University of Cologne

Faculty of Arts and Humanities

Institute of Musicology

Master's Thesis

Primary Reviewer: Jun.-Prof. Dr. Marcus Erbe

Constructing Knowledge about Voice(s): An Exploration of the Relation between Essentialism and the Stereotyping of Social Categories

by

Katharina Makosch

Master's Programme: Musicology

Abschlussarbeit – Philosophische Fakultät

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche unter Angabe der Quelle kenntlich gemacht. Diese Arbeit ist in gleicher oder ähnlicher Form im Rahmen einer anderen Prüfung noch nicht vorgelegt.

Köln, 31.05.2025

Contents

1 Introduction

There are several meanings encompassed by the term „bias“, as exemplified by the corresponding entry in the Cambridge Dictionary: The noun „bias“ can refer to „the action of supporting or opposing a particular person or thing in an unfair way, because of allowing personal opinions to influence your judgment“, „the fact of preferring a particular subject or thing“ or „the fact of a collection of data containing more information that supports a particular opinion than you would expect to find if the collection had been made by chance“[1]. While a bias as a preference is not necessarily problematic, these definitions hint at two main issues: On the one hand, biases in research or scientific processes, such as a publication bias[2] or sampling bias[3], can affect the production of knowledge, for example in terms of which results are publicly available or by potentially leading to inaccurate generalizations. On the other hand, personal opinions influencing judgment and leading to unfair actions add an ethical dimension to these epistemic concerns. In relation to social categories, examples of biases include the undertreatment of Black patients' pain due to racist stereotypes[4] and theories of embryonic development denying active steps in the development of embryos assigned female, which reflect biased assumptions while also reproducing them[5]. Biases can also be found in the context of voice(s), as demonstrated by the Unbias AI revealing a negative correlation between vocals assigned female and high performance[6] or lesbian or gay speakers potentially experiencing avoidant discrimination due to how other people perceive their voice(s)[7].

The term „implicit bias“ is sometimes used interchangeably with the term „implicit stereotype“ in the sense that it describes „the unconscious attribution of certain characteristics to a person [… or]

---

1 Cambridge Dictionary (no date): bias. URL: https://dictionary.cambridge.org/dictionary/english/bias [accessed March 5th, 2025].

2 The term „publication bias“ refers to the fact that people in gatekeeping positions choose which submissions are published and, crucially, which ones are not. One factor may be the strength and direction of study results, which in turn may lead to researchers withholding research that does not reach statistical significance.
DeVito, Nicholas J.; Goldacre, Ben (2018): Catalogue of bias: publication bias. In: *BMJ Evidence-Based Medicine*. 24/2. p. 53.

3 The term „sampling bias“ refers to the phenomenon of recruiting participants for studies from a group of people, for example college students, that may not be representative of the broader population and thus renders generalizations of results questionable.
Worthy, L. D.; Lavigne, T.; Romero, F. (2025): Culture and Psychology. How People Shape and are Shaped By Culture. Phoenix, AZ: MMOER. p. 84.

4 Bibbins-Domingo, Kirsten (2021): Racism and Race: The Use of Race in Medicine and Implications for Health Equity. San Francisco: University of California Health Humanities Press. p. 7.

5 Voß, Heinz-Jürgen (2010): Making Sex Revisited. Dekonstruktion des Geschlechts aus biologisch-medizinischer Perspektive. Bielefeld: transcript Verlag. p. 14.

6 Kamal, Mohamed (2021): Gender Bias in the Music Industry. URL: https://unbias.co/gender-bias-in-the-music-industry/ [accessed March 5th, 2025].

7 Fasoli, Fabio et al. (2021): Stigmatization of 'gay-sounding' voices: The role of heterosexual, lesbian, and gay individuals' essentialist beliefs. In: *British Journal of Social Psychology*. 60. p. 828.

associations between certain characteristics and social categories“[8]. Elsewhere, stereotypes are conceptualized as a specific instance of biases: The Cognitive Bias Codex lists various related phenomena which include stereotyping and essentialism[9]. The relationship between stereotypes as well as prejudice and essentialist beliefs is the subject of research conducted in the field of psychological essentialism, with several studies indicating a link between these phenomena and Haslam and Whelan concluding that essentialist thinking in group perception promotes stereotyping and prejudice[10].

In this thesis, I will seek to answer the question in what respect relations between essentialism and stereotypes are present in knowledge about voice(s). I will start by working towards a theoretical framework (2), in which I will first discuss relevant terms and concepts (2.1), namely voice(s) (2.1.1), essentialism (2.1.2) and knowledge (2.1.3) as well as social categories, stereotypes and prejudice (2.1.4) and make my understanding of them explicit for the present purpose. Next, I will summarize research conducted in the domain of psychological essentialism, including how the concept has been operationalized (2.2). In order to develop a theoretical framework, I will discuss how the previously introduced terms and concepts relate to each other, clarify my conceptualization of their interdependencies and present the approach that I will use to analyze specific case studies (2.3): These case studies (3) will include academic publications (3.1) as well as examples of (in)formal lessons (3.2) to work towards a broader perspective on how knowledge about voice(s) can be produced and transmitted in relation to essentialist beliefs and stereotypes. Finally, I will summarize my results and draw conclusions regarding my research question (4).

My aim is not to give an exhaustive overview on the extent to which essentialism and stereotyping are present in knowledge about voice(s) as opposed to non-essentialist or other approaches, but to explore sites of knowledge production and transmission in relation to the present body of research in the context of my research question.

---

8 Dekanat Math.-Nat. Fakultät Heinrich Heine Universität Düsseldorf (no date): Biases in academia and education. URL: https://www.math-nat-fak.hhu.de/en/deans-office/equal-opportunity-commissioner/translate-to-english-biases-im-wissenschafts-und-bildungsbetrieb [accessed March 6th, 2025].

9 Wikimedia Commons (2018): File:Cognitive bias codex en.svg. URL: https://commons.wikimedia.org/wiki/File:Cognitive_bias_codex_en.svg [accessed March 6th, 2025].

10 Haslam, Nick; Whelan, Jennifer (2008): Human Natures: Psychological Essentialism in Thinking about Differences between People. In: *Social and Personality Psychology Compass*. 2/3. p. 1300.

2 Theoretical Framework

2.1 Terms

2.1.1 Voice(s)

Human voice(s) can be considered bodily phenomena in the sense that the production of vocal sounds results from breath and the compression of air, the vibration or oscillation of the vocal folds and the filtering of the resonance through the pharynx and physical cavities[11]. Apart from bodies, sources of vocal sounds can also include technical systems, which make recorded or digitally created voices perceivable[12]. Concerning social categories, the idea exists that certain attributes of a person, such as gender, age or origin, become audible through vocalizations[13]. In academic research on vocal production, Zimman identifies two main tendencies to approach differences in voice(s) in relation to social categories, namely determinist and constructivist perspectives: Determinist perspectives on voice(s) can be characterized by the premise that vocal output is determined by a given physiology, with differences between social categories being explained in terms of biological difference. Constructivist perspectives can be characterized by the premise that people have agency when it comes to vocalizations, with differences between voice(s) in relation to social categories being explained in terms of social learning, thus taking voice(s) as a practice that can differ over time[14]. In addition to the aspect of vocal production, the aspect of listening, or vocal perception, is also relevant concerning the phenomenon of voice(s)[15]. In this context, there are determinist and constructivist approaches in relation to social categories as well: Determinist perspectives concerning vocal perception can be characterized by presenting a direct link between the physical characteristics of a vocalizer reflected in their vocalizations and the listener's perception of information concerning social categories, such as gender or age[16]. In a constructivist vein, Eidsheim argues that the process of listening is an active one: While „voice is thought to reveal the true nature of the body“[17], it is not innate but social and cultural in the sense that the collective influences

---

11 Sundberg, Johan et al. (2021): One Singer, Two Voices. In: *Acoustical Society of America*. 17/1. p. 43.

12 Frühholz, Sascha; Belin, Pascal (2018): The Science of Voice Perception. In: *The Oxford Handbook of Voice Perception*. S. Frühholz / P. Belin (eds.). Oxford: Oxford University Press. p. 3.

13 Schrödl, Jenny; Kolesch, Doris (2018): Stimme. In: *Handbuch Sound*. D. Morat / H. Ziemer (eds.). Stuttgart: J. B. Metzler. p. 223.

14 While Zimman focuses on the category gender, I would argue that determinist and constructivist perspectives are not exclusive to this category. I will elaborate on this point in the further course of this thesis.
Zimman, Lal (2018): Transgender Voices: Insights on Identity, Embodiment, and the Gender of the Voice. In: *Language and Linguistics Compass*. 12/8. p. 2f.

15 Frühholz; Belin (2018): The Science of Voice Perception. p. 4.

16 see for example Latinus, Marianne; Zäske, Romi (2019): Perceptual Correlates and Cerebral Representation of Voices - Identity, Gender, and Age. In: *The Oxford Handbook of Voice Perception*. S. Frühholz / P. Belin (eds.). Oxford: Oxford University Press. p. 561f.

17 Eidsheim, Nina Sun (2019): The Race of Sound. Listening, Timbre, and Vocality in African American Music. Durham, London: Duke University Press. p. 17.

which select areas of vocal potential a vocalizer attends to and collective pressure and encouragement can be seen as the decisive factor in honing each voice's potentiality in the context of the vocalizer's position within the collective[18]. Listening produces meaning actively in the sense that through listening, people, including vocalizers listening to themselves, name and define what or whom they hear, which leads Eidsheim to conclude that „[T]he voice does not arise solely from the vocalizer; it is created just as much within the process of listening."[19] Constructivistically speaking, voice(s) can therefore be conceptualized in relation to a given collective or sociocultural context, with vocal production as well as perception explainable in terms of social learning, which can take place in formal as well as informal lessons, such as everyday socialization[20].

Thus, the idea that attributes of a person become audible through their vocalizations can be interpreted both through a determinist and a constructivist lens: Deterministically speaking, categories like gender or age have biological or physical correlates that determine vocal output, through which they can be perceived by other people. Constructivistically speaking, there are social norms in relation to which people learn vocal patterns as bodily acts, and vocalizers as well as other people can actively perceive and apply learned associations with social categories. Both approaches identify bodies as sources of vocal production but while determinism implies a certain necessity about the relationship between category membership, vocal output and its perception, constructivism puts a focus on agency and implies that vocalizations could also be performed and listened to in another way. I would argue that the aspect of agency vocalizers have over their vocalizations must not be confused with immediate access to all potentialities encompassed by their vocal range. Instead, vocalizations can be described in the context of vocal habits, implying agency in the potential to change present or to incorporate new habits[21]. Depending on the perspective taken, the term „voice" can therefore refer to multiple aspects of the phenomenon, such as the vocal tract, audible vocalizations, specific sonic qualities (e.g. pitch, formants, volume) or other factors that can be present in vocalizations (e.g. language, word choice).

While there are approaches that combine determinist and constructivist perspectives, for example by putting a focus on one of these perspectives while also mentioning elements of the other[22], constructivism challenges the idea of a direct determinism that denies or ignores sociocultural

---

18 Ibid., p. 24, 11.
19 Ibid., p. 24.
20 Ibid., p. 23; 41.
21 Ibid., p. 41.
22 See Latinus; Zäske (2019): Perceptual Correlates and Cerebral Representation of Voices - Identity, Gender, and Age. p. 570.; Frühholz; Belin (2018): The Science of Voice Perception. p. 9.

influences on voice(s): Zimman argues that „[t]he fact that speakers can learn to alter the perceived gender of their voice – including features that are known to be influenced by physiology like F0 [or a speaker's pitch] – undermines the assumption that vocal anatomy is the direct, determining cause of gender differences in the voice.“[23] Additionally, multimodal or context information can influence auditory perceptions: People can make auditory judgments biased by visual cues and vice versa[24], lending credibility to the argument that listening to voice(s) is an active process as opposed to an objective reception of information about a vocalizer.

I use the term voice(s) to encompass both vocal production and vocal perception as well as their possible mutual influences. Following Zimman, I conceptualize vocal production as physiologically determined in the sense that a given body determines an individual frame in which vocalizations are possible, while recognizing that specific vocalizations only represent an instance of the vocal range a person is capable of[25]. Following Eidsheim, I conceptualize vocal perception or listening as an active process that produces meaning and has the potential to influence vocal production. In this line of thought, I use the term „voice(s)“ to draw attention to the fact that 'the voice' of any individual could be used in more ways than are audible in any specific instance, while also refraining from rendering actual perceivable phenomena abstract in the sense that you can never perceive „the“ (human) voice, just (human) voices. For specific audible uses of voice(s), I will use the term „vocalizations“.

2.1.2 Essentialism

The term „essentialism“ denotates a philosophical concept that can be traced back to philosophers such as Aristotle and Plato and has since been the subject of an ongoing discourse[26]. While there have been many different accounts of essentialism over the years[27], I will try to capture the most widespread ideas about essentialism and won't do justice to every single conception of it that has been proposed within the discourse given that within the context of my thesis, I will focus on essentialism concerning the relationship between social categories and individual members of those categories in relation to the production of knowledge.

---

23 Zimman (2018): Transgender Voices. p. 10.

24 Marchand Knight, Jay; Sares, Anastasia G.; Deroche, Mickael L. D. (2023): Visual biases in evaluation of speakers' and singers' voice type by cis and trans listeners. In: *Frontiers in Psychology*. 14:1046672. p. 3f.

25 Zimman (2018): Transgender Voices. p. 6.

26 Koslicki, Kathrin; Raven, Michael J. (2024): Introduction. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. p. 2f.

27 Ibid., p. 3.

Essentialism is concerned with the idea of essences, which can be characterized as the totality of all essential, as opposed to accidental, properties of an object[28]. Said „object“ can refer to either an individual or a general category or (natural) kind, with the essence of a kind being shared by its members. The concept of a natural kind implies that the grouping of individuals reflects the structure of the natural world rather than the interests and actions of human beings[29]. In this context, essences are taken to exist objectively and mind-independently, with humans being in the position to know or discover them[30]. For a property or feature of an entity to be essential, it has to be a property or feature of that entity for it to exist; other properties can be described as accidental[31]. In this sense, the concept of essences is connected to the concept of necessity: „If it is essential to x that it is F, then it is necessary that x is F.“[32] However, a necessary or essential property or trait is not the same as a sufficient definition or a full statement of essence given that an essence is conceptualized as the totality of all essential traits[33]. On the other hand, it is possible for multiple distinct kinds to share accidental traits without sharing the same essence. There are several implications to the concept of an essence: When an essence is conceptualized as the totality of all essential traits or features of an entity, that entity has these traits or features necessarily, which means that they can not change without the entity ceasing to exist as that entity. In other words, essences are taken to be a criterion of identity, in that they make an entity or members of a kind what they are, which in turn implies that it is possible to draw clear distinctions between different kinds and to conceive of members of a kind as homogenous in regard to their essence. Given that essential traits or features are necessary and unchangeable, they can be said to be inherent or intrinsic to an entity or kind[34]. Additionally, the assumption of essences allows for logical inferences based on group membership: If group A has an essential trait T and the individual X is a member of A, then X necessarily possesses trait T. In other words, if all members of group A have an essential trait T, every single individual that belongs to that group must have trait T. However, given that distinct kinds can share accidental traits without sharing the same essence, it is not logically valid to infer from the observation of trait T that an individual X is a member of A without conceptualizing T as essential. Furthermore, essence is said to play a definitional role as the statement of an essence

---

28 Britannica (2025): essentialism – philosophy. URL: https://www.britannica.com/topic/essentialism-philosophy [accessed March 19th, 2025].

29 Bird, Alexander; Tobin, Emma (2024): Natural Kinds. In: *The Stanford Encyclopedia of Philosophy*. E. N. Zalta / U. Nodelman (eds.). URL: https://plato.stanford.edu/entries/natural-kinds/index.html [accessed March 25th, 2025].

30 Mallozzi Antonella (2024): The Epistemology of Essence. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. p. 211.

31 Koslicki; Raven (2024): Introduction. p. 1.

32 Mallozzi (2024): The Epistemology of Essence. p. 210.

33 Koslicki; Raven (2024): Introduction. p. 7.

34 Kampourakis, Kostas (2023): Ancestry Reimagined: Dismantling the Myth of Genetic Ethnicities. New York: Oxford University Press. p. 31.

or a full account of essential properties of an object or kind can be said to define what said object or kind is[35]. Another function that has traditionally been assigned to essences is the aspect of explanation, which can be summed up by the idea that the essence of a given entity can be used to explain certain other properties and behaviours of that entity, such as accidental properties.

Within the philosophical discourse on essentialism that continues to this day, there have been several objections to essentialism[36]: Hallett questions the conditions of knowing about essences: „If an essence, to qualify as an essence of X, must be present in all X's, then there is no telling from any single member of the class that the essence present in it is the essence of X.“[37] Accordingly, one can either look at all Xs and actually find or empirically validate a common essence, or one must declare a common essence, either as a generalization of some findings or by restricting the term X to individuals that conform to a postulated essence[38]. Hallett argues that it makes sense to agree on a specific terminology and conventional stipulations in the context of scientific endeavours. However, it is less sensible to adopt this uniform way of thinking when it comes to concepts or words used in non-scientific communication[39]: „Essences in the traditional sense are core properties or clusters of properties present, necessarily, in all and only those things which bear the common name. Knowledge is one thing; language is one thing; beauty, meaning, humanity, life, law, justice – each is a single, invariant reality[.]“[40] Given the assumption that if there is one word to describe a class, that word must imply or describe a shared reality, one criticism of essentialism in philosophy concerns the disregard for established word-uses[41]: „Freedom to disregard usage is critical for those who would philosophize in the essentialistic manner. It permits them to eliminate rivals; it permits them to construct their own systems. If one can simply declare what events are, what actions are, what intentions are, and so forth, without regard for how these words are used, one may proceed [...] with apodictic assurance.“[42] In short, to define an essence of a class or kind, one must name this group and thus defines it according to the meaning of the term. The meaning of a term, however, can vary with the use of it. This results in the possibility that other uses of the same term can be dismissed by claiming that they are not genuine instances of the term in the same sense of the

---

35 Torza, Alessandro (2024): Modal Conceptions of Essence. In: *The Routledge Handbook of Essence in Philosophy.* K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. p. 111.
36 Koslicki; Raven (2024): Introduction. p. 3.
37 Ibid., p. 9.
38 See Hallett, Garth L. (1991): Essentialism. A Wittgensteinian Critique. Albany: State University of New York Press. p. 92.
39 Ibid., p. 139.
40 Ibid., p. 2.
41 Ibid., p. 158, 179.
42 Ibid., p. 135f.

word[43]. Thus, essentialism can reflect biases in the sense that „[E]ssences may reflect the essentialistic preferences of those who speak of them [...]“[44]. Additionally, Hallett sees no reason to assume the existence of essences because even if one were to find some set of properties that all members of a class share, it could be described as something all members of that class have in common without implying statements of necessity[45].

While there have been a number of explicitly essentialist theories, essentialist thinking without explicit declarations of essences or necessity can be said to have been a prominent feature of Western thought[46]. Tracing its origin back to philosophical proposals from over thousands of years ago, Hallett describes how essentialist thinking had an influence on thinkers both within the discipline of philosophy and other disciplines such as natural sciences. In this sense, the author describes essentialism as contagious and argues that while essentialist thinking has led to success in the context of natural sciences, this does not mean that essentialist thinking will be equally successful in other contexts[47].

To summarize, I conceptualize essentialism as the assumption of essences, which are the totality of all necessary or essential properties, features or traits that make an entity what it is and can therefore be used as a sufficient definition. Essential properties are necessary, unchangeable and allow for logical inferences in the context of group membership: If an essence or essential trait of a group is known, it is necessary that every single member of that group will possess this essence or essential trait. Accordingly, it is possible to infer from the information that an individual is part of a certain group that that individual will share in all essential features of that group. Essences of natural kinds are taken to exist objectively and mind-independently. Furthermore, essences have been assigned an explanatory function in that they can be used as a reason to explain further properties or behaviours of an entity that follow from an essence.

Potential issues concerning essentialist thinking include (1) the conditions of knowing about an essence in the sense that statements of essences about classes, kinds or groups are necessarily general statements about every single individual encompassed by that category and without having confirmed a declared essence in every single individual one must generalize or restrict the use of the term or the category's name to instances that fit the description, (2) limiting a category according to

[43] Ibid., p. 22.
[44] Ibid., p. 23.
[45] Ibid., p. 128.
[46] Ibid., p. 28f.
[47] Ibid., p. 138f.

the way one interprets the meaning of its name while ignoring other word-uses and (3) that the assumption of essences in terms of necessary and sufficient conditions does not necessarily follow from describing a class, kind or group according to what its members have in common. While issues (1) and (2) can reflect personal biases, (3) displays a bias for essentialism as opposed to other ways of reasoning or explaining. Furthermore, I will not restrict the label „essentialist" to theories or arguments explicitly stating essences but include the possibility that they can be implicitly essentialist in their ways of reasoning.

2.1.3 Knowledge

A common definition of the term „knowledge" holds that „knowledge is *justified true belief*"[48] (emphasis in the original), which means that the term „knowledge" can only be applied if the belief someone holds is in fact true. However, it is possible for people to believe they know something is the case or conforms to reality even though it is not or does not, for example in the context of biases. Nevertheless, the aspect of truth can be described as a normative component of knowledge[49], which is why I will take the term „knowledge" to include cases in which people believe something to be true, independently of whether it actually is. While knowledge is sometimes assumed to be objective, there have also been approaches to characterize knowledge as constructed: Hall argues that knowledge about a topic is represented in a discourse that constructs a topic in a certain way through the use of language. In this context, facts or beliefs held can be (re)presented in different ways[50]. Furthermore, the production and transmission of knowledge can be described as a social process in the way that knowledge is produced in the context of other people in respective communities as well as cooperating circumstances[51]. One element present in the discursive practice of knowledge production is the context of power relations, as the knowledge a discourse produces can be said to be power exercised over the subjects, or people that are known, by those who know. Additionally, if people believe something they know is true they can act on their beliefs, with their actions having real consequences, independently of whether the beliefs in question are true or not[52]. However, the claim that knowledge is constructed does not entail that it can not be true or conform to reality: In a phenomenological context, Husserl argues that sciences and scientific methods result

---

48 Trinkaus Zagzebski, Linda (2020): Epistemic Values: Collected Papers in Epistemology. New York: Oxford University Press. p. 21.
49 Ibid., p. 27.
50 Hall, Stuart (1992): The West and the Rest: Discourse and Power. In: *Formations of Modernity*. S. Hall / B. Gieben (eds.). Cambridge, Oxford: The Open University. p. 291f.
51 Trinkaus Zagzebski (2020): Epistemic Values. p. 32.
52 Hall (1992): The West and the Rest. pp. 293ff.

from human work and therefore take the status of cultural acquisitions. For ideas and scientific models to be communicated, they need to take a form that can be sensually perceived, for example via writing or spoken language. In this sense, one can never experience, communicate or transmit objectivity itself, only representations of knowledge and ideas[53]. While Husserl does not deny that there is one world experiencing subjects are all part of[54], the very act of taking the methods of (natural) sciences to stand in for an objective reality can be termed an active act[55], or a construction. Crucially, by calling knowledge constructed instead of objective, and implying that there is no unmediated access to the objective world, nothing has been said about whether any specific knowledge in question conforms to reality or not.

In the context of knowledge production and transmission, universities and academic practices play a crucial role in defining what counts as scientific knowledge through conducting research, teaching and awarding academic degrees[56]. While research can be and is conducted outside of universities as well, scientific knowledge is sometimes presented as a special or superior kind of knowledge in the sense that it is taken to be an accurate or true account of reality, which leads to the conception of academics as experts in their respective fields whose arguments should be given priority[57]. This account is reflected in distinctions between different kinds of knowledge, such as scientific as opposed to everyday knowledge[58] or codified or explicit knowledge, which can be characterized by being written down, as opposed to tacit knowledge, which can be characterized as being based in experience[59]. While these distinctions can be criticized, for example in terms of what counts as „written down“, the question inhowfar codified knowledge is (not) based in experience and whether it makes sense to treat knowledge in different contexts as different kinds of knowledge[60], they point to multiple possible sites of knowledge production and transmission.

In this thesis, I will use the term „knowledge“ for cases in which people hold beliefs or produce and communicate or transmit information that they believe to be true, independently of whether the information actually conforms to reality or not. In both cases, acting in accordance with one's

[53] Husserl, Edmund (2012): Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. E. Ströker (ed.). Hamburg: Felix Meiner Verlag. p. 26.
[54] Ibid., p. 23.
[55] Ibid., p. 55.
[56] Biesta, Gert (2007): Towards the knowledge democracy? Knowledge production and the civic role of the university. In: *Studies in Philosophy and Education*. 26/5. p. 468f.
[57] Ibid., pp. 471ff.
[58] Ibid., p. 475.
[59] Carayannis, Elias G.; Campbell, David F. J. (2019): Definition of Key Terms: Knowledge, Knowledge Production, Innovation, Democracy, and Governance. In: *Smart Quintuple Helix Innovation Systems. SpringerBriefs in Business.* Cham: Springer. p. 5.
[60] See Biesta (2007): Towards the knowledge democracy? p. 475.

knowledge has consequences. I view knowledge as actively produced or constructed and transmitted in relation to social processes. By claiming that knowledge is constructed, I do not mean to say that it can not be true but merely state that the way knowledge is (re)presented in a discourse depends on sensual experiences of subjects and could be realised in different ways. Scientific or academic as well as non-academic contexts can be described as possible sites of knowledge production and transmission, which have the potential to be influenced by biases and power relations.

2.1.4 Social Categories, Stereotypes, and Prejudice

Social categories are classifications of people based on shared characteristics, for example in the context of age, gender, „race“[61], socioeconomic status or occupation[62]. For terminological clarity, I understand concepts such as gender as a dimension of social categories, while the categories themselves could include terms like „man“, „woman“ or „non-binary“ in relation to this dimension. Banton describes the act or process of categorizing people as follows: There is an assumption that human individuals have distinct characteristics. The attribution of significance to these distinct characteristics results in the creation of social categories because individuals share these characteristics with other individuals. Social categories themselves and the significance attributed to any particular characteristic depends on sociocultural contexts and can therefore be subject to change[63]. For example, Aristotle's ethics distinguished humans according to rationality, claiming that slaves are distinct from fully rational agents because their souls lack a governing rational component[64]. Furthermore, the familiar demographic environment of individuals creates normative ideas in relation to social categories. Characteristics shared between individuals can become the basis of collective action, for example in the context of defending privileges or identity politics. One possible strategy to uphold power relations can be to justify them by stressing the idea of biological differences between social categories[65].

Stereotypes have been defined as cognitive biases that build up prejudice and can lead to

---

61 I will explain how I understand and employ the term „race“ in 2.3.

62 McKee, Adam; Bransford, Scott (no date): social categories. Definition. URL: https://docmckee.com/oer/soc/sociology-glossary/social-categories-definition/ [accessed March 27th, 2025]

63 Banton, Michael (2011): A Theory of Social Categories. In: *Sociology*. 45/2. p. 189ff.

64 Grim, Patrick; Rescher, Nicholas (2023): Theory of Categories. Key Instruments of Human Understanding. London, New York: Anthem Press. p. 138.

65 Banton (2011): A Theory of Social Categories. p. 190f.

discriminatory behaviour[66]. More specifically, stereotypes can be conceptualized as the cognitive representation of a group in respect to the attributes, features and behaviours members of that group are thought to share, with stereotypes being possible both towards a group a given individual is a member of (ingroup) and a group that that individual is not a member of (outgroup)[67]. Stereotypes build up prejudice in the sense that prejudice can be defined as the positive or negative evaluation of a social group and its members based on the characteristics attributed to it. Stereotypes and prejudices can lead to discriminatory behaviours, which can be defined as the rejection or disadvantaging of individuals based on their group membership[68]. Possible functions of stereotypes that have been proposed in social psychology include a positive distinctness of an individual's ingroup as opposed to outgroups, the causal explanation of social phenomena and the justification of behaviours towards outgroups. In this context, social categories can accentuate perceived similarities and differences, as differences within a given category tend to be underestimated while differences over categories tend to be overestimated[69]. Stereotypes can have consequences such as the phenomenon of stereotype threat: „For example, where the cultural stereotype 'Women are bad at math' is known to be commonly held, women are likely to underperform at mathematical activities when their gender is made salient."[70]

Grim and Rescher point out several possible errors in relation to the use of social categories, which they explain as typical errors of stereotypes: One of these errors is the issue of oversimplification in regard to causal explanations of social phenomena, which the authors illustrate in the context of inequality: „The social causes and consequences of inequality in our society are almost undoubtedly complex, with tangled conglomerates of correlations and categories. [...] Search for a single simple explanatory category is bound to fail in our attempts to understand and address that complexity. All too often the error of oversimplification involves latching on to a very specific category – minority status, ethnicity, or race – as "the" explanation, with "true generalizations" offered as if they supported [a] fallacious inference."[71] A related possible error in terms of explaning social phenomena can be to make use of an irrelevant or wrong category. Furthermore, the act of attributing characteristics to individuals on the basis of their perceived group membership can be

66 Employment and Disability Project handicap international (2020): Stereotypes, cognitive biases and discrimination. Factsheet. URL: https://www.hi.org/sn_uploads/MEMO-3_Stereotypes-cognitive-biases-and-discrimination.pdf [accessed March 28th, 2025].

67 Stürmer, Stefan; Siem, Birte (2022): Sozialpsychologie der Gruppe. 3rd edition. München: Ernst Reinhardt Verlag. p. 52f.

68 Ibid., p. 53.

69 Ibid., p. 56ff.

70 Wodak, Daniel et al. (2015): What a Loaded Generalization: Generics and Social Cognition. In: *Philosophy Compass*. 10/9. p. 628.

71 Grim; Rescher (2023): Theory of Categories. p. 133.

subject to a false attribution, or the presupposition of facts that do not conform to reality. Another possible error is the aspect of overgeneralization on the basis of stereotypes, which results in the possibility to think of individuals incorrectly as well as unjustly[72]. Grim and Rescher summarize that „Categorical misapplications with regard to the natural world can result in failed attempts to predict, explain, understand, or intervene in the course of nature. Categorical misapplications in the social realm can carry an additional burden of ethical mistake and negative social consequences.“[73]

Additionally, social categories intersect in the way that every human is assigned to and most likely identifies with multiple social categories, which renders it a simplification to think of social categories as genuinely distinct[74]. In relation to stereotypes, prejudice and discrimination, this idea points to the concept of intersectionality, which is „basically a lens, a prism, for seeing the way in which various forms of inequality often operate together and exacerbate each other. We tend to talk about race inequality as separate from inequality based on gender, class, sexuality or immigrant status. What’s often missing is how some people are subject to all of these, and the experience is not just the sum of its parts.“[75]

In summary, it can be said that social categories are classifications of people based on shared characteristics. Both the relevance of social categories themselves and the significance attributed to certain characteristics in the process of categorizing can be called constructed in the sense that they depend on sociocultural contexts and can therefore be subject to change. In the context of power relations, shared characteristics can become the basis for collective action and the act of stressing biological differences between social categories can be used as a strategy to justify an unequal status quo. I understand stereotypes as the cognitive representation of a group in respect to the attributes, features and behaviours members of that group are thought to share. It is possible to have stereotypes about one's ingroup as well as outgroups. Additionally, they can have consequences such as impairing individuals' performances in relation to their content. Prejudice adds the dimension of an evaluation to stereotypical preconceptions, therefore depends on stereotypes and can lead to discriminatory behaviours. Possible errors in the use of stereotypes in connection with social categories include oversimplification, the use of a wrong or irrelevant category to explain social phenomena, false attribution and overgeneralization. These errors can result in both epistemological and ethical issues. Furthermore, as individuals can be assigned to and identify with

[72] Ibid., p. 133ff.
[73] Ibid., p. 131.
[74] Banton (2011): A Theory of Social Categories. p. 199.
[75] Steinmetz, Katy (2020): She Coined the Term ‘Intersectionality’ Over 30 Years Ago. Here’s What It Means to Her Today. URL: https://time.com/5786710/kimberle-crenshaw-intersectionality/ [accessed March 28th, 2025].

multiple social categories, stereotypes and discrimination can work in intersectional ways and take specific forms depending on the interplay of multiple social categories.

2.2 Psychological Essentialism

There has been a growing body of research on essentialist notions in the context of social categories since the second half of the 20th century[76]: In 1989, Medin and Ortony coined the term „psychological essentialism" to describe the phenomenon that laypeople hold beliefs that many categories have essences[77], or beliefs that members of a category share deep-seated properties that determine their identity[78]. Concerning social categories, psychological essentialism has been described as a pervasive cognitive bias[79]. According to a study from 2000, the social categories that were essentialized most were gender, „race" and ethnicity as well as physical disability categories, which can be interpreted as being rooted in an assumed physical reality and biological theories of category origin[80]. Nevertheless, it is not necessary for people to define what a specific essence entails to believe that there is one[81]. Much of the research conducted in the field of psychological essentialism is concerned with the relationship between essentialist beliefs and stereotypes as well as prejudice towards social categories. In the case of stereotypes, people who essentialize human attributes have been found to be especially likely to endorse stereotypes about social categories independently of other measured predictors of stereotype endorsement[82]. While more evidence has been accumulating that essentialist beliefs about social categories are connected to the stereotyping of and prejudice towards those specific categories in a way that promotes the latter two[83], there are varying conclusions as to whether this connection is a causal one or a correlation[84].

A definitive claim as to the type of connection between essentialist beliefs and stereotypes or prejudice is further complicated by the possibility of its dependency on the specific social categories under scrutiny: In the case of gender, for example, the essentializing of categories can promote

---

76 Haslam; Whelan (2008): Human Natures. p. 1297f..
77 Haslam, Nick et al. (2000): Essentialist beliefs about social categories. In: *British Journal of Social Psychology*. 39. p. 114.
78 Haslam; Whelan (2008): Human Natures. p. 1297.
79 Rhodes, Marjorie et al. (2012): Cultural transmission of social essentialism. In: *PNAS*. 109/34. p. 13526.
80 Prentice, Deborah A.; Miller, Dale T. (2007): Psychological Essentialism of Human Categories. In: *Current Directions in Psychological Science*. 16/4. pp. 202, 205.
81 Haslam; Whelan (2008): Human Natures. p. 1297.
82 Ibid., p. 1299.
83 Ibid., p. 1300.
84 Glazier, Jessica J. et al. (2021): The Association Between Prejudice Toward and Essentialist Beliefs About Transgender People. In: *Collabra: Psychology*. 7/1. p. 2.

inferences and generalizations about how people of a certain gender differ from each other in a way that is assumed to be stable, and support for these notions has been found to be linked to sexist attitudes[85]. On the other hand, in the case of sexual orientation, biologically based essentialist notions have the potential to invalidate views of non-heterosexuality as immoral or sinful because sexual orientation is not conceptualized as alterable or chosen[86]. So while essentialism can contribute to stereotypes, prejudice and discriminatory behaviours, it also has the potential to further acceptance. Another complication is the possible dependency on other factors, such as social status: A study from 2009 found positive associations between essentialism and sexism but only found a link between essentialism and prejudice when the dominant group of participants felt that their privileges were threatened by the prospect of social change[87] and there has been additional evidence that essentialist thinking can serve a justifying function in relation to systems or hierarchies by making existing status inequalities seem natural and inevitable[88].

Additionally, there is no definitive consensus about the operationalization of essentialism, as different studies employ different criteria: Haslam and Whelan point out that „[r]esearchers may often be claiming to find essentialist thinking on the basis of very partial evidence (e.g., evidence pertaining to a single component of essentialism such as perceived inalterability or inductive potential)[. …] Inferring that a group has an internal property is not in itself essentialist; it only becomes so when that property is seen as unchanging, deep-seated, natural, and responsible for the group's observable attributes."[89] One proposition about the structure of essentialist beliefs includes 8 factors, namely that (1) membership in a category is fixed or immutable, (2) there is an inherent nature underlying surface characteristics, (3) a category is discrete and has defining features, (4) category members are uniform, (5) the category is natural, (6) the category is informative about its members and thus allows for inductive potential, (7) the category is identity-determining and thus excludes members from other categories and (8) the category exists independently of human language or societies and is therefore historically invariant[90]. Haslam et al. have proposed that these factors can be considered in terms of two dimensions: Factors (1), (3), (5), and (8) are said to correspond to the philosophical concept of a natural kind, while factors (2), (4), (6) and (7) are said to correspond to the concept of entitativity, which can be described as „the extent to which a group

85 Haslam; Whelan (2008): Human Natures. p. 1302.
86 Ibid., p. 1304.
87 Morton, Thomas A. et al. (2009): Theorizing Gender in the Face of Social Change: Is There Anything Essential About Essentialism? In: *Journal of Personality and Social Psychology*. 96/3. p. 663.
88 Haslam; Whelan (2008): Human Natures. p. 1303.
89 Ibid., p. 1308.
90 Haslam, Nick et al. (2002): Are essentialist beliefs associated with prejudice? In: *British Journal of Social Psychology*. 41. p. 88f.

is understood to be a meaningful unit with deep commonalities“[91]. This distinction led the authors to conclude that in their study from 2002, no prejudice measure relating to sexism, racism and anti-gay attitudes was associated with the natural kind factor, whereas these measures were only strongly associated with the entitativity factor in relation to anti-gay attitudes[92]. However, this points to another complication concerning the comparability of studies in the field of psychological essentialism, namely the measures of stereotypes and prejudice: Haslam et al. used the Attitudes to Lesbians and Gay Men scale to measure anti-gay prejudice, which focuses on a single „condemnation – tolerance“ factor. In terms of sexism and racism, they employed three scales each: Concerning racism, they used (1) the Modern Racism scale, which is made up of three components, namely denial of continuing discrimination, antagonism to African-Americans' demands, and resentment about special favours for African-Americans, (2) the Old-Fashioned Racism scale, which assesses prejudice towards African-Americans in terms of objections to intermarriage, integration and fair housing laws and (3) the Racial Attitudes questionnaire, which contains items concerning acknowledgment of continuing discrimination and support for greater social, economic and cultural recognition of African-Americans as well as attitudes towards the work ethic, drive, social organization and values of African-Americans. Concerning gender, the authors used (1) the Modern Sexism scale, which is comprised of the components denial of continuing sexism, antagonism to women's demands and resentment of special favours towards women, (2) the Old-Fashioned Sexism scale, which measures endorsement of restrictively traditional gender roles and lesser female intelligence and (3) the Ambivalent Sexism Inventory to assess hostile as well as „benevolent“ or paternalistic forms of sexism[93]. While these measures can be criticized in and of themselves, for example for equating racism with prejudice towards African-Americans, it is also questionable in how far they measure comparable forms of prejudice: While anti-gay attitudes are measured in terms of a „condemnation – tolerance“ factor, the scales concerning sexism and racism largely point to political demands. Given that the authors found that Black people and women received relatively high ratings on the natural kind factors, whereas gay men received higher ratings in terms of entitativity, leading them to conclude that these categories are essentialized in different ways[94], it is unclear how the conception of categories as a natural kind is supposed to be related to prejudice in terms of the evaluation of political demands as opposed to prejudices that directly relate to the natural aspect of that conception. Furthermore, other studies have used different measures, for example the ATB scale to measure explicit racism as well as the IAT to measure

[91] Ibid., p. 89.
[92] Ibid., p. 93.
[93] Ibid., p. 90f.
[94] Ibid., p. 92.

implicit racism[95] or avoidant discrimination scales[96], which makes the comparability of different studies questionable.

Despite these complications there is relative consensus that essentialist thinking in group perception is a widespread phenomenon[97]. One reason that has been proposed to play a part in this is the use of generic language: Cimpian and Markman have found that both pre-school children and adults provided more essentialist explanations for biological properties as well as abilities when novel properties were provided in generic sentences as opposed to non-generic sentences, with this effect being stronger in relation to biological properties[98]. Rhodes et al. found that hearing generic sentences about a novel social category led 4-year-old children as well as adults to develop essentialist beliefs about that category, while experimentally inducing parents to hold essentialist beliefs about a novel social category also led them to use more generic language when they talked about the category with their children, thus leading the authors to conclude that generic language serves as a mechanism by which essentialist beliefs can be culturally transmitted[99].

Sentences like „Boys are good at math." can be considered category-referring generic sentences, which imply a broader set of referents than nongeneric sentences such as „This boy is good at math."[100] In this sense, generic language is „more likely to imply that the information it conveys is *essential* […] or *central* to the identity of its referents than an equivalent nongeneric sentence."[101] (emphasis in the original). Concerning social categories, generic language, such as the previous example, has been described as a common vehicle for stereotypes. Wodak et al. differentiate between descriptive and normative generics, with normative generics not necessarily being grounded in an assumed inherent nature but in ideals concerning the category in question, for example „Scientists care about truth."[102] Both versions can be said to refer to a category as an abstract whole. It has been found that more abstract descriptions lead to more essentialist construals[103]. However, while universally quantified sentences, such as „All boys are good at

[95] Mandalaywala, Tara M. et al. (2018): Essentialism Promotes Racial Prejudice by Increasing Endorsement of Social Hierarchies. In: *Social Psychological and Personality Science*. 9/4. p. 462.
[96] Fasoli et al. (2021): Stigmatization of 'gay-sounding' voices. p. 831.
[97] Haslam; Whelan (2008): Human Natures. p. 1302.; Haslam et al. (2002): Are essentialist beliefs associated with prejudice? p. 88.
[98] Cimpian, Andrei; Markman, Ellen M. (2011): The Generic/Nongeneric Distinction Influences How Children Interpret New Information About Social Others. In: *Child Development*. 82/2. p. 487.
[99] Rhodes (2012): Cultural transmission of social essentialism. p. 13526ff.
[100] Cimpian; Markman (2011): The Generic/Nongeneric Distinction Influences How Children Interpret New Information About Social Others. p. 471.
[101] Ibid.
[102] Wodak et al. (2015): What a Loaded Generalization. p. 628.
[103] Cimpian; Markman (2011): The Generic/Nongeneric Distinction Influences How Children Interpret New Information About Social Others. p. 487ff.

math.“, can be falsified by counterexamples, this is not necessarily the case for generic sentences[104]. As generic statements lack explicit quantification (e.g. all, some, most), it is open to the interpretation of listeners or readers whether any sentence was meant as a universal statement or not, and, in turn, in how far such statements are taken as the attribution of essential properties. In this context, Wodak et al. make the point that generics without explicit universal quantification leave room for retractions: „Say a speaker asserts that ‘Latinos are lazy.’ You could respond by presenting an onslaught of counterexamples, but then the speaker can accept that ‘although many Latinos aren’t lazy, they tend to be — thus embracing the characteristic generic’ [...]. Alternatively, you could respond by arguing that ‘Latinos show no greater tendency towards laziness than any other group. The speaker can then suggest that, although it is not part of the nature or essence of Latinos to be lazy, most are’ [...]. Or say a speaker asserts that ‘Women are submissive.’ This could be intended as a descriptive generic; but even if one shows that submissiveness is not a common, striking or characteristic property of women, the speaker could accept this and suggest it is an ideal of womanhood. […] This slide back and forth between different interpretations of the utterance allows speakers to avoid taking responsibility for the implications of their claims [...].“[105] Additionally, there is some evidence suggesting that „generics are by default understood as characteristic property generics – that is, as being true in virtue of the intrinsic nature of the kind“[106] or that generic language can contribute to essentialist beliefs by this option merely being considered[107].

While most research in the field of psychological essentialism has focused on the essentializing of social categories themselves and how they relate to prejudice and stereotypes, there is a smaller line of research concerning whether and which traits are thought of as essences[108]. In this context, Fasoli et al. have conducted a study about the relationship between essentialist beliefs about voice(s) as a cue to social categories concerning sexual orientation and the perpetration and experience of stigma[109]. The authors conceptualize voice(s) as a trait that triggers social categorization and can therefore be subject to essentialist beliefs in a category-trait association[110]. I will review this study in more detail in 3.1.1.1.

---

104 Ibid., p. 472.
105 Wodak et al. (2015): What a Loaded Generalization. p. 630f.
106 Ibid., p. 627.
107 Ibid., p. 631.
108 Haslam; Whelan (2008): Human Natures. p. 1306.
109 Fasoli et al. (2021): Stigmatization of 'gay-sounding' voices. p. 826.
110 Ibid., p. 841.

In summary, it can be said that the field of psychological essentialism is concerned with people's beliefs about whether social categories have essences, or whether their members share inherent properties that determine their identity. Psychological essentialism has been described as a cognitive bias. One focus of research has been the relationship between essentialist beliefs and stereotypes and prejudice. While evidence has been accumulating that there is a connection between the two, there are several aspects complicating definitive claims about this relation: Among these aspects are varying ways of measuring both essentialist beliefs and stereotypes or prejudice, doubts concerning the validity of the measures employed and possible dependencies on which social categories are the subject of interest as well as other factors contextualizing the relationship between essentialist beliefs and stereotypes or prejudice, such as social status or endorsement of the status quo. I will discuss how I will proceed concerning measures of essentialism and stereotypes in greater detail in 2.3 in order to develop criteria that I will use to analyze case studies. Apart from research concerning essentialist beliefs about social categories themselves, there is also research on category-trait associations, which includes the possibility that traits, such as voice(s), can be subject to essentialist beliefs as well. With relative agreement that it is not uncommon to essentialize social categories, one possible reason that has been proposed is the use of generic language. Generic language, which can be descriptive or normative, has been described as a common vehicle for stereotypes and there is some evidence to suggest that it can promote and transmit essentialist beliefs. As generic statements lack explicit quantification they allow for multiple interpretations.

2.3 Building a Theoretical Framework

To reiterate, in this thesis I seek to answer the question in what respect relations between essentialism and stereotypes are present in knowledge about voice(s).

Knowledge about Voice(s)

Drawing from my characterizations of knowledge and voice(s) respectively, knowledge about voice(s) can be produced and transmitted at multiple possible sites: In academic contexts, to which an authoritative role can be assigned, the production and reception of voice(s) can be the subject of research and publications in a variety of disciplines such as musicology, linguistics, phonetics, biology, neuroscience, evolutionary studies, acoustics and engineering[111]. Knowledge about voice(s) is also produced and transmitted in formal and informal lessons, such as vocal-pedagogical processes or everyday socialization, through which people learn about voice(s) and form vocal

[111] Eidsheim (2019): The Race of Sound. p. 15; Frühholz; Belin (2018): The Science of Voice Perception. p. 10.

habits in relation to their sociocultural contexts[112]. As I conceptualized knowledge as actively constructed and not necessarily true, I include the possibility that knowledge about voice(s) can be biased. Given that multimodal or context information can lead to biases in auditory perceptions, the process of listening to others or oneself depends on sensual experience and is therefore not objective, so the situation in which listeners perceive voice(s) matters. I will differentiate between three possible scenarios:

1. *The presence of voice(s) without knowing about or seeing the body of the vocalizer*
   (e.g. voice chats or phone calls, listening to auditory media, digitally synthesized or designed voice(s))
2. *The presence of voice(s) with the body of the vocalizer also present or perceivable*
   (e.g. face-to-face communication, live performances)
3. *The presence of voice(s) that are matched to another perceivable body than the one of the vocalizer*
   (e.g. voice-overs in audiovisual media)

By conceptualizing knowledge about voice(s) as constructed as well as actively produced and transmitted, I imply that discourses can (re)present the topic in different ways through the use of language and can take different forms, as exemplified by determinist and constructivist approaches in academic publications about voice(s): In this context, I argued that determinist perspectives imply a certain necessity about the relationship between category membership, vocal output and its perception in terms of a direct determination of voice(s) due to biological factors. While I will not fully discuss how determinism and essentialism relate to each other due to the scope of this thesis, I will entertain the possibility that there is an overlap: In the context of philosophy, I characterized essentialism as the assumption of identity-determining essences, which are the totality of all essential or necessary properties of an object or a general category, such as a natural kind. The concept of a natural kind implies that it reflects the structure of the natural world and is therefore objective and mind-independent. Essences in this sense are inherent or intrinsic and have been assigned an explanatory function concerning other properties of category members. As members of a natural kind category are necessarily homogenous in regard to their essence, though not necessarily in regard to further accidental properties, knowledge about the category membership allows for inferences concerning any individual member in regard to essential properties. Determinism and essentialism share the element of direct determination or necessity, which

[112] Eidsheim (2019): The Race of Sound. p. 41.

fundamentally contradicts constructivist perspectives that imply that voice(s) could be produced and perceived in other ways. I am taking a constructivist perspective in relation to voice(s) as I conceptualized vocalizations as specific instances of the voice(s) a person is capable of and argued that people have agency in this regard, as exemplified by the fact that people can learn to alter their perceived vocal gender.

Social Categories and Natural Kind Essentialism

I conceptualized social categories as classifications of people based on shared characteristics. These classifications are not necessarily essentialist, as according to Hallett, one can communicate about common characteristics without assuming essences and their implications. However, in the context of psychological essentialism it has been found that there can be essentialist beliefs in regard to social categories, of which some are essentialized more than others, with the dimensions that were essentialized most rooted in assumed biological theories of category origin. As I conceptualized voice(s) as bodily phenomena, this renders it likely that if there are essentialist perspectives concerning voice(s) in relation to social categories, they would be based on biological theories of category origin, similar to determinist perspectives. I will therefore focus on natural kind essentialism and in turn on social categories that have been discussed as natural kinds: Among these are the dimensions of gender and „race“, which have also been found to be essentialized most. According to Mallon, racial essentialism is common in everyday thought and can be characterized as the view that „races“ possess or are characterized by an essence that is conceived as natural or nonsocially contingent. These essences are thought to play explanatory roles in the sense that they both explain what it is to be a member of a „race“ and other properties that are considered „race-typical“ as caused by essential properties[113]. There is considerable interdisciplinary agreement that racial essentialism is a false view, as no viable biological candidates for such an essence have been found and essentialism about biological kinds like species, and „racial kinds“ in turn, is nearly universally rejected in the discipline of biology[114]. Following Roberts, I view „race“ as an invented political categorization system. Despite there being no scientific evidence for categorizing people according to biological „races“, the prevalence of the concept points to its relevance as a social category that shapes people's perceptions and experiences[115]. There is also ample evidence that

[113] Mallon, Ron (2024): Race. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. p. 361f.

[114] Ibid., p. 362ff.

[115] Roberts, Dorothy quoted in Bibbins-Domingo (2021): Racism and Race. p. 5ff.
I use the term „race“ in quotation marks to point to the fact that I do not share the assumption that there are human „races“.

gender categories are commonly essentialized[116]. In this context, the idea of binary developmental trajectories in humans is commonly made salient, with the categories „female“ and „male“ conceptualized as natural kinds in that there are intrinsic and necessary biological features that determine kind membership prior to socialization. However, universal claims in the essentialist fashion do not account for people who do not have typical „female“ or „male“ attributes or whose self-identification does not match such attributes[117]. Concerning self-identification, there is a common distinction between sex, referring to biological or physical characteristics, and gender, referring to social and cultural aspects. However, the concept of sex can be viewed as constructed as well: As the binary categorization precedes modern biological theories of category origin[118] and sex can be described as constructed in the sense that it encompasses various aspects, such as anatomy, chromosomes, or hormone levels, which are normalized in relation to the binary categories of „man“ and „woman“ despite different possible manifestations or combinations[119], I will treat references to both sex and gender as constructed in relation to social categories. While I consider it important to note that individuals identify with and are assigned to multiple categories and that stereotypes as well as discriminatory processes work in intersectional ways depending on the interplay of multiple social categories, I am interested in stereotypes about groups of people in relation to essentialism, wherefore I will focus on the way categories are presented in isolation in knowledge about voice(s).

Stereotypes and Natural Kind Essentialism

I conceptualized stereotypes in a formal way, namely as the cognitive representation of a group in respect to the attributes, features and behaviours members of that group are thought to share. Prejudice depends on stereotypes in that it adds the aspect of evaluation to this representation. I will stick to this formal characterization in order to not exclude possible contents in relation to voice(s) a priori. As seen in 2.2, several studies in the field of psychological essentialism have proclaimed a link between essentialist beliefs and the stereotyping of social categories but definitive claims as to their relationship are complicated by possible dependencies on the categories under scrutiny as well as other factors such as social status and varying methods concerning the measuring of essentialism as well as stereotypes or prejudice. However, there is relative consensus that there are certain social

[116] Haslam; Whelan (2008): Human Natures. p. 1302.
[117] Rosario, Esther (2024): Sex and Gender. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. p. 377.
[118] See Voß (2010): Making Sex Revisited. Dekonstruktion des Geschlechts aus biologisch-medizinischer Perspektive. p. 15f.
[119] Butler, Judith (2014): Gender Trouble. Feminism and the Subversion of Identity. London / New York: Routledge. p. viiff.

categories that are commonly subjected to essentialist beliefs. In this sense, I will include the possibility of implicit essentialism, as beliefs of this sort can be characterized in relation to the philosophical concept of essentialism, but it would be a stretch of the imagination to assume that it is a requirement for people to be familiar with essentialism as a philosophical concept and be aware of its implications to hold essentialist beliefs. In relation to the ways I characterized stereotypes and essentialism, they share a structural similarity: Stereotypes as the representation of a group or social category in respect to attributes its members are thought to share imply that it is possible to infer said attributes from group membership. Additionally, groups are rendered homogenous in relation to these attributes. While this characterization does not necessarily imply essences, it is entirely compatible with natural kind essentialism, which adds the dimension of necessity due to natural or inherent essences that exist objectively and mind-independently. It may not be possible to infer essentialism from stereotypes but it would appear that if one were to take essentialism about social categories as natural kinds seriously, stereotypes that generalize about social categories and allow inferences about individuals due to their group membership would be a logical consequence of that conception. However, this formal approach does not necessitate specific stereotypes, as the content of stereotypes would have to depend on the content of the essence in question.

Relations between Social Categories and Voice(s) as Trait(s)

Discourses around voice(s) in relation to social categories center around the idea that certain attributes of a person become audible in their vocalizations. While this idea can be explained both deterministically and constructivistically, it can be conceptualized as a category-trait relation in both cases. In relation to natural kind essentialism, this relation can take different forms:

1. *Voice(s) as accidental properties of social categories*

   If an essence is assigned to a social category, voice(s) can either be an essential or accidental property. As nonspeaking individuals could still be classified in relation to social categories, I would argue that voice(s) as perceivable phenomena would not be an essential property of social categories. In this sense, voice(s) would take the role of an accidental property. As voice(s) are bodily phenomena in that they result from the vocal tract, biologically conceptualized essences of a natural kind would take an explanatory function: Both the biology of the vocal tract and specific aspects of perceivable vocalizations that result from it can be explained in relation to the essence of a supposed general or natural kind.
2. *Voice(s) as essential traits of social categories*

   Were one to ignore the fact that nonspeaking individuals can also be classified in relation to

social categories or focus on features of the vocal tract without the need to include vocalizations, voice(s) can be considered essential traits of social categories: In this sense, aspects of the vocal tract or aspects of the vocalizations resulting from it would take the role of properties or features that are necessarily present in each member of the social category in question. However, it is also possible that essential features of voice(s) are conceptualized in regard to non-biological essences.

3. *Voice(s) as essences*

   Voice(s) themselves can be the subject of essentialist beliefs. I consider this option because of the line of research that considers traits as thought of as an essence itself. However, the conceptualization of voice(s) as possessing essences does not necessitate a relationship to social categories, which could therefore take a multitude of forms.

Essentialism allows for logical inferences: In the case of voice(s) as *accidental properties* of social categories, essences can take an explanatory function, which will explain voice(s) as caused by the essence or essential properties of a social category. In the case of voice(s) as *essential properties*, it is necessary that all members of group G share in the essential aspects A of voice(s), which also allows for the inference that if individual X is a member of group G, their voice(s) will necessarily feature certain aspects A. Following Johnson et al., I will call this a predictive inference[120]. In the case of voice(s) as *accidental properties*, it is not logically valid to infer group membership from listening to vocalizations because accidental properties are not identity-determining and can be shared between groups. In this sense, it would be logically invalid to infer membership in group G from the observation of aspects A of voice(s) in individual X without the additional premise that certain aspects A of voice(s) are *essential properties* or genuinely distinct between different groups and therefore signal group membership. In this case, it would be impossible that an individual X shares in certain aspects A regarding their voice(s) without also being a member of G unless A are viewed as accidental properties, which would render them non-informative regarding group membership. In this sense, if different categories are taken to be fundamentally different and certain aspects of voice(s) mutually exclusive in regard to social categories, inferring group membership from voice(s) necessitates the conceptualization of (aspects of) voice(s) as essential traits. Following Johnson et al., I will call the act of inferring group membership from certain aspects of voice(s) a diagnostic inference[121]. In taking this formal approach, I intend to not exclude aspects of voice(s) that can lead to predictive or diagnostic inferences.

[120] Johnson, Samuel G. B. et al. (2016): Explanatory Biases in Social Categorization. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. 38/0. p. 776.

[121] Ibid.

Approach

Moving forward, I will analyze case studies concerning knowledge about voice(s) in regard to the question in which respect there are relations between essentialism and stereotypes. I will conduct these analyses according to the following pattern: First, I will provide a summary of the case in question, especially focusing on the presentation of relationships between voice(s) and social categories. I will then point out the specific site of knowledge production or transmission to situate the case and name the social categories in question as well as the listening situation in relation to the three scenarios I proposed above. Next, I will discuss the relationship between essentialism and the way knowledge about voice(s) is presented: As I do not expect to find (m)any explicit statements of essences, I need criteria that allow me to identify implicit natural kind essentialism. According to Haslam et al., four factors concerning the structure of essentialist beliefs that correspond to the philosophical concept of a natural kind are that (a) membership in a category is immutable, (b) a category is discrete and has defining features, (c) that category is natural and (d) historically invariant, therefore existing independently of human language or societies. While these factors correspond to aspects of my characterization of natural kind essentialism, I would argue that the further factors of essentialist beliefs Haslam et al. assigned to the concept of entitativity follow from ideas (a) to (d): These state that (e) the category has an inherent nature underlying surface characteristics, which describes the idea that a category is natural and its essence can explain other characteristics, (f) category members are uniform, which is a consequence of defining features that must be necessarily present in each member of the category, so in relation to this essence, the members of a category must be uniform, (g) the category is informative and allows for inductive potential, which is due to the necessity of defining features and (h) the category is identity-determining and thus excludes members from other categories, which follows from the conceptualization of category membership as immutable and discrete. So while it is possible that essentialist beliefs correspond to the concept of entitativity without also corresponding to the concept of a natural kind, I would argue that the factors concerning entitativity are entailed by the factors concerning natural kinds. In this sense, I will use the following criteria for (implicit) natural kind essentialism: To view a social category as a natural kind implies that the category is *natural* in that it reflects the structure of the natural world and is therefore objective and mind-independent, rendering membership in the category immutable. The category is furthermore *discrete*, implying defining features and excluding members from certain other social categories. Members of the social category in question necessarily possess the same essence or essential properties in regard to which they are presented as *uniform*. This necessary uniformity allows for *inferences* regarding category-trait relations. Based on this characterization, I expect that (implicitly) essentialist

knowledge would be presented by using generic language. However, it is possible that I will encounter generic language without (implicitly) essentialist intentions. Given that generic language leaves room for the interpretation of readers or listeners and essentialism can be implicit, I will discuss whether the case in question allows for an essentialist reading or interpretation as proposed in the ways relationships between voice(s) and social categories can be essentialized instead of postulating that these four factors need to be explicitly present to label cases (implicitly) essentialist. In this sense, I will focus on the question of whether knowledge about voice(s) is presented in a way that contradicts my criteria of (implicit) natural kind essentialism, for example by mentioning counterexamples (contradicting the aspect of uniformity) or presenting social categories as mutable or human constructions (contradicting the aspect of naturalness). This procedure will not allow me to gain insights into what the people who produce and communicate knowledge about voice(s) intended. Instead, the possibility of an essentialist reading or interpretation points to a possible transmission of essentialist beliefs. As essentialism and stereotypes share a structural similarity, but stereotypes are not neccesarily based on natural kind essentialism, I will furthermore discuss inhowfar specific knowledge about voice(s) meets the formal characterization of stereotypes and which contents are present in this form in addition to the question whether the specific case allows for an essentialist reading before discussing their relationship. In this context, I will discuss possible inferences.

3 Case Studies

There are multiple sites of knowledge production and transmission in relation to voice(s) that include academic publications and research, formal lessons and informal lessons. I will include and explore examples from each of these sites to work towards a broader perspective concerning the construction of knowledge about voice(s), considering (implicit) natural kind essentialism possible at all of these sites. Nevertheless, my distinction between academic and other, therefore non-academic sites of knowledge production and transmission by implication, is not clear-cut: Academic publications and research can be concerned with non-academic contexts, while non-academic sites can be influenced by knowledge produced and transmitted in academic contexts. However, given that a certain status or potentially special authority can be assigned to different sites of knowledge production, I will hold on to this somewhat arbitrary distinction. I will use singular quotation marks while paraphrasing sources to signal that I present a formulation instead of my own position.

My choice of case studies is a result of my exposure to some of the material during studying voice(s) before starting to work on this thesis, as well as researching relationships between voice(s) and social categories in regard to my research question. It is a limitation of my approach that I can make no claim as to whether or inhowfar this choice is representative of knowledge about voice(s) in general. As my analysis is based on how instances of knowledge can be interpreted, it would be fair to assume that I underly a confirmation bias, or am biased towards interpreting essentialism into my case studies, resembling the third potential issue Hallett describes in relation to essentialist thinking. However, as I am interested in whether and / or in what respects there are relations between essentialism and stereotypes in knowledge about voice(s) in terms of the possibility of an essentialist reading and I will actively look for information that contradict this interpretation, stating that instances of knowledge about voice(s) allow for an essentialist reading does not exclude other interpretations.

3.1 Academic Knowledge about Voice(s)

I will first consider academic knowledge about voice(s) and start by reviewing the two examples of publications specifically concerned with essentialism and voice(s) that I found in the context of my research for this thesis (3.1.1). While these cases do not fit my theoretical framework in regard to implicit essentialism as relationships between essentialism and attitudes concerning voice(s) in relation to social categories are made explicit as the subject of research, I nevertheless consider them relevant in regard to the question in what respects relations between essentialism and stereotypes are present in knowledge about voice(s). I will compare the presented approaches to my theoretical framework concerning how essentialism, social categories, voice(s), stereotypes and their interdependencies are conceptualized. As a next step, I will use my theoretical framework to analyze an academic publication (3.1.2) in regard to how the knowledge it presents allows for an implicitly essentialist reading and relates to stereotypes.

3.1.1 Academic Knowledge about Essentialism and Voice(s)

3.1.1.1 Stigmatization of 'gay-sounding' voices

In their article *Stigmatization of 'gay-sounding' voices: The role of heterosexual, lesbian, and gay individuals' essentialist beliefs* from 2021, Fasoli et al. point out that while voice-based judgements of sexual orientation can prompt group-based discrimination, there has not been any research on the

relationships between stigmatization and essentialist beliefs about vocal cues to sexual orientation up to the point of publication. Their overarching research question is whether essentialist beliefs about 'the voice' are relevant to how stigma is enacted, anticipated and experienced[122]. The authors position their research within the context of auditory gaydar: The idea of an auditory gaydar describes the use of vocal cues to infer other people's sexual orientation. The authors argue that belief in an auditory gaydar presumes the existence of real differences between gay and straight people in regard to discernable vocal traits (voice discreteness). Additionally, auditory gaydar implies that differences in voice(s) are deep-rooted and fixed (voice immutability) and raises the question inhowfar people can intentionally change their voice(s) to emphasize or conceal their sexual orientation (voice controllability). While immutablity and controllability beliefs may seem logically incompatible, the authors point out that previous research has demonstrated that people can be aware of voice controllability as a means to communicate identity while also believing in discrete vocal differences between groups[123]. The authors conceptualize voice(s) in this context by claiming that „[s]ome acoustic cues are defined by physical factors cues (e.g., frequency parameters) whereas others (e.g., pitch, duration) are somehow controllable and influenced by context“[124].

The article is comprised of two studies in order to investigate the process of stigmatizing as well as the experience of stigmatized people: In study 1, Fasoli et al. ask how essentialist beliefs concerning voice(s) amongst the heterosexual majority give rise to stigmatizing processes in the form of prejudice and avoidant discrimination against lesbian and gay people[125]. Participants who did not identify as heterosexual were excluded from this study; the remaining participants completed measures of voice essentialist beliefs in regard to discreteness, immutability and controllability as well as prejudice and avoidant discrimination. The results of study 1 include the finding that essentialist beliefs, especially discreteness, concerning voice(s) as a trait of sexual orientation were endorsed by heterosexual participants and strongly related with measures of avoidant discrimination[126]. Both measures of voice essentialist beliefs and avoidant discrimination were stronger when applied to gay men than to lesbian women. Based on previous research concerning essentialist beliefs and categories of sexual orientation, the authors had hypothesized that immutability beliefs would be associated with lower prejudice. While immutability beliefs did not

[122] Fasoli et al. (2021): Stigmatization of 'gay-sounding' voices. p. 826f.
[123] Ibid., p. 827f.
[124] Ibid., p. 828.
[125] Ibid., p. 829.
[126] Ibid., p. 841.

predict measures of anti-gay prejudice, there was a positive relationship to avoidant discrimination[127]. Fasoli et al. conclude that „[...] this result underscores the point that stigmatization has a different relationship to essentialist thinking when category-trait associations are considered.“[128] In study 2, the authors focused on the question of how lesbian and gay individuals' essentialist beliefs about their own voice(s) shape their expectations of rejection and vigilance as a form of stigmatization[129]. Participants who did not identify as gay or lesbian were exluded from this study; the remaining participants completed measures of essentialist beliefs (the authors used the same scale as in study 1), self as well as external perception of their voice(s) in regard to whether they sound gay or lesbian and gender-typical, stigma expectancy and vigilance[130]. The results of study 2 include the finding that „believing in discrete immutable differences between LG [lesbian and gay] and heterosexual voices was related to LG individuals' beliefs that they sounded LG themselves and their expectations of rejection, particularly amongst gay men“[131]. While beliefs in immutability and discreteness were positively associated with self-perceptions of sounding gay or lesbian and negatively associated with sounding gender-typical, controllability beliefs were positively associated with perceptions of one's voice(s) as gender-typical and did not predict perceiving one's own voice(s) as sounding gay or lesbian. Self-perceptions of one's own voice(s) as sounding gay or lesbian positively predicted expectations of rejection, with a significant link between expectations of rejection and vigilance[132]. In the final discussion, Fasoli et al. point out that „stereotypes communicated through mass media particularly describe voice as a [sexual orientation] cue for men [...]. Some gay men associate a negative connotation with sounding gay and try to avoid such stereotypical speech […] Gay men tend to avoid ingroup members portrayed in a stereotypical way when they feel pressure to conform to norms [...]. If sounding gay is seen as a 'stereotypical speech' [...] that represents a deviation from norms and elicits negative attitudes [...], then controllability beliefs may be associated with stigmatization of gay-sounding ingroup members.“[133]

The scales and items regarding essentialist beliefs can be found in the supplementary materials and were adapted from previously published essentialist belief scales[134]. For exmple, the items include „Gay/lesbian people sound gay/lesbian, and there is not much they can do to really change that.”

[127] Ibid., p. 833f.
[128] Ibid., p. 834.
[129] Ibid., p. 829.
[130] Ibid., p. 836.
[131] Ibid., p. 841.
[132] Ibid., p. 839.
[133] Ibid., p. 842f.
[134] Ibid., p. 831.

(immutability), „Gay/lesbian people can choose to sound gay or straight depending on the situation." (controllability) and „When listening to a person it is possible to detect his/her sexual orientation from his/her voice very quickly." (discreteness)[135]. Given the context of my research for this thesis, I had expected that the items measuring controllability would focus on the idea of agency in relation to a person's vocalizations. While this aspect is not exactly absent, it nevertheless surprised me that several of the controllability items frame this aspect in a way that seems to imply being perceived as gay or lesbian on the basis of vocalizations is due to gay or lesbian speakers making the conscious decision to signal their sexual orientation and not adhere to heterosexual norms, as in „People try to sound gay/lesbian because they want everyone to notice immediately that they are gay/lesbian.", „People usually do not sound gay/lesbian unless they want to sound that way to others." or „A voice coach or occupational therapist can help most gay/lesbian people to change their voices to sound less gay/lesbian." These points could be interpreted as prejudicial in that they could be taken to justify treating speakers who are perceived as gay or lesbian differently because of the implication that they have access to other vocalizations and sounding gay or lesbian, which is equated with being perceived as gay or lesbian, is not only viewed as a choice but potentially undesirable, as in the context of vocal therapy helping people to sound less gay or lesbian, reminiscent of conversion programs for homosexual people. By presenting controllability items in a way that can be said to imply prejudice, I would argue that it is not surprising that agreement with these items is linked to prejudice or avoidant discrimination. Furthermore, I was surprised to find qualifiers throughout the scale that I had not expected in the context of essentialist beliefs, such as "there is *not much* they can *really* do to change that" (my emphasis) or "Gay/lesbian voices *probably* only exist in certain cultures." (my emphasis) concerning controllability, especially because the scale also contains items without these qualifiers, such as "People can talk in ways that could sound either gay/lesbian or straight, but the essence of their voice as either gay/lesbian-sounding or straight-sounding is clear-cut." (immutability). In regard to the relation between sexual orientation categories and voice(s) as trait(s), some items present the perception of voice(s) as gay, lesbian or straight as determined by the sexual orientation of speakers (e.g. „Gay/lesbian people sound gay/lesbian, and there is not much they can do to really change that.", „There exist gay/lesbian and straight voices and, thus, it is possible to figure out a speaker's sexual orientation relatively quickly."), while others present a non-necessary relationship between a speaker's sexual orientation and their vocalizations (e.g. the reversed items „Judgements on people's sexual orientation from voice are not accurate." or „A person's voice can never tell you a lot about his/her

[135] Fasoli, Fabio et al. (2021): Gaydar Beliefs and Stigma – Supplementary Materials. URL: https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjso.12442 [accessed May 5th, 2025]. p. 2f.

sexual orientation.")[136]. While an essentialist argument one could make on the basis of these beliefs fits the context of an auditory gaydar in that direct or essential relationships between a person's category membership, their vocalizations and their perception are presented, the question of why voice(s) are considered traits that signal category membership is not addressed in this study.

Concerning knowledge production and transmission, Fasoli et al.'s article can be situated in the context of academic research and publications. The social categories in this case study are „gay", „lesbian" and „straight / heterosexual" in relation to sexual orientation, with an additional focus on the intersection with the categories „men" and „women". As no actual voice(s) were featured, the listening situations I proposed in my theoretical framework are not applicable. Vocal production and vocal perception are not clearly differentiated in this article, as the items measured people's general beliefs about voice(s) in relation to sexual orientation. Nevertheless, the article seems to put a focus on how voice(s) are perceived in the framework of auditory gaydar research, both in terms of external judgements and such relating to the participants' own voice(s). In the final discussion, the authors also included potential effects of stereotypes concerning 'gay-sounding' voice(s) in terms of gay speakers avoiding stereotypical speech patterns, therefore also discussing potential influences of vocal perceptions on vocal production.

Comparison of Conceptualizations of Voice(s)

The authors' approach to voice(s) differs from my conceptualization of the term in that they claim some acoustic cues are defined by physical factors cues, such as frequency parameters, while other parameters are controllable and influenced by context. In contrast, I argued that the possible range of vocalizations is defined by an individual's anatomy and that every specific vocalization only represents an instance of the possible vocalizations a person is capable of, therefore not implying that any acoustic cues in specific vocalizations are necessary or could not be different. While the authors are rather vague about the physical factors cues that define acoustic cues, frequency parameters such as the fundamental frequency and formants have in fact been demonstrated to be parameters of vocal production that people can modify[137]. As the authors bring up their conceptualization of voice(s) in the context of controllability and people believing in discrete vocal differences while also being aware of voice controllability[138], this implies the possibility of specific defined vocal markers that could potentially signal group membership. While the authors'

[136] Ibid.
[137] Xu, Mingdi et al. (2020): Unconscious and Distinctive Control of Vocal Pitch and Timbre During Altered Auditory Feedback. In: *Frontiers in Psychology.* 11/1224. p.
[138] Fasoli et al. (2021): Stigmatization of 'gay-sounding' voices. p. 828.

conceptualization of voice(s) could be discussed in the context of implicit essentialism, they do not make reference to specific social categories in this context. This discussion is therefore beyond the scope of this thesis.

Comparison of Conceptualizations of Essentialism in Category-Trait Relations

According to my framework, essentialism can play different roles in category-trait relations. I entertained the possibility that voice(s) can be subject to essentialist beliefs but pointed out that in this case, this approach does not necessitate any relation to social categories. Fasoli et al. situate their research in the context of psychological essentialism and category-trait relations. They research essentialist beliefs about voice(s) that are furthermore conceptualized as a cue to speakers' sexual orientation (SO). The authors use formulations like „essentialist beliefs referring to a SO-trait that triggers SO categorization"[139] and „essentialist beliefs about voice as a cue of SO"[140] and therefore treat voice(s) as essentialized while adding the aspect that people infer sexual orientation categories from this trait[141]. I would argue that in the context of category-trait relations, the possibility to infer categories from essential traits (diagnostic inference), as implied by the concept of an auditory gaydar, points to an essence being ascribed to the category, not the trait. The trait could be either essential or accidental, with the further possibility that only specific markers or aspects of voice(s) are viewed as essential. Additionally, this conceptualization would eliminate the mentioned contradiction between immutability and controllability beliefs, since the category could be conceptualized as having an immutable essence that gives rise to essential or accidental properties of voice(s), therefore including the possibility that while people can control certain aspects of their voice(s), some immutable essential markers might signal group membership. In this conceptualization, immutability and controllability beliefs in regard to voice(s) could be held simultaneously. Fasoli et al. identified discreteness, immutability and controllability as factors of interest concerning essentialist beliefs about voice(s). In my framework, I focused on social categories conceptualized as natural (and therefore immutable), discrete and uniform, including the potential to draw inferences. While not explicated, I would argue that these aspects (except for the aspect of naturalness that could be replaced by inherency in non-natural general kinds) are present in the items of the essentialist beliefs scales, such as uniformity being implied in „People can talk in ways that could sound either gay/lesbian or straight, but the essence of their voice as either gay/lesbian-sounding or straight-sounding is clear-cut." Furthermore, I would argue that the items

[139] Ibid., p. 841.
[140] Ibid., p. 826.
[141] While it is theoretically possible that Fasoli et al. thought of voice(s) as possibly being considered to be essential traits of sexual orientation categories, I consider it more likely that they conceptualized voice(s) as possibly being considered as possessing essences due to the formulations used in this article.

used to measure essentialist beliefs about voice(s) refer to slightly different conceptualizations of category-trait relations: The reversed items, such as „Judgements on people's sexual orientation from voice are not accurate.” can be read as decidedly non-essentialist by displaying a non-necessary and non-informative relationship between categories of sexual orientation and 'voice' as a trait. If people disagree with the item, this could be because of several different reasons and therefore different (non-)essentialist conceptualizations, such as an essentialist view of sexual orientation categories that determine vocal output or the assumption that people will talk in a way to signal category membership even though they could choose not to. Other items, such as „When listening to a person it is possible to detect his/her sexual orientation from his/her voice very quickly.” imply the possibility of a diagnostic inference which would only be logically valid if the trait is essential to the group, therefore fitting the framework of conceptualizing the category as essentialized instead of the trait. However, the item „Gay/lesbian people can talk differently, but their voices can't really change.” does not explicate inhowfar the possibility to talk differently without your voice(s) really changing is thought to be characteristic of gay or lesbian people instead of a general belief about voice(s) of people and therefore fits the conceptualization of voice(s) as essentialized. So in regard to essentialist category-trait relations, these items can be said to measure different conceptualizations. These can of course co-exist in principle, however, I would argue that in order to adress the connection between essentialist beliefs and stereotypes, it makes more sense to conceptualize categories as essentialized and voice(s) as essential or accidental traits, as this approach allows for investigating the relationship thought to exist between categories and traits, making use of the explanatory function assigned to essentialism. The question of why voice(s) are thought to signal category membership remains unexamined in this article.

Comparison of Conceptualizations of Social Categories

The authors do not make explicit how they conceptualize the categories they employ in their article. In terms of gender categories, they employ the binary conception of „men” and „women” while presenting sexual orientation categories as „gay”, „lesbian” and „straight” and raising data according to the participants' self-identification, therefore excluding other sexual orientations that do not fit these clear-cut labels. While this can be interpreted as a specific interest in only these categories, the reason for this choice is not made explicit. Additionally, they equate the process of stigmatizing with heterosexual people and the experience of stigma with gay and lesbian people while also pointing out that stereotypes can exist towards one's ingroup as well in the context of gay men potentially avoiding stereotypically gay speech patterns. However, by pointing out that there are differences surrounding processes of stigmatization in regard to gender categories, the authors

add an intersectional layer to their research by choosing to not group gay and lesbian people into the category „homosexual". As I focused on social categories that have historically been viewed as natural kinds, I did not explicitly consider sexual orientation categories. However, this study demonstrates that it is not necessary for social categories to be viewed as natural kinds for people to hold essentialist beliefs in the context of category-trait relations.

Comparisons of Conceptualizations of Stereotypes

While the authors explicitly mention stereotypes, for example the stereotype that gay or lesbian people are not gender-typical or the in context of gay men associating negative connotations with sounding gay and therefore avoiding stereotypical speech patterns[142], they focus on the concept of stigmatization and employ the concepts of prejudice and avoidant discrimination. While they differentiate these concepts in terms of the scales they use to measure them, it is not made explicit inhowfar avoidant discrimination is thought to differ from prejudice. In relation to my framework, I would argue that both avoidant discrimination and prejudice depend on the idea that you can draw inferences concerning people based on their group membership. Whether they elicit negative evaluations or actions such as avoiding, a crucial aspect of this related to essentialism and not explicitly researched in the article is the aspect of uniformity that renders a category informative in certain regards.

Relations Between Essentialism and Stereotypes

The authors demonstrate that their framework and method led to significant results that link essentialist beliefs about voice(s) as a cue to sexual orientation to avoidant discrimination and experiences of being stigmatized in terms of expectations of rejection and vigilance. These effects differ depending on whether gay men or lesbian women are considered. While there are differences to the framework I developed to analyse case studies in terms of conceptualizations of essentialism in regard to category-trait relations and I am interested in stereotypes instead of stigmatization, I have argued that essentialist beliefs about voice(s) do not necessitate references to social categories. Therefore, in treating voice(s) as essentialized instead of the category in the context of auditory gaydar, it is not possible to differentiate whether voice(s) or aspects thereof are viewed as essential traits or accidental ones that could nevertheless be explained in relation to a supposed essence.

[142] Fasoli et al. (2021): Stigmatization of 'gay-sounding' voices. p. 835, 843.

3.1.1.2 The Race of Sound

In *The Race of Sound*, Eidsheim discusses essentialism in relation to voice(s). I will focus on the theoretical groundwork to summarize the author's approach to the relationship and discuss it in relation to my framework. Eidsheim writes that in listening to human voice(s), the fundamental question asked in this act is about the identity of the vocalizer, independently of whether the vocalizer is also visible or not. In this sense, the author brings up the concept of acousmatic sounds, described as sounds one hears without seeing the causes behind it[143]: „[T]he acousmatic question arises from the assumption that, in asking, it is possible to elicit an answer. It is assumed that if I listen carefully to a sound — in the absence of a visually presented or otherwise known source — I should be able to identify a source, and that any limitations are due to inexperience or ignorance. For instance, through attentive and informed listening, I should be able to know a lot about the vocalizer, and possibly about his or her identity. […] In the context of the human voice, this assumption about the possibility of knowing sound in the first place extends to a second assumption: that it is possible to know a person. The acousmatic situation arises from the assumption that voice and sound are of an a priori stable nature and that we can identify degrees of fidelity to and divergence from this state. This position is grounded in a belief — and truth claims — about the voice as a cue to interiority, essence, and unmediated identity. We assume that when we ask the acousmatic question we will learn something about an individual. We assume that when we ask the acousmatic question we inquire about the essential nature of a person.“[144] The first premise Eidsheim identifies is about the possibility of gaining knowledge or information from sound, including human voice(s). The second premise is described as an extension of the first in that vocalizations arise from persons and so it is possible to gain knowledge or information about the source of sound, a person, from human voice(s). The third premise is not named as such but as an assumption concerning an a priori stable nature of 'voice'. The fourth premise, called a belief and truth claim, entails the idea of voice(s) being cues to interiority, essence, and unmediated identity. In this context, Eidsheim does not reference social categories. Instead, 'voice' is conceptualized as a cue to a person's essence. Concerning social categories, Eidsheim writes that „the thick event [of vocalizations...] is reduced to socially and culturally categorized and evaluated vocal sounds, such as pitch and voice, as essential markers.“[145] Here, it is clarified that vocal sounds are conceptualized as essential markers or properties of sociocultural categories, with the implication that the essence in question is therefore the essence of the categories. This point is also taken up at a later point in

[143] Eidsheim (2019): The Race of Sound. p. 1f.
[144] Ibid., p. 2.
[145] Ibid., p. 8f.

the book: „The sound categories that can be further identified include distinct pitches, adult voices (versus, say, children's voices), male versus female, "ethnic" versus "nonethnic," "authentic" versus "inauthentic." However, because the premise of listening is identification, we do not question the likelihood of the a priori existence of the identified categories."[146] In this sense, Eidsheim stresses the similarity between essentialist processes of listening concerning individuals and categories. The author mainly focuses on the category „race", described as „essence", and timbre as an aspect of voice(s), described as a „trait[...] believed by a given society to be essential to people"[147]. In this context, the author points out the following inconsistency: „One of the many paradoxes related to timbre is that while vocal timbre is understood as essential, classical vocal pedagogy is built upon the very notion that it is possible to construct timbre."[148] Eidsheim goes on to explain that while some aspects of timbre are understood as essential, others aspects are understood as acquired and performed[149], opening up the possibility that not only are there aspects of voice(s) that are understood as more essential than others, but also aspects of these aspects can be thought of as essential while others are not. Throughout the book, Eidsheim argues that voice(s) are in fact not essential traits but can be listened to in terms of style and technique of the vocalizer[150], thus proposing to listen to voice(s) in terms of what people do instead of who they are. In regard to the passages cited above, it seems relatively clear that Eidsheim conceptualizes category-trait relations concerning voice(s) by presenting the view that an individual or a category is thought to have an essence, with voice(s) thought of as a cue to this essence in that aspects of voice(s), such as timbre, are thought to be essential to people. However, there are several formulations that imply a different conceptualization: For example, „the belief in race as an essential trait"[151] is not the same as describing „race" as an essence (see above) as it points to „race" as a trait of an individual, „as long as we understand voice as essence"[152] states 'voice' as essence as opposed to 'voice' as a cue or essential property and „vocal patterns may be seen as style and vocal technique rather than as an essential voice that points toward an essential self"[153] describes both 'voice' and the 'self' as essential on the same level.

Concerning knowledge production and transmission, *The Race of Sound* can be situated in the context of academic research and publications. As implied by the title, the main focus is on the

[146] Ibid., p. 56.
[147] Ibid., p. 25.
[148] Ibid., p. 52.
[149] Ibid., p. 55.
[150] Ibid., p. 23f., 34.
[151] Ibid., p. 51.
[152] Ibid., p. 154.
[153] Ibid., p. 186.

category „race“, while Eidsheim argues that the basic principle of asking the acousmatic question and expecting to find an answer can be applied to other categories as well. Concerning vocal production and perception, the author focuses on the process of listening in the context of diagnostic inferences. While the term „acousmatic question“ implies an acousmatic listening situation, there is no explicit listening situation or presence of voice(s) in the theoretical part that I examined.

Comparison of Conceptualizations of Voice(s)

Eidsheim conceptualizes voice(s) in terms of non-essential style and technique. This is compatible with my framework of vocalizations only representing an instance of the vocalizations a person is capable of as well as factoring in agency and vocal habits, which should not be surprising since I derived my conceptualization of voice(s) from this source.

Comparison of Conceptualizations of Essentialism in Category-Trait Relations

Eidsheim presents individual people, social categories, 'voice' as well as certain aspects of voice(s) as potentially viewed as essences, while also discussing voice(s) and social categories as essential traits. While it is possible that all of the above can be thought of as possessing essences, I would argue that this is not necessarily the case: For example, in the context of the 'paradox' concerning the premise of vocal lessons that voice(s) and timbre can be trained and the assumption of an essential timbre, the author points out that not all aspects of timbre must be understood as essential. While this matches the conceptualization of timbre as possessing an essence and therefore essential as well as accidental properties, as soon as social categories such as „race” are considered, there is no necessary relation between the category and timbral aspects without one being considered a trait of the other. In this sense, I would argue that the category would have to be thought of as possessing an essence for a trait or parts of a trait to be viewed as an essential or accidental property. The conceptualization of a social category thought of as possessing an essence that has an explanatory function concerning other accidental traits, such as timbre, would also lift the paradox. Therefore, essentialist category-trait relations can be said to require the category to be essentialized, while the belief that traits have an essence is optional. So in relation to the idea that attributes of vocalizers, including membership in social categories, become audible in their vocalizations, I would argue that while it is possible that people hold essentialist beliefs about voice(s) or certain aspects thereof, social categories and individual people simultaneously, an essentialist relationship between vocalizations and the group membership people infer from them requires the social category in question to be essentialized, which would give rise to an explanatory function concerning traits such as voice(s) and allow for the possibility that generalizations are made without viewing certain

aspects of voice(s) as necessary.

Comparison of Conceptualizations of Social Categories

Eidsheim conceptualizes social categories in relation to voice(s) as defined a priori and implies that the likelihood of their existence could be questioned, which is compatible with a reading of these categories as constructed. Furthermore, it is pointed out that vocal sounds are socially and culturally categorized and evaluated, with the categorization pointing to the basic premise of identification through listening and the evaluation to normative ideas in relation to these categories. While this is compatible with my conceptualization of social categories, it adds the dimension of individuals making sense of the world through listening and categorizing what they hear, which goes beyond the focus on social categories in relation to human voice(s).

Comparisons of Conceptualizations of Stereotypes

While stereotypes are not explicitly conceptualized in *The Race of Sound*, the premise of identification through listening to voice(s) can be read as being based on stereotypes as to identify which social categories an individual is a member of rests on an assumption concerning the attributes, features and behaviours members of a category are thought to share. However, this approach does not explicate on its own why there are stereotypes concerning voice(s) in relation to some social categories, such as along gender or „race“ dimensions, as compared to social categories in relation to dimensions such as occupation.

Relations Between Essentialism and Stereotypes

In Eidsheim's conceptualization, there is an overlap of essentialism and stereotypes in regard to the general idea that the act of listening is carried out in relation to the assumption that through listening, it is possible to identify. While this process does not necessitate stereotypes when applied to individuals or non-human sound sources, voice(s) or aspects of voice(s) are thought of and listened to as informative in regard to social categories. Eidsheim proposes that by listening to voice(s) in terms of style and technique instead of identification, associations between social categories and voice(s) could no longer be naturalized in an essentialistic manner[154].

In relation to my research question, I would argue that the respects in which stereotypes about social categories and essentialism are present in knowledge about voice(s) are not ideally addressed by viewing voice(s) as essences, as the actual relations in category-trait relations, such as between a

[154] See ibid., p. 24.

social category and voice(s), remain unexamined this way. Moving forward, I will therefore focus on the conceptualization of essentialism in terms of social categories potentially viewed as possessing essences that give rise to an explanatory function in regard to (certain aspects of) voice(s) that can be viewed as either essential or accidental properties in this context. This does not exclude the possibility that people hold essentialist beliefs about voice(s) simultaneously, but given that this approach does not necessitate a relation to social categories, I will not focus on this possibility any further. As I have pointed out above, this framework also eliminates contradictions or inconsistencies that arise from conceptualizing voice(s) as possessing essences, such as the facts that people can believe voice(s) to be controllable to some degree while also conveying category membership or that people can believe certain aspects of voice(s) to be essential while also being aware of the potential to train and change voice(s).

3.1.2 Implicit Essentialism and Academic Knowledge about Voice(s) - Handbuch der Tonstudiotechnik

In *Handbuch der Tonstudiotechnik*, which can be translated to „Handbook of Recording Studio Technology“, there is a chapter on sound sources with a subchapter titled „Menschliche Stimme“, or 'Human Voice'[155]. I will refer to the 9th edition of the book, published in 2023. In the preface, the volume is described as having become a classic handbook for professional audio technology and a standard reference for generations of students and people working in the field of audio technology since the first edition was published in 1976[156]. At the beginning of the chapter, 'the human voice' is described as a versatile sound generator that can produce sounds and noises of varying acoustic compositions. In this context, vocal production is described in terms of physiological components of the vocal tract and bodily features that modify resonances[157]. While these descriptions do not make reference to social categories, specific information concerning frequency and volume parameters is presented for men and women respectively throughout the chapter: Dickreiter describes the fundamental frequency of the vibration of the vocal folds as constantly changing during speech and defines its range as „about 120 to 160 Hz“ for men and „about 220 to 330 Hz“ for women and children[158]. These ranges are also visualized in a figure combining speaking and singing frequency levels, with singing frequency levels not made in reference to the categories men,

[155] Dickreiter, Michael (2023): Menschliche Stimme. In: *Handbuch der Tonstudiotechnik*. M. Dickreiter / V. Dittel / W. Hoeg / M. Wöhr (eds.). 9th edition. Berlin, Boston: Walter de Gruyter GmbH. p. 67.
[156] Ibid., p. VII.
[157] Ibid., p. 67.
[158] I chose „about“ as a translation for the German word „etwa“.
Ibid., p. 68.

women or children but in terms of the voice types of bass, baritone, tenor, alto, mezzo-soprano and soprano. While these categorizations of singing frequency levels is visualized with an overlap between the categories, there is no overlap between the categories „men“ and „women and children“ in terms of speaking frequency levels[159]. Dickreiter writes that given the large variety of speech sounds, it makes sense to refer to statistically derived average speech spectra instead of singular speech sounds and goes on to put forward that components below 80 Hz for 'male speech' and below 100 Hz for 'female speech' can be filtered out without sonic distortion because they are not audible in a regular distance between listener and speaker. As the low frequencies in amplified speech that was originally spoken quietly are relatively stronger compared to actual loud speaking, Dickreiter describes that normally inaudible components of 'male speech' below 80 Hz lead to a droning effect in the amplified condition and claims that 'male speech' is more sensitive in regard to this effect than 'female speech', wherefore 'female voices' are better suited for announcements in public areas or noisy environments, such as traffic announcements on the radio[160]. Furthermore, Dickreiter describes characteristic maxima in statistical speech spectra in relation to loudness and the maximum concerning speech pitch as gender specific[161], with the maximum for 'male speakers' around 150 Hz and around 250 Hz for 'female speakers'. Further characteristic maxima are described as around 500 Hz and 1.500 Hz for women and around 1.500 Hz for men speaking loudly as well, with key components of speech being limited to the frequency band between 100 Hz and 10 kHz for men and 200 Hz and 10 kHz for women. These statistical speech spectra are visualized in two graphs for German speech of male speakers and female speakers respectively[162]. In the context of lavalier microphones, Dickreiter writes that raising the volume level by about 6 dB results in a colouring of the sound in a small frequency band around 700 Hz for men and slightly above that for women[163]. Concerning volume, Dickreiter describes the sound level of speakers as 60 dB with a distance of 60 cm to the mouth of a speaker, claims that it rises by about 6 dB with loud speaking and that it reaches levels of 76 dBA with men and 68 dBA with women in the context of unnaturally loud speech[164]. In the context of singing voice(s), high 'female voices' (which are equated with sopranoes) and high 'male voices' (which are equated with tenors) are said to reach the highest volume levels, namely 50 to 60 dB. Concerning 'the female voice', formant frequencies are said to be adapted to respective pitches in singing, generally resulting in a darkening of vowel character. Dickreiter points out that the singing formant between 2,8 and 3 kHz is of major importance for the

[159] Ibid., p. 72.
[160] Ibid., p. 69.
[161] „geschlechtsspezifisch“, ibid., p. 70.
[162] Ibid.
[163] Ibid., p. 72f.
[164] Ibid., p. 71.

timbre of 'the male singing voice' and, connected to a general amplification of higher sound components, allows 'the voice' to stand out even against a loud orchestra[165].

The knowledge presented in „Handbuch der Tonstudiotechnik“ can be situated in the context of academic publications, with the additional feature that it is a handbook or textbook that has been described as a standard reference for students and people working with audio technology, which can therefore be a source people seek out to gain specific knowledge in order to apply it. The social categories mentioned in relation to voice(s) in the chapter summarized above are 'men', 'women' and 'children', while additional vocal 'types' in relation to singing voice(s) imply that singing is conceptualized according to Western classical music traditions as opposed to other singing styles. The relations between these categories and the knowledge about voice(s) presented here can be summarized as differences in terms of frequency parameters and volume levels, with voice(s) of men presented as lower in regard to pitch as well as characteristic maxima of average speech spectra and able to produce higher volume levels in 'unnaturally' loud speech compared to voice(s) of women. The category 'children' only appears in the context of fundamental frequency levels in speech and only in combination with the category 'women'. The question of the listening situation does not apply in this context as Dickreiter names and describes relations between voice(s) in terms of measurable vocalizations and social categories without referencing specific situations or vocalizations. There is also no distinction between vocal production and perception, which I assume is the case due to the overarching context in which voice(s) are viewed as sound sources and acoustic signals, implying a direct relationship. Interestingly, while there are references to literature and sources are made explicit at some points in the chapter[166], no sources are provided in the context of relations between vocal features and social categories. While Dickreiter makes clear that the knowledge presented refers to statistically derived average speech spectra as opposed to singular speech sounds, it is unclear whether the statistically derived average only refers to speech as opposed to specific vowels and consonants or general averages presented in relation to social categories as well. However, as Dickreiter neither provides sources for proclaimed differences between voice(s) in relation to the categories made salient nor makes explicit why there are differences, it is not possible to tell from reading the text how this knowledge was produced; it is therefore presented in an a priori manner.

I would argue that Dickreiter's „Menschliche Stimme“ allows for an essentialist reading: The

[165] Ibid.
[166] e.g. ibid., p. 68.

conceptualization of 'the human voice' as a versatile sound generator that can produce sounds and noises of varying acoustic compositions can be interpreted in terms of agency people have in relation to their voice(s) and various possible instances of vocalizations, which contradicts the aspect of uniformity and therefore renders it unlikely that 'the human voice' is thought of as possessing an essence itself. On the other hand, there are no contradictions to an interpretation of the categories of 'men' and 'women' as natural kinds with an inherent essence: The aspect of discreteness is amplified by the constant juxtaposition of different vocal features of these groups, as seen for example in the seperate graphs for male and female speakers or the figure concerning speaking and singing frequency levels that shows no overlap between the categories of 'men' and 'women and children'. While Dickreiter does not make explicit why women and children are grouped together in this figure, which I perceive as confusing as this is the only instance in the chapter that children are mentioned, I would argue that Dickreiter implicitly refers to testosterone levels in puberty that have been shown to influence the development of the vocal tract[167]. However, if testosterone was intended as the relevant aspect, this begs the question of how Dickreiter would conceptualize people who take exogenous testosterone at a later point in life, which also results in laryngeal growth[168]. In combination with the explicit references to physiological features at the start of the chapter and the context of the handbook being audio technologies which could be taken to imply objective measuring of natural phenomena such as voice(s), this lends credibility to the reading of men and women as natural kinds defined by 'objective', physiological features. While a belief in inherent differences between these categories is not made explicit in the text, I would argue that it is heavily implied. The aspect of uniformity is debatable: On the one hand, Dickreiter mostly refers to ranges and therefore includes the possibility that voice(s) can differ within their respective categories. On the other hand, the discreteness points to a uniformity in terms of vocal possibilities that are defined in relation to men and women. So while Dickreiter does not imply a uniformity in relation to every aspect of vocalizations, there is uniformity in regard to general possibilites of voice(s), which can be interpreted as ideas about men and women as natural kinds serving an explanatory function in regard to accidental properties, namely voice(s) or certain aspects thereof. While there are some qualifiers, such as „around“, in relation to the specific Hz and dB values Dickreiter puts forward, differences between social categories are presented in generic language without additional qualifiers, following the pattern of an aspect of voice(s) being discussed and presented in the context of different value ranges for men and women respectively. By equating social categories with certain vocal features, it is logically valid to infer that if an individual is a

[167] See Zimman (2018): Transgender Voices. p. 3.
[168] Ibid.

man or a woman, they will display the range of vocal features of their respective group. Additionally, for example by arguing that certain frequencies are inaudible with a 'normal' distance between speaker and listener, presenting these frequencies as different for men and women and concluding that due to this difference, female voice(s) are better suited for announcements in public spaces or noisy environments, Dickreiter also implies that in the context of audio technology, there are different features of men's and women's voice(s), that men's and women's voice(s) should be treated or processed differently and that there are contexts in which women's voice(s) should be preferred over men's voice(s). The potential to draw inferences is amplified by the fact that this chapter is featured in a textbook, so the implication is that people can apply practically what they learn theoretically. This further supports predictive inferences regarding how to work with voice(s) based on perceived group membership instead of actual vocalizations, therefore possibly serving as a reason for gender-specific actions that have the potential to reproduce normative ideas concerning the sound of voice(s).

As the differences between men and women are presented in generic language in regard to social categories and the possibility of exceptions or counterexamples is not made explicit, the knowledge presented in this chapter meets the formal requirement of stereotypes as it can be termed a representation of groups in regard to the attributes, features and behaviours members of that group are thought to share. In this respect, generic language concerning social categories provides the basis for stereotypes as well as for an essentialist reading in regard to the aspects of naturalness, discreteness, uniformity and the potential to draw inferences. While it is possible that the knowledge presented by Dickreiter might hold true in regard to averages or most people in a certain category, it would have been possible to present it without referencing social categories. For example, instead of presenting distinct frequency ranges for men and women and children, one could present one range of fundamental frequency parameters for human speaking voice(s), as I consider it likely that there are human individuals habitually speaking in the gap range of 160 to 220 Hz. Furthermore, instead of concluding that 'female voices' are more suitable for public announcements as they possess certain acoustic features, it would have been possible to spell out the exact parameters that make an individual's voice(s) suitable for such contexts instead of replacing it with a social category in generic language. In this sense, both epistemical and ethical issues may arise from the way knowledge about voice(s) is produced, presented and transmitted.

As I was curious about the lack of sources, I decided to compare the chapter I analyzed to an earlier edition, namely the 3rd edition from 1979. While the chapter is also called *Menschliche Stimme* and

presented as a supchapter of sound sources, it is shorter, thus containing less information. In regard to the presentation of social categories, the Hz values of fundamental frequencies in regard to the categories „men“ and „women and children“ were identical and expressed verbatim with the exception of the fundamental frequency of the vibration of the vocal folds being clarified as denotating the „Sprechmelodie“, or speech melody, in the 9th edition[169]. The aspect of components to filter out without audible differences and 'male speech' being more sensitive in regard to a droning effect are present in the 3rd edition whereas the conclusion that 'female speech' is more suitable for certain contexts is not. In this context, social categories are also referenced generically, with the difference that the terms „male speech“ and „female speech“ are included in brackets behind the Hz values instead of being incorporated without brackets[170]. While the descriptions differ slightly, the figures presented in relation to average spectra of German speech for male and female speakers respectively are identical[171]. There were no sources or references regarding how these values were determined in the 3rd edition, either. However, in the preface it is made explicit that the book is not intended as a „Lehrbuch“ or textbook but as a systematically structured presentation of modern studio technology, in which derivations and explanations have been omitted as far as possible because the main emphasis is on a summary of facts in regard to working in a studio[172]. It would seem that while there have been updates in combination with the publication of new editions, there were no updates concerning the Hz value ranges presented and no significant updates concerning the presentation of social categories. As a study of German speakers from 2017 measured a mean fundamental frequency of 111,9 Hz for male speakers and 168,5 Hz for female speakers in regard to conversational voice(s)[173], it is questionable inhowfar specific values that are reproduced over several editions and the course of 44 years still hold up.

3.2 (In)formal Knowledge about Voice(s)

I will next consider (in)formal knowledge about voice(s). I understand formal knowledge about voice(s) (3.2.1) as knowledge produced and transmitted in the context of vocal-pedagogical processes and informal knowledge about voice(s) (3.2.2) as everyday socialization, or an umbrella

[169] Dickreiter (2023): Menschliche Stimme. p. 68; No author (1979): Menschliche Stimme. In: *Handbuch der Tonstudiotechnik*. Schule für Rundfunktechnik (eds.). Revised by Michael Dickreiter. 3rd revised edition. München: K. G. Saur Verlag KG. p. 46.

[170] Dickreiter (2023): Menschliche Stimme. p. 69; No author (1979): Menschliche Stimme. p. 46f.

[171] Dickreiter (2023): Menschliche Stimme. p. 70; No author (1979): Menschliche Stimme. p. 47.

[172] Schule für Rundfunktechnik (eds.) (1979): Handbuch der Tonstudiotechnik. Revised by Michael Dickreiter. 3rd revised edition. München: K. G. Saur Verlag KG. No page.

[173] Berg, Martin et al. (2017): The Speaking Voice in the General Population: Normative Data and Associations to Sociodemographic and Lifestyle Factors. In: *Journal of Voice*. 31/2. DOI: http://dx.doi.org/10.1016/j.jvoice.2016.06.001. p. 257.e16.

term for contexts in which people are exposed to or express beliefs about voice(s) that take place in neither an academic nor a vocal-pedagogical setting. As I make these divisions in relation to the site in question, I will use academic publications among other sources when they are concerned with (in)formal lessons. I will explore two case studies by analyzing arguments concerning the relation between social categories and voice(s) based on my theoretical framework. As (in)formal lessons concerning voice(s) can be said to be a very broad site of knowledge production and transmission, I do not claim that these case studies are representative of ways knowledge about voice(s) is produced and transmitted. They nevertheless can be said to explicate potentials of how knowledge about voice(s) can be constructed.

3.2.1 Formal Knowledge about Voice(s) - Personal Experiences

In the winter semester of 2024/2025, I participated in a course about voice(s) and presentation held at the University of Cologne. At the start of the course, I had a rough idea about the topic of my thesis and thus my motivation of participation was to gain perspective on how knowledge about voice(s) can be taught practically. As I did not initially intend to include my experiences in this thesis, I did not communicate this intention to the participants or the instructor, so I did not get consent for explicit naming and thus will refer to people by their roles and use the pronoun „they“ for everyone. There was one instructor and about 20 participants, with varying numbers of attendance between sessions. While I did not have a clear research question at the beginning of the course, I was already interested in connections between voice(s) and social categories. The following is based on notes I took during and after the sessions of the course, including specific wordings and my personal experiences. I will focus on notes I took in regard to social categories, include the German wordings in footnotes and discuss aspects of my framework in regard to specific situations before elaborating on these at the end of this subchapter.

For context, I consider it relevant to articulate my positionality at this point. I identify as a white, cis-gender and able-bodied queer woman. I started singing as a teenager in the context of busking and remember feeling vulnerable about my voice(s) in that context, as performing in a public space encouraged numerous people to give me feedback on my performance and negative feedback had the potential to make me insecure and thus influenced how I perceived and used my own voice(s). I started taking an academic interest in voice(s) around 2020 and have been taking private singing lessons for about 5 to 6 years now. My singing teacher tends to focus on bodily actions to perform

while singing, which has led me to experience producing vocal sounds I had never thought possible when I was younger because of the habits I had developed earlier in my life. In combination with insights from my studies that biases surrounding voice(s) are possible, for example in the context of social categories, I have reached a point at which despite not being a professional vocalizer, I am comfortable vocalizing and am aware that any specific feedback is a perspective instead of 'the whole truth'. Because of this background, I did not participate in this course with the intention to learn about how to use my voice(s) but to gain perspective on possible processes of teaching and did not perceive the instructor as an absolute authority concerning knowledge about voice(s).

In the first session on October 8th, 2024, the instructor described voice(s) as as individual as a fingerprint and stressed that on that basis, voice(s) could be shaped and modulated. This idea was supported with the argument that even a very good imitation of voice(s) would not be identical with the original if it was measured and compared. This introduction was followed by participants sharing motivations to participate in the course. At this point, I communicated my motivation to gain perspective for my thesis. Other people expressed insecurities about their voice(s), for example in relation to teaching contexts or feeling inauthentic due to their voice(s) sounding different in different situations. The instructor furthermore implemented a vocal warm-up and told the participants to stand up and place our feet at shoulder width, with the additional instruction that especially women should take a stance wider than that[174]. My assumption was that gender categories were made salient because of the sociocultural norm that women keep their legs closer together than men and this instruction should serve as a reminder to not fall into that habit. I had already taken up the stance I usually take in the context of private singing lessons. As I identify and am usually read as a woman, I widened my stance in response to the instruction, just to realise that I was overdoing it and corrected for that again. So instead of acting on the individual, body-related instruction of shoulder width, I acted on the fact that a social category I identify with was made salient. While „especially“ can be said to be a sort of qualification in that it does not exclude people identifying with other social categories, the instruction nevertheless addressed women directly. As the instruction was voiced at the very beginning of the warm-up before voice(s) were even used or all people had stood up, it has a predictive character. In hindsight, I find it interesting that as shoulder width is an individual bodily marker, it would have been possible to keep the instruction about this factor without making reference to social categories.

On October 15th, 2024, we focused on experiencing resonance chambers or spaces. The instructor

[174] „Insbesondere Frauen über schulterbreit“

told us about how certain vowels resonate more in specific areas of 'the body' (e.g. associating the vowel 'A' with chest resonance and the vowel 'I' with head resonance), which I found slightly irritating because during my singing lessons, I had become familiar with the concept of an even vowel production or vowel balancing and thus considered the strong focus on where certain vowels are supposed to resonate simplified at the least. Participants were singled out to read out verbal exercises so their voice(s) could be listened to by the other participants. After one participant had read out the exercise, the instructor pointed out another participant and, before hearing their voice(s), told us that we had just heard a woman's voice and were about to hear a man's voice in which the 'A' will sound more powerful and clearer[175]. This can be said to be a predictive inference based on gender categories, since the participant in question had not spoken up to that point. The participants' gender was seemingly identified on the basis of visual appearance, since during the course, we were never asked for our pronouns. Formally, the utterance that the vowel 'A' sounds more powerful and clearer when it is produced by a man as opposed to a woman can be said to be a stereotype, as it is a representation of a group in respect to an attribute, feature or behaviour members of that group are thought to share. After the prediction, the participant read out the verbal exercise and while I cannot write about what other people heard, I did not hear a clearer or more powerful 'A' since the participant spoke nasally and therefore proved the instructor's assumption to be wrong in my perception. While the instructor remarked on the nasal resonance, they did not take up their previous inference again. From my perspective, it was very wholesome to hear an immediate counterexample to a generic prediction based on gender categories. While I did not talk to the participant in question, from my own experience I assume that it can also cause distress to disconfirm expectations in this context because it is implied that their voice(s) are gender-atypical, which has the potential to be hurtful.

On October 29th, 2024, there were more exercises concerning vowels and resonance spaces. In this context, the instructor described the act of keeping a sound in the neck and chest and called it a typical problem of men[176]. Later, they made a similar point about the tendency of men to keep a sound 'at the back'[177]. While a case could be made that these remarks are stereotypical in that they associate a vocal behaviour with the social category „men“, the qualifications of 'typicality' and 'tendency' imply that this is not a necessary relationship and therefore contradict a reading of these behaviours as essential traits. While I wrote down these remarks as they reference social categories,

[175] „Wir haben jetzt gerade eine Frauenstimme gehört, jetzt hören wir gleich eine Männerstimme, da wird das 'A' kraftvoller und klarer klingen.“

[176] „[...] dass ich das hier nur im Hals und in der Brust lasse, was ja ein typisches Männerproblem ist“

[177] „[...] wenn ich dann gerade als Mann dazu neige, das mehr hinten zu lassen“

I felt like they were not relevant for my use of voice(s) in the practical work of the exercises as the category made salient was one I do not identify with.

On November 5th, 2024, the instructor conducted another vocal warm-up, with a focus on an exercise termed „Stimmorgel“ during which we were instructed to hum the sound /m/ while sliding the pitch up and down and through different registers in a continous cycle. The instructor said that especially men should focus on the transition from head to chest register, as it was often the case that for men the chest resonance is strongest[178]. While „especially“ and „often“ are qualifications that contradict a reading of these vocal behaviours as necessary or essential traits, the instructions we were given as participants nevertheless varied according to gender categories regardless of specific vocalizations. While I do not identify with the category made salient, I also paid attention to the transition from head to chest register because I had found this aspect useful in vocal warm-ups before.

On November 19th, 2024, there was another vocal warm-up in the context of which women were instructed to take a wider stance than normal[179]. While this resembles the instructions concerning stances from the very first session, there was no reference to bodily markers such as shoulder width in this case. Instead, the underlying assumption seems to have been that women normally take a stance that is not wide enough for the purpose of the warm-up. This time, I did not adapt my stance because I had learned from the first lesson that my usual stance works for me in this context. In the same session, we worked with resonance spaces again by reading out exercises. The instructor introduced us to the concept of the mask resonance and contrasted it to the tendency to speak 'at the back'. Instead of only listening to individual participants, the instructor grouped us into men and women and let each group recite the exercise in unison, commenting on whether we heard the difference afterwards[180]. This grouping reflects the assumption of similarity or common characteristics in regard to gender categories, with the grouping being predicitive in the first place, then tested and „confirmed“. It also excluded people who do not identify binarily. As I only learned later in personal conversations and free talks people gave as an exercise that some participants of the course did in fact not identify binarily, I have no data on how they perceived this grouping or reacted to it.

[178] „[...] insbesondere Männer auf Übergang von Kopf- zu Brustregister achten. Es ist oft bei Männern so, dass die Brustresonanz am stärksten ist.“

[179] „Frauen spezifisch angesprochen einen breiteren Stand als normal anzunehmen“

[180] „Hört ihr?“

On December 17th, 2024, there was another vocal warm-up including the „Stimmorgel“. While men were instructed to pay attention that 'not just dirt' sounded in lower areas, women were explicitly instructed to explore low pitch areas[181]. These can be called predictive instructions or advice based on assumptions about how voice(s) of people of certain gender categories differ from each other and therefore meet the formal criteria of stereotypes. For me, it was strange to be asked to explore my low pitch areas because I am more comfortable with and used to my lower pitch areas than my high pitch areas, which is why I decided to ignore the advice and focused more on my high pitches.

On January 7th, 2025, there was yet another warm-up during which women were asked to take a stance that was a little wider than hip width and wider than usual while men were instructed to take a normal stance[182]. While this resembles the instructions of the sessions on October 8th and November 19th, shoulder width was replaced with hip width and it was made explicit that only women were to adapt their stance while men were instructed to stand as they normally do. Once again, this is a predictive instruction based on assumptions about gender categories. Additionally, by referencing the idea of normality it can be called a normalization of the socioculturally influenced idea that women keep their legs closer together than men and thus have to correct for that when working with their voice(s).

On January 14th, 2025, the „Stimmorgel“ featured heavily in the warm-up again. Before the actual exercise, the instructor said that 'the head voice' mostly works on its own for women, so women were instructed to focus on low pitch areas while men were instructed to focus on high pitch areas[183]. Therefore, the instructor gave different advice for people identifying with different social categories within the same exercise. In this case, it seemed very clear to me that not only are the instructions based on assumptions of difference between gender categories, this assumed difference can also have a direct influence on which areas of vocal potential people are asked to train and focus on. Personally, I did not like the implication that 'the head voice' basically comes to women automatically since I used to struggle with that area of my voice(s) before I started taking singing lessons regularly and so it was a significant amount of work for me to get comfortable with it. Since I am more comfortable with the lower pitch areas of my voice(s), I decided to ignore the advice again. In the same session, we worked on exercises concerning the mask resonance. In that context, the instructor said that men were allowed to feel chest resonance as well and added „men, or other

[181] „Männer darauf achten, dass unten nicht nur Dreck klingt“; „Frauen explizit dazu aufgefordert, Tiefe zu exploren“
[182] „wieder etwas über hüftbreit und Frauen breiter als sonst, Männer normal“
[183] „Kopfstimme geht bei Frauen meistens von alleine, daher den Fokus mehr auf Tiefe legen, Männer Fokus hoch“

people who tend to speak more at the back“[184]. While I was rather taken aback by the wording „allowed to“ in the context of the category „men“ due to the implication that women were not „allowed to“ feel chest resonance in this exercise, I was surprised by the addition that followed: The act of equating „men“ and „other people who tend to speak more at the back“ put my focus on the vocal habit in question instead of social categories that I do or do not identify with. It also seemed clear to me that in this case, the category was first used to stand in for a vocal habit and thus generalizes about category members before being replaced by the habit in question itself. As I do not perceive a tendency to speak 'at the back' in regard to my voice(s), I did not actively focus on (not) feeling my chest resonate.

While the site of knowledge production and transmission of this case study could be identified as academic due to the course taking place at the University of Cologne, I chose to include it in regard to formal lessons because it was about practical work with voice(s). The social categories that were made salient during the course were binary gender categories, namely „men“ and „women“. The relations between voice(s) and social categories that were presented throughout the course were assumptions of difference in regard to certain vocal features. While the focus of the course was on vocal production, this was entangled with vocal perception: Concerning the listening situation, voice(s) were present with the bodies of the vocalizers also present. There were many predictive instructions in relation to gender categories that were thus not based on specific vocalizations or voice(s), as well as diagnostic inferences about participants' gender in binary terms in combination with or solely based on visual cues instead of self-identification.

I would argue that conceptualizing voice(s) as individual like a fingerprint, stressing the potential to modulate voice(s) and claiming that there would be measurable differences between a vocal imitation and the original allows for an essentialist reading of voice(s) in regard to an individual. In relation to a person, there is a supposed core to voice(s) that cannot be imitated exactly and is thus closely linked to identity. The potential to modulate voice(s) can be read in terms of accidental properties. Even very good imitations, so 'modulated' voice(s), are said to be different from the original, which points to a core that cannot be changed, or some features of voice(s) that are beyond a person's agency. While it is possible that voice(s) are thought of having an essence and this is related to individual identity (perhaps also thought of having an essence of which voice(s) are traits), this does not make reference to or necessitate a relationship to social categories. Inhowfar

[184] „Männer dürfen bei der Maskenübung auch Brustresonanz spüren […] Männer oder auch andere Personen, die dazu neigen, weiter hinten zu sprechen“

this case study presents the categories „men“ and „women“ in the context of natural kind essentialism is debatable: Throughout my notes and memories of the course, I could not identify any indicators of whether these categories are thought of as natural kinds or not. In this sense, there is no contradiction to the aspect of naturalness. Concerning discreteness, qualifiers presented a non-necessary relationship to the vocal features in question. Nevertheless, these could be read as accidental properties, as the salience of binary gender categories could also imply that there is an essence to these respective categories that justifies this distinction. In terms of uniformity, qualifiers contradicted this aspect in relation to certain vocal features that might me viewed as accidental, while statements without qualifiers, such as that certain vowels will sound different depending on whether a man or a woman produces them, open up the possibility that in relation to certain other, possibly essential vocal features or capabilites, these groups can be said to be uniform. Regardless of whether qualifiers were present or there were generic statements in relation to gender categories, there were several predictive and diagnostic inferences based on gender categories that led to different instructions for people identifying with different categories. All in all, this case study allows for an essentialist reading of the gender categories „men“ and „women“ as distinct natural kinds that are uniform in relation to an essence that is not made explicit. While no vocal feature discussed in the course was explicitly presented as a necessary or essential component of the categories, there were assumptions of difference in regard to the use of voice(s) according to gender categories. While this could also be explained in terms of sociocultural influence or statistical tendencies, the absence of non-binary gender categories makes it seem more likely that „men“ and „women“ are conceptualized as based on the idea of sex instead of gender. If men and women are viewed as natural kinds in this context, this would be a reason for this categorization and possible sociocultural influences, with the mix of these ideas allowing for inferences with and without qualifiers because aspects of voice(s) as traits can be conceptualized as accidental in this context. While qualifiers contradict a necessary relationship between vocal features and social categories, they do not necessarily contradict stereotypes in that necessity is not a requirement in this context and even with qualifiers, references were made to vocal features or behaviours as common characteristics of a group. In this sense, natural kind essentialism can be interpreted to be an assumed basis for inferences and to have an explanatory function regarding groups and therefore stereotypes, while statements including qualifiers demonstrate that it is not necessary for essentialist notions to be present to put forward generalizing assumptions or stereotypes.

Interestingly, the gender categories made salient throughout the course seemed to stand in for vocal habits or behaviours that could also be articulated without reference to social categories, as

exemplified by the addition of „or other people who tend to“ speak in a certain way to the category „men“. The same goes for a stance that could be described in terms of shoulder or hip width, which is an individual feature that can vary significantly between members of a category, or attending to pitch areas that one tends to neglect. Even if the identification of vocal habits and certain categories is or were accurate for most members of a category, it is both less accurate than focusing on the vocal habit itself and can potentially have negative effects on people for whom the identification is not accurate, such as causing distress in relation to gender-typicality or training one's voice(s) according to categorical instructions instead of individual vocalizations and vocal habits. In this sense, instructors working with voice(s) produce and transmit knowledge and I would argue that especially because people expressed insecurities at the beginning of the course and an instructor can be seen as a figure of authority or an expert, teaching people about how to use their voice(s) is not just an epistemic but also an ethical responsibility.

3.2.2 Informal Knowledge about Voice(s) - „Race“-Conscious Casting and Voice Acting

In 2008, there was a scandal in the French film industry when news spread about a racist policy regarding the casting of dubbing actors: „[I]t was reported that casting directors deemed White dubbing actors as having „universal voices“ suitable for dubbing any kind of actor, and Black and Asian dubbing actors as having a distinct tone of voice suitable only for dubbing Black and Asian actors, respectively. The unfairness of such a policy is twofold: first, because it excludes particular groups of individuals on the basis of their racial background, and second because it rests upon a false premise – there is no scientific evidence for dividing up humanity on the basis of presumed racial differences, there is likewise no necessary biological relationship between phenotypic traits such as skin color and the quality of the voice.“[185] The idea that members of certain „races“ have distinct tones of voice(s) while White people have „universal“ voice(s) can be called implicitly essentialist in that it presents certain voice(s) as traits of social categories as discrete and uniform in regard to a distinct tone, with inferences being drawn regarding which dubbing actors can voice which characters. According to Eidsheim, the idea of a distinct timbre according to „race“, here in regard to African American timbre, can be traced back to the pseudo-science of craniometry, in which timbral differences were attributed to racialized cranial dimensions[186]. While this context

[185] Nardi, Carlo (2016): Sound and Racial Politics: Aural Formations of Race in a Color-Deaf Society. In: *Sound as Popular Culture. A Research Companion.* J. G. Papenburg / H. Schulze (eds.). Cambridge, MA: The MIT Press. p.77.

[186] Eidsheim (2019): The Race of Sound. p. 40.

would allow for a reading of „racial“ differences in relation to natural kind essentialism, Nardi argues that the rationale of racism has shifted „from a biological to a cultural explanation that justifies discrimination as a consequence of intrinsic and unchangeable cultural differences.“[187] While the conceptualization of differences in regard to cultural factors contradicts a reading of natural kind essentialism in regard to the aspect of naturalness, it can count as an example for essentialism concerning social categories attributed to non-natural factors that are nevertheless conceptualized as intrinsic and immutable. In both cases, stereotypical inferences can be drawn concerning individuals in relation to ideas held about groups they are perceived to be a member of.

In recent years, there has been a public discourse about „race“-conscious casting in audiovisual media, based on the idea of 'matching' the actor's and the character's „race or ethnicity“, in the vein of which several White voice actors have stepped down from voicing non-White characters[188]. Romano positions „race“-conscious casting in the context of representation and points out that both specific jobs in voice acting and other roles in the production of audiovisual media, such as writing and producing, are predominantly occupied by White people: „Such an extreme lack of diversity on the production side of things manifests in a lack of diversity in the kinds of TV shows, films, and games that are made. The question of whose stories get told and who gets to tell them is left up to a production that often lacks a plethora of viewpoints from the start.“[189] In summary, Romano describes how the lack of representation in voice acting is tied up with the lack of diversity in the characters presented: Especially in children’s animation, there have been „ethnic and racial“ stereotypes for a long time, for example concerning cartoon villains' foreign accents. When non-White voice actors were hired to work on cartoons, this was often on vocally performing demeaning stereotypes, which made the voice actors in question increasingly reluctant to play such stereotypical roles. According to Romano, Hollywood's reaction was to write less parts for non-White voice actors, thus increasing the rarity of „diverse parts for diverse actors“ while the stereotypes persisted in animation for children, often performed by White voice actors. Additionally, there are stereotypes concerning the vocal potential of non-White voice actors, as exemplified by Deven Mack's description of working on the 2006 TV show Grossology: „There were lots of good people there, but the assumption still the whole time [was], ‘Oh, he’s a 17-year-old Black kid,’[...] ‘The only thing he can play is the Black kid. He can’t lose his accent. We just assume that. So we’re not even going to give him other stuff to do.’ So everybody else on the show was playing additional

[187] Nardi (2016): Sound and Racial Politics. p. 78.
[188] Romano, Aja (2020): How voice actors are fighting to change an industry that renders them invisible. In: *Vox.* URL: https://www.vox.com/2020/7/22/21326824/white-voice-actors-black-characters-cartoons-whitewashing [accessed May 15th, 2025].
[189] Ibid.

voices ... and making more money than me as a result, because they just assumed I couldn't do it."[190] Given this background of stereotypes about what people sound like and related experiences of discrimination, the idea of „race"-conscious casting can be read as an issue of representation because non-White voice actors tend to be considered for fewer roles than White voice actors to begin with, and making sure that a more diverse representation of characters in audiovisual media is carried out by a more diverse cast of people behind the scenes could potentially lead to more and more accurate representations.

The idea of „race"-conscious casting is also conceptualized in terms of authenticity, as in casting characters 'authentically' in connection with „race"[191]. The idea of authenticity is also made salient on the website of Blue Wave Voiceover, a marketing collective of people working in voiceover spaces[192]: With the rider „Search Voices" on their website, voice actors can be filtered by social categories, such as „Male Voice Talents", „Asian American Voice Talents" or „Spanish Speaking Voice Talents", the latter of which leads to a subpage titled „Hispanic & Latino Bilingual Voiceovers", on which it is claimed that „[W]orking with a professional and *authentic* Hispanic/Latino voices from Blue Wave Voiceover ensures you receive a high-quality recording that you can be confident to use right away."[193] (my emphasis). Van Gerven et al. have suggested that the concept of authenticity can be accounted for by essentialism, which they conceptualize as a psychological phenomenon and bias: In the context of individual objects, the authors discuss how objects such as the Queen's spoon as opposed to any other spoon can be valued for unobservable properties. The value arises from an unobservable reality that sets them apart from similar-looking objects and thus determines the identity of a given spoon as the Queen's spoon. The authors propose that this underlying reality could take the place of a conceptual placeholder, which allows for inductive inferences and causal effects. In this sense, there are discrete boundaries between an object in relation to whether it is 'authentic' or not, with „authenticity" referring to a quality inherent in the object[194]. While the authors mention the limitation that their account refers to individual objects only[195], I would argue that ideas surrounding essentialism can be mapped to the concept of authenticity in other contexts as well: For example, to claim that a real person can voice a fictional

[190] Mack, Deven quoted from ibid.
[191] e.g. ibid.
[192] BLUE WAVE Voiceover (no date): Frequently Asked Questions. URL: https://www.bluewavevoiceover.com/faqs/ [accessed May 19th, 2025].
[193] BLUE WAVE Voiceover (no date): Hispanic & Latino Bilingual Voiceovers. URL: https://www.bluewavevoiceover.com/voicescategory/hispanic-latino-bilingual-voices/ [accessed May 19th, 2025].
[194] van Gerven, Dylan J. J. et al. (2019): From Hitler's Sweater to Dinosaur Fossils: An Essentialist Outlook on Authenticity. In: *Review of General Psychology*. 23/3. p. 371ff.
[195] Ibid., p. 372.

character in audiovisual media authentically implies that there is something or someone to be authentic to, which renders it relational. As an absolute matching of identity seems neither relevant nor possible in this context, the idea of authenticity can be read to mean that there are certain characteristics to a character to be voiced that should be matched by the person voicing it for the performance to be authentic. While it is not necessary for this conception to include natural kinds, the possible distinction between 'authentic' and 'inauthentic' performances or casting implies discreteness in regard to the proclaimed markers of authenticity, which allows for a grouping of people based on these markers and allows for the inference that if an individual is authentic in regard to certain ideas, they will enact these ideas. There are different possibilities as to what these markers could be: For example, voice actress Kristen Bell stepped down from voicing a „biracial“ character with the argument that „Casting a mixed race character w/a white actress undermines the specificity of the mixed race & Black American experience.”[196] While this does not include a direct reference to the concept of authenticity, it is thinkable that markers of authenticity could relate to experiences people make as members of social categories in an environment in which these categories are made relevant. However, it is also possible to postulate different markers, such as claiming, for example, that cranial dimensions give people a specific timbre, these dimensions are shared by „races“ and so to cast a character 'authentically', it is necessary to cast a voice actor of the same „race“ to deliver a vocal performance thought of as 'authentic' based on biological and therefore vocalized sameness to the character. As Nardi notes, auditive and visual components in voice-overs „need to reproduce the illusion of issuing from the same source. Since the basis for this audiovisual illusion is the result of sedimented discourses and practices, the criteria that guide choices in dubbing can reveal how media professionals conceive of race.“[197] In this sense, the concept of authenticity allows for multiple interpretations that proclaim some common core between the idea of what it is to be authentic and a person getting evaluated in regard to the markers that are thought to be relevant for this relation. By sharing the basic formal principle of some markers of an individual whole being made relevant in relation to broader ideas about categories and what members of these categories have in common, the concept of authenticity does not only share a similarity with essentialism but with stereotypes as well, wherefore the idea of authenticity and stereotypes are compatible: As Mack points out, stereotypes and discriminatory practices in voice acting can persist if skin color is taken to be the marker of authenticity: „[W]hen you do have people [of color] who have a specific way of speaking — if you have a lot of Black actors right now

[196] Bell, Kristen quoted from Romano (2020): How voice actors are fighting to change an industry that renders them invisible.
[197] Nardi (2016): Sound and Racial Politics. p. 79.

who sound white, quote-unquote — they're not going to book the Black roles as often."[198] While this may seem like speculation, there have been cases of voice actors experiencing a disconnect concerning expectations as to what their voice(s) are 'supposed to' and actually sound like on the basis of their social category: "Sometimes I get auditions for projects that have specifically Asian characters, and these would be required to have an Asian accent […] If a character is Asian then somehow they must be from that country, whereas my fellow voice actors and I who grew up in the States don't necessarily have those nuances – those come from our parents."[199] In such cases, there might be no argument for matching the „race" of the voice actor to the character as the relevant marker is supposed to be an accent that is demonstrably not a vocal feature people in a social category necessarily share. So while it would still be possible to engage in identity politics based on the lack of representation, the concept of „race" as a marker of authenticity seems to be based on expectations concerning the marker of specific accents in this context, and is potentially used to justify invitations to auditions. Additionally, non-White voice actors may end up reproducing stereotypes that the idea of a more diverse representation is supposed to break down: „Let's face it, the majority of people making and writing for games are white, and the voice they have in mind is a stereotypical one that doesn't accurately represent our community […] You can go into a casting and do it in your voice – therefore a black voice – but they had it in their mind as what *they* imagine as a black voice. So we're often forced to ham up an accent that we're probably not very happy performing because that's the only way we can get roles."[200] (emphasis in the original). In this case, it becomes clear that the stereotypical ideas people hold in regard to voice(s) of groups such as in accordance with the category „black", are not necessarily related to what members of the group actually sound like. As people can be confronted with this fact and still hold on to their stereotypical ideas, it would seem like the category „black" is standing in for desired vocal markers in the context of a normative instead of descriptive generic. While it is unclear whether this is a case of stereotypes being held without essentialist beliefs being present or whether other Black voice(s) are dismissed as ingenuine instances of the essence of an ideal, this idea points to audiovisual media as a possible site on which norms and stereotypes about the relation between visual and vocal markers can be reproduced. Furthering the idea that representation and authenticity are two different lines of argument, voice actor Dante Basco discusses effects of „race"-conscious casting on his career: „I always felt sometimes like they brought me in, like, 'Oh, Dante, you're an Asian celebrity, you can

---

198 Mack, Deven quoted from Romano (2020): How voice actors are fighting to change an industry that renders them invisible.

199 DaMisanthrope (2020): Liberating or boxed in: being a voice actor and person of colour in video games. URL: https://damisanthrope.wordpress.com/2020/04/15/liberating-or-boxed-in-being-a-voice-actor-and-person-of-colour-in-video-games/ [accessed May 19th, 2025].

200 Osili-Wood, Elle quoted from ibid.

play this Asian role,' [...] And so I was probably one of the earliest ones [to get hired when productions] were accurately trying to cast the character close to their ethnicity […] The pendulum swings both ways […] I don't know if [racially conscious casting] is necessarily important on every single thing ... I always tell people that if I had to wait around for a Filipino American role, I would not have a career."[201] While social categories can lead to more job opportunities for non-White people when they are considered relevant, a casting practice trying to achieve a close „racial" matching is not presented as beneficial. This also points to the question of where category boundaries are drawn: While Basco experienced an increase in job opportunities in relation to the broad social category „Asian", he also claims he would not have a career if the narrower category „Filipino American" was the only one to get auditions for, therefore stressing that representation in voice acting depends on which roles are made available for which voice actors. In this case, more representation for Asian American voice actors is not to be confused with authenticity in regard to more specific categories.

„Race"-conscious casting is also discussed on the platform Reddit, as for example in the r/changemyview section: 5 years ago, a user wrote the following original post, titled „The 'a character should be voiced by a voice actor of the same race' issue going around doesn't make sense and is hypocritical.": „With actors such as Jenny Slate and Allison Brie stepping down from their roles on cartoon shows Big Mouth and Bojack Horseman respectively, I feel like I have to speak out against this current trend. The argument that a voice actor can only play a character of his/her same race is fundamentally flawed in the current and historical state of voice acted media. What I've noticed is that the only voice actors getting called out for this are white. But if this was truly a "all race" issue, then why haven't these other voice actors been called out? Samurai Jack is played by Phil LaMarr, a black voice actor. Most of the main cast of Avatar the Last Airbender are played by white voice actors. If this discussion was truly about, "all races must be played by their respective race", then why haven't these shows been called out? In my opinion, this is a "whitewashing" issue, but it's not even really whitewashing. All I'm asking is for the people pushing this forward should be consistent in their judgement. But I feel like if there was an animated show with an all Asian cast played by Mexican voice actors, no one would bat an eye, because the voice actors arent white."[202] While the view of the user has officially been changed, the main argument presented here in terms

[201] Basco, Dante quoted from Romano (2020): How voice actors are fighting to change an industry that renders them invisible.

[202] no author (no date): CMV: The The „a character should be voiced by a voice actor of the same race" issue going around doesn't make sense and is hypocritical. URL: https://www.reddit.com/r/changemyview/comments/hjtk8r/cmv_the_a_character_should_be_voiced_by_a_voice/?rdt=52666 [accessed May 19th, 2025].

of inconsistency as only White people get called out for voicing characters that do not match their „race“ is taken up by other people on the page: For example, ekill13 writes that „[...] I can understand the argument of wanting minorities to get more chances, however, I think that people should be hired based on merit, not race. If a black person is a better voice actor and/or fits the voice needed for a white character, he/she should get the job. The same goes for an Asian voice actor, a white voice actor, or a Latino voice actor. It should be based purely on who is the best/most qualified for the job.“, Moston_Dragon writes that „It appears to me that this is an extension of the affirmative action argument; since there's a lot of white voice actors, they should be replaced by non-white voice actors to "level the playing field." Isn't that an example of racial bias when hiring?“ and Seebeedee writes that „Providing opportunities to “non-white” actors simply because they are not white is literally the exact same result as refusing to cast white actors because they are white. The roles are scarce and opportunity should be awarded based solely on merit. Race should not be considered in this hiring process and now you’re judging people based on their race rather than their skillsets and qualifications. You were correct with your original viewpoint and now I think you’re perpetuating racism and division by caving to the mob. Good luck and I hope you reconsider.“[203] The argument that is presented here is that the aim of „race“-conscious casting is to match the „race“ of voice actors to the characters they are voicing and that people who support this idea are inconsistent and hypocritical because the rule only seems to apply to White voice actors, which is described as racism as casting decisions should be made in terms of merit and skill instead of „race“. While this argument does explicitly not address issues of lacking representation and unequal working opportunities are only mentioned in terms of „wanting minorities to get more chances“ in one of these replies, it can be read as a response to an extension of a claim of authenticity in regard to an undefined essence: If the argument were made that only people of a „race“ can voice characters of that „race“ 'authentically' because of vocal markers or experiences inherent in every and only members of that „race“, it would indeed be inconsistent to only call out White voice actors for a 'mismatch'. The idea of a close matching of character and voice actor, which can be read in terms of 'authenticity', is ridiculed by the user HalfcockHorner, who writes that „The sentiment you quoted in the title doesn't go far enough. It is selective about the delineation it demands. What we need is for cartoon characters to be voiced only by cartoon people. That is the only way to eradicate the problem.“[204] I would argue that claims about „racism against White people“ in this context are based on a shift from the idea of addressing existing inequalities in the context of which voice actors are considered for which roles by demanding more representation to the idea of authenticity,

[203] Ibid.
[204] Ibid.

which is a vague concept as it does not necessarily explicate what makes a person or voice(s) authentic in regard to which markers. Ironically, the decidedly anti-essentialist position that there is nothing about voice(s) that justifies casting according to „race“ and so it is unfair if people are not cast based on merit and skill, seems to be centered around the idea that it would be fair if everyone had the same opportunities to audition for every role, a sentiment that non-White voice actors probably share.

The site of knowledge production in regard to „race“-conscious casting in voice acting is best described as informal, as discourses about this topic take place in the context of processes in the creation of audiovisual media as well as on platforms like Reddit. This also points to various actors and roles, such as production teams, voice actors and audiences. Therefore, multiple listening situations are present in this context: While audiences of audiovisual media perceive voice(s) that are matched to another perceivable body or character, the casting processes can also include auditions in which voice actors are present with their own bodies. The social category in question is „race“. Due to the vast amounts of opinions on this topic, I focused on some key arguments that I considered relevant regarding my research question. These include that „race“-conscious casting in voice acting can be viewed as (1) addressing unequal working opportunities and a lack of (accurate) representation of non-White characters as well as voice actors in the field, (2) serving the purpose of making sure characters are voiced 'authentically' and (3) an unfair ideal or practice because casting decisions should be made based on merit and skill instead of „race“. While I would argue that (1) does not directly relate to essentialism in relation to voice(s), (2) can be read as essentialist as the concept of authenticity implies a property or an essence-like cluster of properties in regard to what makes an individual authentic in regard to a given idea. Even though the markers of authenticity have not been defined or explicated in any of the contexts I reviewed, the implication that groups of people are distinct according to whether they are considered able to voice characters authentically or not, what makes individuals authentic is a common characteristic and therefore renders them uniform in regard to these markers allows for an essentialist reading. Whether the concept of authenticity in regard to „race“-conscious casting is based on the view that different „races“ are different natural kinds heavily depends on the conceptualization of the markers that are made relevant. While views of biological difference are possible in this context, I consider it more likely that, as Nardi describes, views of differences between „races“ in relation to authenticity point to a conception of cultural inherence or the experiences individuals make as members of social categories. On the other hand, argument (3) can be read as non-essentialist in relation to voice(s) as identity or the social categories individuals are members of are treated as irrelevant, with merit and

skill being brought up as the relevant factors. By implication, this argument also claims that there are no discernable differences in vocal performances according to „race“ that would render any given performance more accurate or authentic than another. In this sense, voice(s) are treated as non-essential traits of people that are neither distinct, uniform, nor allow for inferences in relation to the category of „race“. While it is still possible that people who make this argument believe social categories to possess essences, this does not give rise to an explanatory function in relation to voice(s) concerning „race“. I stated that audiovisual media have the potential to (re)produce norms and stereotypes in regard to the relation of visual and vocal markers and that stereotypes are formally compatible with the concept of authenticity. Stereotypes have not only been present in regard to the audiovisual depiction of characters in regard to „race“ but also on the production side, as in the context of White voice(s) being considered universal while Black and Asian people are thought to have a distinct timbre or examples of voice actors reporting being confronted with stereotypes about their vocal potential, such as the assumptions that people can not lose their accents or are expected to perform a specific accent based on their „race“. In the context of „race“-conscious casting, the fact that voice actors are invited to auditions for characters that match their „race“, which could be interpreted in terms of authenticity, seemingly leads to situations in which voice actors are confronted with stereotypes and normative ideas about what they and the characters they audition for are supposed to sound like. In this sense, reproducing stereotypes by performing them vocally has been viewed as the only way for non-White voice actors to get roles. As to the relationship between essentialism and stereotypes in this case, unequal hiring processes based on stereotypes about people's vocal potential that can also be read as essentialist seem to exist in parallel with stereotypical depictions of characters in audiovisual media. The idea of more or more accurate representation of non-White characters and voice actors can be read as rooted in identity politics. The goal of representation could be said to break down stereotypes or the uniformity of presentations of certain characters. In this context, argument (1) can be read as a non-essentialist argument with that goal. Argument (2) could also share this goal, but the idea of authenticity relies on common characteristics, however they may be defined. In relation to previous research in the context of psychological essentialism, the aspect of social status seems to be relevant in this case: The unification of people in relation to certain, 'authentic' characteristics could be read as strategic essentialism in relation to political demands. On the other hand, this also has the potential to (re)produce stereotypes as the factors of disreteness, uniformity and the possibility to draw inferences about people based on their social category allow for generalizations about people in regard to characteristics they are thought to share. Depending on the markers of authenticity, these generalizations might be more positive than existing stereotypes but not necessarily accurate for

every individual grouped into a social category. However, conceptions of authenticity have the potential to be filled in with people's preexisting beliefs if the markers of authenticity are not made explicit. While argument (3) does not display a direct relation to stereotypes, in the context of social status it can be read as an endorsement of anti-essentialism due to the perception of the status quo being threatened: The argument leans on the idea that it is fair to not consider the „race“ of voice actors and focus only on merit and skill in casting decisions while not addressing the issues that it seems to be a privilege of White voice actors to have been given opportunities to get cast for characters independently of „race“, that non-White characters are represented less than White characers and that stereotypes influence casting decisions, as seen with Mack's account of being reduced to „the Black kid“. The status of not only being considered for certain roles is threatened by the idea of „race“-conscious casting, as it would result in White voice actors not being considered to voice non-White characters. In this sense, the rejection of 'authentic' vocal performances in regard to „race“ and describing „race“-conscious casting as racist without addressing current issues seems to criticize this idea in order to justify existing power relations. Arguments (1) to (3) could be presented in generic language in regard to „race“ categories that could therefore have the potential to transmit essentialist beliefs: In argument (1), such possible essentialist beliefs would be grounded in the experiences people make in relation to the social categories they are assigned to and identify with. Arguments (2) and (3) do not explicate what a possible essence might be grounded in, as the concept of authenticity and the focus on „race“-conscious casting as a general rule based on the distinctness of identities instead of people's lived experiences could be read in various ways related to presupposed differences.

In summary, it can be said that essentialist ideas in regard to the performances of voice actors appear both in racist assumptions about vocal potential and possibly in regard to authenticity. Non-essentialist ideas are present in the way Nardi counters stereotypical views about vocal potential as well as in arguments that justify the status quo by claiming the dimension of „race“ is or should be irrelevant in casting decisions. Essentialist notions can therefore be connected to stereotypes against members of „racial“ categories as well as to identity politics. While identity politics in regard to this topic are also conceivable in relation to a lack of representation and unequal opportunities, adding authenticity can be thought of as strategic essentialism in advocacy for change but also has the potential to (re)produce stereotypes. In relation to stereotypes, it seems like the idea of „race“-conscious casting might not lead to more accurate representation as long as the people who cast voice actors hold stereotypical and normative ideas about how their characters and the people who voice them should sound: If characters are presented in stereotypical ways that is an ethical issue

independently of the identity of the voice actor that performs them. While casting a non-White voice actor in such a case would be an increase in job opportunities, it would also carry the potential to normalize stereotypical presentations of relationships between vocal and visual markers.

4 Conclusion

In this thesis, I set out to answer the question in what respects relations between essentialism and stereotypes are present in knowledge about voice(s). I approached this question by building a theoretical framework on the basis of the interdependencies between relevant terms and concepts that I used to analyze case studies from several sites of knowledge production and transmission in relation to voice(s). First, I compared my theoretical framework to previous academic research concerning essentialism in knowledge about voice(s) and argued that in order to explicate how essentialism and stereotypes relate to each other in regard to the idea that attributes of vocalizers become audible in their vocalizations, it is relevant to explicate social categories in question as possibly thought of as possessing essences that give rise to an explanatory function in relation to essential or accidental traits such as voice(s) that category members are thought to share. This approach focuses on the conceptualization of the relation between social categories and voice(s) and lifts contradictions that arise from conceptualizing voice(s) as possibly possessing essences. Next, I analyzed three case studies, drawing from academic as well as (in)formal sites of knowledge production and transmission in regard to whether they allow for an essentialist reading in the context of how relations between voice(s) and social categories are presented and how they relate to stereotypes: In *Handbuch der Tonstudiotechnik* (3.1.2), generic language was used to express the relation between social categories and vocal features, with specific vocal features presented in combination with qualifiers while social categories were presented as distinct and uniform. In this case, generic language served as a point of intersection between stereotypes, understood as representations of a group in respect to the attributes, features and behaviours members of that group are thought to share, and an essentialist reading concerning category-trait relations as there were no contradictions to an interpretation of social categories as natural, discrete, uniform and allowing for inferences based on group membership. Regarding my personal experiences (3.3.1), generic language was mostly used in combination with qualifiers that directly contradicted a reading of vocal features as essential and therefore necessary traits of social categories. While there were no contradictions to a reading of the social categories made salient as essentialized natural

kinds, their presence in combination with qualifiers that pointed to a non-necessary relationship revealed the possiblity that when social categories are made salient, they are used to stand in for specific vocal features or behaviours that could be addressed without referencing social categories. Given that voice(s) or vocal features can be conceptualized as accidental traits in category-trait relations that would logically allow exceptions or counterexamples to the relations presented, the salience of social categories can be explained if the category is considered informative, which matches the idea of essences being assigned an explanatory function. In 3.2.2, I pointed out that stereotypes about what social categories vocalize or sound like do not necessarily depend on essentialism as these stereotypes can persist when faced with counterexamples that undermine the idea of essential vocal markers. Furthermore, I theorized that essentialist arguments in relation to the concept of authenticity can be read as strategical essentialism in the context of identity politics, while non-essentialist arguments can be brought forward as a justification of the status quo, lending credence to the idea that the endorsement of essentialist beliefs about social categories depends on the social status people perceive to have.

In summary, it can be said that knowledge about voice(s) is produced and transmitted in discourses, which are shaped by the use of language: Essentialism can serve as a basis on which social categories are made salient in regard to vocal features or behaviours as social categories are presented as informative. In this context, the possibility of an essentialist reading and the formal characteristics of stereotypes intersect in the use of generic language due to their structural similarity. While qualifiers in the context of generic language contradict the reading of vocal features as essential traits of social categories, they do not contradict the view that social categories possess essences as qualified vocal features can be conceptualized as accidental properties of a category. Accordingly, when social categories are made salient in relation to vocal features or behaviours without the presence of necessary relations, their informativeness is rendered dubitable when general social categories are used to stand in for specific vocal features or behaviours. The salience of social categories in this context can be said to serve as a basis for stereotypical generalizations without a presentation of relations between social categories and voice(s) as necessary.

Natural kind essentialism can be described as a bias in the sense that assumptions of naturalness or immutability, discreteness, uniformity and the potential to draw inferences about individuals on the basis of the group membership attributed to them rests on necessary generalizations that allow for necessary inferences despite the lack of empirical validation of presupposed properties being

present in every single member of an essentialized group. Essentialist notions, as examples of biased assumptions, in knowledge about voice(s) can be said to be both epistemically and ethically detrimental: In regard to my case studies, when relations between social categories and voice(s) were presented in generic language, the presentation of truth claims without a mention of exceptions is less accurate than an inclusion of counterexamples. When social categories stand in for vocal features or behaviours, their salience is less accurate than the explication of relevant vocal features or behaviours themselves. These issues carry an ethical dimension because a presentation of knowledge about voice(s) that allows for an essentialist reading can serve as a basis for stereotypes about social categories in regard to vocal features or behaviours and either excludes people's voice(s) that do not match the relations presented or at least renders them non-normative. Stereotypes in the context of relations between social categories and voice(s) can not only lead to discriminatory behaviours, as demonstrated by Fasoli et al., but also influence which areas of vocal potential people train and attend to, which can be said to naturalize relationships between voice(s) and social categories. This could potentially feed back into essentialist assumptions of difference between social categories that can affect the production and transmission of knowledge. While essentialist notions can be epistemically detrimental as well as connected to stereotypes and discriminatory behaviours, they can also be used strategically in the context of identity politics and emancipatory processes, whereas non-essentialist notions can also be used to justify the status quo. Due to the epistemical and ethical implications, both essentialism and stereotypes as well as their relation have implications concerning how knowledge about voice(s) is produced, transmitted, and acted upon.

The limitations of this account include that my analysis rests on interpretation; I did not gain insight into what the people who produced and transmitted knowledge about voice(s) in these cases intended or whether they also hold essentialist beliefs about social categories when the presentation of knowledge allows for an essentialist reading. However, given the possibility that essentialism can be implicit, as demonstrated by research into the phenomenon of psychological essentialism, I chose to determine whether the way knowledge is presented allows for an essentialist reading by looking for contradictions to my conceptualization of natural kind essentialism. This approach points to the possibility of essentialist beliefs being transmitted in regard to the use of generic language, which includes the possibility that essentialist beliefs can be transmitted without the person who uses generic language actually holding essentialist beliefs. Furthermore, by pointing out the possibility of the transmission of essentialist beliefs, I have not gained any insight into whether essentialist beliefs are actually transmitted. In choosing an interpretative approach with its

limitations, I showed that conceptualizations of social categories, voice(s) and further concepts such as authenticity often remain implicit. I conclude that knowledge about voice(s) can be produced and transmitted more accurately by making conceptualizations of social categories and their relevance concerning the relations presented in regard to voice(s) as well as the potential to draw inferences explicit, including inhowfar voice(s) are conceptualized as the product of phsyical or sociocultural processes: For instance, when social categories stand in for vocal features or behaviours, the potential to (re)produce stereotypes about social categories could be eliminated by focusing on the vocal features or behaviours in question instead of social categories, which would also be more accurate. This could be applied in formal lessons concerning voice(s), as well as in decisions concerning which voice actors are asked to audition for specific roles and in the context of studying effects of biological, anatomical or physiological factors on voice(s). For example, one could study how exposure to testosterone influences the vocal tract and take the exposure to testosterone as a common characteristic for a group of people. This would be more accurate on the epistemic level as it would not allow for inferences concerning what men's or women's or children's voice(s) are like and thus eliminate the need to include exceptions or counterexamples to make statements that are true. In this context, it would eliminate the need for generalizations that include the possible reading that binary social gender categories are natural kinds. On the other hand, Fasoli et al. demonstrated that it is possible to study members of social categories in respect to their experience as self-identified members of social categories and thus explicate a way in which social categories are thought to be relevant. Further limitations include that my number of case studies was too small to determine whether different social categories can be essentialized in different ways, for example in regard to the question whether a reading of social categories as natural kinds or other kinds with inherent identity-determining clusters of properties makes more sense in a given context. While I theorized that due to the bodily component of vocal production, natural kind essentialism might be a useful framework for analysing discourses about relations between voice(s) and social categories, the aspect of naturalness did not stand out as more likely opposed to other possible conceptions of essences in my case studies. Last but not least, my conclusions refer to relations between essentialism and stereotypes in knowledge about voice(s). It is possible that different conclusions emerge when other traits are considered. In order to determine inhowfar people who produce knowledge about voice(s) hold essentialist beliefs, inhowfar essentialist beliefs are transmitted in relation to people's interpretations and the extent to which knowledge about voice(s) allows for an essentialist reading beyond the limited number of case studies I analyzed, further research would be necessary.

References

Bibbins-Domingo, Kirsten (2021): Racism and Race: The Use of Race in Medicine and Implications for Health Equity. San Francisco: University of California Health Humanities Press.

Banton, Michael (2011): A Theory of Social Categories. In: *Sociology*. 45/2. pp. 187 – 201. DOI: 10.1177/0038038510394013.

Berg, Martin et al. (2017): The Speaking Voice in the General Population: Normative Data and Associations to Sociodemographic and Lifestyle Factors. In: *Journal of Voice*. 31/2. pp. 257.e13–257.e24. DOI: http://dx.doi.org/10.1016/j.jvoice.2016.06.001.

Biesta, Gert (2007): Towards the knowledge democracy? Knowledge production and the civic role of the university. In: *Studies in Philosophy and Education*. 26/5. pp. 467 – 479.

Bird, Alexander; Tobin, Emma (2024): Natural Kinds. In: *The Stanford Encyclopedia of Philosophy*. E. N. Zalta / U. Nodelman (eds.). URL: https://plato.stanford.edu/entries/natural-kinds/index.html [accessed March 25th, 2025].

BLUE WAVE Voiceover (no date): Frequently Asked Questions. URL: https://www.bluewavevoiceover.com/faqs/ [accessed May 19th, 2025].

BLUE WAVE Voiceover (no date): Hispanic & Latino Bilingual Voiceovers. URL: https://www.bluewavevoiceover.com/voicescategory/hispanic-latino-bilingual-voices/ [accessed May 19th, 2025].

Britannica (2025): essentialism – philosophy. URL: https://www.britannica.com/topic/essentialism-philosophy [accessed March 19th, 2025].

Butler, Judith (2014): Gender Trouble. Feminism and the Subversion of Identity. London / New York: Routledge.

Cambridge Dictionary (no date): bias. URL: https://dictionary.cambridge.org/dictionary/english/bias [accessed March 5th, 2025].

Carayannis, Elias G.; Campbell, David F. J. (2019): Definition of Key Terms: Knowledge, Knowledge Production, Innovation, Democracy, and Governance. In: *Smart Quintuple Helix Innovation Systems. SpringerBriefs in Business.* Cham: Springer. pp. 5 – 15.

Cimpian, Andrei; Markman, Ellen M. (2011): The Generic/Nongeneric Distinction Influences How Children Interpret New Information About Social Others. In: *Child Development*. 82/2. pp. 471 – 492.

DaMisanthrope (2020): Liberating or boxed in: being a voice actor and person of colour in video games. URL: https://damisanthrope.wordpress.com/2020/04/15/liberating-or-boxed-in-being-a-voice-actor-and-person-of-colour-in-video-games/ [accessed May 19th, 2025].

Dekanat Math.-Nat. Fakultät Heinrich Heine Universität Düsseldorf (no date): Biases in academia and education. URL: https://www.math-nat-fak.hhu.de/en/deans-office/equal-opportunity-commissioner/translate-to-english-biases-im-wissenschafts-und-bildungsbetrieb [accessed March 6th, 2025].

DeVito, Nicholas J.; Goldacre, Ben (2018): Catalogue of bias: publication bias. In: *BMJ Evidence-Based Medicine*. 24/2. pp. 53 – 54. DOI: https://doi.org/10.1136/bmjebm-2018-111107

Dickreiter, Michael (2023): Menschliche Stimme. In: *Handbuch der Tonstudiotechnik*. M. Dickreiter / V. Dittel / W. Hoeg / M. Wöhr (eds.). 9th edition. Berlin, Boston: Walter de Gruyter GmbH. pp. 67 – 74.

Eidsheim, Nina Sun (2019): The Race of Sound. Listening, Timbre, and Vocality in African American Music. Durham, London: Duke University Press.

Employment and Disability Project handicap international (2020): Stereotypes, cognitive biases and discrimination. Factsheet. URL: https://www.hi.org/sn_uploads/MEMO-3_Stereotypes-cognitive-biases-and-discrimination.pdf [accessed March 28th, 2025].

Fasoli, Fabio et al. (2021): Gaydar Beliefs and Stigma – Supplementary Materials. URL: https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjso.12442 [accessed May 5th, 2025].

Fasoli, Fabio et al. (2021): Stigmatization of 'gay-sounding' voices: The role of heterosexual, lesbian, and gay individuals' essentialist beliefs. In: *British Journal of Social Psychology*. 60. pp. 826 – 850.

Frühholz, Sascha; Belin, Pascal (2018): The Science of Voice Perception. In: *The Oxford Handbook of Voice Perception*. S. Frühholz / P. Belin (eds.). Oxford: Oxford University Press. pp. 3 – 14.

Glazier, Jessica J. et al. (2021): The Association Between Prejudice Toward and Essentialist Beliefs About Transgender People. In: *Collabra: Psychology*. 7/1. pp. 1 – 17. DOI: https://doi.org/10.1525/collabra.25528.

Grim, Patrick; Rescher, Nicholas (2023): Theory of Categories. Key Instruments of Human Understanding. London, New York: Anthem Press.

Hall, Stuart (1992): The West and the Rest: Discourse and Power. In: *Formations of Modernity*. S. Hall / B. Gieben (eds.). Cambridge, Oxford: The Open University. pp. 276 – 320.

Hallett, Garth L. (1991): Essentialism. A Wittgensteinian Critique. Albany: State University of New York Press.

Haslam, Nick et al. (2000): Essentialist beliefs about social categories. In: *British Journal of Social Psychology*. 39. pp. 113 – 127.

Haslam, Nick et al. (2002): Are essentialist beliefs associated with prejudice? In: *British Journal of Social Psychology*. 41. pp. 87 – 100.

Haslam, Nick; Whelan, Jennifer (2008): Human Natures: Psychological Essentialism in Thinking about Differences between People. In: *Social and Personality Psychology Compass*. 2/3. pp. 1297 – 1312. DOI: 10.1111/j.1751-9004.2008.00112.x.

Husserl, Edmund (2012): Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. E. Ströker (ed.). Hamburg: Felix Meiner Verlag.

Johnson, Samuel G. B. et al. (2016): Explanatory Biases in Social Categorization. In: *Proceedings*

of the Annual Meeting of the Cognitive Science Society. 38/0. pp. 776 – 781.

Kamal, Mohamed (2021): Gender Bias in the Music Industry. URL: https://unbias.co/gender-bias-in-the-music-industry/ [accessed March 5th, 2025].

Kampourakis, Kostas (2023): Ancestry Reimagined: Dismantling the Myth of Genetic Ethnicities. New York: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780197656341.001.0001

Koslicki, Kathrin; Raven, Michael J. (2024): Introduction. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. pp. 1 – 13.

Latinus, Marianne; Zäske, Romi (2019): Perceptual Correlates and Cerebral Representation of Voices - Identity, Gender, and Age. In: *The Oxford Handbook of Voice Perception*. S. Frühholz / P. Belin (eds.). Oxford: Oxford University Press. pp. 561 – 584.

Lee, Carole J.; Schunn, Christian D. (2011): Social Biases and Solutions for Procedural Objectivity. In: *Hypatia*. 26/2. pp. 352 – 373. DOI: https://doi.org/10.1111/j.1527-2001.2011.01178.x

Mallon, Ron (2024): Race. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. pp. 361 – 374.

Mallozzi, Antonella (2024): The Epistemology of Essence. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. pp. 209 – 225.

Mandalaywala, Tara M. et al. (2018): Essentialism Promotes Racial Prejudice by Increasing Endorsement of Social Hierarchies. In: *Social Psychological and Personality Science*. 9/4. pp. 461 – 469.

Marchand Knight, Jay; Sares, Anastasia G.; Deroche, Mickael L. D. (2023): Visual biases in evaluation of speakers' and singers' voice type by cis and trans listeners. In: *Frontiers in Psychology*. 14:1046672. DOI: 10.3389/fpsyg.2023.1046672.

McKee, Adam; Bransford, Scott (no date): social categories. Definition. URL: https://docmckee.com/oer/soc/sociology-glossary/social-categories-definition/ [accessed March

27th, 2025].

Morton, Thomas A. et al. (2009): Theorizing Gender in the Face of Social Change: Is There Anything Essential About Essentialism? In: *Journal of Personality and Social Psychology*. 96/3. pp. 653 – 664.

Nardi, Carlo (2016): Sound and Racial Politics: Aural Formations of Race in a Color-Deaf Society. In: *Sound as Popular Culture. A Research Companion.* J. G. Papenburg / H. Schulze (eds.). Cambridge, MA: The MIT Press. pp. 77 – 87.

No author (1979): Menschliche Stimme. In: *Handbuch der Tonstudiotechnik*. Schule für Rundfunktechnik (eds.). Revised by Michael Dickreiter. 3rd revised edition. München: K. G. Saur Verlag KG. pp. 45 – 48.

No author (no date): CMV: The „a character should be voiced by a voice actor of the same race“ issue going around doesn't make sense and is hypocritical. URL: https://www.reddit.com/r/changemyview/comments/hjtk8r/cmv_the_a_character_should_be_voiced_by_a_voice/?rdt=52666 [accessed May 19th, 2025].

Prentice, Deborah A.; Miller, Dale T. (2007): Psychological Essentialism of Human Categories. In: *Current Directions in Psychological Science*. 16/4. pp. 202 – 206.

Rhodes, Marjorie et al. (2012): Cultural transmission of social essentialism. In: *PNAS*. 109/34. pp. 13526 – 13531.

Romano, Aja (2020): How voice actors are fighting to change an industry that renders them invisible. In: *Vox*. URL: https://www.vox.com/2020/7/22/21326824/white-voice-actors-black-characters-cartoons-whitewashing [accessed May 15th, 2025].

Rosario, Esther (2024): Sex and Gender. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. pp. 375 – 387.

Schrödl, Jenny; Kolesch, Doris (2018): Stimme. In: *Handbuch Sound*. D. Morat / H. Ziemer (eds.). Stuttgart: J. B. Metzler. pp. 223 – 229.

Schule für Rundfunktechnik (eds.) (1979): Handbuch der Tonstudiotechnik. Revised by Michael Dickreiter. 3rd revised edition. München: K. G. Saur Verlag KG.

Steinmetz, Katy (2020): She Coined the Term 'Intersectionality' Over 30 Years Ago. Here's What It Means to Her Today. URL: https://time.com/5786710/kimberle-crenshaw-intersectionality/ [accessed March 28th, 2025].

Stürmer, Stefan; Siem, Birte (2022): Sozialpsychologie der Gruppe. 3rd edition. München: Ernst Reinhardt Verlag.

Sundberg, Johan et al. (2021): One Singer, Two Voices. In: *Acoustical Society of America*. 17/1. pp. 43 – 51. DOI: https://doi.org/10.1121/AT.2021.17.1.43.

Torza, Alessandro (2024): Modal Conceptions of Essence. In: *The Routledge Handbook of Essence in Philosophy*. K. Koslicki / M. J. Raven (eds.). New York, Oxon: Routledge. pp. 105 – 123.

Trinkaus Zagzebski, Linda (2020): Epistemic Values: Collected Papers in Epistemology. New York: Oxford University Press. DOI: 10.1093/oso/9780197529171.001.0001.

van Gerven, Dylan J. J. et al. (2019): From Hitler's Sweater to Dinosaur Fossils: An Essentialist Outlook on Authenticity. In: *Review of General Psychology*. 23/3. pp. 371 – 381.

Voß, Heinz-Jürgen (2010): Making Sex Revisited. Dekonstruktion des Geschlechts aus biologisch-medizinischer Perspektive. Bielefeld: transcript Verlag.

Wikimedia Commons (2018): File:Cognitive bias codex en.svg. URL: https://commons.wikimedia.org/wiki/File:Cognitive_bias_codex_en.svg [accessed March 6th, 2025].

Wodak, Daniel et al. (2015): What a Loaded Generalization: Generics and Social Cognition. In: *Philosophy Compass*. 10/9. pp. 625 – 635.

Worthy, L. D.; Lavigne, T.; Romero, F. (2025): Culture and Psychology. How People Shape and are Shaped By Culture. Phoenix, AZ: MMOER.

Xu, Mingdi et al. (2020): Unconscious and Distinctive Control of Vocal Pitch and Timbre During Altered Auditory Feedback. In: *Frontiers in Psychology*. 11/1224. pp. 1 – 15. DOI: 10.3389/fpsyg.2020.01224.

Zimman, Lal (2018): Transgender Voices: Insights on Identity, Embodiment, and the Gender of the Voice. In: *Language and Linguistics Compass*. 12/8. pp. 1 – 16.