universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Masculinity articulated: /s/-variation as a sociophonetic resource in transmasculine German-speakers discursive co-construction of gender

verfasst von | submitted by

Julia Baldessari BA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 899
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Angewandte Linguistik
Betreut von | Supervisor:	Univ.-Prof. Dr. Jürgen Spitzmüller

Acknowledgements

First and foremost, I thank the six participants who made this study possible. I hope you find it representative and insightful, your openness and cooperation is truly what made this thesis possible. Secondly, I want to thank my academic guides: my thesis advisor Prof. Jürgen Spitzmüller for his patience, feedback, encouragement and guidance in exploring this topic with me; Carina Lozo for her confidence in my abilities, her literature recommendations, (decaf) coffee, and open ear in times of uncertainty. Wholeheartedly, I want to thank my friends Moritz Salzer and Darragh Winkelman - the former for making the *Kunstgeschichte* library a place I really wanted to go every day as well as for believing in me and being there for me when life happened halfway through the writing process; the latter for proofreading my work, making me feel smart and taking my mind off the hard things for a little while. I also thank the *Studienbeihilfenbehörde und Stipendienstelle Wien* as well as the *ÖH QueerFem Referat* and my family for their financial support. To the Department of Linguistics at the University of Vienna I am truly grateful for shaping me intellectually and also for letting me use their facilities and equipment. The professors Alan Yu and Susan Gal at the University of Chicago also provided invaluable academic insights and training which I cannot thank them enough for. Lastly, I thank all of my friends and family members who patiently listened to me rave on about the spectral qualities of /s/ without a clue what I was talking about and then cry about my choice to write about something I cared so deeply about. I sincerely thank all of you for making this thesis what it is. Without you, this would not have been possible.

ÖH
UNI WIEN

Gefördert von der Hochschüler*innenschaft an der Universität Wien

"Those who are deemed 'unreal' nevertheless lay hold of the real, a laying hold that happens in concert, and a vital instability is produced by that performative surprise. This book is written then as part of the cultural life of a collective struggle that has, and will continue to have, some success in increasing the possibilities for a livable life for those who live, or try to live, on the sexual margins." - Judith Butler (2007: xxviii)

1. Introduction

While transgender people have existed, exist and will continue to exist everywhere in the world, only few societies today explicitly protect their rights by law and even fewer acknowledge the discrimination and oppression they often face. This, however, does not make them any less real or take away from their humanity. The question of how 'real' the experience of not identifying with the gender one is assigned at birth might be haunts the conversation in the media, the public and political but also the private sphere. While the 47th President of the United States issued a decree effectively forcing around 15 thousand employees in the US military services to resign in January of 2025 simply due to their gender identity, trans rights are simultaneously attacked and stripped away around the world. Despite their proponents' ideological justifications, these attacks are real, have real consequences on real people. And much like Butler's above quote from the preface to the 2007 edition of *Gender Trouble* suggests, this reality continues to be of life-altering importance to those who live a marginalized life in 2025.

Naturally, this issue is complex and shall not be resolved in a single linguistic thesis as this. Nevertheless, what the present study hopes to accomplish is to shed more light on specific aspects of the complexity at the core of gender identity by creating a scientific contribution to the ever-present discourse. On the one hand, it does so by challenging conventional ideas about the origins of gendered differences in voice

phenomena, particularly articulation of /s/. On the other, it takes up the aspects of "performative surprise" and of this happening "in concert" as Butler worded it in the above quote: it asks how gender is not only performed but co-constructed in context. Specifically, the present study explores how /s/-variation functions as an effective resource for the discursive co-construction of gender in transmasculine German-speakers living in Vienna.

Crucially, this is not achieved by merely writing about transness and trans voices from a theoretical perspective but by taking trans individuals' experience as its focal point. Instead of positing unidirectional causalities or determinist correlations, this study employs mixed-methods exactly with the aim of painting a more holistic picture of the origins, objectives and effects of gendered /s/-variation. Self-descriptions and insights on the relationship to (their own) gender gained by interviewing six transmasculine participants created the present study's point of departure for how best to understand these individuals' articulations of /s/ which contribute crucially to the gendered voice. This thesis is, therefore, situated right at the intersection of Sociophonetics and Sociolinguistics, investigating the phenomenon of the gendered voice as it occurs in a broader societal context.

The field of *voice* is, in itself, complex and often lacks concrete boundaries. What is it that we investigate when we talk about voice? Of course, on a more metaphorical level, the term has taken on meaning especially through Bakhtin's (1981) writings on literary scholarship and was adopted into Linguistics by scholars such as Agha (2003) and Silverstein (2023). This concept, though extremely useful, is not, however, the topic of the present study. *Voice* refers here to a cluster of acoustically and articulatorily measurable phonetic features but also their semiotic lives and sociocultural settings that make the individual speaker recognizable to the interlocutor. This complex is baptized the *phonosonic nexus* and forms the basis of the present investigation. Still, this definition does not tell us much about the phenomena that stand to be investigated under this umbrella term and nothing concrete about

their origins and effects. Thus, the subsequent literature review concerns itself with a survey of the issues which the field of *voice* encompasses before gradually narrowing the scope to gendered voice phenomena and eventually the issue of gendered sibilant articulation.

In order to allow for a profound understanding of the gendered voice, the issue of *gender* requires further explication. Hence, an analysis on what *gender* is, how it can be thought of and implemented into theory for the purposes of investigating voice and how it has been conceptualized in Linguistics to begin with forms the first part of the theory section. Much has been written about *gender* and, yet again, the issue holds controversies and complexities within the discipline of Gender Studies and beyond (for a summary see Connell & Pearse 2014 or Stryker & Whittle 2006). Therefore, what is introduced in this section pertains either to previous linguistic research on the topic of gender that helps frame the current study or to frameworks employed in the qualitative analysis of the obtained data.

The second, but equally important part of the theory section is constituted by an introduction to the phonetic frameworks employed in the process of finding evidence for /s/-variability in the interview partners' language use. To begin with, previous studies on that same phenomenon illuminate how best to conduct such research and what to watch out for. Then, theoretical writings about spectral properties of sibilants and how to best measure them are surveyed in order to carve out an appropriate methodological approach.

The methods section then illustrates how exactly these methods were implemented in practice. As the study employs mixed methods and is designed as a two-pronged investigation, the first part of this section reports methods and aims pertaining to production of /s/-variation (i.e. interviews and map-tasks) and the second part for perception of /s/-variation (i.e. perception study). Subsequently, the results are described in this two-part structure before summarizing them and considering their

implications and shortcomings as well as future avenues for research in the discussion section.

This thesis sees itself as a unique contribution to a variety of fields. Its methodology can be described as sociophonetic as well as sociolinguistic, which are also the fields within which I have received formal education and have formed my expertise. On the other hand, theories on gender and masculinity form a large part of its approach to understanding how gender is co-constructed in contextual discourse which is why these fields also stand to gain insight from the present findings. What's more, the focus on context and indexicality, which can undoubtedly be detected in discussions of gender here, stems from intensive studies in Linguistic Anthropology. Hence, different approaches are consolidated in the present investigation of the research question: *(how) is /s/-variability employed as an effective phonetic resource in transmasculine German-speakers' discursive co-construction of gender.* Any inconsistencies that may arise from this approach are hopefully outweighed by the integrated nature of its findings.

2. Literature Review

a. Voice

"The voice is a complicated bundle of reactions […] no one has succeeded in giving a comprehensive account of what the voice is and what changes it may undergo." - Edward Sapir (1927: 895)

As the above quote suggests, defining how voice can be understood has been an ongoing effort in linguistic scholarship as well as various other fields. Still, it is not so easy to neatly define what we mean when we use the term *voice.* Surely, we can posit, from a conventional phonetic standpoint, that what we investigate is the physical process of sound production by the vocal folds and its acoustic results that make the speaker recognizable to the interlocutor (cf. Garellek 2017). But this narrow view on

the matter leaves out essential contributions to what sets one individual's voice apart from another's as well as the entirety of the social context in which this process takes place. For one, the physical body influencing the idiosyncratic sound of a person does not begin or end at the vocal folds (cf. Kreiman & Sidtis 2011). What's more, how we learn to engage with all our articulators and physical affordances matters greatly in what we sound like to the ear of others, maybe just as much as how these ears learned to interpret voice in the first place (cf. Harkness 2014).

On the one hand, it is essential to understand how voice production happens at a physical level – what processes underlie it and how these can vary. On the other, no serious linguistic theory is complete without accounting for a receiver or some kind of uptake – how voice becomes intelligible and therefore meaningful in the communicative act. And naturally, different dimensions and processes of meaning making can be investigated. The palpable aspect of physiological voice production has often predisposed these endeavors to biological determinism: the voice can convey information about our age, health status, sex, height, weight and many other physical realities via acoustic features which we (ostensibly intuitively) know to decode (c.f. Kreiman & Sidtis 2011). These aspects are often investigated by speech pathologists, phoneticians and psycholinguists.

But what about our position in society, psychological disposition, geographical origins, sexual orientation, group identity or gender? How do these find their way into the voice and how do we decode them, make social meaning out of them? Research on these issues has increasingly found its way into Sociophonetics (e.g. Podesva 2007; Levon 2014; Podesva & Callier 2015; Zimman 2018), Social Semiotics (e.g. van Leeuwen 2017), Sociolinguistics in general (e.g. *Cameron 1997*; Eckert & McConnell-Ginet 2013; Bucholtz & Hall 2016) and Linguistic Anthropology (e.g. Harkness 2014, 2017). These are also the fields within which the present study is broadly located. Therefore, a sociophonetic account of voice as this present study

represents is well equipped to take into account acoustic phonetics while not falling short of a semiotic approach to voice like some older approaches have done.

In short, voice cannot be disengaged from its semiotic potential if we want to do justice to its complex effects and use. Harkness (2014) extensively ponders and investigates the perceived dichotomy of physical and social properties of the voice before arriving at the concept of a *phonosonic nexus* in his observations on voice practices in Korea. He emphasizes that there is an ongoing scientific idea of voice as separate from, prior to or underlying its semiotic, cultural, communicative or social functions and that this view is condensed in what Sapir (1927) calls the "voice proper". The phantasy of a voice outside of sociality that fuels Sapir's search for this "voice proper", Harkness argues, is inherently flawed as the way one learns to produce different qualities of voice, to engage with and calibrate one's own articulatory apparatus is culturally (and therefore contextually) mediated and cannot be divorced from those social practices.

To understand this, we must first look at the underlying mechanisms of voice production and acoustic phonetics. It is hard to define where voice as mere physical sound starts and ends due to the complexity of the different processes which produce it. The airstream mechanism that underlies this process can surely be posited as the source of phonic production but what, then, about non-pulmonic sounds and stops? Do they not belong to voice? And is articulation relevant in how the voice sounds or do we just take into account the phonation process that takes place in the larynx? If so, what makes our voice sound nasal when we are sick or imitating a posh person then? Is that quality irrelevant to the larger phenomenon of voice? Let us take a closer look at the speech mechanism in order to gain a better understanding of where and why problems of categorization of the voice might occur. As already alluded to, there are four basic processes that form the speech mechanism that underlies voice: the airstream, phonation, articulatory and oro-nasal process (Ladefoged 2011: 6). As

we will see, all four of these contribute to the phenomenon of *voice* on a physical level and can hardly be isolated as the sole source thereof.

In pulmonic sound production, the airstream comes from the lungs, rises through the subglottal cavity and arrives at the glottis where it either passes unobstructed into the supraglottal pharyngeal cavity (breathing) or, through different settings of the vocal folds, begins to vibrate and hence produces sound. The sound that is produced in this way has a fundamental frequency (f0) that, for adult humans, lies between approximately 80 and 230 Hz depending on the length and thickness of the vocal folds. This basic frequency as well as multiples thereof are then amplified, which multiples are amplified depends on the shape and length of the vocal tract it passes through. These different overtones (multiples of the basic frequency) which arise and are intensified by their environment are called *formants*. The lowest three formants are what make vowel sounds recognizable as distinct sounds but also recognizably differ based on sex/gender as they derive from vocal tract length. Lastly, the voice gains its qualities from the final stage of constriction it goes through: the oral or nasal cavity and their respective articulatory settings.

Simply put, a rounded mouth produces a different sound than a lax, open jaw due to the shape and length of the resonating cavity which amplifies and dulls specific frequencies. All of this is also known as the Source-Filter Theory (cf. Ladefoged 2011) because it conceptualizes the airstream and larynx as the source of a very basic sound (fundamental frequency) which the vocal tract then inscribes its idiosyncrasies on ('filters' it) before it is eventually also shaped by the environment after output. In much the same way, a nasal vowel that has to pass through a specifically shaped nasal cavity, thereby taking on specific acoustic properties, can not only carry differences in grammatical meaning as opposed to a non-nasal one (depending on the language) but can also lend a spoken segment social meaning.

What is often conceptualized as *phonation type* (Podesva 2013) herein deserves a separate mentioning as it constitutes what is generally often referred to as *voice*

quality. The glottis, of course, does not simply operate on an 'open-closed' binary but rather on a continuum resulting in various constellations of openness. Modal voice, which can be described as the most common setting for speaking voices, is produced when the cartilages that hold the vocal folds (arytenoid cartilages) are close together, creating a closed glottis that allows air to stream through a rather tight constriction between the vocal folds, thereby creating sound. The tightness of his constriction and the resulting tension in the vocal folds during phonation as well as the position of the cartilages relative to each other create various different voice qualities which will be introduced in short here. The following descriptions of physical processes resulting in certain phonation types are according to Laver (1980: 93-140) and what he calls "phonatory settings".

Creaky voice (as is conventionally associated with phrase finality in different languages (Davidson 2020)) is also created by the cartilages being very close together like in modal voice, creating tension in the vocal folds but instead of an opening the length of the same, the air is only released through a small glottis opening towards the back of the larynx which leads to much more pressure than modal voice. Creaky voice is hence produced with compressed and thick vocal folds which results in low f0 and a low air-flow rate. Breathy voice is produced with the cartilages apart and less tension in vocal folds. This means that the glottis is kept open along most of its length and the vocal folds vibrate without ever closing, resulting in high air flow. Falsetto results from stretched and adducted vocal folds which then results in rapid vibrations and, consequently, a high f0. Whispery voice and harsh voice result from a high degree of laryngeal tension and are acoustically recognizable by jitter or aperiodic f0. What's more, lowering and raising the larynx also create voice qualities by shortening or lengthening the vocal tract which results in formant differences. The effects of both will be discussed later as they highly pertain to gendered speech production.

It is, in conclusion, not just the laryngeal phonation process that we perceive on an acoustic level when we hear a person's voice and even that process in itself bears complexities which go beyond a simple pitch range. The entirety of this speech mechanism and how its outcomes are achieved can carry meaning. Think only of a crying person's breathing when they speak and what that transports to the interlocutor, the sound of a pouty mouth and what we associate with it, how whispery voice can immediately index a social setting or how the raised larynx of someone imitating a small girl is unmistakably recognizable. Consequently, if we want to do justice to its complexity in production as well as perception, voice cannot be thought of as a mere range of sonic frequencies produced by the larynx in the phonation process as scholars in acoustic phonetics often have (e.g. Garellek 2017). Because every part of the voice production process filters these frequencies in a unique way and is thereby inscribed in the final soundwave it must be conceptualized as a complex cluster of sound qualities that form a whole in the mind of the listener.

Harkness (2017) conceptualizes these 'qualities' as *qualia* in a Peircean sense: culturally communicated potentialities (types) that are instantiated in speech as qualities (tokens). The fundamental frequency as well as the peaks of resonating acoustic energy (formants) which are shaped by the composition of the aforementioned multitude of points of articulation (the entirety of the vocal tract) are perceived in particular acoustical spaces where specific arrangements of the former are set as the norm. Think of whispering at a funeral, preaching in a church, hollering at sports events or seducing someone at a bar – these are not ubiquitous but instead learned practices of the voice as opposed to simple vocal fold settings. The way one learns to arrange and use one's articulators, oro-nasal cavity, larynx, vocal folds and breathing is not a process that can be divorced from the local, cultural and contextual life of the individual. Anyone who has grown up in a loud family and has ever been asked to lower their voice in a public place will attest to that.

This also ties right into the next step in the speech chain: voice perception. Psycholinguists have long discovered that our brains learn to perceive sounds categorically at a very young age (Kuhl 1979, 1987). This means that the sound waves that reach our eardrum undergo yet another filtering but this time by our brains. The McGurk effect (McGurk & MacDonald 1976) is one of the most poignant examples thereof: visual information such as mouth movements influence which sound we think we hear. Similar effects have been reported not only for mouth movements: social priming effects regarding speaker origin influence hearer judgement of vowel quality in Niedzielski (1999) and Hurring et al. (2022), Munson et al. (2017) show that perceived speaker gender has an effect on perceived sibilant quality and Johnson et al. (1999) report that speaker gender even influences ratings of vowel quality when the listener is merely asked to imagine a male/female speaker.

Another problem of the voice is that it is not a uniform occurrence over time. Foulkes (2010) also remarks that what makes the individual recognizable through their voice is not one distinct feature or even a cluster of features that are permanently present. He argues that voices are rather plastic and, as we will focus on in the section *Context and Indexicality*, can only ever function to convey individuality through a complex of interacting features which are not ubiquitously available to index identity. Using, for example, creaky voice as a young woman in the United States in 2010 may have indexed a certain upwardly mobile female persona to a specific group of people (Yuasa 2010) while doing so in 2040 might not yield the same results anymore. Individuals change and adapt their voices over a lifetime according to a great many social factors, be they conscious or unconscious, which makes it hard to think of the voice as a singular identifying factor. This was recently revealed to me personally when a friend I hadn't seen in years crossed my path and after some conversation, remarked "Wow, I hardly recognize your voice!".

Hence, while some of the factors influencing the voice might be more permanent (i.e. biological variation), these factors rather constrain variation than determine it. One of

the most important points we can, therefore, take away from Harkness' (2014) ethnographic work is that all aspects of the voice should be considered as a culturally meaningful whole instead of separate phonetic, sonic and social occurrences. Whether variation in the voice is biologically or socially acquired, none of its idiosyncratic and dynamic qualities or *qualia* make sense outside their context. I follow Harkness (2014: 16) when I say that "there is a culturally conceptualized, sociohistorically normative, prototype-based voice" and we position ourselves, with the use of our own voices, towards this prototypical voice practice.

Just as much as speakers can position themselves in society using their lexicon, syntax, morphology, body language and other semiotic modalities and practices, the voice must be understood as one such modality instead of a mere complex of physiological affordances. It follows that, just like other modalities, the voice is acquired by an individual in an array of social settings and has the potential for continuous development, simultaneously adopting, embodying and shaping discourse on itself. The individual is hence neither limited to nor free from normative voice practice in their respective environment. Going forward, I will, therefore, make use of the idea of the *phonosonic nexus* as coined by Harkness (2014). I do this with respect to the multifaceted nature of the present study encompassing quantitative as well as qualitative research and data. A detailed account of how gendered socialization is reflected in different types of voice practices will be explored in the next section.

b. Voice and Gender

The frequency code (Ohala 1994) is an oft-employed tool in attempting to understand the meaning and origins of gendered phonetic variation. It is based on the assumption that all vocalizing species intuitively understand and make use of an ostensibly natural correlation between pitch and size as well as a subsequent link

between size and displays of power and threat. This code is further assumed to apply not only to natural sound but verbal language, too. Higher pitch, in this line of reasoning, is consequently linked to smallness, non-threatening behavior and positive intention. Lower pitch, on the other hand "conveys largeness, threat, self-confidence and self-sufficiency" (Ohala 1994: 343).

As Eckert (2017) argues, claiming this connection to be universal or even natural in language should be seen as an overstatement at best. Like any symbolism, sound symbolism, too, is subject to culturally mediated conventionalization and does not directly import natural sound into language. These conventions not only pertain to the relation between language and sounds in nature but also to the characteristics taken to be represented by them. The physiological opposition between large and small becomes an affective one through salient social evaluations and judgements, which size has culturally acquired, when this association between size and stance is entirely a matter of language ideology (Silverstein 2023).

Nonetheless, the frequency code has persisted in informing sociolinguistic and, in extension, sociophonetic research and has, thereby, largely contributed to a common view of sexual dimorphism as the origin of all gendered phonetic variation. For any voice phenomenon that is reliant on frequency (such as pitch, voice quality, vowel space, formants or sibilants), this has resulted in a binary framework within which high frequency correlates to smallness and female sex, low frequency to largeness and male sex. And even with a lack of significant evidence pointing towards biological differences between male and female palate size (e.g. Fuchs & Toda 2010) as well as a considerable number of reliable accounts of learned articulatory behavior (as will be exemplified below) many scholars seem to refuse to abandon such deeply rooted, determinist assumptions in their discussions (e.g. Fuchs & Toda 2010, Weirich et al. 2016).

Theorists concerned with language ideology (Gal & Irvine 2019; Silverstein 1979, 2023; Woolard & Schieffelein 1994) assume that such sociocultural beliefs as the frequency code being the source of gendered phonetic variation rely on indexical

iconization (Silverstein 2003; 2005). Hearers ideologically rationalize and naturalize semiotic practice as essential, often by locating its origins purely within the body. However, metapragmatic discourse on language use and its origins is not only shaped by actual language use in a bottom-up way, it also shapes language use as a top-down function itself (cf. Gal & Irvine 2019). Yet, this view is not always considered in discussions on the origins of gender differences in the voice. A selection of relevant literature will be elaborated upon in the present section with two aims: for one, it is important to evaluate what is already known about gendered voice phenomena in order to get closer to our research interest. What's more, it is equally important to critically examine how these findings were accrued and to question biological determinism and sound symbolism as they are possible sites of indexical iconization. These steps not only bring us closer to the research gap of this study, they illuminate different approaches to the field and their problems at the same time. We begin by looking at formants since they allow us a representative insight into how difficult it is to tease apart metapragmatic discourse, language socialization and biology. As formants heavily depend on the length of the pertaining vocal tract, assuming the lowering of the larynx in pubertal boys as the origin of gendered differences in formant structure seems a sensible approach at first glance. The question of language socialization has, however, become increasingly relevant in this regard as more and more findings report these differences to already be present in prepubescent speakers where no hormonally based sex differences are expected (Busby & Plant 1995; Lee et al. 1999; Perry et al. 2001).

Zimman (2018) makes the point that, while formant differences surely emerge quite clearly during and after puberty, the above findings suggest that a broader effect of language socialization might be at work here: the lowering or raising of the larynx in childhood. By lowering the larynx, for example, prepubescent boys might elongate their vocal tracts and produce formants which are more normatively masculine than those of girls. Respectively, it is possible that girls of the same age group raise their larynxes to create a shorter vocal tract and more normatively feminine formants.

Simpson et al. (2017) indeed find that such practices come into play when children are asked to speak like a girl/boy. This process might depend on language socialization as an implicit force as well as explicit metapragmatic discourse on voice pitch differences since lowering and raising the larynx also produces lower and higher pitch.

Various perception studies also point in the direction of gender identity being indexed by children through voice cues (e.g. Perry et al. 2001; Munson et al. 2022; Funk & Simpson 2023;). It is, therefore, highly likely that a process which initially depends on hormonal development is preponed in linguistic development as a part of gender socialization. This does not mean that hormones have no influence on vocal tract length but that the effect they do have might be exaggerated due to the iconicity the process takes on and the place this iconicity is (implicitly or explicitly) granted in language acquisition.

As for voice quality, a similar pattern emerges. Podesva (2013) observes that not only is the use of different voice qualities of much more intersectional and interactional nature than previously thought, it is also important to rethink which voice qualities we consider as marked when it comes to gender. In his study, Podesva suggests that creaky voice had previously been thought to occur more frequently in male speakers due to iconic associations of its low frequency with lower pitch. This association is then indexically linked to stances usually associated with masculinity like toughness as Mendoza-Denton (2007) describes in her ethnographic work with Latina gang girls. Similar indexical iconization may take place in associations of falsetto with higher pitch and, in extension, femininity and stances associated with the same. An especially apt example thereof comes from Podesva himself (2007) – an exploration of how a gay man uses falsetto to construct a gay diva persona. Again, we see a naturalization of the form-meaning connection which is embedded in language ideology (Gal & Irvine 1999) and must therefore be examined accordingly. Unsurprisingly then, the real use of creaky voice as well as falsetto is less straightforward. In Podesva's (2013) survey of White and African American residents

of Washington DC, creak is found more predominantly in women than men, regardless of race. Falsetto is reported to occur more often among women than men, but this is attributed to especially high falsetto use among African American women. Whispery voice, however, was shown to be used more by White women in comparison with White men. However, this gender difference disappears in combination with race – African American women and men use whispery voice to the same extent. These findings illustrate two important points: firstly, it does seem to be indexical iconicity which commonly links voice qualities to gender and not some biological determinant. Secondly, Podesva comments that, had he not investigated race as a possible factor, the picture of whispery voice as the marked style in females as opposed to males may have emerged. By including this information, however, the data suggests that the White male under-use of whispery voice seems to be the marked style in comparison with both White and African American women and African American men.

Similar observations regarding pitch and gender are made by Yuasa (2010). Her finding that young, educated women in California employ creaky voice much more often than their male as well as Japanese counterparts underscores that voice quality underlies sociocultural factors as well as gender socialization. She also elaborates on the indexicality of creaky voice when used by these women: listeners judged them to be hesitant, nonaggressive, and informal but at the same time educated, urban-oriented, and upwardly mobile. Such evaluations together with listeners' mentioning of regularly encountering this voice quality in young women point in the direction of creaky voice as a register in US women's use of voice which was forming at the time. Yuasa (2008) also argues that voice pitch and range differences between men and women are in part the product of sociocultural expectations regarding gender performance as well as interactional considerations such as politeness. Taken together, what all of these findings make clear is that the gendered voice is the site of multiple semiotic processes: language acquisition, metapragmatic discourse and biology. Both Zimman (2018) and Foulkes (2010) also argue that sexual dimorphism

can be posited as the playing field within which voice functions, a restriction rather than a determinant. It is in this light that the present study investigates how the gendered voice is not only produced but also perceived.

In order to do this, we will assume, as Bucholtz & Hall (2016) remark, that from the viewpoint of semiosis, there is no conceptual difference between discursive and embodied resources. Hence, it does not so much matter *where* certain features actually originate but rather *how* they gain meaning. The processes by which they are viewed and employed in certain ways are far more relevant to understanding the gendered voice than their exact origins. A great many of these semiotic features related to gender in the voice have been investigated and a selection shall be reviewed here to create a picture of where the present research is situated, beginning with formants.

Although formants are not only related to vocal tract length but also vowel quality they have been shown to vary between genders beyond what is attributable to mere physiology: differences in vowel space manifest in females having higher f1 values for open vowels and higher f2 values for closed front ones in Dutch, English, German and Swedish (cf. Henton 1995). Wassink (2001) also finds that such differences in vowel quality depend on gender, theme and setting in Jamaican. Gender differences in vowel space as measured in formants have been reported quite early on in (socio)phonetic research (Peterson & Barney 1952; Hillenbrand et al. 1995; Henton 1995; Diehl et al 1996) and have continued to be of interest to the field (Simpson & Ericsdotter, 2007; Whiteside, 2001; Weirich & Simpson, 2014).

Furthermore, vowel space differences have also been found regarding sexual orientation (Pierrehumbert et al. 2004; Munson et al. 2006a). While queer women produced lower f1 or f2 frequencies than straight women for the vowels /ɛ/ and /oʊ/ respectively, queer men produced higher f1 frequencies for /ɛ/ and /æ/ than straight men (Munson et al. 2006a). Pierrehumbert et al. (2006) find that vowel space is altogether larger and more distinct for queer speakers when compared to their straight counterparts although this effect was not attributable to all vowels in the

same measure: for queer women the back vowels contributed more prominently to larger vowel space while for queer men it was three out of five tested vowels that showed larger dispersion with /æ/ showing the most distinct articulation. Rogers & Smyth (2003) also report that the more 'gay-sounding' men were, the more likely they were to produce vowels closer to the periphery of the vowel space than less 'gay-sounding' men.

Taken together, the trend is that female and queer individuals produce more 'extreme' vowels, albeit not all of them to the same degree. One avenue of explanation for this phenomenon has consistently been that of 'clear speech': it is possible that hyperarticulation is indexical of feminine speech while hypoarticulation is associated with masculine speech. Much of the literature on the subject equates 'clear speech' with hyperarticulation (de Jong 1995) and posits it as the 'marked' style as opposed to 'unmarked' male speech. This 'marked' style can then be extended through indexical iconicity to sexual orientation – features of 'clear' articulation that are usually associated with femininity become indexical of sexual orientation by transferring their meaning potential in an act of fractal recursivity (cf. Gal & Irvine 2019). This means that the features associated with feminine speech rely on the 'feminine-masculine' axis of differentiation which is applied to all individuals equally. When men are consequently evaluated on this same axis of differentiation, gay men more frequently fall on the 'feminine' side and their speech is heard as more 'feminine' than that of straight men. What must, however, be considered is that a similar effect as in the findings concerning whispery voice by Podesva (2013) could be assumed: if queer men and women as well as straight women show some level of hyperarticulation, we must operate under the assumption that straight male speech is the marked, underarticulated style.

Clear speech has been linked to longer segment durations and slower speech rates (Byrd1992), larger vowel space (Diehl et al. 1996; Weirich & Simpson 2014) and larger f0 excursions (Mennen et al. 2012). Male speech, on the other hand, has been described as 'mumbling' which is linked to sounding 'macho' in contextual gender

construction (Heffernan 2010). More studies on this phenomenon show that men do in fact show a lesser degree of jaw opening resulting in lower sonority (Harrington 2000). Weirich et al. (2016) also find that women show greater variability in jaw opening than men which allows them to mark different accent conditions more distinctly.

What all of these studies suggest is the importance of context. Not only does the range of the one feature that is so commonly associated with gender in the voice, f0, vary significantly based on language (Mennen et al. 2012), the difference between male and female f0 is also sensitive to cultural context in the same way. Van Bezooijen (1995), for example, finds larger gender differences for f0 in Japanese than Dutch and Weirich et al. (2019) find similarly large gender differences for German vs. Swedish not only for f0 but also vowel space. While Wassink (2001) locates gendered differences at an intersection of gender, social status and context, Heffernan (2010) also argues for mumbling as part of a gendered 'macho' identity that is constructed during discourse.

As for vowel space, Weirich & Simpson (2014) find that larger vowel space can be observed for men in accented conditions as opposed to unaccented ones. The abovementioned findings by Yuasa (2008; 2010) also exemplify how sensitive f0 and voice quality are to cultural expectations: which quality and pitch a woman's voice exhibits is not simply a function of her gender identity but more so of how this identity is negotiated in different macro and micro contexts. Naturally, the male voice and its often underarticulated features are subject to and results of these same processes.

In summary, viewing either gender, sex or sexual orientation as ubiquitous causes of differences in the voice is a reductionist enterprise. Linguistic and ethnographic research on the subject must take into account the language, cultural environment and the different relevant social contexts it operates in. In order to do justice to this complexity, we now turn to what we know about gendered voice phenomena in German.

Funk & Simpson (2023) find evidence that the abovementioned f0 differences between prepubescent boys and girls are also quite pronounced in German-speaking children. They also underline that self-reported gender conformity has a large effect on how children articulate and are perceived. A cluster analysis revealed that children with higher self-identification as female were also more likely to produce higher f0, larger vowel space and sibilants with higher Center of Gravity (CoG) and lower negative skewness in sibilants. What's more, children who were strongly identified by listeners as male had lower f0 than those with more ambivalent ratings who showed average high f0 as well as high speech tempo (typically associated with girls) while exhibiting sibilants with low CoG and high (positive) skewness values (typically associated with boys).

What these findings underline is that there is a range of acoustic parameters that may facilitate gender perception for prepubertal German-speaking children. However, such perceptions do not seem to rely simply on one parameter or a specific cluster thereof. What emerges from this study is the observation that either one of those parameters (f0, speech tempo, CoG and skewness of /s/) can be used as an index of gender identity for children in German but the logic by which the indexical process is anchored in one feature rather than the other is not as straightforward as a simple form-meaning connection. This is also in line with findings on the gay voice in German.

Vorberger (2024) undertakes this field by investigating gay speech production as well as perception and finds that different phenomena can contribute in different measure to a voice being perceived as gay: nasalization of non-nasals as well as denasalization of nasals, (punctually extreme) increased f0, raised f1 and f2 in /ɛ/, (punctually extreme) elongation of vowels, higher articulatory precision (resulting in longer consonants, full realization of otherwise reduced vowels, fortification of lenis obstruents and elevated aspiration) as well as increased general intensity in speech. Kachel et al. (2020) additionally find that multiple cues for sexual orientation in men

led to higher accuracy in judgements thereof, underlining the complexity of the indexical process.

These findings resonate what Levon (2007) also claims for English gay speech: a combination of pitch range and sibilant duration may be used to index gay identity albeit contingent on other sociolinguistic factors. All of these insights again show that the gendered voice is always a product of a multitude of linguistic and contextual factors. Levon (2014) pertinently discusses that perceptions of gayness and masculinity being indexed through sibilants and pitch are highly contingent on listener attitudes towards gender roles as well as existing stereotypes. It is in this light that we can interpret the findings of Weirich & Simpson (2018), too: male speakers who identified themselves as more feminine showed larger vowel space and higher mean f0, the resulting gender identity was perceived more distinctly by female participants. This means we cannot simply postulate that gender identity or sexuality is indexed to every listener in the same way through the same features, not even if all listeners are German speakers.

Therefore, processes of indexicality are highly contingent not only on speaker identity but also on listener attitudes and other contextual factors pertaining to both speaker and listener. These processes are at play in German as well as English (as well as any other language for that matter) which reifies the necessity for moving beyond simple sound-meaning mappings and towards emerging meanings for specific interlocutors in specific contexts. In this light, we assume that sibilants may be *part* of how speakers index gender and/or sexual orientation in German as well. The research question, therefore, comes into closer reach. But first, we must ask: how exactly do gendered meanings for /s/ emerge in German and how much do we know about them so far?

c. /s/

In his 2001 collection of short stories *Me talk pretty one day*, David Sedaris remembers how, in his boyhood, he was selected by his teacher to undergo speech therapy due to his supposed 'lisp'. He relays how all of the children who were diagnosed with such a 'lisp' were boys with a shared disinterest in college sports, never popular kids and never a single girl. The conclusion he draws from this experience is that the speech therapist may as well have put a sign on her door reading "future homosexuals of America". With this story, Sedaris (born in 1956) opens a door to the lived experience of many gay and queer men in the US – back then as well as today. The phenomenon of the 'gay lisp' is the stuff of commonplace as well as scientific discourse on gay male speech. But while we know that a fronted articulation of /s/ is perceived as gay in English (e.g. Munson 2007; Munson et al. 2006a, 2006b, 2017), other languages demonstrate different indexicalities associated with this sound or simply have not been investigated on the matter.

Let us start by looking at /s/ and its indexical field. The special case of /s/ in Sociophonetics is illustrated not only by multiple earlier studies concerned with the topic (e.g. Heffernan 2004; Stuart-Smith 2007; Holmes-Elliot & Levon 2014; Boyd et al. 2021) but also by a special issue of *Linguistics* (Levon et al. 2017) entitled *The Sociophonetics of* /s/. In it, Eckert (2017) argues, first and foremost, for a critical view on sound symbolism as mentioned above and investigates the semiotic potential of the phoneme. Eventually, she reaches the conclusion that /s/, from a standpoint of Peircean semiotics, may represent the *most perfect of signs* as it encompasses not only symbolic but also iconic and indexical sign dimensions at the same time. Yet, in order not to get lost in the semiotics of this potent linguistic feature, we will focus here on previous work that focusses on gendered aspects of its social use.

To begin with, we revisit seminal findings of an imaging study by Fuchs & Toda (2010) who find that women produce a more front place of articulation for /s/ than men, resulting in a shorter front cavity and consequently higher frequency. In an

auditory analysis, Flipsen et al. (1999) also find a larger number of dentalized or fronted tokens in women's speech when compared to men. And concerning perception, Strand (1999) finds women with fronted /s/ to be perceived as more feminine. Munson (2007) also underscores this with his findings that more fronted /s/ is associated with femininity and argues that the constructs of perceived sexual orientation and gender are therefore not to be confounded. Interactions of the two, however, shall be discussed later.

In their study with Mandarin speaking children, Li et al. (2016) make the more detailed observation that gender differences in sibilant articulation arise in early childhood and persist throughout childhood and adolescence. What's more, they observe that gender identity has significantly more effect on this than height or age. Li (2017) also reports that gender differences in sibilants emerge clearly at age 4 and become more distinguished at age 5 for Mandarin speaking children. However, the adults in the study did not show the same gendered patterns for sibilants. This suggests that a change in progress regarding gendered indexicality of sibilants is taking place in Mandarin. This leads us to the question if such gender differences are also present or evolving in German.

Unfortunately, not much research has been conducted on the matter so far. Simpson & Weirich (2015) report that sibilant contrasts are stronger for women than men in much the same way for German as for English. What's more, Funk & Simpson (2023) suggest that prepubescent children's gender self-identification is significantly correlated to their production of /s/, yielding results (similar to Li et al. (2016)) showing higher CoG for girls than boys. They also find that CoG for /s/ may be one of a cluster of features which are used to ascribe gender to children. For adults, Weirich & Simpson (2018) cannot find any evidence for sibilants indexing gender identity. These findings regarding gender indexicality are slowly leading us to the research gap for the present study. First, however, let us look at how gendered

meanings for /s/ might interact with other identities as we will be looking at sexual orientation and gender in concert in this study.

Stuart-Smith (2007), for example, investigates how gendered meanings of /s/ interact with class and age in Glaswegian. She argues that no single category fully accounts for /s/ variation based on her findings of consistent interactions of gender, age and class. The construction of particular social identities which relies on various constellations of (non) working-class, young or old and male or female identity results in patterns of fronted /s/ use that are not easily explained by a simple gendered category. As we have seen before, /s/ is hence no exception to the pattern that simple sound-meaning mappings are to be generally questioned. Wherever we detect gendered meanings of /s/, they inevitably come packaged in a larger envelope of identities and contexts associated with or explicitly opposed to gender. It is, therefore, important to view /s/-variation as a resource that may do more social semiotic work than to simply index a singular identity.

Pharao et al. (2017) also find that not only is fronted articulation of /s/ indexical of gayness and femininity in Danish, it is also associated with a 'street' identity that relies on the iconization of this fronted /s/ as foreign and non-Danish. The indexicality of the feature as either gay/feminine or street/immigrant is contingent on the cluster of other linguistic features it co-occurs with (i.e. stops and prosody). As Levon (2014) also finds, more salient phonetic features can block indexical perception of /s/. Other studies, however (Levon 2007; Campbell-Kibbler 2011) suggest that clusters of features can evoke additive or alternative judgements instead of producing blocking effects. It is, hence, not entirely clear how indexicality of /s/ functions. Levon (2014), therefore, argues that the indexicality of such features needs to be individually assessed and investigated for each linguistic phenomenon instead of making blanket claims which may miss their mark.

It is, therefore, necessary to clearly define which indexicalities of /s/ we are hoping to investigate here and how these might interact. This now finally brings us to the

research question and research gap the present study sets out to fill. As Zimman (2017a, 2017b) finds that /s/-variability in his transmasculine English-speaking participants may not be, as aforementioned, straightforwardly mapped to varying gender identities or sexual orientations, he concludes that /s/ seems to be used in a manner of stylistic bricolage. Its indexical potential can be used to construct a gendered persona and/or achieve discursive goals related to the speaker's desired perception. Fronted /s/, for example may be used (in varying degrees) to index gender or sexual orientation in concert with other such features (like f0, voice quality, etc.) just like one can decide to wear high-heeled shoes (of varying height) to index femininity in concert with other accessories.

In short, Zimman argues that, for transmasculine speakers, /s/-variability is one setting in the mode of voice that can be employed in order to construct an identity which encompasses gender identity, gender presentation and sexual identity. /s/ can, therefore, be understood as a building block in the indexical field of the gendered voice. Although Zimman (2017a) takes very good care to understand his participants' identities and how they might relate to their use of /s/, his studies so far do not broach the issue of uptake. Naturally, as the above literature suggests, fronted /s/ seems to be robustly enregistered in English as gay and/ or feminine sounding for men. Yet, the intersection of the two identities might yield even more interesting results on a perception level.

All of these findings led me to wonder: if fronted /s/ is such a well-documented phenomenon regarding gender and sexual orientation in English production as well as uptake, is there a possibility that the phenomenon has made it into German? Boyd et al. (2021) find that, although fronted /s/ is produced by gay cis male German speakers, no uptake of it sounding gay seemed to take place on the listeners' side. To the contrary, Kachel et al. (2017) observe that, although speakers were perceived as sounding straighter the lower their CoG in /s/ was, the actual production of /s/-variability could not be detected for queer male German speakers. It seems then that

some gendered identities (gender self-identification) may be in indexed through /s/-variability (Funk & Simpson 2023) but that the feature neither indexes other such identities (sexual orientation) nor has become salient enough to be enregistered.

These conflicting results left me with the question if there might be a change in progress happening which is not yet detectable or stably enregistered. Boyd et al. (2021) indeed argue that fronted /s/ may be waiting for its baptismal moment to gain social indexicality in German, not yet fulfilling its gendered indexical potential. This argument is underlined by the observation that one of their German-speaking participants mentioned knowing of the gay lisp in English but denied the phenomenon's existence in German. In light of this, my research interest was to investigate how different gendered identities may be indexed through /s/-variability in German.

By trying to replicate Zimman's (2017a) findings of gendered /s/ production for German as well as replicating Boyd et al.'s (2021) study design in hopes of finding newer information on perception of the phenomenon, I aim to fill a gap left by the literature regarding how /s/-variability is used to index gendered aspects of identity in German. The choice to investigate this phenomenon by interviewing transmasculine individuals is reflective of my interest in the complexity of this process. Hence, the research question emerged:

(How) is /s/-variability used by transmasculine German-speakers as an effective sociophonetic resource to discursively co-construct gendered identity?

The hypothesis that it indeed is points to my larger observation that we might be witnessing a change in progress regarding the indexicality of /s/-variability in German.

3. Theory

a. Gender

i. Sociolinguistics and Gender

Sociolinguistics has historically taken different approaches regarding the field of gender. We will, therefore, briefly summarize which understandings of gender and language have shaped the present discourse and how this helps us understand the same. Eckert (2012) conceptualizes Sociolinguistics as having developed in three waves: the First Wave as the classic variationist framework, the Second Wave with an ethnographic approach and the Third Wave with its focus on social indexicality. Although Eckert does not speak of a Fourth Wave, more recent developments in the field have focused on addressing ideologies and power structures through an intersectional lens. The three waves broadly map to important phases in the investigation of gender and language and will, therefore, be used as landmarks in the retelling of this field's history. These waves can and should be problematized since the above foci did not emerge chronologically and in a straightforward, clearly delineated manner. Yet, for the present study I will use the framework in order to briefly summarize the different approaches to language and gender that are needed to frame the present approach.

Early scholars such as Lakoff (1975) have argued that women and men use language fundamentally differently and have positioned women's language as weak, uncertain or polite. Such writings centered male speech as the norm, operated on binary and essentialist terms and also viewed identity as monolithic and underlying the observed variation in a straightforward way. Of course, many of these early works were conducted by feminist scholars with an aim to make women and female language practices visible. They can, therefore, be said to have laid the groundwork for later scholars seeking to further the understanding of gender differences in language.

Ethnographic approaches to the issue were pioneered by scholars like Ochs & Schieffelin (1984; Ochs 1992). As anthropologists, they changed the general perspective on how gender roles, motherhood, care, kinship and social roles are organized and acquired through language. In *Indexing Gender* (Ochs 1992) makes the case that gender and gendered roles (like mother, father, etc.) are not prior to language but instead constructed and maintained through it. These approaches are especially important because they illustrate the large role language socialization plays in acquiring gender roles and how these are consequently maintained linguistically. Many of these approaches can be argued to better fit Eckert's definition of Third Wave Sociolinguistics due to their focus in indexicality, ideology and discursive construction of gender roles. This critique of the framework becomes even more valid when we consider that Ochs & Schieffelin (2011) also speak to the importance of language ideology in the process of language socialization.

Language ideologies can be conceptualized as "representations, whether explicit or implicit, that construe the intersection of language and human beings in a social world" (Schieffelin et al. 1998: 3). Language ideologies are what the field became interested in because they illustrated that it does not so much matter what is objectively true (if this can ever be determined at all) but rather what users of a language *believe* to be true about language (c.f. Schieffelin et al. 1998; Silverstein 2023). This approach helps us understand gender in a more ethnographic way, too: we can ask ourselves how language users understand their own gender rather than examining if this is in some way warranted or, as is a central argument of this study, real. In an interview I conducted with Susan Gal in 2024 she revealed to me that language ideology was a concept which was highly influenced by female scholars in the field of Linguistic Anthropology and therefore encompassed many feminist approaches to power structures and inequality. This is also why much of linguistic ethnographic work focuses on how speakers construct, understand and maintain their gendered roles in society and what ideologies underlie these processes. Although the present study does not conduct an entirely ethnographic fieldwork, the

construction of the participants' gender is examined through such a semiotic lens here, looking at meaning making processes instead of causalities.

Throughout the Second Wave of Sociolinguistics, Linguists and Linguistic Anthropologists also developed two different schools of thought concerning language and gender: one based on dominance of men over women, the other on differences between the two groups as cultural socialization. Prominent scholars of the former approach were, for example, Fishman who famously claimed that "Women do the shitwork of conversation" (Fishman 1983: 100) and Spender (1980) who argued that language reproduced a male dominated world and therefore made women invisible. These were the early days of talking about gender inclusive language, discursive visibility and structural disadvantages being represented through language. The latter approach is represented by writings such as *You just don't understand: Women and men in conversation* (Tannen 1990) which raise the issue of differing male and female language socialization and the resulting linguistic difficulties.

The Third Wave of Sociolinguistics is characterized by a poststructuralist approach which veers away from static and monolithic conceptualizations of binary gender or sexuality and instead investigates social practice through language. Most prominently, the writings of Butler (2007, 2011) became influential for their view on gender as a practice rather than a fixed identity. This is exemplified by the highly influential paper *Doing gender* by West & Zimmerman (1987) which makes the argument that (gendered) identity is an emergent property of discourse rather than an a priori fact of human psychology. Other important work came from scholars like Cameron (1997, 2005, 2007), Eckert (2000, 2008), Eckert & McConnell-Ginet (2013), Bucholtz (1999), Hall (1997, 2014), Bucholtz & Hall (1995, 2004, 2005) or Ehrlich et al. (2014). Cameron (2005) characterizes this turn as one away from binary gender and towards diversity of identities and practices. For a more complete review see Zimman & Hall (2016).

This approach is largely what we will adopt here, too. As participants talk about their gender, it can never be assumed that they position themselves in this way as a general rule but rather in interaction with me, their interviewer, and regarding what they want to reveal about themselves, which gendered identity or goal they aim for in this particular context. Gender is, therefore, also understood as an emergent property rather than a binary, monolithic, a priori identity here.

The last turn of the field is one that encapsulates contemporary issues of power structures, inequality, ideology, embodiment and intersectionality. It is characterized by a variety of different methods and approaches. However, these works, just like the present study, all see themselves as active participants in discourses creating and perpetuating power structures. Many of the above scholars continue to expand their work under this new aspect. Pertinently for the present work, the Fourth Wave of Sociolinguistics also focusses on trans and queer perspectives not as iterations of binary ideologies but instead as communities of practice negotiating their place in society. Zimman (2013, 2017a, 2017b, 2018, 2020a, 2020b, inter alia) provides much of this perspective in his studies on transness and transmasculinity. So finally, to transition this section to the topic of Trans Linguistics, let me quote Zimman (2020b: 256) who so poignantly summarizes this movement: "The fourth wave turns attention toward language as a site where ideologies of the body, identity, and social transformation are produced and resisted."

ii. Trans Linguistics

An important field that started to develop out of the above discourse in the Sociolinguistics of the 1990ies is Queer Linguistics. Early writings such as the collection *Queerly Phrased: Language, Gender, and Sexuality* by Livia & Hall (1997) initiated the move away from investigations of queer language as sexolects, or discrete language varieties which were viewed as separate and distinct from those employed by cisgender heterosexual speakers (Calder 2024). Instead, the field aimed at combining queer theory and feminist theory to arrive at a contemporary account of how language was used in and constitutive of different gender non-conforming and sexually non-normative (i.e. queer) communities and how the larger societal discourse positioned their practices as other and reified power imbalances. In doing so, Queer Linguistics took a global approach to studying the concrete instantiations of language use among and about queer communities in a variety of linguistic and cultural contexts.

This approach continues to echo in many seminal writings on queerness in the field. They include considerations of grammatical gender in language (e.g Livia 1997, 2001), ethnographic work on gender non-conforming communities of practice (e.g. Cromwell 1995; Kulick 1998; Besnier 2003; Jackson 2004) as well as investigations of transness (e.g. Zimman 2013; Edelman & Zimman 2014) among other subjects. Most recently, a collection entitled *Linguistics Out of the Closet* (Kibbey 2024) investigates different dimensions of queer linguistics while also challenging the field and its practices. The theme of sexuality in Linguistics has resonated so widely that this has led to an *Oxford Handbook of Language and Sexuality* (Hall & Barrett 2018) that can be consulted online. In it, we find an especially relevant piece of writing for the present purpose: a call to an approach of Trans Linguistics instead of mere Queer Linguistics by Zimman (2020). In the following, we will explore how Trans Linguistics differs

from Queer Linguistics and what this approach offers to the field as well as the present study.

To begin with, let us look into the term *transgender* as opposed to *transsexual.* Zimman elaborates that, in the English-speaking world, the term *transsexual* gained recognition mainly through Harry Benjamin's (1966) book *The Transsexual Phenomenon*. This text was designed to outline criteria, which were to be applied to persons who did not identify with the sex they were assigned at birth and sought medical interventions and hormone therapy. These criteria were highly normative and required individuals to essentially embody stereotypes of straight men or women. Zimman states that it was in opposition to this heteronormative medical context that the term *transgender* emerged through activists like Virginia Prince. The term served as a non-medical alternative to the normative and binary term *transsexual* and was used by and for individuals who did not seek medical interventions but instead changed their gender roles and presentation socially.

Activists like Holly Boswell (1991), Sandy Stone (1992), and Leslie Feinberg (2006) later proposed a new vision of *transgender*: the label could be seen not simply as an alternative to *transsexual* in opposition to to the latter label's demand for surgical alterations but instead an alternative to the gender normativity that lay at the core of the issue. The consensus slowly emerged that compulsory gender normativity and predetermined paths of transition which had to inevitably end in surgery were neither what the community needed nor what should define the new label. Over time, *transgender* became an umbrella term which now largely relies on self-identification and challenging gendered norms. It encompasses different labels and strives to question binary gender and its norms.

This distinction is important because it ties into the problematic nature of much of the previous research on the topic. Zimman explains that, although Queer Linguistics includes an array of identities on the spectrums of gender and sexuality, the majority of studies on the subject of transness in this field stems from ethnographic

approaches. These often understand *transgender* as a "homogenously gender normative, essentialist, genital-focused trans community consisting only of male-assigned individuals" (2020: 5), which in turn leads to the term not being used for communities that might well be described as transgender. These anthropological studies have focused on gender non-conforming groups in the global south, like hijras in India (e.g. Hall & O'Donovan 1996; Hall 1997), travestís in Brazil (e.g. Kulick 1998), Nigerian *'yan daudu* ("men who act like women") (e.g. Gaudio 1997), Tongan *(faka)leitis* (e.g. Besnier 2003); *kathoey* in Thailand (who balance femininity and masculinity at the same time) (e.g. Jackson 2004) or Filipino *bakla* in the United States (gay men who employ feminine practices for identity construction) (e.g. Manalansan 1995). All of these works, though highly valuable, fail to view the groups they observe as trans.

This, according to Zimman, is one of the prominent differences between Queer Linguistics and Trans Linguistics: the former does not understand *transgender* as a label which can encompass a variety of gender non-conforming practices but instead adheres to outdated ideas of transness being equal to being *transsexual.* Much of this literature is saturated with problematic ideologies on the topics of authenticity and passing which reveal this misconception of transness as essentialist and normative. A pertinent example thereof comes from Livia (1997, 2001) as she discusses how grammatical gender is employed in the French autobiography of a trans woman. Therein, Livia takes gender-neutral and gender subverting uses of grammatical gender seriously but at the same time problematizes trans people's desire for bodily transformation as essentialist and contrary to the idea of gender as a spectrum.

Since much of the Queer Linguistic work on transness investigates male-assigned individuals, many such "feminist" voices have continued to raise concerns about trans identities (e.g. Lorber 1994). Other scholars have either criticized trans women for not adhering to misogynistic gender norm (Knight 1992) or echoed the problematic ideology that trans people have ultimately failed to fulfil their function

of destroying the gender binary (e.g. Kulick 2019). This cis-centered view that trans people perpetuate essentialist gender roles which circulates in Queer Linguistics only illustrates how little the field really encompasses trans people's own views and approaches. An exception to this is Jason Cromwell who, in his 1995 paper describes trans people's use of protective language. One of few transmasculine researchers at the time, he refrains from focusing on the subject of authenticity and from asking how well trans people serve cis people's narrative of gender politics and instead looks at and recognizes trans people for their use of the resources available to them.

In summary, what a Trans Linguistic approach can offer specifically is this: it does not simplify transness but instead works to make visible the unique perspectives trans researchers and speakers add to our understanding of gender. What's more, transness is not viewed here as an essentialist or medicalized complex of problems; it is a continuous and ever-evolving challenge to these exact lines of thinking. This is also why it does not impose Western colonial categories such as gay or transsexual on communities but tries to understand speakers' own self-identification in their respective culture and context. It does so by taking seriously the claim that *transgender* stands as an umbrella term for all conforming and non-conforming gendered practices – be they subversive of the gender binary or not. Most importantly, it recognizes that gender is complicated and therefore problematizes the scrutinizing cis eye on trans people's authenticity – after all, what does this really add to the endeavor of observing trans communities' language use?

Trans Linguistics may, therefore, produce new and different perspectives on and interpretations of trans experience than those that emerged early on in Queer Linguistics. This is not to say that trans scholars have not published their own works under the latter label or that trans speakers have not been given a voice in these publications. However, Trans Linguistics sees itself as a movement to actively educate on, uncover and fight transphobia in Linguistics. It is the ideology of authenticity, which is continuously raised by cis researchers, that comes under

scrutiny here. Instead of asking, as expressed in the initial Butler quote (2007), how real the experience of transness can be, Trans Linguistics takes this reality as a given and treats trans speakers not as exceptions from the norm. Trans speakers and their experiences are centered in this line of research instead of marginalized or analyzed in relation to cis people's practices as well as demands. In this way, we can begin to ask different questions that extend beyond those of authenticity and passing and move towards explorations of lived realities.

For the present study, this framework becomes particularly useful as it allows us to ask research questions without having to worry whether they are worthwhile or add to a dismantling of the gender binary. It allows us to take seriously the persons involved in the making of this study and uses categories as makeshift crutches for understanding local contexts instead of imposing them on the speakers. It is in this light that uses of categories such as *hegemonic masculinity*, *trans(masculine)* and others can be understood. What the approach adds to this research is the conviction that trans people, just like any other group of people, are neither a homogenous community nor do they need to justify their transness. It is the interest in gendered language use by transmasculine speakers that guides the analysis of this data and not the aspiration to uncover new markers of (in)authenticity. The next section consequently looks into how such gendered language use can be conceptualized by way of Gender Theory.

iii. Gender Theory

As we have already seen, Gender Theory and Queer Theory have played a significant part in the trajectory of sociolinguistic approaches to gender and sexuality. Therefore, we will summarize the essential findings and arguments that the two make, especially looking at the seminal writings by Judith Butler – *Gender Trouble* (2007) and *Bodies that Matter* (2011). The concepts that are introduced in these texts are

mirrored in this thesis' approach to gender as performative and contextually co-constructed.

To begin with, Butler (2007) posits that gender is not innate or essential to individuals. Instead, they argue that gender is constructed and reproduced through ritualized acts which we anticipate and produce in a way that may seem generic to us as we are so used to this stylized presentation. Therefore, structural differences in, for example, gendered language are not expressions of an inner truth but instead effects of discourse and power. At the extreme, they write, gender can be seen as "an hallucinatory effect of naturalized gestures" (xvi). It is with this idea that we now come to understand gender as performative: instead of *having* a gender, we *do* gender - continuously and often subconsciously. This doing never happens in isolation; it is contingent on all the parties involved, the contexts they find themselves in and their positions therein. This is why we speak of gender as being co-constructed in discourse: identities do not simply appear or disappear and are never apparent in the same way to all speakers. Instead, they are indexed and taken up (or failed to) in contexts which each interlocutor is differently aware of and positioned in.

With this turn, a lot of the previous work on gender differences in language had to be re-evaluated: if gender is not in some way anchored in the person and all speakers do is (refuse to) reproduce ritualized power structures, then what exactly is the angle through which we look at the differences between male and female speech? Most importantly, what has changed through Gender Theory in sociolinguistic terms is that we can now understand identities not as monolithic and unchangeable but instead as transient, malleable and context dependent. And naturally, the gender binary comes into scrutiny when the concept of gender as an innate characteristic is taken seriously.

In this line of reasoning, Butler also criticizes the epistemological differentiation between sex and gender. This poststructuralist approach questions whether binary biological conceptualizations of sex are not also discursively constructed and

consequently produced through normative processes, not unlike gender. The constructivist ideology imagines sex as separate from gender, and in sociophonetic terms posits certain phonetic variation as either a physical reality unrelated to and unchanged by the discourse surrounding them or as a learned behavior unrelated to physicality. But as we have seen, ideas of which voices are allowed as male or female, ideologies of how the gendered voice must be employed and overt as well as tacit norms, which circulate through discourse, influence how sex and gender are constructed and recognized in the voice. It is therefore that I take up Butler's argument that the body is not prior to sociality.

In *Bodies that Matter* (2011) Butler discusses this problem further. They carve out the argument that materiality of the body cannot be conceptualized outside of language and power structures. The way bodies are understood is not universal or pre-contextual. Instead, Butler argues that hegemonic discourses shape which bodies we deem understandable: preconceived categories determining what is intelligible as, for example, female or male influence how we understand the physical realities we encounter in human form. This argument rests on the concept of the regulatory norm, which describes how societal expectations not only guide behavior but also determine which bodies are recognized as capable of existing and therefore seen as real in the first place.

Again, this concept refers us back to the initial Butler quote – it illustrates the mechanisms by which certain bodies are robbed of their claim to realness and thereby expunged from existence. Butler's main claim here is that regulatory norms have very material consequences for bodies which are not deemed real or viable: inter, trans and non-binary bodies are excluded from claims to health, happiness and humanity. Butler here asks which bodies do in fact count as livable and which ones are devalued and ostracized. In *Gender Trouble* (2007), they give the example of their own uncle who was institutionalized for their crime of not being intelligible to gender norms of society at the time. In this sense, ritualized gender performance is

not unsolicited – deviations from the norm are seen as transgressions. Living in a body which does not count as intelligible and therefore loses its claim to realness has, ironically, real physical consequences for the individual's (and in extension the community's) life.

Therefore, societies enforce gender norms on bodies from birth - through language as well as other semiotic means. One is, in a sense, neither free to do gender in all the ways one might want to nor free from gender as a whole. In a sociophonetic sense, we think of voices as unmistakable indexes of binary gender and explicitly or implicitly associate certain articulatory gestures with this sound. Yet, when a speaker's pitch drops due to testosterone therapy, their voice does not always become immediately intelligible as male. Maybe they front their /s/, use larger vowel space or faster speech tempo, maybe their (now lower) f0 is still too exuberant in its excursions or they employ too much whispery voice. No matter the cause, a voice may become unintelligible to a binary system and lose its privilege of being understood as male or female. This, in turn, can have very real implications for the speaker: will their whole gender presentation be admissible to qualify as male/female when their voice is giving away that they are not? And what happens if they are not deemed real for this offense? These are the questions Gender Theory asks, and which inform the present analysis.

These issues, of course, also plays into Queer Theory: if bodies are not intelligible through a binary system then the idea of heteronormativity becomes redundant. The conceptualization of sexuality as two complementary genders which relate to one another in heterosexual ways as the norm comes under scrutiny here. Instead, deviations from this regulatory norm may be seen not as outliers but as examples of non-normative practices which escape the enforcement of a gender binary.

Butler's important contribution to the present study in this regard is a focus on diversity and contradiction within categories. Although they question the existence of innate gender categories to begin with (which has often been criticized by trans

activists such as Susan Stryker [1994] or Sandy Stone [1992]), they also opt for leaving open the answer to the question what constitutes each of these categories. Therefore, one might ask what constitutes the male voice and not receive so much a clear answer as an exploration of what we have so far found on how this question *can* be answered. If we take Butler seriously, we will try to understand how exactly the phonosonic nexus of a voice can become gendered, in what contexts this applies and how to make sense of the process. Therefore, the next section will focus on the concepts of indexicality and context.

iv. Context and Indexicality

As previously discussed in the Sociolinguistics & Gender section, through various phases, Sociolinguistics has carved out the idea that identification in discourse does not depend on rigid categories but rather on contingent models of identity. To understand this, we must take a short look at Peircean semiotics. In simple terms, Peirce (c.f. 1931-1958) postulates three relations a sign can have to the signed: the symbolic, iconic and indexical. A symbol is related to the meaning of the sign via convention, an icon via likeness and an index via causal or contiguous information. The index can, therefore, point to nearly anything in any given context. This means that anything can be a sign for anything provided it is taken as such by someone. Sign relations are consequently not fixed but rather culturally and contextually emergent. Understanding linguistic signs in a semiotic way, therefore, means investigating all possible contexts which the sign may draw on to gain meaning. This is often done by looking for contextualization cues which make salient the context that is evoked.

Wortham (2006) exemplifies this very aptly by showing that metapragmatic models (widely circulating types which are attributed certain characteristics) such as 'loud black girl' (which become salient in the observed settings) may depend on historical

and local gendered and racialized types. Yet, these dimensions cannot account for all of the interactional work that is done: Wortham asserts that interactional identities depend on event-level metapragmatic models as well as sociohistorical & intermediate ones. What this tells us is that understanding gendered identities is not simply the work of listing existing stereotypes and metapragmatic models in order to derive from them the meaning a particular linguistic sign carries. Instead, participants in any discourse must use relevant contextual information to infer which identity, goal or stance is being indexed.

I, therefore, follow Bucholtz & Hall (2005: 592) in asserting that "identities encompass (a) macro-level demographic categories; (b) local, ethnographically specific cultural positions; and (c) temporary and interactionally specific stances and participant roles". What is important here is to understand that signs of particular identities such as *gay* or *trans* do not directly establish these identities. Rather, configurations of signs limit which contexts become relevant and thereby point to particular metapragmatic models which can be used to interpret the signs. This means that one can neither abstract a sign from its context nor view it at the present moment only. Signs of identity gain meaning through publicly circulating models of identity, which can thicken through increased use until they, in the most extreme case, become stereotypes.

For the present analysis this means that /s/-variability is not in itself an index of gendered identity. Instead, we can view it as a type of language use associated with different metapragmatic models and stereotypes. This type then instantiates in tokens of real-life discursive use which can point to any one of these models. Through other contextual cues like voice phenomena, stylistic or linguistic choices, the listener can access the context which is being evoked and understand that the speaker is using /s/-variation to construct a gendered identity instead of, for example, street credibility. This is why uptake is so important: if we cannot understand which

identity a speaker is trying to index, social meaning is constructed entirely differently.

Eckert (2008:453) also argues that "variables do not have static meanings, but rather general meanings that become more specific in the context of styles." She points out that these meanings can be said to form an indexical field encompassing all possible contexts that the variable (or sign) may evoke. In this sense, /s/-variation in German may exhibit an indexical field which is underinvestigated at the present moment. Therefore, I hope to contribute to an understanding of transmasculine German speakers' indexing of gendered identity through their use of /s/-variability. I ask if such meaning can be accessed through /s/-variability (in concert with other features) and, if so, which dimensions of this identity are evoked in the interview context.

v. Hegemonic Masculinity

The concept of *hegemonic masculinity* was introduced and elaborated in a seminal piece of writing for the field of Gender Studies entitled *Masculinities* by R.W. Connell (1995). In it, Connell formulates the idea that a certain type of masculinity is dominant (i.e. hegemonic) in a particular society at a particular point in time. This masculinity is, however, not necessarily embodied by the majority of men but instead serves as a cultural blueprint, an ideal which subordinates other forms of masculinity as well as femininity. This, of course, means that the concept is not a static category but is instead historically malleable, adapts and changes through time and is contingent on the specifics of cultural environments. Types of masculinities that can be seen as hegemonic in one context may be subordinate or marginalized in others. The concept, therefore, is especially adept to pointing out changing ways in which patriarchal power structures are historically and locally present or absent.

With this concept, Connell argues that masculinity is not one ubiquitous category appearing in a uniform way but is hierarchically structured. In this way, hegemonic

masculinity ensures the patriarchal dominance of men over women but also the power of certain types of men over other men. Four types of masculinity are carved out in the writing: hegemonic, complicit, marginalized and subordinated masculinity. The first, as mentioned above, is a cultural ideal which subordinates the other forms of masculinity. Complicit masculinity profits from the hegemonic norms that underlie the latter masculinity (for example by laying claim on male privileges) without fully embodying the hegemonic form. Marginalized masculinity is the case of men who are excluded from the hegemonic ideal due to factors such as class, ethnicity or disability although they can still represent some aspects thereof. Subordinated masculinity, finally, has no claim whatsoever on hegemonic ideals. Subordinated men are devalued and ostracized (e.g. gay and trans men) and often labelled as unmasculine entirely.

Importantly, the mechanisms by which hegemonic masculinity functions do not necessarily have to be violent. The ideal does also not primarily function through coercion but mainly through societal acceptance: media, education, upbringing and other institutions like marriage or capitalistic structures propagate hegemonic masculinity by positing it as natural, normal or God-given. This means that, similar to what we have encountered in indexical iconicity in linguistic terms, pre-existing (power) structures are naturalized and thereby legitimized. This is important for two reasons: firstly, it counters the problematic idea that hegemonic masculinity is to be equated with violent oppression which locates patriarchal power only in overt displays thereof and secondly, it illustrates how one person may embody several aspects of these masculinities in different contexts depending on which level of acceptance of hegemonic masculinity the environment displays and allows for.

In a later text on the subject, Connell (2005) remarks that, although the cultural norms that constitute hegemonic masculinity in many western contexts have largely changed, "the policing of heterosexuality has been a major theme in discussions of hegemonic masculinity" (837). This is also where we can neatly tie the concept into

the present study. Zimman (2013) argues that transmasculine speakers, although they experience a significant drop in voice pitch due to testosterone therapy, often still exhibit learned articulatory behaviors which make them sound less normatively masculine. This often leads to them being heard as 'gay' due to transmasculine voices diverging from the hegemonic norm. Zimman here argues that hegemonic masculinity privileges a very narrow range of acceptable gender expression in linguistic behavior. Meanwhile, almost any form of non-normative masculinity is stigmatized and relegated to the category of 'gay'.

This argument also supports my earlier claim that (straight, White, cis) male articulation could be seen as rather underarticulated and these men's linguistic gender expression as rather curtailed when compared to other gendered identities. In this sense, I use the concept of hegemonic masculinity here to make this exact claim: that transmasculine individuals curtail their own linguistic expression when they do not wish to be perceived as non-hegemonically masculine. This does not mean that these speakers enact violent power structures, dominate others or fail to be authentic. It rather suggests that the default expectation of men or masculine presenting individuals to linguistically adhere to hegemonic practice is one that these speakers can adhere to or not – depending on how comfortable and safe they feel in their role as non-hegemonic men. Or, as Zimman (2013: 2) phrases it: "As individuals whose male identities may be socially precarious and whose lived experience diverges from that of *cisgender* (i.e. non-trans) men, these speakers promise to tell us much about the linguistic negotiation of hegemonic masculinity."

Naturally, the degree to which employing these strategies is a conscious process cannot be fully determined, and the question will also not be undertaken here. Taking up Butler's Gender Theory, one could argue that gender socialization is so internalized a force that it is near impossible to completely escape the influence it has on one's own linguistic strategies. I call these strategies *choices* in the sense that they reflect some wish to be heard in a certain way, but do not assume that speakers are

fully aware of or even at all acquainted with the sociophonetic phenomena they seem to embody. In this sense, the subsequent exploration of spectral properties of sibilants shows how fine-grained these linguistic choices are, which makes it highly unlikely that the whole process is of conscious nature.

b. Spectral Properties of Sibilants

Stuart-Smith (2020) summarizes how acoustic properties of sibilants can be measured in dynamic or static acoustic analysis. *Static*, in this case, refers to measurements of one single point in time or an average across the sound whereas *dynamic*, conversely, refers to measurements which "capture the time-varying nature of segmental acoustics" (Stuart-Smith 2020: 2). As the present study hopes to replicate findings based on the static measurement of center of gravity (CoG), a measurement indicating where most of a sound's energy is concentrated, I will also be investigating this measure in my analysis. While extremely useful, dynamic approaches are not yet found in enough research to allow for fruitful comparison with other results. Therefore, I will use the static measurements which are proposed by Jongman et al. (2000) for English sibilants. These three measurements are called *spectral moments* and include CoG, skewness, and kurtosis. In what follows, I will provide a detailed discussion of the production of sibilants and the method used for their acoustic analysis.

Sibilants are produced by an airstream exiting the vocal tract through the constriction between the tongue and its coronal place of articulation (i.e. alveolar ridge, teeth, palate) (Ladefoged & Johnson 2011: 175). The size of the 'front cavity,' formed between the tongue constriction and the teeth, defines the frequency range visibile on a spectrogram, which represents the energy distribution over both time and frequency (Shadle 1985; 1991). For smaller cavities, a higher frequency profile (i.e. spectrum) can be observed, while larger cavities yield lower frequencies. Forrest et al. (1988) model this spectrum using a normal distribution. The main distributional

properties of each sound are reflected in the three spectral moments mentioned above: Center of Gravity, skewness and kurtosis. These can be calculated and have been found to distinguish the sibilants' place of articulation (Jongman et al. 2000).

This means that the place of articulation and the resulting spectral properties are essential for sibilant recognition. Sibilants such as the voiceless and voiced postalveolar /ʃ/ (like in the word *shoe)* and /ʒ/ (like in the word *vision*) usually exhibit a mean spectral peak at around 2,500–3,000 Hz while the voiceless and voiced alveolar /s/ and /z/ show mean frequencies around 4,000 – 5,000 Hz. For the sound which interests us, namely /s/, the bulk of acoustic energy falls within the range of 4,000-10,000 (Shadle 1990). The three spectral moments, therefore, help to understand whether a speaker's tongue was placed at their lower teeth, between their teeth, at their alveolar ridge or their palate. Another cue for categorical perception is, of course, voicing, which distinguishes /s/ from /z/ and /ʃ/ from /ʒ/. For the present study, we are only interested in voiceless sibilants, since they are known to differ based on gender and sexual orientation as we have previously carved out in the Literature section.

The first spectral moment we investigate here, Center of Gravity (CoG), is described by Jongman et al. (2000) as reflecting the weighted average frequency of the sibilant and is measured in Hertz (Hz). Concentration of energy at higher frequencies (resulting in higher CoG values) points to a more fronted articulation whereas concentration of energy at lower frequencies results in lower CoG values and therefore indicates more retracted place of articulation. Using CoG measurements, women's means have been reported to range from 6,400–8,500 Hz while men's have been reported to range from 4,000–7,000 Hz (Avery & Liss 1996; Fuchs & Toda 2010; Flipsen et al. 1999; Nittrouer 1995; Nittrouer et al. 1989; Stuart-Smith 2007; Tjaden & Turner 1997). Variation in Center of Gravity has therefore often been linked to gender and sexuality (e.g., Stuart-Smith 2007; Munson 2007). These findings, however, also show that there is a gender ambiguous range in CoG measurements, that lies

between approximately 6,400-7,000 Hz. This range will consequently be used in the present experiment to create gender ambiguous speech.

Skewness has also been linked to gender and sexuality (e.g., Stuart-Smith 2007; Munson 2007). This measurement indicates the asymmetry of the spectral energy distribution: while there is a peak of energy on the spectrogram for sibilant acoustics, there are naturally also other frequencies that still carry energy to varying degrees. Imagining the energy that is produced as a curve with a clear center, the frequencies to the right of the same are higher frequency and the ones to the left lower frequency than this peak. This means that energy concentration on higher frequencies draws a right-leaning curve (negative skewness) and energy concentration on lower frequencies draws a left-leaning curve (positive skewness). Negative skewness is characteristic of /s/ while positive skewness is typical typical of /ʃ/. Skewness is, therefore, also a robust measurement to investigate gender and sexuality. To exemplify this, let us consider spectral slices (extracted from Praat) of the lowest (2,366 Hz) and highest (10,755 Hz) CoG /s/-tokens of the present study:

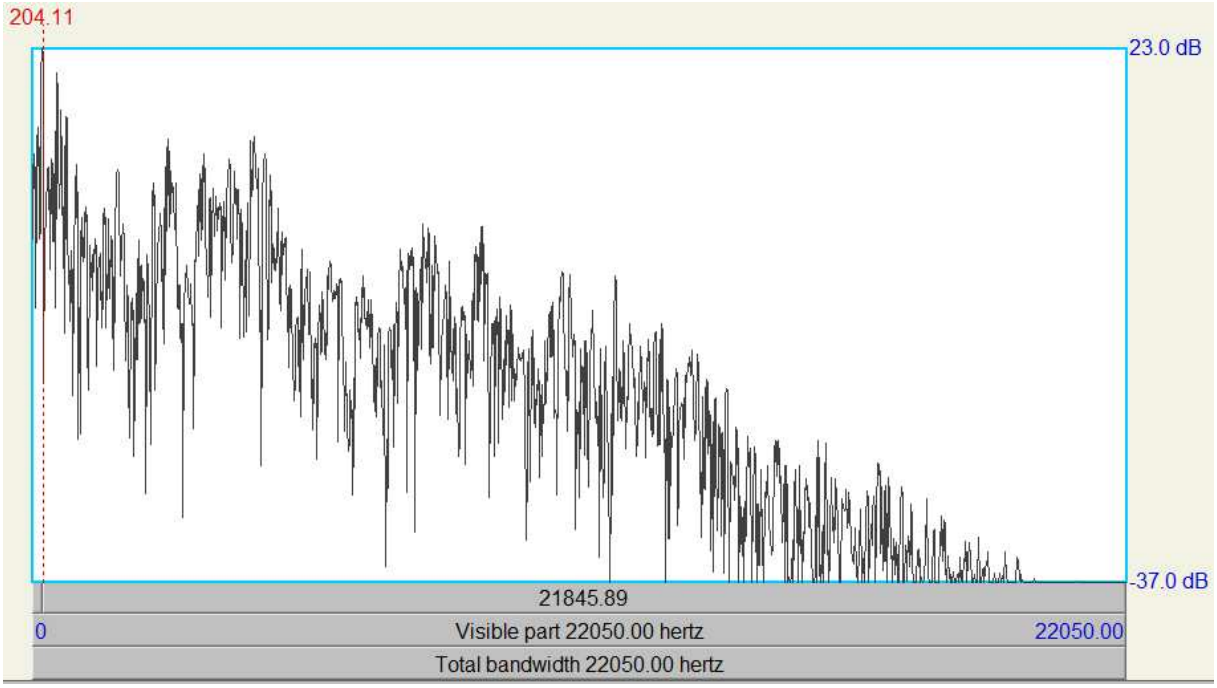

Figure 1: spectral representation of an /s/-token with a CoG of 2,366 Hz

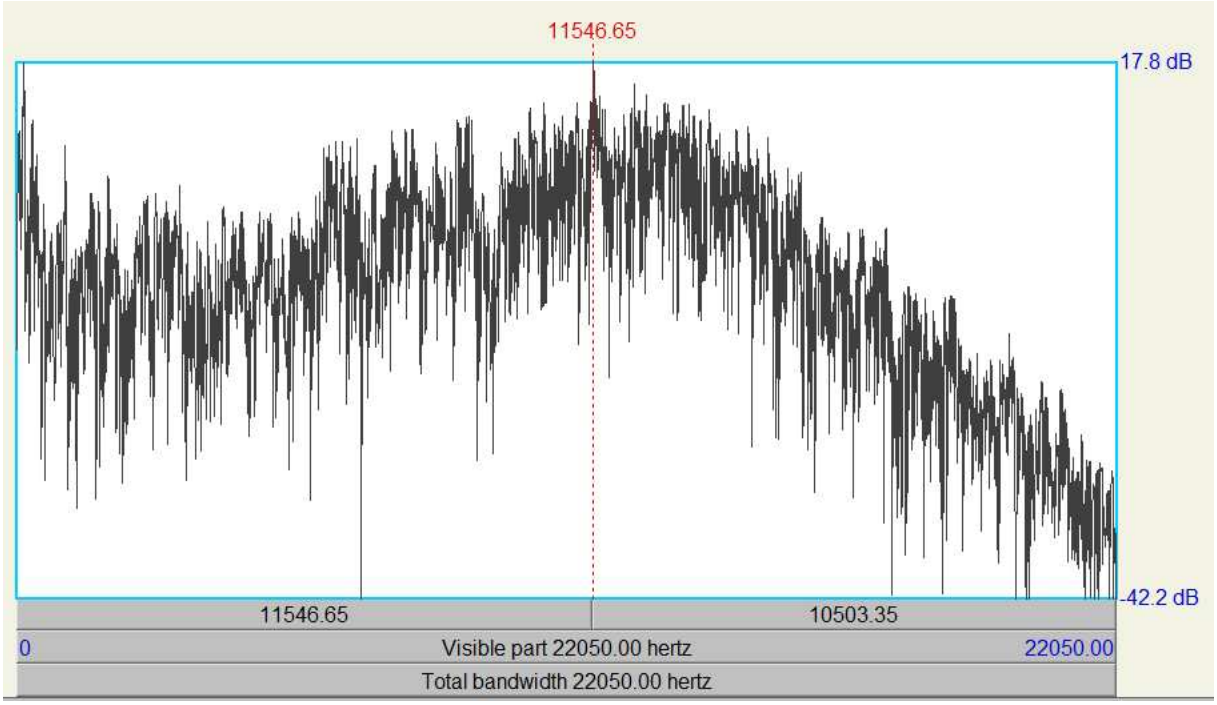

Figure 2: spectral representation of an /s/-token with a CoG of 10,755 Hz

Figure 1 clearly shows the skewness of the sibilant leaning towards the left with a peak (at approximately 204 Hz) and energy concentration in the lower frequency range. This results in positive skewness value (1.33) while the sibilant in Figure 2 shows a skewness value of negative 1.24 and is skewed more to the right. This sibilant has two peaks, one at a low frequency and one at about 11,500 Hz. However, considering the energy distribution of this sound, the negative skewness value makes sense as there seems to be a lot of energy in the higher frequencies.

Kurtosis assesses the 'peakedness' or sharpness of the spectral distribution. As we have seen, spectral energy is not only concentrated on one peak frequency but forms a curve that can be flat or narrow, depending on how energy is distributed across frequencies. If the curve is flat, this indicates that the energy is spread out among a larger frequency range whereas a steeper curve suggests a stronger energy concentration around a narrow frequency band. Higher kurtosis values are indicative of sharper peaks, as is characteristic for /s/, while lower kurtosis values suggest a flatter curve and, therefore, a broader spread of energy across frequencies, as is typical in /ʃ/. In the present dataset, for example, the /s/-token in Figure 1 has a

kurtosis value of 2.18 while the one in Figure 2 shows a value of 1.42. This reflects the stronger concentration of energy around the 204 Hz peak in Figure 1. While this is not necessarily a large difference, other tokens in the datasets show extreme kurtosis values of up to 2.94 and negative 1.45.

Naturally, one could simply measure the peaks of energy as these have also been reported to vary between genders. English-speaking women have been reported to show peak frequencies in the range of approximately 5,500–9,000 Hz whereas English-speaking men exhibit a range of approximately 4,700–8,000 Hz (Fuchs and Toda 2010; Schwartz 1968; Stuart-Smith 2007; Yeni-Komishian & Soli 1981). It is, however, more informative to have a cumulative measurement which can encompass more information on the complex nature of the sibilant. However, Zimman (2017b: 1007) argues that one must not necessarily look at CoG, skewness and kurtosis in concert to find meaningful correlations to gender and sexuality. Rather, he suggests that CoG "provides enough complexity on its own for a full exploration of sex, gender, and sexuality as practiced within these communities". This is also why the present study uses CoG as its sole measurement and does not take into account skewness or kurtosis.

The methods used in the below experiment rely largely on those employed in other, similar, studies on gendered differences in sibilants (e.g. Munson 2007, Stuart-Smith 2007, Pharao et al. 2014, Zimman 2017b). Sibilants are hence extracted through a Praat Script (DiCanio 2013) which creates a weighted mean (CoG). Other measurements could surely be incorporated into such work, especially as Stuart-Smith (2020) reports differences in results for static and dynamic measurements. For now, these more complex and statistically elaborate and time-consuming endeavors shall be left for future investigations. As this study hopes to replicate findings by Zimman (2017b) for German, it suffices to investigate and map variation in CoG just like he does.

4. Methods

Based on the introduced literature and theory, we will now turn to the methods of data collection and analysis. This section gives a general overview of the aims of the present empirical research as well as a short introduction to the methods employed before providing more detail on data collection and processing for each of the three experiments.

In this two-pronged study design, the aim was to determine if transmasculine German speakers employ variation in articulation of the phoneme /s/ as an effective stylistic resource to discursively co-construct their gender. Special emphasis was put on exploring how effective this resource might be should it be employed. That is precisely why the present study employs mixed methods to determine if /s/-variation is produced by speakers as well as perceived by listeners. A qualitative approach is first needed to understand the complexity of speakers' relationship to (their own) gender to make sense of the pattern of /s/-articulation which then emerges from acoustic analysis. For the reasons previously discussed in the theory section, the latter quantitative acoustic analysis can only be meaningfully interpreted in juxtaposition with qualitative data that grants an insight into the possible objectives behind the speakers' stylistic choices. Furthermore, assessing influences on listener perception relies on another quantitative approach that can best be interpreted in concordance with both quantitative but also some qualitative aspects of participant data.

What emerges then is the following: a mixed-methods study making use of qualitative interviews, the insights from which are consulted to interpret quantitative acoustic data procured from the same six speakers in a map task. This first part of the study aims to elucidate the patterns of /s/-articulation of transmasculine German speakers. These patterns can then be compared to such results for English (Zimman 2017a) to see if a similar picture emerges. For the second part, a perception study is carried out in order to determine if variation in /s/ is an effective indexical resource for German. This

part of the study strives to determine whether German speakers assign indexical meaning regarding gender and sexual orientation to variation in articulations of /s/.

The first two experiments (constituting the *production* part of this thesis) aim at better understanding individuals' relationships to their own gender in order to determine if the complexity of this relationship was mirrored in the place of articulation for /s/. Acoustically, this place of articulation can best be measured through the Center of Gravity (CoG) which was the method employed here. In the interest of doing justice to the qualitative as well as quantitative nature of this research interest, it was necessary to elicit authentic as well as comparable language data. To achieve this, semi-structured interviews were conducted, recorded and transcribed with all six participants. Based on these interviews, a chart was devised that compiled all relevant information pertaining to gender and sexuality which made it possible to keep oversight of the large amount of qualitative data. The same individuals were then also asked to participate in a map task which was specifically designed to elicit target words with /s/-tokens suitable for acoustic analysis.

The CoG of these tokens, which were obtained with the help of a Praat script (DiCanio 2013), were then plotted in RStudio. The intraspeaker as well as interspeaker variability was best visualized as CoG-range in a boxplot and the total range of speaker variability as CoG-means in a chart. Independent quantitative variables like time on testosterone were tested for significant correlations to CoG before the aforementioned qualitative chart was consulted along with specific participant statements (on the subject of gender). These qualitative variables were used to draw additional post-hoc conclusions on the correlations between speakers' relationships to (their own) gender and their articulations of /s/ manifest in the two plots. This first part of the study that concerns itself with production of /s/-variability was inspired by the aforementioned study by Zimman (2017a). It was consequently designed in such a similar way to the latter study that it was possible to compare results, albeit taking into account the necessarily smaller scale of this thesis.

For the second part of the study, the goal was to determine if listeners would assign different indexical values to variations in the articulation of /s/. For this purpose, a 4-second segment by one of the interviewees was selected based on spectral analysis of the participant's dataset (obtained in the previous step) as well as other linguistic criteria like prosody, mean f0 and phrase final intonation. This segment was then used in a perception survey (conducted online with participants recruited over social media and personal contacts), which aimed at finding differences in listener perception based on /s/-variation. Three manipulations of said segment, each only encompassing one frequency of /s/ per manipulation, were created by splicing three /s/-tokens out of different utterances by the same speaker. These tokens were chosen to replace the original /s/-tokens as they represented the lowest, highest and most ambiguous frequencies in the speaker's dataset. This resulted in four different versions of the same segment which could then be used in testing. An online survey was employed in order to find if listeners would rate the four samples as more or less "masculine" or "homosexual". This part of the study design was inspired by the previously discussed perception study by Boyd et al. (2021) which also investigated the indexicality of /s/-variation. A few differences necessarily emerged from the restrictions posed on an MA thesis as well as contextual vagaries but comparing results was still worthwhile.

In the following, I introduce each of the three experiments in detail: the interview, the map task and the perception study. Each section first introduces the participants and how they were recruited, subsequently moves on to the implemented methods of data collection and rationale behind them before discussing the resulting methods of data analysis that were employed. The three different experiments make use of quantitative as well as qualitative methods in data collection as well as analysis in order to test the research question "*do transmasculine German-speakers employ /s/-variation as an effective phonetic resource for discursive co-construction of gender?"*.

The underlying hypothesis of the first part of this study is that a varied pattern of /s/-articulation indeed emerges in transmasculine German-speakers. Hence, this section aims to paint a complex picture of quantitative as well as qualitative factors that

possibly contribute to this stylistic variety. The second part hypothesizes that uptake of higher frequency /s/ as indexical of non-hegemonic masculinity (e.g. gayness and/or femininity) is observable and likewise aims to explore different avenues of explanation.

a. Interviews

i. Participants

To begin with, participants were recruited through a flyer that was posted at different parts of the University of Vienna, public spaces like bars, bus stops etc., but also online on social media. The flyer was designed in German and deliberately worded as looking for *trans-masculine individuals* ("transmaskuline Personen") in order to avoid deterring any assigned female at birth (AFAB) trans-people who did not identify as trans *men* specifically.

As a result of this, six individuals contacted me via Instagram or email and received the participation criteria together with general information about the experiment in a PDF-document. This course of action was taken to ensure that prospects could take enough time to decide whether they were certain of their wish to participate before they were finally invited to an interview. Furthermore, it seemed obvious that my own (perceived) sexual orientation and gender presentation would influence how participants would engage with me and the interview process, especially in how safe they would feel opening up about sensitive topics and, consequently, how likely they were to do so. Therefore, the document prospects received identified the researcher as a queer cis woman and assured the reader that the former was acutely aware of the highly personal and discrete nature of such an interview. A copy of the flyer and PDF-document can be found in the annex to this thesis.

In order to qualify for the study, participants had to identify as transmasculine, speak German as (one of) their first language(s) and either speak an Austrian Dialect or Austrian Standard German in order to not invite too many confounding factors. Yet,

as two of the six individuals were eager to participate but spoke Bavarian or Southern German respectively, the decision was made to broaden the linguistic scope a little. As Bavarian is part of the same Austro-Bavarian dialect group (Österreichische Akademie der Wissenschaften 1983), including this speaker did not seem like a large risk and was therefore allowed. As for the Southern-German speaker, after consulting with colleagues, I concluded that there was no significant difference in articulations of voiceless sibilants between the two regional varieties and decided to invite this speaker to interview as well. Based on these considerations, the inclusion of the *dialect* and *nationality* categories into the datasets was deemed necessary. Should stark inter-speaker differences emerge, these factors were codified in any case and could be taken into account in data analysis.

The interview partners were 18-25 years old, held different occupations and were born into different familial constellations. Four of the six were university students or had finished studying, one was a trained employee and one was still in school. Only one of them was born and raised in Vienna, the rest came from Lower Austria, Upper Austria, Southwestern-Germany and Bavaria. Ideally, a more age-diverse group would have emerged in my search for participants in order to assess if there were generational differences between speakers, but time constraints precluded me from broadening my search. Luckily, the time on testosterone for the recruited persons ranged from 12 to 72 months and age at beginning of t-therapy from 17 to 24. This variety allows for at least a few interesting differentiations and observations in testing.

All participants identified as transmasculine and five used he/him pronouns in both English and German at the time of the interview. One person used *er/ihn (he/him)* pronouns in German and *they/them* pronouns in English, one stated they were also ok with *they/them* in English. Nonetheless, the six interview partners were highly varied in their self-descriptions regarding gender expression and style as well as their self-described sexual orientation.

Four participants spoke Austrian Standard German and/or an Austrian dialect, two spoke German dialects and/or German Standard. One participant grew up in Lower

Austria with a parent who spoke Carinthian German dialect and stated that he could comfortably understand the latter although he could not speak it. None of them were L1 speakers of another language while all of them said they spoke English very well and five of them stated that they spoke at least one more language to some extent.

As some of the participants' attention was drawn to the study through personal contacts, two of them shared acquaintances/friends with me. This fact was also discussed in the interviews and surely has to be taken into account in interpretation of elicited data. More ethical concerns regarding this topic are discussed in the Consent & ethics section.

ii. Data Collection

1. Consent and Ethics

In the interviews, I placed much emphasis on making the participants feel that they were in a safe space and could pass up a topic or stop the interview altogether without any negative consequences. This was also emphasized in the process of going through the consent form together and reiterated before starting the interview. The consent form included information about the interview process as well as the resulting data and was verbally paraphrased to participants before they got as much time as they needed to read through the form on their own and sign it. Participants were also given some time to ask any questions that might arise before starting the interview and, after concluding the interview, were asked if they had felt comfortable with the questions to which all of them reported that they had.

The atmosphere of trust that I aspired to create in this way was born out of two considerations: firstly, my personal and scientific ethics around the topic of queerness and, secondly, a review of some of the existing literature on the practice of the ethnographic interview (De Fina 2019; Spradley 2016). For one, personal experience

informed the assumption that not all participants would show the same level of comfort when revealing details about their gender identity and sexuality to a stranger. It seemed necessary to take into account that high levels of anxiety around this encounter might be battling the wish to participate in the study simultaneously. It was, therefore, of great importance to me to not abuse the power imbalance at hand (Mishler 1991) but instead build the interview process on transparency and consent. This is why the aforementioned precautions and measures seemed essential.

Furthermore, the much-discussed problems of researcher positionality (Creswell 2024) and the relationship of the interviewer with the informants (Copeland 2019) made me acutely aware of my own position as a researcher engaging with an at-risk population. Hence, my objective was to broadly follow some of Spradley's (2016) suggestions and provide an environment for our conversations that would (1) put participants' minds at ease as to how their personal information would be handled, (2) clearly delimit the roles and responsibilities of each interlocutor and (3) create the kind of rapport that is necessary to engage in discursive co-construction of identities.

Per the last point, De Fina discusses that "traditional interviewing methods imply a de-humanisation of interviewees through questions that are standardised and divorced from local contexts and experiences" (2019: 161). In order to avoid treating my interview partners as such mere sources of information, I made a strong point of not scripting my questions, contextually revealing details of my own life as well as referring to commonalities and shared acquaintances. Naturally, this might raise yet another question of ethical behavior similar to the one Howell (2004) discusses in regard to the relationship context certain information is revealed when talking about gender and sexual abuse. Nonetheless, I felt that, in contrast to Howell's (2004) research, the clear delineation of the interview-interviewee relationship in the present study was unambiguous enough not to be mistaken for a casual chat among acquaintances.

What's more, it is possible that my own position in the queer community was a prerequisite for the trust that my interview partners placed in me. This is why

revealing personal and shared experiences was necessary in order to build confidence. Additionally, the fact that the channels through which they had become aware of the study were trusted ones (shared friends and institutions) may have been a mitigating factor in any reservations participants may have had. In conclusion, it is not entirely possible to tease apart the personal and the scientific when conducting ethnographic interviews. Creating a safe and confidential atmosphere ultimately took priority over constantly monitoring myself in regard to coercive speech or maintaining relationship boundaries implying objectivity. Yet, as all participants reported being comfortable with the conversation after the fact, this can be taken as an indication of consent before as well as after the interview and should allay such ethical concerns on this matter. It is, however, necessary to view the obtained information in light of this researcher's positionality.

2. Scope

The relation between interviewee and interviewer, of course, raises the issue of performative speech. As discussed in the theory section, a variationist, Labovian search for the most authentic "vernacular" speech is not the goal of the investigation at hand. The local, contextual co-construction of gender which takes place in these interviews should not be regarded as generalizable or representative of an entire speech community. Instead, one should think in Butler's terms of gender performance (cf. Butler 2007), which relies on exactly such local and contextual factors. An expansive critical discussion of the interview process and its performative nature can be found in Briggs (1986, 2007). The way individuals linguistically and multimodally present themselves to the researcher is contingent on the perceived identity of the researcher herself as well as her self-presentation towards the interview partners. How, then, to treat the data that is obtained in such a situation?

Zimman (2017b) discusses that the self-consciousness which is possibly evoked in a transgender person by a conversation on the specific subject of gender and sexuality with a cisgender speaker could be viewed as "a potential opportunity for participants

to produce a kind of self-conscious gendered performance" (2017b: 348). It can, therefore, be interesting to investigate gendered performance as pertaining to a specific interlocutor in a specific context and it is from this angle that the theory and methods explicitly allow us to view the contents of said performance. This does not suggest that speakers consciously manipulate their speech but implies a situational awareness that has to be taken into consideration when interpreting the data.

3. Location and Interview Content

The interviews were recorded in July/August of 2024 with a Zoom H2n Handy Recorder on a tripod and transcribed after the fact using the AI-powered transcript creator Transkriptor[1]. They all took place in a staff room of the Linguistics department of the University of Vienna with only the researcher and the interview partner present. As this was a semi-structured interview, the goal was to engage with the interview partner in a way that didn't seem decontextualized or like a series of strict questions. Therefore, my research notes at the beginning of each interview mostly encompassed themes I wanted to talk about and the occasional distinct question. The interview questions themselves mostly emerged contextually and were chosen in an order that would allow the interviewer to slowly build rapport with the interviewee and to not ask them to reveal too much sensitive information about themselves right away. The topics ranged from the initial, less sensitive ones like languages and dialects over role models and family to gender and transitioning.

Initially, I asked my interview partners to reiterate their name and age, then asked them which pronouns they used, where they were from and if applicable how long they had been living in Vienna and why they had moved there. In some cases, a

[1] https://app.transkriptor.com

conversation unfolded over the topic of pronouns and further questions presented themselves, making the conversation meander and come back to the other topics later. It is, therefore, not so straightforward to list all of the interview questions here. Yet, what can definitely be said is that most questions arose organically from conversation and aimed at helping me understand my interlocutor as well as their relationship to (their own) gender more deeply.

iii. Data Analysis

Of course, the obtained transcripts had to be thoroughly reviewed as AI-generated transcription is still quite faulty. After this review, I was already well acquainted with the intricacies of each interview but still decided to additionally file and edit them in MAXQDA[2] for better overview and easier access to thematic segments.

Then, the interviews along with their transcripts and interview notes were reviewed and additional notes were made on participants' self-descriptions in regard to the relevant gender categories. These were based off Zimman's (2017a) method in which he initially distinguishes between *sex, gender, gender expression* and *sexual orientation* and also notes *age, place of origin, class* and *race/ethnicity*. For my own analysis, these categories were largely adopted with the exception of two which had to be adapted.

Class was replaced by the category of *profession/occupation* due to two factors; firstly, Zimman's analysis, though taking *class* into account, did not find a very large direct impact of this factor on articulation. If anything, social stratification is to be seen as underlying one's identity in a much more complex way than could be teased apart in a study that does not explicitly concern itself with this issue. Secondly, it seemed unwise to ask this question in a face-to-face interview lasting between 30 and 55 minutes when a big concern was creating rapport and confidence in the researcher. *Class* seemed too risky a subject to easily incorporate it into a short conversation while the reward seemed too low.

[2] https://www.maxqda.com

Furthermore, the category of *race/ethnicity* did not seem relevant due to all participants being perceived as White and reporting having been raised in Austria/Germany. So instead, two more categories, namely those of *nationality* and *dialect(s)* were added. As mentioned above, these became pertinent as two of the participants who were included were German and spoke Southern German dialects.

In a second, more detail-oriented step, intricacies of participants' gender and gender expression were added. These included detailed additions to the category of *gender presentation* like more direct quotes and visual appearance in order to create a fuller picture. This second step also resulted in the additional categories of *role models, pronouns, time on testosterone* and *age at beginning of t-therapy*. The latter two were added regarding the aforementioned low age distribution while *pronouns* and *role models* aimed at adding to the range of possible avenues in qualitative analysis. All of this information was then compiled into an expansive chart with the aim of providing a backdrop to the quantitative data which would be acquired in the next step.

On the one hand, linguistic data was the main source of information in this process. Almost all of what was compiled in this qualitative chart were self-descriptions of the respective participants. On the other, looking at Zimman's (2017a) chart which encompassed more detailed descriptions of gender presentation, it seemed obvious that this level of fine-grained self-description was acquired over the course of a long ethnographic study. As my own research project was limited by the time constraints of an MA thesis, such deep insights into participants' self-construction were not possible in the given timeframe. Therefore, some of the participants' appearance-based stylistic choices within which their linguistic stylistic choices were situated were also added (in italics) to the dataset. This was done in respect to participants' own mentioning or verbal highlighting of said features. No features or stylistic choices which were not explicitly discussed are mentioned in order to avoid researcher bias, discriminatory language and most importantly to protect the participants' privacy.

A more detailed discussion of how these features add to the understanding of the individual's gender performance and consequently their articulation patterns can be found in the results section.

b. Map-Task

i. Participants

The same six participants were asked to come in a second time for the map task a few weeks later with the exception of one participant (P6) whose interview was conducted on the same day as the map task. As the map task required two people to converse with each other, the two pertaining participants were introduced to each other at arrival. Incidentally, P1 and P5 knew each other which became apparent right away. Of course, concerns about a conversational convergence effect (Pardo 2006) could be brought up regarding these two individuals. Yet, as with the general concern about performative speech, this fact simply must be considered in analysis. Of course, it is possible that these two individuals approximate each other's /s/-variation due to their more intimate relationship than that between other participants. However, such considerations only underscore that /s/-variation is employed stylistically in a specific context and that it is not a deterministic view on the matter that will help us understand the phenomenon.

ii. Data Collection

The map task was an elicitation task designed to encompass a large number of target words with suitable /s/-tokens which would not show any coarticulation effects. Initially, the following targets were included:

- Flussufer-, Messe-, Hauptstraße ([ˈflʊsʔuːfɐ], [ˈmɛsə], & [ˈhaʊ̯ptʃtraːsə])
- Alt- & Parkgasse ([ˈalt] & [ˈparkgasə])
- Messegelände ([ˈmɛsəgə,lɛndə])
- Fluss ([ˈflʊs])
- Schloss(allee) ([ˈʃlɔsʔale:])
- Fußgängerzone ([ˈfuːsgɛŋɐ,t͡soːnə])
- S-Bahn ([ˈɛsbaːn])[3]

As the task was designed not to attract too much attention to these targets in order to avoid speaker bias, a few targets did not contain any or as many /s/-tokens as others (e.g Neuer & alter (Fluss), Brücke, (Messe-) Gelände, Alt- & Parkgasse, Hauptstraße). The design was inspired by Pharao et al.'s (2014) map task that was also created with the aim of eliciting /s/-tokens. The room and recorder were identical to the interview setting.

Participants were introduced and sat across from each other before a divider was placed between them and they each received a printed copy of the map task. Instructions were then given to explain to the other how to get from point A to point B on the map. Both participants had received an identical map with the exception of point A and B and the resulting route between these points, which was marked in pink for one participant and yellow for the other. The task was to explain to the other the route that was marked on one's own paper while the other had to follow along with a marker on their own map, eventually creating two maps with two identical routes marked on them. When they had understood the assignment, the microphone was placed in front of the person whose turn it was to describe the route on their map to the other. A copy of the blank map can be found in the annex to this thesis.

In preparation for subsequent analysis of the map tasks data, transcripts of the task (which were also created using Transkriptor[1] and reviewed) were force-aligned with

[3] German orthography represents the phoneme /s/ as "s", "ss" or "ß" depending on environment.

the corresponding audio files in WebMaus[4]. Thereby a TextGrid file that matches sound structure to orthographic form (SAMPA[5]) was created. This text grid was then annotated by hand by adding another tier for suitable /s/-tokens which would later be run through a script. The length of each token had to be manually adjusted based on visual impressions of the waveform as WebMaus tends to make minor misjudgments about where exactly a sibilant starts and ends.

The selected /s/-tokens were taken from utterances that satisfied the following constraints:

- VsV → Straße, Gasse, Messe, Flussufer, dassi, mussi
- Vs# → was#, bis#, dass#, es#, Fluss#
- Vsf → Kreisverkehr, Ausfahrt
- #sV → #so, #sorry, #Süden, #südlich
- Vsb → S-Bahn

These restraints (which excluded word *like links, rechts, östlich, westlich*) were chosen in regard to possible coarticulation effects as well as regional variety in articulations of /st/ and /sp/ (Ammon et al. 2011). However, a few exceptions were allowed: *passt* (since no variety of German includes [ˈpaʃt]) and *ist* when articulated as /*is*#/ which often satisfied the second constraint.

When all the applicable /s/-tokens' beginning and ending were marked in the TextGrid file (and all the boundaries moved to the nearest zero crossing as is standard procedure), a Praat script (DiCanio 2013) was run over the file in order to extract the spectral properties of each token. A more detailed discussion of this process can be found in the next section.

[4] https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/WebMAUSBasic

[5] Speech Assessment Methods Phonetic Alphabet, the standard orthographic representation in Praat

iii. Analysis

The analysis of the map task data aimed at finding and visualizing the participants' patterns of /s/-articulation (mean CoG and CoG-range) in order to determine if any of the quantitative variables in the chart (namely dialect, nationality, time on testosterone, age at beginning of t-therapy) had statistically significant impact on this pattern as well as to provide the quantitative counterpart to the qualitative chart devised in the previous experiment. As previously discussed, the most useful heuristic for place of articulation is CoG which is what this experiment focused on. As previously discussed, in a broader scope skewness and kurtosis could also be taken into account for this analysis but as the present study is limited by the constraints of an MA thesis and CoG is a robust measure, this analysis will solely rely on the latter. CoG values were first obtained using a Praat script by DiCanio (2013) which extracted every /s/-token from the pertaining tier in the TextGrid file and measured its spectral properties (CoG, skewedness, kurtosis, standard deviation) as well as intensity (in Decibel) and duration (in milliseconds). This resulted in six separate CSV files encompassing all of these values for the respective participants. These were then compiled into one big dataset (distinguishing the participants by their initials) which provided the basis for analysis.

As this file did not include any orthographic representation of the word or utterance the token was extracted from, this information had to be added by hand. Other categories were likewise added to the file manually in order to create the independent variables necessary for significance-testing and mapping in RStudio: *participant alias* (P1-P6), *dialect group* (BA for *Bairisch* which encompasses Austro-Bavarian dialects and SD for *Süddeutsch* which was used for the single Southern German speaker), *nationality* (AT for Austrian and DE for German), *age at beginning of t-therapy* (in years) and *time on testosterone* (in months).

Dialect group, nationality, age at beginning of t-therapy and *time on testosterone* were added to the dataset as control variables. None of them were expected to have a significant effect on CoG frequency. If any of them did, however, the articulation pattern that was

expected to emerge could be connected to some aspect of these variables. This insight, in turn, could point to a multicomplex array of physiological and social factors underpinning /s/-variability just as Zimman (2017a) suggests. Stylistic use of /s/-variability draws, in his view, on a combination of physiological and socially acquired resources. This is why statistical testing could potentially clarify which factors might influence this process.

The following quantitative analysis concerns itself with these statistical significance tests and the process of plotting the CoG patterns. The subsequent qualitative analysis elaborates how the chart devised from the interview data was used to complement the articulation patterns that emerged from the quantitative analysis.

1. Quantitative Analysis

The complete dataset containing all the above-mentioned variables and CoG values for all six participants was used as the basis for analysis in RStudio. The means, token numbers, minimum and maximum values and resulting range per participant as well as the range of means were compiled into a table for later analysis. Then, normal distribution of CoG values had to be determined in order to conduct statistical tests. This was done in RStudio through a Shapiro-Wilk-test which showed that normal distribution was given (p-value = 0.2442). The distribution is plotted in Figure 3.

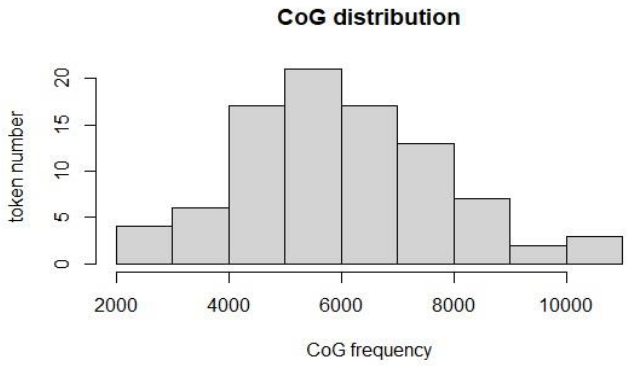

Figure 3: CoG frequency distribution by number of tokens

Then, variables for statistical testing were defined: *dialect group, nationality, age at beginning of t-therapy* and *time on testosterone,* as independent variables and *CoG* as the dependent variable. The hypotheses for all of them were that they would show no significant effect on *CoG.* After testing the nominal variables for variance with a Levene test, *dialect group* (p-value = 0.3191) and *nationality* (p-value = 0.4232) were tested for significance in a t-test. For the metric variables of *age at beginning of t-therapy* and *time on testosterone,* a Pearson-correlation-analysis was carried out (confidence interval 95% for all tests).

In order to visualize the articulation patterns, a boxplot was created using RStudio. For this purpose, the y-axis was defined as *CoG (Hz)* and the x-axis as *participant.* This resulted in Figure 4.

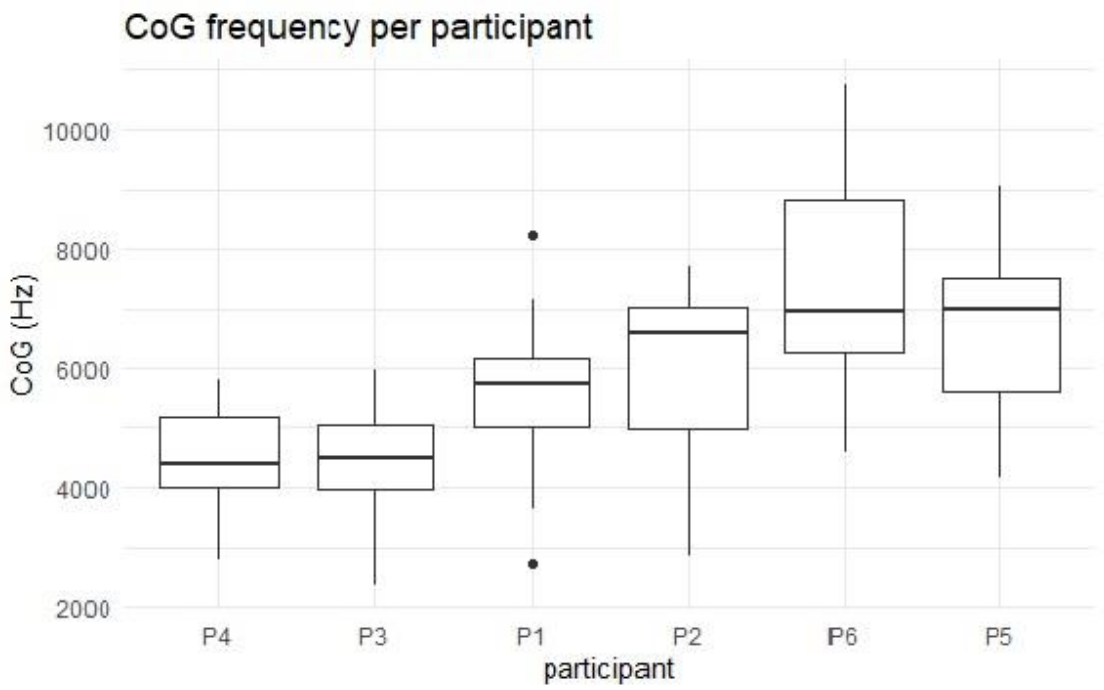

Figure 4: means & distribution of CoG per participant (in increasing order of means)

2. Qualitative Analysis

In order to find possible correlations between qualitative variables (i.e. self-descriptions) and CoG, I compared the interview chart to the boxplot in Figure 4. A trend became apparent as can be seen in Figure 4: two participants showed more normatively "masculine" CoG (P4 & P3), two were positioned in a more gender ambiguous range (P1 & P2) while two showed relatively high, more normatively "feminine" values (P6 & P5).

These results and their implications will be discussed in more detail in the Results section but need to be mentioned here in order to illuminate the process of qualitative analysis. The aim of comparing Table 1 to the qualitative chart was to find which aspects of the individual's relation to their gender and gender as a concept might play a role in place of articulation of /s/. Thus, I revisited each interview, this time with the CoG-distribution in Figure 4 in mind. In light of this pattern, participants' words gained more and more performative character until I finally formed a working hypothesis on which aspects of their gender identity might be most conducive to such intra- as well as interspeaker variety. The concept of hegemonic masculinity (Connell & Messerschmidt 2005) was combined with the differentiations Zimman (2017a) makes between aspects of gender to devise a chart encompassing the aspects within which participants performed (non-) hegemonic masculinity.

c. Perception Study

i. Participants

Participants for the perception study were acquired by sharing the corresponding link to the online survey with friends, family and on social media. The whole of the survey was conducted in German and almost all questions (with the exception of one) were mandatory. Demographic questions aimed at potential factors that might correlate

with the way individuals interpret the indexicality of and interact with the speech samples.

These factors were operationalized in the following way:

- Age (categorical, options in table below)
- Self-identified gender (options: male, female, non-binary, inter, genderfluid, other:[open ended])
- Trans- vs. cis-identity ("Was this gender assigned to you at birth?")
- Occupation (options in table below)
- Highest education (options in table below)
- Region of origin (open ended, aiming at spoken varieties/languages)
- Urbanity vs. rurality (options in table below)
- Evaluation of perceived regional origin of the speaker's variety
- L1 ("Is German (one of) your first language(s)?" – Yes/No)

The total number of participants was 37, 26 of which were female, 10 were male and one person was non-binary. 36 of these were cisgender and one was trans. The age distribution as well as the participants' occupations can be seen in the tables below.

age group	**n**
18-24	7
25-30	7
31-35	15
36-40	1
41-45	4
51-55	1
61-65	2

Table 1: age distribution ranked by age group

occupation	**n**
employed	19
student	11
unemployed	2
retired	2
other:self-employed	2
other: mini-job	1

Table 2: participant occupation ranked by number

Two of the age-group options had no answers and were thus omitted from the chart: *46-50* and *66+*. This also applies to two of the occupation options: *in school* and *apprentice*.

In Table 3, "Matura/Abitur" stand for the Austrian & Swiss/German degrees awarded when finishing high school after 12/13 years of education. *AHS*/*BHS* stand for *Allgemeinbildende Höhere Schule*/*Berufsbildende Höhere Schule* (general education school/ professional education school) which are the two most common forms of high school in Austria. The option *higher education* includes all academic degrees awarded in the DACH region (Bachelor, Master, Magister, Diplom, Doktor, PhD). More categories (with fewer years of schooling) were available but got no answers. The term *Fachschule* describes an older form of professional education that has recently been replaced by the *Fachhochschule* in Austria with this newer form granting a bachelor's degree while the older awarded Diplomas.

parish size (inhabitants)	**n**
up to 5.000	6
up to 10.000	0
up to 20.000	3
up to 100.000	5
more than 100.000	23

Table 4: parish size of current main residence ranked by size

region of origin	**n**
Vorarlberg	12
Vienna	7
Baden-Württemberg	4
Bavaria	3
Lower Austria	3
Styria	2
Germany	2
Hessen	1
Carinthia	1
Budapest	1
Bosnia & Herzegowina	1

Table 5: region of origin ranked by number

highest education	n
higher education	25
Matura/Abitur	
AHS	9
Matura/Abitur	
BHS	2
other: Fachschule	1

Table 3: highest education ranked by number

Region of origin was selected as a factor in descriptive statistics (1) as a possible confounding factor in how speaker variety was evaluated and (2) in order to have as much information as possible for post-hoc statistical testing. The question was posed as "Which region(s) did you mainly grow up in?" in order to elicit the most dominant variety/varieties/language(s). *Parish size* was chosen to constitute the factor of rurality/urbanity which could then be tested for in data analysis.

Of the 37 participants only 3 stated that they were not L1 speakers of German but were included in the dataset anyway as an exclusion/further analysis based on L1 was still possible at a later time. Of these three, two were Hungarian L1 speakers and one was a BKMS (Bosnian, Croatian [German: "Kroatisch"], Montenegrin, Serbian) L1 speaker. Three of the L1 German-speakers also gave English/Portuguese/BKMS as their L1, making them multilingual.

The evaluation of speaker variety was treated as a possible confounding factor for the later evaluation of /s/-variation and can thus be treated as demographic information. Nonetheless, the coding and testing of this variable was imbedded in the larger process of data collection and will therefore be discussed in the next section.

ii. Data Collection

The perception study looked for listener judgement of the indexicality of /s/-variation. In order to elicit these judgements, a listening task was devised, in which participants would hear different versions of the same segment wherein only the /s/-frequency was changed. To find a fitting segment for this experiment, I looked for a participant whose F0 was consistently not much higher than 100Hz to rule out listener bias on the basis of pitch. This pitch measurement was conducted in Praat for each of the map tasks. After selecting a suitable speaker, a fitting segment was determined. This was done based on the following criteria: the segment had to (1) encompass many /s/-tokens which could be exchanged (2) not include many other sibilants which might function as confounding factors (3) was intonationally relatively neutral so as to not evoke too many other strong evaluations (especially regarding 'gay speech' [Levon 2007]) and (4) ended on a downward intonation in order to avoid another possible bias associated with uptalk (Ritchart & Arvaniti 2014; Tyler 2015).

With the appropriate speaker and segment selected, the process of finding /s/-tokens which could be used to substitute the original tokens in this segment began. The goal was to find tokens with normatively masculine (/s-/), normatively feminine (/s+/) and gender ambiguous (/s~/) /s/-frequency (comp. Zimman 2018) that could then be spliced out of their regular environment to replace the original recording's /s/-tokens. After some experimentation, I decided to only use /s/-tokens by the same speaker and recording the segment was originally taken from in order to avoid any interference of background noises or harmonics from surrounding vowels. Luckily, the selected speaker (P5) showed great variation in his CoG-frequency which made finding three suitable tokens in his dataset quite easy.

The three tokens that were selected placed at approximately 4,100Hz for /s-/, 6,500Hz for /s~/ and 8,500Hz for /s+/. All /s/-tokens in the original recording were then replaced with the respective tokens, creating four versions. Target token length placed at 0.092 second (/s+/ & /s~/) and 0.07 seconds (/s-/) and sibilant length was not adjusted which

means the original tokens were uniformly replaced with the target token in order to not add another confounding factor. This way, four versions of the same segment were created: the original without any manipulation, the high-frequency, low-frequency and ambiguous-frequency version.

Participant 5, who was determined as the most suitable speaker for this experiment, was an Austrian Standard speaker. Considerations concerning the relevance of speaker variety also emerged in the process of finding the best segment to manipulate. Listener evaluation regarding other variables may differ based on perceived speaker variety/L1. Therefore, an option to evaluate the speaker's variety/origin was added to the list of evaluation items for the unaltered sample. On the one hand, if many of the listeners judged the speaker to be either an L2 speaker or a non-Austrian L1 speaker, this might be useful information to determine if this evaluation had an impact on other perceptions. On the other hand, asking participants to speak to variety was also a useful tactic to divert attention away from the target scales.

Finally, an online survey was created on Sco-Sci and sent out to participants on different social media platforms. Demographic data (as mentioned above) was collected through single choice questions and an open question about the speaker's perceived regional accent or dialect was posed after playing the first, unaltered recording. All questions were posed in the German politeness form of *Sie* and required an answer except for the dialect/accent evaluation. This was an oversight on my part as I added this question to the questionnaire later and overlooked that it was not coded as mandatory. Then, eight questions on semantic differential scales were posed right after hearing each sample. Six of the eight scales drew attention away from the two target scales but also functioned as control variables: should any surprising differences emerge, these might be just as informative as the ratings on the target scales. Values on the left end of the scale were assigned the value of 1 while the other extreme was assigned the value of 100. This design was chosen to replicate the study design of Boyd et al. (2021).

Just as in the latter study, the target scales for rating the guise (in German) were:

- "heterosexuell / homosexuell" (heterosexual / homosexual)
- "maskulin / feminin" (masculine / feminine)

Other scales included:

- "faul / fleißig" (lazy / hardworking)
- "unnatürlich / natürlich" (unnatural / natural)
- "unangenehm / angenehm" (unpleasant / pleasant)
- "ungebildet / gebildet" (uneducated / educated)
- "unfreundlich / freundlich" (impolite / polite)
- "unglaubwürdig / glaubwürdig" (not credible / credible).

In a last step, listeners were also asked if they had ever heard of the 'gay lisp' phenomenon and to please elaborate on it if they had. This question was included to assess whether explicit knowledge of the phenomenon had any impact on ratings on the target scales. The idea to test for this possibility emerged from an observation Boyd et al. (2021) make: even though one participant (in in-person testing) noted that they had heard of the 'gay lisp' in English, they stated that this was not an occurrence in German. This, as was mentioned earlier, is also reflected in the study's perception data but not mirrored in production. Testing for the effect explicit knowledge might have on listeners' indexicality judgements could therefore be a step on the path to understanding the phenomenon's scope and to determining if it might be on its way to enregisterment. The question was asked in the last step of the survey to avoid the Halo effect influencing how subsequent questions would be answered.

A standard optional question asking for comments, suggestions, issues and/or ideas concluded the survey. The study was presented as part of an MA thesis in Linguistics at the University of Vienna but the aim and topic as well as the title were omitted.

iii. Data Analysis

After a month, the study was taken offline and ScoSci provided a dataset that could be used for statistical testing in RStudio[6]. As a large number of tests were run, these will be listed here in chronological order. Correlations were tested not only for the two target scales but also for the other scales. This was done in order to determine if listeners made any significant differentiations between the variants, even if not on the target scales. The hypothesis was that the two target scales would show significant group differences while the other scales would not. If no significant differences for the target scales emerged, this would suggest that the samples may indeed have been too similar to distinguish. If any of the other scales showed significant differences, this information might be pertinent to post-hoc analysis. The first two tests were run in order to determine if there was a clear tendency regarding the study's research question. The other six tests were explorative and served the purpose of better understanding the data at hand.

In another step, the four variables *Gay Lisp*, *Age*, *Rural/Urban* and *Gender* were chosen for explorative analysis as well. *Gay Lisp* was interesting due to the explicit mentioning of the gay lisp phenomenon in English by a German-speaking participant in Boyd et al. (2021). The question arose if knowing of the practice would have any influence on how speakers were rated in regard to their gender and sexual orientation. The pertaining hypothesis was that explicit knowledge influences perception on the two target scales. *Age, Rural/Urban* and *Gender* were chosen as these seemed like plausible variables that might influence perception. What's more, all of these variables were easy to code as binary without forfeiting too much complexity.

In *Gay Lisp*, the binary coding was done based on the answer to the question (*Yes* or *No*). In *Age,* there was a large skew towards participants being under 35. This meant that the more deciding factor would have to be generational differences. The split at 40 was chosen with regard to the generally high level of education. I therefore chose

[6] RStudio Version 2024.12.0

40 as the cutoff (with all groups up to and including *36 to 40* taken together) as I thought this younger group might be more inclined towards public discourse on gender and sexuality. As for *Rural/Urban*, the choice of taking the two largest categories (*up to 100.000 inhabitants* and *over 100.000 inhabitants)* together and testing them against the smaller categories taken together was made based on the larger Austrian University cities: Vienna, Graz, Linz, Salzburg, Innsbruck and Klagenfurt all fall roughly into this definition of 'urban', while smaller cities and towns do not. For *Gender*, as there was only one non-binary participant, I coded groups as *Male* and *Non-Male*. This was done in regard to 'maleness' generally being a strong predictor for how one might react to gendered speech (Eckert & McConnel-Ginet, 2013).

For the first eight tests, the variable */s/-variant* refers to the version of the spoken segment the participants listened to. The 8,500Hz CoG version will be shortened to /*s*+/, the 4,100Hz version to /*s*-/, the 6,500Hz version to /*s*~/ and the unaltered version to /*s*/. All data analysis methods were run in RStudio with a confidence interval of 95%.

1. **Kruskal-Wallis test for group differences (/s/-*variant*) on scale *masculine/feminine***

The dependent variable "masculine/feminine" was tested for group differences (/s+/, /s-/, /s~/, /s/), in the grouping variable "/s/-variant". The "masculine/feminine" variable was rated on a metric scale between 1 and 100 while there were four groups for /s/-variant: /s+/, /s-/, /s~/ and /s/. The values for all four groups were combined into one group to test for data distribution. A Shapiro-Wilk test was administered which revealed that the data was not normally distributed (p-value = 1.193e-06). In order to determine if there were any prominent anomalies, the data was plotted on a histogram as can be seen in Figure 5.

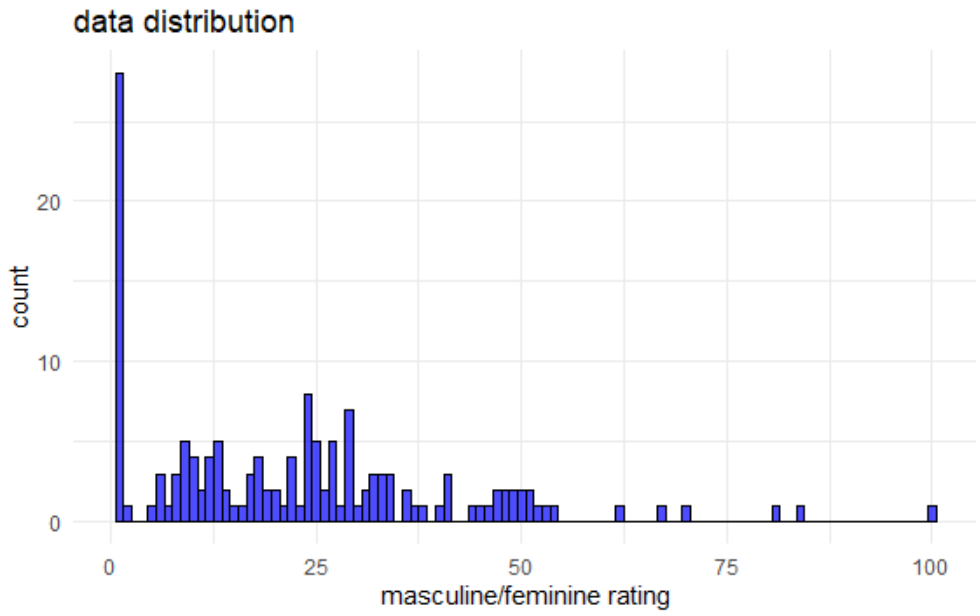

Figure 5: ratings on the masculine/feminine scale for all four variants combined, n = 37

The value *1*, which corresponds to the most 'masculine' rating, was evidently selected most often. This was not surprising; however, I decided to explore if this data might be exaggerated by singular unreliable or out of the ordinary raters. While going over the dataset, I discovered that four participants had rated each sample with a value of 1 as opposed to other participants whose ratings varied, even if just slightly. Two possibilities emerged: either those four participants were unreliable raters, or they represented a distinct group of raters who genuinely could not make out any differences between the variants. In order to avoid bias, two datasets were created – one with n =37, one with n = 33 excluding those four participants. Both datasets were tested to see if differences would emerge, which could in turn be interpreted in regard to the four excluded participants' characteristics.

The second dataset showed a less extreme distribution pattern (Figure 6), albeit still not normal distribution (Shapiro-Wilk test: p-value= 1.193e-06). However, as the speaker was chosen specifically for his low, normatively masculine F0, this was neither surprising nor bad. Therefore, a Kruskal-Wallis test was administered for both datasets. The hypothesis was that significant group differences would be observed.

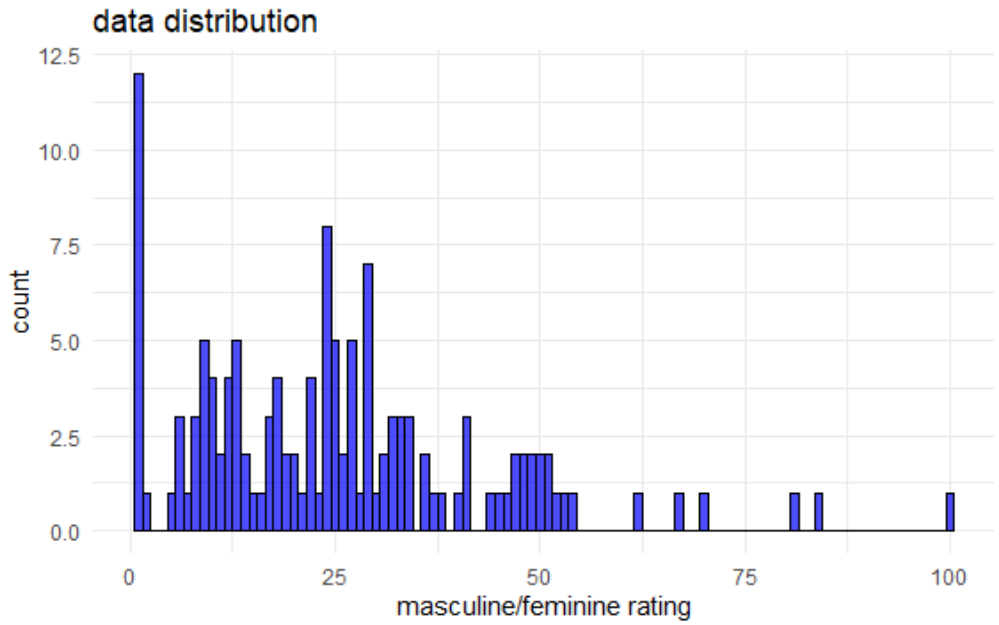

Figure 6: ratings on the masculine/feminine scale for all four variants combined, n = 33

2. **Kruskal-Wallis test for group differences (/s/-*variant*) on scale *heterosexual/homosexual***

The dependent variable *heterosexual/homosexual* was tested for group differences in the independent variable /s/-*variant*. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (p-value = 0.0006844 for either). Consequently, a Kruskal-Wallis test was administered. The hypothesis was that significant group differences would be observed.

3. **Kruskal-Wallis test for group differences (/s/-*variant*) on scale *impolite/polite***

The dependent variable *impolite/polite* was tested for group differences in the independent variable /s/-*variant*. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (p-value = 0.009072 for either). Consequently, a Kruskal-Wallis test was administered. The hypothesis was that no significant group differences would be observed.

4. **Kruskal-Wallis test for group differences (/*s*/-*variant*) on scale *uneducated/educated***

The dependent variable *uneducated/educated* was tested for group differences in the independent variable /*s*/-*variant*. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (p-value = 0.001727 for dataset n = 37, p-value = 0.003106 for dataset n = 33). Consequently, a Kruskal-Wallis test was administered. The hypothesis was that no significant group differences would be observed.

5. **Kruskal-Wallis test for group differences *(/s/-variant)* on scale *lazy/hardworking***

The dependent variable *lazy/hardworking* was tested for group differences in the independent variable /*s*/-*variant*. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (p-value = 1.049e-07 for either). Consequently, a Kruskal-Wallis test was administered. The hypothesis was that no significant group differences would be observed.

6. **Kruskal-Wallis test for group differences *(/s/-variant)* on scale *unnatural/natural***

The dependent variable *unnatural/natural* was tested for group differences in the independent variable /*s*/-*variant*. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (p-value = 0.0009475 for either). Consequently, a Kruskal-Wallis test was administered. The hypothesis was that no significant group differences would be observed.

7. **ANOVA for group differences (/*s*/-*variant*) on scale *not credible/credible***

The dependent variable *not credible/credible* was tested for group differences in the independent variable /*s*/-*variant*. A Shapiro-Wilk test revealed that the distribution was

normal for both datasets (p-value = 0.3126 for either). A Levene test determined that variances between the groups were normal, too (p-value= 0.9431 for n = 37, p-value = 0.9286 for n = 33). Consequently, a single measure ANOVA was administered. The hypothesis was that no significant group differences would be observed.

8. **ANOVA for group differences (/s/-*variant*) on scale *unpleasant/pleasant***

The dependent variable *lazy/hardworking* was tested for group differences in the independent variable /s/-*variant*. A Shapiro-Wilk test revealed that the distribution was normal for both datasets (p-value = 0.2778 for either dataset). A Levene test determined that variances between the groups were normal, too (p-value= 0.5335 for n = 37, p-value = 0.656 for n = 33). Consequently, a single measure ANOVA was administered. The hypothesis was that no significant group differences would be observed.

9. **Mann-Whitney U test for group difference (*Gay Lisp)* on target scale *masculine/feminine***

To test if explicit knowledge had an effect on the ratings on this target scale, the answer to the question if one had ever heard of the phenomenon of the 'gay lisp' was treated as an independent variable with two groups (Yes/No). The *masculine/feminine* rating was treated as the dependent variable. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (group *Yes*: p-value = 0.00000365 for n =37, p-value = 0.0006233 for n = 33, group *No*: p-value 0.0000345 or n = 37, p-value = 3.451e-05 for n = 33). Consequently, a Mann-Whitney U test was administered. The hypothesis was that significant group differences would be observed.

10. **Mann-Whitney U test for group difference (*Gay Lisp*) on target scale *heterosexual/homosexual***

Again, "Yes/No" was treated as the independent grouping variable. This time, the ratings on the target scale *heterosexual/homosexual* were treated as the dependent

variable. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (group *Yes:* p-value = 0.0000326 for n =37, p-value = 0.0211 for n = 33, group *No*: p-value 0.00191 for n = 37, p-value = 0.001911 for n = 33). Consequently, a Mann-Whitney U test was administered. The hypothesis was that significant group differences would be observed.

11. Mann-Whitney U test for group difference (*Age*) on scale *masculine/feminine*

For the independent variable *Age*, the pre-defined age categories were taken together to create a binary grouping of *<40* and *>40*. Again, *masculine/feminine* was treated as the dependent variable. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (group *<40*: p-value = 0.0000000899 for n =37, p-value = 0.000000589 for n = 33, group *>40*: p-value 0.0204 for either dataset). Consequently, a Mann-Whitney U test was administered instead of a t-test for two-group variables. The hypothesis was that significant group differences would be observed.

12. Mann-Whitney U test for group difference (*Age*) on scale *heterosexual/homosexual*

Once more, the dependent variable was *heterosexual/homosexual*. The independent grouping variable was, as before, age with the two same age groups. A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (group *<40*: p-value = 0.00000457 for n =37, p-value = 0.000525 for n = 33, group *>40*: p-value 0.0733 for either dataset). Consequently, a Mann-Whitney U test was administered. The hypothesis was that significant group differences would be observed.

13. Mann-Whitney U test for group difference (*Rural/Urban*) on scale *masculine/feminine*

For this test, the dependent variable was *masculine/feminine*. The independent grouping variable was *Urban/Rural* with the two groups *R* (rural) and *U* (urban). A Shapiro-Wilk

test revealed that the distribution was not normal for either dataset (group *R:* p-value = 2.053e-05 for n =37, p-value = 0.0001199 for n = 33, group *U*: p-value = 0.0000205 for n = 37, p-value = 0.000120 for n 0 33). Consequently, a Mann-Whitney U test was administered. The hypothesis was that significant group differences would be observed.

14. **Mann-Whitney U test for group difference (*Rural/Urban)* on target scale *heterosexual/homosexual***

For this test, the dependent variable was *heterosexual/homosexual.* The independent grouping variable was *Urban/Rural* with the two groups *R* (rural) and *U* (urban). A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (group *R*: p-value = 0.0778 for n =37, p-value = 0.392 for n = 33, group *U:* p-value = 0.0000294 for n = 37, p-value = 0.000384 for n 0 33). Consequently, a Mann-Whitney U test was administered. The hypothesis was that significant group differences would be observed.

15. **Mann-Whitney U test for group difference (*Gender*) on target scale *masculine/feminine***

Here, the dependent variable was again *masculine/feminine.* The independent grouping variable was *Gender* with the two groups *NM* (non-male) and *M* (male). A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (group *NM*: p-value = 0.0000000508 for n =37, p-value = 0.00000316 for n = 33, group *M*: p-value = 0.0405 for either dataset). Consequently, a Mann-Whitney U test was administered. The hypothesis was that significant group differences would be observed.

16. **Mann-Whitney U test for group difference (*Gender)* on target scale *heterosexual/homosexual***

Here again, the dependent variable was *heterosexual/homosexual.* The independent grouping variable was *Gender* with the two groups *NM* (non-male) and *M* (male). A Shapiro-Wilk test revealed that the distribution was not normal for either dataset (group *NM*: p-value = 0.0000169 for n =37, p-value = 0.00184 for n = 33, group *M*: p-value = 0.0405 for either dataset). Consequently, a Mann-Whitney U test was administered. The hypothesis was that significant group differences would be observed.

5. Results

c. Quantitative production results

- Dialect group: The mean in group *BA* was 6,064.016 Hz, the mean in group *SD* 5,961.328 Hz. The t-test for this independent variable showed that this difference was not significant (p-value = 0.8022). The hypothesis was accepted.
- Nationality: The mean in group *AT* was 6,315.813 Hz, the mean in group *DE* 5,440.452 Hz. The t-test for this independent variable showed that there was a significant correlation between participants' nationality and their CoG frequency (p-value = 0.02945)**.** The hypothesis was rejected.
- Age at beginning of t-therapy: The Pearson-correlation-analysis for the independent variable *age at beginning of t-therapy* showed that there was a significant correlation between participants' age at the beginning of testosterone therapy and their CoG frequency (p-value = 2.455e-08). The hypothesis was accepted.

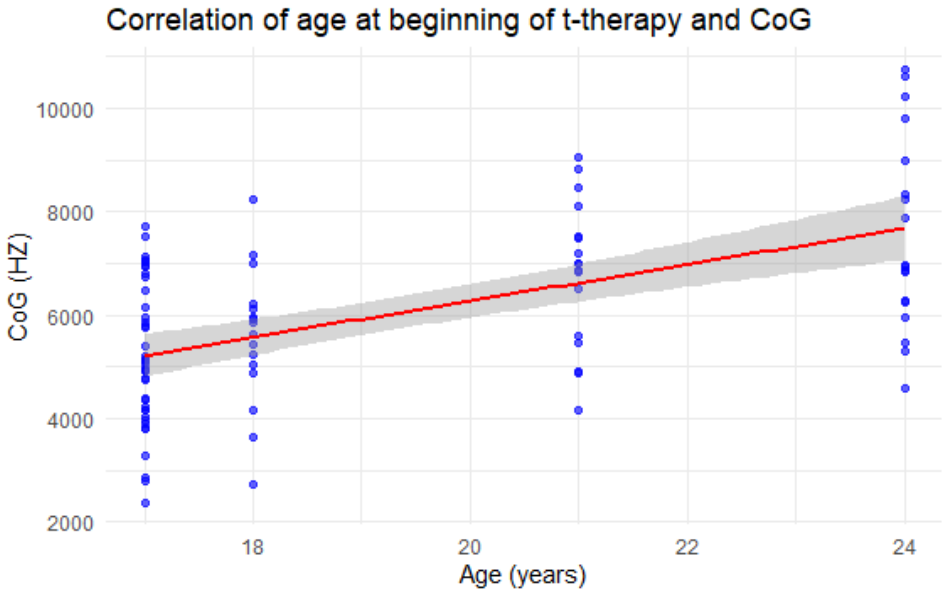

Figure 7: scatterplot of age at beginning of t-therapy (years) by CoG (Hz)

- Time on testosterone: The Pearson-correlation-analysis for the independent variable *time on testosterone* showed that there was no significant correlation between participants' time on testosterone and their CoG frequency (p-value = 0.3334). The hypothesis was accepted.
- Map task data:

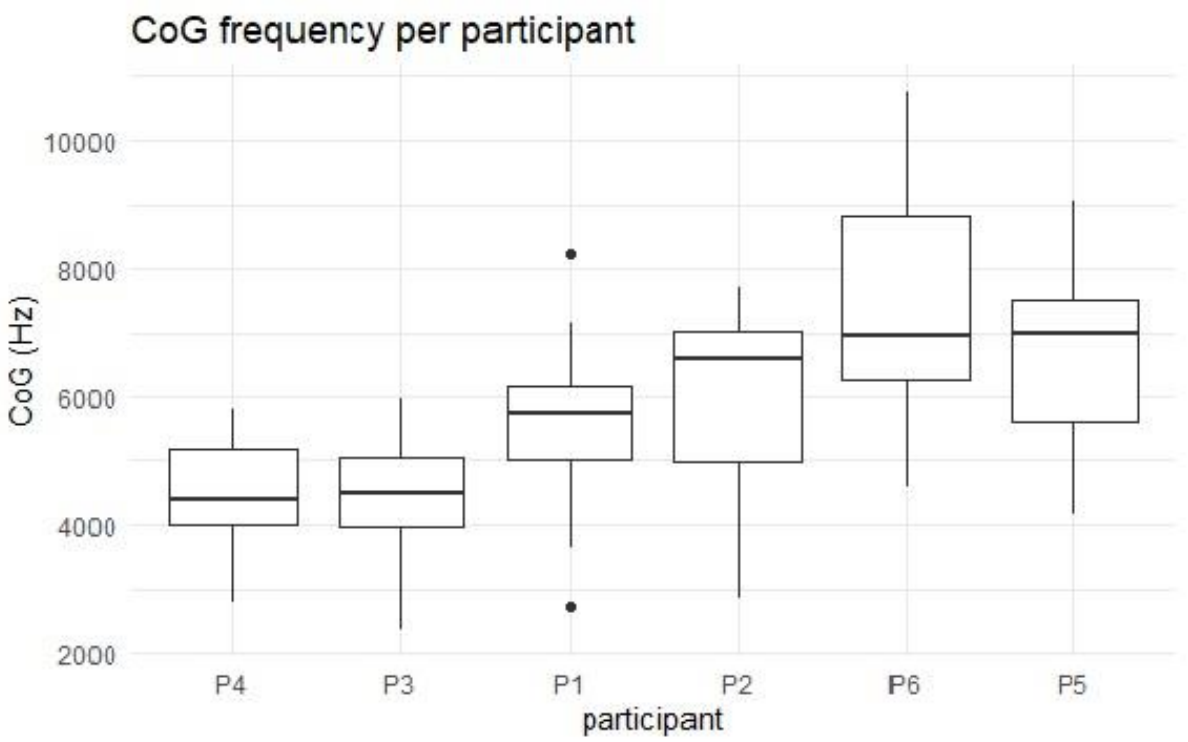

Figure 8: means & distribution of CoG per participant (in increasing order of means)

Participant	tokens	speaker min & max in Hz	speaker range as difference in Hz	mean COG in Hz
P1	16	2,729 – 8,250	5,521	5,579
P2	18	2,869 – 7,730	4,861	5,961
P3	11	2,366 – 5,948	3,582	4,503
P4	11	2,780 – 5,796	3,016	4,542
P5	16	4,157 – 9,049	4,892	6,697
P6	18	4,587 – 10,755	6,168	7,577
range of means			4,503 – 7,577	

Table 7: token number, CoG mean, minimum & maximum, range as difference between minimum and maximum per participant & range of means total

Firstly, what can be said about these results is that they very much echo those of Zimman (2017a), showing a highly varied pattern of /s/-articulation. Although, in the latter study, participants' CoG mean is generally set a little higher (between 5,226 Hz and 9,116 Hz) than in the present one (between 4,503 Hz and 7,577 Hz), the inter-speaker CoG range is similar in both. The range of means spans a difference of 3,074 Hz in the present results, which is similar to the 3,890Hz difference Zimman finds.

Participants 4 and 3 show very normatively 'masculine' CoG (as discussed in the Theory section) with means around 4,500 Hz. These speakers' respective minimums are set at 2,366/2,780 Hz and their maximums at 5,948/5,796 Hz. This creates a similar range for the two participants at 3,582/3,016 Hz. These are also the two speakers with the fewest tokens (11 respectively).

Participants 1 and 2 also show similar means closer to the more 'gender ambiguous' range at 5,579/5,961 Hz and produced 16/18 appropriate /s/-tokens. Their minimum and maximum also looked quite similar: for P1 they were set at 2,729 – 8,250 Hz, for P2 at 2,869 – 7,730 Hz. This also means that their ranges approximate each other at 5,521/4,861 Hz.

Participant 5 also shows a similar range of 4,892 Hz. The difference between P5 and the former two, however, is that his minimum and maximum are set higher at 4,157 – 9,049 Hz. This is quite similar to P6 whose minimum and maximum are also set in a more normatively 'feminine' range at 4,587 – 10,755 Hz. Both of these Participants show higher means than the others with 6,697 Hz for P5 and 7,577 Hz for P6.

d. Qualitative production results

The core result in the qualitative production part is Table 8 which encompasses quantitative as well as qualitative participant features. The interpretation of this table which was carried out by consulting Table 7 and Figure 4 forms the qualitative result of this production part.

The subsequent sociolinguistic analysis encompasses all findings of the production part and uses them as resources to understand the semiotic process at hand.

Participant	age	nationality	dialect(s)	Gender identity & presentation	Sexual identity	time on testosterone in months	age at beginning of t-therapy	pronouns	role models
P1	21	AT	Lower Austrian Standard	non-binary, "5/8 man", trans, *performs as drag queen*	queer/bi (NOT straight)	30	18	er (DE) they (EN)	Jinx Monsoon, father
P2	23	DE	Southern German Standard	man ("I don't think about my gender") *rarely gets misgendered*	bi, used to want to be straight to pass as male	72	17	he/him	trans kids
P3	21	DE	Lower Bavaria	man, *wears make-up since passing as male*	bi, tendency towards men because easier	48	17	he/him	the Gender Role of "man"
P4	18	AT	Lower Austria/ Carinthia	stereotypical man (trans identity not important), *wears nail polish sometimes*	queer (not important), on asexual spectrum	12	17	he/him	favorite teacher, Morgan Freeman's voice
P5	25	AT	Vienna	man/trans masculine "reluctantly male" "a man but barely"	queer/bi (NOT straight)	42	21	he/him, ok with they	"soft men who don't look soft", Grandfather
P6	25	AT	Upper Austria	trans man/ trans masculine *does not care for gender norms in clothing too much*	queer (no cis men)/ attracted to trans men, politically lesbian	13	24	he/him	Elliot Page, trans men who have biological kids

Table 8: Qualitative & quantitative self-descriptions of all six participants

To begin with, as has been discussed in the Theory section, the varied CoG pattern that can be seen in Table 7 and Figure 4 could not simply be explored by looking for singular causality or even clear correlations that become apparent at first sight. Instead, I was looking to understand if any of the factors in Table 8 might come together to produce a meaningful cluster of features that functions as the field of stylistic bricolage within which the individual's /s/-variation gains discursive function.

Considering the findings from the quantitative analysis, *nationality* and *age at beginning of t-therapy* already formed a basis upon which to operate. On average, German citizens had a significantly lower CoG frequency than Austrian ones and the later the participant started testosterone-therapy, the higher their CoG would be. Therefore, the next step was to compare P2 and P3 who were German. They seemed relatively distinct from each other on a quantitative level: P2 had a much higher mean CoG (5,961 Hz) than P3 (4,503 Hz) and their range was also quite different at 4,861 Hz for P2 and 3,582 Hz for P3. What they did have in common was that their minimum was set relatively low (2,869 Hz for P2 and 2,366 Hz for P3). It seemed, therefore, possible that Germans make more use of the lower end of the spectrum which had to be taken into account in the present analysis.

Looking again at Table 8, I noticed that both of these Participants had started testosterone therapy relatively early at 17 and had also been on testosterone the longest. Both of these individuals had emphasized in conversation that they preferred not to divulge their trans identity if not absolutely necessary. A third participant, P4, had also made such a statement. In light of this, I noticed that there were a lot of qualitative as well as quantitative factors tying these three Participants together. For one, they were the three youngest transitioners, all starting at 17. What's more, they had all expressed that the label "trans" was not important to them and that they, to the contrary, identified as men without any additional labels. As mentioned, they all preferred not to reveal their trans identity.

I also noticed that P2 had said that he had wanted to pass as straight when he was younger in order to help him pass as cisgender. Sexual identity was also a topic P3 and I talked about in detail. He stated that, although he saw himself as bisexual, he had a tendency towards men because he did not want to fulfill women's expectations of him always having to be the more proactive, strong and dominant partner. P4, on the other hand, stated that he saw himself as somewhere on the asexual spectrum but generally did not have any gender preference when it came to romantic and sexual partners.

Here, the concept of hegemonic masculinity as discussed in the Theory section became useful. I found these three individuals, on the one hand, to embody non-hegemonic masculinity as they were trans and queer. Yet only one of them also deviated from normatively masculine /s/-articulation in their own language use. While P3 and P4 both showed relatively low CoG-range and -mean, P2 seemed to make use of a much larger part of the spectrum. Both P3 and P4 exhibited minimums and maximums entirely within the normatively masculine range, never venturing into the range that is associated with non-hegemonic masculinity. Hence, the question of what might set those two participants apart from P2 led to the proposal of participants performing hegemonic vs non-hegemonic masculinity on different dimensions of masculinity.

All three had expressed their wish to be seen as cis and did not feel attached to the label *trans*. Yet, one distinguishing factor became apparent in revisiting the interviews: P2 stated that he was rarely or never misgendered anymore while P4 and P3 expressed that being misgendered was still an uncomfortable reality in their lives. The latter two were also younger (18 & 21) than P2 (23) and had consequently spent less time on testosterone and living openly in their male gender role. It seemed, therefore, possible that P3 and P4 were less sure of their outward passing as cis which in turn might manifest as a more rigid linguistic range of expression. While all three had stated the desire to express their masculinity in a non-hegemonic form (to wear make-up/nail polish, to be openly queer/asexual, to pursue hobbies that were seen as 'feminine' like drawing and crafting), they also strongly associated themselves with the hegemonic

masculinity inherent in passing as cis. Therefore, the working hypothesis emerged that the individual's level of confidence that they could pass as cis might be just as important a factor in their stylistic choices as the importance they placed on being seen this way.

With this observation, I looked to the other three participants to explore if their gendered performance of (non-) hegemonic masculinity might also be reflected in their CoG-patterns. For P1, although they had also started to transition earlier than P5 and P6 (at 18), their linguistic as well as non-linguistic stylistic choices seemed to aim at non-hegemonic masculinity. Fistly, they identified more as trans and non-binary than male, performed drag and also very much emphasized that they were not straight. What's more, they also stated that they had chosen their name based on how easily it could be viewed as non-male and said that the label *trans* was definitely important to them. When asked how they wanted to be perceived from the outside they replied, "rather as a man" ("eigentlich schon als Mann"). Yet, their gender expression seemed more important to them than their perceived gender as they reiterated, when asked if they would identify as a man, that they would rather identify as non-binary and trans. When asked what label they would give themselves, they said that they saw themselves as "5/8 man and the rest is something airy, indefinite".

P5 and P6 also stated that being trans masculine was important to them. Although they both said that they would not divulge their trans identity in every situation, being perceived as trans was not something they claimed to be generally uncomfortable with. P5 said that nowadays a lot of people thought he was trans feminine when they learned he was trans and that he found this funny. He also stated that he had "never wanted to be a man but just happened to be one" and was "reluctantly male" and "a man but barely". He continued to explain that, now that he passed as male, it was important to him to actively let people in on his trans identity as he didn't want to be seen as "some random dude" especially around AFAB people. "I like it when people know but I also like to use the privilege of people not knowing in certain situations",

he elaborated. Contrary to P2, P3 and P4, he relayed that he wanted people he got to know better to "figure out I'm not quite cis" on their own.

Although P6 said that he thought his voice still confused people and said he wanted it to get deeper still, he stated that he "mainly thought of himself as himself rather than as trans masculine" and "could not really relate to cis masculinity at all". P6 explained that this was also the reason he had initially identified as non-binary before starting his transition. For three years he had used no pronouns and identified as non-binary because he had found cis masculinity "boring, rigid and unemotional" and could not identify with this whatsoever. He stated that this was why it was important that he had gotten to know and see different kinds of masculinity (queer and trans): because he had learned that masculinity could also be "loud and emotional and that I didn't have to change my personality by 180 degrees to be a man". What's more, he said that there were situations where he wouldn't out himself like at work or with total strangers but that, generally, he wanted people he knew to know about his gender identity.

In order to make sense of all of these aspects, I devised a table trying to capture all these dimensions of (non-)hegemonic masculinity. Following Zimman's (2017a) division of different facets of gender, I divided the participants' masculinity into three dimensions within which they wished to be perceived a certain way: *passing, gender expression* and *sexual orientation*. I replaced the original category *sex* with *passing* as this was the phrase the participants used to describe the effect their gendered performance should have on their perceived sex. For each of those dimensions, I noted if the individual wanted to be seen in a hegemonically masculine (h.) or non-hegemonically masculine (n.h.) way. Wanting to be perceived as queer, trans, or in some way non-normatively masculine was counted toward non-hegemonic masculinity while the opposite was true for the wish to be perceived as straight, cis and normatively masculine. When participants had emphasized certain aspects, I highlighted the corresponding answers in order to make them stand out.

desired perception in:	*passing*	*gender expression*	*sexual orientation*
P1	h.	n.h.	n.h.
P2	h.	n.h.	h. --> n.h.
P3	h.	n.h.	n.h.
P4	h.	n.h.	n.h.
P5	n.h.	n.h.	n.h.
P6	n.h.	n.h.	n.h.

Table 9: Participants' desired perception in each aspect of (non-) hegemonic masculinity

I thus formulated the proposition that it was not only important in how many of these dimensions the individual wanted to be perceived as performing hegemonic masculinity, but also how significant a given dimension was to them.. While P1 wanted to be perceived as cis male concerning their *passing*, they did not place a lot of importance on this aspect and instead emphasized their trans and queer identity much more. P2, on the other hand, made a point of wanting to pass as cis and having wanted to pass as straight, too, while the non-hegemonic aspects of his masculinity did not seem to be quite as important to him. P3 placed no importance on his trans identity and said it was important to him that other people also viewed him as "a typical man" while mentioning his desire to perform in a non-hegemonic way concerning his *gender expression* and *sexual orientation* without giving this the same importance. For P4, a very similar picture emerged: he wished to be perceived as cis and accordingly performed hegemonic masculinity in the aspect of *passing* but also mentioned that he did wear nail-polish and felt rather asexual and queer, thereby performing non-hegemonic masculinity but not giving these dimensions of *gender expression* and *sexual orientation* the same emphasis. P5 reported that his non-hegemonic masculinity was important to him, especially in the dimension of *passing* since he did pass as male and wanted to make it clear he was not "some random dude". He also talked about queerness being very important to him as opposed to P3 and P4. P6 performed non-hegemonic masculinity in all three aspects and all three of them seemed almost equally important

to him. Most notably, however, he highlighted that he was not straight and still considerably identified with the political concept of lesbianism.

Relating all of this to the pattern of CoG-variation visible in Figure 4, I found it most likely that what could be observed was not only the importance that individuals place on these aspects but also the degree to which they felt they had achieved the desired perception. While P2 placed the same importance on performing hegemonic masculinity in the aspect of *passing* as P3 and P4, P2 seemed to be much more confident in the outcome of this performance. It is possible, then, that his rather varied and non-normatively masculine articulations of /s/ could be seen as the freedom he allowed himself to perform non-hegemonic masculinity without running the risk of being seen as trans. This could be seen as a trade-off since the risk of being put into the 'gay' or non-normative category is thereby increased, which might result in associations with transness, yet the participant might feel this risk is worth the freedom to express himself.

P3 and P4, on the other hand, might have wanted to be perceived as cis as well but with less confidence in that general perception than P2. Therefore, their CoG patterns might allow for less freedom and display more rigidity. Being perceived as performing non-hegemonic masculinity in one aspect might still result in being perceived as less hegemonically masculine in the aspect of *passing*. In order to avoid this, fewer risks might be taken.

P1 showing such an ambiguous CoG distribution also fits into this picture. While they wish to be perceived as a man, this desire is much less pronounced than for the P2, P3 and P4. Allowing themselves to perform non-hegemonic masculinity and thereby embrace much more of the spectrum of possible /s/ articulations runs a much lower risk of being exposed because the importance that is placed on wanting to be seen as male is subordinate to the desire to be perceived as non-binary and queer.

The fact that P5 and P6 produce CoG frequencies which lie in the more normatively 'feminine' range would then also point to how important it is to both of them to not be

viewed as cis and to hence perform non-hegemonic masculinity. Neither of them run the risk of being exposed as a non-hegemonic man since they actively seek to perform this identity already. Considering that these are also the two latest transitioners, it seems possible that both of them had already found aspects of their gender identity outside of *passing* to be more pivotal to their identity. Although *passing* looks to be the biggest factor in how much variability individuals allow themselves in their articulations of /s/, how important this aspect is to the individual's gendered performance might also vary.

Therefore, the varied picture we see could be traced back to (1) which aspects of masculinity participants want to perform in an outwardly (non-)hegemonic way (2) the importance participants place on each of those aspects and (3) how successful they feel they are at performing the desired identity.

e. Perception Results

Table 10 and 11 report the means of the ratings on each scale for each variant in each dataset.

n = 37	*8,500Hz (/s+/)*	*4,100Hz (/s-/)*	*6,500Hz (/s~/)*	*unaltered (/s/)*
masculine/feminine	21.81	25.89	24.16	19.43
hetero/homosexual	49.27	48.27	46.32	50.05
(im)polite	60.05	58.86	55.46	67.40
(un)educatd	57.43	57.43	54.62	67.38
lazy/hardworking	46.65	51.73	53.03	52.30
(un)natural	63.67	46.83	42.97	66.43
(not) credible	63.24	58.08	55.67	70.81
(un)pleasant	52.70	46.05	41	62.83

Table 10: mean rating on each scale for each /s/-variant in the original dataset (n = 37)

n = 33	*8,500Hz (/s+)*	*4,100Hz (/s-)*	*6,500Hz (/s~/)*	*unaltered (/s/)*
masculine/feminine	24.33	28.91	26.96	21.66
hetero/homosexual	49.45	48.30	45.81	50.18
(im)polite	59.24	58.48	55.18	66.69
(un)educatd	56.33	56.39	53.91	66.54
lazy/hardworking	44.76	51.88	52.85	51.03
(un)natural	64.87	47.87	42.51	66.45
(not) credible	62.24	57.88	56	69.64
(un)pleasant	53.09	46.24	42.21	62.33

Table 11: mean rating on each scale for each /s/-variant in the adjusted dataset (n = 33)

Table 12 shows the results of all 16 statistical tests. “SW” stands for Shapiro-Wilk, “KW” for Kruskal-Wallis, and “MWU” for Mann-Whitney U. Highlighted results are significant, if applicable the groups that exhibit significant differences are also highlighted.

variables	n = 37	n = 33
/s/-variant x masculine/feminine	SW test: p-value = 1.193e-06 KW test: p-value = 0.568	SW test: p-value= 1.193e-06 KW test: p-value = 0.4785
/s/-variant x heterosexual/ homosexual	SW test: p-value = 0.0006844 KW test: p-value = 0.8174	SW test: p-value = 0.0006844 KW test: p-value = 0.7566
/s/-variant x (im)polite	SW test: p-value = 0.009072 KW test: p-value = 0.04916 Dunn test: /s+/ x /s-/: p-value = 1.0000000 /s+/ x /s~/: p-value = 1.0000000 /s-/ x /s~/: p-value = 1.0000000 /s+/ x /s/: p-value = 0.6900068 /s-/ x /s/: p-value = 0.2546666 /s~/ x /s/: p-value = 0.0430857	SW test: p-value = 0.009072 KW test: p-value = 0.04832 Dunn test: /s+/ x /s-/: p-value = 1.0000000 /s+/ x /s~/: p-value = 1.0000000 /s-/ x /s~/: p-value = 1.0000000 /s+/ x /s/: p-value = 0.26117790 /s-/ x /s/: p-value = 0.00691076 /s~/ x /s/: p-value = 0.04146456
/s/-variant x (un)educated	SW test: p-value = 0.01452 KW test: p-value = 0.001727 Dunn test: /s+/ x /s-/: p-value = 1.0000000 /s+/ x /s~/: p-value = 1.0000000 /s-/ x /s~/: p-value = 1.0000000 /s+/ x /s/: p-value = 0.04593788 /s-/ x /s/: p-value = 0.061161144 /s~/ x /s/: p-value = 0.00111077	SW test: p-value = 0.01452 KW test: p-value = 0.003106 Dunn test: /s+/ x /s-/: p-value = 1.0000000 /s+/ x /s~/: p-value = 1.0000000 /s-/ x /s~/: p-value = 1.0000000 /s+/ x /s/: p-value = 0.063472205 /s-/ x /s/: p-value = 0.095338882 /s~/ x /s/: p-value = 0.001985676
/s/-variant x lazy/hardworking	SW test: p-value = 1.049e-07 KW test: p-value = 0.1963	SW test: p-value = 1.049e-07 KW test: p-value = 0.1419
/s/-variant x (un)natural	SW test: p-value = 0.0009475 KW test: p-value = 0.0002837 Dunn test: /s+/ x /s-/: p-value = 0.06880639 /s+/ x /s~/: p-value = 0.0147422 /s-/ x /s~/: p-value = 1.0000000 /s+/ x /s/: p-value = 1.00000000 /s-/ x /s/: p-value = 0.014221463 /s~/ x /s/: p-value = 0.00240166	SW test: p-value =0.0009475 KW test: p-value = 0.0003087 Dunn test: /s+/ x /s-/: p-value = 0.07198806 /s+/ x /s~/: p-value = 0.00826291 /s-/ x /s~/: p-value = 1.0000000 /s+/ x /s/: p-value = 1.00000000 /s-/ x /s/: p-value = 0.027994715 /s~/ x /s/: p-value = 0.002624251
/s/-variant x (not) credible	SW test: p-value = 0.2778 Levene test: p-value= 0.5335 ANOVA: p-value= 0.00606 Turkey HSD test:	SW test: p-value = 0.2778 Levene test: p-value = 0.656 ANOVA: p-value = 0.0154 Turkey HSD test:

	/s+/ x /s-/: p-value = 0.6706445 /s+/ x /s~/: p-value = 0.3493025 /s-/ x /s~/: p-value = 0.9523501 /s+/ x /s/: p-value = 0.3493025 /s-/ x /s/: p-value = 0.0301173 /s~/ x /s/: p-value = 0.0062510	/s+/ x /s-/: p-value = 0.7686911 /s+/ x /s~/: p-value = 0.5125136 /s-/ x /s~/: p-value = 0.9756279 /s+/ x /s/: p-value = 0.3613077 /s-/ x /s/: p-value = 0.0498622 /s~/ x /s/: p-value = 0.0159308
/s/-variant x (un)pleasant	SW test: p-value = 0.3126 Levene test: p-value = 0.9431 ANOVA: p-value= 0.000837 Turkey HSD test: /s+/ x /s-/: p-value = 0.6238747 /s+/ x /s~/: p-value = 0.1506481 /s-/ x /s~/: p-value = 0.7958291 /s+/ x /s/: p-value = 0.2595133 /s-/ x /s/: p-value = 0.0145395 /s~/ x /s/: p-value = 0.0006673	SW test: p-value = 0.3126 Levene test: p-value = 0.9286 ANOVA: p-value = 0.00146 Turkey HSD test: /s+/ x /s-/: p-value = 0.5744193 /s+/ x /s~/: p-value = 0.1787157 /s-/ x /s~/: p-value = 0.8739646 /s+/ x /s/: p-value = 0.3106735 /s-/ x /s/: p-value = 0.0160335 /s~/ x /s/: p-value = 0.0013824

Table 12: results of all multiple group statistical tests in chronological order

variables	n = 37	n = 33
Gay Lisp (Yes/No) x masculine/ feminine	SW test: p-value = Yes: p-value = 0.00000365 No: p-value = 0.0000345 MWU test: p-value = 0.1244 group mean Yes 26.8 No 24.7	SW test: Yes: p-value = 0.0006233 No: p-value = 3.451e-05 MWU test: p-value = 0.3354 group mean Yes 26.8 No 24.7
Gay Lisp (Yes/No) x heterosexual/ homosexual	SW test: Yes: p-value = 0.0000326 No: p-value = 0.00191 MWU test: p-value = 0.002203 group mean Yes 56.2 No 44.0	SW test: Yes: p-value = 0.0211 No: p-value = 0.001911 MWU test: p-value = 0.0003841 group mean Yes 56.2 No 44.0
Age (<40/>40) x masculine/ feminine	SW test: <40: p-value = 0.0000000899 >40: p-value = 0.0204 MWU test: p-value = 0.01054 group mean <40 20.5 >40 32.9	SW test: <40: p-value = 0.000000589 >40: p-value = 0.0204 MWU test: p-value = 0.06086 group mean <40 23.5 >40 32.9
Age (<40/>40) x heterosexual/ homosexual	SW test: <40: p-value = 0.00000457 >40: p-value = 0.0733 MWU test: p-value = 0.8607 group mean <40 49.6 >40 44.2	SW test: <40: p-value = 0.000525 >40: p-value = 0.0733 MWU test: p-value = 0.8607 group mean <40 49.6 >40 44.2
Rural/Urban (R/U) x masculine/ feminine	SW test: p-value = 2.053e-05 R: p-value = 0.00000282 U: p-value = 0.0000205 MWU test: p-value = 0.01303	SW test: p-value = 0.0001199 R: p-value = 0.0000744 U: p-value = 0.000120 MWU test: p-value = 0.1276

	group mean R 18.1 U 24.4	group mean R 22.9 U 26.2
Rural/Urban (R/U) x heterosexual/ homosexual	SW test: R: p-value = 0.0778 U: p-value = 0.0000294 MWU test: p-value = 0.006285 group mean R 40.3 U 50.6	SW test: R: p-value = 0.392 U: p-value = 0.000384 MWU test: p-value = 0.006285 group mean R 40.3 U 50.6
Gender (NM/M) x masculine/ feminine	SW test: M: p-value = 0.0405 NM: p-value = 0.0000000508 MWU test: p-value = 0.002505 group mean M 29.8 NM 20.3	SW test: M: p-value = 0.0405 NM: p-value = 0.00000316 MWU test: p-value = 0.04664 group mean M 29.8 NM 23.6
Gender (NM/M) x heterosexual/ homosexual	SW test: M: p-value = 0.244 NM: p-value = 0.0000169 MWU test: p-value = 0.008817 group mean M 55.8 NM 45.2	SW test: M: p-value = 0.244 NM: p-value = 0.00184 MWU test: p-value = 0.008817 group mean M 55.8 NM 45.2

Table 13: results of all dual group statistical tests in chronological order

1. ***/s/-variant x masculine/feminine***

The Kruskal-Wallis test showed that there was no significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *masculine/feminine* in either dataset (p-value = 0.568 for n= 37, p-value = 0.4785 for n= 33). The hypothesis was rejected.

2. ***/s/-variant x heterosexual/homosexual***

The Kruskal-Wallis test showed that there was no significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *heterosexual/homosexua*l in either dataset (p-value = 0.8174 for n= 37, p-value = 0.7566 for n= 33). The hypothesis was rejected.

3. ***/s/-variant x (im)polite***

The Kruskal-Wallis test showed that there was a significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *impolite/polite* in either dataset (p-value = p-value = 0.04916 for n= 37, p-value = p-value = 0.04832 for n= 33). The hypothesis was rejected.

4. ***/s/-variant x (un)educated***

The Kruskal-Wallis test showed that there was a significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *uneducated/educated* in either dataset (p-value = 0.001727 for n = 37, p-value = 0.003106 for n= 33). A subsequent Dunn test showed that in the original dataset, the variants that showed significant differences were /s+/ and /s/ (p-value = 0.04593788) as well as /s~/ and /s/ (p-value = 0.00111077). For the adjusted dataset only the latter case applied (p-value = 0.001985676). The hypothesis was rejected.

5. */s/-variant x lazy/hardworking*

The Kruskal-Wallis test showed that there was no significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *lazy/hardworking* in either dataset (p-value = 0.1963 for n = 37, p-value = 0.1419 for n= 33). The hypothesis was accepted.

6. */s/-variant x (un)natural*

The Kruskal-Wallis test showed that there was a significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *natural/unnatural* in either dataset (p-value = 0.0002837 for n = 37, p-value = 0.003087 for n= 33). A subsequent Dunn test showed that, in both datasets, the variants that showed significant differences were /s+/ and /s~/ (p-value = 0.0147422 fpr n = 37, p-value = 0.00826291 for n = 33), /s-/ and /s/ (p-value = 0.014221463 for n = 37, p-value = 0.027994715 for n = 33) as well as /s~/ and /s/ (p-value = 0.00240166 for n = 37, p-value = 0.002624251). The hypothesis was rejected.

7. */s/-variant x (not) credible*

The single measure ANOVA showed that there was a significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *not credible/ credible* in either dataset (p-value = 0.00606 for n = 37, p-value = 0.0154 for n= 33). A subsequent Turkey HSD test showed that, in both datasets, the variants that showed significant differences were /s-/ and /s/ (p-value = 0.0301173 for n = 37, p-value = 0.0498622 for n = 33) as well as /s~/ and /s/ (p-value = 0.0062510 for n = 37, p-value = 0.0159308 for n = 33). The hypothesis was rejected.

8. */s/-variant x (un)pleasant*

The single measure ANOVA showed that there was a significant correlation between the /s/-variant participants heard and their rating of the speaker on the scale *unpleasant/pleasant* in either dataset (p-value = 0.000837 for n = 37, p-value = 0.00146 for

n= 33). A subsequent Turkey HSD test showed that, in both datasets, the variants that showed significant differences were /s-/ and /s/ (p-value = 0.0145395 for n = 37, p-value = 0.0160335 for n = 33) as well as /s~/ and /s/ (p-value = 0.0006673 for n = 37, p-value = 0.0013824 for n = 33). The hypothesis was rejected.

9. Gay Lisp (Yes/No) x masculine/feminine

The Mann-Whitney U test showed that there was no significant correlation between participants' explicit knowledge regarding the Gay Lisp phenomenon and their rating of the speaker on the scale "masculine/feminine" in either dataset (p-value = 0.1244 for n = 37, p-value = 0.3354 for n= 33). Although participants with explicit knowledge rated the speaker as more *feminine* than those without explicit knowledge, this trend was not significant. The hypothesis was rejected.

10. Gay Lisp (Yes/No) x heterosexual/homosexual

The Mann-Whitney U test showed that there was a significant correlation between participants' explicit knowledge regarding the Gay Lisp phenomenon and their rating of the speaker on the scale *heterosexual/homosexual* in either dataset (p-value = 0.002203 for n = 37, p-value = 0.0003841 for n= 33). Participants with explicit knowledge rated the speaker as significantly more *homosexual* than those without explicit knowledge. The hypothesis was accepted.

11. Age (<40/>40) x masculine/feminine

The Mann-Whitney U test showed that there was a significant correlation between participants' age *(<40/ >40)* and their rating of the speaker on the scale *masculine/feminine* in the original dataset but not in the adjusted one (p-value = 0.01054 for n = 37, p-value = 0.06086 for n= 33). The younger group (*<40*) rated the speaker as more *masculine* than the older group (*>40*) in both datasets, but the trend was only

significant in the original one. The hypothesis was accepted for the original dataset and rejected for the adjusted one.

12. Age (<40/>40) x heterosexual/homosexual

The Mann-Whitney U test showed that there was no significant correlation between participants' age *(<40/>40*) and their rating of the speaker on the scale *heterosexual/homosexual* in either dataset (p-value = 0.0733 for either). Although the younger group *(<40)* rated the speaker as more *homosexual* than the older group (*>40*), this trend was not significant. The hypothesis was rejected.

13. Rural/Urban (R/U) x masculine/feminine

The Mann-Whitney U test showed that there was a significant correlation between participants' rurality/urbanity and their rating of the speaker on the scale *masculine/feminine* in either dataset (p-value = 0.01303 for n = 37, p-value = 0.1276 for n = 33). The rural group rated the speaker as significantly more "masculine" than the urban group. The hypothesis was accepted.

14. Rural/Urban (R/U) x heterosexual/homosexual

The Mann-Whitney U test showed that there was a significant correlation between participants' rurality/urbanity and their rating of the speaker on the scale *heterosexual/homosexual* in either dataset (p-value = 0.006285 for n = 37, p-value = 0.006285 for n = 33). The urban group rated the speaker as significantly more *homosexual* than the rural group. The hypothesis was accepted.

15. Gender (NM/M) x masculine/feminine

The Mann-Whitney U test showed that there was a significant correlation between participants' Gender (male/non-male) and their rating of the speaker on the scale *masculine/feminine* in either dataset (p-value = 0.002505 for n = 37, p-value = 0.04664 for

n = 33). The non-male group rated the speaker as significantly more *masculine* than the male group. The hypothesis was accepted.

16. Gender (NM/M) x heterosexual/homosexual

The Mann-Whitney U test showed that there was a significant correlation between participants' Gender (*male/non-male*) and their rating of the speaker on the scale *heterosexual/homosexual* in either dataset (p-value = 0.008817 for either). The male group rated the speaker as significantly more *homosexual* than the non-male group. The hypothesis was accepted.

These results can be summarized as follows: significant group effects for /s/-variant could be observed for all but three dependent variables. No significant effects could be observed for the variables *masculine/feminine, heterosexual/homosexual* and *lazy/hardworking.* This means that hypotheses for the two target scales were rejected.

The guises which were judged to sound significantly different from each other for the dependent variable *(im)polite* were /s-/ and /s/ (n = 33) as well as /s~/ and /s/ (both datasets). The same judgements applied in both datasets for the scales (*not) credible*" and *(un)pleasant.* For the variable (*un)educated* the guises /s+/ and /s/ (n = 37, p-value almost significant with 0.06 in n = 33) as well as /s~/ and /s/ (both datasets) were rated as significantly different from each other. The variable *(un)natural* showed significant differences between the guises /s+/ and /s~/, /s-/ and /s/ as well as /s~/ and /s/ in both datasets.

Furthermore, significant effects were found for the dependent variable *heterosexual/homosexual* in the three independent variables *Gay Lisp, Rural/Urban* and *Gender*. Listeners judged participants significantly higher on this scale (i.e. as more *homosexual*) if they were male, urban and had explicit knowledge of the 'gay lisp' phenomenon.

What's more, the independent variables *Age, Gender* and *Rural/Urban* also showed significant effects for the dependent variable *masculine/feminine..* Younger, non-male

and rural listeners judged the speaker lower on this scale (i.e. more *masculine)* throughout.

6. Conclusion

From the outset of the present study, its goal was to contribute in a meaningful way to the currently ever-present discourse on transness and, with that, acknowledgement of trans people's struggles as well as protection of their rights. What I hope to demonstrate here is that, on a linguistic level, investigations of trans voices often fail to take into account the intertwined and often interdependent nature of metapragmatic categories and their material instantiations. This false dichotomy positing a fundamental separation of discourse and materiality is called into question here in order to understand how semiosis takes place in such a complex field as sociophonetics encompassing sex as well as gender.

The present research question which focusses on one specific issue of the above field is if (and consequently *how*) transmasculine German-speakers employ /s/-variation as an effective sociophonetic resource in discursive co-construction of gender. This question was conceived through reviewing the available literature on the phenomenon of gendered /s/-variation (also sometimes referred to as the 'gay lisp') and discovering that, while Zimman (2017a) finds extremely varied use of /s/ by transmasculine US English-speakers, no corresponding data had been collected for German. Other languages and varieties, such as Mandarin (Li 2017), Danish (Pharao et al. 2017), Glaswegian (Stuart-Smith 2007) and British English (Holmes-Elliott & Levon 2017) show evidence of gendered /s/-variation and, therefore, the question arose if the phenomenon might extend to German as well.

Further literature review led to the findings by Boyd et al. (2021), which show that, although German-speaking cis gay men use fronted /s/ in discourse, the feature showed no uptake among German-speaking listeners at the time of the study. The authors, hence, speculate that fronted /s/ may be waiting for its baptismal moment for enregisterment in German, not yet realizing this specific indexical potential in a broader societal context. Considering both Zimman's (2017a) and Boyd et al.'s (2021) findings together sparked the idea of a potential change in progress concerning the

indexicality of /s/ in German. In order to test this, a two-pronged study was devised, taking into account production as well as perception of /s/-variation. Transmasculine individuals were interviewed in order to compare results to those of Zimman and a perception survey geared at comparing results to those of Boyd et al. was created. The hypothesis was that on both levels – production and perception – /s/-variability would be used to index gendered identities.

Results did not ubiquitously present themselves as expected but showed some significant tendencies and interesting articulation patterns. Regarding production, a pattern of /s/-variability similar to that in Zimman (2017a) emerged from map-task data. Participants in this study produced /s/-tokens with CoG values situated between 2,366 Hz and 10,755 Hz, thereby encompassing the entire spectrum that is typically reported for male as well as female speakers, just like Zimman reports for his English-speaking participants. The range of means for German, though generally set slightly lower, is also quite similar to that of English: while the English-speakers exhibit a range between 5,226 and 9,116 Hz (a difference of 3,890 Hz), the German-speakers are situated between 4,503 Hz and 7,577 Hz (a difference of 3,754 Hz).

In a qualitative analysis of this data, I employ the framework of hegemonic masculinity by Connell (1995) to show that the speakers may use their varied articulations of /s/ to avoid/accomplish discursively positioning themselves as embodying hegemonic masculinity. I argue that this process relies not simply on a unitary concept of masculinity but instead split masculinity into the aspects of *passing*, *gender expression* and *sexual orientation*. What's more, I suggest that it is not only the wish to be perceived in a hegemonic vs. non-hegemonic way in regard to these aspects but also the emphasis that is placed on this wish and to what degree the desired identity is already perceived to be achieved that play into articulatory practice. In this framework, the potential grounds for the linguistic choices leading to the evident pattern of /s/-variability seem more palpable: they do not correspond in a binary way to a monolithic identity or goal but are rather contextually and individually fluid.

While different participants voiced different concerns about being seen as cis vs. trans or straight vs. queer, these concerns were explicitly mentioned as being sensitive to their personal history: while the participant who had been undergoing testosterone therapy for the longest time stated that he used to want to pass as straight in order to seem more like a typical man, this concern seemed to have been mitigated over time by the fact that he now passed as male more reliably and was rarely to never misgendered. In short, the trade-off of discursively constructing a queer persona and thereby activating an indexical field that might at the same time activate associations of transness and non-normative masculinity did not seem threatening to him anymore. The likelihood of his non-normatively masculine use of /s/ being interpreted as indexing transness (which he explicitly wanted to avoid) had become so low that his desire to index a queer persona in discourse could be realized now without threatening his desire to be perceived in a hegemonic way regarding his passing.

Similarly, a participant who did state that they wanted to be perceived as male when it came to passing simply did not place quite as much importance on this and instead emphasized how much they wanted to avoid being seen as straight and talked at length about their non-binary identity and said they had chosen their name with regard to how easily it may be interpreted in a non-binary way. This participant also showed CoG-values right at the center of the scale and almost evenly distributed towards either end of the spectrum. It is, therefore, not the case that /s/-variation can be mapped onto one distinct, unchangeable and impermeable gender identity but should instead be understood to span an indexical field with the help of which specific aspects of masculinity can be transported.

Another important finding of the production part suggests that there is a significant effect of age at the beginning of testosterone therapy on CoG. Namely, the lower the age, the lower the CoG. This can be interpreted in different ways: a biologically determinist perspective might suggest that the younger a person is when they begin hormonally transitioning, the more malleable their physiology still is and the more

prone to the effects of testosterone they are. Hence, their palate might still grow, producing a lower CoG /s/. But while Fuchs & Toda (2010) do not find any palate size difference between men and women, I also want to suggest that this avenue of explanation is simply not fruitful in light of the above findings.

As Foulkes (2010) and Zimman (2018) suggest, physiological actualities rather restrain than determine the use of an individual's voice. It seems more likely, then, that the individuals who started hormonally transitioning at a younger age (17-18) in this study also were the four individuals wanting to be perceived as hegemonically male in terms of passing, resulting in a statistically lower CoG. This theory is also supported by the data. The question of correlation between the two factors cannot be conclusively answered here and only be speculated upon and is hence not undertaken in the present thesis.

Another piece of statistically significant production data emerged that should briefly be mentioned here: the effect of nationality on CoG. An actual effect seems unlikely, however significant the result may be, and I would suggest a statistical fallacy here. Both Germany and Austria encompass a wide range of varieties which are not easily distinguished by national boundaries. It seems, therefore, not plausible that a sample of n = 6 should be representative of whole countries and their complex constellations concerning varieties and realizations of the phoneme /s/. This statistical fallacy is therefore treated as irrelevant to the research question. Taken together, however, the above findings create strong evidence that /s/-variability is indeed used by transmasculine German-speakers to discursively co-construct their gender identity as hypothesized.

Concerning perception, the results are less straightforward. For one, the present data do not suggest that the /s/-variant listeners hear (/s+/, /s~, /s-/ or /s/) generally has a significant effect on ratings of masculinity/femininity or sexual orientation. What did, however, emerge were significant group effects on these scales. The group which had heard of the phenomenon of the "gay lisp" rated all four samples as significantly more

homosexual than the group who had not (p-value = 0.002203), the group under 40 rated all four samples as significantly more masculine than the group over forty (p-value = 0.01054), the rural group rated the samples as more masculine (p-value = 0.01303) and more heterosexual (p-value = 0.006285) than the urban group and the male group generally rated the speakers as more feminine (p-value = 0.002505) and more homosexual (p-value = 0.008817) than the non-male group.

The tendency, then, seems to be that younger, more rural and non-male individuals rate speakers as more masculine than their counterparts while those with knowledge of the 'gay lisp', urban and male speakers tend to rate speakers as more homosexual. Possible explanations for these findings are, of course, post-hoc speculations but still seem needed here. Concerning the ratings of sexual orientation, I would argue that it seems probable that the three categories show some overlap. Those who are familiar with the phenomenon of the 'gay lisp' might be more likely to be gay men themselves. What's more, as the queer community does tend to flourish in more urban centers, exposure to metapragmatic discourse on queer language might also circulate more prevalently in urban circles, possibly amongst gay men themselves. In an attempt not to ask participants for too much personal information, the question of their own sexual orientation was omitted. Retrospectively, however, this information could have contributed to a more complete picture of the perception data and should have been posed.

Concerning masculinity ratings, the tendency of rural individuals to rate samples as more masculine was not surprising considering their tendency to also rate them as more heterosexual. It seems, then, that rural speakers tended to rate all four variants as more hegemonically masculine. This could either suggest that they have a higher window of tolerance for what linguistically constitutes masculinity and heterosexuality or that they heard fewer differences between the variants, creating a clearer statistical tendency with less intra-group deviation. The same, of course, applies to non-males and younger participants. However, taking into account that,

unlike the other two groups, the younger age group only predicted a higher masculinity rating but not higher heterosexuality rating, it could be argued that the younger group might actually be more sensitive to differences as they do not rate all samples as equally hegemonically masculine in all respects.

What's more, the listeners did assign significant differences to the variants they heard, albeit on the control scales of (im)polite, (un)educated, (un)natural, (not) credibly and (un)pleasant. The variants which were most often rated as significantly different from each other were /s-/ vs /s/ and /s~/ vs. /s/. The variants which were never rated as significantly different from each other were /s+/ vs. /s-/ and /s-/ vs. /s~/. These findings suggest that it is possible that, though differences can be heard in general, not all samples were equally distinguishable for listeners. This issue might then be at the core of the lack of evidence for perception differences on the scales targeting gender and sexual orientation.

Furthermore, the scale of unnatural/natural showed three variant pairings that were rated as significantly different from each other: /s+/ vs. /s~/, /s-/ vs. /s/ and /s~/ vs. /s/. This finding suggests that some of the samples might sound too manipulated which might overshadow potential indexicality on other scales. This would also be in line with group effects that span all four samples being observed for the target scales but no effects of variant on these same scales.

Taking all of these findings on the production and perception of /s/-variation by transmasculine German-speakers together, it seems plausible that this variation is produced by these speakers in order to index different dimensions of their gendered identity while the uptake of this linguistic feature might still be limited to certain groups. While we can see significant tendencies here, showing that linguistic perception of sexual orientation generally differs depending on explicit metalinguistic knowledge, the group of participants who had heard of the 'gay lisp' was too small (n = 15) to reliably test for significant effects of /s/-variant on perception of sexual

orientation and/or gender in this subgroup. Hence, we do not know if this group might perceive the four variants as indexing gender and sexual orientation differently.

This is, however, an avenue for future research in this direction. It is very possible, given the present group differences in perception of sexual orientation and gender, that metalinguistic knowledge, age, urbanity, gender and possibly other factors could have an effect on the perception of the indexicality of /s/-variants. Therefore, I would argue that, because the present sample is too small, future research should concentrate on acquiring a bigger sample in order to investigate group differences in the perception of indexicality of /s/-variants.

What the present study contributes is a basic understanding of /s/-variation in transmasculine German-speakers. It finds that, while this variation seems to be produced, the question of indexical uptake is neither ultimately denied nor supported. It seems likely that the quality of the samples may have imposed limitations on how much difference could be perceived between the variants and that a bigger sample is needed in order to make reliable statements about the indexicality of /s/-variation and group differences thereof.

In conclusion, I hope to have shown that gendered /s/-variation is present in German just as it is in many other languages. It may be used contextually to convey different aspects and dimensions of gendered identity to certain groups and is possibly connected to a change in progress concerning the indexicality of /s/-variation. This claim is supported by a significant tendency of those with explicit metapragmatic knowledge of the phenomenon to rate speakers as more homosexual. This tendency could possibly point to an ongoing process of enregisterment of /s/-variation as indexing different dimensions of masculinity to an ingroup of urban men. Future research should concentrate on creating more reliably natural-sounding samples, accumulating a bigger and more gender diverse participant sample as well as asking more detailed questions about participants own sexual orientation and involvement in the queer community.

7. References

Ammon, Bickel, Ebner, Esterhammer, Gasser, Hofer, Kellermeier-Rehbein, Kyvelos, Löffler, Mangott, Moser, Nyffenegger, Oehler, Schloßmacher, Schläpfer, Schmidlin, Vallaster, Bickel, Hans, Ebner, Jakob, … Vallaster, Günter. 2011. *Variantenwörterbuch des Deutschen: Die Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Südtirol*. De Gruyter.

Agha, Asif. 2005. Voice, footing, enregisterment. *Journal of Linguistic Anthropology* 15(1), 38–59. https://doi.org/10.1525/jlin.2005.15.1.38.doi.org+6.

Avery, Jack D. & Julie M. Liss. 1996. Acoustic characteristics of less-masculine-sounding male speech. Journal of the Acoustical Society of America 99(6). 3738–3748.

Bakhtin, Mikhail M. 1981. Discourse in the novel. In Michael Holquist (ed.), *The dialogic imagination: Four essays*, trans. Caryl Emerson & Michael Holquist, 259–422. Austin: University of Texas Press.

Benjamin, Harry. 1966. *The transsexual phenomenon*. Ann Arbor, MI: Julian Press.

Besnier, Niko. 2003. Crossing genders, mixing languages: The linguistic construction of transgenderism in Tonga. In Janet Holmes & Miriam Meyerhoff (eds.) *The handbook of language and gender*, 279–301. Malden, MA: Blackwell.

Bezooijen, Renée van. 1995. Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech* 38(3), 253–265. https://doi.org/10.1177/002383099503800303.repository.ubn.ru.nl+3

Boswell, Holly. 1991. The transgender alternative. *Crysalis Quarterly* 1(2), 29–31.

Boyd, Zac, Josef Fruehwald & Lauren Hall-Lew. 2021. Crosslinguistic perceptions of /s/ among English, French, and German listeners. *Language Variation and Change 33*(2), 165–191. https://doi.org/10.1017/S0954394521000089

Briggs, Charles L. 1986. *Learning How to Ask: A Sociolinguistic Appraisal of the Role of the Interview in Social Science Research.* Cambridge: Cambridge University Press.

Briggs, Charles L. 2007. Anthropology, Interviewing, and Communicability in Contemporary Society. *Current Anthropology 48*(4), 551–580.

Bucholtz, Mary. 1999. You da man: Narrating the racial other in the production of white masculinity. *Journal of Sociolinguistics* 3(4), 443–460.

Bucholtz, Mary & Kira Hall (eds.). 1995. *Gender articulated: Language and the socially constructed self*. New York: Routledge.

Bucholtz, Mary & Kira Hall. 2005. Identity and Interaction: A Sociocultural Linguistic Approach. *Discourse Studies* 7(4/5), 585–614. https://doi.org/10.1177/1461445605054407.

Bucholtz, Mary & Kira Hall. 2016. Embodied sociolinguistics. *Sociolinguistics*, 173–198. Cambridge University Press.

Busby, Peter A. & Geoff L. Plant. 1995. Formant frequency values of vowels produced by preadolescent boys and girls. *Journal of the Acoustical Society of America* 97(4), 2603–7.

Butler, Judith. 2007. *Gender Trouble: Feminism and the Subversion of Identity.* 4th edition. London New York: Routledge.

Butler, Judith. 2011. *Bodies That Matter: On the Discursive Limits of "Sex".* 2nd edition. London New York: Routledge.

Byrd, Dani. 1992. Sex, dialects, and reduction. *Proceedings of the 2nd International Conference on Spoken Language Processing*, 827–830. Banff: ISCA.

Calder, J. 2024. Towards a queer and trans sociophonetics : Connecting sociophonetics to a queer and trans linguistics. In Tyler Everett Kibbey (ed.) *Linguistics Out of the Closet : The Interdisciplinarity of Gender and Sexuality in Language Science*, 117-137. First edition. Berlin, Germany: Walter de Gruyter GmbH.

Cameron, Deborah. 1997. Performing gender identity: Young men's talk and the construction of heterosexual masculinity. In Sally Johnson & Ulrike Hanna Meinhof (eds.) *Language and masculinity,* 47–64. Oxford: Blackwell.

Cameron, Deborah. 2005. Language, gender, and sexuality: Current issues and new directions. *Applied Linguistics* 26(4), 482–502.

Cameron, Deborah. 2007. The myth of Mars and Venus: Do men and women really speak different languages? Oxford: Oxford University Press.

Campbell-Kibler, Kathryn. 2011. Intersecting variables and perceived sexual orientation in men. *American Speech* 86(1), 52–68.

Connell, R.W. 1995. *Masculinities*. Cambridge: Polity.

Connell, R.W. & James W. Messerschmidt. 2005. Hegemonic masculinity: Rethinking the concept. *Gender & society* 19(6), 829-859.

Connell, Raewyn & Rebecca Pearse. 2014. *Gender: In World Perspective*. 3rd edition. Cambridge: Polity Press.

Copland, Fiona. 2019. Ethics. In Karin Tusting (ed.), *The Routledge Handbook of Linguistic Ethnography*, 184-197. New York: Routledge.

Creswell, John W. & Cheryl N. Poth. 2024. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. 5th edition. Thousand Oaks London New Delhi: SAGE Publications Ltd.

Cromwell, Jason. 1995. Talking about without talking about: The use of protective language among transvestites and transsexuals. In William L. Leap (ed.) *Beyond the*

lavender lexicon: Authenticity, imagination, and appropriation in lesbian and gay languages, 267– 295. Amsterdam: Gordon and Breach.

Davidson, Lisa. 2020. The versatility of creaky phonation: Segmental, prosodic, and sociolinguistic uses in the world's languages. *Annual Review of Linguistics* 6, 89–107.

De Fina, Anna. 2019. The ethnographic interview. In Karin Tusting (ed.), *The Routledge Handbook of Linguistic Ethnography*, 154-167. New York: Routledge.

DiCanio, Christian. 2013. Haskins Laboratories & SUNY Buffalo.

Diehl, Randy L., Björn Lindblom & Keith R. Kluender. 1996. Speaker normalization: A critical review. *Speech Communication* 20(1), 1–35. https://doi.org/10.1016/0167-6393(96)00010-2.

Eckert, Penelope. 1989. *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.

Eckert, Penelope. 2000. *Linguistic variation as social practice*. Malden, MA: Blackwell.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4), 453–476.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41. 87–100.

Eckert, Penelope. 2017. Comment: The most perfect of signs: Iconicity in variation. *Linguistics* 55(5), 1197–1207.

Eckert, Penelope & Sally McConnell-Ginet. 2013. *Language and Gender*. 2nd edition. Cambridge: Cambridge University Press.

Edelman, Elijah Adiv & Lal Zimman. 2014. Boycunts and bonus holes: Trans men's bodies, neoliberalism, and the sexual productivity of genitals. *Journal of homosexuality* 61(5), 673-690.

Ehrlich, Susan, Miriam Meyerhoff & Janet Holmes. 2014. *The handbook of language, gender, and sexuality*. Second edition. Malden, MA: Blackwell.

Feinberg, Leslie. 2006. Transgender liberation: A movement whose time has come. In Susan Stryker & Stephen Whittle (eds.) *The transgender studies reader*, 205–220. New York: Routledge.

Fishman, Pamela M. 1983. Interaction: The work women do. In Barrie Thorne, Cheris Kramarae & Nancy Henley (eds.) *Language, gender and society*, 89–101. Rowley, MA: Newbury House.

Fitch, W. Tecumseh & Jay Giedd. 1999. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America* 106(3), 1511–1522. https://doi.org/10.1121/1.427148.

Flipsen, Peter, Jr., Lawrence Shriberg, Gary Weismar, Heather Karlsson & Jane McSweeny. 1999. Acoustic characteristics of /s/ in adolescents. *Journal of Speech, Language and Hearing Research* 42, 663–677.

Forrest, Karen, Gary Weismer, Paul Milenkovic & Ronald N. Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: preliminary data. *The Journal of the Acoustical Society of America.* 84, 115–123. https://doi.org.uaccess.univie.ac.at/10.1121/1.396977.

Foulkes, Paul. 2010. Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology* 1(1), 5–39. https://doi.org/10.1515/labphon.2010.003.

Fuchs, Susanne & Martine Toda. 2010. Do differences in male versus female /s/ reflect biological or sociophonetic factors? In Susanne Fuchs, Martine Toda & Marzena Żygis (eds.), *Turbulent sounds: An interdisciplinary guide*, 281–302. Berlin: De Gruyter Mouton.

Funk, Riccarda & Adrian P. Simpson. 2023. The Acoustic and Perceptual Correlates of Gender in Children's Voices. *Journal of Speech, Language, and Hearing Research* 66(9), 3346-3363.

Gal, Susan & Judith T. Irvine. 2019. *Signs of Difference: Language and Ideology in Social Life*. First edition. Cambridge: Cambridge University Press.

Garellek, Marc. 2019. The phonetics of voice. In William F. Katz & Peter F. Assmann (eds.), *The Routledge handbook of phonetics*, 75–106. New York: Routledge. https://doi.org/10.4324/9780429056253-5.

Gaudio, Rudolf Pell. 2009. *Allah made us: Sexual outlaws in an Islamic African city*. Malden, MA: Blackwell.

Hall, Kira. 1997. 'Go suck your husband's sugarcane!': Hijras and the use of sexual insult. In Anna Livia & Kira Hall (eds.) *Queerly phrased: Language, gender, and sexuality*, 430–460. New York: Oxford University Press.

Hall, Kira. 2014. Exceptional speakers: Contested and problematized gender identities. In Susan Ehrlich, Miriam Meyerhoff, and Janet Holmes (eds.) *The handbook of language, gender, and sexuality*, 48–67. Second edition. Malden, MA: Blackwell.

Hall, Kira & Veronica O'Donovan. 1996. Shifting gender positions among Hindi-speaking hijras. In Veronica L. Bergvall, Janet M. Bing, and Alice F. Freed (eds.) *Rethinking language and gender research: Theory and practice*, 228–266. London: Longman.

Hall, Kira & Rusty Barrett (eds.). 2018. *The Oxford Handbook of Language and Sexuality* (online edn., Oxford Academic), https://doi.org/10.1093/oxfordhb/9780190212926.001.0001, accessed 9 April 2025.

Harkness, Nicholas. 2014. *Songs of Seoul: An ethnography of voice and voicing in Christian South Korea*. Berkeley: University of California Press.

Harkness, Nicholas. 2017. The open throat: Deceptive sounds, facts of firstness, and the interactional emergence of voice. *Signs and Society* 5(S1), 21–52. https://doi.org/10.1086/690041.

Harrington, Jonathan, Janet Fletcher & Mary E. Beckman. 2000. Manner and place conflicts in the articulation of accent in Australian English. In M. B. Broe & J. B. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, 40–51. Cambridge: Cambridge University Press.

Hebdige, Dick. 1984. *Subculture: The Meaning of Style*. New York: Methuen.

Heffernan, Kevin. 2004. Evidence from HNR that /s/ is a social marker of gender. *Toronto Working Papers in Linguistics* 23(2), 71-84.

Heffernan, Kevin. 2010. Mumbling is macho: Phonetic distinctiveness in the speech of American radio DJs. *American Speech* 85(1), 67–90. https://doi.org/10.1215/00031283-2010-003.

Henton, Caroline. 1995. Cross-language variation in the vowels of female and male speakers. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, vol. 4, 420–423. Stockholm: ICPhS.

Hillenbrand, James, Linda A. Getty, Michael J. Clark & Kimberly Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5), 3099–3111. https://doi.org/10.1121/1.411872.

Holmes-Elliot, Sophie & Erez Levon. 2014. East End Boys and West End Girls:/s/-Fronting in Southeast England. *University of Pennsylvania Working Papers in Linguistics* 19(2), 110-120.

Holmes-Elliott, Sophie & Erez Levon. 2017. The substance of style: Gender, social class and interactional stance in /s/-fronting in southeast England. *Linguistics*, 55(5), 1045–1072.

Howell, Jayne. 2004. Turning out good ethnography, or talking out of turn? Gender, violence, and confidentiality in Southeastern Mexico. Journal of Contemporary Ethnography, 33(3), 323–352. https://doi.org/10.1177/0891241604263588

Hurring, Gia, Jennifer Hay, Katie Drager, Ryan Podlubny, Laura Manhire & Alix Ellis. 2022. Social priming in speech perception: Revisiting kangaroo/kiwi priming in New Zealand English. *Brain Sciences* 12(6), 684. https://doi.org/10.3390/brainsci12060684

Irvine, Judith T. & Susan Gal. 2000. Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of Language: Ideologies, Polities, and Identities*, 35–84. Santa Fe, NM: School of American Research Press.

Jackson, Peter A. 2004. Gay adaption, tom-dee resistance, and kathoey indifference: Thailand's gender/sex minorities and the episodic allure of queer English. In William L. Leap & Tom Boellstorff (eds.) *Speaking in queer tongues: Globalization and gay language*, 202–230. Urbana: University of Illinois Press.

Johnson, Keith, Elizabeth A.Strand & Mariapaola D'Imperio. 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*(4), 359–384. https://doi.org/10.1006/jpho.1999.0100

Jong, Kirsten J. de. 1995. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America* 97, 491–504. https://doi.org/10.1121/1.413007.

Jongman, Allard, Ratree Wayland & Serena Wong**.** 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108(3),1252–1263. https://doi.org/10.1121/1.1288413.

Kachel, Sven, Melanie C. Steffens, Sabine Preuß & Adrian P. Simpson. 2020. Gender (Conformity) Matters: Cross-Dimensional and Cross-Modal Associations in Sexual Orientation Perception. *Journal of Language and Social Psychology* 39(1), 40–66. https://doi.org/10.1177/0261927X19883902.

Kibbey, Tyler Everett. 2024. *Linguistics Out of the Closet : The Interdisciplinarity of Gender and Sexuality in Language Science*. First edition. Berlin, Germany: Walter de Gruyter GmbH.

Knight, H. Merle. 1992. Gender interference in transsexuals' speech. In Kira Hall, Mary Bucholtz & Birch Moonwomon (eds.), *Locating power: Proceedings of the 2nd Berkeley Women and Language Conference*, 312–317. Berkeley, CA: Berkeley Women and Language Group.

Kreiman, Jody & Diana Sidtis. 2011. *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. Malden, MA: Wiley-Blackwell. https://doi.org/10.1002/9781444395068.

Kuhl, Patricia K. 1979. Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America* 66(6), 1668–1679. https://doi.org/10.1121/1.383639.

Kuhl, Patricia K. 1987. Perception of speech and sound in early infancy. In P. Salapatek & L. Cohen (eds.), *Handbook of Infant Perception: From Perception to Cognition*, vol. 2, 275–382. New York: Academic Press.

Kulick, Don. 1998. *Travestí: Sex, gender, and culture among Brazilian transgendered prostitutes*. Chicago: University of Chicago Press.

Kulick, Don. 2019. Foreword. *International Journal of the Sociology of Language* 256, 1– 3.

Ladefoged, Peter & Keith Johnson. 2011. *A course in phonetics.* 6th edition. Boston, MA: Wadsworth.

Lakoff, Robin. 1975. *Language and woman's place*. New York: Harper & Row.

Laver, John. 1980. *The phonetic description of voice quality*. Cambridge: Cambridge University Press.

Lee, Sungbok, Alexandros Potamianos & Shrikanth Narayanan. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America* 105(3), 1455-1468.

Leeuwen, Theo van. 2017. Parametric systems: The case of voice quality. In Carey Jewitt (ed.), *The Routledge handbook of multimodal analysis*, 68–77. London New York: Routledge.

Levon, Erez. 2007. Sexuality in Context: Variation and the Sociolinguistic Perception of Identity. *Language in Society* 36(4), 533–554. https://doi.org/10.1017/S0047404507070431

Levon, Erez. 2014. Categories, Stereotypes, and the Linguistic Perception of Sexuality. *Language in Society* 43(5), 539–566.

Levon, Erez, Marie Maegaard & Nicolai Pharao (eds.). 2017. Special issue: The sociophonetics of /s/. *Linguistics* 55(5).

Li, Fangfang. 2017. The development of gender-specific patterns in the production of voiceless sibilant fricatives in Mandarin Chinese. *Linguistics*, 55(5), 1021–1044.

Livia, Anna. 1997. Disloyal to masculinity: Linguistic gender and liminal identity in French. In Anna Livia & Kira Hall (eds.) *Queerly phrased: Language, gender, and sexuality*, 349–368. New York: Oxford University Press.

Livia, Anna. 2001. *Pronoun envy: Literary uses of linguistic gender*. New York: Oxford University Press.

Livia, Anna & Kira Hall (eds.). 1997. *Queerly phrased: Language, gender, and sexuality*. New York: Oxford University Press.

Lorber, Judith. 1994. *Paradoxes of gender*. New Haven, CT: Yale University Press.

Manalansan, Martin F., IV. 1995. Performing Filipino gay experiences in American: Linguistic strategies in a transnational context. In William L. Leap *(ed.) Beyond the lavender lexicon: Authenticity, imagination, and appropriation in lesbian and gay languages*, 249– 266. Philadelphia: Gordon and Breach.

McGurk, Harry & John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264(5588), 746–748. https://doi.org/10.1038/264746a0.

Mendoza-Denton, Norma. 2007. Homegirls remember: Discourse and literacy practices among U.S. Latina gang girls. In P. Hodkinson & W. Deicke (eds.), *Youth Cultures: Scenes, Subcultures and Tribes*, 129–144. London: Routledge.

Mennen, Ineke, Frank Schaeffler & Sven Kranich. 2012. Cross-language differences in fundamental frequency range: A comparison of English and German. *Journal of the Acoustical Society of America* 131(3), 2249–2259. https://doi.org/10.1121/1.3685650.

Mishler, Elliot G. 1991. *Research interviewing: Context and narrative*. Cambridge: Harvard University Press.

Munson, Benjamin. 2007. The Acoustic Correlates of Perceived Masculinity, Perceived Femininity, and Perceived Sexual Orientation. *Language and Speech* 50(1), 125–42. https://doi.org/10.1177/00238309070500010601.

Munson, Benjamin, Elizabeth C McDonald, Nancy L DeBoe & Aubrey R White. 2006a. The Acoustic and Perceptual Bases of Judgments of Women and Men's Sexual Orientation from Read Speech. *Journal of Phonetics* 34(2) 202–40. https://doi.org/10.1016/j.wocn.2005.05.003.

Munson, Benjamin, Elise McDonald, Megan DeBoe & Kelsey White. 2006b. The influence of /s/ quality on ratings of men's sexual orientation. *Journal of Phonetics* 34(1), 20–33. https://doi.org/10.1016/j.wocn.2005.05.001.

Munson, Benjamin, Kayleigh Ryherd & Sara Kemper. 2017. Implicit and explicit gender priming in English lingual sibilant fricative perception. *Linguistics* 55(5), 1073–1108. https://doi.org/10.1515/ling-2017-0021.

Munson, Benjamin, Natasha Lackas & Kiana Koeppe. 2022. Individual differences in the development of gendered speech in preschool children: Evidence from a longitudinal study. *Journal of Speech, Language, and Hearing Research* 65(4), 1311-1330.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1), 62–85. https://doi.org/10.1177/0261927X99018001005.

Nittrouer, Susan. 1995. Children learn separate aspects of speech production at different rates:Evidence from spectral moments. *Journal of the Acoustical Society of America* 97(1), 520–530.

Nittrouer, Susan, Michael Studdert-Kennedy & Richard. S. McGowan. 1989. The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research* 32(1), 120–132.

Ochs, Elinor. 1992. Indexing gender. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context: Language as an interactive phenomenon*, 335–358. Cambridge: Cambridge University Press.

Ochs, Elinor & Bambi B. Schieffelin. 1984. Language acquisition and socialization: Three developmental stories. In Richard A. Shweder & Robert A. LeVine (eds.) *Culture theory: Essays on mind, self and emotion*, 276–320. Cambridge: Cambridge University Press.

Ochs, Elinor & Bambi B. Schieffelin. 2011. The theory of language socialization. In Alessandro Duranti, Elinor Ochs & Bambi B. Schieffelin (eds.) *The handbook of language socialization*, 1–21. Malden, MA: Wiley-Blackwell.

Ohala, John J. 1994. The frequency code underlies the sound-symbolic use of voice pitch. In Leanne Hinton, Johanna Nichols, and John J. Ohala (ed.), *Sound Symbolism*, 325–347. Cambridge: Cambridge University Press.

Österreichische Akademie der Wissenschaften. 1983. *Bayerisch-österreichisches Wörterbuch.* Oldenbourg: Verlag der Österreichischen Akademie der Wissenschaften.

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(23), 82–93.

Peirce, Charles Sanders. 1931–1958. *Collected Papers of Charles Sanders Peirce*. Volumes 1–8. Charles Hartshorne, Paul Weiss (Vols. 1–6) and Arthur W. Burks (Vols. 7–8) (eds.). Cambridge, MA: Harvard University Press.

Perry, Theodore L., Ralph N. Ohde & Daniel H. Ashmead. 2001. The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America* 109(6), 2988-2998.

Peterson, Gordon E. & Barney, Harold L. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24(2), 175–184. https://doi.org/10.1121/1.1906875.

Pharao, Nicolai & Marie Maegaard. 2017. On the influence of coronal sibilants and stops on the perception of social meanings in Copenhagen Danish. *Linguistics*, 55(5), 1141–1167.

Pharao, Nicolai, Marie Maegaard & Janus Spindler Møller. 2014. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society*, 43(1), 1-31.

Pierrehumbert, Janet B., Tessa Bent, Benjamin Munson, Ann R. Bradlow & J. Michael Bailey. 2004. The influence of sexual orientation on vowel production. *Journal of the Acoustical Society of America* 116(4), 1905–1908. https://doi.org/10.1121/1.1788729.

Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics*, 11(4), 478-504.

Podesva, Robert J. 2013. Gender and the social meaning of non-modal phonation types. *Proceedings of the annual meeting of the Berkeley Linguistics Society, 37(1), 427-448.*

Podesva, Robert J. & Patrick Callier. 2015. Voice quality and identity. *Annual Review of Applied Linguistics* 35, 173–194. https://doi.org/10.1017/S0267190514000270.

Ritchart, Amanda, & Amalia Arvaniti. 2014. The form and use of uptalk in southern Californian English. *Proceedings of the International Conference on Speech Prosody*, 331–336.

Rogers, Henry & Ron Smyth. 2003. Phonetic differences between gay- and straight-sounding male speakers of North American English. In *Proceedings of the 15th International Congress of Phonetic Sciences*, 1855–1858. Barcelona: ICPhS.

Sapir, Edward. 1927. Speech as a personality trait. *American Journal of Sociology* 32(6), 892–905. https://doi.org/10.1086/214279.

Schieffelin, Bambi B., Kathryn Ann Woolard & Paul V. Kroskrity. 1998. *Language Ideologies : Practice and Theory*. New York Oxford: Oxford University Press.

Schwartz, Martin F. 1968. Identification of speaker sex from isolated voiceless fricatives. *Journal of the Acoustical Society of America* 43(5), 1178–1179.

Sedaris, David. 2000. *Me Talk Pretty One Day*. New York: Little, Brown and Company.

Shadle, Christine H. 1985. The acoustics of fricative consonants. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Shadle, Christine H. 1990. Articulatory-acoustic relationships in fricative consonants. In William J. Hardcastle & Alain Marchal (eds.) *Speech production and speech modelling*, 187–209. Dordrecht: Kluwer Academic.

Shadle, Christine H. 1991. The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics* 19(3–4), 409–424.

Silverstein, Michael. 1979. Language structure and linguistic ideology. In Paul R. Clyne, William F. Hanks & Carol L. Hofbauer (eds.), *The elements: A parasession on linguistic units and levels*, 193–247. Chicago: Chicago Linguistic Society.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23(3–4), 193–229. https://doi.org/10.1016/S0271-5309(03)00013-2.

Silverstein, Michael. 2005. Axes of evals: Token versus type interdiscursivity. *Journal of Linguistic Anthropology* 15(1), 6–22.

Silverstein, Michael. 2023. *Language in culture: Lectures on the social semiotics of language.* Cambridge: Cambridge University Press.

Simpson, Adrian P. 2001. Dynamic consequences of differences in male and female vocal tract dimensions. *The Journal of the Acoustical Society of America* 109(5), 2153–2164.

Simpson, Adrian P. & Caroline Ericsdotter. 2007. Sex-specific differences in f0 and vowel space. *Proceedings of the 16th International Congress of Phonetic Sciences*, 933–936. Saarbrücken: ICPhS.

Simpson, Adrian P. & Melanie Weirich. 2015.*Gender-specific differences in sibilant contrast realizations in English and German. ICPhS*.

Simpson, Adrian P., Riccarda Funk & Frederik Palmer. 2017. Perceptual and Acoustic Correlates of Gender in the Prepubertal Voice. *Interspeech*, 914-918.

Spender, Dale. 1980. *Man made language*. London: Routledge & Kegan Paul.

Spradley, James P. 2016. *Participant observation*. Long Grove: Waveland Press.

Stone, Sandy. 1992. The empire strikes back: A posttranssexual manifesto. In Julia Epstein & Kristina Straub (eds.) *Body guards: The cultural politics of gender ambiguity*, 280–304. New York: Routledge.

Stryker, Susan. 1994. My Words to Victor Frankenstein Above the Village of Chamounix: Performing Transgender Rage. In M. Jacobus, E. Fox Keller & S. Shuttleworth (eds.) *Body guards: The cultural politics of gender ambiguity*, 237–250. New York: Routledge.

Stryker, Susan & Stephen Whittle (eds.). 2006. *The transgender studies reader*. New York: Routledge.

Stuart-Smith, Jane. 2007. Empirical evidence for gendered speech production: /s/ in Glaswegian. *Laboratory Phonology: Phonology and Phonetics (9),* 65-86.

Stuart-Smith, Jane. 2020. Changing perspectives on /s/ and gender over time in Glasgow. *Linguistics Vanguard* 6(1). https://doi.org/10.1515/lingvan-2018-0064.

Szakay, Anita. 2006. Rhythm and pitch as markers of ethnicity in New Zealand English. *Proceedings of the 11th Australian International Conference on Speech Science & Technology*, 421–26. Auckland: University of Auckland.

Tannen, Deborah. 1990. *You just don't understand: Women and men in conversation*. New York: William Morrow.

Tyler, Joseph C. 2015. Expanding and Mapping the Indexical Field: Rising Pitch, the Uptalk Stereotype, and Perceptual Variation. *Journal of English Linguistics 43*(4), 284–310. https://doi.org/10.1177/0075424215607061

Vorberger, Lars. 2024. „Das klingt echt schwul "–Eine soziophonetische Untersuchung zum stereotypen schwulen Sprechen im Deutschen. *Zeitschrift für Angewandte Linguistik* 80, 136-180.

Wassink, Alicia Beckford. 2001. Theme and variation in Jamaican vowels. *Language Variation and Change* 13(2), 135–159. https://doi.org/10.1017/S0954394501132023.

Weirich, Melanie & Adrian Simpson. 2014. Differences in acoustic vowel space and the perception of speech tempo, *Journal of Phonetics* 43, 1-10.

Weirich, Melanie & Adrian P. Simpson. 2018. Gender identity is indexed and perceived in speech. *PLoS One* 13(12).

Weirich, Melanie, Susanne Fuchs, Adrian Simpson, Ralf Winkler & Pascal Perrier. 2016. Mumbling: Macho or morphology? *Journal of Speech, Language, and Hearing Research* 59(6), 1587–1595. https://doi.org/10.1044/2016_JSLHR-S-15-0040.

Weirich, Melanie, Adrian P. Simpson, Jasmine Öjbro & Christine Ericsdotter Nordgren. 2019. The phonetics of gender in Swedish and German. *Proceedings of Fonetik, 49-53.*

Weirich, Melanie, Stefanie Jannedy & Gediminas Schüppenhauer. 2020. The Social Meaning of Contextualized Sibilant Alternations in Berlin German. *Frontiers in Psychology*, 11.

West, Candace & Don H. Zimmerman. 1987. Doing gender. *Gender & Society* 1(2), 125–151.

Whiteside, Stephen P. 2001. Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences. *Journal of Phonetics* 29(3), 281–300. https://doi.org/10.1006/jpho.2001.0134.

Woolard, Kathryn A. & Bambi B. Schieffelin. 1994. Language ideology. *Annual Review of Anthropology* 23, 55–82.

Wortham, Stanton Emerson Fisher. 2006. *Learning Identity : The Joint Emergence of Social Identification and Academic Learning*. Cambridge : Cambridge University Press.

Yeni-Komishian, Grace H. & Sigfrid D. Soli. 1981. Recognition of vowels from information in fricatives: Evidence of fricative-vowel coarticulation. *Journal of the Acoustical Society of America* 70(4), 966–975.

Yuasa, Ikuko Patricia. 2008. *Culture and gender of voice pitch: A sociophonetic comparison of the Japanese and Americans.* London: Equinox.

Yuasa, Ikuko Patricia. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech* 85(3), 315–337. https://doi.org/10.1215/00031283-2010-018.

Zimman, L.al. 2013. Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language and Sexuality 2*(1), 1-39.

Zimman, Lal. 2017a. Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and/s/. *Language in Society 46*(3), 339–370. https://doi.org/10.1017/S0047404517000070

Zimman, Lal. 2017b. Variability in /s/ among transgender speakers: Evidence for a socially grounded account of gender and sibilants. *Linguistics* 55(5), 993–1019. De Gruyter Mouton. https://doi.org/10.1515/ling-2017-0018

Zimman, Lal. 2018. Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass 12*(8). https://doi.org/10.1111/lnc3.12284

Zimman, Lal. 2020a. Transgender Language, Transgender Moment: Toward a Trans Linguistics. In Kira Hall & Rusty Barrett (eds.) *The Oxford Handbook of Language and Sexuality (online edn., Oxford Academic), https://doi.org/10.1093/oxfordhb/9780190212926.013.45, accessed 15 April 2025.*

Zimman, Lal. 2020b. Trans self-identification and the language of neoliberal selfhood: Agency, power, and the limits of monologic discourse. *Language in Society* 49(2), 255–279.

Zimman, Lal & Kira Hall. 2016. Language, gender, and sexuality. In Mark Aronoff (ed.) *Oxford bibliographies in linguistics*. New York: Oxford University Press.

8. Annex

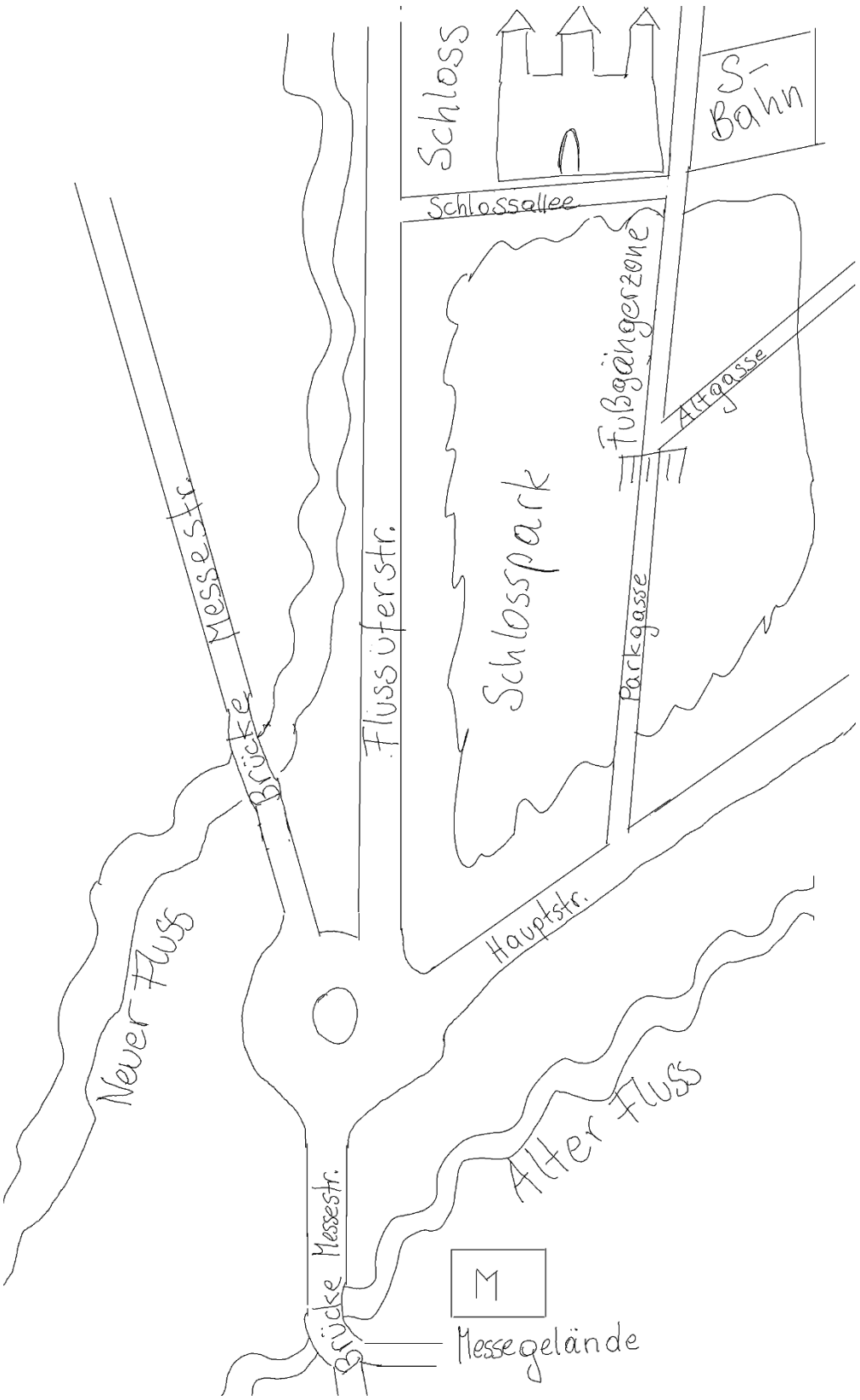

Figure 9: blank version of map-task

Transcripts

The transcripts can be found under the following link as they were too long to attach:

https://docs.google.com/document/d/1GuR3SR1koXU29YYffbINqd712Qnt1tNU-SjG4xz7mWE/edit?usp=sharing

Abstracts

Masculinity articulated: /s/-variation as a sociophonetic resource in transmasculine German-speakers discursive co-construction of gender

Within the large indexical field associated with /s/-variation, gender and sex have often been explored as its origins. This search for causes of gendered /s/-variation has mostly taken precedence over exploration of its social indexicality and contribution to discursive gender co-construction. The present study proposes that two important concepts must play a role in analysis of this so-called 'gay lisp' phenomenon: the **phonosonic nexus** (Harkness 2014) and **hegemonic masculinity** (Connell 1995). By employing these two frameworks, we can approach the gender indexicality of /s/-variation from a semiotic standpoint. Instead of looking for origins we can thereby ask what the phonosonic nexus of /s/-variability contributes to negotiations of (non-)hegemonic masculinity in discourse. Inspired by the findings of Zimman (2017), who demonstrates that a complex picture of *stylistic bricolage* emerges for transmasculine speakers' use of /s/ (as their Centers of Gravity [CoG] encompass the entire range typically reported for male *and* female speakers of English), this thesis explores if similar findings can be made for German. While enregisterment of fronted, high-CoG /s/ (/s+/), as 'gay' or 'feminine' in English-speaking men is found by Boyd et al. (2021), German-speaking gay cis-men may produce /s+/, but no uptake could be observed. Based on these findings, this study investigates whether /s/ is employed as an effective sociophonetic resource in transmasculine German-speakers' discursive co-construction of gender, suggesting an ongoing change in the indexicality of /s/-variation. This is achieved through interviews and a perception study. Results suggest that /s/-variation may indeed be used in German to contextually navigate different dimensions of gender which results in a picture of stylistic bricolage similar to the one displayed by Zimman's English-speakers. Perception thereof may yet be limited to specific groups or contexts. Results show that urban, male listeners and those with explicit knowledge of the English phenomenon were more likely to rate speakers as more 'gay' sounding while non-

male, rural and younger listeners were more likely to rate speakers as more 'masculine'. This points towards a change in indexicality of /s/-variation which is not yet salient enough for broad enregisterment.

(keywords: Applied Linguistics, Sociophonetics, Semiotics, /s/-variation, transmasculinity)

References

Boyd, Zac, Josef Fruehwald & Lauren Hall-Lew. 2021. Crosslinguistic perceptions of/s/among English, French, and German listeners. *Language Variation and Change 33*(2), 165–191. https://doi.org/10.1017/S0954394521000089

Connell, R.W. 1995. *Masculinities*. Cambridge: Polity.

Harkness, Nicholas. 2014. *Songs of Seoul: An ethnography of voice and voicing in Christian South Korea*. Berkeley: University of California Press.

Zimman, Lal. 2017. Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and/s/. *Language in Society 46*(3), 339–370. https://doi.org/10.1017/S0047404517000070

Artikulierte Männlichkeit: /s/-Variation als soziophonetische Ressource in der diskursiven (Ko-)Konstruktion von Gender bei transmaskulinen Deutschsprechenden

Im weiten indexikalischen Feld, das mit /s/-Variation verbunden ist, wurden Gender und biologisches Geschlecht häufig als dessen möglicher Ursprung untersucht. Diese Ursachenforschung zu gegenderter /s/-Variation nimmt meist Vorrang vor der Erforschung ihrer sozialen Indexikalität und ihres Beitrags zur diskursiven (Ko-)Konstruktion von Gender. Die vorliegende Studie schlägt vor, dass zwei wichtige Konzepte in der Analyse des sogenannten ‚Gay Lisp'-Phänomens einfließen müssen: der **phonosonische Nexus** (Harkness 2014) und die **hegemoniale Männlichkeit** (Connell 1995). Mithilfe dieser beiden theoretischen Rahmen kann die gegenderte Indexikalität von /s/-Variation semiotisch betrachtet werden. Anstelle einer Suche nach Ursachen kann so gefragt werden, welchen Beitrag der phonosonische Nexus von /s/-Variabilität zur Aushandlung (nicht-)hegemonialer Männlichkeit im Diskurs leistet. Inspiriert von den Ergebnissen Zimmans (2017), der zeigt, dass sich bei transmaskulinen Sprechern ein komplexes Bild stilistischer Bricolage ergibt (da ihre Center of Gravity [CoG] den gesamten Bereich umfassen, der typischerweise für männliche sowie weibliche Sprecherinnen des Englischen beobachtet wird), untersucht diese Arbeit, ob sich ähnliche Befunde für das Deutsche finden lassen. Während Boyd et al. (2021) ein *enregisterment* eines nach vorn

verlagerten, hochfrequenten /s/ (/s+/) als ‚schwul' oder ‚feminin' bei englischsprachigen Männern feststellen, produzieren deutschsprachige schwule cis-Männer zwar /s+/, jedoch konnte keine soziale Zuschreibung (Uptake) beobachtet werden. Auf Basis dieser Erkenntnisse untersucht diese Studie, ob /s/ als effektive soziophonetische Ressource in der diskursiven (Ko-)Konstruktion von Geschlecht bei transmaskulinen Deutschsprechenden verwendet wird – und legt nahe, dass sich die Indexikalität der /s/-Variation im Wandel befindet. Dies wird durch Interviews und eine Perzeptionsstudie überprüft. Die Ergebnisse deuten darauf hin, dass /s/-Variation im Deutschen tatsächlich genutzt werden kann, um kontextuell unterschiedliche Dimensionen von Gender auszuhandeln – wodurch ein Bild stilistischer Bricolage entsteht, das dem von Zimmans englischsprachigen Sprechern ähnelt. Die Wahrnehmung dessen scheint jedoch auf bestimmte Gruppen oder Kontexte beschränkt. Die Ergebnisse zeigen, dass städtische, männliche Hörer sowie solche mit explizitem Wissen über das englische Phänomen Sprecher häufiger als „schwul klingend" einstuften, während nicht-männliche, ländliche und jüngere Hörer:innen die Sprecher eher als „maskulin" wahrnahmen. Dies weist auf eine Veränderung in der Indexikalität der /s/-Variation hin, die jedoch (noch) nicht breit genug ist, um als *enregistered* zu gelten.

(Schlagwörter: Angewandte Linguistik, Soziophonetik, Semiotik, /s/-Variation, Transmaskulinität)

Quellen

Boyd, Zac, Josef Fruehwald & Lauren Hall-Lew. 2021. Crosslinguistic perceptions of/s/among English, French, and German listeners. *Language Variation and Change 33*(2), 165–191. https://doi.org/10.1017/S0954394521000089

Connell, R.W. 1995. *Masculinities*. Cambridge: Polity.

Harkness, Nicholas. 2014. *Songs of Seoul: An ethnography of voice and voicing in Christian South Korea*. Berkeley: University of California Press.

Zimman, Lal. 2017. Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and/s/. *Language in Society 46*(3), 339–370. https://doi.org/10.1017/S0047404517000070