Universitat Oberta de Catalunya

artnodes
E-JOURNAL ON ART, SCIENCE AND TECHNOLOGY

https://artnodes.uoc.edu

ARTICLE

NODE "MATERIOLOGY AND VARIANTOLOGY: INVITATION TO DIALOGUE"

How to problematize a border, or: some ongoing notes on transient methodologies

Pedro J. S. Vieira de Oliveira
Independent scholar and artist

Date of submission: February 2024
Accepted in: June 2024
Published in: July 2024

Recommended citation

Vieira de Oliveira, Pedro J. S. 2024. «How to problematize a border, or: some ongoing notes on transient methodologies». In: Siegfried Zielinski and Daniel Irrgang (coords.). Node «Materiology and Variantology: invitation to dialogue». *Artnodes*, no. 34. UOC. [Accessed: dd/mm/yy]. https://doi.org/10.7238/artnodes.v0i34.424935

Abstract

In this paper I rehearse a method for studying media that is engaged by what I am calling a *transient methodology*. This work with "transients" emerges as much as an analytical tool as it does as a material engagement, in direct connection with my encounter with a GDR-era synthesizer prototype, the Subharchord I/II, and its use of a specific audio filter design which is found again only forty years later as a building block for most state-of-the-art speech recognition algorithms. I am particularly interested in one such algorithm, namely the so-called "Dialect Recognition Software", a proprietary machine learning solution in use since 2017 by the German Office for Migration and Refugees in cases of undocumented asylum seekers in the country. By bringing these two stories closer and mobilizing them through the work of transients – a concept that works in sound as much as it does in movement – I seek to activate an experimental, speculative media archaeology that focuses on the haptics of listening. Transient methodologies, I argue, offer the potential to critically listen to these two technologies together while at the same time slowing them down for critical inquiry and analysis.

Keywords

machine listening; sound studies; migration studies; media archaeology

Cómo problematizar una frontera o: algunas notas continuas sobre metodologías transitorias

Resumen

En este artículo ensayo un método para estudiar los media que están comprometidos con lo que llamo una metodología transitoria. *Este trabajo con «transitorios» surge tanto como una herramienta analítica como un compromiso de material, en conexión directa con mi encuentro con un sintetizador prototipo de la era de la RDA, el Subharchord I/II, y su uso de un diseño de filtro de audio específico que se encuentra de nuevo, solo cuarenta años después, como componente básico para la mayoría de algoritmos de reconocimiento de voz de última generación. Estoy especialmente interesado en uno de estos algoritmos, a saber, el llamado «Dialect Recognition Software», una solución de aprendizaje automático patentada, en uso desde 2017 por la Oficina Alemana para la Migración y los Refugiados en casos de solicitantes de asilo indocumentados en el país. Al acercar estas dos historias y movilizarlas a través del trabajo de los transitorios, un concepto que funciona tanto en sonido como en movimiento, busco activar una arqueología de medios experimental y especulativa que se centre en los hápticos de la escucha. Digo que las metodologías transitorias ofrecen el potencial de escuchar de forma crítica estas dos tecnologías juntas, al tiempo que las ralentizan para la consulta y el análisis críticos.*

Palabras clave

escucha automática; estudios de sonido; estudios de migración; arqueología de medios

Note:

A (temporary) link to an audio file related to the submission is here: https://soundcloud.com/ppedrooliveiraa/transients-hkw-mixdown/s-eHLDBR9zWEq

I

In 1956, the East German *Rundfunk- und Fernsehtechnisches Zentralamt* (Central Technical Office for Radio and TV, from here on RFZ) was granted funding to establish a state-of-the-art experimental Electronic Music Studio in Adlershof, Berlin. Gerhard Steinke and his newly founded *Labor für akustisch-musikalisch Grenzprobleme* (Laboratory for the boundary problems of acoustics and music) started to investigate "issues of electronic sonic production" (Steinke 1993, 16),[1] in close connection with the desire to advance socialist content through avant-garde artistic and sonic cultural production. The Labour's program was part of not only the development of media entertainment for the GDR public but also to support, conceptualize and produce what could become "*Wege zu einer neuen Klangkunst*" – pathways towards a new Sound Art (Steinke 2015).

Among the many of their developments between 1956 and 1970 was the design, construction and mass production of an electronic organ, inspired by earlier German devices such as Oskar Sala's Trautonium and the Siemens E-Studio Hohnerola from Munich. The Subharchord, designed by Ernst Schreiber and Steinke, was prototyped and produced between 1961 and 1965 as a "small series of six or seven units" (Steinke 2007) that were then sold to studios in Germany, Norway, Poland, the USSR and (then) Czechoslovakia for a price of around 30000 German/East German Marks (Miersch 2003, 120).[2] Much like its aforementioned predecessors, the Subharchord from the RFZ gave composers the possibility of working with the "subharmonic series" – musical tones that are divisions rather than multiplications of a fundamental frequency. This provided a different way to deal with the synthesis of sounds that would not mimic that of an acoustic instrument, thus producing tones that would be unfamiliar to the ear.[3]

Differently from its aforementioned predecessors, however, these subharmonic frequencies could be then further sculpted by a 14-band filterbank[4] specially tuned to the Mel Scale, a psychoacoustic concept from 1940 that works with the assumption that differences in pitch perception cannot be properly mapped to differences expressed in Hertz (Stevens & Volkmann 1940). Basing itself on the functionalities of the human inner ear as well as the subjective perception of listening, the

1. in the original, *Probleme der elektronischen Klangerzeugung.*
2. Manfred Miersch (2003, 120) speaks about seven to eight units instead.
3. The GDR was initially suspicious of the subharmonic series. In letters exchanged between Steinke and DDR officials between 1954 and 1956, they called it a "fiction". (Steinke 1993, 17)
4. Common audio filters such as lowpass, high-pass, or bandpass often contain an adjustable center frequency which determines the scope and strength of attenuation of filtering. A filterbank on the other hand, as the name itself implies, is an audio filter in which multiple frequency bands (ranges of frequencies, for instance from 120 to 240 Hertz, 320 to 640 Hertz and so on) can be simultaneously attenuated, emphasized or removed from the signal.

Mel-Scale offered a supposedly more harmonically pleasant spectral cut. Accordingly, then the subharmonic divisions of the Subharchord were split into these tuned frequency bands, giving the composer the ability to play with their activation or not at will, thus emphasizing timbral playing (*Klangfarbenspiel*) rather than melodic progression.

II

To be able to recognize speech, an algorithm must be modeled in such a way so that it can characterize the mathematical vectors that define an input signal to be speech against one that is not. This is one of the most fundamental steps of speech recognition, one with which any understanding of how machines "listen" must begin. The question here is not at all of a philosophical nature but instead material: what are the acoustic features, as in vibrational components, that make up a signal so that it can be properly read and processed as speech? Most contemporary speech recognition algorithms respond to that question by dealing with how power is contained and distributed across a given signal. An algorithm performs an analysis of a snapshot captured by a recording device, a snapshot that reproduces the path of airflow through a tube as well as the folds, nooks and cranes such tube can be shaped into to articulate a specific distribution of peaks. That is to say: the vocal tract, when regarded as a computable, mathematical model, can be measured and probed into not at the source but at the *event* of utterance, be it live or recorded. This snapshot, usually far less than a second long, is understood by data scientists to contain the entirety of acoustic information of the given signal.

One next possible step within a speech recognition algorithm, that is, after being able to recognize speech as speech, is to determine who is speaking. This process can take many shapes and might lead to different modes of determination, for instance, maybe not the legal identity of the person speaking (as, say, a fingerprint would do), but rather their ethnicity, gender, age, even height or health status (Li & Mills 2019, Kanervisto *et al.* 2014). Our snapshot is now treated as a harmonically rich signal that we already know to be speech. The next step is to decompose it according to the harmonic information it contains; to do so, it is split into a series of up to thirteen logarithmic power vectors, following the divisions of the Mel Scale. If we stay with harmonics, these power vectors act as a filter. In this way, we can call this component of an algorithm for recognizing speech a Mel filterbank.

A Mel filterbank in the digital domain, that is, when employed in speech recognition, computes a series of coefficients to achieve "robust acoustic feature extraction." This mode of vectorizing data, done so logarithmically, is called Mel-Frequency Cepstral Coefficients (MFCCs).[5] The Mel filterbank helps the algorithm sort out the intensity with which frequencies are distributed in listening. The Mel-scale provides a more accurate measurement device for the perception of differences in pitch, that is, differences between frequencies, and as such not only models the vocal tract of a specific speaker or class of speakers (Dusan & Deng 2000; Nuance 2015) but does so by modeling and at the same time employing a "universal" human ear. It seems, then, that implied in the question "who is speaking?" is not so much the speaker but the ear that listens. The borders between what is said and what is understood – by humans or machines – seem to fall into the traps of an assumed universality in the event of speech before recognition – or at its expense.

III

Nuance Communications Inc. brands itself as playing "a foundational role in the emergence of conversational AI." (Nuance nd.) Among the solutions offered by Nuance is "voice biometry", which can take the shape of "voice passwords," as well as "speaker recognition." (Nuance 2015, 2-4) They claim their machine learning-based model to be particularly successful due to its reliance on analyzing and determining the "unique physical traits of the [speaker's] vocal tract" (ibid., 2). Based in the United States, Nuance Communications is the contractor responsible for developing the so-called "Dialect Recognition Software" (*Dialekterkennung*, sometimes also *Stimmbiometrie* or *Sprachbiometrie*, or DIAS) for the *Bundesamt für Migration und Flüchtlinge* – the German Federal Office for Migration and Refugees (BAMF). The DIAS Software was developed in 2017 by request of the BAMF, and since 2020 it has been working as a license (BAMF 2022, 13). The databases that trained their software, mostly Arabic variants, were initially purchased from the Linguistic Data Consortium in the University of Pennsylvania; according to a report from 2020, the software also began using data acquired from BAMF's own recordings of interviews, hearings and tests as a way to decrease the error margin of the software and optimize its use for adjacent languages such as Urdu, Pashto and Farsi (ibid.).

Since April 2017, the DIAS machine learning solution has been deployed in more than 40000 cases of asylum seekers without the right papers. The software is used to determine the probabilities of otherwise "undocumented" applicants of being citizens from the country they claim they are from, thus establishing a "truth" conveyed – and oftentimes negated – by the applicant's own voice. The error margin as of this writing is of 15% for Levantine Arabic and up to 27% for the three aforementioned adjacent languages (ibid. 13-14).

(interlude)

"Du weißt, es ist das Heiligtum"[6]

This is what the Studio director says to me as he opens the door. Yet I am blissfully unaware of what stands before me, or what I am

5. While a technical explanation of MFCCs is way beyond the scope of this text, suffices to say that MFCCs figure as the most effective, reproducible and scalable method for dialect recognition algorithms, in particular, those tuned for Arabic dialect variants (see e.g. Ma & Fokoué 2014; Ali et al. 2015; Hanani *et al.* 2015).
6. "You know, this is the sanctuary."

about to encounter. I must say from the start that it honestly felt to me like looking at an odd, somewhat clunky prop from a 1960s sci-fi movie. Square-ish, stiff, functional-modernist design and in a flat shade of blue bearing a panel full of wooden-colored knobs. It indeed drew my attention, but more for its weirdness than anything else. Long rows of black switches adorned the entire top and bottom width of the front panel; they looked to me rather uninteresting visually but perhaps quite pleasant in terms of their haptics. It is not until I look closer at the bottom row of black switches – not least because they were all perfectly aligned – that something changes: a "Mel-Filter".

Immediately, I was struck by that little piece of information engraved in the metal plate, and all my disinterest in it turned upside down. Something had shifted dramatically in a matter of a millisecond, a coalescing of a number of years into that otherwise passing moment of weird nostalgia display, media-archeological prowess or simply curiosity. I was looking for it without knowing it; like many things with my research, it actually found me. From that brief encounter – of which I initially said nothing – an entirely new thread, or a wormhole maybe, opened up in my work. A wonderfully chaotic detour of juxtapositions, historical (and poetic) coincidences, serendipitous and unexpected connections that were, sometimes, just too fitting to be true.

Figure 1. The Subharchord I/II Prototype at the E-Studio of the Akademie der Künste Berlin
Source: photo by the author

Weighing maybe a hundred kilos if not more, extremely sensitive and housing columns of components that oftentimes barely hang from dusty circuit boards, this Subharchord I/II is unique. Allow me to explain: this is an "in-between" device, a prototype for the transition between Types I and II of the instrument, meant to test out design improvements that would later be manufactured and distributed in series – of six, seven or eight, depending on who is counting. This transition was not only sonically important, or avantgarde in itself: it was also extremely political as it sought to shape the creation of "extraordinary acoustic Worlds" (Steinke 1958, 4) for a listening subject up to par with the aspirations of the socialist dreams of the GDR.

In the words of Gerhard Steinke:

> "[das Subharchord] wird zunächst auch dem Hörer überraschen […] der Hörer [erwartet] (und hier ist der tatsächlich 'qualifizierte,' geistig, weltanschaulich und kulturell gebildete Hörer gemeint), auch wenn er dem Komponisten mit aller Aufgeschlossenheit entgegenkommt, eine gewisse 'Vorgabe', die ihm das Werk überschaubar macht." (ibid. 4-5)

An actually qualified, intellectually, ideologically and culturally educated listener whose desires and openness would propel the crossing of different boundaries than those possible or even imaginable at that time.

IV

A speech recognition algorithm is as effective as its capacity to perform comparison and probabilistic inference. Yet the methods with which to characterize speech – that is, to provide the grounds from which comparison and probabilistic inference are thought – can be traced back to how, historically, listening as a method for knowing actively contributed to notions of being and belonging, to otherness and the right to be and feel oneself as "human". Scientific listening, human or machinic, is expressed through the elimination of contingency, so that it can respond to a most fundamental question, figuring at the core of any ontological project: "who are you?" In Germany, for instance, early 20th-century scientific listening defined "*fremde Völker*", (Doegen 1925 cited in Vieira de Oliveira 2021)[7] ranked undesirable otherness, and laid grounds for the development of methods by which people could be discerned, and have their fates decided through manner and matter of speech.

Models of scientific listening, mathematical or not, are as effective as their capacity to reduce contingency. The acoustic feature model briefly described above, for instance, performs best the more transient information – consonants – can be smoothed out. A sudden, short, sharp impulse that peaks without coherent harmonic information, that is, a transient, is **sonic information that cannot provide "correct" or "precise" information about itself**. Too many transients and the algorithm cannot perform its intended function; it malfunctions or stops completely. Thus, privilege is given to formants, that is, frequencies that resonate sympathetically – vowels, for instance –, that thus provide readable and recognizable harmonic information. We can say, then, that consistency is at the core of speech recognition algorithms towards well-distributed harmonics.

We do know, however, that in daily speech the boundaries between a vowel and a consonant are always a matter of contingency. It seems that the fuzziness of life flips the script on its head, and it becomes

7. "foreign populations"

increasingly difficult to come up with an ideal method for separating a transient from a formant, or at least defining a workable threshold for its usefulness, as something to be integrated into the recognition process. Defining precisely the types of filtering that should happen so that consistency is achieved is a longstanding challenge for developers, engineers and many bureaucrats. This is partly because transients are unruly signals: they refuse to be contained within a specific threshold. Rather, **transients are the threshold itself**. You can call it a boundary or a border – what is certain is that they do present a problem for speech recognition and dialect determination. A different type of *Grenzproblem* emerges.

V

The Subharchord I/II, that is, the prototype that resides as this writing in the Electronic Music Studios of the Akademie der Künste in Tiergarten, Berlin, exists precisely at the slash. This in-betweenness is more than its charm or a feature: it is its character. When we listen to the way it is imparting its composition, compressing it together with what might be otherwise thought as a supposedly "pure" signal being sent in, we may notice how it is tainting it, coloring it a different shade and spitting out an ensemble composed of the initial sound, filter – and with it every single component and its materials –, amplification and the hair cells of our inner ear. In attending to the haptics of the sounds coming out of this specific Subharchord, that is, how they excite the listener's ears with a specific kind of material quality that is as much felt as it is heard, we begin to perceive them as containing more than what meets the ear, something that might help us make sense of what is the role played by its material composition into whatever signal is sent in. Be it electrical components or lines of code, we can provoke a mode of announcing the materiality of not only what a filter *filters* out but also what gets *added in by the filter*, thus troubling any notion of "purity" that might be assumed to exist.

This can be evoked by a simple gesture: that of striking. Like playing a drum, with striking we produce transients, and with transients we can call forth the materiality of a filterbank, while arresting the process of speech and dialect recognition on the other. Listening to short, sudden and sharp impulses, it seems, is to listen to electrical components being called into being; it makes them announce themselves as raw matter, as brief oscillatory hiccups that fade away as quickly as they emerge, all the while leaving a thick, resonant surplus. Technical language can only do so much to this gesture; yet it becomes a stepping stone from where we can attend to what this form of filtering is actually doing to the signals being sent to it. We gain insight into the ways transients excite and arrest a filter: its spectral arrangement starts dancing ever so slowly, our ears tingling with trying to grasp, to comprehend, to ascertain the jittery bouncing of multidirectional peaks and troughs, but ultimately never settling for any of them, until they disappear.

In other words, striking calls the materiality of one filterbank into being while allowing us to momentarily bring the other into a halt. In striking we provoke a (spectral) ringing, so that this movement, which contains the materiality of these histories as well as the materiality and the energy contained in the striking mechanism – and the particular ways in which it decays – can be listened to. With each struck band of the filterbank, different forms of onset and decay are activated, teased and brought into the world momentarily so that we can attend to its porous yet thick material composition. In enabling a polyrhythmic, material understanding of the process of filtering we can make the case that filtering does as much addition as it does subtraction, that is, it produces something of a different order than a simple input-output relationship.

VI

Attending to the sound of a Mel filterbank – something indeed possible with the Subharchord – allows us to propose a durational character for computation. I am, in here, not looking at the traces left by a technology of the past. I am trying to think of the Subharchord and the DIAS Software together, at the edge of the constraints of linear history, to talk about the ways in which the body listens in an always uneven arrangement with other bodies, machines, jurisprudences, texts, commands, formulations and, finally, institutional bodies. A tracing, and at that one that is doing work of a different nature than a palimpsest; it is a tracing that contains the old, the new, and the next – all at once. With it, we travel back from the frequency to the time domain. Listening with the materiality of this filter, that is, attending to its emergence as infrastructure and listening to the ways in which sound is constantly being re-composed and imparted by it is a way of thinking with it, of expanding it beyond what its mediated sound can contain.

Focusing on the Subharchord as a living archive and test site for one of the building blocks of the DIAS software is an active exercise in doing and undoing the temporalities inherent to the dialect test. It becomes a mode of study that attends to the materiality of the object in real time, engaging with it in the present so as to understand its emergence in another time, but also with a listening that seeks to project this design to the current moment. Because here the narrative cannot be reduced to one of linear causality; the Mel filterbank in the Subharchord does not cause the Mel filterbank of most speech and dialect recognition algorithms, including the one in use by the German Migration Authorities. Yet, by drawing them closer together, by bringing them in momentary contact with one another, I am interested in seeing what forces pull them apart, and with which strength they do so. More so, I want to examine what lies in the touching point between the two.

In practice this means that thinking the Subharchord with the DIAS Software might change both trajectories at once. To attend to how these histories come together does not mean to imply that they cause each

other but rather that they contain the traces, echoes, and specters of one another. It allows us to trace the discontinuities that emerge through the lenses of *actually listening* to the Mel filterbank in both stances. Because we listen to it we can attend to the ways in which it materially imparts sound. This listening strikes me as the sheer impossibility that exists in a filter standing for a more "pure" experience of sound, be it in a musical impulse for artistic creation or contained in a migrant voice for the determination of their origin. Exactly because we listen to it, we can take the opportunity of arresting and ask other questions about what it is that it is doing, as it does, when it does.

VII

What interests me here is to mobilize a form of studying media that is engaged by what I am calling a *transient methodology*. This work with "transients" emerges as much as an analytical tool as it does as a material engagement with the Subharchord and the DIAS Software. Ambiguity and non-linearity here become devices in both form and content: I am working with double meanings, double-entendres, and blurring the boundaries between technical and discursive language in both English and German as a methodological device. Materially, it works by exciting the electrical components of the Subharchord I/II's Mel filterbank with transient sounds (i.e., short and sharp impulses that peak and rapidly decay) to elicit not a "filtered sound" but the **sound of the filterbank itself**, that is, its **material composition**. Analytically, I exercise the work of methodological transience by "striking" the (filter) design of the DIAS Software, that is, by putting it in movement with a short analytical and, therefore, historicizing, non-linear impulse, while at the same time bringing its workings to a halt so that they can be carefully put apart and interrogated. These two forms of engaging transients, of course, blend and influence one another, exactly because the work of transients provokes ringing: that is, a transient methodology allows me to excite the Filterbank of the DIAS Software as much as to strike the history of the Subharchord I/II, so that they ring together, that is, announce themselves; in this way they can be held accountable (Amoore 2020).

A listening in transience to the Subharchord, as its own form of media archeology, effects a movement in the opposite direction of what technical and cultural devices such as this synthesizer were actually designed to do. Rather than subjecting (or subsuming) the body to technology, in here, we arrive at a composition with the body but not necessarily through it. The work of the body here starts similarly, that is, with the gesture of engagement, an invitation to activate, to come closer, to touch. That is where they stop and take different directions; touching, here, sets a trajectory for a specific form of orientation but does not propel, only nudge it. This nudging is the refusal of subsuming the body to technology, to let the body be fully transformed or fully engaged with technical devices; the work here is more of a teasing, a scraping or scratching of the surface, the contact that creates the connection, defines the nature of the relation, but does not engulf it in a totalizing gesture. It is not a grasping, but rather a caress, a holding closer without holding tight. Body and technology, in this speculative exercise of coupling without forming, come into relation to one another but refuse to be amalgamated into a constructed sameness. This does not mean, however, that they remain separate; instead, they form something other – a continuity rather than an entanglement (Ferreira da Silva 2016); a vector with a specific orientation that does not contain the whole but is definitely more than the sum of its parts because it can never remain the same as before the event. It has affected, while simultaneously being affected by the encounter.

It is not my intention, however, to develop a general theory of what a transient methodology is or can be, but rather to put it in motion as the work of theorizing. This happens for a particular reason, namely, that a transient methodology is not a method per se but a product of the contingent material arrangements that make it possible to be engaged as such. It is not only because the Subharchord I/II exists *there*, at the Akademie der Künste in Berlin, in its current form and state, and that it can be played; it is all of these and the fact that I found it while thinking of the DIAS software. In other words, the work of transients makes sense here because of the material and analytical possibilities offered by the shared and serendipitous co-presence of the Mel filterbank in both the Subharchord I/II and the DIAS Software, as much as it does by the former's explicit and the latter's implicit concern with *Grenzprobleme*.

(coda)

A last remark: I kept the exact German formulation throughout because of the speculative room it allows. In German, the word Grenze might be used both for *boundary* as for *border*, and to think of the GDR together with the BAMF makes this *double-entendre* even more evident. Thus, activating the unruly work of transients mobilizes another kind of historical and spectral (de)composition; one that can be used to strike the many coalescings of history that haunt every corridor of every immigration office, every *Ankunftszentrum*,[8] and every problem created by borders.

Acknowledgments

This text was first presented as a lecture-performance and subsequently developed for publication during my Fellowship at the Leuphana Institute for Advanced Studies in Culture and Society (LIAS) in Lüneburg, Germany. This research has been partially funded by the Junge Akademie der Künste Berlin (2022), and LIAS (2023–2024).

8. This is the name given to reception centers for refugees in Germany. Literally "Arrival Center."

References

Ali, Ahmed, Najim Dehak, Patrick Cardinal, Sameer Khurana, Sree Harsha Yella, James Glass, Peter Bell and Steve Renals. "Automatic Dialect Detection in Arabic Broadcast Speech". *arXiv* (2015, 23 September). DOI: https://doi.org/10.48550/arXiv.1509.06928

Amoore, Louise. *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham: Duke University Press, 2020. DOI: https://doi.org/10.1215/9781478009276

Bundesamt für Migration und Flüchtlinge. "Drucksache 20/3238: Einsatz von Dialekterkennungssoftware Im Bundesamt Für Migration Und Flüchtlinge". *Bundestag* (2022, 31 August). https://dip.bundestag.de/vorgang/einsatz-von-dialekterkennungssoftware-im-bundesamt-f%C3%BCr-migration-und-fl%C3%BCchtlinge/290504

Dusan, Sorin and Li Deng. "Recovering Vocal Tract Shapes From Mfcc Parameters". *Proc. ICSLP*, (2000): 3087-3090.

Ferreira da Silva, Denise. "On Difference Without Separability". *Incerteza Viva (Living Uncertainty): Catálogo Da 32a São Paulo Art Biennial*. (São Paulo, SP, Brasil, 2016), 57–65.

Hanani, Abualsoud, Hanna Basha, Yasmeen Sharaf and Stephen Taylor. "Palestinian Arabic Regional Accent Recognition". *2015 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, (2015): 1-6. DOI: https://doi.org/10.1109/SPED.2015.7343088

Kanervisto, Anssi, Ville Vestman, Md Sahidullah, Ville Hautamaki and Tomi Kinnunen. "Effects of Gender Information in Text-Independent and Text-Dependent Speaker Verification". *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. (New Orleans, LA: IEEE, 2017), 5360-5364. DOI: https://doi.org/10.1109/ICASSP.2017.7953180

Li, Xiaochang and Mara Mills. "Vocal Features: From Voice Identification to Speech Recognition by Machine". *Technology and Culture 60*, no. 2 (2019, 18 June): S129–60. DOI: https://doi.org/10.1353/tech.2019.0066

Ma, Zichen and Ernest Fokoué. "A Comparison of Classifiers in Performing Speaker Accent Recognition Using MFCCs". *Open Journal of Statistics*, no. 4 (2014, 20 June): 258. DOI: https://doi.org/10.4236/ojs.2014.44025

Miersch, Manfred. "Die Geschichte des Klangerzeugers 'Subharchord' (2)". *Keyboards*, (2003).

Nuance Communications. "The Essential Guide to Voice Biometrics". *Nuance Communications*, (2015, 11 December). DOI: https://doi.org/10.1016/S0969-4765(15)30042-4

Nuance Communications. "What We Do - Our Mission". *Nuance Communications*, (nd.). https://www.nuance.com/company-overview/what-we-do.html

Steinke, Gerhard. *Auf dem Wege zu einer neuen Klangkunst*. Berlin: Akademie der Künste, 2015. DOI: https://opac.adk.de/00/bvnr/BV043446027

Steinke, Gerhard. "Die Wiedergeburt Des Klangerzeugers Subharchord". In: Berno Odo Polzer and Thomas Schäfer (eds.). *Katalog Wien Modern*. (Saarbrücken: Pfau, 2007), 162-164.

Steinke, Gerhard. "Probleme der elektronischer Klangerzeugung (II)".61 1958. ADK 0-0486. Akademie der Künste. https://archiv.adk.de/objekt/2397817

Steinke, Gerhard. "Rückblick auf die Arbeit des 'Studios für künstliche Klang - und Geräuscherzeugung' im Rundfunk - und Fernsehtechnischen Zentralamt (Deutsche Post) 1960 - 1970, Berlin - Adlershof". *Deutsche Sektion der Internationalen Gesellschaft für Elektroakustische Musik*, no. 10 (1993, 9 January): 15-26.

Stevens, S. S. and J. Volkmann. "The Relation of Pitch to Frequency: A Revised Scale". *The American Journal of Psychology,* vol. 53, no. 3 (1940): 329-353. DOI: https://doi.org/10.2307/1417526

Vieira de Oliveira, Pedro J S. "'...The Table Was Set, and We Were Never Dead': On the Persistence of Colonial Listening in Germany". *The Journal of Media Art Study and Theory,* vol. 2, no. 2 (2021): 89-101.

CV

Pedro J. S. Vieira de Oliveira
Independent scholar and artist
pedro@oliveira.work

Researcher and sound artist whose work is committed to an anticolonial study of listening and its material intersections with the violences of the European border. His ongoing research concerns specifically with the long histories of technical listening in Germany. In his academic practice, he has held Fellowships at the Leuphana Institute for Advanced Studies and the Helsinki Collegium for Advanced Studies, and taught at the Humboldt University Berlin, Linnaeus University Växjö (Sweden), and the Heinrich-Heine University Düsseldorf. Currently, he is part of the guest faculty in Sound Studies and Sonic Arts at the Universität der Künste Berlin. He holds a PhD, also from the Universität der Künste Berlin.