[WLG]
WIENER LINGUISTISCHE GAZETTE

# Bridging interdisciplinary demands

## From bias to balance in integrating interactional sociolinguistics and acoustic phonetics

*Carina Lozo*

Sonderdruck aus: *Wiener Linguistische Gazette* 95 (2024): 7–30

Universität Wien · Institut für Sprachwissenschaft · 2024

**Eigentümer und Verleger:**
Universität Wien, Institut für Sprachwissenschaft
Sensengasse 3a
1090 Wien
Österreich

**Herausgeberschaft:**
Jonas Hassemer, Florian Grosser & Carina Lozo (Angewandte Sprachwissenschaft)

**Erweiterte Redaktion:**
Markus Pöchtrager (Allgemeine Sprachwissenschaft)
Stefan Schumacher (Allgemeine und Historische Sprachwissenschaft)

**Kontakt:** wlg@univie.ac.at
**Homepage:** *http://www.wlg.univie.ac.at*

**ISSN:** 2224-1876
**NBN:** BI,078,1063

Die *Wiener Linguistische Gazette* erscheint in loser Folge im Open-Access-Format.
Alle Ausgaben ab Nr. 72 (2005) sind online verfügbar.

# Bridging disciplinary demands

## From bias to balance in integrating interactional sociolinguistics and acoustic phonetics

*Carina Lozo*[*]

*Wiener Linguistische Gazette* (WLG)
Institut für Sprachwissenschaft
Universität Wien
Ausgabe 95 (2024): 7–30

**Abstract**

Dieser Beitrag befasst sich mit den Herausforderungen, die sich aus den unterschiedlichen Ansätzen zur Datenerhebung in der akustischen Phonetik und der interaktionalen Soziolinguistik ergeben, sowie mit dem Versuch, diese Bereiche in dieser Hinsicht zu verbinden. Das Hauptaugenmerk liegt auf der Zusammenführung unterschiedlicher Datenerhebungsmethoden und der Beleuchtung ihrer methodologischen Lücken. Der Beitrag reflektiert kritisch die Integration von akustischer Phonetik und interaktionaler Soziolinguistik innerhalb eines Forschungsprojekts, das die Rolle der Stimme in Gender-Performances untersucht.

**Schlagwörter:** Interdisciplinarity, reflexivity, phonetics, interactional sociolinguistics

* Carina Lozo, Department of Linguistics, University of Vienna, carina.lozo@univie.ac.at

## 1 Introduction

"I've never really thought about authenticity, we all just assume that the interactions we put the speakers in are genuine," remarked an accomplished phonetician when I asked about how they ensure authentic interactions in their data collection process. This nonchalant response left me contemplating, as my PhD project seeks to investigate "authentic" or "real life" interactions from a phonetic perspective and at this point, I had invested quite some time in deciphering the essence of data authenticity.

The phenomenon of the human voice, with its intricate interplay of mechanics, functions, limitations, and its diverse roles in social interactions, has been a subject of interest across disciplines for decades. This fascination with the voice extends from medical domains to philosophical discourse, where the voice's enigmatic nature has often fueled multidisciplinary investigations. In the framework of my PhD project rooted in applied linguistics, I intend to assimilate different perspectives with the aim of understanding the voice as a social practice. In this effort, I draw on various disciplines such as philosophy, linguistic anthropology, phonetics, and interactional sociolinguistics; my primary focus, however, remains constant: understanding the voice as a mechanism intricately tied to gender performance. This paper focuses on acoustic phonetics and interactional sociolinguistics, which form the cornerstones of my corpus collection and, consequently, should be the primary focus of this contribution.

My academic journey into this field, which I tentatively term "sociophonetics of the voice," was driven by both personal curiosity and scholarly interest. This blend of personal intrigue and academic pursuit paved the way for my exploration of the nuances of human interaction and the physical manifestation of identities through voice. On a personal level, I have always been fascinated by the intricacies of these phenomena. Simultaneously, my academic interest has been drawn to the intersection of voice with culture and identity and how it is reflected on a physical level, one that could be quantified. With my formal academic training in applied linguistics and a background in phonetics,

I felt prepared for the tensions I expected to encounter. However, with time, it became evident that these tensions required a deeper engagement than I initially anticipated. It was at this point that I realized that the inherent reflexivity of applied linguistics had been somewhat lost on me as I wanted to transcend disciplinary boundaries.

This contribution addresses the unique challenges that arise from the contrast in data collection methods between acoustic phonetics and interactional sociolinguistics, as well as the effort to bridge these fields in this regard. Acoustic phonetics is rooted in experiments within controlled settings, whereas interactional sociolinguistics places a premium on the unembellished interpersonal communication found in unscripted, spontaneous conversations. The latter posits a fundamental aversion to staged interactions, stressing the significance of uncontrived communicative exchanges. This juxtaposition describes the challenge at hand – how to harmonize the different demands of acoustic phonetics and interactional sociolinguistics.

To achieve this goal, I strived to design an experimental data collection setup that could serve the requirements of both fields. The benefit of the integration of acoustic phonetic analysis within interactional sociolinguistics can be seen as crucial for enriching the understanding of conversational data. Addressing an underexplored territory within sociolinguistics, the challenge lies in quantifying nuanced sociolinguistic phenomena, such as the intricacies of vocal gender performance. Ultimately, my project aspires to contribute to the development of a gender-neutral synthetic voice, and thereby highlights the need to understand gender-linked metrics that phonetic analyses could provide. This convergence of requirements prompts a reevaluation of disciplinary boundaries, advocating for an inclusive, interdisciplinary approach. This shift challenges common hierarchical preferences for quantitative over qualitative methods, aiming for a more balanced research framework. Integrating acoustic phonetic analysis in sociolinguistics and vice versa, infusing sociolinguistic insights into phonetic

analysis, uncovers hidden dimensions in interactional data and prompts a reassessment of entrenched disciplinary norms.[1]

In what follows, I will first present the theoretical framework that underpins my corpus collection (Section 2). Section 3 delves into a reflexive account, addressing the navigation of biases encountered in bridging the disciplinary demands stemming from sociolinguistics and phonetics. In Section 4, I discuss the conditions required for building corpora and the methodological choices that guided this process. Moving forward to Section 5, I discuss the data generizability of the collected corpus. Finally, in Section 6, I conclude this paper by summarizing implications drawn from this interdisciplinary exploration.

## 2 Theoretical framework

This section navigates the interdisciplinary blend of interactional sociolinguistics and acoustic phonetics. Subsection 2.1 delineates the contrasting methodologies and perspectives of these fields: interactional sociolinguistics focusing on social dimensions in language use, and acoustic phonetics emphasizing quantifiable speech analysis. The discussion further explores, in 2.2, the multifaceted concept of "authenticity" within experimental setups, acknowledging its dynamic nature and relevance in capturing genuine social interactions.

### 2.1 Contrasting approaches

The decision to combine interactional sociolinguistics with acoustic phonetics in the study of gendered voices is rooted in the comprehensive nature of these two disciplines. The juxtaposition of these fields not only enhances the understanding of how gender is constructed and

1 It must be noted that phonetic analyses have historically repeatedly been employed in sociolinguistics, particularly within the variationist framework. However, I seek to differentiate my approach from this tradition, as the objectives of variationist sociolinguistics diverge from the broader epistemological foundations that guide my research project.

expressed through language use and its acoustics but also paves the way towards a gender-inclusive synthetic voice landscape.

Interactional sociolinguistics emerged in the latter half of the 20th century and builds upon the foundations of sociolinguistics and conversation analysis. The field stems from the recognition that language is a social practice embedded in various contexts and focuses on examining the different sociocultural factors that influence language use, including gender, social identities, and social norms. The primary objective of interactional sociolinguistics is to shed light on the complex relationship between language and society by analyzing the nuances of communication in "real life" situations and interactions (Hinnenkamp 2018; Imo & Lanwer 2019; Rampton 2020). "Real life" interactions in this sense signify the unscripted exchanges among individuals, which occur within their everyday environments.

Exploring gendered voices within interactional sociolinguistics unveils complex relationships among voices, gender identity, and the fabric of social dynamics. Methodologies such as discourse analysis, playback interviews, and microanalysis play an important role in decoding the nuances of gender performance embedded within vocal expressions. These approaches serve as invaluable tools, offering insights into the multifaceted dimensions of how gender manifests and evolves within the realm of spoken language. Playback interviews serve as a complementary tool, offering participants an opportunity to reflect upon their linguistic choices, intentions, and underlying motivations during recorded interactions. They provide a retrospective lens, for the researcher and the participant, allowing for the examination of conversational snippets, unveiling hidden layers of gender expression within speech. Microanalysis explores the granular elements of speech, meticulously scrutinizing phonetic and prosodic features to unravel the gendered nuances imprinted in vocal communication. Collectively, these methodologies within interactional sociolinguistics offer a comprehensive toolkit for understanding the complexities inherent in gendered voices.

In tandem with interactional sociolinguistics, acoustic phonetics, the study of the physical properties of speech sounds, provides a quantifying framework to explore the acoustics of gendered voices. Its primary goal

is to understand the mechanisms and variations in speech production, as well as the auditory aspects of speech perception (Harrington 2010). Assessing the acoustics of human voices involves evaluating various domains within the voice's acoustic signal, including *periodicity*, *noise*, and *amplitude relations*. Periodicity, representing the frequency of the glottis' opening and closing cycle, finds expression primarily in the fundamental frequency (Fo), associated with the perceived pitch of the voice. When considering noise within the voice signal, measures focus on quantifying the additive noise present alongside the harmonic segments. One common metric used for this purpose is the harmonics-to-noise ratio (HNR), which delineates the relationship between harmonic and non-harmonic portions of the signal. This measure is particularly relevant in analyzing voices characterized by creakiness for example, where a heightened degree of noise is evident. Amplitude relations provide further insights, albeit multifaceted, depending on the specific relations under consideration. In the sociophonetics of the voice, examination often involves assessing amplitude differences between specific harmonics, such as the first and the second harmonic, H1-H2, and the second and the fourth harmonic, H2-H4. These amplitude relationships offer intricate details: H1-H2 variations, for instance, signify pronounced glottal opening and tend to manifest in breathy voices. Conversely, H2-H4 measurements are indicative of glottal tension, typically observed to be higher in male speakers compared to females (Keating et al. 2015).

Interactional sociolinguistics delves into the social dimensions of everyday language use, employing qualitative methods that prioritize contextual understanding of nuanced social interaction. While sacrificing statistical accuracy or broad generalizations, these methodologies aim to grasp the nuances of social dynamics in interactional settings. In contrast, acoustic phonetics relies on quantitative measurements to derive quantifiable outcomes and broad insights into speech acoustics. However, this emphasis on quantifiability may sometimes overlook the organic nature of social interactions, focusing more on physical measurements and potentially missing the richness of everyday experiences.

Despite their differing approaches – quantitative and qualitative, respectively – both fields converge in their exploration of language, albeit from distinct vantage points.

By merging these disciplines, I aim to holistically investigate how gender identity is shaped socially and manifested acoustically in vocal communication. This interdisciplinary perspective bridges the gap between the social and physical aspects of communication, merging quantitative and qualitative approaches to understand language and social interaction.

## 2.2 "Authenticity" in experimental settings

This section is dedicated to outlining my stance on "authenticity" and presenting how I will approach the concept within the scope of my work. In my pursuit of spontaneous and unscripted interactions, conventional setups for phonetic research, which typically include sociolinguistic interviews, proved inadequate. While De Fina & Perrino (2011) advocate for "authenticity" within an interview context, as it presents a legitimate social interaction, I could not presume that such controlled environments would encapsulate the "real life-ness" I strive to capture. My focus gravitated towards capturing speech that is unrestricted and casual to a level that goes beyond what can be replicated in a controlled sound booth environment. Hence, I am interested in interactions that occur in the daily lives of speakers, where the "speaker does not reflect on their existence but merely exists" (Kramsch 2012: 486). The quest for "authentic" or "natural" everyday interactions stands at the intersection of debate and complexity. These seemingly straightforward terms laden with ideological underpinnings, have meanings shaped by individual beliefs, societal norms, and cultural contexts. The understanding of what is deemed authentic or natural varies across different research cultures but usually refers to the quality of being "genuine" or "real."

The notion of "authenticity" in sociolinguistic research is often intertwined with discussions about identity, power dynamics, and

representation, rendering it a nuanced and ideologically charged concept. This concept, however, acquires a new dimension when applied to linguistic data itself in the context of my study, where I combine two fields with ontological and epistemological differences. The data we collect, once considered a reflection of "unfiltered" linguistic expression (Labov 1972), is now recognized as a product of conscious and context-dependent choices made in its collection and interpretation (Eckert 2014). To critically examine "authenticity" in social interactions and what it means for my data collection process, is to unravel its nuanced nature, recognizing it not as a fixed entity but as a dynamic and context-dependent phenomenon (Lacoste et al. 2014). The "authenticity" of social interactions is embedded not just in the linguistic choices of the speakers but also in the methodological decisions of the researcher. As highlighted by Buchholtz's (2003) problematization of the concept, the claim for "authenticity" is inherently relational, never total but always partial. Further, it is produced through contextually situated and ideologically informed configurations of self and other.

The importance of "authenticity" extends beyond social interactions. The transition from the "authenticity" of social exchanges to the data collection settings demands precision without sacrificing the spontaneity inherent in everyday discourses. "Data authenticity," then, becomes the conduit translating interactions into a corpus reflecting spontaneous linguistic expressions, contextual intricacies, and everyday conversations. This transition embodies a dichotomy: data authenticity must represent both the social interactions and the setting they emerge from.

Acknowledging the changing concept of authenticity in sociolinguistics (cf. Androutsopoulos 2015; Coupland 2003; Lacoste 2014), which primarily focuses on speakers and their linguistic expressions, I extend scrutiny to the authenticity of the data collection process itself. The "deconstruction of authenticity" in sociolinguistics, as termed by Androutsopoulos (2015), extends beyond recognizing authenticity as a socially constructed phenomenon instead of an inherent attribute. It also prompts a reevaluation of the conventional understanding of an "authentic speaker," as this deconstruction advocates for a significant expansion in the scope of analytic objects. It thus encompasses various

interactional settings where authentication processes are observable and influential. This shift in focus recognizes that data is a construct shaped by the researcher's decisions and the surrounding context.

To better describe the setting I seek, I want to broaden the concept of an "inherent authenticity," which categorizes objects or settings as either authentic or inauthentic, through a notion of ethereality. This notion describes the intangible essence arising from the interconnectedness of the researcher, participants, and the experimental setting. Consequently, it acknowledges that the authenticity of both data collection setup and the data itself is discursively co-created. This notion addresses the evolving complexities of understanding "genuine" social interactions, capturing their dynamic and context-dependent nature within this interplay. It emerges as pivotal in describing the social interactions I seek for my corpus. It underpins my methodology, bridging the gap between conventional data collection approaches and the ever-evolving dynamics of human engagement. In this sense, authenticity with an ethereal aspect refers to a distinctive genuineness that eludes replication. This expansion of the conventional understanding of authenticity underscores the irreplicable nature of the interactions I seek. It acknowledges the unique, ephemeral quality that permeates real life interactions, emphasizing their connection to the everydayness and transience of human engagement.

## 3 Navigating biases and interdisciplinary collaboration

My project seeks to blend acoustic phonetics and interactional sociolinguistics by uncovering a shared ground where both fields can coalesce, thereby shedding light on their methodological blind spots. To navigate this intersection effectively, it is crucial to address the biases and challenges that arise.

First, it is essential to acknowledge that, as a researcher with a more positivist background steeped in phonetics, my initial inclination naturally gravitated towards constructing a corpus that catered more to phonetics' requirements. This gravitational pull was, in part, driven by

my inherent comfort and familiarity with the acoustic intricacies, controlled environments, and instrumental setups characterizing phonetic data collection. This familiarity inevitably led to a degree of caution that was reflective of my training and perspectives within the discipline of phonetics. This self-acknowledged bias towards phonetic criteria necessitated an introspection – an examination of how this orientation might inadvertently overshadow the nuances cherished by sociolinguistics, particularly those stemming from unscripted and everyday interactions. The potential for my biases to overshadow the methodological choices informed by sociolinguistic principles served as a clarion call for critical self-awareness.

Despite being advocated and encouraged, the practice of self-reflection is conspicuously absent from formal academic training. However, as a PhD candidate, I frequently find myself immersed in reflective contemplation and a continuous reevaluation of stances and choices remains an enduring aspect of my academic trajectory, at times arising from an intrinsic motivation to preempt potential critique.

In the beginning, as I stepped into my role as a researcher, I underestimated the impact it would have on my perspectives and my understanding of the research process. Especially in the context of data collection, as my background in experimental phonetics provided ample prior experience. At the outset, I wanted to be a mere vessel for results, a neutral figure devoid of personal inclinations or needs. In this view, I saw myself as a smooth, unobtrusive surface, akin to a conductor in the realm of electrical circuits – simply transmitting data.

However, reality took a different shape. The challenge arose when I found it increasingly difficult to cleanly separate my personal identity from my research persona. These two facets of my life, which I had assumed could be neatly compartmentalized, became inexplicably intertwined. This confluence of identities is not a novel experience for seasoned academics and it raises questions about the extent to which we, as researchers, are truly "objective and detached" from our subjects of study. It forces us to confront the complex interplay between our personal motivations, biases, and the pursuit of academic knowledge.

As I navigated the line between catering to the phonetic demands and incorporating the spontaneity valued by sociolinguistics, it became evident that each decision I made was underscored by an ongoing dialectic. This dialectic was marked by a constant negotiation of the tensions arising from my inherent biases while striving to remain faithful to the broader interdisciplinary goals of the study. Not being aware of my inherent biases, I proceeded with the presumption that gathering sociolinguistic interviews within the confines of an acoustic laboratory would suffice for the sociolinguistic aspect of the corpus. Also, a certain perseverance influenced my perspective, as I maintained that the notion of "requirements" was best suited to the realm of experimental phonetics – a domain I felt more aligned with.

Through exchanges with other scholars from the field of phonetics, such as the quote in the beginning of Section 1 shows, a realization emerged that the pursuit of "real life" speech ultimately is secondary to precise, reliable, and quantifiable data. It became evident that, while the desire for "real life" data remains commendable in the field, the pragmatic realities of research often fail to find a balanced compromise between the aspiration for real life data and the practical considerations of ensuring data reliability and measurability. Through these academic discussions, I recognized that, while authenticity's inherent value is acknowledged, it is often the pursuit of methodological accuracy that shapes the contours of research practice. This is when I realized that the standards I had established for my project stood in contrast to the principles upheld within phonetics.

While I appreciated the discussions and outcomes of exchanges with phoneticians, my reflections were notably enriched by engaging with researchers from other linguistic domains – dialogues that directed me to a more comprehensive grasp of linguistic interaction. Through these conversations, my understanding of the sociolinguistic perspective deepened, as I came to recognize the merit of unscripted conversations as vital portals into understanding the intricate fabric of the human voice. This perspective led to a recalibration of my approach – one that placed a high value on unstaged discourse as an indispensable conduit for exploring vocal variation in its myriad forms.

Now, to create a corpus that would effectively integrate both acoustic phonetics and interactional sociolinguistics, a precursor involved the reflective account above. This account marked a departure from my initial inclinations. I needed to acknowledge that my past perspectives, while rooted in principles of scientific objectivity, had the potential to obstruct my further exploration of the voice, which I was planning.

This awareness underscores the pivotal role of reflexivity in scholarly endeavors. It captures the journey from one's own momentarily obscured biases toward a realization that emerges through a dynamic interchange of perspectives, conversations with colleagues, and the recognition of one's own academic foundations.

The outcome of this thought process, the corpus described below, is in itself emblematic of the inherent growth that academia fosters. It reminds us that, through our academic endeavors, maintaining an open mindset and the willingness to recalibrate are essential companions on our intellectual journey.

Ultimately, the journey from biases to a more balanced and appreciative standpoint was catalyzed by dialogues within the research community. Engaging with other scholars not only grounded my research within a broader linguistic narrative but also underscored the profound significance of interdisciplinary collaboration. Through this exchange of insights, the study's methodological approach evolved from an individual self-dialogue into a collective endeavor.

## 4 Divergent approaches to data collection

This section explores contrasting approaches in linguistic data collection. It starts by examining the paradox of observing human behavior in the social sciences, pivoting towards contrasting perspectives on performance (4.1). Then the section transitions to two lenses in 4.2, acoustic phonetics and interactional sociolinguistics, by highlighting disparities in data collection methodologies, revealing intentional deviations from each other's foundational principles. Section 4.3 concludes with the

methodological choices made and the presentation of the data collection set up.

## 4.1 Contrasting performance

Viewed through a quantum physics lens, we confront the unsettling notion that "the observed is not objective," a concept that challenges the foundation of our social scientific endeavors, which aspire for objectivity. A realization that forces us to grapple with what is known as the *observer's paradox.* This dilemma, summarized by Labov, underscores the inherent predicament in collecting linguistic data: "The aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain this data by systematic observation" (Labov 1972: 209).

With Goffman, however, a nuanced perspective can be introduced. His theory, which is aptly captured in the German translation of his seminal work "The Presentation of Self in Everyday Life" (1956), portrays social interaction as a dramatic stage: "Wir alle spielen Theater" ('we all play theatre'). In this metaphor, individuals consciously or subconsciously perform roles for others, meaning that our actions and expressions are performative, similar to actors on a stage. This perspective acknowledges that while the act of observing inevitably influences behavior, it also underscores the inherent performative nature of social interaction itself. Just as actors on a stage engage in performances, our interactions in the social realm inherently bear a performative aspect, influenced by the awareness of being observed. While not entirely escaping the observer's paradox, this does offer a new lens through which to consider the intricacies of human interaction. Within this framework, the desired "natural" or "authentic" interactions sought by sociolinguists are consistently veiled by a certain level of performativity on the part of the speakers, which remains constant and cannot be "turned off."

Having explored the complexities of observation within social interactions, I want to shift the focus now to contrasting perspectives

on performance. Viewed through both phonetic and interactional sociolinguistic lenses, distinct focal points emerge, highlighting diverse approaches to understanding this concept.

Phonetics, emphasizing the physical and acoustic properties of speech sounds, scrutinizes the articulatory mechanics and acoustic manifestations of speech production. It investigates the physiological aspects of how sounds are formed within the vocal tract, the variations in phonemes, and the acoustic signatures of individual speech units. Performance, within phonetics, revolves around the execution of speech sounds – its articulatory precision and acoustic properties – often in controlled settings, aiming to comprehend the mechanics of sound production.

On the other hand, interactional sociolinguistics takes a broader, socially embedded approach to performance. It views performance as more than the mere production of speech sounds; instead encompassing the ways individuals navigate social contexts through language use. Interactional sociolinguistics scrutinizes how language performance reflects and shapes social identities, how individuals strategically employ linguistic resources within conversations to negotiate relationships, convey social meanings, and adapt speech patterns according to the situational and contextual demands.

Conversely, while phonetics aims to understand the physical and acoustic intricacies of speech production, interactional sociolinguistics widens the scope, exploring the social and cultural dimensions within diverse social contexts. These perspectives together offer a comprehensive understanding of performance, integrating the physiological mechanics of speech production with the socially embedded nature of linguistic interactions.

## 4.2 Disparities in corpus collection

Building on these conceptual reflections, I will now explore the practical requisites for data collection. In attempting to reconcile these approaches, it appears that each field purposefully veers away from the

core principles upheld by the other. The requirements of an "ideal" corpus in interactional sociolinguistics surpass the participant's role as a mere subject of study. For instance, inviting individuals into a controlled laboratory setting solely to record their speech within an interview situation presents inherent challenges. Central to my approach is the principle of minimizing staging and artificiality during data collection, with the goal of capturing unscripted and spontaneous interactions. This aspiration encounters the practical challenge of capturing everyday human communication.

In contrast to this, acoustic phonetics operates within the realm of structured interactions. Here, the emphasis is placed less on the content of the language used and more on the manner in which it is articulated. Striving for precision in the recording of speech, phonetic experiments often rely on controlled settings featuring sound-dampened environments and specialized microphones engineered to capture speech at a high quality. This approach is pivotal to ensure that the data remains easy to process when extracting physical measurements.

As mentioned above, acoustic phonetics adheres to stringent criteria due to the significant susceptibility of data to contamination, e.g., by background noise. These disturbances manifest in the acoustic signal and can have an impact on spectral measurements, complicating the differentiation between desirable and undesirable noise components. To address this challenge, controlled experiments are conducted within sound attenuated booths, ideally with sole occupancy by the speaker with minimal movement, while the experimenter operates externally, talking to and instructing the speaker through a microphone-headphone conduit. This arrangement creates a controlled environment that facilitates the capture of speech acoustics in as "clean" a state as possible. By "clean," I refer to acoustic signals that are free from distortions, clipping, or excessive background noise.

Within the scope of my corpus, I tried to uphold the "no background noise" principle primarily due to practical considerations, as the process of cleaning noisy speech data is time-intensive and can also be quite challenging. Concurrently I was trying to transition from the confined

laboratory environment towards a more open and unrestrictive environment, also allowing more participants into the setting.

In my exchanges with scholars from the field of speech acoustics, it was frequently emphasized that, over time, participants in controlled experimental settings, exhibit, to some extent, a tendency to become less conscious of the experimental environment, replete with cables, microphones, and screens. This perspective suggests that this setting becomes "authentic" to speakers or participants over time, although it may not fully align with the criteria typically emphasized within an interactional sociolinguistic framework. This is why the objective shifted towards finding an interactional context that would facilitate meaningful multiparty engagement, marked by minimal movement, a quiet background, unobtrusive microphone use, and a crucial emphasis on the inherent everydayness of the setting. In my data collection setup, I wanted to capture types of interaction that are direct, emergent, and inherently spontaneous and unscripted. The objective was to peel away the layers of artificiality that can accompany scientific framing and experimental goals, creating a space where spontaneous exchanges can thrive. Thus, the experimental setting should be stripped of any excessive staging, a conscious decision rooted in the belief that too much structure can impede the flow of interactions. This approach recognizes that an overly structured environment, laden with predetermined expectations or experimental constraints, can potentially stifle interpersonal encounters. An obvious first choice for the setting was the recording of a friend group who might overlook cables and microphones due to their familiarity with each other. However, the recording a friend group was initially dismissed, given the inherent challenge of providing each speaker with an exclusive microphone channel since cross-talk interference posed a significant concern, potentially undermining the quality of subsequent phonetic analyses. Thus, another essential criterion emerged, stipulating the need for individualized signal channels for each speaker to ensure the fidelity of the data. Finding a compromise between the desire for a spontaneous interactional setting and the technical constraints was essential. Recognizing the necessity of individual

microphone channels, I sought a solution that would enable the collection of conversational data while preserving the data's reliability and the ability to conduct thorough phonetic analyses.

Still casting a shadow, the pandemic paved the way for one of the initial viable suggestions that were presented to me:[2] the recording of a Zoom meeting – an action that would not have been considered as an everyday interaction just a few years ago. Yet, the idea of recording a friend group hinged on the requirement that the interactions happen organically, without any scripted prompts. Consequently, the option to record a Zoom meeting was ruled out rather quickly, as using Zoom for a group conversation resembled the staged environment of laboratory recordings. Nevertheless, this online setup did provide a valuable feature: Individual microphone channels, which I found essential.

## 4.3 Methodological choices and resulting corpus

After exploring various technical options, an online gaming setup was devised that allowed for both spontaneous interaction and the recording of clear, distinct audio signals from each speaker. This enabled the research to maintain its focus on interaction while ensuring the methodological accuracy required for robust phonetic analyses. In this way, the research project evolved, incorporating the insights from both interactional sociolinguistics and acoustic phonetics, and finding a balanced approach that respects the core principles of both fields.

A pillar of my corpus lays on the technical competence of the participants. The requirement for individual microphoning added a layer of complexity to the participant selection process, necessitating a more intricate approach beyond a simple random grouping of friends. The participants were required to possess a degree of technical proficiency, enabling them to independently manage their recording processes, ideally with a sensitivity for audio processing.

2 Another idea presented to me was, "let's make couples fight" – something which, from a scientific perspective at least, was worth considering, as conflict often proves to be an effective distractor in laboratory settings.

At the outset, I found myself grappling with an inclination to cease from personal involvement in the corpus – a reluctance stemming from the scientific ideologies that I, as an experimental phonetician, had internalized. This originated from the belief that a researcher's disengagement from the subject matter lent a particular "purity" to the investigation. The transformative pivot came in acknowledging that my dual identity as a researcher and participant did not compromise the scientific rigor I held dear, but rather elevated it. An immersion into the experiment served as a profound counterpoint to my earlier opposition.

With the immersion I transitioned into an ethnographic research context. However, it is worth noting that this immersion was not in a traditional ethnographic sense; instead, it was an attempt to exert a degree of influence on the setting. This transition comes with its considerations. As I adopted a more immersive stance, I recognized the impact of my presence and actions on the experiment environment. Balancing involvement and control became pivotal to ensure that my influence did not disrupt the flow of the interactions.

This change in approach notably shaped the recording process, where eventually a close-knit online gaming friend group became the focal point of the study. The group consisted of five people, three female (one of which was me) and two male, all around the age of 30. Each speaker recorded their audio in their own home, using the same setup they were accustomed to during their typical online gaming sessions. Through a Discord server (Discord Inc. 2023),[3] all speakers were connected and were able to talk to each other. Within Discord, users can create or join servers, which are dedicated spaces for discussions or collaboration. Servers can be customized with different channels for text or voice communication, allowing users to discuss specific topics or to engage in certain activities. The recordings were set up to resemble a gaming session, mirroring the environment familiar to the speakers. This setup included a specific voice channel within the Discord server,

3 Discord is a communication platform primarily designed for creating communities and enabling voice, video, and text communication among users. It is commonly used by gamers.

allowing the speakers to engage in real-time conversation. They each employed an external microphone connected to their computer along with a recording software to capture the sessions. The communication was exclusively through voice chat, without the use of text chat or any other channels.

In this manner, on three different occasions (excluding one pilot run), we played a survival game that the friend group regularly played, *Valheim*. The corpus emerged from conversations that transpired within this gaming/Discord server environment. While the gameplay acted as the catalyst for gathering, the conversations swiftly shifted towards everyday topics. The participants navigated discussions that traversed beyond the realms of gaming mechanics or strategies, delving into personal anecdotes, reminiscences from the past, work-related matters, and various facets of everyday life. These conversations provided a window into the flow of conversations among friends, capturing the nuances of language use, social dynamics, and the interplay of speech elements that transcend the gaming context. Despite the initial gaming premise, the resultant corpus reflects real life interactions among friends, showcasing the intricate fabric of everyday communication and the multifaceted nature of linguistic interactions.

Regarding ethical considerations, I ensured that the speakers were well-informed about the consent process, with the understanding that they could retract their consent at any time. They were explicitly notified that their transcribed conversations would undergo both discourse analysis and phonetic analysis. Additionally, any personal information within the output text has been either removed or pseudonymized in accordance with privacy measures.

Despite the intention to shift away from the confines of the laboratory, my efforts unintentionally but inevitably transplanted certain laboratory conditions into participants' homes. This is particularly evident in the complex role participants play in instrumenting the setup. While promising, this approach is constrained by the demand for professional expertise or necessitates substantial participant training. Moreover, the diversity of home environments introduces various external interferences, such as the varying noise levels emitted by

laptops as they heat up. Consequently, considerable post-processing of signals remains a necessary step. However, in light of the "everyday" attributes I seek, this post-processing is a compromise I must accept. Regardless of these challenges, my optimism remains, as this corpus merely marks an initial step toward future linguistic inquiries, offering a chance to capture the essence of "speech in the wild."

## 5 Methodological impact on data validity and generalizability

At first glance, the collected data appeared somewhat disorderly, necessitating various post-processing steps such as aligning the separate signal channels in order to get the accurate timing of the interactions.

In order to be able to mix five different channels, a time-stamping procedure was carried out, as well as a noise profile assessment for each individual speech signal to facilitate a proper denoising procedure in the post-processing stage. For the desired acoustic precision, these post-processing procedures were essential. Through the noise profile, I could identify a baseline noise originating from each participant's computer. However, complexities emerged due to the intensifying computer noise throughout the recordings, necessitating a denoising process in different parts of the signal over each recording. This is when the question of generalizability surfaced.

Further, the inherent spontaneity of the recorded interactions introduced other unpredictable elements, such as keyboard strokes, mouse clicks, and human noises like laughter and exclamations, which pose challenges for phonetic analyses. These exclamations and laughter, however, can also be referred to as *response cries*, a concept introduced by Goffman (1978). Response cries refer to involuntary and spontaneous vocalizations or verbal expressions that individuals produce in response to a situation or stimuli. These are not premeditated or consciously planned utterances but are rather immediate and often reflexive responses to a particular moment. Goffman argues that these response cries are socially significant and are important for understanding the dynamics of social situations. The insights provided by

Aarsand & Aronsson (2009) in their analysis of gaming interactions add depth to understanding the dynamics of unscripted interactions. The authors explored how response cries, along with active noising (e.g., singing along, sound effects) and metacomments, contribute to the establishment of intersubjectivity and a sense of drama in gaming interactions. In essence, response cries serve as the sought-after markers of the interactional atmosphere I aim to capture. These vocalizations testify to the spontaneous, unscripted nature of the setting and present an acoustic phenomenon that one field attempts to evade while the other requires it.

Of particular interest is the heightened variability in the acoustic dimensions, which serves not merely as an artefact but rather as evidence of the everydayness and ethereal authenticity in this home-based setting. This observation suggests a reimagining of phonetic research – one that transcends laboratory confines and embraces the expansiveness of "real world" contexts. But how is this high variability to be treated? Normalizing the data would be probably the answer in many scholarly exchanges. However, normalizing data in this context is unsubstantiated. Firstly, the intention of normalizing is often to minimize variations and standardize data to facilitate comparisons and statistical analyses. However, in a dataset characterized by diverse contexts, interactions, and individual speaking styles, normalization could potentially remove crucial nuances and unique features that are inherent to the real world communication setting I try to capture. Secondly, attempting to normalize highly variable data might lead to distortion. Spontaneous conversations are replete with idiosyncratic features that contribute to their richness. Normalization could inadvertently erase these distinctive elements, rendering the data less representtative of actual communicative experiences. Moreover, normalization presupposes a certain level of consistency or regularity within the dataset, which may not align with the inherent variability of spontaneous, unscripted everyday interactions. Applying normalization techniques might result in artificially homogenized data that does not accurately reflect the complexity and diversity present in "real world" speech. In

this context, recognizing dataset variability as an asset rather than a limitation may lead to more insightful results.

## 6 Conclusion

In summary, this paper was motivated by the pursuit of a form of authenticity in my data collection process, with a focus on highlighting a non-replicable aspect of interactional context. This exploration into the sociophonetics of the voice emerges from a blend of personal intrigue and scholarly interest, resulting in an interdisciplinary endeavor marked by both challenges and opportunities that transcended disciplinary boundaries. A central challenge encountered revolves around reconciling the disparate approaches to data collection in acoustic phonetics and interactional sociolinguistics. While the former thrives on controlled experiments, the latter values unscripted and spontaneous discourses, rejecting staged interactions all together.

As technology continues to reshape our interactions, linguistic data collection in "real life" environments becomes increasingly feasible, with virtual and physical boundaries blurring. This digital integration offers new opportunities for observing and recording conversations in online spaces, creating a platform to address the tension between "real life-ness" and precision in data collection. The pandemic-induced shift to online interactions has accelerated these opportunities, though it has also introduced inadvertent laboratory-like conditions in participants' homes. Yet, as highlighted by Coupland already in 2003, "electronically" (i.e. digitally) mediated communication via remotely mediated networks, much like the corpus I have amassed, offers unique avenues for fostering intimacy and social connections, complementing traditional face-to-face interactions (Coupland 2003: 426). This observation illuminates the aspect of this shift, presenting opportunities for interdisciplinary exploration and emphasizing the potential of digital contexts in providing a platform for nuanced spoken interactions. As we navigate this evolving landscape, the fusion of "real life" and digital environments

could significantly contribute to understanding the complexities of spoken human communication.

Dedicating both time and space to document this reflexive process marks a significant step in my academic journey. It sheds light on an essential aspect of my scholarly involvement that often stays concealed. This way, I can show the seemingly linear progression of my research, an illusion that dissipates and gives way to a recognition of the myriad of ups and downs along the way. By incorporating this reflective process into a written narrative, I can, as a researcher, assume a critical stance, offering readers insight into the intricacies of my experiences and positioning my work within a broader academic discourse.

## Acknowledgements

I would like to thank Brigitta Busch, Florian Grosser, Jonas Hassemer, Jan Luttenberger, Dimitrios Meletis, Khoi Nguyen, Roshanak Nouralian, Vinicio Ntouvlis, Mandana Piroozfar, Jürgen Spitzmüller, and Jenia Yudytksa. Their collective support has been an integral part of this research, and I am immensely thankful for their encouragement, engaging discussions, and contributions.

## References

Aarsand, Pål André & Karin Aronsson. 2009. Response cries and other gaming moves—Building intersubjectivity in gaming. *Journal of Pragmatics* 41(8). 1557–1575.

Androutsopoulos, Jannis. 2015. Negotiating authenticities in mediatized times. *Discourse, Context & Media* 8. 74–77.

Bucholtz, Mary. 2003. Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics* 7(3). 398–416.

Coupland, Nikolas. 2003. Sociolinguistic authenticities. *Journal of Sociolinguistics* 7(3). 417–431.

De Fina, Anna & Sabina Perrino. 2011. Introduction: Interviews vs. 'natural' contexts: A false dilemma. *Language in Society* 40(1). 1–11.

Discord Inc. 2023. Discord. *https://discord.com*.

Eckert, Penelope. 2014. The trouble with authenticity. In Véronique Lacoste, Jakob Leimgruber & Thiemo Breyer (eds.), *Indexing authenticity*, 43–54. Berlin, München & Boston: De Gruyter.

Goffman, Erving. 1956. *The presentation of self in everyday life* (University of Edinburgh. Social Sciences Research Centre. Monograph No. 2). University of Edinburgh, Social Sciences Research Centre.

Goffman, Erving. 1978. Response cries. *Language* 54(4). 787–815.

Harrington, Jonathan. 2010. Acoustic phonetics. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The handbook of phonetic sciences*, 1st edn., 81–129. Chichester & Malden: Wiley Blackwell.

Hinnenkamp, Volker. 2018. Interaktionale Soziolinguistik. In Frank Liedtke & Astrid Tuchen (eds.), *Handbuch Pragmatik*, 149–162. Stuttgart: J.B. Metzler.

Imo, Wolfgang & Jens Lanwer. 2019. *Interaktionale Linguistik: eine Einführung* (Lehrbuch). Berlin [Heidelberg]: J.B. Metzler.

Kramsch, Claire. 2012. Imposture: A late modern notion in poststructuralist SLA research. *Applied Linguistics* 33(5). 483–502.

Labov, William. 1972. *Sociolinguistic patterns* (Conduct and Communication 4). Philadelphia: University of Pennsylvania Press.

Lacoste, Véronique, Jakob Leimgruber & Thiemo Breyer. 2014. Authenticity: A view from inside and outside sociolinguistics. In Véronique Lacoste, Jakob Leimgruber & Thiemo Breyer (eds.), *Indexing authenticity*, 1–13. Berlin & München & Boston: De Gruyter.

Rampton, Ben. 2020. Interactional Sociolinguistics. In Karin Tusting (ed.), *The Routledge handbook of linguistic ethnography* (Routledge Handbooks in applied linguistics), 13–27. London & New York: Routledge.