Kersta's report on "Voiceprint Identification", in: Nature 196 (1962), 1253 ff. but differentiates the fingerprint as "direktes objektives Abbild" from the more contingent external conditions which envelop the recorded voiceprint <314>. Koristka defines the human Hearing Analysis (*Höranalyse*) as "subjective method", and the Measuring Analysis as "objective" <313>.]

An expert report on voice analysis defends the *Höranalyse*, since such heuristic *resonance* only results from empathetic listening.[12]

---

9 Typescript Gärtner 1989, p. 57 (BStU p. 57), quote from West German journal *Der Spiegel* (No. 26, 23 June, 1986)

10 Christian Koristka, Stimmanalysen - eine neue Methode der kriminalistischen Personenidentifizierung, in: Forschungen und Fortschritte, 41. Jg. <year ???>, Heft 10, 310-316, in: File MfS Abt. 26, Nr. 820

11 Christian Koristka, Magnettonaufzeichnungen und kriminalistische Praxis, Berlin (Ost) (Ministerium des Innern, Publikationabteilung) 1968, 24

12 "Kein Analysegerät ist in der Lage, die subjektiven Einschätzungen von solchen Paramtersn wie Gesamteindruck, stimmfülle, Klangfarbe usw. in objektiven Tatbeständen vorzunehmen und darzustellen." File MfS - HA XXII, Nr. 17247, typescript "Planaufgabe 2051. Thema: Konzeption für die Schulung von Mitarbeitern operativer Diensteinheiten zur sachkundigen Beschaffung von Ausgangs- und Vergleichsmaterial für die Personenidentifizierung anhan dder Stimme und Sprache", chapt. "Meßanalyse", p. 243

As has been argued in Licklider's seminal article on "Human Machine Symbiosis" in 1960, it is the differential combination of the "narrow band", but parallel signal processing human brain with the accuracy of the digital, though sequential computer which results in most efficient human-machine communication.[13]

But even voice sonagrams only serve for an intuitive first interpretation; future voice identification will be data (rather than signal) based, allowing for "automatic analysis" <Koristka: 316>. In this algorithmic approach, surveillance practice and recent methods called "Digital Humanities" converge in media ethically ambivalent, through knowledge-generating alliance.

[There is a temporal gap in the correlation between an actual voice recording with a signal from the *Stimmenarchiv*. In remarkable clarity - after removing circumstantial noise in the recording media and in background acoustics - an individual voice stays invariant against delay on the time axis[14], and its acoustic-phonetical characteristics are robust even against the speaker's intentional dissimulation.[15] This became media-theatrical drama in Samuel Becket's one-act play *Krapp's last tape* from 1959.]

[Hintergrundgeräusche und Störungen im Übertragungskanal (Mikrophon, Raumakustik) fordern automatische Spracherkennung heraus, während das menschliche Gehör in hohem Maße daraus zu filtern vermag (siehe Cocktail-Party-Effect, spektrale Verdeckungseffekte, neuronale Kurzzeitadaption, Refraktärzeiten und die von Jonathan Sterne beschriebenen Verfahren der MP3-Kompression). Die Autoren entwickeln daher ein Verfahren zur Verarbeitung von Sprachsignalen für einen Hidden-Markov-Spracherkenner nach dem Modell (als Simulation) der menschlichen Gehörperipherie. "Das auditorische Modell erzeugt eine interne Repräsentation des Sprachschalls, welche dem nachgeschalteten Erkenner als Eingang dient."[16] "Hidden-Markov-Modelle (HMM) haben

13 J. C. R., Licklider, Man-Computer Symbiosis, in: IRE Transactions on Human Factors in Electronics. HFE-1 (März 1960) No. 1, 4-10

14 See Fig. 3 "Relative Übereinstimmung der Ausgangsinformation (AI) und der Vergleichsinformation (VI) des Stimmspektrums einer Person, die über Telefon gesprochen hatte <...>. Zeitspanne der Aufnahme der AI und der VI 12 Wochen", in: Christian Koristka, Die Verwendung der menschlichen Stimme zur Identifizierung einer Person, in: Forum der Kriminalistik 1 (1965), Heft 3, 32-36 (34)

15 Gerhard van der Giet / Hermann J. Künzel (Bundeskriminalamt Wiesbaden) "Rechnergestützter Stimmenvergleich für forensische Anwendungen", in: Kriminalistik 9/81, 341-346 (345)

16 J. Tchorz / T. Dau / B. Kollmeier, Gehörtgerechte Signalvorverarbeitung zur robusten Spracherkennung in Störgeräuschen, in: Dieter Mehnert, Historische Schallaufnahmen. Das Lautarchiv an der Humboldt-Universität zu Berlin, in: ders. (Hg.), Elektronische Sprachsignalverarbeitung, Dresden (Ges. für Signalverarbeitung und Mustererkennung) 1996, 46-51 (46, *abstract*)

das Ziel, ein statistisches Modell zur Beschreibung der Auftretenswarscheinlichkeit von Vektorfolgen aufzubauen und seine Parameter zu bestimmen."[17]]

[Esp. in the case of radio archives, the tapes preserved by the sender significantly differ in quality from the sound quality actually received (and occasionally recorded on magnetic tape by amateurs) in concrete radio sets, esp. in the case of international radio via short wave transmission.]

The focus on the message of sound archives as technical medium significantly differs from the focus on the cultural "content" of such sound recordings. Is the historical reading an "othering" (Vivian Sobchack[18]) or even obscuration of the arte-factual sound archive by the discourse of political correctness?

[A document from 1988 in file MfS OTS No. 1635, differentiates between "Auditive und messanalytische Parameter zur Sprecherklassifikation" - which is human *listening* (performance) vs. non-human signal recognition (operative); another file <MfS Abt. 26 Nr. 790: Bl. 38 / 39> explicitely differentiates between "Höranalyse" (human listening) and "Meßanalyse" (machine measuring of the physcial articulation) by sonography; both methods are explicitely put into a complementary (if not even "dialectic") relation <c Bl. 50 / Bl. 51 Hör- und Meßanalyse / Sonagraph>. In addition to this channel of transmission, possible technical noise sources are taken into account <c Bl. 45 / 46 mgl. technische Störquellen>.]

[Although the aim of any data ontology is objectivity, "it is important to underscore that a human listener decided what to index and what not to index; a human listener decided what indexing term to use and what indexing term not to use; and a human listener decided if a given narrative segment could be described by a keyword or not. This is a fundamentally interpretative process that must necessarily remove some of the potentialities of the narrative in the application of the data ontology. In the end, it has the effect of turning the narrative into data." <Presner 2012, 17>]

„Listening" *without* the human ear leads to a different kind of sonic hermeneutics, a differnt kind of understanding. Machines do not "hear" language but measure audio signals - which is its weakness and strenght. They have a different insight into sound and voice recording, exactly because they do not "listen" but radically analyze.

This is in full accordance with Mara Mill's findings on previous physiognomic uses of sound recordings which resulted in the development of machinic speech recognition.

---

17 Ch.-M. Westendorf, Experimentelle Sprachverarbeitung mit signaLab, in: Mehnert (Hg.) 1996: 76-85 (82)

18 Sobchack 2011, "Afterword"

[*Vocal Features: From Stimmphysiognomik to Speech Recognition by Machine. The physiognomic uses of sound recordings between 1926 and 1953*. Abstract: "Collections of phonographic 'vocal portraits' - such as the recordings of 'criminal' speech in Berlin's Lautarchiv - prompted investigations into the features of the individual voice, for the purposes of characterology, lie detection, and speaker identification. Visual recordings of speech (oscillograms and spectrograms) were at first applied to the same purposes. Voiceprint identification" was eventually abandoned as hopelessly inexact, but the generic speech features described in the era of Stimmphysiognomik subsequently enabled the beginnings of speech recognition by machine."]

Ironically, machine-based voice surveillance results in tools which can now be applied for scientific research of sound archival heritage as well. Surveillance and research are two sides of one algorithmic coin in audio signal processing.

Historians tend to read archival files on the application of speech analysis tools by former GDR State Security in their political context which has been a totalitarian state. While our conference *Listening to the Archive* is sub-titled "Histories of Sound Data in the Humanities and Sciences", let me replace "histories" by archaeology.

<§ "the sonic archive": XXX>

The decipherment of audio records not as „historical" documents but as sonic monuments with media-archaeological interest rather asks: To what degree did (and still does) surveillance paranoia result in analytic technologies which actually create knowledge?

[According to Foucault, paranoid power is productice for producing knowlege.]

The MfS "Speech archive" project raises the question: To what degree does inhuman listening provide insight which is otherwise hidden by sympathetic human listening to the archival voices? Paranoia, as we learned by Michel Foucault, does not only lead to a suppressive power regime, but is productive for knowledge as well.

So why not use the clever algorithms applied by MfS or NSA for voice recognition since the 1950 on computational basis for cultural analysis? Or does this rather degrade "digital" humanities to a second-hand justification and ornament of non-academic practices? Would it be naive to apply algorithmic tools for scientific research without being aware that such tools have been developed for forensic, surveillance or military application (as "Heeresgerät" in the Kittlerean sense)? Is there a "good" cultural use of "evil media"?

"Cold listening" to sensitive sound archives

Such "cold listening" that uses speech features automatically extracted by computer algorithms, today, results in experimental sonifications of large sound collections.

[For a computer, a sound file is only an array of binaries. A computer uses “low-level” information to "interpret" the sound. A semantic gap opens when it comes to "understand" the musicality of such audio signals. Trying to “close the semantic gap” <...> is one of the motivations for using multiple features", such as detecting similarities in sound files. We learn the formal way of "hearing" from algorithms whicih identify what a sound represents.]

Lev Manovich's software mining of big data explicitely adapts Franco Moretti's approach to hundreds of literary texts from past centuries, requiring "distant reading". What Manovich has developed in his so-caled "cultural analytics", though, is focused on visualization.[19]

[Versinnlichung hoch-dimensionaler Daten heißt nicht automatisch Visualisierung, sondern verlangt ebenso nach sonischer Algorhythmisierung i. S. Miyazakis]

This may be modified for acoustic space, to "distant hearing". Most digital analysis of files from sound archives is still restricted to one piece, but large-scale digitisation projects open the option for "big sonic data" analysis.

[cp. application Tkaczyk *Sound Archives and Auditory Memories* at the Netherlands Organisation for Scientific Research and the Council for the Humanties]

The true impact of digitisation is the message of algorithmic computing as medium. The application of sonic data mining un-covers hidden, archivally implicit knowledge which musicological research would not even ask for. Mathematical intelligence of algorithms serves for developing new strategies of audio archival findings like so-called "deep machine learning" - a set of algorithms which uses a deep graph with multiple processing layers for automated speaker and speech recognition.[20] Whith the "deep" suggesting a stratigraphic layers, it is still not an archaeological metaphor but a pointer to the radical way of *operative* Archaeology of Knowledge (Foucault) which is the implementation of advanced mathematics in high-frequency computing

19 See Lev Manovich, Data Science and Computational Art History, in: International Journal for Digital Art History, no. 1 (2015), 12-35

20 As deeply described in the Wikipedia entry *online* (accessed July 19th, 2016)

power: media archaeology as techno-mathematical alliance.

While musicological understanding re-creates cultural memories which are clearly addressed to the human ear, algorithmic sonic data mining provides insight into *implicit sonicity*.

[*Archivological* analysis is using such software tools not for individual sound recordings but for a large, trans-individual array of sound files.]

[In the 20th century, listeners to phonographic archives typically compared small numbers of sound records, "and the use of our human cognitive capacities unaided by machines was considered to be sufficient. Even if the number of phonographic and gramophonic records in sound collections like the Lautarchiv looks impressive and therefore has been rather individually researched so far[21], it is small when compared to born-digital sound files around the globe today. This has resulted in the development of a kind of mass-statistical ear. In social web portals like YouTube, tens of thousands of files with sonic expression are born digital as user generated content which poses a different challenge to individual human research by its shere scale. Even "digitized collections of historical artifacts can be also quite large" <Manovich 2015: 33>.]

[In a paraphrase of Lev Manovich's term for computer-based "cultural analytics", there is now the option to apply computational, which is: algorithmic intelligence ("intelligence" in both its meanings). "Listening to the digitized sound archive" in contemporary culture not only offers but even requires the usage of information science.]

But at that point of analysis, "algorithmic criticism" is required.[22] The real sound archive is not its recordings but the *archive* in Foucault's sense: the underlying algorithms of digital audio signal processing tools themselves.

Key methods applied here are core opeations of data science such as feature extraction, measuring distance in feature time, and dimension reduction.

Oral testimony from Holocaust survivors (like the Yale Archive), out of necessity (the material deterioration of the magnetic tapes), has been digitized. The audio tracks can be extacted and used as data bank for new algorithmic creation of so far hidden knowledge. A proposal by Amit Pinchevski, media scholar at the Department of Communication of Hebrew University in Jerusalem, calls for an experimental data mining of Holocaust testimony. If all these voices are thrown - contrary to their hyper-

21 See Nepomuk Riva (ed.), Klänge aus der Vergangenheit, xxx

22 See Stephen Ramsay, Reading Machines. Towards an Algorithmic Criticism, Urbana, Chic. (Univ. of Illinois Pr.) 2011, Kap. 1 "An Algorithmic Criticism", 1-17

individualization - into one data pool and tumbled algorithmically, step by step, out of the phonetic chaos patterns will emerge like in a minimal music composition by Steve Reich (phase shifting / "phrase shifting"): syntactic formulars and repetitive expressions. It will become apparant to what degree the oral history interview dispositive already creates narrative conventions which are sublimanlly at work even if the actual recollection is the most dramatic individual experience.

[Audio: Reich-Piano-Phase.mp3, Wergo]

There is a specific reason why I refer to Steve Reich's musical composition *Piano Phase*,

since the same composer created the piece *Different Trains* where he correlated phonetic parts of oral testimony of former American train porters to "different" train sounds before, during, and after World War II.

[Audio: Reich-Trains.mp3, 3x]

<passage in xxx VOICE-JERUSALEM>

<see KITTLER-KOV-CARLE.wav; video project Herwig Weiser>

<see Fig. Lautarchiv-Oszilloskopie-Brandl-1928>

<see Fig. LAUTANALYSE-PLANCK>

Todd Presner <2012: 26> comments on the idea of "social tagging" of digitally accessible online survivor testimonies: "Narratives would be heard and echoed in their polyphony, with some listeners hearing some things and others hearing quite different things."

What if that "other listener" is non-human, in fact: a techno-mathematical ear?

[cp. as visual equivalent <copy ALGORITHM-ETHICS-PRESNER> the "computer-generated visualizations of a portion of the history of the Holocaust: the Digital Monument to the Jewish Community in the Netherlands.[23] "The monument has no physical or built counterpart; it only exists on the web. It is a digital image consisting of about 831,432 colored pixels. Each little box of pixels represents a single person, and they vary in size according to the age of the victim: Tall blue bars are adult men; tall red bars are adult women (these are 4x1 pixels); half-length green bars are boys 6-21; half length yellow bars are girls 6-21 (2x1 pixels); and light blue and pink represent children (1x1 pixels). The monument is a raster graphic, or bitmap, which is comprised of a rectangular grid of pixels viewable in a web-browser on a computer monitor. The graphic represents the nearly 100,000 Dutch Jews who were killed <...>. Clicking on an individual color box brings a viewer

23 http://www.joodsmonument.nl/?lang=en

to a webpage containing information about the victims, including their names, dates of birth and death (if known), place of birth, and family members, including information about whether they survived the war or not."]

["The graphic organization of the monument is based on the alphabetical order of the place of residence of the victims when they were deported."]

[Such an active computer-graphical diagramm might be developed (in the Signal Laboratory) for opening up the Phonogrammarchiv and the Lautarchiv to so-called Digital Humanities - even at the risk that this might end up in algorithmic inhumanity. What if not humans but algorithms "listen" to the recorded voices? Does the "cold ear" of media archaeological *listening to the archive* correspond with the telephone voice surveillance practice of former GDR State Security? Here a critical question arises. Does the active application of "digital humanities" tools to audio collections like the Lautarchiv with recordings from World War I, or to oral history projects of former Jewish prisoners in German camps immediately after World War II such as David Boder's wire recorded narratives in Displaced Persons Camps in 1946[24]

[or the Yale Video Archive for Holocaust Testimonies. Videotaping Holocaust survivors here began in 1979, later named the Fortunoff Archive, resulting in today more than 4400 testimonies and some ten thousand hours of video]

or finally, the "Stimmarchiv" of the State Security of former GDR, miss the ethically "sensitive" issues of such archives?]

[What happens if a collection of individual voices (be it the voice of Max Planck in the Lautarchiv,

<Fig.: Lautanalyse-Planck.jpg (Braguinski)>

or Holocaust survivors) is transformed into big data? Is the ethic momentum of historically "sensitive archives" lost by sonic data mining?]

The original recording, archiving, and dissemination of video testimony may have been widely defined through an ethics of listening. With the listener being replaced by software for audio data mining, does this ethical dimension get lost in favor of a non-human analysis of testimony?

[One of the leading critics of Digital Humanities, Todd Presner, has delivered a paper "The Ethics of the Algorithm: Close and Distant Listening to the Shoah Foundation Visual History Archive"[25].]

---

24 See Alan Rosen, The Wonder of Their Voices: The 1946 Holocaust Interviews of David Boder, Oxford (Oxford UP) 2010
25 Draft March 2012; *online* http://www.toddpresner.com/wp-

Turning speech evidence into data is the opposite of the use of figurative description "adding to" the factual reality of the audio signal as acoustic event; data analytics rather "subtracts from" or "abstracts of" the narrative as told by the human testimonies.

"In other words, what goes missing in the 'pursued objectivity' of the database is narrativity itself: from the dialogical emplotment of the events in sentences, phrases, and words <...> to the tone, rhythm, and cadence of the voice, to the physical gestures, emotive qualities, and even the face itself. Of course, this is because databases are not narratives or people telling narratives; instead, they are formed from data (such as keywords) arranged in relational tables which can be queried, sorted, and viewed in relation to tables of other data. [Database relationships are foremost paradigmatic or associative]" <Presner 2012, 17>.

Therefore algorithmic analytics shall not aim at creating relational data banks but rather signal correlation which pays attention to the subliminal tonal gestures.

"Cantometrics"

- critique of (re-)visualizing the sonic archive by tools like "Sonic Visualizer", e. g. in computational ethno-musicology

<L> George Tzanetakis et al., Computational Ethnomusicology, in: Journal of Interdisciplinary Music Studies, Fall 2007, vol. 1, issue 2, 1-24

<begin copy § "Tonorganisation. Fallstudie Lautarchiv und Alan Lomax "in MEDPHIL-HUMAN>

- <L> Svec, Lomax' Deep Digitality

<see § "Alan Lomax' "digital humanities" of folk song archives": DIGITALE-ZEITZEUGENSCHAFT-FORSCHUNG>

<see DATADRIVE / Braguinski>

- At Humboldt University, Rainer Kluge wrote his thesis (Dissertation B) on *"Faktorenanalytische Typenbestimmung" als Beispiel des Rechnereinsatzes in der systematischen Musikwissenschaft*.

<begin copy DATADRIVE>

Automated feature description as a media-archaeological tool is close to the methods of Systematic Musicology:

---

content/uploads/2012/09/Presner_Ethics.pdf

"<...> large electronic collections of music in a symbolically-encoded form <...> have enabled music researchers to develop and test <...> empirical theories of music on large data sets." The availability of such music data "creates a new perspective for Systematic Musicology, which <...> often sets out to explain or describe music through the induction of empirical laws, regularities or statistical correlations in relation to music objects or music related behaviour"[26]

<such as the servo-motoric feedback between the *guslar* singer and his string instrument, the *gusla*>

For the M^{4}S project (Modelling Music Memory and the Perception of Melodic Similarity), hosted by the Computing Department of Goldsmith's College (University of London), the term "symbolically-encoded music" means

"<...> music in a computer-readable format where the fundamental unit of representation is the note." <Müllensiefen / Wiggins / Lewis ebd., 133>

"Symbolic formats can be contrasted with audio formats which, instead of capturing notes explicitly, encode the sonic **<!>** aspect of a musical performance by representing sound as a complex waveform. The best known formats are audio CD, the WAV and AIFF formats used primarily in computers and iPods, and MPEG-1 Audio-Layer 3 (mp3) as a compression format used for web-based and portable applications." <Müllensiefen / Wiggins / Lewis ebd., 133>

The "semantic" listening (concentrating on *musical objects* like a melody) makes the difference to the media-archaeological "(h)ear(ing)" which focuses on the *sonic object*. Whereby a melody is basically a contour kept and recognized in memory "over time" (in both senses), the time-critical approach of media archaeology rathers concentrates on non-harmonic micro-figurations of temporality within the sonic event. Thus special algorithms are needed which identify such temporal qualities (such as dynamic time warping[27]), and efficient algorithms "for extracting the repetitive structure of an audio recording"[28].

<...>

"The most ambitious corpus-based musicology project was one based

---

26 Daniel Müllensiefen / Geraint Wiggins / David Lewis, High-level feature descriptors and corpus-based musicology: Techniques for modelling music cognition, in: Systematic and Comparative Musicology: Concepts, Methods, Findings, hg. v. Albrecht Schneider, Frankfurt am Main u. a. (Peter Lang) 2008 (= Hamburger Jahrbuch für Musikwissenschaft 24), 133-153

27 See Müller 2007: 69

28 Meinard Müller, Information Retrieval for Music and Motion, Berlin / Heidelberg / New York (Springer) 2007, 165

in Princeton University concerned with Josquin scholarship. From 1963 to the beginning of the eighties, researchers, led by Arthur Mendel and Lewis Lockwood, generated electronic scholarly editions of the complete works of Josquin <...>, many including concordances, and relevant related works. From this, statistics for cadential progressions and modal indicators were compiled and subjected to statistical analysis primarily in order to study issues of authorship and stemmatic filiation (see <...> various papers in Computers in the Humanities between 1969 and 1978). The ambitions of this project, though great, never extended to revealing cognitive processes, being limited, essentially, to style analysis."[29]

"In folk music research, feature extraction and the use of computers have been employed as a means for the (automatic) classification of songs (mainly melodies) according to their musical characteristics. In a comprehensive study Steinbeck (1982) classified European folk melodies into six homogeneous groups by employing Ward's classification algorithm with 35 relatively simple features derived from the monophonic melodies. He was able to show that this classification was in close correspondence with the melodies' regional origin and functional uses."[30]

At the borderlines of the *semantic gap*, to most musicologists

"the evaluation of musical relationships is not a task amenable to automation. The quantification discussed above is a statistical one and, whilst its usefulness will be greater as more information is provided to the system, it is cognitive experiment and musicological reasoning that must prove the final arbiter of the system's performance. Furthermore, such an approach can only offer limited assistance to those wishing to perform detailed analyses of single works—which is the standard paradigm in traditional music analysis."[31]

<see § "Case Study: Kurenniemi's audio-diary, re-activated by Constant": MEDARC-CCA>

"The power of the quantitative <...> lies in the new kinds of access to interpretation it provides." *Distant hearing* (as slighty modified in respect to Fitzpatrick's review of Moretti's approach) inevitably raises the question not of *whether* one ought to hear distantly, but of *what* one can hear *only* distantly, and what one requires closeness in order to capture.

<comparable to the media-archaeological gaze is Friedrich Nietzsche's notion of "passionate distance" ("Pathos der Distanz")>

<end copy MEDPHIL-HUMAN>

---

29 Muellensiefen et al.: 136
30 Muellensiefen et al.: 136
31 Muellensiefen et al.: 139

- but while such an analysis focuses on an individual song recording (which can be "big data" in itself, for hour-long oral poetry performances like the *guslari* performances in Bosnia and Montenegro in the tradition of Homeric epics,

Fig.: VISUALISIERUNG-BICIRBEY-BRAGUINSKI

stochastic signal analysis covers "big data" of a whole sound archive.

- a non-historicist approach to sound archives takes its point of departure from the recorded signals themselves which stay invariant towards the passing and changing historical and cultural contexts happening in the temporal progress around the durable material record. Cultural contexts change, while the physical laws of speech articulation and recording change according to a different temporal rhythm which is not historical time.

- Sound archives are "sensitive" not only in ethical and other meaning but in being non-symbolical, non-alphabetical: signals are the most indexical "sensitive" traces of physical, here: acoustic events.

- A new option of audio signal mining opens when the term "digitizing" analogue sound carriers is not simply understood as an act of saving recordings and making them able for "open access". The real opening of access is the use of audio data in terms of mathematical intelligence.

(Here the world-play of "surveillance" and "research" starts)

The digitization of analogue sound signals is not a re-alphabetization of music, but its mathematization. Different from classical computational linguistics, algorithms themselves become the non-human agencies of knowledge. This starts with simple analytic tools like the "Silence Finder" in the free audio software Audacity.

In algorithmic techno-memory practice, the "Silence finder automatically, that is: algorithmically, taggs intentional and non-intentional pauses in speech or sound files.

<see below, sw "Silence Finder">

Even if this tools has been developed for sound editing and opens the option of "remove" the moments of silence, it can be used as a research tool to identify in large data banks the moments when enunciation hesitates - for reasons which then require hermeneutic, i. e. human, context-intense interpretation.

Is there a "sound of the archive"?

"Listening" to the archive?

<see auch MEDARC-CCA, on "archive" / "sound">

There is a discursive (cultural historical) listening vs. the audio-listening, as is evident from the present conference papers which fall into two groups: one performs "close listening" to the materiality of both sound and its recording technologies, and the other de-centers such non-discursive practices by widening the scope of analysis in cultural and historical (text-based) contextualisation. There are "two bodies" of the archives here: the real sound archive, and the historical archive of which the sound archive is not an active agency but an object of research.

<begin copy HISTORY-OR-RESONANCE-JVC>

Different from the archive as symbolic order composed of records in historiographic, that is: in alphabetic notation, there is an para-archival modality of sub-textual, signal-based recording of the past: the sound of times past. The BBC World Service has launched the "Save our Sounds" project, looking to "archivize" sounds that may soon be lost due to the post-industrial world.[32]

The traditional sonic experience in real archives is silence; historical imagination (as expressed by the Romantic French historian Jules Michelet[33]) though hallucinates the voices of the dead here. The media-archaeological ear, on the contrary, is trained to actually stand archival silence, gaps, voids.

But silence itself can become part of the archive. The software for sound analysis *Audacity* actually provides an algorithm called "Silence Finder". The shere endurance of periodic frequencies is a Bergsonean time which passes. While an empty space within a painting positively endures with time, silence in acoustics is always a temporal (though negative) event itself which I call its "sonicity". This term reminds of the fact that explicit sound is just a thin slice of a wider spectrum which is audible to humans. But below and beyond this phenomenological range, sonicity mediua-epistemologically refers to implicit sound as an object of knowledge about temporal forms of the vibrational event, to time signals as such.[34]

32 But caution, this is not an archive: As long as an algorithm is missing which rules the transition of sound provenience to permanent storage, it is just an idiosyncratic random collection.

33 See W. E., Das Rumoren der Archive, Berlin (Merve) 2002

34 For a related approach see Steve Goodman, The Ontology of Vibrational Force, in: same author, Sonic Warfare. Sound, Affect and the Ecology of Fear, Cambridge, Mass. (MIT Press) 2009, 81 ff., and Peter Price, Resonance. Philosophy for Sonic Art, New York / Dresden (Atropos Press) 2011

Historians remind us that there is no unmediated access to the past. But in the negative sound of the archive, its silence, we listen to the past in its truest articulation. Let us pay respect to absence instead of converting it into the specters of a false memory. Written records or printed texts necessarily miss sound matter. But in the deeper sense there is implicit sonicity even in images, diagrams and graphs which are derived from sound sources; any sonagram keeps an indexical relation to the sonic event.

There is sound even from the digital archive. When an ancient "Datassette" is being loaded from external tape memory into the ROM of a Commodore 64 computer, we are actually listening to data music. What we hear is not sound as memory content like an old percussion-assisted song[35], but rather the sound of computer memory itself, that is: a software program which is "scripture" (though in the alphanumeric mode). We are listening to the data archive which is not sonic memory but sonicity.

<end copy HISTORY-OR-RESONANCE-JVS>

<see MEDMEM-OSLO>

Audio-recordings and their media-archaeological understanding

<begin copy HISTORY-OR-RESONANCE-JVC>

Dis-covering the temporal implications (rather than metaphorical "layers") of the archive is not just an operation of the mind or the eyes, but of hearing and literally archival "understanding" as well (German *verstehen* refers to auditory as well as to cognitive perception).

The spatial, text-based archive is familiar as a radically silent place. Acoustically, this silence might be re-interpreted as an enduring negation of time-based sound, as performed in John Cage's piece *4'33*.[36] Whereas the-classical archive which is based on alphabetic scripture is a static array of records (like parchments and papers) on the grand scale and letters on the microscale, which can be set into motion only by the act of human reading line by line, the Edison phonograph at first glance looks like the first form of "archive in motion", since its "records" (notably the early ethnographic field recordings around 1900, leading to the Vienna Phonograph Archive and the Berlin Phonogramm Archive) are based on a continuously rotating, technically moving apparatus

35 For an analysis of the interplay between technical memory and affective remembrance see Ben Anderson, Recorded music and practices of remembering, in: Social and Cultural Geography, vol. 5, No. 1, March 2004, 3-19

36 On the occasions which led to this composition see Seth Kim-Cohen, In the Blink of an Ear, New York (Continuum) 2009, 160ff

both in recording and in re-play.

Strictly speaking, the phonographic record which consists of intinitesimally continuous signals instead of alphabetic or other elementary symbols is no "archive" at all - with the archive being both composed of and itself representing the symbolical order of discrete elements (letters on the lower level, archival tectonics on the upper oganizational level). Phonographic inscription is different from cinematographical recording and projection of visual momevent which is based in discrete, mechanically interrupted frames.

When we listen to an ancient phonographic record, *the audible past* (alluding to Jonathan Sterne's book title) very often refers rather to the noise of the recording device (the ancient wax cylinder) than the recorded voice or music. Here, the medium talks both on the level of enunciation and of reference. What do we hear most: the cultural content (the formerly recorded songs) or the medium message such as limitations in vocal bandwith, even noise (the wax cylinder scratch and groove)?

With digital sampling and processing of audio-signals, analog noise is usually significantly filtered, thus: silenced. But the former noise is being replaced by an even more endangering challenge: the "quantizing noise" on the very bit-critical (technical) level of signal sampling, and the migration problems of digital media data and the physical vulnerability of electronic storage media in terms of institutional (cultural) sound tradition. This is not just a technical question, it has an epistemological dimension as well.[37]

<end copy HISTORY-OR-RESONANCE-JVS>

37 See Arild Fetveit, Medium-Specific Noise, in: Liv Hausken (ed.), 189-215