

Chapter Title: The Weaponization of Datafied Sound: The Case of Voice Biometrics in German Asylum Procedures

Chapter Author(s): Daniel Leix Palumbo

Book Title: Doing Digital Migration Studies

Book Subtitle: Theories and Practices of the Everyday

Book Editor(s): Koen Leurs, Sandra Ponzanesi

Published by: Amsterdam University Press. (2024)

Stable URL: <https://www.jstor.org/stable/jj.11895524.23>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



This book is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>.



Amsterdam University Press is collaborating with JSTOR to digitize, preserve and extend access to *Doing Digital Migration Studies*

13. The Weaponization of Datafied Sound: The Case of Voice Biometrics in German Asylum Procedures

Daniel Leix Palumbo

Abstract

Since 2017, German border authorities have used voice biometrics to analyse the accents of undocumented asylum seekers to pinpoint their country of origin and, consequently, determine their eligibility for asylum. This chapter analyses voice biometrics in the framework of sonic weaponization and in the sociotechnical imaginary that supports today's voice biometric industry, with the aim to direct new critical attention to the auditory realm of top-down governmentality digital practices in the field of digital migration studies. Finally, I argue how the very sonic nature of voice and its capacity to establish intimacy is datafied and weaponized to construct digital identities and control borders—alienating the value and meaning of voice as a site to affirm one's subjectivity and political agency.

Keywords: sound; voice; biometrics; asylum procedures; datafication.

13.1 Introduction

Sound is an essential element underpinning the relational nature of voice (Cavarero, 2005). It activates language into utterances, affording the capacity of voice to establish relations between subjects and objects (Bakhtin, 1979/1986). Sound also symbolically connects voice to identity. It affords the uniqueness of one's voice through a sonorous self-revelation that overcomes linguistic registers or signification, exemplified, for example, by announcing oneself on the phone through the pronouns "I" or "me" (Cavarero, 2005). At the same time, affective, social, ethical and political forces colour the sound

Leurs, K, and S. Ponzanesi (eds.), *Doing Digital Migration Studies: Theories and Practices of the Everyday*. Amsterdam: Amsterdam University Press, 2024

DOI 10.5117/9789463725774_CH13

of voice, making it the channel for a person to express their subjectivity and form links and groupings (Kanngieser, 2012). Thus, the sound of voice accentuates “individual identity as a relational project” and allows individuals “to foster confrontations between one and another, and to infuse language with degrees of intimacy” (LaBelle, 2010, p. xxi).

As such, voice also plays a primary role in machine-mediated local and transnational communication—from the telephone to the oral/aural renaissance occurring in today’s digitized media context (Gallego, 2021). The case of migrants is no exception, as online voice calls and messages on messaging platforms are present in everyday digital practices (Greene, 2020; Weitzel, 2018; Zijlstra & van Liempt, 2017). The sound of voice can thus be argued to be constitutive of migrants’ bottom-up digital practices of everyday meaning-making, which the field of Digital Migration Studies recognizes as a form of “cosmopolitanism from below” (Leurs & Ponzanesi, 2018). Through digital technologies such as smartphones and social media, migrants take part in practices of boundary-making by maintaining bonds with diaspora communities—including family and friendships at a long distance—while simultaneously crossing boundaries through local, intercultural networking and integration with the host society. Such practices create diasporic spaces of belongingness while providing individual and collective intercultural, cosmopolitan experiences (Leurs & Ponzanesi, 2018), also imbued with the intimacy elicited by the sociopolitical and affective forces of the soundings of voice. But what if this moment of intimacy established by the sound of voice is weaponized in the name of policing the digital fortress of Europe?

This question brings us to the reception centres of Germany, where the German Bundesamt für Migration und Flüchtlinge (BAMF, or Federal Office for Migration and Refugees) started an ambitious project for the massive digitization of its whole administrative infrastructure and border policing system in 2016 (BAMF, 2018). This was initiated in response to the so-called “European refugee crisis” and premised on a technocratic imaginary, favouring the implementation of advanced technologies (such as blockchain and machine learning) to guarantee efficiency, transparency and control in migration management (Witteborn, 2022). If the introduction of blockchain technology aimed to facilitate data sharing and communication among the different units of BAMF in German’s federal system (Witteborn, 2022), at the core of the digitization process was the implementation of IT tools for the assessment of asylum seekers’ country of origin and identity. These tools were aimed at making up for the inability of Germany’s administration to cope with the high number of asylum applications, operationalizing the call for more restrictive measures to decrease the inflow of asylum seekers and

also ease forced returns (BAMF, 2018; Kreienbrink, 2018). According to official information, most asylum seekers arrive in Germany without presenting a passport; therefore, the BAMF implemented an IT toolbox to handle asylum procedures to reinforce the deterrence policy toward migrants not covered by the Geneva Convention (Kreienbrink, 2018).

In the IT package embraced by BAMF, the controversial tool of voice biometrics stands out. It is built around automatic dialect recognition and named the Language and Dialect Identification Assistance System, or DIAS, by BAMF. This was presented as the flagship of the BAMF digital renewal, and has been in full use since the first testing in 2017, despite criticism of its lack of accuracy (Bellanova & Fuster, 2019; Biselli, 2018). Voice biometrics are deployed when decision-makers do not trust asylum seekers' claims about their origin, for instance, if the documentation is suspected to be counterfeit or is not provided (Tangermann, 2017). In these cases, the voice biometric software records the voice of the asylum applicant speaking in her native language, focusing on the acoustic qualities and inflexions to analyse their accent and determine the country of origin. As a result, the software releases a report that assesses the applicant's accent in probability percentages—which might confirm or contradict their claims—assisting the decision-maker in determining the eligibility for asylum.

The implementation of voice biometrics in Germany can be considered the meeting point between the logic of techno-solutionism (Madianou, 2019) and the increased market interest in voice biometric technology (Gallego, 2021; Kang, 2022; Turow, 2021). The former refers to the techno-solutionism and hype that leads governments to attempt to solve complex social issues through the latest technological innovations, without first meticulously understanding situations that may be not suited to digital disruption (Madianou, 2019). Consequently, asylum seekers are used as a testing ground for experimentation with data-driven practices such as biometrics which, against border authorities' claim of mechanical efficiency, cause the risk of failures that further endanger the most vulnerable groups (Madianou, 2019). The latter trend instead concerns the recent large investments made in speech recognition to make voice a central medium of interaction within networked technologies and online services, with companies interested in measuring and translating the information of the human voice into data for digital profiling and other purposes (Gallego, 2021; Kang, 2022; Turow, 2021).

The interest has also been shared by migration offices and state agencies in general. This is evidenced by the development of the Speaker Identification Integrated Project (SiiP), the first international and interoperable database of voice biometrics to support law enforcement investigations into

transnational threats, terrorism and organized crime (Jansen et al., 2021); the use of speaker recognition technology by the US National Security Agency (NSA) for counter-terrorism operations (Kang, 2022); the implementation of voice biometrics in the German border system; but also by the more recent programme started in Turkey in 2019, called Capacity Building for Effective Nationality Determination (Bellanova & Fuster, 2019). Due to a commitment to the EU, Turkey also enlisted automated language tests to detect a person's country of origin among its strategies aimed to slow down the inflow of migrants and refugees. However, at the end of the trial phase, the software for language recognition has not been implemented in Turkey due to unsatisfactory results (Ozkul, 2023).

In this respect, the collection and analysis of voice data have already raised many concerns about questions of surveillance. These accounts have mainly focused on the literal translation of spoken audio captured by AI voice technologies (Alepis & Patsakis, 2017; Woods, 2018). However, this scholarship has largely overlooked how the soundings of voice can also be used to manage digital identities. In this chapter, I direct critical attention to the information embedded in the sonic aspect of voice, indicating how its interpretation can be instrumental in constructing knowledge about subjects, creating power imbalances between the speakers, and the political actors behind the use of AI voice technologies. To do so, I look at the example of the German border system, where, through voice biometrics, sound becomes an emerging factor in the datafication of life, determining the identity and country of origin of asylum seekers.

While framing BAMF's accent recognition technology within the broader discussion on voice biometrics and data-driven security practices, I address this case from the specific perspective of sonic weaponization. With this term, I refer to the broad range of techniques that manipulate physical and affective properties of sound by converting it into a "weapon": intending to coerce, manage and control subjects (Goodman, 2012). Literature on the use of sound in contexts of aggression, torture and war shows how material and affective forces of sound can be harnessed by political actors—despite usual conceptions of acoustic pureness and abstraction. This scholarship also draws on the rise of affect studies, where affect is applied to sonic weaponization in its many different connotations. In particular, sound can be weaponized through its capacity to modulate moods and induce psychological effects, provide a sense of communion and belonging, reshape surroundings, and alter the sense of the immediate, elicit alert, or intimacy (Birdsall, 2012; Goodman, 2012; Thompson & Biddle, 2013). By incorporating these studies, I build a nuanced framework to make it possible to grasp the logic and operation

of voice biometrics in asylum procedures, while providing an integrated approach to address orality/aurality in contemporary modes of governance.

13.2 Methodology

Access to knowledge regarding developments in policing is notoriously difficult (Brayne & Christin, 2021), and much information is withheld from the public by the BAMF on the grounds of national security. For instance, it has not yet been declared what algorithms and data are used to train the voice biometric software, how many languages it can recognize, and what the error rate for the different languages it recognizes is. Therefore, in my argumentation I draw from various publicly available sources and secondary data. These include official documents released by the BAMF, such as the *Digitisation Agenda* and *The stages of the German asylum procedure*. Information drawn from this documentation does far more than provide an understanding of the steps and legal ground of asylum applications, or the framing of the use of voice biometrics and its rationale within the procedures. It also embodies aspirations, motives and broader cultural imaginations that foster the experimentation with new technologies for decision-making, which “successfully” allowed the agency to meet the challenges presented by the European refugee crisis and posit it as “a new leader in digitisation” and “a digital, breathing public authority” (BAMF, 2018, p. 4). Along this line, the documentation shows the agency’s willingness to reaffirm the geopolitical role of Germany within the EU, and in the general digital society. It states that systematic examination with migration authorities from other countries is taking place to discuss “interesting opportunities for working together, especially in terms of dialect recognition” (BAMF, 2018, p. 35).

My argumentation also incorporates a discussion of a part of the databases on which BAMF’s voice biometrics rely—the only ones that are publicly available by the immigration agency at the present moment. This information was provided with figures on the use and error margin of voice biometrics, together with other documents, following a parliamentary inquiry in the Bundestag, the German Federal Parliament, and various freedom of information requests lodged in the inquiry platform *FragDenStaat* by journalist and computer scientist Anna Biselli. The content analysis of the meta-information of this speech training data—sold to BAMF by the University of Pennsylvania’s Linguistic Data Consortium (LDC)—allows for an understanding of the system’s decision-making, and the role played by affective capacity of sound in it. The other documentation obtained by the

freedom of information requests includes information on official internal regulations regarding document verification, the establishment of identity, and training documents for BAMF employees. This information specifically grants insight into the key features of voice biometrics, and the compositional steps of their use by decision-makers in the context of the personal interview for asylum. Finally, the analysis includes information from the work of Anna Biselli and sound artist Pedro Oliveira, who have been on the frontline in reporting the forms of racial profiling and errors caused by voice biometrics.

13.3 Weaponizing Sound

It was not long ago that outrage was sparked online by the news of Greek police testing the use of long-range acoustic devices, or “sound cannons,” to deter migrants from crossing into Europe. Tested during the quiet months of the corona pandemic in 2021 along the border with Turkey, these devices mounted on trucks can emit noise matching the loudness of a jet engine, causing permanent deafness and other health issues. The reporting of this news by journalists and human rights activists not only provoked indignation, but also implicitly shed a light on more general public discussion about the coercive qualities that acoustic power can have (Euro-Med Human Rights Monitor, 2021). While the public outcry is recent, there is a long tradition of scholarship we can draw from to explore how sound can be harnessed by political actors.

The study of music and sound in the context of sonic violence and torture presents us with an important area of inquiry. This scholarship allows for a discussion on sound and power, adding to how the politics of sound—besides generating subjectivities and collectivities—can also be repurposed to maintain grids of power along intersecting hierarchies of class, gender, sexuality, race, among others (Revill, 2000; Stoeber, 2016). The analysis of sonic materiality is one of the main productive entry points, which serves to dismiss any idea of ineffability connected to sound. As will be discussed in the next section, operations of identification and authentication through voice biometrics reduce voice to a partial ontology—solely understood as a fixed sound material devoid of any socio-cultural dimension, and where the continuous becoming of a speaker’s vocal identity is denied (Kang, 2022). As such, sound needs to be understood not as pure or neutral, but as a complex entity whose use is informed by political dispositions. The material and affective qualities of sound contributed—for example, through sound system cultures—to forms of identification in the networked, diasporic community of the Black Atlantic (Gilroy, 1993; Goodman, 2012; Henriques,

2011). At the same time, sonic materialities are also the basis of a multitude of coercive acoustic weapons (Daughtry, 2014; Goodman, 2012). This idea of the abstraction of sound is often harnessed by political actors to conflate violent and brutal uses of sound with notions of “no-touch” control and “non-lethal” weaponry (Cusick & Joseph, 2011). In response, these studies have denounced the problematic use of sonic material and affective capacities to perpetrate physical and psychological violence. I build on this framework to situate the turn to voice biometrics in asylum procedures, where the sonic nature of voice serves to propose the same ideas of sound pureness—where no coercion is involved, and to offload accountability from human actors.

In particular, musicologist Suzan Cusick has carried out pioneering work on the use of loud sound in the detention camps of the United States’ “global war on terror.” Relying on first-person witnesses of interrogators and former detainees from US detention camps in Afghanistan, Iraq and Cuba, Cusick (2008) describes how music and sound are an integrant part of various techniques of sensory manipulation to force detainee’s tendencies towards compliance during interrogations. Cusick & Joseph (2011) argue that sound is weaponized within contemporary practices of torture because of its foundation on the so-called principles of “no-touch” control. This refers to principles to control bodies without leaving readily identifiable physical traces, and where the one-to-one relationship between torturers and the one tortured is truncated—with no person to be blamed. The ephemeral qualities of sound, therefore, are key to its understanding as seemingly innocuous and malleable.

Secondly, Steve Goodman (2012) draws from affect and social theory to discuss the material quality of sound, focusing on the role of “vibrational” force within sonic power relations. The understanding of sound as a vibrational force—the outcome of a combination of frequencies and loudness—is crucial for grasping the invisible materialities and physical capacities by which sonic power can be weaponized at a coercive and affective level (Goodman, 2012). Consider, for instance, the exploiting of sound vibrations occurring at frequencies that are beyond the human standard hearing range, but which are still perceivable by humans at a tactile level as a physical rumbling. Goodman (2012) calls such frequencies at the periphery of acoustic perspective “unsound.” Along with sonic loudness exceeding the human threshold for pain, unsound is harnessed for developing brute, non-lethal acoustic weaponry in war scenarios and torture techniques in detention centres. Law enforcement agencies also weaponize low-frequency infrasonic tones¹ to arouse fear, anxiety and bad vibes to disperse demonstrations

1 Frequencies which are below the lower limit of human audibility.

or riots (Goodman, 2012). Addressing these extremely brutal acoustic phenomena, Martin Daughtry (2014) coins the term “thanatosonics” to indicate that level of intensity where sound’s “inherent polysemy is sacrificed to the unequivocal demands of acoustics” (p. 39). In the abovementioned cases of acoustic violence, the sheer sonic materiality reaches such a level of intensity that it stops behaving like sound. It cannot be even listened to or witnessed, but turns the passive bodies into pure victims of vibration, inflicting permanent and profound physiological damage (Daughtry, 2014).

Scrutinized by this critical framework, and supported by increased interest in voice within the tech sector and the trust placed in big data and the body as guarantors of identity, sound becomes, through voice biometrics, a newly available singularity for allowing the biometric assemblage (Madianou, 2019) and algorithmic augmentation of borders (Ajana, 2015). Voice biometrics, along with the sound cannons in Greece, thus represent a new potential sonic weapon within the digital fortress of Europe.

13.4 Biometric Technology and the Voice as Evidence

Resulting from growing anti-immigration rhetoric and political pressure to deter migration arisen with the refugee crisis, the “Fortress Europe” approach has enforced migrations policies within and beyond Europe, especially furthering the incremental adoption of data-driven technologies (Dijstelbloem et al., 2011). This development is premised on the assumption of a self-evident relationship between data and identities (van Dijk, 2014), where the aggregation and algorithmic processing of a large amount of information enables a surveillant knowledge infrastructure (Bollier & Firestone, 2010). Such an infrastructure operates by performing tasks of identification and identity authentication (Ajana, 2013), or constructing profiles according to the logic of risk, to speculate about future behaviours or threats and act pre-emptively (Amoore, 2011). In this respect, biometric data has been particularly instrumental in developing migration policies, thanks to authorities’ confidence in the informatization and digitization of the body (Kloppenborg & van der Ploeg, 2020). In other words, biometrics are accredited by the belief in a specific equation between bodies and identities, and in registered and processed digital data for managing risk.

It is in this broader framework that the voice-body-identity equation needs to be examined, to understand the sociotechnical imaginary that informs the current global multibillion-dollar voice biometric industry. Edward Kang (2022) argues that this ferment in the voice biometric industry

is supported by normative ideologies and sociotechnical beliefs which frame voice, body and identity as fixed and correlative objects. As a consequence, voice is treated as a fixed, extractable and measurable “scientific” sound object (Schaeffer, 2017) to which identity can be attached through “scientific” methods employed to make socially constructed judgments about the speaker. Practices of identification through voice biometrics therefore treat voice as a site where individual identities can be measured in terms of race, gender, nationality, etc., and thereby as “a reliable and stable carrier of knowledge about the body” (Kang, 2022, p. 13). By equating identity with a collection of physically measured information, which is then rendered as processable digital data, the subject’s own experience is ignored (Wevers, 2018) and voices are mistaken for fixed objects (Magnet, 2011). However, ways of speaking and soundings of voice should be understood in relation to the cultural, social and institutional contexts in which speakers are communicating (Hall, 1976). There is no universal form of speaking or a fixed one, but the soundings of a person’s voice need to be considered as the result of her community and sociocultural context. Instead, the logic of voice biometrics assumes voice to be a permanent and coherent phenomenon that stays the same over time, without ageing, suffering changes or undertaking self-initiated alterations. Nevertheless, this logic can be invalidated, for example, by singers. They go through meticulous training regimes to shape and refine their vocal physiology, which indeed partly determines the sound of one’s voice, but it “is not fixed, and like our cultural identities, is always in the process of becoming” (Kang, 2022, p. 12).

The reliance upon biological understandings of identity complicates the claims of biometric technology to mechanical objectivity and infallibility, instead causing the misrecognition or misidentification of individuals and having serious consequences on their mobility. According to this framework, voice biometrics errors reflect long-standing normative cultural assumptions and beliefs, which lead to the techno-solutionist and experimental use of speech recognition that further oppresses marginalized groups like asylum seekers.

13.5 Voice Biometrics and the Datafication of Sound

Language analysis to determine origin and identity is not a new phenomenon within asylum procedures. In many countries inside and outside the EU, under forensic linguistics programmes such as LADO, it has been the long-term, specialized domain of language experts hired to assess individual cases

(Patrick, 2012; Pfeifer, 2023). The case of Germany, however, indicates a major shift to outsourcing this task to speech recognition systems to generate faster, automated results. But this shift is not followed by improved thoroughness or accuracy in asylum procedures. Despite the claims of success made in the agency's Digitization Agenda, the BAMF revealed that voice biometrics have an error margin of 15%—which improves to 10% when it comes to recognizing Levantine Arabic (Deutscher Bundestag, 2018). This may result in many applicants having their right to asylum denied on the basis of distorted results (Biselli, 2018). Despite BAMF's promises to implement improvements, voice biometrics errors prompted criticisms similar to those levelled at LADO—namely, that using language as a marker of geographic origin is problematic, as it does not take into account the sociolinguistic biography of an individual, and the context-dependent nature of language (Rosenhouse, 2013). That means language analysis, whether automated or not, cannot be used as a reliable method to definitely indicate an individual's region of socialization or origin (Pfeifer, 2023). The BAMF claims that voice biometrics are only an assistant tool to provide an initial assessment, that it has no direct consequences on the final decision, and does not replace the evaluation by the decision-makers (2018). However, the research of Anna Biselli (2018) has reported cases where voice biometrics were used to make decisions, compromising asylum seekers' applications.

In the asylum application procedure, the personal interview arguably represents the most important step—where decision-makers require asylum applicants to describe their story and biography by providing evidence to support their claims (BAMF, 2019). However, if their story is not supported by valid documentation, voice biometrics are introduced. These perform the task of identity authentication, creating a biometric template from the asylum seekers' speech that is checked against different databases of stored biometric templates (BAMF, 2017). In other words, a person's speaking aspects are captured, processed and then confronted with those stored in the software's databases to verify the truthfulness of their claims, and scientifically pinpoint their identity and origin. The provenance of most databases, or speech corpora, has not been revealed by the BAMF. However, the Federal Interior Ministry has only indicated that those for Levantine Arabic, in its different variations, were purchased from the LCD of the University of Pennsylvania (Deutscher Bundestag, 2018). The voice biometric test for the asylum applicant who is being questioned is performed in a dedicated room. In this room (see Figure 13.1), the applicant picks up a phone handset, in which she is required to speak following a signal. She is asked to describe, in the fullest possible detail and without any interruption, an image



Figure 13.1. Illustration of the use of voice biometrics on asylum applicants.
Note. The slide is taken from the training documents for BAMF personnel. It provides an overview explaining in which cases voice biometrics are used and illustrates the procedure for the analysis of asylum applicants' speech (BAMF, 2017).



Figure 13.2. Sample of a voice biometrics result report.
Note. The slide shows what a result report produced by BAMF's voice biometrics looks like. The report consists of three different sections: the first lists the dialects/accents assessed for the asylum applicant in probability percentages, as well as the Log Likelihood Ratio (LLR); the second indicates the details of the recording, namely its duration (*Aufnahme-Dauer*) and the amount of spoken audio in the recording (*Netto Sprachdauer*)—both values should ideally diverge only a little according to the report; the third and final section concerns the technical details of the assessment, which are, however, dismissed as not relevant for the language assessment (BAMF, 2017).

for two minutes (BAMF, 2017). After just a few minutes, the system releases a result report, which assesses the speaker's dialects or languages of origin in probability percentages (see Figure 13.2). Depending on the conformity of the applicant's claims about her origin with the probability percentages produced by the accent recognition software, the fate of the asylum seeker is determined—allowing or denying access to a life in Germany.

The strategies adopted in conducting the accent recognition process are not left to chance. A phone handset used to capture the applicant's speech, or the specific choice of the image they are meant to describe, are instrumental within the evaluation process. As identified by Pedro Oliveira, BAMF tries to operationalize certain strategies by following sociolinguistic guidelines on how to elicit a speaker to speak naturally and in a more "sincere" manner when conducting an analysis of a dialect—albeit not considering the ethical problems imposed by the discipline (2019). The "success" of voice biometrics is dependent on strategies that would elicit the speaker to provide the most natural account of a person's speech in terms of prosody, pronunciation and vocabulary, in order to ensure that the speaker, consciously or not, adapts their speech. This strategy is operationalized by replacing the figure of the researcher (or a recognizable recording machine) with a telephone device, which usually suggests intimacy or familiarity (Oliveira, 2019). The adoption of this setting is also in line with findings in sociolinguistic research that the evocation of dialects is most successful with topics with which speakers are emotionally involved (Meyerhoff et al., 2012).

Indeed, the image that needs to be described via the phone recalls something familiar—for instance, in the depiction of a Muslim family eating together in a domestic setting (Oliveira, 2019). In the practices of BAMF, therefore, the phone device is instrumentalized as an object that evokes a connection to an elsewhere for the applicant: in this case, home, family and the intimate sphere. The pattern is also exemplified when looking at the name of the speech corpora on which the accent recognition software relies. These are named *CALL FRIEND* and *CALL HOME*, consisting of captured conversations occurring between Arab speakers in moments where intimate reconnections are established on the basis of voice (Oliveira, 2019). The LCD harvests these corpora by involving participants resident in North America to call friends and families overseas, so as to increase its repository created for language-related education, research and technology development. BAMF purchased the corpora, but no research or development project was initiated with the University of Pennsylvania (Bewarder, 2019). The use of this corpora by BAMF might be considered an example of function creep (Madianou, 2019), which refers to how data collected for one purpose is

used instead for a very different one—in this case, speaker authentication for border control.

The enactment of these strategies, from the phone handset to the choice of the speech corpora, indicates a further step within the broader tendency by authorities to exploit and control the domain of affectivity. Digital processes of top-down governmentality already exploit migrants' affective bondings by surveilling the data traces of their diasporic and cosmopolitan digital practices of everyday meaning-making (Leurs & Ponzanesi, 2018). In this specific case, sound as a vehicle of affectivity is instead harnessed to set up an imagined moment of transnational intimacy and is weaponized against the speaker, reminding “an asylum seeker that home is always elsewhere” (Oliveira, 2019, p. 6). In addition, what needs to be observed is that what the BAMF's system embodies is:

a certain way of governing through data that does not even pretend to translate reality. It is unconcerned with what asylum seekers actually say, or, more exactly, it is grounded on the assumption that whatever they might say is not worth being taken into account before the trustworthiness of their belonging is ascertained (Bellanova & Fuster, 2019, p. 360).

Voice biometrics in this way alienate the political and symbolic nexus between voice and identity by turning it into a biological equation. Asylum seekers attempt to communicate their stories, but these are essentially ignored and made dependent upon the decision of voice biometrics—focused on translating sound into data to establish their identity and country of origin. The digital identity is, in other words, sought in how the asylum seekers' voice sounds, not in what they are saying. Voice loses its power as a means of self-affirmation, and is treated solely as fixed sound material—where the socio-cultural dimensions that colour its composition and determine its continuous becoming are ignored. This process implies the idea of an identity that is not attached to the individual, although it is paradoxically extracted from the sonic materiality of their speaking voice. The same sound of an accent which would normally be the reflection of one's life (but not one's country of birth) is weaponized for the assignment of a fixed identity—or at least, the fantasy of a clear-cut identity. Thus, voice biometrics in asylum procedures weaponize sound as a vehicle to exploit the domain of intimacy by treating it as a fixed, measurable acoustic material to impose their truth.

At the same time, the sonic nature of voice biometrics also works to propose the technology in public discourse as a cutting-edge and mechanically

efficient tool for securing the German borders with no coercion involved. As opposed to fingerprinting, which is commonly associated with criminality and violent extrapolation (Aas, 2011; Ajana, 2013), the capture of voice can be presented as a quick and neutrally efficient check of one's identity—where there is no processing of bodily parts if not of an abstract vocal sonority. However, the datafication of one's voice, in reality, enables a new form through which sound can be used for the control of individuals, one which takes part in the making of the digital selves and the knowledge that composes them. Voice biometrics enable sound as a new factor through which people are categorized and assigned a status of identity affecting their life and im/mobility (Cheney-Lippold, 2017). In particular, the datafication of the acoustic features of a person's voice becomes another available weapon to exclusively criminalize, discriminate and marginalize targeted, unwanted migrant populations. In this way, borrowing from Michelle Weitzel (2018), the intrinsic political key value of the voice to determine yourself—to voice yourself according to your own individual subjectivity—is thus wrested from and turned against the speaker.

13.6 Conclusions

Practices of decision-making through voice biometrics involve new implications in the discussion on the value and meaning of voice in digital Europe. In an investigation of institutional European initiatives that seek to digitally bring to the foreground migrants and refugees telling their stories, Myria Georgiou focuses on voice to discuss the complexities of representational politics in migration. Although providing an “alternative form of mediation against the voiceless and threatening *Other*” predominating European mainstream media (2018, p. 20), digital Europe normatizes migrants' voices according to the colonizing gaze of mainstream representation of migration. These voices are only framed in narratives of care, where their rights are limited only to humanitarian aid—but not to political and legal achievements. Thus, representations of migration in hyperspace perform bordering power by not acknowledging the political agency of migrants' voices—whereas Hannah Arendt (1958) describes voice as the privileged way for individuals to identify themselves to others and define themselves as political subjects. At the same time, as discussed in this chapter, top-down governmentality of digital practices through voice biometrics reaffirm bordering power by alienating the symbolic relationship between voice and identity, where voice should represent a channel of self-expression which illustrates a

person's identity based on their claims and experiences (Couldry, 2010). Voice biometrics colonize this space of self-expression, harnessing the sound of voice to turn it into a site of identity construction for purposes of border control.

The weaponization of sound in asylum procedures occurs on different levels: by exploiting its affective capacities to establish intimacy, setting up fictional moments of transnational belongingness to elicit applicants to speak naturally; by turning voice and its political value into a fixed, extractable sound material—denying one's subjectivity and the sociocultural dimensions that aesthetically define one's voice, describing a life; and finally, by denying its inherently political nature and proposing a pure and ineffable one, which facilitates technosolutionist experimentations on asylum seekers in line with the current hype for the voice biometric industry. My preliminary considerations, and the framework built above, can serve future research in the field of digital migration studies that seeks to address orality/aurality in contemporary modes of governance. The use of voice biometrics in German asylum procedures represents a novelty. However, efforts by BAMF to partner with other immigration agencies, along with the most recent news of Turkey implementing voice biometrics in the EU-funded project Capacity Building for Effective Nationality Determination, and the development of the SiiP project, lead me to believe that Germany is not going to be an isolated case. These recent developments suggest the emergence of a sonic surveillant knowledge infrastructure in Europe, which will require going beyond the paradigmatic logic of vision that has long served as a metaphorical lens to address power and authority (Carmi 2020; Hsieh 2021; Weitzel 2018). With this different form of datafied knowledge production serving as a new paradigm for making deductions and predictions on individuals and their behaviours, future research will need to go beyond the epistemic limitations of ocularcentrism, and start a deeper discussion of the politics of sound.

Such discussion in digital migration studies might draw from the opposite pole of the framework of sonic weaponization. If I have looked at the use of sound for purposes of control and subjugation, sonic material and affective dimensions can also be deployed as a form of resistance (Goodman 2012; Weitzel, 2018). New accounts on voice biometrics can originate from this different perspective, which I have only briefly mentioned in this chapter. The fieldwork conducted by Weitzel (2018) interviewing undocumented sub-Saharan African migrants in Morocco suggests a starting point to do this: they report how the interviewees, thanks to previous conversations with other migrants, were aware of the modes of interrogation and identification awaiting them at the border, and had already planned counterstrategies

which involved performing muteness or altering their utterances in front of officers, knowing that voice could be a dangerous indicator of identity. As sound, in this respect, represents for migrants a means to contrast processes of identification carried out by officers at the border, future research might look at how undocumented asylum seekers can find ways to "game", "cheat" and side-track voice biometric systems by adopting similar strategies. The manipulation of sound by asylum seekers to their own advantage would further indicate the limits of a technology which already struggles to identify the sonic complexity of voice, defined in a continuous becoming by sociocultural processes and context.

In addition, further contributions might also bridge interdisciplinary dialogue with needed accounts from legal and computational linguistics perspectives. Finally, a critical discussion on the politics of sound in asylum procedures can allow for a useful framework to make sense of recent developments occurring in the voice biometric industry, considering the increasing aspirations of conglomerates in the datafication of voice as a site of profit and control.

References

- Aas, K. F. (2011). "Crimmigrant" bodies and bona fide travellers: Surveillance, citizenship and global governance. *Theoretical Criminology*, 15(3), 331–346. <https://doi.org/10.1177/1362480610396643>
- Ajana, B. (2013). *Governing through biometrics: The biopolitics of identity*. Palgrave Macmillan.
- Ajana, B. (2015). Augmented borders: Big data and the ethics of immigration control. *Journal of Information, Communication and Ethics in Society*, 13(1), 58–78. <https://doi.org/10.1108/JICES-01-2014-0005>
- Alepis, E., & Patsakis, C. (2017). Monkey says, monkey does: Security and privacy on voice assistants. *IEEE Access*, 5, 17841–17851. DOI: 10.1109/ACCESS.2017.2747626
- Amoore, L. (2011). Data derivatives: On the emergence of a security risk calculus for our times. *Theory, Culture & Society*, 28(6), 24–43. <https://doi.org/10.1177/0263276411417430>
- Arendt, H. (1958). *The human condition*. University of Chicago Press.
- Bakhtin, M. (1979/1986). *Speech genres and other late essays* (V.W. McGee, Trans.). University of Texas Press.
- BAMF (2017). *Integriertes Identitätsmanagement-Plausibilisierung, Datenqualität und Sicherheitsaspekte. Einführung in die neuen IT-Tools*. Bundesamt für Migration und Flüchtlinge. https://fragdenstaat.de/dokumente/9653-schulung_idms_bamf/

- BAMF (2018). *Digitisation agenda 2020. Success stories and future digital projects at the Federal Office for Migration and Refugees (BAMF)*. Federal Office for Migration and Refugees.
- BAMF (2019). *Ablauf des deutschen Asylverfahrens. Ein Überblick über die einzelnen Verfahrensschritte und rechtlichen Grundlagen*. Bundesamt für Migration und Flüchtlinge.
- Bellanova, R., & Fuster, G. G. (2019). Composting and computing: On digital security compositions. *European Journal of International Security*, 4(3), 345–365. <https://doi.org/10.1017/eis.2019.18>
- Bewarder, M. (2019, May 24). *Viele Behörden können Identität von Asylbewerbern nicht prüfen*. WELT. <https://www.welt.de/politik/deutschland/article176623392/Fluechtlingskrise-Viele-Behoerden-koennen-keine-Fingerabdruecke-ueberpruefen.html>
- Birdsall, C. (2012). *Nazi soundscapes: Sound, technology and urban space in Germany, 1933-1945*. Amsterdam University Press.
- Biselli, A. (2018, August 20). *Eine Software des BAMF bringt Menschen in Gefahr*. Vice. <https://www.vice.com/de/article/a3q8wj/fluechtlinge-bamf-sprachanalyse-software-entscheidet-asyl>
- Bollier, D., & Firestone, C. M. (2010). *The promise and peril of big data*. The Aspen Institute.
- Brayne, S., & Christin, A. (2021). Technologies of crime prediction: The reception of algorithms in policing and criminal courts. *Social Problems*, 68(3), 608–624. <https://doi.org/10.1093/socpro/spaa004>
- Canavan, A., & Zipperlen, G. (1996). *CALLFRIEND Egyptian Arabic* (Speech LDC96S49) [Data set]. Linguistic Data Consortium. <https://doi.org/10.35111/nnm5-kp69>
- Canavan, A., Zipperlen, G., & Graff, D. (1997). *CALLHOME Egyptian Arabic* (Speech LDC97S45) [Data set]. Linguistic Data Consortium. <https://doi.org/10.35111/d8yb-9m13>
- Carmi, E. (2020). *Media distortions: Understanding the power behind spam, noise, and other deviant media*. Peter Lang.
- Cavarero, A. (2005). *For more than one voice: Toward a philosophy of vocal expression*. Stanford University Press.
- Cheney-Lippold, J. (2017). *We are data*. New York University Press.
- Couldry, N. (2010). *Why voice matters: Culture and politics after neoliberalism*. Sage publications.
- Cusick, S. G. (2008). "You are in a place that is out of the world ...": Music in the detention camps of the "global war on terror." *Journal of the Society for American Music*, 2(1), 1–26. DOI:10.1017/S1752196308080012
- Cusick, S. G., & Joseph, B. W. (2011). Across an invisible line: A conversation about music and torture. *Grey Room*, 42, 6–21. https://doi.org/10.1162/GREY_a_00024

- Daughtry, J. M. (2014). Thanatosonics: Ontologies of acoustic violence. *Social Text*, 32(2), 25–51. <https://doi-org/10.1215/01642472-2419546>
- Deutscher Bundestag (2018, December 19). *Drucksache 19/6647: Einsatz von IT-Assistenzsystemen im Bundesamt für Migration und Flüchtlinge*. <https://dserver.bundestag.de/btd/19/066/1906647.pdf>
- Dijstelbloem, H., Meijer, A., & Besters, M. (2011). The migration machine. In H. Dijstelbloem, & A. Meijer (Eds.), *Migration and the new technological borders of Europe* (pp. 1–21). Palgrave Macmillan.
- Euro-Med Human Rights Monitor (2021, June 3). *EU use of hi-tech to deter asylum seekers is condemnable and dangerous*. <https://euromedmonitor.org/en/article/4445/EU-use-of-hi-tech-to-deter-asylum-seekers-is-condemnable-and-dangerous>
- Gallego, J. I. (2021). The value of sound: Datafication of the sound industries in the age of surveillance and platform capitalism. *First Monday*, 26(7). <https://doi.org/10.5210/fm.v26i7.10302>
- Georgiou, M. (2018). Does the subaltern speak? Migrant voices in digital Europe. *Popular Communication*, 16(1), 45–57. <https://doi.org/10.1080/15405702.2017.1412440>
- Gilroy, P. (1993). *The black Atlantic: Modernity and double consciousness*. Harvard University Press.
- Goodman, S. (2012). *Sonic warfare: Sound, affect, and the ecology of fear*. MIT Press.
- Greene, A. (2020). Mobiles and “making do”: Exploring the affective, digital practices of refugee women waiting in Greece. *European Journal of Cultural Studies*, 23(5), 731–748. <https://doi.org/10.1177/1367549419869346>
- Hall, E. T. (1976). *Beyond culture*. Anchor Books.
- Henriques, J. (2011). *Sonic bodies: Reggae sound systems, performance techniques, and ways of knowing*. Continuum.
- Hsieh, J. C. (2021). Making noise in urban Taiwan: Decibels, the state, and sonosociality. *American Ethnologist*, 48(1), 51–64. <https://doi.org/10.1111/amet.13003>
- Jansen, F., Sánchez-Monedero, J., & Dencik, L. (2021). Biometric identity systems in law enforcement and the politics of (voice) recognition: The case of SiiP. *Big Data & Society*, 8(2). <https://doi.org/10.1177/20539517211063604>
- Kang, E. B. (2022). Biometric imaginaries: Formatting voice, body, identity to data. *Social Studies of Science*, 52(4), 581–602. <https://doi.org/10.1177/03063127221079599>
- Kanngieser, A. (2012). A sonic geography of voice: Towards an affective politics. *Progress in Human Geography*, 36(3), 336–353. <https://doi.org/10.1177/0309132511423969>
- Kloppenborg, S., & Van der Ploeg, I. (2020). Securing identities: Biometric technologies and the enactment of human bodily differences. *Science as Culture*, 29(1), 57–76. <https://doi.org/10.1080/09505431.2018.1519534>
- Kreienbrink, A. (2018). Restriction, pragmatic liberalisation, modernisation: Germany’s multifaceted response to the “refugee crisis.” In I. Sirkeci, E. Lana

- de Freitas Castro, & Ü. Sezgi Sözen (Eds.), *Migration policy in crisis* (pp. 31–51). Transnational Press London.
- LaBelle, B. (2010). *Acoustic territories: Sound culture and everyday life*. Continuum.
- Leurs, K., & Ponzanesi, S. (2018). Connected migrants: Encapsulation and cosmopolitanization. *Popular Communication*, 16(1), 4–20. <https://doi.org/10.1080/15405702.2017.1418359>
- Madianou, M. (2019). The biometric assemblage: Surveillance, experimentation, profit, and the measuring of refugee bodies. *Television & New Media*, 20(6), 581–599. <https://doi.org/10.1177/1527476419857682>
- Magnet, S. A. (2011). *When biometrics fail*. Duke University Press.
- Meyerhoff, M., Adachi, C., Nanbakhsh, G., & Strycharz, A. (2012). Sociolinguistic fieldwork. In N. Thieberger (Ed.), *The Oxford handbook of linguistic fieldwork* (pp. 120–146). Oxford University Press.
- Oliveira, P. (2019). On the endless infrastructural reach of a phoneme. *Transmediale Journal*, 3, 1–6. <https://archive.transmediale.de/content/on-the-endless-infrastructural-reach-of-a-phoneme>
- Ozkul, D. (2023). Automating Immigration and Asylum: *The Uses of New Technologies in Migration and Asylum Governance in Europe*. Refugee Studies Centre, University of Oxford
- Patrick, P. L. (2012). Language analysis for determination of origin: Objective evidence for Refugee status determination. In L. M. Solan & P. M. Tiersma (Eds.), *The Oxford handbook of language and law* (pp. 534–546). Oxford University Press.
- Pfeifer, M. (2023). “The Native Ear”: Accented testimonial desire and asylum. In P. Rangan, A. Saxena, R. Tharoor Srinivasan, & P. Sundar (Eds.), *Thinking with an Accent. Toward a New Object, Method, and Practice* (pp. 192–207). University of California Press.
- Revill, G. (2000). Music and the politics of sound: Nationalism, citizenship, and auditory space. *Environment and Planning D: Society and Space*, 18(5), 597–613. <https://doi.org/10.1068/d2>
- Rosenhouse, J. (2013). Assessing acoustic features in the speech of asylum seekers. *Acoustical Society of America*, 19(1), 1–9. <https://doi.org/10.1121/1.4798413>
- Schaeffer, P. (2017). *Treatise on musical objects. An essay across disciplines*. University of California Press.
- Stoever, J. L. (2016). *The sonic color line: Race and the cultural politics of listening*. New York University Press.
- Tangermann, J. (2017). *Documenting and establishing identity in the migration process: Challenges and practices in the German context; Focused study by the German national contact point for the European Migration Network (EMN)*. [Working Paper No.76]. Federal Office for Migration and Refugees.

- Thompson, M., & Biddle, I. (Eds.). (2013). *Sound, music, affect: Theorizing sonic experience*. Bloomsbury Publishing.
- Turow, J. (2021). *The voice catchers: How marketers listen in to exploit your feelings, your privacy, and your wallet*. Yale University Press.
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208. <https://doi.org/10.24908/ss.v12i2.4776>
- Weitzel, M. D. (2018). Audializing migrant bodies: Sound and security at the border. *Security Dialogue*, 49(6), 421–437. <https://doi.org/10.1177/0967010618795788>
- Wevers, R. (2018). Unmasking biometrics' biases: Facing gender, race, class and ability in biometric data collection. *TMG Journal for Media History*, 21(2), 89–105. <http://doi.org/10.18146/2213-7653.2018.368>
- Witteborn, S. (2022). Digitalization, digitization and datafication: The three D transformation of forced migration management. *Communication, Culture and Critique*, 15(2), 157–175. <https://doi-org/10.1093/ccc/tcac007>
- Woods, H. S. (2018). Asking more of Siri and Alexa: Feminine persona in service of surveillance capitalism. *Critical Studies in Media Communication*, 35(4), 334–349. <https://doi-org/10.1080/15295036.2018.1488082>
- Zijlstra, J., & Liempt, I. V. (2017). Smart (phone) travelling: Understanding the use and impact of mobile technology on irregular migration journeys. *International Journal of Migration and Border Studies*, 3(2-3), 174–191. <https://doi.org/10.1504/IJMBS.2017.083245>

About the Author

Daniel Leix Palumbo is a PhD candidate at the Department of Media and Journalism Studies at the University of Groningen, the Netherlands. His NWO-funded PhD research looks at the use of voice biometrics in asylum procedures for processes of identity construction and border control. He is also a musician who composes widely for videos and performances.