

TWO SIDES OF INTELLIGIBILITY:
THE PRACTICE AND PERCEPTION OF PERFORMED ACCENTS ONSTAGE

by

ELLEN LOUISE KRESS

A DISSERTATION

Presented to the Department of Theatre Arts
and Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2021

DISSERTATION APPROVAL PAGE

Student: Ellen Louise Kress

Title: Two Sides of Intelligibility: The Practice and Perception of Performed Accents
Onstage

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Theatre Arts by:

Theresa J. May	Chairperson
Nelson Barre	Core Member
John B. Schmor	Core Member
Melissa M. Baese-Berk	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies
-------------------	-----------------------------------

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded December 2021

© 2021 Ellen Louise Kress
This work is licensed under a Creative Commons
Attribution- NonCommercial-NoDerivs (United States) License



DISSERTATION ABSTRACT

Ellen Louise Kress

Doctor of Philosophy

Department of Theatre Arts

December 2021

Title: Two Sides of Intelligibility: The Practice and Perception of Performed Accents Onstage

The profession of voice and dialect is built upon the premise of maximum understanding for the audiences attending theatre. This maximum understanding, or intelligibility, has historically driven the practice and continues to shape the profession today. Intelligibility has been used as an objective measure for countless performers throughout the history performance. However, intelligibility may not be an objective threshold of listening, but a socially constructed term used for both the practice and perception of voices onstage. The work of this dissertation unpacks the idea of audience intelligibility from two perspectives—a critical examination of the relatively short history of the profession of voice and dialect in English-speaking countries, and an empirical investigation into the audience’s role in building intelligibility for actors. Intelligibility is in fact susceptible to social structures and individual’s preconceived normative ideas towards language.

Analysis in the history of voice and dialect reveals two recurring goals throughout the past two centuries. One goal of the practice was to eliminate any non-standard language usage in actors and students, to eliminate and traces of linguistic lived experiences for students onstage. The second goal is to replace these non-standard language varieties with sanitized or stereotyped versions of acceptable language varieties,

appearing as either a general standardized accent, or stereotypical versions of foreign or regional dialects.

The main results of the series of linguistic experiments appear in three main themes. The first main theme is the context of language (e.g., listening to a performance) will necessarily change how listeners perceive language. The second theme is that there are multiple ways to achieve maximum constructed intelligibility, which makes way for more diverse voices in performance. The third theme uncovers the ambiguous relationship between authenticity, imitation, and stereotype, which leads to bigger questions of the role authenticity continues to play in performance.

I then offer modifications to a profession by taking seriously the notion of intelligibility as a socially constructed judgment that has a real-world effect on perception. The findings from the history and the experiments contribute to my position about the state of contemporary voice and dialect practices. I use the findings from the body of this dissertation to grapple with my own position as a white theatre maker and advocate for practices that respect the linguistic autonomy of students and actors while honoring the needs of theatrical production.

CURRICULUM VITAE

NAME OF AUTHOR: Ellen Louise Kress

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of New Mexico, Albuquerque
Augustana University, Sioux Falls

DEGREES AWARDED:

Doctor of Philosophy, Theatre Arts, 2021 University of Oregon
Master of Arts, Linguistics, 2016 University of Oregon
Bachelor of Arts, Theatre and Linguistics, 2014, University of New Mexico

AREAS OF SPECIAL INTEREST:

Speech Perception
Actor and Voice Pedagogy
New works development

PROFESSIONAL EXPERIENCE:

Teaching Fellowship, University of Oregon, 2014-2021
Dialect Coach, Oregon Contemporary Theatre 2019-2021

GRANTS, AWARDS, AND HONORS:

Graduate school Special “Opps” grant. (2018, 2019). University of Oregon.
Meritorious Achievement Award in Direction for *Machinal*. (March 2019). Kennedy Center American College Theatre Festival Region VII.
General Excellence in Communication, Best website Design, and Best Social Media on behalf of GTFF, local 3544. (May 2018). American Federation of Teachers – Oregon
Arnold, Isabel, and Rupert Marks Scholarship. (2017-2018, 2018-2019, 2019-2020). Department of Theatre Arts. University of Oregon.
1 Hour Panel Presentation Panel Winner. (May 2017). University of Oregon Graduate Student Forum 8.

PUBLICATIONS:

Tonning-Kollwitz, Melissa, Joe Hetterly, and Ellen Kress. "The Current Use of Standard Dialects in the United States Theatre Industry." *Voice and Speech Review* (2021): 1-15.

Kress, Ellen, et al. "Embodiment and Social Distancing: Performances." *Journal of Embodied Research* 3.2 (2020).

Gillooly-Kress, Ellen. "Review: *Building Character* by Amy Cook." *New England Theatre Journal* (2019).

Gillooly-Kress, Ellen. (2018). "#HEWILLNOTDIVIDEUS: Weaponizing Performance of Identity from the Digital to the Physical." *The Journal of American Drama and Theatre*. Vol. 30, 2 (2018).

Kress, Ellen. "An Interdisciplinary Exploration of Language through Theatre and Linguistics" *UNM Department of Theatre and Dance Honors Thesis*. University of New Mexico (2014).

ACKNOWLEDGMENTS

I am profoundly indebted to and grateful for the Indigenous Nations and communities whose peoples, cultures, and customs continue to nurture my life. I was raised primarily on traditional lands of nineteen pueblos and the Apache. I completed this dissertation on the lands of the Kalapuya peoples. In this acknowledgement, I hope to build the capacity for solidarity with movements like #LandBack and #BlackLivesMatter and the activists engaged every day in the hard work of social justice.

I also like to acknowledge completing this dissertation in the global COVID-19 pandemic and the lives lost and livelihoods interrupted due to this global event. May we build a stronger, more resilient, and sustainable profession that acknowledges the dignity and creativity of the people who make the theatre profession so great.

For their patience, insight, and mentorship I am incredibly indebted to the members of this dissertation's advisory committee. To Theresa J. May, for your lessons in persistence and resiliency; to John Schmor, for jumping into this project feet first and reminding me there's joy to be had in this work; to Nelson Barre, for a fast friendship and stimulating inquiries; and to Melissa Baese-Berk for fostering and encouraging my interdisciplinary inquiry.

Much of this dissertation would not have been completed without access to a graduate employee union that fights tirelessly for the dignity of graduate students every day. To my comrades who served with me for three years on the executive board of the GTFF; to the 2018-2019 bargaining team for becoming beyond comrades and fighting to truly make

difference in the world (hail Satan!). Finally, to Michael Marchman for becoming a de facto mentor and treasured friend.

And to Liz Fairchild, Zeina Salame, Waylon Lenk, Anna Dulba-Barnett, Tricia Rodley, Chelsea Couch, Brendon Zuel, Stephen Armijo, Lilly Josten, Ashley Baker, Mica Pointer, Nate Severance, Harrison Sim, Arielle Owens, and Loren Billington thank you for your friendship, advice, your creative collaboration and space you create for me when I need it. To anyone I have worked with in the theatre, you inspire me more than you know. To Ryan Sayegh and Ilan Weinschelbaum, I love you and I would not be where I am today without you.

Finally, to my family, Joel and Barbara Kress, for raising an outgoing kid with a strong sense of justice who loves both art *and* science.

DEDICATED TO

My mom and dad
for in every thing
I do, they support
me all the way.



TABLE OF CONTENTS

Chapter	Page
I. INTELLIGIBILITY AND AUDIENCE MEANING-MAKING	1
1. Overview	1
1.2 Motivation and Structure	7
2. Literature Review: Different Disciplines, One Room	13
2.1 Thread One: The Foundations of the Voice Profession	15
2.2 Thread Two: A Cognitive Account of How Audiences Construct Intelligibility	28
3. Chapter Summary	37
II. A CRITICAL HISTORY OF VOICE AND DIALECT	42
1. Overview	42
2. Learn to Speak “Good” American English: Elocutionists 1900-1950	50
2.1 William Tilly Teaches World English.....	50
2.2 Edith Warman Skinner Bridges Performance	55
2.3 Mimesis Vs. Semiosis: Establishing Dialect Coaching as a Profession.....	60
3. Freeing Tension and the Rise of Regional Theatre	65
3.1 Berry and Linklater and the British Voice Training Revolution.....	68
3.2 A New Professional Organization Establishes Itself.....	74
3.3 The Hunt for the Perfect Dialect: Midcentury Dialect Coaching.....	78
4. Voice Practitioners Join the Internet	82
4.1 Monich Coaches for the Movies.....	86
4.2 Digital Approaches to Voice and Dialect	87

Chapter	Page
5. Towards a Cognitive Conception of Voice Training	93
III. CONSTRUCTING AUDIENCE INTELLIGIBILITY USING EMPIRICAL	
INQUIRY	101
1. Rationale for Empirical Approach.....	101
1.2 Language Perception: General Mechanisms, Several Models	107
1.3 Top-down Processes Help Organize the Speech Signal.....	114
1.3.1 What is a Dialect Anyway? Social Construction of an Accent	115
1.3.2 Social Construction of an Accent Affects Perception	120
1.3.3 Measures of Factors Affecting Accented Speech Perception.....	122
2. Preliminary Study	125
2.1 Participants	125
2.2 Stimuli	125
2.3 Listening Groups	126
2.4 Procedure	127
2.5 Results (Two Alternative Forced Choice Task)	127
2.6 Results (Free Response Question).....	129
2.7 Interim Discussion.....	131
3. Experiment 1	133
3.1 Method: Participants, Stimuli, Procedure.....	133
3.2 Results	134
3.2.1 Intelligibility (Accuracy of Recall)	134
3.2.2 Accentedness and Comprehensibility.....	136

Chapter	Page
3.2.3 Accentedness and Comprehensibility by Expectation	139
3.2.4 Who is the Actor?	141
3.2.5 Who Has the Most Stereotypical Accent?	142
4. Experiment 2	143
4.1 Method: Participants, Stimuli, Procedure	143
4.2 Results	145
4.2.1 Adjectives by Speaker	145
4.2.2 Adjectives by Expectation	148
4.2.4 Who is the Actor?	151
4.2.4 Who Has the Most Stereotypical Accent?	151
5. Discussion: What Can Practitioners Take from This Chapter?	152
 IV. TOWARDS A NEW TRAINING PARADIGM FOR VOICE	
PROFESSIONALS.....	158
1. Enacting the Expansive Imagination in Voice	158
2. Case Study: Contemporary Linguistic Needs for Latinx Actors and Directors	166
3. Pragmatic Answers to Utopian Questions	170
4. My Own Practice Recommendations for Dialect	173
4.1 Who Let this Dialect: Pre-production and Season Selection.....	175
4.1.1 A Note on Casting	182
4.1.2 Questions to Ask the Play (and Production Team)	184
4.2 The Heart of the Work: One-on-one with the Actor	185
4.2.1 Finally Meeting the Actor	189

Chapter	Page
4.2.2 Questions to Ask an Actor.....	193
4.3 Audience Outreach: Working Within the Community with the Dramaturg(e)	193
4.3.1 Questions to Ask a Dramaturg(e)/Community Outreach	197
5. Challenges that Remain, Where do We Go from Here?.....	198
APPENDIX	206
REFERENCES CITED	208

LIST OF FIGURES

Figure	Page
1. Histogram of responses by keyword, and again with keywords by speaker.....	129
2. Histogram of responses by keyword, by expectation condition	131
3. Box and whisker plot that shows the median accentedness scores for all four speakers	137
4. Box and whisker plot that shows the median comprehensibility scores for all four speakers	139
5. Box and whisker plot for accentedness with expectation conditions.....	140
6. Box and whisker plot for comprehensibility with expectation conditions.....	141
7. Comparison of expectation conditions for selection of speaker most likely to be the actor	142
8. Comparison of expectation conditions for selection of speaker most likely to be the most stereotypical speaker of Russian-accented English.	143
9. Box and whisker plots of the results from the Likert rating of all five adjectives, by speaker.....	146
10. Box and whisker plots for the five adjectives that comparison the two listening conditions.	150
11. Comparison of expectation conditions for selection of speaker most likely to be the actor in experiment 2.	151
12. Comparison of expectation conditions for selection of speaker most likely to be a stereotypical speaker of Russian-accented English in experiment 2.....	152

LIST OF TABLES

Table	Page
1. Four different listening groups in the experiment.....	127
2. Results of two alternative forced choice task.....	128
3. Accuracy of transcription of sentences for each speaker.	135
4. Accuracy of transcription of sentences by expectation by listening condition...	136
5. Adjective alignment on the Likert scales for experiment 2.	145
6. Mean and Standard Deviation for all five adjectives, by speaker.....	148

CHAPTER I INTELLIGIBILITY AND AUDIENCE MEANING-MAKING

“You are a linguist. You think everything is about linguistics.” - Julia Cho, *The Language Archive*

1. Overview

When performers or presenters speak to an audience that has gathered for the purpose of listening to what the speaker has to say, audience members generally expect to understand what the performer is saying. Indeed, this assumption on the part of the audience/listener is more fundamental to the performance event than any other concern, including whether the listener will agree with the message, enjoy the story, or (in the case of dramatic performances) have empathy for the circumstances of the character portrayed. Because of this expectation, performers create an extra effort to produce speech over and above what counts as understood speech in day-to-day scenarios. An entire constellation of professions that cater to the voice in performance cater to this seemingly simple and objective expectation of being able to understand speakers in public performance. Voice professionals have dedicated their lives to defining what it means to be understood in these public speaking contexts, and in turn can lend their expertise to speakers of all stripes, including public speaking for CEOs of large companies to actors in theater productions. This dissertation will focus on the latter group and examine the role that expectations of intelligibility plays in performance.

Dudley Knight, a practitioner of voice and speech, claims that a way to measure this understanding is with intelligibility, or a measurement of the amount of information that is communicated to the listener or audience member (Knight 20). This expectation of intelligibility appears at first blush to be an objective measure of communication;

common sense dictates there ought to be some objective threshold the speaker must meet in order to be understood by the listener. However, this logical idea about how speech perception might work does not necessarily square with the cognitive reality of the act of speaking and listening. An entire field of linguistics investigates this very assumption, and researchers have theorized that an objective threshold of understanding might not exist. In fact, audiences' expectations for understanding performed language is an under-examined microcosm of much larger social forces at play. According to current research in cognitive linguistics, intelligibility is not merely a feature of the speech itself, but a result of a combination of factors that encompass the speaker, listener, and the context in which speech is perceived (Bakanic 12). Similarly, this attention to listener or audience context is reflected in current audience research in both performance and theatre (Sedgman 103). For both fields, the relative privilege and cultural context of audience members, along with implicit attitudes about how language “ought to sound,” quickly starts to affect this seemingly objective measure of intelligibility. According to research, attitudes and prior held beliefs can predict the behavior of listeners, which poses a problem for the assumption that intelligibility of language onstage is somehow separate from the context in which that same language is presented.

Driven by the assumption that to speak on stage means an actor must deliver the text in a way that the audience understands the words (as well as the actions), performers are often asked to “speak clearly.” (Indeed, this may be the most common note given to actors in all levels of acting training, voice classes, and production.) But what does it mean to be “clear” or “intelligible” on stage? Practitioners assume that a basic level of effort is required so that the audience may understand not only what is being said, but

how it relates to the characters and situations in the play they are witnessing. Audience members expect a basic or privileged level of understanding, whether they label this quality of speech as “clarity,” “intelligibility,” or “authenticity.” Audiences assume that performed speech has an objective set of linguistic forms (e.g., volume, enunciation, breath). The question of what and how audience members understand what they hear is at the heart of a debate in theater praxis, and, more generally, in the field of linguistics. Both fields examine closely the role that expectations and cultural contexts play in perception for audience members. These expectations of properly intelligible speech are prescriptive in their approach, which measures the speech of performers to some predetermined ideal approach. Often, prescriptive expectations of ideal speech that mask as mere descriptions of linguistic forms are still subject to the overarching ideologies and beliefs of society. Performed speech, which is a type of communication that is limited to certain venues and contexts in society, is not immune to these expectations.

The social makeup of performers, characters and audience members combine to influence the social factors constructing intelligibility. Specifically for theatre, this expectation of intelligibility is directly connected to the types of bodies that have had the privilege to be onstage. This means the practice of western theatre has constructed intelligibility around the white body. Historically, in the United Kingdom and The United States, practitioners have privileged a certain sound or way of speaking that was indicative of a certain race, class, and gender of speaker (Skinner ix). This privilege contributed to the continued marginalization of theatre practitioners in most arenas of performance, with overt discrimination against those who do not speak close to the

standard accepted language in performance. Those who do not sound like privileged white, middle-class, male cis-gender actors from the correct part of the United States or England were judged as speaking an inferior type of dialect and were not featured on the stage. Indeed, actors and theatre professionals, like new broadcasters and politicians, are often encouraged to work to lose their regional or cultural accents in order to conform to a normative sound. The results of this exclusion reinforced the types of actors invited to train in the various official institutions of theatre. Those who were able to overcome institutional barriers found hostility from their voice and acting teachers. For example, Stan Brown, an Associate Professor in Theatre at the University of Nebraska, Lincoln, recounts his early experiences with a voice teacher while training to become a voice teacher himself,

As a young acting student, I was told by one of my voice teachers that the English language didn't belong to me. I am an African American. Her exact words were, 'Well Stan, you know the English language doesn't really belong to you...your culture.'... my initial response to her assessment was one of total silence and stillness (Brown 17).

Such overt racist ideas stem from this expectation of total intelligibility in performance yet are irrefutably affected by prejudiced ideas of the culturally dominant ideal English speaker.

Contemporary voice and dialect professionals have inherited these overtly racist ideas in their practice and continue to grapple with both the legacy of these origins and the expectations of audience members for intelligibility of actors on stage. Both ideas of elocution and audience expectations of intelligibility spring from the deeply held ideas of

a society that privileges certain language usage over other usage—and by extension, privileges users of more standard language, determined by voice professionals, over non-standard or colloquial language users. This means that vocal professionals, in pursuit of presumed objective measures like intelligibility or clarity, reproduce harmful structures of language classification that help perpetuate bias and further enforce the flawed idea that some speakers do not get to claim to be users of a language. Similar to the idea of the white gaze, this raciolinguistic perspective is “attached...to a listening subject who hears and interprets the linguistic practices of language-minoritized populations as deviant based on their racial (or socioeconomic) position in society as opposed to any objective characteristics of their language use” (Flores and Rosa p. 151). The focus of both the theorizing of these raciolinguistic practices and the bulk of the work of this dissertation is on the *listening subject* (both the highly trained listening subjects in voice and dialect trainers, and the listening subject of the average theatre audience member) rather than the empirical practices of *speaking subjects* or performers.

The flexible and social nature of language perception and standard language ideologies in listeners combine to produce real-world consequences for speakers who do not speak standard dialects or speak their second language with an accent of their first language. While an audience may perceive as harmless an actor using a foreign-accented English dialect to punctuate their villainous character, these types of social expectations that accompany accent perception in real-world scenarios can perpetuate harm. This harm comes from the stereotyped ideas that non-standard speakers are somehow inferior to speakers closer to the overall societal ideal (read as white) speaker. As this dissertation will demonstrate, in the same way explicitly racist ideologies influenced strict

gatekeeping in early elocution approaches that morphed into practices throughout the voice profession, these insidious social expectations influence everything from courtroom proceedings to education and many other contexts of language usage. Linguists have theorized that harmful social stereotyping as a result of normative language expectations has influenced perception in many different social continua and can lead to consequences that extend beyond initial judgments about these accents. Real-world consequences are documented in many different linguistic studies. For example, in courtroom proceedings, perceived accent affects the credibility of eyewitness statements in the courtroom (Frumkin 317). Six videos of identical eyewitness testimony, varying by accent and ethnic background of the eyewitness, were presented to participants. Listeners perceived speakers with non-native and non-standard accents as less credible than native speakers with more standard accents, over and above ethnic background of the speaker alone, and these speakers were also more likely to be perceived as deceptive. Further, listeners were overall less accurate in their recollections when they perceived accented eyewitness testimony. Theatre, in these cases, can offer a site for intervention that disrupts these negative views, and offers audiences a chance to re-configure their perceptions about non-standard language speakers.

It follows, then, that practitioners in voice and dialect ought to consider how audience members who are seeing (and hearing) performance in which language is a central participant are using their biases to construct intelligibility, and therefore meaning and experiences they are drawing from the performed event. However, the way voice professionals are constructing intelligibility does not often take audience expectation and meaning creation from these expectations into account. Overwhelmingly, voice

professionals rely upon their expert experience and predictions of audience intelligibility to guide how they train theatre performers. The work of this dissertation demonstrates that this assumption contributes to the continuing use of standard language as a guide for voice training. Understanding how audiences use their own context and expectations to construct intelligibility as a subjective measure is an essential piece of updating the training model for voice professionals, where such understanding will create space for performers and speakers who historically have been excluded from performing.

This dissertation opens audience expectations of intelligibility and standard language usage to assess the subjective nature of this measurement of speech. This dissertation will combine a critical history of the voice profession, a cognitive consideration of how audiences use social context and expectations to create meaning on stage, and a specific empirical linguistic case study to specifically build the argument that audience members, while perceiving intelligibility as an objective reality of language, create their own subjective parameters around language perception on stage. I will demonstrate that because these parameters are sensitive to social context, intelligibility and speech perception can be used in thoughtful ways to counteract the overarching societal expectations around the use of standard language, denying a large portion of speakers who do not share characteristics with those historically who have had control.

1.2 Motivation and Structure

I propose a direct challenge to the objective view of intelligibility as a space for the audience and practitioner to explore their own linguistic biases, which privileges the white, male cis-gender middle-class listener as the preferred audience member. Examining cultural norms regarding language in this way opens opportunities and

possibilities in theatre and entertainment production that allows for speakers who might not otherwise have a chance of participation in the process of production. The role of the audience member in constructing intelligibility for the stage is a site to challenge the idea that the way a person sounds immediately indicates or confirms their innate character traits. Often when a speaker is referred to as accented, with rare exceptions, the context is negative. This pernicious idea also figures heavily in the profession of accent reduction or modification, which is a clear example of the burden of expectations of understanding lying squarely on the speakers, while absolving listeners of their responsibility for their role in constructing intelligibility. Approaches for training, both in performance and in real life, advocate for an appropriateness-based type of language education, claiming there are appropriate avenues for different types of speech. These models of training “advocate teaching language-minoritized students (or performers) to enact the linguistic practices of the white speaking subject when appropriate” while denying that the white listening subject or audience member may still continue to perceive linguistic markedness even with the best training that is available to the student or performer (Flores and Rosa 149). Performed language qualifies as a type of privileged language and thus is highly susceptible to this type of appropriateness-based training. Performed language must therefore be conceptualized not as objective linguistic categories but as a set of racialized ideological perceptions perpetuated by those who create and maintain the structures of power in formal entertainment structures in stage, television, and film. While performed language appears in many forms, from public speaking to formalized performances, the work of this dissertation will focus on the formal voice profession structures that have

arisen to support professional forms of performed speech—found in professional theatre, television, and film.

This dissertation also spans between fields serving as a model of interdisciplinary inquiry that takes seriously the notion of approaching a problem from both a humanistic and a scientific point of view. Theatrical performance remains a uniquely structured human behavioral activity, which means the practice is an arena rife with the possibility of scientific and linguistic inquiry. My position as an artist and scientist allows me to see a problem or assumption in one discipline such as audience perception of language onstage and approach the problem as both a theatre and a linguistics scholar, creating significant contributions to both fields in the process. For example, this dissertation not only holds the potential to influence pedagogical approaches of voice and dialect practitioners, but also stands to contribute significantly to conversations in linguistics and cognitive science about social models of speech perception more generally. Answering this question of audience expectation of understanding feeds directly back into best practices for voice professionals to confront the overtly racist ideas that serve as the foundation of this practice, as the results of examining this phenomenon feeds directly in the form of a practical guide for theatre makers.

One of the core tensions of this dissertation is disentangling the difference between knowledge that is considered objective versus subjective experience. In my own experience, appealing to objectivity for a socially constructed phenomenon such as intelligence or intelligibility automatically reinforces built-in biases of evaluators in mild cases, and creates active harm for those evaluated in the most severe cases. Specifically, I am interested in the issue of intelligibility (in both senses as linguistics research and voice

practitioners) and how it can be considered an objective measure, when the substance that intelligibility is measuring is as nebulous as information or content of the speech, when human speech carries more meaning than the words that are uttered. Intelligibility used by these practitioners is, in my estimation of the literature, a synecdoche for the perceptual experience and subsequent quality judgment of the expert listener.

Intelligibility, for a voice or dialect coach, stands for their estimation of the perceived understanding of the dialect in the context of performance (Knight 72). For the sake of empirical research, linguists have been able to side-step that larger epistemology by creating a working definition of intelligibility that quantifies words understood and recreated by listeners (Flege 2020). In consideration of the objective, voice teachers have also adopted this approach to a degree, yet do not measure content produced as narrowly as asking each individual audience member to regurgitate the content that they have just witnessed. If we were to recreate this experimental paradigm in the theatre, imagine participating as an audience member in one such experiment, with a researcher asking you to rewrite each line as you witnessed *Midsummer Night's Dream* by William Shakespeare.

My goal with the research in this dissertation is to use the definition of intelligibility and the constellation of similar terms narrowly and precisely from linguistics to explore the more common denotations of intelligibility in the voice profession that uses this term as a marker for success in theater. Carefully parsing terms will reveal a gap in the assumptions that govern the voice profession and give clarity to existing linguistic literature on intelligibility. The linguistic experiments I have designed for my investigation initially aim to isolate the moment of perception of the average

audience member when they encounter a speaker or a voice minus the richly complicated immediate context of live performance. This approach requires the explicit assumptions that the basic auditory, visual, and linguistic perception in individual audience members continues to function as normal, even in a highly specialized context and environment (McConachie 34). From this base, I can then expand the paradigm to capture richer and more complicated contexts in which these voices are heard. In the *linguistic* sense, intelligibility and the accompanying factors audiences use to judge an accent will provide feedback or confirmation for the concrete goals for the expert listener.

To examine this social construction of intelligibility, I must approach this topic using two main threads of inquiry. The first thread asks, what are the assumptions of audience understanding guiding the principles of voice and dialect coaching? How do these assumptions shift from their historical origins to influence the contemporary profession of voice and dialect? To answer these questions, I create a short critical history of this relatively young profession in theatre making, answering these questions for different eras of voice and dialect. Interspersed with this critical history is an examination into how audiences construct meaning onstage, more specifically theatre found on educational and professional stages in the United States. How does the audience conception of voice (promoted and influenced by key players in voice and dialect training) and intelligibility contribute to this meaning making process? This first thread draws the answers from the legacy of American Realism, the establishment of actor training in the United States and the United Kingdom, and Ireland draws upon both historical and contemporary cognitive audience reception/perception research. A central theme of this exploration is the idea that words uttered onstage in their rich linguistic and

social context of performance are *doing things* far and above the base meaning of the assertions found on the page or in performance, providing each audience member the opportunity to create meaning for themselves (Austin 6). I will use the inheritors of J. L. Austin's ordinary language theory, George Lakoff's and Mark Johnson's embodied realism, to critique the often-detached ways that historical practitioners have conceived of language and perception in their pursuit of the profession of voice.

The other thread builds upon the theories of meaning-making presented in the previous thread by using a very specific empirical case study to advance our cognitive understanding of audience perception and how they use intelligibility to construct meaning on the stage. This thread examines in quite literal terms the assumptions of listeners when they encounter performed speech. In this thread, I ask, what cognitive processes are audiences accessing in the moment when encountering performed accented language? How do those expectations affect how speech is perceived, specifically in terms of intelligibility or clarity? I build a brief and useful linguistic primer for practitioners before describing in detail the experiments and their results, ending with a collection of takeaways that will influence the conclusion of this dissertation. The conclusion brings together the prior cognitive audience research and this specific empirical linguistic inquiry to speculate on the future of the voice profession, offering best practices informed by the experiments conducted within the body of this dissertation. Mirroring other efforts in updating representation on stage, I speculate towards a new field of voice and dialect coaching that critically grapples with audience meaning making as a reflection of the cultural context of theatre production in a society that explicitly privileges some language varieties over others. A critical examination of the underlying

meaning-making processes of audiences means we as practitioners may unlock the potential to intervene in harmful stereotype creation and reinforcement of dominant varieties of language. A re-tooling of voice and dialect and its role in theatre creation can expand the notion of who gets to sound like whom onstage, creating room at the table for an expanded variety of lived linguistic experiences.

2 Literature Review: Different Disciplines, One Room

This dissertation draws its foundation from two disparate fields to investigate the role of intelligibility—one humanistic, the other scientific—as both fields contain unique approaches to knowledge that complement each other. Both bodies of knowledge, and ways-of-knowing (one aesthetic, the other empirical) are crucial to deciphering the specific ways in which voice professionals inadvertently reinscribe biases and standard attitudes, and to suggest proactive countermeasures to mitigate the damage already done by the profession. I begin in the first thread by providing a history of the profession of voice and dialect, that critically engages with the ideologies of over one hundred years of voice professionals, ending with recent prominent voices through the professional organization Voice and Speech Trainers Association (VASTA). Along with the voices of professionals themselves, I describe the material conditions of theatre production as a possible source for the shape of the profession. Voice practitioners have, from time to time, pondered the ethical considerations of their craft, with fewer still producing academic literature. One publication, *Standard Speech: Essays on Voice and Speech* (2000), printed as the initial issue of the Voice and Speech Review by VASTA, has served as the model for discussion about ethics of producing voice and dialect work. Precious few of these voices in the literature of voice professionals are concerned with

audience perception and reception, which is the conversation into which I inject the research of this dissertation. I build upon the implicit arguments of Patsy Rodenburg work against what she calls “vocal imperialism” in her first book *The Right to Speak* (2015).

The history of this profession is offered through a lens of philosophical approaches to language via cognitive audience studies that will roughly divide the different approaches to voice and dialect training throughout the brief history of this craft. In order to engage with this history, I am using two prominent audience listening theories. One area is audience reception theory, introduced by Susan Bennett’s book *Theatre Audiences: A Theory of Production and Reception* (1997), and the other is contemporary iterations that have followed other humanities scholars and adopted a cognitive turn towards scholarship through Bruce McConachie’s *Engaging Audiences: A Cognitive Approach to Spectating in the Theatre* (2008) and other scholars that are already borrowing from psychology, linguistics, cognitive science, and embodied realism to explore the processes of audience reception/perception. Bennet has essentially laid the groundwork for future cognitive humanities investigation into theatre by arguing for deeper systematic research that accounts for different contexts for audience reception (89).

The second thread answers Bennet’s call for deeper systematic research by examining a very specific instance of audience members constructing intelligibility through speech perception of non-native dialects on stage. To establish this thread, I examine the up-to-date theories of language perception that voice practitioners may directly use in training. I will specifically draw upon Rosina Lippi-Green’s work *English*

with an Accent: Language, Ideology and Discrimination in the United States (2012) where she uses empirical evidence to demonstrate what she calls “standard language ideology,” a phenomenon affecting listeners as they reconcile their expectations of how a speaker should sound based upon visual cues such as perceived race or gender (Lippi-Green 64). This standard language ideology is particularly pertinent to performed language as the language is often spoken by an actor in a space who is read in multiple social levels (race, age, socioeconomic status, gender). I add to this by summarizing the most recent findings in the field of cognitive linguistics, by summarizing the most recent theories of non-native speech perception, second dialect acquisition, and speech adaptation. I start by specifically drawing upon the work of Kevin McGowan (2015), Donald Rubin (1990, 1992, 2013), James Flege (1995, 2020), Munro and Derwing (1997), and the work of the very lab of which I am a member and where I have conducted the research of this dissertation, under the direction of Dr. Melissa Baese-Berk.

To view these two threads as an intertwining braid, I will draw parallels between these two fields of study by highlighting overlapping vocabulary terms used in both fields. By defining terms such as *accent*, *perception* and even *intelligibility*, I can create a space with this dissertation that combines parallel conversations and offers best practices as a result of these overlapping fields of study. I will use the main takeaways from both threads to directly advocate for a better approach to a profession that has historically contributed to active harms of marginalized people.

2.1 Thread one: The foundations of the voice profession

The first thread concerns itself with a thematic historical overview of the voice and dialect profession, beginning with elocutionists at the turn of the twentieth century

and leading to the working model of contemporary practitioners that lend their services to many forms of entertainment, from live theatrical performance to film and television. Vocal professionals work in terms of meaning-making by teaching performers to access their voice as an integral aspect of conveying language on stage. Most theatrical audiences expect to be able to easily understand these performers, which is an aspect that voice professionals have identified as a key area for voice work. In this thread, I connect the expectations of both voice professionals and theatrical audiences to raciolinguistic ideology—both of these behaviors stem from the inferred expectation of the white listener, which automatically and systematically labels the marginalized voice and body as an Other, leading to real-world consequences for performers and marginalized communities. The chapter will demonstrate that even contemporary voice training ascribes to the appropriateness model of education, popular in many different language education models. The arguments in this thread preview the notion that the so-called objective listening criteria offered for performed speech is subjectively constructed between the speaker and the listener (i.e., the audience). The training apparatus of theatre is, indeed, the white listening apparatus and the accompanying economic system made manifest.

In addition to conveying the written language of the piece, vocal choices the performer makes also do work to create meaning or context of this language. For example, an artistic director for the company may choose to hire a dialect coach for a production of *Good People* and ask that actors in the production perform using a Boston dialect, to convey a sense of meaning of place in production. Regardless of location of performance, audiences take this acoustic signal as part of the process of meaning-

making in theatre. However, use of a Boston Southie dialect on the West Coast of the United States may be interpreted differently than audiences on the East Coast (especially if that playhouse is in Boston). The audience's threshold for dialect accuracy may be wider for those on the west coast of the United States since audience members have likely less direct experience with a Boston dialect in the real world outside of the performance. This difference in perception of authenticity between individual audience members affects how they are perceiving the show. Even with a stereotypical dialect, audiences expect ease of access to understanding the language onstage as a seamless part of meaning making in theatrical production, calling for an accent that is both recognizable as authentic but clear in delivery. However, clear in this instance is defined by the narrow experience of the small slice of socioeconomically privileged demographics that Susan Bennett in *Theatre Audiences* demonstrates attends the theatre (Bennett 114).

Practitioners have established authority in this growing field by opting for engagement with more general audiences for their scholarship. These practitioners cultivate access to their work not only through creating institutes and offering workshops, but to espouse their philosophies and approaches to voice through publications that are meant as accessible guides for speakers and performers of all stripes. In these general audience publications, some practitioners grapple with the built-in inequalities of the profession of performance. In her 1993 book, *The Right to Speak: Working with the Voice*, Patsy Rodenburg writes about the right for every person to have their voice heard through her system of work of the voice. This is a simple declaration, yet to make such a declaration requires an honest admission of who has historically had the right to speak

and in which arena, and who has not. In her introduction, she introduces the notion of “vocal imperialism” with respect to the notion that when a person opens their mouth to speak, they are exposed to snap judgments from others of their geographical and socioeconomic origin and subsequently their capabilities as speakers (Rodenburg 5). In this society, certain voices are privileged above others as a result of overlapping prejudices and beliefs, which is then reflected in the media, performance, and entertainment that this society consumes. Rodenburg warns against how these types of snap judgments lead to a loss of voice or vocal power in a speaker. Rodenburg contributes a culturally sensitive addition to the usual idea of vocal practice as being a deeply unique and individual practice that aims to investigate the speaker’s own limits and restrictions. The fact that she names “vocal imperialism” as the first obstacle to declaring one’s right to speak reveals volumes about the society in which speakers and listeners find themselves.

Naming these biases forms the basis of inquiry into how practitioners and audiences alike conceive of intelligibility, whether it is called vocal imperialism in the field of voice or standard language ideology in the field of linguistics. This deceptively simple belief in a standard language continues to hold massive implications for the use of dialect onstage. The fault alone does not rest in this field of voice; belief in standard language is a pervasive, common-sense idea that is deeply and subconsciously ingrained in how most of society views communication. However, the field of voice and dialect is in the unique position to actively push against how standard language beliefs affect individual audience members’ perceptions of how intelligible and clear an actor must sound in order to be accepted as a good performer or speaker. However, these

expectations are not only the responsibility of the untrained listeners in the audience, voice experts' opinions and trained ears are also responsible for shaping the conception and use of intelligibility.

While some vocal practitioners like Patsy Rodenburg push back against this prevailing idea, the fact remains that vocal and dialect training has been shaped by these biases towards an abstract idealized language. These expert expectations are responsible for appeals to standardization of language use on stage. This thread traces a critical history of the harmful usage of these language standards throughout the profession that still find their ways into contemporary linguistic practice onstage, in film, and on television. Standard language varieties often reflected a neutral mode of speech that was anything but neutral, as these varieties often favored white, middle class cis-gender actors because the variety acoustically most resembled these actors. (Lippi-Green 14). Often in these books from early elocutionists, these attitudes towards modes of speaking would give way to overt discussions of racist ideas like restricting immigration to the United States. Published in 1924, Marguerite DeWitt writes in *EuphonEnglish*, explicitly highlighting why she believed white American speakers of English were the most superior speakers in general, “ignorance may be condoned, lack of dexterity may be excused, but faulty speech and foreign accent are indelible signs of social inferiority” (DeWitt as quoted in Knight 40). Building from these explicitly racist structures, other voice and dialect professionals incorporated these prejudices and systemic inequities throughout the following century so that these explicitly racist ideas are now hiding as common sense or implicit approaches to vocal production. The use of “Good American Speech” as introduced by Edith Skinner in training, for instance, necessarily implies there

are versions of Bad American Speech that do not deserve to be heard onstage (Skinner 1990, ix). These effects are borne disproportionately by those actors with marginalized identities, and there are many instances where actors are encouraged to either drop their home dialect or play up the ethnic aspects of their speech in order to secure roles (Sullivan).

The professional manifestation of these practices includes the organization Voice and Speech Trainers Association (VASTA) that houses both practical resources like lists of trainers that are available in the institutional ecosystem, and also a scholarship wing where critical conversations have shaped the profession in its twenty years of existence. I engage here with two of the most influential voices throughout the history of VASTA and the publication *Voice and Speech Review* because I have seen their work influence the community of voice and speech trainers in lasting and damaging ways. With explicitly racist foundations that this profession has seldom acknowledged—apart from a growing contemporary call from a small group of scholars—I choose these voices as critical entries into understanding the historic legacy of racism and sexism in the contemporary iteration of this profession. Like voice and dialect coaches in production, these scholars are admired as respected authorities and gatekeepers of access to the voice and bear the responsibility to honestly reckon with the dark history and continued oppressive practices of this profession. Some of these influential scholarly and professional voices still defend the use of standardized dialects or Skinner’s “Good American Speech” as a pedagogical tool, contending the importance for actors to learn about their unique voices through learning a different accent or mode of speech (Robbins 55). While many voice professionals may not use standardized dialects or accents or “Good American Speech,”

there are other insidious ways where normative language attitudes seep into their practices. Again, while these normative attitudes are on the surface aimed at the marginalized speaker or actor, these attitudes spring from subjective ideologies that actually privilege the white listener.

One of the ways that voice and dialect practitioners have attempted to circumvent issues of standardization of language practice is to appeal to the objective sounding measure called intelligibility, as described at the beginning of this introduction. This is the phrase that Dudley Knight used in his article on standard language usage in voice pedagogy in the initial journal of *Voice and Speech Review* in 2000 that was so influential there was a reprint in 2012. His claim appeals to the commonsense notion that there must be some objective threshold for understanding sound and language onstage. Surely, Knight argues, there is some absolute baseline of minimum understanding from the audience when it comes to communication (Knight 65). The work of this dissertation dismantles this appeal to common sense and demonstrates that intelligibility is socially constructed; thereby demonstrating that no such objective threshold exists in the way conceived by Knight. Conveniently, Knight does not offer a direct definition of intelligibility, instead appealing to a know-it-when-you-see-it approach by saying, “most theatre accent coaches have a keen experiential awareness of what intelligibility is, because they have had to modify the accuracy of accent all the time to accommodate it” (Knight 75). Using appeals to authority is a common theme for voice trainers throughout the short history of this profession, which accomplishes two things—establishing this profession as having legitimate expertise, and gatekeeping speakers of non-intelligible accents from the profession. He then goes on to claim that intelligibility is fully the

responsibility of the speaker, “A standard based on indelibility is not tied to any prescriptive pattern. Rather is it based solely on the speaker’s ability to transmit to the listener the appropriate amount of linguistic information to the level of detail and specificity appropriate to the event” (Knight 75). This is a curious approach, because he appears to set up the responsibility for intelligibility to either the expert listener or the speaker, but not at the same time. The responsibility for intelligibility during rehearsal lies with the expert listener or voice coach, while intelligibility rests on the shoulders of the actor during performance.

Knight also appeals to this objective measure as a way to circumvent the idea that normative language attitudes about race and gender affect all forms of communication and especially language perception. Knight is referencing specifically the linguist Rosina Lippi-Green, who was also invited to contribute to the initial volume of *Voice and Speech Review* on Standard Language as an outside expert. In her article “The Standard Language Myth,” Rosina Lippi-Green explains a phenomenon with which voice and speech trainers must contend called *standard language ideology*—introduced briefly here but discussed more in depth in the following section—where a listener believes that a homogeneous or perfect version of language exists, and they are comparing what they hear with that expectation (Lippi-Green 24). In order to create a standard objective measure of intelligibility, Knight had to reconcile the subjective nature of standard language ideology, since every listener in the audience has the potential for a slightly different version of ideal language. His solution was to discredit Lippi-Green’s theory that listeners share responsibility for intelligibility, even going as far as accusing Lippi-Green for cherry-picking anecdotal evidence to support her case by offering his own list

of anecdotal evidence himself. However, Knight's own cherry-picked anecdota of exceptions all skew heavily white cis-gender male (who could initially be perceived as standard speakers by their looks alone) which ironically serves to confirm Lippi-Green's theory that listeners are using social cues from their beliefs of standard language to construct intelligibility. It takes incredible hubris to criticize a linguist, whose entire job is to systematically investigate language usage, as picking and choosing evidence while ignoring others in building a theory of language usage for voice professionals. Knight's mistake in 2000, and since reprinted in 2012, was refusing to take into account the audience member's role in measuring intelligibility of speakers onstage.

In the intervening time since 2012, Dudley Knight has since softened his appeal to intelligibility as a wholly objective measure, describing intelligibility of the work he offers on his website as "not a fixed property of some idealized and prescribed accent model, but a constantly negotiated process between speaker and listener, within conditions set by the acoustics of the space and the familiarity of the audience with the language style" (Knight and Thompson). However, they still suggest the use of objective measure through referencing the acoustics of a given location, along with an implied ability to measure the audience's familiarity with the language style of the piece. For example, familiarity itself cannot be measured as a fixed quantity, since the audience's experiences with the piece itself affects familiarity proportionate to the amount of time they spend experiencing the production. In other words, the audience's familiarity with the language styles will increase with every minute of the production.

I engage with Knight's arguments here in detail to make a point about how voice and dialect practitioners approach many kinds of objective or scientific measures. That is

to say, the history in the first thread will reveal that Knight and other voice and dialect practitioners appeal to scientific authority on topics by using objective-sounding measures without following through on *how* to use such measures. The use of intelligibility as an objective measure of linguistic ability can be compared to the use of IQ tests; at the outset, there appears to be objective measures of intelligence, but scratch beneath the surface and one can find many instances of normative and racist judgments that accompany these types of measures (Kendi 311). To balance this appeal to objectivity, as Knight often does in his article, voice and dialect practitioners appeal to their extensive and subjective experience of their profession. As a voice practitioner myself, I do not have a large issue with using personal experiences as evidence *per se*. I do, however, criticize when subjective observation is then passed off as objective or scientific evidence. Fortunately, other practitioners in this field of voice approach the issue of standard language through a lens that accounts for cultural differences in practitioners and audiences, laying the foundation for an alternative discipline of voice training.

This deep interrogation of the foundation of this field is necessary if we are to include voices that have been pushed aside, voices that continue to be marginalized in both society at large but also in our performance spaces. After all, as Rodenburg says, “voice work is for everybody...your voice belongs to you, it is your responsibility and right to use it fully” (Rodenburg xiv). A new generation of voice professionals have taken this quote to heart, as they create systems that attempt to open space for speakers who do not speak perceived standard varieties of English or other languages. Professor Melissa Topping-Kollwitz and Joe Hetterly have pioneered work both in the professional sphere

and through their continued scholarship of industry needs and the documented shift away from standard dialect in *The Voice and Speech Review*. From these and other professional observations, Tonning-Kollwitz has crafted a new system of voice work that accounts for biases, and she has adopted actively antiracist stances within the entertainment industry. Another voice professional Daron Oram advocates for a very explicit decolonization of the “linguistic imperialism” described by Rodenburg (Oram 280). Contemporary voice and dialect professionals have begun the difficult work of transforming this profession through their own years of negative experiences. While these viewpoints are previewed in the initial thread of inquiry, the final chapter of this dissertation is in direct conversation with these contemporary scholars and practitioners.

Interspersed with discussion of these voice practitioners and different eras of voice training are cognitive considerations of how trainers and audiences alike are constructing meaning using the social contexts of the voice. Part of this thread is a consideration of the use of the word “voice” and the many different permutations that govern this profession. This chapter also offers cognitive explanations behind the assumptions—established as common sense principles—of practitioners in the field of voice. Through audience studies, cognitive humanities, and the philosophy of language, I can examine why these assumptions feel common sense, and the explicit role intelligibility of speech plays in the larger context of meaning creation in production.

Interest in audience studies, also known as audience reception, did not rise to scholarly prominence until Susan Bennett laid the foundation for a compelling case for studying audience reception systematically within theatre production in 1989 through her book *Theatre Audiences: A Theory of Production and Reception*. In Bennett’s estimation,

theatre, while beginning in the west as an act of communal religious and political gathering for most of the populace (an embodied act or performance of democracy), has shifted instead to serve the interests of the middle class or bourgeois society (Bennett 3). "Naturalist theatre," which is what Bennett labels Realism in the American tradition that arose out of the work of Stanislavski, can be the culprit most guilty of catering to the middle-class tastes of these audiences. Bennett observed that treatment of audiences as homogenized and sanitized masses did a massive disservice to the types of theatre that was pushing against the dominant or mainstream approach to theatrical production. These theatre companies, often sharing and promoting work created by marginalized theatre creators, often approach theatrical spectating as a group of individuals who are "productive and emancipated spectators" as part of a vibrant cultural ecosystem (1). While an imaginary or stage world is still at the center of the model, it is the stage world that is concentrically wrapped in audiences' cultural expectations that constitutes audience reception. The cultural context in which theatre is created belongs directly to the audience members who are witnessing these acts of performance; the performance by its very nature cannot be separated from the expectations of those who witness it. Thus, according to Bennett, to study performance is to study the audience and the various contexts in which they encounter theatre. Audiences make meaning not just from the imaginary world on stage, but the real world in which they find themselves encountering this act.

Because of the serious consideration of these environments, audience reception studies makes room for empirical consideration of the audience experience and thus engages with different fields such as sociology, anthropology, and even philosophy of

language. Bennett builds a foundation for her model by reviewing the available empirical studies on theatrical audiences, which all confirm the initial conceit of the introduction; namely, audiences polled who attend performances occupy a narrow swath of socio-economic and political opinions, regardless of the geographical area that is sampled (Bennett 92). While often sharing dominant societal habits and attitudes in general, this small band of politically like-minded middle-class theatre attendees also share similar held ideas about expectations of intelligibility's role in meaning on stage. To this group, theatre and art for the middle class were made for relatively easy consumption, which requires an access to understanding with little to no effort on the part of the theatrical audience. Theatre has been built for this particular type of white listener, which is a subject that is found in other "appropriate" contexts, including academic and classroom instruction of language (Flores and Rosa 145). Contemporary audience research scholars such as Kirsty Sedgman follow this tradition of empirical investigation, calling for rigor in theatre audience research that rivals serious social scientific inquiry in other fields. This dissertation answers that call by posing uniquely specific empirical questions of audience perception in the theatre, by incorporating research from the field of linguistics and detailing a series of empirical experiments that advance cognitive audience studies.

Bennett borrows heavily from semiotics and philosophy that regards humans in the tradition of Locke and his "tabula rasa" rational mind, while still attempting to rectify the immediate and material influence of context on that mind. In contrast, Bruce McConachie uses recent cognitive scholarship to embrace both nature and nurture in the quest to describe human meaning-making in performance. McConachie points to actor training as a huge influence on the mode of meaning-making for theatrical audiences,

particularly in the United States of America and the United Kingdom and Ireland. I will take up the arguments from McConachie and demonstrate that voice and dialect coaches are responsible in large part for replicating and disseminating ideas and modes of meaning-making for the voice and the use of dialects onstage, and that this, in turn, replicates social bias. The context of Realism or “naturalist theatre” in which audiences have grown accustomed to encountering theatre governs the practice of voice and dialect directly (Bennett 2). Both audiences and theatre producers in particular are subject to this naturalist approach to the dominant epistemology of theatre creation in the twentieth and twenty-first century. As the voice profession that was created in the twentieth century, reasoning extends that the voice and dialect profession ascribes to these same ideas about knowledge creation. This thread closely examines these ideas about knowledge creation through distinct generations or approaches to voice training. This thread not only offers an alternate lens into the guiding assumptions of voice and dialect professionals, but I also lay the foundation for a very specific application of empirical investigation into the exact construction of intelligibility in which audience members participate.

2.2 Thread two: A cognitive account of how audiences construct intelligibility

One of the unique contributions of this dissertation is the critical engagement with linguistic theories that parallel discussions that scholars in theatre are having regarding the role of communication. This thread will serve as a specific instance of cognitive humanistic inquiry that details the cognitive mechanisms behind audience perception in order to assess its role in meaning-making for performance. Since the focus of this dissertation is on audience experience and perception of voices and accents on stage, the best expertise that is available on the topic arrives directly from psycholinguistics and

related fields, where researchers have examined the perceptual, mechanical, and psychological processes that accompany language perception more generally. Sociolinguistic theories in non-native or foreign accent perception also serve as a critical basis for investigation of standard language ideology, which is a root cause of the historic and continued use of standardized dialects or accents in actor training. Also found in sociolinguistics is a very useful model with which to frame various approaches to voice and dialect training called *raciolinguistics*, a model that describes how speakers are unfairly held up as flawed examples of the language or accent they are attempting to acquire, and the prevailing belief that some forms of language have appropriate avenues and are only appropriate during certain contexts. This dissertation acts as a bridge in multiple ways, connecting conversations between fields as wide apart as the social sciences and the humanities, and within discipline between subfields of linguistics. To decipher the ways in which normative vocal training may reinscribe implicit bias and inadvertently prejudice the entire practice of theatre, I make use of ideas and terms from the field of linguistics. While these are not always credited to one particular author (as is often the case in the humanities), these terms are central to my project, and indeed they are the tools I apply reflectively to the field of vocal training for the theater. I begin here with the term at the nucleus of my inquiry, standard language ideology, and branch to the theory and experimental evidence from multiple scientists throughout this literature review.

Much like Rodenburg's "vocal imperialism" Rosina Lippi-Green introduces the term "standard language ideology" in her book *English with an Accent*. This term is defined as "a bias toward an abstract, idealized, homogeneous language" (Lippi-Green

64). The use of standard language ideology captures the audience belief that there is an idealized or perfect language to which speakers must measure themselves. This bias appears both in positive feelings towards the “idealized” language variety and in negative feelings towards non-standard or non-native varieties of language. For instance, many listeners register positive feelings towards prestigious accents such as British-accented English, or French-accented English, and more negative attitudes towards Spanish-accented English, especially varieties from countries other than Spain (Lindemann 187). This belief itself is often the root cause of linguistic discrimination of all stripes; these beliefs do lead to discrimination and consequences in real-world scenarios that can lead to loss of economic and cultural opportunities. For example, listeners will rate a local Catalanian dialect as less trustworthy than a standardized Spanish accent when listening to speakers on the radio, leading these listeners to disregard important information (Renaires-Lara et al. 16). Throughout this dissertation, I will compare “vocal imperialism” as described by Rodenburg with Lippi-Green’s “standard language ideology” as proposed in the field of sociolinguistics. The parallels are so striking, I believe that both scholars are talking about the same phenomenon through different ways to access that knowledge. Rodenburg coined “vocal imperialism” using her years of personal embodied experience of ushering hundreds of students towards vocal freedom in performance. Lippi-Green points to the plethora of empirical evidence found by numerous linguistics researchers and scholars as the basis of her theory. The goal of this dissertation is to create conversation between these different modes of knowledge, highlighting the diverse paths that one can take towards discovering the underlying assumptions of a profession such as voice and dialect.

In some cases, standard language ideology can reinforce an accepted, explicit standard dialect that has been championed by those in power. In the United Kingdom, Received Pronunciation, a region less and constructed dialect championed by the British Broadcasting Company and encouraged to be used by its on-air personalities, is favorably associated with competence, education, self-confidence, and intelligence (Brown et al.). Received Pronunciation is a dialect that has been created for performance and for media, yet still influences listeners' language attitudes about how English, at least in the media, ought to be pronounced. Audiences tuning into the evening news expect anchors and reporters to be easily understood and expect a high level of intelligibility. Efforts by dialect coaches and voice professionals in the United States of America have attempted a type of standardized dialect similar to Received Pronunciation for performance, with one dialect, sometimes named Mid- or trans-Atlantic, becoming a popular dialect with which to train actors (Skinner). Actors and performers trained for both stage and television have long been taught standard dialects with an eye towards prestige, reinforcing who gets to sound like whom onstage. These dialects and the prestigious institutions that created them also enforce ideas about what accent is right or standard in any given culture, leading to a kind of feedback loop that affirms the confirmation bias of audiences and practitioners alike. These reflections are so prevalent, one can actually trace the changes in prestigious forms over years. Performed speech may reflect the standard or idealized speech of the dominant time in which the media is produced (Elliott 105).

Earlier generations of voice and speech practitioners claim that the speech of actors (especially explicitly trained in dialect or voice for the stage and film) can represent an ideal or standard style of speech to which all speakers ought to conform.

Further, standard language ideologies of non-standard dialects and accents on stage influence stereotyped representations, where representations of accent and dialect in media can be conceived as cognitive shortcuts for characterization of the characters (Bakanic 14). All of these forms of language package the idea that certain types of speech are more appropriate for public-facing contexts, presenting these styles of speaking as objective linguistic fact, when the bulk of the identity of these styles rests in the listener. Both the speaker and the listener are creating these styles of speech through cooperation—the speakers' continuous use of these styles, and the listeners' continued expectation of these styles. To examine how these styles are built without trainers constructing this interaction misses a large chunk of the story, which will be the focus of this thread.

The field of accent perception offers the strongest rebuttal to Dudley Knight's arguments for an "objective" measure of intelligibility. There is a significant amount of prior research that has examined the factors that affect accented speech perception previewed here and explored in depth in this thread. The three factors of accented speech perception that many researchers use to build their theory of accented speech perception are accentedness, comprehensibility, and intelligibility. Accentedness is a subjective measure that refers to how "strong" a listener believes an accent to be. Comprehensibility is also a subjective measure, which asks the listener how easily they can understand the speaker. Both of these first two factors are measured on a Likert scale while asking for the opinion of the respondent. For example, respondents giving their comprehensibility opinion will be asked "how easy was it for you to understand this speech?" and given the

choice between 1 as “not at all easy to understand” to 9 “extremely easy to understand” (Derwing and Munro 1).

In contrast to accentedness and comprehensibility, intelligibility is measured in a different matter. Linguists refer to intelligibility as an explicit measure of the amount of information a listener gathers from the speech signal. However, instead of being left to the impression of an expert listener such as a dialect coach, researchers devise paradigms where listeners are expected to write down the exact words that they heard from speech. In this way, whether or not the listener guesses “high” at the end of the sentence “the ball bounced very high.” becomes a matter of objective achievement. Crucially, however, these researchers do not only include this measure of intelligibility in their investigations into speech perception. Research in this area operationalizes and uses intelligibility as one of many factors to examine the cognitive processes behind accent perception. These three factors, including intelligibility, are highly sensitive to different modes of context. Many factors often determine these scores, including factors intrinsic to the speaker, intrinsic to the listener, or related to the environment in which the language is perceived (Moyer 192). Of interest to this particular research is the environment, the factors affecting perception of accented speech are not fixed within a listener, as these factors can be influenced by the context in which the speech is being perceived, including expectations of the listener (Kang & Rubin 441).

There are many specific empirical instances where expectation affects perception of non-native accent, which can point to theatre or performance being a specialized social context for language perception. For example, listeners’ perception of the vowel space that speakers use is sensitive to explicit labeling of regional dialects on testing materials,

shifting listeners' perception depending on the regional label (Niedzielski 80). Explicit mention of geographical areas may not be necessary as listeners are so sensitive to their environments that they can be affected by relatively minor influences on the testing environment (e.g., stuffed toys in the testing area, Hay and Drager 867). That is, listeners' perception of vowels was influenced by the toys the listeners had seen before the experiment, driving listeners to label these vowels as originating from different regions, even as they heard the *exact same vowels* in each experiment. This means that audience members are sensitive to all types of cues in the performance environment, from their expectations regarding the bodies that are onstage, to the decisions that designers make for costumes, set, lighting and general ambiance.

This sensitivity can lead mismatched expectations that can also have perceptual consequences. In other work, listeners were less accurate in transcribing information when they experienced a mismatch between what they were seeing and what they heard (e.g., seeing a photo of a white woman while hearing Mandarin-accented English, McGowan 515). This evidence supports a model in which linguistic and non-linguistic information (e.g., social expectations) are intertwined (Hay & Drager 866) and one in which socially weighted perception of spoken words that encompass both linguistic and social factors: where listeners map acoustic patterns to linguistic and social representations in tandem (Sumner et al. 1015). The results of storing and later accessing these representations means listeners are directly encoding their judgments and prejudices in the very apparatus of language perception that they use in everyday life. Language use literally cannot be disentangled from the social context in which listeners and speakers find themselves, including power structures inherent in the dominant society.

Consequences of standard language ideology have been demonstrated in educational environments which point to real-world effects, demonstrating that intelligibility can be affected by listeners' attitudes toward a speaker. In a seminal study on perceived accentedness, degree of accent correlated negatively with undergraduate students' perception of teaching competence of international teaching assistants. (Rubin and Smith 351). Comprehensibility ratings were measured after playing 4-minute lessons either in a "moderate" or "strong" accent for 92 undergraduate students while displaying the photograph of one of two "lecturers"—a white or an East Asian instructor. In a follow-up, a standard American accent was used as the audio signal, students who saw a picture of an East Asian woman while listening to the lecture performed more poorly on the content exams in the post test, thus affecting intelligibility (Rubin and Smith 348). In later research, this phenomenon is referred to as "reverse linguistic stereotyping" which refers to a listeners' difficulty navigating a seemingly neutral accent being produced by a speaker who appears to not be from the area (Kang & Rubin 441). An example of this type of reverse linguistic stereotyping was the relentless accusations leveled at Barack Obama for not talking "black enough" throughout his presidency (Graham). Direct implications for casting and theatre can be added to the complicated equation of listeners' use of intelligibility to perceive language.

One final term from linguistics that will contribute to the foundation of this dissertation is how models of language or dialect acquisition can be conceived as deficit models of language production. Deficit models take the basic assumptions made by standard language ideology, that there exists a "perfect or homogeneous" version of language and apply this ideology to language instruction. The deficit model assumes

speakers or learners are flawed or incomplete in their acquisition of the target language and are subsequently judged by the degrees to which they are assumed to be flawed (Modiano 525). The ideal form is the unattainable yardstick by which all speakers are measured, and teachers are free to treat students by their expert estimation of their skill acquisition in the classroom. Even liberal or culturally sensitive approaches to language instruction demonstrate shades of this deficit model. For example, a movement to explicitly teach code switching between non-standard home dialects and school-approved standard dialects still implies there is an appropriate space for one variety of speech over the other (Modiano 527). When the “appropriate” environment for the acceptable standard way of speaking is also the institutionally reinforced environment of school, the student learns their home dialect does not belong in the institution, thus reinforcing individual ‘deficits’ in their way of speech (Rosa and Flores 145).

Given this background, I designed a series of experiments using performed speech to test the hypothesis that objective measures of speech are indeed socially constructed in the minds of listeners. Both of these experiments manipulate listeners’ conscious expectation of performance, which disambiguates the role of the context of performance from language perception in general. These experiments both replicate classic experiments (e.g., exploring accentedness like Munro and Flege) and build a performance-specific inquiry into how audiences judge accents on stage. Results from these experiments are incorporated into a subsequent discussion of the implications for cognitive humanities research, along with practical results for voice and dialect practitioners. The combination of empirical findings of the specific linguistic inquiry of this project and the systematic exploration of the psychological and philosophical

underpinnings of conceptions of the voice contributes to meaning-making and reinforce an emerging view of how to approach performance in the conclusion of this dissertation.

3. Chapter Summary

In this dissertation, I refer to each section as a thread, as each section is both independent and dependent upon the other in narration construction and in chronology. That is to say, each thread is thematically organized, and within each thread is its own unique chronological and thematic progression. References within individual sections that point towards discussion elsewhere in the dissertation will be noted as the dissertation progresses.

The first thread contains a critical history of voice practitioners and their guiding assumptions of audience experience; I will be sketching three general generations of voice and dialect pedagogy based upon three approaches to voice instruction. Starting from the foundations with elocutionists and inventors of standard dialects such as Received Pronunciation, through practitioners who were concerned with “freeing the natural voice,” to more recent practitioners who have embraced the science of voice in their approach to training the actor’s voice in VASTA, I examine the assumptions that lie at the core of each of these eras of voice pedagogy. To end this thread, I preview efforts by recent voice practitioners to decouple this practice from using standard language in voice pedagogy, which aligns with efforts in other areas of theatrical production to expand representation both onstage and off. Interspersed through this historical discussion is a discussion about how the theatrical apparatus influences audience members’ meaning-making by examining how American Theatre has been shaped by over 100 years of the tradition of Realism, a theatrical movement that has its roots in

Moscow Art Theatre established by Konstantin Stanislavski in 1897 (Benedetti 12). Reliance upon “authenticity” in this mode of production leads to expectations of “real life” onstage. I will also discuss the material circumstances and the economics of American theatre making that influence the approach to voice training. The thread is bolstered by scholarship in cognitive humanities that takes seriously the notion that cognition in performance and the arts arise from the cognitive structures that are already in use by each human. This thread also offers cognitive philosophical underpinnings of the metaphors actively in use by voice professionals. This establishes a foundation by which I examine a smaller piece of this context of performance, namely what linguistic perception and expectation of intelligibility of voices and speakers on stage offers to the production and meaning-making process at large.

The second thread interrogates the assumptions made by voice and dialect practitioners in their work about audience experience using linguistic experimentation as a second lens for critical inquiry. In this thread, I examine the linguistic literature that traces the origins and effects of standard language ideology, which then serves as additional background to my own empirical investigation into audience perception of accents on stage. I background other research of interest to voice and speech practitioners that demonstrate the nuances of acquiring a second dialect (Siegel), perceptions about regional and non-native accents (Moyer), and the mechanics behind clear speech (Smiljanić and Bradlow 4020, Bradlow and Bent 707). Very specifically, I focus on the factors that surround audience judgments of imitated accents and their effect on intelligibility, both in the broad voice practitioner sense, and the more narrowly defined linguistic sense. Instead of assuming about how audiences perceived a performed accent,

I ask audience members what criteria of judgment they were using when they perceived these types of accents. This thread ends with interpretation of these experiments and exciting implications of how these data will contribute to the re-configuration of the profession of voice and dialect. These experiments will assist in defining what clear speech regarding audience expectations of understanding—and even “intelligibility” as conceived by Dudley Knight—means in the context of performance, which will lead to a more thoughtful pedagogy that accounts for these audience expectations.

These two threads form the basis for the conclusion, where I grapple with my position as a voice professional and offer practices that are informed by both the historic notion and the cognitive notion of audience expectation are proposed as a specific guide for voice and dialect professionals. Part of this work is discussing the critical pitfalls of working in such a profession situated within a society that has such strong standard views towards language, and a large part of this chapter addresses the mismatch between the judgments that audience members make in the experiments with the judgments voice practitioners believe audience members make. I specifically begin with a question that guides my work as vocal professional, borrowing from Amy Cook’s *Building Character: The Art and Science of Casting*. She asks of character creation more generally, “What does it mean to build characters from the ecosystem up, rather than a more psychologically focused method of character creation?” (117). This thread answers this question more specifically about building characters as part of an embodied dramaturgical framework that treats dialect and accent selection as seriously as other aspects of theatrical production. As a dialect coach myself, I use specific examples from my own practice that address assumptions about audience perceptions and how they

might create meaning while watching theatre onstage. As a scholar, I hope that my focus on voice for this dissertation does not lead the reader to the false conclusion that I do not care about embodiment of voice. In fact, this dissertation deeply considers how voices are perceived as they are attached to bodies onstage, and the combination of these two experiences carries social meaning. I am advocating for a deep consideration of these topics in voice precisely because I think these considerations have been left out of our conversation about representation.

By pushing back against the more general assumptions about rationality and cognition, I am able to create a space that more deeply considers the intertwined, sometimes contradictory nature of meaning making and artistic creation in theatre (and in entertainment more widely). As Mark Johnson asserts in *The Meaning of The Body* (2007)

Our “body” and “mind” are dimensions of the primordial, ongoing organism-environment transactions that are the locus of who and what we are.

Consequently, there is no mind entity to serve as the locus of reason. What we call “reason” is neither a concrete nor an abstract thing, but only embodied processes by which our experience is explored, criticized, and transformed in inquiry (vi).

This dissertation aims to illuminate assumptions to the contrary of these embodied processes to propose a new approach to voice and dialect practice that both rejects harmful, explicitly racist practices and builds a model that reflects our understanding of human cognition and meaning making. This dissertation begins with the assertion that we

all carry within us a voice that has been shaped and created by the location we grew up in, those who we called peers, and every linguistic interaction since acquiring language from a young age. My aim is to create a paradigm that teaches students and actors about the complex lived experiences that accompany dialect on stage, while actively working to counteract the harm caused by the problematic aspects of the profession. In essence, I am proposing a type of deeply situated dialect dramaturgy that honors different forms of objective/subjective knowledge and that accompanies the pragmatic aspect of learning the sounds of a new dialect as an essential part of the character creation process.

CHAPTER II

A CRITICAL HISTORY OF VOICE AND DIALECT

“... Physically crossing ethnic borders was relatively easy for me until I entered the world of theatre. There cultural and monetary capital was acquired by entering the dominant culture. To gain entrance, I abandoned my voice.” Micha Espinosa, “A Call to Action: Embracing the Cultural Voice or Taming the Wild Tongue”

1. Overview

Robert L. Hobbs published the book *Teach Yourself Transatlantic* in 1986, teaching a dialect that was created for the stage as a secret to becoming a successful individual in society. As an appeal to his authority as voice expert, Hobbs, a “well-known teacher in the field” (xii), claims that his system of working not only is advantageous for a student of acting, but also his system of working would indeed lead to success in more fields than just in performance. Hobbs uses his authority as voice teacher to combine the two major goals of voice practice in his arguments for Transatlantic as an appropriate dialect for all aspects of life.

Can accent really make a difference? Yes—some people claim that it’s the major difference between managements at the upper and lower levels...*the way you speak gives an impression—for better or worse—far more lasting than the clothes you wear or the design of your home.* If you have upward mobility on your mind, speaking transatlantically can help you blend more successfully into the particular social or professional group of your choice. (Hobbs, 1986, X, emphasis my own)

Hobbs presents his own standard language ideology as immutable fact and wraps that ideology with his authority as a respected vocal coach to sell his book to people who may be feeling self-conscious about how they sound compared to their peers.

Using this book to teach oneself a stage and film dialect from the 1940’s is an extreme example of the idea of appropriateness—a theory that posits that speakers

believe certain varieties of the language or dialects that a community speaks are *appropriate* in certain contexts and not others. Hobbs and those who subscribe to these approaches conceptualize standardized linguistic practices such as stage dialects as an objective set of linguistic forms that are appropriate for an academic, work or otherwise successful setting (Flores and Rosa 149). Appropriateness-based approaches center the idealized white listener as the target and aims to create a maximum intelligibility based around expectations of that listener. Marginalized workers and performers are required to learn the appropriate linguistic forms and assess *when* to use these forms, rather than ask a listener to accommodate the speaker. Hobbs uses this implicit framework by placing the onus of communication on the speaker and at no point does he suggest that bosses and other listeners ought to practice listening to the plethora of dialects of their employees. Hobbs inherited this approach to Transatlantic and the belief that changing one's speech patterns can lead to advances in life from a long line of voice professionals.

This thread lays out three major waves or generations of voice and dialect practices in United States theatrical production in the 20th and 21st centuries.⁹ Key philosophies and approaches to voice and dialect training shape each of these waves, and they reflect approximate successive generations of teachers and students who apprenticed under their preceding teachers, developing their own materials from their prior training. In part due to this loose apprenticeship structure, the three segments of voice training do not necessarily have distinct chronological borders. Each successive generation inherited the voice philosophy of the last, which fueled their own problematizing and creating their own view on the voice. Because of this, this chapter is roughly divided chronologically

⁹ I will also be addressing British voice and dialect norms to the extent of their direct influence on practices in the United States.

but the sections will focus on key practitioners from the era. Like oral histories of families, strict chronologies do not matter as much as who taught whom, and who learned to accept the ideology from their teacher and who decided to push back against their teacher. Coupled with these shifts in approaches and teaching philosophies are the changes in the material circumstances of each successive generation, with the establishment of industry and educational norms. The approaches to voice philosophy and to theatrical production in the United States are intertwined enough that presenting them side by side will demonstrate their effect on each other. Standards in both philosophy of voice and material circumstances of production of one era of training become the essential problems and questions of the next era.

Each of these waves contain key influential practitioners that help define the overarching approach to voice. My historical study is limited to voice practitioners who have written instructional materials and scholarship that document their particular approaches and are often cited as touchstone approaches to the practice of voice. This study includes both approaches to voice instruction more broadly, and practitioners whose area of instruction includes dialect coaching. Dialect coaching is defined as training an actor in a dialect or accent that is not their own (including dialects that are intentionally created and do not have a real-world equivalent) and can be a specialization of voice teachers who practice more broadly. I am defining voice practitioners as any instructor that trains performers in any aspect of the voice, including vocal anatomy, breath work, movement (especially as it pertains to preparing the body for performance), and articulation of the vocal apparatus. The definition of voice practitioner is broad and can include theatrical, film, singing, and dialect. This history includes both voice

practitioners in the broad sense and dialect coaches in specificity to set the stage for the linguistic experiments and subsequent best practices that specifically focus on the issue of dialect training. To attempt a history of dialect coaching without situating it within the larger voice profession would be an exercise in futility, especially since so many dialect coaches employ general voice practice in their work.

The first generation of practitioners, spanning the first half of the twentieth century, I will name the Elocutionary phase, due in part to the influence of several elocutionists who did not start specifically in performance or theatre, though may have transitioned later in their careers. Philosophical advancements of the 1800s, including Semiotics and new scientific approaches to acting, heavily influence the thinking of these practitioners. This elocutionary era saw the foundation of several influential schools and departments in higher education, establishing the authority of the brand-new profession of voice and elocution. Practitioners in this era would not shy away from their explicit stances towards linguistic supremacy of English spoken by the white middle class majority that created most popular entertainment. This phase also saw a change in preferred entertainment and media tastes in performance, shifting from live performance to film (Elliott 140).

The second generation, students of the first generation who benefitted from key structural changes to the institution of theatre making and became “master practitioners” in their own right, created a response to the strict standardization of speech through exploring psycho-social and anatomical approaches to freeing students from tension. This generation’s focus on freeing tension coincided with an explosive growth of regional theatres, thus further embedding this profession of vocal training into the vast network of

regional theatres and the larger economic arm of theatrical production in the United States. These regional theatres have become an integral part of the theatrical landscape in the United States in particular through the birth of the League of Resident Theatres (LORT) system of theatrical creation (Zazzali 192). The philosophical underpinnings of these instructors in this time period attempts to break free of the partitioning of the mind and body into separate entities.

The third era takes the idea of freeing tension and claims to use a more scientific approach to the practice of voice. Successive practitioners in the first twenty years of the twenty-first century have shifted their pedagogy to create context- and cultural-specific approaches to voice pedagogy that embrace embodied realism, where the practitioner and audience member alike are constructing meaning through individual and collective enacted experience of the world around them. The intergenerational shift between these successive voice practitioners are in part a result of the heavy use of the master/apprenticeship model of knowledge transmission, and thus some overarching conceptions of voice have succeeded in influencing the practice even today.

While other authors, including Dudley Knight (2000, 2012) and Derek Mudd (2014) have written versions of the history of Voice and Dialect practitioners, this dissertation's version has a specific critical focus. The critical lens of this brief history of voice and dialect practitioners arises in two main themes, both directly related to "standard language ideology" of Lippi-Green (9) and "vocal imperialism" of Rodenburg (14). The goal of voice instruction as a profession (regardless of generation) supports two explicit goals—one where a student is stripped of their particular idiolect or unique way of speaking, and the other where that student is then encouraged to use one or more

dialects that have been simplified or sanitized. Both of these goals are named in the name of audience intelligibility in the estimation of the trained voice professional. Each generation of voice professional has had these two goals, either explicitly or implicitly, which has had the potential to harm students in the process. For example, the first goal in the elocutionary generation manifested in practitioner Edith Skinner's classroom as her infamous day one exercise, where she would invite each student to pronounce their name and then she would correct the name into her proprietary "Good American English" implying that the student's way of pronunciation was unacceptable (Skinner et al. 20). This first goal manifests differently in contemporary practices and can include various microaggressions and social and economic barriers that institutions put into place that prevent a student actor from accessing voice training in the first place. Each era, through its relationship with higher institutions of learning and the economic realities of theatrical production, presents a way to affect students' voices in a manner in which their home dialect or accent is not welcome on stage.

The second goal for these voice and dialect professionals appears more explicitly when a second dialect or accent is needed for the stage, whether it is a standardized dialect or a foreign or regional dialect. Historically, voice coaches taught imitated dialects in a way that strips the individual dialect of its nuance and complexity in the name of intelligibility or audience understanding. These sanitized dialects are reflections of stereotypes, or standardized linguistic ideas, that society creates through associating meaning-making with how people of a certain race, class, gender, or language background sound. Students whose own voices do not match the expectations of intelligibility in performance are often faced with two issues of linguistic representation,

being silenced in the voice training classroom, and then sometimes being asked to exoticize their own accents to fit the stereotype expectations of dialect coaches, casting coaches, and directors.¹⁰ For example, Asian American actors like John Cho are still asked to put on highly stereotypical East Asian accents to portray Asian characters in the movies in which they are cast. Aware of the historical problematic representation of Asian characters in particular, explaining his acting choices, Cho did not, “want to do this [one] role in a kid’s comedy, with an accent, because I don’t want young people laughing at an accent inadvertently” (Sullivan). Black, Indigenous and Actors of Color like Cho are often keenly aware of the harmful stereotypes they are asked to perpetuate by voice and dialect professionals and directors in theatre, film and television. Marginalized actors are often put into the unenviable position of advocating for themselves and their identity groups against the overarching power structures of performance creation that favors stereotypical presentations of foreign and regional dialects that are accompanied by negative presentations of race, gender, and class. In this environment, marginalized actors do not possess enough power in their workplaces to push back against the tendency to simplify and present stereotypical accents.

These two goals in dialect and accent (subtracting undesirable linguistic traits or behaviors and replacing with sanitized or stereotyped versions of language) are ways of erasing authenticity and therefore embodied linguistic knowledge for the voice student or performer under the care of voice professionals. These goals have remained prominent in part due to an unchallenged authority of the vocal or dialect coach as a singular expert in all matters of the voice. This authority rests on a semi-scientific knowledge and appeals

¹⁰ I discuss a case study of contemporary Latinx students the conclusion of this dissertation.

to audience intelligibility or ease of understanding for the audience. This appeal to intelligibility also feeds back into linguistic supremacist ideas that some speakers are already more naturally intelligible than other speakers, which often align with other supremacist ideas about race, gender, and class. This approach often claims that there is a voice that is appropriate for performance that does not match the voice of the actor, and thus they must learn the linguistic forms and policies that govern this appropriate way of speaking. Throughout this history of voice practitioners, I will highlight the ways in which key practitioners use their authority as experts in language to further their agenda.

The following history of voice and dialect foregrounds the assumptions that practitioners in each successive generation used to build their profession and also highlights the material circumstances behind each approach to voice training. The material circumstances are an important piece of the story of this profession, as theatre has become an institutionalized piece of an ideal listening apparatus, having been privileged as a form created specifically for overwhelmingly white, middle-class audiences (Bennett 114). I trace the lineage of these assumptions through the different generations, to construct the base for the contemporary understanding and approaches of voice and dialect work. The following sections will demonstrate the way voice training is structured historically has catered nearly exclusively to an appropriateness-based approach to audience, at times explicitly privileging constructed linguistic forms over natural or spontaneous in service to maximum intelligibility of audience members. Theatre as practiced in the United States during this time is the ideal white listening apparatus made manifest and is therefore a site ripe for intervention against appropriateness-based approaches to contemporary voice training and in the future.

Creating this opportunity to push back against prior generations of training (in the tradition of those who have come before) will set up a different view of how audiences may create meaning through their social expectations of voice independent of voice professional intervention, which establishes the foundations for the experiments in the following thread.

2. Learn to speak “Good” American English: Elocutionists 1900-1950

2.1 William Tilly teaches World English

The first generation of modern voice professionals begin with the elocutionists of voice, which is an era that begins at the turn of the century and continues through approximately the 1950s. This generation is marked by elocutionary teachers who were not necessarily specifically associated with theatre or film performance but were imposing strict standard English practices with their students in an effort to create speakers who were successful in life as well as performance. Most practitioners’ goal was to create permanent good speech in their students according to their own standards, reflecting their own view on what good speech ought to sound down to the tiniest minute phonetic detail, without producing much evidence on why that speech was supposed to be better than other types of speech.

The progenitor of this era, with foundational writing and training in speech, is William Tilly, who through his obsession of capturing fine phonetic detail from speakers of English, also contributed to the creation of the International Phonetic Alphabet¹¹ (IPA) by the International Phonetic Association (Knight 32). In this era, practitioners used strict phonetic transcription as the learning model for students. Many students of William Tilly,

¹¹ This alphabet is still in use by thousands of linguists around the world.

such as Edith Skinner and Marguerite DeWitt, practiced narrow transcription of IPA as a way to capture precise detail of speech and advocated for a strict approach to the English language through this transcription style. Narrow transcription is the practice of transcribing linguistic sounds with as much phonetic detail as possible, where each letter that represents a linguistic sound (e.g., a phoneme, discussed in detail in the next chapter) can feature diacritic marks that indicate slight differences in pronunciation due to the location of the letter within the structure of words, and of course due to differences in dialect of the speaker. With this system, each phoneme represents the ideal pronunciation of the sound, and every diacritic added represents a failure at achieving that ideal.

While elocution as a formal profession arose near the turn of the twentieth century, it owes most of its origin to practice of rhetoric, the art of persuasive speech in the realm of public speaking. Rhetoric begins with Aristotle formally and the art of instruction for public speaking takes many forms throughout history. The direct English descendent of rhetoric that contributes to elocutionary studies begins in the middle 1700s. *The Art of Speaking* published by James Burgh in England in 1762 kicked off an elocutionary movement in the United States. This book would inspire other texts where the goal was to inspire proper persuasive speaking in the public sphere.

Two competing schools of elocution would arise at this time in training, with competing philosophies or approaches to the role of the voice in public performance. One such school, known as the Mechanical School, taught countless students to align gestures and expression with speech in order to appear persuasive in the public sphere, though without much emphasis on connecting to the emotions underlying the speech. As a result of this school of elocution, in 1827, James Rush, a U.S. medical doctor, published *A*

philosophy of The Human Voice, based off his work on anatomy and physiology of the human voice, according to medical knowledge of the time (Mudd 31). This would mark the first instance where a voice practitioner would appeal to the authority of the fields of science and medicine, to lend a veneer of authenticity to the claims made within the book. Proponents of the mechanical school of acting were influenced highly by Denis Diderot's *The Paradox of Acting* (1883). In his treatise, Diderot claims that the role of the actor is to recreate the forms, gestures, and habits of characters without the actor becoming emotional themselves (Roach 116). Soon after, another school of "expression" would rise in opposition to the Mechanical School of elocution. The expressionism school of elocution was influenced heavily by Romanticism, Naturalism, and other philosophical movements that arose in the same era. These teachers were interested in the interior expression of a speaker and warned against the external and mechanical nature of the school before them. The expression school for elocution further split into two fields, namely oral interpretation and actor training (Mudd 31).

At the same time as these competing schools of thought, a young William Tilly formed his school of elocution in the late 1800s, which leaned into the Mechanical aesthetic. William Tilly grew up in Australia in the 1860s and 1870s and moved to Germany. Having already established his school for elocution in Germany in 1890s, Tilly moved to the United States in 1918, right when practitioners in the school of expression split the profession between oral interpretation and acting (Knight 32). While the school of expression found homes in English and speech departments and eventually fledgling theatre departments in universities in major cities along the east coast of the United States, Tilly's sights were set on the scientific interpretation and precise expression of

language and subsequently found the humanities limiting to his vision. Through his school, Tilly attracted many students who wished to master English as a second *spoken* language in all arenas of life, believing that they were carrying linguistic deficits. The United States at this time was undergoing a massive shift in demographic, with more than 15 million immigrants arriving in the years between 1910 and 1915, which was a number equal to the number of immigrants who had arrived in the previous 40 years before this date (“Immigrants in the Progressive Era”). Due to this, Tilly had many eager students striving to assimilate to their new homeland. Dudley Knight describes Tilly’s chief reform—one that was passed down to many of his students, most of whom became very influential in the field of voice—as, “his attempt to teach the pronunciation of English as a spoken language, and not as a written one” (Knight 33).

To assist with his goals, William Tilly was one of the first elocutionist practitioners to create a wholly artificial dialect that he taught to his students. Subsequent students of Tilly’s would call his system World English or even World Standard English (Knight 34). Through his advocacy for narrow transcription, Tilly set the stage for fanatical adherence to how English *should* sound in every arena of public performance (and by extension, private communication). His students, including Marguerite DeWitt, Margaret Prenderghast McLean, and later Edith Warman Skinner, would carry this fanaticism through their own teaching and strict adherence to detail. With books such as *Speak With Distinction* (Skinner et al.) and *EuPhon English* (DeWitt), elocution teachers framed their ideologies with their narrow instructions on how to sound in real life. Implicit biases against speakers whose first language was not English often became explicit, as evidenced by the introductions in several of these books. The explicit

ideologies professed by voice practitioners and other disciplines that were examining language in a systematic and scientific way.

The debate between those professing that language ought to be pronounced in a certain way and those who merely wished to observe and describe the varieties of language found in the world became particularly fierce in the 1920s. On one side of this debate was Tilly and his students' fanatical adherence to speech standards, while on the other side were linguists and linguistic anthropologists beginning to establish their fields and academic departments in universities. Anthropologist John Kenyon, having just published his influential textbook *American Pronunciation*, advocated fiercely for the equality between different dialects heard in the United States, and argued fiercely against the use of standard dialects, especially ones promoted by Tilly based on class and access to different elocutionary and training techniques (Mudd 35). Through this feud between sides of standards in United States English pronunciation and dialect, many followers of William Tilly revealed their own biases against speakers with no formal training, especially those for whom English was not their first language. Elocution practitioners were expressing fear of a polyglot United States where diversity and difference are valued over unity and homogeneity. These ideas bore strong resemblance to racist ideas expressed in other social spheres in this tumultuous time in United States history (Kendi 145). One passage from Tilly's student Marguerite DeWitt in the introduction of her book *Euphon English* highlights this explicit prejudice, even fearing the dissolving of the United States entirely:

To squander national vitality and money on that which will but cause biological disintegration of a nation is not the philanthropy; to infuse into a body politic

blood that destroys the racial blood of a nation is not the deed of a rational healer; to foster the growth of parasites on a national tree of education and knowledge is not the work of advanced sociologist. (DeWitt, qtd. in Knight 40)

DeWitt refers here specifically to her reluctance to educate immigrants to this country by using particular supremacist phrases like “racial blood of a nation,” claiming that such time spent in education is equivalent to fostering parasites that would otherwise be unwelcome to a racially pure United States. These racist ideas, however, were not explicitly limited to the practice of elocution, though to find them at the very root of the origins of this discipline should not be dismissed as incidental to the time in which these authors found themselves. To demonstrate how these explicitly racist ideas became ingrained to the work of performance, I turn now to a practitioner who bridges elocution in general to performance in theatre specifically.

2.2 Edith Warman Skinner bridges performance

One student, who can trace her lineage directly to William Tilly and his school, will become particularly influential in future approaches to voice in the twentieth century, and will make the connection between elocution generally to performance specifically. Edith Warman Skinner, like many in the expression school of rhetoric, realized the power of vocal training in the life of an actor (Skinner et al. xi). Simultaneously with this jump to speech training in performance, acting training has been shifting from the teacher-apprentice model of various touring troupes of theatre ensembles where actors are immersed in on-the-job training, and towards a more formal site of education through

new partnerships with higher education institutions. For instance, the first theatre¹² department in the United States was established in a similar timeline as Tilly's growing influence in 1925, at Yale University (Berkeley 23). In these burgeoning theatre departments, the expressive school of elocution was winning, and theatre practitioners were employing new techniques by Konstantin Stanislavski to couple emotion and action with text. The actor was to consider the word as "verbal action" and no longer as literary form (Moore 69).

Skinner had been trained as an actress at the Powers school, so she had an interest in combining the elocutionary lesson she was learning from her teacher Margaret Prenderghast McLean with her work as an actor (Knight 43). Skinner became the speech instructor at Carnegie Tech's theatre training program in 1937. Skinner established herself as one of the premier speech trainers for theatre in America, not only because of the large number of famous actors she worked with but also the number of speech trainers that she taught, as well. Her legacy as a speech trainer can be seen in the generations of contemporary speech trainers who can and often do trace their lineage directly back to her and her work at Carnegie. Edith Skinner would go on to hold two appointments at Carnegie Mellon school and the Juilliard school at Yale. After World War II, many soldiers returned home from the war and took advantage of G.I. Bill benefits, thereby flooding the American university system, and allowing the freedom to pursue disciplines in the humanities that were not necessarily immediately lucrative.

¹² Standards for language extend to the never-ending debate of the ending of the practice of theatre/theater. Apart from regional differences (British versus American spelling preferences), some American institutional bodies use Webster's dictionary as an appeal to authority to argue for the -er spelling. My preference is to refer to the practice as "theatre" and the space or location as "theater"

Theatre departments developed throughout the country and would often focus on training for actors. Within this environment, Edith Warman Skinner found the necessary conditions to create a system similar to her teachers before her, a system that would become the standard for theatrical production and would heavily influence speech in film. Skinner's system, named *Good American Speech*, presented what Skinner described as the most intelligible type of speech for performance (Knight 44). Skinner also taught her own version of the International Phonetic Alphabet, often combining her own cursive symbols with standard symbols for sounds, demanding her students practice an exact copy of her own work.¹³ Through this, Skinner was able to create a proprietary system that required rigorous study with her and her designated students, thus establishing the practice of creating exclusive systems that require particular access to training.

Though Skinner's text *Speak with Distinction* was published posthumously, her unofficial notes and voice approach were shared between departments, always with careful attribution to Skinner and her system. Skinner's work eventually formed the basis for Mid- (or Trans-) Atlantic dialect, a dialect that was used by many film and stage performers throughout the twentieth century, eventually losing popularity in the mid 1980's (Elliott 105). This dialect is recognizable in many Hollywood stars such as Audrey Hepburn, Judy Garland, and Marlon Brando. Skinner constructed this dialect with the aim for maximum intelligibility and the dialect is constructed out of a mixture of dialect features from British varieties of English, most notably lacking /r/ in particular environments and the use of broad /a/ or the initial vowel sound in "father," and the

¹³ See Michael J. Barnes "A Critique of Phonetic Transcription in American Actor Training" in *Standard Speech: Essays on Voice and Speech* on page 100 for diagrams of Skinner's use of the IPA.

rhythm of the speech of higher class white residents of the Eastern coast of The United States. Eventually, this accent enjoyed the prestige status of being the most intelligible and preferable dialect for both stage and screen, due to the fanatical advocacy of Skinner's students. Skinner's system would remain popular with students even until the twenty-first century; the first edition of the *Voice and Speech Review* would feature no less than six articles and rebuttals to Skinner's system in 2000.

Skinner continues to be one of the most influential early figures for voice professionals, whether they ascribe to her ideologies or push against them. Notable of Skinner's students are Tim Monich, a dialect coach beloved by Hollywood and covered later in this chapter, and Sanford Robbins and David Hammond, both of whom would figure heavily in the eventual formation of the professional organization for voice professionals. Skinner's Transatlantic is often still held up as the proper or correct approach to Shakespeare¹⁴ or classical work in particular (Hammond 143). The association between Skinner's Good American English and performing classical works such as Shakespeare is incredibly strong; facets of this accent can be heard in stereotypical pseudo-British Shakespearean dialect that students and those poking fun at Shakespeare often drop into while performing classical texts. Lippi-Green's standard dialect ideology is at play here, since many audience members and inexperienced actors

¹⁴ I often point out that stereotypical approaches that pokes fun at a Shakespearean theatrical accent sounds relatively close to the Transatlantic accent Skinner taught in her classrooms. This is especially ironic given David Crystal's work in Original Pronunciation (OP) of Shakespeare which—as it has been historically reconstructed from Early Modern English—does not share many linguistic traits with Skinner's system. For an excellent audio comparison between RP/Transatlantic and the OP of Shakespeare, see the video "Listen to a demonstration of the original pronunciation of Shakespeare's English and how it differs from modern English" on Encyclopedia Britannica's website: <https://www.britannica.com/video/187707/David-Crystal-pronunciation-Ben-Elizabethan-English-British>

strongly associate a prototypical accent with a particular type of dialect or accent.

Linguistic stereotypes can be strongly associated with context when enough language users employ the same dialect consistently.

Concurrent to Skinner's strong adherence to linguistic standards is the work of Arthur Lessac, who would become the progenitor of the next generation of voice professionals that eschew overt standards for an individualized psychosocial approach. Edith Skinner and her contemporary Arthur Lessac would rise to be one of the most preeminent theatre voice teachers in the United States (Mudd 30). The major difference between Skinner and Lessac was Skinner's emphasis on "standardized speech" while Lessac advocated for a more individualized approach to voice in performance that examined each voice student. Arthur Lessac's work as a voice teacher would mark a shift from standards and elocutionary practice to a more individualized approach to voice, though Lessac's goal to "produce beautiful voices" and, "clear, articulate speech" still bore the hallmarks of standardization (Lessac 114). Lessac's standardized voice did not have to conform to a narrow phonetic or overly prescriptive approach to voice. Lessac's class had more room than Skinner's for voices outside of the narrow band of approved students, but not to the extent where every voice was welcome in the training classroom. Each speaker still had to adhere to a standard of "clear, articulate speech" though now that standard was not made explicit through precise use of phonetic symbols and rote drills. Lessac's work, however, did pave the way for the next era of voice training, guided for the most part by Kristin Linklater's landmark work *Freeing the Natural Voice*, covered in the next generation.

2.3 Mimesis vs. Semiosis: Establishing dialect coaching as a profession

Near the end of this elocutionary era, a different set of voice professionals guided by similar overt language ideologies as Tilly and his students staked their expertise through dialect instruction for voice in performance, specifically with dialects that were trained and did not necessarily originally belong to the performers themselves. The second goal of historic dialect training, to produce sanitized and easily acquired stereotypical accents, guided these practitioners as they established their authority on dialect coaching. In this era, the husband and wife team Lewis Herman and Marguerite Shalett Herman published two seminal texts on dialects, *Foreign Dialects* (1943) and *American Dialects* (1947). The original subtitle for both of these texts reads *A Manual of Dialects for Radio, Stage, and Screen*, implying the dialects presented were not only representative of the countries and regions they claimed but also suitable for a plethora of public performance scenarios over and above theatrical presentation. The text on foreign dialects in particular were intended to, “help the actor prepare for the most difficult foreign role and offer the director or producer a convenient aid for correcting actors and evaluating applicants for authenticity and dialect ability” (Herman and Herman back cover). Presented as an authority on these dialects, Herman and Herman compiled this material during more than twenty-five years of acting, writing, and teaching across Europe, New York, and Hollywood.

Their texts became canonical dialect and accent texts for producers, writers, and actors alike. In their *Foreign Accents* text, Herman and Herman not only describe how the accents mechanically worked, but also described the stereotypical stress patterns (described as “lilt”) of an average speaker. They also provide grammar expertise for common mistakes produced by these speakers, in service of providing what they believed

was a more authentic example of dialects for playwrights and screenwriters. What Herman and Herman believed were mistakes in the dialect are actually valid and real differences in dialect that approached grammar differently from American English. Because they were approaching dialect from an ideology that held American English up as the standard, any deviations from this particular dialect were described as mistakes. Despite this, Herman and Herman frame this book as a helpful or neutral guide to foreign dialects. However, in their introduction, Herman and Herman reveal their preference for American English for stage and film,

The art of the dialect is the twin art of being consistent with the fundamental and radical changes and of being consistently inconsistent with the less-important [sic] changes...if the dialect is to be very light, the radical changes may also become inconsistent. But if this point is reached, the character will be speaking an almost *perfect American speech*. (Herman and Herman 15, emphasis is mine)

Written in 1942, Herman and Herman selected accents that were in demand from producers of stage and film and revealed their biases for and against several varieties of accented English through their selection and organization of these accents. For example, Herman and Herman divide British English into two chapters, assigning Cockney English its own chapter, and then assigning Australian English, Bermuda English, and the “Dialect of India” to the other chapter (Herman and Herman 65). That Herman and Herman consider Indian English¹⁵ a form of British English is a damaging and implicit reflection on their ideas on the colonization of India.

¹⁵ According to Babel, there are 22 official languages in India, and well over 19,000 languages and dialects, so reducing a subcontinent to only one dialect reveals multitudes about Herman and Herman’s attitude towards this country.

In addition to the grammar and rhythm considerations, the original publication of the book also described the characterization of an average speaker of that particular dialect. For example, British speakers are characterized as stolid, resistant to change, and “unbrilliant... With temperate habits, and temperate emotions” (Herman and Herman 52). Chinese speakers were characterized as industrious, frugal, devoted to family and country, and have “a proverb for every occasion and a wide grin to accompany it” (Herman and Herman 245). Each chapter, therefore, explicitly trains speakers and indoctrinates them into accepting that speakers of a certain dialect were inextricably linked with character traits both positive and negative. These characterizations are particularly egregious for the various East Asian accents presented in the book, as when this book was first published one of the enemies of the American army was the Japanese in World War II (Royde-Smith). Angela C. Pao writes specifically about this portrayal of Japanese accents written in the original 1943 publication,

The opening lines of the chapter on the Japanese dialect blatantly signal the war-time substitution of subjective prejudices for more objective observations:

‘Unfortunately, the Japanese military has caused the people of the other nations to brush the cherry-blossoms from their eyes and to thinking of these little, yellow men in unmentionable expletives. Their overpowering politeness has currently taken on a sinister threat, and their wide toothy grin, an ominous leer.’ (Herman and Herman 225, qtd. in Pao 358).

By painting these prejudices and overall approach to dialect with an objective or scientific veneer, Herman and Herman established the practice of dialects for film and theatre as a neutral or even beneficial contribution to entertainment. Using appeal to

authority with nearly twenty-five years in the business, Herman and Herman could parrot explicit racial prejudice as scientific fact, and necessary information for producing commercially viable yet ultimately harmful dialects.

These accents and this training manual, therefore, contributed to popular entertainment's construction of ethnicity and race. Accents and dialects are no longer neutral indicators but carry a constellation of meaning that includes stereotypes that audience members can incorporate into their artistic experience, according to what Robert Hodge and Gunther Kress describe as *Social Semiotics* (1988). In this publication, Hodge and Kress posit that due to the human capacity to assign meaning to every experience, every perceptible detail is available for meaning-construction. Pao states,

We are in a society that assigns character traits (i.e. meaning) to how people sound and to deny that fact is negligent behavior on behalf of voice and dialect professionals everywhere...Succinctly stated, accents of all kinds (foreign, regional, class) function not on the mimetic plane (to which dialect coaches refer on almost all fronts) but on the semiotic plane (the production of meaning.) (Pao 359).

Herman and Herman built the argument that dialect coaches are offering their services as a reflection of the mimetic conception of accent, while denying or deemphasizing the semiotic use of accents to produce meaning in the minds of audiences. Nevertheless, Herman and Herman still offered stereotypical characterizations of the speakers

While overt racial characterization descriptions are omitted from the 1997 re-publication of this text, implicit characterizations of these accents remain throughout the text, through descriptions of lilt, mouth position, and Herman and Herman's attempt at

delineating different types of accents within each chapter (Pao 364). Anti-Asian sentiment in particular is present in surprising places, including descriptions of the Filipino dialect that “resembles Pidgin English. In fact, it has, for this reason, been called *Bamboo English*. But the pronunciation is based mostly on the Spanish with some infiltrations of Malaysian” (Herman and Herman 190, emphasis author’s own). Herman and Herman, while careful to separate Chinese and Japanese dialects, flatten the diverse history of trade in Southeast Asia by privileging the influence of the European colonial language over and above a neighboring country. Herman and Herman then continue to flatten the taxonomy of different speech areas by assigning the “Portuguese dialect” a sub-area of the chapter on Spanish dialects.

Perhaps the most egregious examples of overt racism and overgeneralization exist in the practice monologues provided at the end of each chapter. For example, the monologue used in the Chinese dialect chapter still refers to an obedient Chinese owner of a Dry-Cleaning business, “grinning widely at a customer” as he says, “Ticket, please? Thank you. Me got wash finish... Maybe you put change in China Relief box, no? That for China people Japan make hurt. That for make world safe for democracy, no?” (Herman and Herman 258, qtd. in Pao 359). Editors at Routledge erased the overt racial linguistic imperialism, yet still subscribe to ideologies through implicit means.

Herman and Herman established what they thought was a neutral and empirical approach to dialect coaching, without acknowledging that they were relying on their authority as experts and building guides for accents according to their subjective perception of the speakers of these dialects. Dialect coaches have denied their participation in the semiotic meanings of these accents by focusing on the mimetic

aspects of these dialects. Performance is created in a society that assigns character traits to how people sound. To deny that fact is a misstep for not only performers but also voice and dialect professionals. Later dialect coaches and other authors of books similar to *Foreign Dialects* and *American Dialects*, grapple with this dichotomy of mimesis/semiotics to varying degrees of success. What remains from this time period, however, is a staunch adherence to the idea that the art of dialect coaching is a precise and near-scientific pursuit, masking the more dangerous prejudice and racial stereotyping that this profession replicates and perpetuates throughout entertainment. In the next section, even dialects presented as neutral contribute these dangerous ideas about the voices that produce these types of accents.

3. Freeing Tension and the rise of regional theatre

The second wave or generation follows the first era, but many of these ideas are attempting to push back against the overt standards of the previous generation. This second wave of thought came to prominence in the 1950s and remained popular until the early 1990s, when voice practice was changed dramatically by the advent of the internet and relative ease of access of knowledge. Practitioners Cicely Berry, Patsy Rodenburg, and Kristin Linklater were key players in transforming the profession from strict adherence to standard stage dialects to an individualized approach that aimed to consider the actor as a whole psycho-social being that requires individualized care and consideration (Mudd 40). Specifically, Linklater's text *Freeing the Natural Voice* (1976) became a touchstone text—partially due to other practitioners who did not write and widely disseminate their materials—for this generation of voice professionals. In this era, practitioners pointed to the need for deep physiological and sometimes psychological

work in order to free the body from tension, thus releasing the voice from the newly untensed body, implying that the ideal body and voice vessel is without cultural and individual habits that are carried in actors' bodies. The underlying ideology from this era presents an interesting new re-interpretation of the first goal of voice practitioners that aimed to erase individual identifying characteristics for actors. As Rockford Sansom notes, "...their work demonstrates a seismic shift in ideology away from elocution to a praxis desiring an actor's authentically personal expression and interpretation" (Sansom 159). From this type of personal expression arose the idea of authenticity onstage, which shares an uneasy relationship with mimesis/semiotic representation for the audience. The audience can read the same linguistic performance as simple mimesis and simultaneously as a stand-in as part of the gestalt of the larger meaning apparatus of the performance as a whole. For this generation, I have chosen to analyze in depth the philosophies of three practitioners in particular as the most widely used examples of the philosophy that governed the voice profession at this time, through the work of Cicely Berry, Patsy Rodenburg (already introduced in the previous chapter), and Kristin Linklater.

These three practitioners claim that to achieve authentic personal expression and by extension success in performance, students must release chronic tension in their bodies, since after all, the voice of the actor cannot be separated from their body. The success of this release still determined by the expertise of these practitioners within their individual proprietary systems. Actors who are more bodily able to release tension get to enjoy the benefits of voice work. Voice professionals, therefore, privilege the voice that inhabits some idealized, unmarred vision of non-tension. In other words, the work of these practitioners still privileges some standard, idealized homogeneous body release

from tension to produce optimal voice work¹⁶. Often, this type of voice work is accompanied by body/breath awareness work such as yoga, Feldenkrais¹⁷, and Alexander¹⁸ technique all in the service of release of muscle tension, and fuller awareness of the body (Moore 101). It should be noted that Feldenkrais and Alexander techniques both seek to eliminate what has been deemed as “harmful” tension in the body, while practitioners of these systems still claim that some tension (e.g., when an actor’s vocal folds are activated in phonation) is still acceptable. The idea of acceptable approaches to body support of the voice is still heavily regulated by students of these practitioners. This era is also marked by the unprecedented boom of the profession in two key performance markets, regional theatres all across the United States and the United Kingdom, and also the professionalization of the position of voice/movement professor in higher theatre education (Zazzali 47). This professionalism was due in part to two sources: the

¹⁶ See Louis Colaianni’s interview pp, 69-81 in *Voice and Speech Training in the New Millenium* by Nancy Sakland for an in depth discussion of the implications of what type of voice or body gets to be perceived as free of tension.

¹⁷ The Feldenkrais Method, created by Moshe Feldenkrais, is a system of body movement and awareness that uses gentle movements to ease tension. According to the method’s website <https://feldenkrais.com/about-the-feldenkrais-method/>, “The Feldenkrais Method is based on principles of physics, biomechanics, and an empirical understanding of learning and human development.” Through learning this system, “Since how you move is how you move through life, these improvements will often enhance your thinking, emotional regulation, and problem-solving capabilities.” Feldenkrais, like many practitioners of this era, used claims to science to sell a proprietary movement method as life-enhancing for their students.

¹⁸ Similar to the Feldenkrais Method, the Alexander Technique is another proprietary system of movement that makes similar claims of easing tension and creating movement “as nature intended.” According to the Alexander Technique official website <https://alexandertechnique.com/fma/>, Australian actor F.M. Alexander developed this approach to movement when he was experiencing chronic laryngitis whenever he performed. Alexander credited relieving his tension in his neck and body as the secret for his recovery from his laryngitis and developed a system to ease muscle tension from his personal experience. There is a second dissertation’s worth of commentary about these movements systems that also cover their proprietary systems in pseudo-science veneer and take advantage of marginalized populations in similar fashion to the voice practitioners of this era.

explosion of federal and grant funding available to regional theatre companies, and to the creation of the Voice and Speech Trainers Association that assisted in codifying expectations for these positions. The creation of such positions thus legitimizes the authority already created in the previous generation and the need for voice and movement practice for aspiring actors in the United States. This generation, like Skinner and her students previous, is also marked with the branding of specific approaches to voice training that will take hold in theatre departments across the United States.

Continuing on the work of Herman and Herman, Jerry Blunt created his own system for dialects, releasing *Stage Dialects* in 1967. Accompanying this guide was an innovation in dialect study; the book was released with practice tapes for the dialects featured within the chapters. The debate between approaches to dialect instruction in this era was between two camps: one camp advocated for use of spontaneous dialects “from the field”—using tapes of actual dialect speakers—while the other camp still advocated fiercely for the use of example dialects produced by dialect coaches themselves (Blunt viii). The implications of this debate would reverberate in the following years of dialect coaching into the twenty-first century. Blunt also made claims about the qualities of the accents and dialects he sought out for his definitive yet limited guide on dialect, revealing an interesting paradox about the types of accents and dialects that are privileged above others. Blunt carried many standard language ideologies established by Herman and Herman into this generation of dialect instruction.

3.1 Berry and Linklater and the British voice training revolution

Kristin Linklater, who was a student under Iris Warren, who taught at many different institutions including London Academy of Music and Art and NYU, would

publish her first book *Freeing the Natural Voice* in 1976, and kickstarted a revolution in voice training. This book would become the foundation of the Linklater system of voice training, which has over one hundred master teachers working today (*Linklater Voice*). How Linklater's book became one of the most influential texts on voice is a result of a combination of the number of prestigious positions Linklater would hold throughout her fifty-year career as a voice professional, and the emerging actor training philosophies that privileged the individual psychological experiences of the actor. Linklater credited her individual approach to her work with legendary voice teacher Iris Warren, often quoting Warren's training philosophy and mantra, "I want to hear the person, not the voice" (*Linklater Voice*).

Other individual approaches were heralded as the new definitive way to train the voice and other absolutely influential practitioners in this era to precede Linklater were Cicely Berry (*Voice and the Actor* 1973) and Patsy Rodenburg (*The Right to Speak* 1964). All three practitioners promoted the idea of psychological work for the actor to "free" them from the habitual tensions that society has placed upon the actors' body. Rodenburg's "vocal imperialism" constitutes the most socially aware ideation of this concept, while Berry and Linklater chose to reference this type of tension more obliquely as something more value neutral that leaves out intersections of class, gender and race. This shift from practicing a constructed standard for acceptability for onstage in elocution and towards true individual freedom signaled attempts at a radical re-imagining of the voice profession. Practitioners in a previous generation wrote openly about their linguistics supremacist ideas where this generation worked harder to include more voices in their work and classrooms. This shift was part of a larger movement towards a goal of

the democratization of theatre, as theatre makers such as Peter Hall, Peter Brook and Trevor Nunn influenced the new “so-called radical ” Royal Shakespeare Company of the 60’s and 70’s (Knowles 95). Linklater, Berry and Rodenburg sought to democratize the voice in response to what they thought were oppressive practices by the elocutionists before them. While they often sought freedom, they would often employ notions that accomplished the opposite, by seeking authenticity and insisting on their own version of intelligibility.

In her book, *Voice and The Actor*, Cicely Berry constructs intelligibility as the direct natural result of the Actor’s conscious work towards relaxation and freedom in their training. Within her book, actors are repeatedly urged to “allow the words to do their own work,” so that if successful, “the meaning will be clear” (Berry 108). In this case, regarding this book, Berry defines meaning in these words of the play as the original intention of the playwright or author, which means the actor’s job through voice work is to become the most neutral vessel possible through which the original intentions of the playwright may be read by the audience. This goal, therefore, privileges bodies and voices that are already closer to what an audience member may consider neutral, which often means a white speaker from a region that does not have any noticeable accent. The goal of this privilege allows the bodies producing these voices to melt into the background, foregrounding the text and meaning constructed through the script. Following this logic, Berry constructs intelligibility that is highly valued, most crucially, as meaning constructed by the words actors speak and not the context in which these words are uttered. In other words, the goal is to become neutral enough to utter words mimetically and not contribute as an actor to any meaning construction in which the

audience may partake.

Berry privileges and rests her authority on the text to be spoken above all other embodied approaches. Peter Brook, in his forward to Berry's book, confirms this privilege by praising her ability to neutralize or free actor's voices, "all is present in nature; our natural instincts have been crippled from birth...by the conditioning, in fact, of a warped society" (Berry iv). Both Brook and Berry commit to the idea that voice professionals must return the acting student to a *tabula rasa* for optimal intelligibility of the text, without explicitly constructing what that blank slate space looks like. In her subsequent book, *The Actor and His Text* Berry privileges this blank slate or maximum ease of tension to promote "intuitive response" (24) to text to access what she believes is "a physical level, deeper than the intellect" (27). In other words, she does not trust the intellectual access to meaning, and advocates for a subconscious approach to text and voice. This simultaneous freedom and instinctual response while pursuing an ill-defined definition of intelligibility of text (particularly Shakespeare) marks many of the influential practitioners in this era.

While still grappling with the particular instinct/intellect dichotomy, Linklater acknowledges that meaning is constructed through a combination of both the individual and the text they are speaking. In her book *Freeing the Natural Voice* (1976), her process revolves around the individual through whom the text is "revealed," and around processes through which "interpretation of the text [is] released from within" (185). No longer neutral mouthpieces through which the authority of the author emanates, actors in Linklater's approach now possess the important job of *interpretation* of the text. Linklater's experience with method acting and American psychotherapeutic approaches

leads to a conception of language that mirrors contemporary linguistic approaches, that incorporates physiological and theoretical meanings of language: “Words have a direct line through the nerve endings of the mouth to sensory and emotive storehouses in the body...That direct line has been short-circuited, and the beginning work to release the build-in art of eloquence must be to re-establish the visceral connect of words to the body” (*Freeing Shakespeare’s Voice* 174). Her consequent privileging of imagery that arises over the text reflects a conception of language in opposition to Berry’s detached definition of language. To Linklater, words are places and are embodied conceptions of a speaker’s interaction with the world. However, the issue with Linklater’s approach is that the goal for ultimate subjective human truth leads to a hyper individualistic approach to voice and text interpretation that may smooth over particular cultural conditioning.

Richard Knowles summarizes this point, “In attempting to *transcend* cultural conditioning en route to ‘the atmosphere of universal experience’ (Linklater 186) she allows for the *effacement* of cultural and other kinds of difference and is in danger of throwing the particular baby with the generalizing wash of her rhetoric” (103). The question remains in Linklater’s training whose “universal experience” is privileged over others, and the answer is still a homogeneous group of actors who had the geographic and economic means to access the training.

Linklater’s conception of voice training reflects overarching approaches to actor training, as Realism or Bennett’s “naturalist mode” of theatre making takes hold over both the British and American stages. Both approaches trade on the audience's knowledge and desire to see authenticity onstage, while eschewing the semiotic power of meaning-making in the theatre. In *Engaging Audiences: A Cognitive Approach to Spectating in the*

Theatre, Bruce McConachie interrogates the apparent tension between the training of the profession and exploration of the actual cognitive processes behind this training, as a way to access meaning-making in the naturalist mode of artistic production:

One rough parallel between therapy and science in our own field is the relationship between the teaching of acting and scientific research into how actors actually pursue their art. Most acting instructors will affirm that Stanislavski's "system," developed between 1906 and 1938, still has much to offer actors. This does not mean, however, that Stanislavski's explanation for why his system worked—a curious psychological stew dependent on the theories of Pavlov and Ribot—retains scientific credibility among research psychologists today. While there appear to be good reasons to continue to work with actors on the basis of "objectives," "obstacles," and "emotions," the acting class alone cannot become a laboratory to test for this scientific basis of Stanislavski's ideas (11).

This particular acting system still has some kind of efficacy and holds a prominent place in American acting and voice training in the Linklater fashion. Both acting and voice training from this time period make heavy use of the container metaphor of knowledge creation, where actors are conceived as empty vessels ready to be filled with knowledge. The teacher's responsibility was to confer acting as a skill into their student. The actor-as-empty-vessel metaphor is very similar to the Lockian blank slate metaphor and contributes directly to the notion that acting and other artistic skills can be created as pre-formed modular units that can only be conferred onto actors who have done the necessary prep work to become blank or neutral. In this sense, practitioners of this era assumed voice teaching (like the number of proprietary approaches) or dialects (sanitized and ready for ease of use) become discrete attainable chunks for the actor to master. While

many theorists have posited and continue to move past this ontological divide, voice and dialect professionals in the Linklater fashion continue to work in this actor-as-empty-vessel mode even while outwardly claiming their resistance against this mode.

3.2 A New professional organization establishes itself

To determine how pervasive these container approaches are in the profession of voice and dialect without examining the establishment of higher education theatre practices in the United States would be painting an incomplete picture of this practice (Zazzali 45). Theatre instructors adopted voice instruction focused on “freeing tension” to compliment the dominant acting style of Realism that was already being used in different theatre departments. The League of Resident Theatres (LORT), formally established on March 18, 1966 (one decade prior to *Freeing the Natural Voice*), sought actors who had enough vocal stamina to perform in different styles for extended periods of time (Calta 26). In response, several university theatre departments began training actors who could meet this demand and could enjoy long contracts as part of established repertory theatre programs, thus cementing the need for vocal coaches not only in these regional theatres, but in higher education institutions as well.

Training in higher education would explode at about the same time as regional theatres would be founded using grants from the government during the middle part of the twentieth century (Zazzali 47). Suddenly, theatre programs were concerned with training actors for the conservatory-style seasons of realism plays that regional theatres were building throughout the nation. Institutions were enjoying unprecedented financial support in the form of government supported grants and private foundations, offering incentive for higher education departments to train a legion of actors for stamina and longevity in their acting. This need for flexibility in acting style meant that an actor

trained to be a performer through not only acting classes, but through a newly founded discipline of voice and movement work. Establishing the intimate connection between institutions like the academy and newly established LORT cemented a type of authority that would create new modes of respect for the profession of voice training. This type of authority would lead to the creation of the first professional organization Voice and Speech Trainers Association (VASTA) to shape the expectations of the profession and further define their authority on the topic of voice training.

This professional organization bloomed from a series of casual gatherings that happened to coincide with the establishment of the League of Resident Theatres, in a way that helped to establish VASTA's prominence not only as a professional organization in higher education, but also as a profession that was required to produce quality theatre in the United States. To protect the integrity of this emerging profession, key players created the Voice And Speech Trainers' Association (VASTA) in 1968. Dorothy Mennen (VASTA'S first president) describes the first academic gathering of speech professionals in 1968 as "a dynamic session which fired the spark that initiated a new group called (at the time) Theatre Voice and Speech" (Moore 100). VASTA as a formal organization was established in 1986, nearly two decades after these gatherings began, by five women at the National Educational Theatre Conference in New York City. (Moore 101). Some of the functions VASTA would eventually fill were to issue evaluation guidelines for voice and speech trainers, a code of ethics and guidelines for training voice and speech trainers, and even advocate for promotion and tenure procedures for newly created voice professionals working in higher education. VASTA also established a publication, *Voice and Speech Review*, that has since become a prominent source of interdisciplinary voice and dialect research. This organization, in its first 10 years of existence, would become

the preeminent authority on who counted as a properly credentialed authority on voice and dialect. Membership has grown from 150 in the first five years of existence, to over 750 active members¹⁹. VASTA also draws its prestige and power through the many associations with other established academic organizations and networks, including the American Speech-Language-Hearing Association, The National Communication Association, The Voice Foundation, and the Association for Theatre in Higher Education (ATHE). Alliances with existing and well-respected organizations led to an air of authority and authenticity for the newly formed organization. VASTA continues to draw many different professionals across the voice training spectrum, from voice teachers in theatre, to singing instructors, and even linguists and speech pathologists.

Many members credit the success of VASTA to the role the annual conferences have played as a means of connection between various vocal professionals and their proprietary voice systems. Many key vocal professionals (many of whom appear in this manuscript) have presented workshops and keynotes, including Cicely Berry, Catherine Fitzmaurice, Jan Gist, Arthur Lessac, Patsy Rodenburg, Dudley Knight, David Crystal, Rocco Dal Vera, among others. VASTA conferences have become famous for their mixture of academic work and practical sessions, including sessions called “Things That Work,” a round-table that shared techniques, tools, and tactics, and “The Identity Cabinet,” a session where members can perform work that is “close to the heart” (Moore 103). Presenting at these conferences have become an unofficial requisite for acceptance into this organization, and to present at these conferences means one’s particular approaches had to be approved by the organizers in the first place. VASTA’s practices

¹⁹ Estimates using Adrienne Moore’s numbers in her 2019 article “The History of Voice and Speech Trainers Association (VASTA)” in the *Voice and Speech Review*.

are particularly insular because they sit at an intersection between academic and outside professional organization, meaning that VASTA can use both gatekeeping mechanisms for academic institutions and private professional organizations.

Another form of authority establishment came from the communication networks that VASTA's members established. One former president, the very same Dudley Knight, seeing a rising need for communication amongst members, established a listserv and email newsletter named VASTA Vox²⁰, which quickly became a location for members to exchange training tips, but more crucially revealing their various stances on different issues, such as incorporation of body techniques into voice instruction, but more crucially the use of standard dialects in voice instruction. This discussion, particularly around the status of standard dialects, flared up occasionally, as the topic inhabits a contentious space. In this case, the arguments often revolve around the use of created standard stage dialects, like Skinner's *Good American Speech*, and not necessarily regional or foreign dialects used onstage. Some practitioners who can trace their voice training lineage directly to Skinner defending *Good American English* to others as a relatively value-neutral "tool to teach phonetics," while others still claim that to teach these standard dialects continues to oppress marginalized actors and introduces white language ideologies into the classroom (Moore 104). Oblique references to heated discussion can be found in VASTA's outward publications in *Voice and Speech Review*, such as Rockford Sansom's 2016 article "The unspoken voice and speech debate [or] the sacred

²⁰ Early VASTA publications are built out of the debates that were had in this list serv. I have since tried to find an archive or other ways to access this list serv through asking early participants in VASTA Vox and it appears that all storage has been wiped out and no institutional repository of this debate exists anymore, apart from perhaps archived emails in members' own personal inboxes. This is the peril of referencing online discourse in the early 2000's; here one day, and gone the next when your school updates its IT capacity.

cow in the conservatory” that summarized discussions on VASTA Vox. Though the unspoken speech debate to which Sansom is referring centers around instructional styles, Sansom still references fierce online discussions via VASTA Vox in his brief outline of historical conflicts within the organization, including the central question of the use of standard dialects in the voice classroom (160). In the intervening years, the substance of this email listserv has since been lost entirely, and multiple attempts to retrieve the data have been unsuccessful. In these listserv conversations, often more senior members, officers and board members would often exercise their control over the conversation by chiming into more heated debates, essentially silencing minority opinion and junior members in their discussions, thus creating a hierarchy within the organization itself (Sansom 160).

3.3 The hunt for the perfect dialect: Midcentury dialect coaching

Following the success of Herman and Herman’s manuals of dialects, the next generation of dialect professionals sought the next evolution in dialect coaching. The innovation came in the form of easier access to audio tapes as supplements to the written manuals. These audio tapes were often pre-recorded exercises by the dialect coach to aid in acquisition of dialects, and thus were sanitized and stereotyped versions of the dialects in question. Access to audio tapes also aided dialect coaches in collecting samples from spontaneous speakers of the target dialects, which would lead to one of the biggest debates in the field. Dialect coaches often debated the use of standardized dialects versus the use of spontaneous real-world examples of dialect. On one hand, synthesized dialects contained fewer target sounds for the actor, which encouraged faster training, on the other hand, practitioners claimed that use of spontaneous dialects created a more “authentic dialect” (Blunt, xx). These authentic dialects were still filtered through actors

who often were not given an adequate amount of time to acquire these new skills, which means that the use of authenticity in this regard should be regarded cautiously. The use of spontaneous dialect would present its own challenges, as often dialect coaches compared their idea of a regional or foreign dialect to the speakers they found in those areas. Jerry Blunt describes this struggle between intelligibility and authenticity in his own manual *Stage Dialects* (1967).

Jerry Blunt published his own book titled *Stage Dialects* in 1967. In this book, he details eleven of what he perceives as most used dialects in the literature of theatre. In a move that already privileges white listeners and speakers, Blunt provides instructions for ten dialects of European or Anglo descent (Regional American, British English and a few foreign accented European dialects), and a Japanese dialect. The prime feature of this book is access to practice tapes for students of all stripes, and his dedication to creating authenticity in each dialect by featuring work done with spontaneous audio samples from across the world. By the design of his chapters, and the discussion of the sources of his dialects, Jerry Blunt reveals his position towards speakers of non-standard and foreign-accented varieties of English, which, while not nearly as explicitly antagonistic as Herman and Herman still bears the hallmarks of a normative approach to language in performance. Blunt carries with him standard dialect attitudes similar to Herman and Herman, where his concern is the assumed white listening audience of stage and film.

One of the ways in this manual that bears this explicit normative stance is “Standard English,” which Blunt claims is the Received Pronunciation or constructed British dialect of Daniel Jones, who documented the dialect phonetically in his *English Pronouncing Dictionary* (52). That Blunt does not include “British” in his title for this dialect speaks to his opinion on how this particular dialect ought to be considered as the

gold standard for pronunciation in production. Even in his claims, however, Blunt admits that this Standard English holds a paradoxical position as the dialect of the educated high class speakers in England,

Standard English is self-conscious. The habitual user is as aware of his speech as he is of his posture or his social deportment..When a speech with this characteristic is used inappropriately, its basic nature is changed from what it is to what it was never intended to be, and a falseness or affectation results (52).

In contrast, Blunt's reference to American English points to his dismissal of the use of a generalized American dialect in terms of authoritative uses on stage. The American English dialect in performance "has no authoritative standing, but is needed to specify the most widely employed form of American speech. It is the dialectal utterance of Midwestern and, more recently, Western groupings" (2). Blunt appeals to education and class despite no real-world use when advocating for Standard English in his chapter while simultaneously appealing to frequency of use when he admits that general American English is needed for performance. In other words, he does not uniformly apply criteria for standards of use for his dialects.

For other accents, Blunt provides tapes as optimal or perfect examples to accompany the text, demonstrating the vowel and consonant shifts for each dialect. His tapes are sanitized versions of the dialects that are in the book; the speakers on the tapes are imitating each dialect after much training. While training these dialects, Blunt details collecting numerous primary sources as reference. Blunt admits to frustration in collecting these primary sources, particularly for the non-native accented English he sought out in Europe. He blames formal education (which he previously celebrated in the

Standard English chapter) for denying him the stereotypical accents he came to expect while overseas,

The core of the problem lay in the fact that the Italian living in Italy learns his English in school under the eye and ear of a teacher who places emphasis upon correct grammar and pronunciation. In contrast, the foreign accent we Americans have come to know is a slightly taught patois developed by the foreign-born living in America (4).

Blunt exalts education as reasons for including Standard English, but education becomes an obstacle when Blunt is pursuing primary sources for foreign accents for his own book. Blunt positions speakers whose English is their second language as second-class speakers of their own dialect. He seeks to extract accents untouched by formal language education without admitting that one of the only ways that these speakers can advance in society is seeking formal education, an education for which Blunt has a position in creating standards. In the introduction of his book, he admits, “more usable foreign accents can be found at home than abroad,” tacitly admitting that he is seeking a stereotype of the foreign-born immigrant whose “need for communication must of necessity bypass rules of grammar and the niceties of pronunciation” (4). Faced with an impossible position, immigrants to the United States must be able to communicate in English, but not sound educated enough according to Blunt’s expectations of intelligibility for a foreign accent. While uneven in his application of criteria about what qualifies as an accent worthy of study, Blunt did innovate the field with the use of practice audio tapes. This sets the stage for the next generation of dialect coaches, who will employ more than just audio, and more than just the samples directly collected by the dialect coaches themselves.

4. Voice practitioners join the Internet

The succeeding contemporary generation of voice professionals who were students of those promoting their own proprietary systems of voice have access to far greater amounts of knowledge than any generation previously. While this generation still focuses on voice, body and breath work like in Linklater's *Freeing the Natural Voice*, their access to information via the internet enforces their authority by appealing to cutting edge empirical knowledge in voice (Bartoskova 2). As a consequence, practitioners make more use of the internet as both repositories of knowledge and advertisements for their individual approaches to voice.²¹ Practitioners also saw the regional theatres and theatre departments that once enjoyed unprecedented governmental and private support shrivel in the wake of Neoliberal policy making of the 1980s (Zazzali 201). Grants dried up with the resurgence of the Neo-liberalist approach to funding, leaving these regional institutions with large built-up capital and buildings, and shrinking budgets for actual production work. Regional theatres that had built immense amounts of capital construction and fossilized into an unsustainable funding model reliant on these grants had to scramble to appease as many individual donors and patrons as possible, limiting the type of work presented to classics and easy-to-digest works of theatre (Zazzali 202). As a result, actors who had enjoyed relatively stable months-long contracts with theatres were now faced with a creative economy that would only offer job security for the length of one production before an actor would have to find more work. Voice and dialect coaches who worked for LORT houses also found their jobs becoming less stable. This destabilization forced contemporary voice coaches to seek jobs in higher education (slow

²¹ Evidence of these websites can be seen in the shift from use of printed publications like books to more internet sources throughout this section.

to respond to the declining demand for conservatory-style actors and still a relatively safe occupation) and create side businesses of accent modification and dialect work through the use of the internet. Beginning with this generation, expectations of voice teachers include finding themselves with a plethora of choices when considering the type of training they want to add to their curriculum vitae to remain competitive as voice experts.

Several key practitioners define this generation through a combination of published work and holding leadership positions within VASTA. One key and influential practitioner in this era include Louis Colaianni, who, as a nod to the elocutionists and their establishment of the International Phonetic Alphabet, released his book *Joy of Phonetics and Accents* (1995), a book that joins the practice of embodied tension release with an emphasis on the use of the phonetic alphabet. Another duo, Dudley Knight and Philip Thompson, began work that similarly borrows psychological research and concepts such as Neuro-linguistic Programming, which makes popular the idea that certain people have certain modalities in which they learn best (Knight and Thompson iii). Like the previous generation of practitioners, Dudley Knight and Philip Thompson offer training certificates in official Knight-Thompson training that lends a veneer of authority to any voice trainer trying to gain an edge in the tightening labor market.

Louis Colaianni can directly trace his lineage to Kristin Linklater and her system, which influences his own approach to voice profoundly. Colaianni adapted Linklater's system of freedom of tension and added what he calls his signature exercise, the Phonetic Pillows approach, through his book *The Joy of Phonetics* (1995). Colaianni's approach has proven incredibly popular, as he has taught in many higher education institutions, along with coaching in some of the largest performance institutions in the United States

(“About”). Riding the medium line between strict standardized learning of sounds and the freedom of imagination has become a lucrative approach to voicework, as Colaianni has offered several workshops in conjunction with the Linklater Center For Voice. In contemporary times, having a Linklater stamp of approval lends authority to Louis Colaianni and his approach to voice. Colaianni, in a credit to his popularity and system, has also made the jump to film, naming many famous actors including Bill Murray and America Ferrera for film (“About”).

Through turning his attention to fine phonetic detail (in a way similar to Edith Warman Skinner) and maintaining freedom from tension (via his training through Linklater), Colaianni has attempted to skirt issues with standard stage dialect and oppressive practices by attempting to strike a balance between these approaches of previous generations. His phonetic pillows, a set of large stuffed phonetic symbols derived from the International Phonetic Alphabet and used as embodied stimulation in the voice classroom, have become a part of his successful voice system. He claims,

One dimensional phonetic symbols, printed on paper, tell our eyes what sounds they represent, tell our ears what sounds we are expected to utter, but make little or no appeal to our imaginations... In an effort to bring phonetics into the same physical world as other performance classes I have worked with student actors for many years on ways to get the symbols to jump from the page, enter our bodies and demand us to express them (“About Phonetic Pillows”).

Colaianni favors the embodied experience and approach to voice work, and implicitly reinscribes the split between “intellectual” pursuit of voice and “embodied” pursuit of performance. Colaianni seems to equally use empirical pursuits, like phonetics and

linguistics, as much as he uses embodied individual knowledge in his work. He is pushing the needle from individualized psychological and social work in the actor back to empirically minded, a subjective/objective balance that this dissertation also attempts to strike.

In contrast, Dudley Knight and Phil Thompson declare that their system of training does not use standards in the same way as training from previous eras. Dudley Knight and Philip Thompson, after developing their working relationship in the founding of VASTA, created a system of training called Knight-Thompson work, publishing *Speaking with Skill: An introduction to Knight-Thompson speech work* in 2012. This approach claims to be standard English agnostic, as Knight and Thompson claim that intelligibility is the ultimate goal of the work they offer (“About the Work”). As analyzed in the introduction, Knight’s use of the term intelligibility references directly an appeal to scientific authority, by claiming there is some objective measure that is resistant to social construction of what it means to be understood easily. Knight-Thompson work also uses accents as a natural extension of their analytical approach to accent. The K-T approach introduces the actor to “the four p’s” *person, posture, prosody* and *pronunciation*,

By addressing characteristic sounds with reference to the speaker’s system of sound categories, the inherent variability in the realization of these sounds, and the relation of these sounds to the speaker’s vocal tract posture, actors can more confidently achieve an accent performance that authentically represents the speech of the character (“About the Work”).

Authenticity again appears as the ultimate goal for voice work, both within generalized

voice and within dialect, while the remaining question is, to whom does the voice work sound authentic? The possible answers about authenticity include the original speaker of the dialect, the expert voice practitioner or director, and the theatrical audience. The implied answer from their work seems to indicate that authenticity is determined by the practitioner, thus again giving authority to the voice coach in this work. Knight and Thompson present authenticity in this work as the ultimate goal, without qualifying *why* practitioners, actors and audiences alike ought to make authenticity the goal. The contemporary generation of voice professionals will challenge this implied goal.

4.1 Monich coaches for the movies

One of the most prolific and famous dialect coaches from this same time period is Tim Monich, whose Hollywood pedigree includes students such as Brad Pitt, Shia LeBeouf, Gerard Butler and many others (Wilkinson). His entry in the Internet Movie Database (IMDb) includes over 192 credits as dialect coach in a career that expands from 1983 to current productions slated for release in 2021 (“Tim Monich”). Like practitioners before him, Tim Monich can trace his voice and dialect training lineage directly to Edith Skinner, who was his teacher at Carnegie Mellon (Wilkinson). Monich even helped Skinner edit her text *Speak With Distinction* (1990), which speaks to his practice with precision phonetic symbols in training actors. While the lion’s share of Monich’s work is for television and film, however, he has hundreds of credits for theatre as well. Because of his prominent place in the film industry, Monich might be the most recognizable American dialect coach in the contemporary era.

Like Blunt, Monich is an avid collector of samples for his dialect work, possessing enough recordings for fifty three consecutive days of listening (Wilkinson).

He works with a number of famous actors, often giving them options for linguistic models with which to work. Monich, while trained in the elocutionary style of Skinner and her proteges, has adapted his technique for dialect training to rise to the challenge of the fast-paced world of entertainment and film. This means he is adapting his dialects to the skills of the actor, the desires of the director, and the overall look and feel of the film. His approach has earned him accolades from several highly influential actors and directors, including Martin Scorsese, and as a result enjoys references for hundreds of accent and voice jobs in television and film. This means Monich has often become the most requested dialect coach in Hollywood, in an increasingly tight market that favors a few coaches at the top of the production hierarchy, leaving many more to carve out a living as coaches for independent films, television and theatrical productions. Because of this structure, many accent and voice coaches have turned to modern social media and technology to piece together enough work to live as a voice professional.

4.2 Digital Approaches to voice and dialect

From the foundation of these key practitioners, contemporary voice and dialect professionals are beginning to approach voice in new ways that attempt to dismantle the harms of this practice. Harm reduction begins with increased collaboration between voice and dialect coaches and voice scientists and other experts in language, via research collaborations that often appear in the pages of *Voice and Speech Review* (Bartoskova 1). These collaborations reflect Colaianni's practice, where the voice is treated as both a creative vehicle for expression, but also a scientific object of study. This type of research celebrates the tension between objectivity and subjectivity of the voice. The modes by which this work is disseminated has shifted in the next generation, mainly with the use of

personal websites, social media, YouTube and other digital means of communication. Accent and dialect coaching also benefits from this online explosion, yet remains unregulated. The explosion of online voice coaches represents a threat to the model that has been established in the second half of the twentieth century by voice and dialect coaches who have established themselves as part of an organized professional association, draw upon the authority of higher education, or the depth of experience as a seasoned coach for stage and film. Online voice coaching, in other words, is not subject to the tight gatekeeping or controls of the previous generations of voice professionals.

A cursory search on YouTube reveals hundreds of channels that are dedicated to the topic of accent and dialect coaching, where some of the more popular personalities have amassed over 100,000 followers combined.²² Not all channels that appeared in this search are entertainment or actor training focused; there are plenty of channels, like Dr. Geoff Lindsey and Linguistix Pronunciation, where the main goal is to help non-native speakers learning a second or third language to achieve more native-like pronunciation in their everyday lives (“Linguistix Pronunciation”). These practitioners exist through a large and profitable business called accent reduction, also known as accent modification or accent neutralization, which is an Anglo-centric profession aimed at speakers of English as a second language to acquire more native-like pronunciation (Hope 10). These professionals and videos are catering directly to the idealized white listener and will often demonstrate different contexts for appropriate registers of language. Other dialect and voice trainers participate in what is called affirming voice therapy. A person’s desire to change may stem from the desire to have a voice that matches the gender identity a

²² I searched the term “dialect coach” for channel names on YouTube on May 3, 2020.

speaker may want to project to the world. In this sense, treating all accent reduction or modification as inherently bad may exclude individuals who may wish to change their voices to match their gender identities (Nolan et al.1368).

For actor training in both film and stage, these channels and videos are part of a larger trend of individualized and freelancing vocal experts, of which the largest sector of voice professionals is voice and singing instructors (“Find a Voice Pro”). Often, videos in the tradition of accent and dialect coaching feature such language as “learn to improve or neutralize your accent!” and, “sound like a native speaker FAST!” which preys both on the precarious position of marginalized speakers and the accelerated nature of theatre and film production for actors. No YouTube video, no matter how thorough and engaging, holds the secret or key to accent neutralization or improvement, because the work of neutralization always caters to the normative white English-speaking listener. Dialect and accent coaches who have created these videos are engaging with raciolinguistic ideas of how to sound without interrogating the damaging normative language attitudes behind these sentiments.

These dialect and accent coaches are often held up as popular culture icons, a type of expert to reference when discussing accents in entertainment. They often derive their authority from the sheer popularity of their videos and materials that are available on the internet. One of the most popular dialect coaches found on YouTube does not have his own channel, however, but is often called upon by popular general media producers like *Wired* and *Insider* as a particular expert in movie dialects and accents. Erik Singer has starred in over 10 videos that amass 1–13 million views each.²³ He stars in a series called

²³ Number of views was assessed on May 3, 2020 using YouTube search for “Erik Singer” and he has since published more videos.

“Technique Critique” where he breaks down several different accents in different scenarios in film and Television. The most popular video in which he is featured is titled “Movie Accent Expert Breaks Down 32 Actors' Accents” where he analyzes both accents considered good and bad in his own opinion (*Wired*). In these videos, Singer showcases his expertise and positions Singer as the undisputed authority on accents in popular entertainment. Subsequent to the original video, Erik Singer has also been featured in another video titled, “Movie Accent Expert Breaks down 31 Actors Playing Real People,” where he breaks down actors' attempt at ideolects, or specific individual accents (Singer *Wired*). The practice of imitating specific individual accents seems to be a genre peculiar to film and television, and strict adherence to recreating painstaking details of individuals' lives is often rewarded both financially and critically. The popularity of this video has led to another video where Singer breaks down 17 more film performances of ideolects in film.

Often, YouTube channels are an external-facing part of the advertising apparatus for individual freelance voice and dialect coaches. Coaches will create viral-like videos in the format of a “talking head” explanatory video, or a demonstration of linguistic prowess through demonstrations of particular accents. These explanatory videos will be uploaded to YouTube, Instagram, Facebook and other social media and link to further services that they offer through their own personal websites. Services offered through this viral-like online presence can sometimes include one-on-one coaching via video conferencing with actors or speakers who wish to change their accent²⁴ In this way, dialect and accent coaches are offering their services to a larger geographical and wider demographic than

²⁴ The popularity of this type of training seems to have risen with the advent of the COVID-19 pandemic.

would otherwise be possible if they were limited to more traditional approaches to training. Social media in this way are used as a type of personal brand advertisement and part of an independent business model. In this way, the digital profession of dialect and accent coaches recreate the inequality and power structures inherent within “influencer” styled professions. This means that a small number of popular or well-known dialect coaches enjoy the benefits of this structure while constantly navigating an ever-changing algorithmic landscape amassing followers and creating income, while a large number of dialect coaches do not (Cotter 904).

Dialect coaches also use the power of YouTube and other audio corpora for various stages of research for different dialects in the tradition that Blunt established. Various videos and snippets of audio are available online for a coach to sift through and utilize in their work, though they may need to be careful because YouTube videos are not always accurately labeled. Coaches use videos uploaded by the speakers of the communities or accent of interest, with no explicit connection to accent or dialect work. A coach must consider the ethics while using the vast trove of internet resources that are available to dialect coaches and would-be actors who are looking for reference accents for their own work. Technically, users of popular social media websites YouTube, Instagram, and Facebook all must agree to user agreements that hold that material that is uploaded publicly does not necessarily belong to the user. However, there are perceptions in this work that material uploaded (especially to locked, unlisted or private accounts) should not be used by other users on the website (Grover et al. 772). While not specifically illegal, use of these audio and video sources can create friction with the perception of ownership of the original material that has been uploaded. Despite legal

murkiness and perception of ownership, the ethical question remains whether dialect and accent coaches may freely use spontaneous language and accent material for their own work, considering that these speakers may not necessarily explicitly consent to their lived linguistic lives being adapted for stage or film. Like using people's images in film and stage, dialect coaches ought to adhere to stricter ethical codes in use of people's linguistic likenesses for dialect and accent usage.

The internet has facilitated both a rise in access to this type of work and an influx of digital influencer-styled accent, voice and dialect coaches without much ethical oversight. Coaching via the internet provides larger access to actor training and to resources, but also presents an issue of licensing and qualifications for doing this kind of actor training. In contrast to the establishment of VASTA in the generation prior, online dialect and accent coaching requires no formal membership nor formal training certification to create a business that caters to actors and speakers. Both approaches present their own advantages regarding access and professional ethics. On one hand, more people have access to resources and coaches online, while those who have access to officially vetted or VASTA members are limited to students in higher education or participants who can pay the fees associated with official voice training programs such as the Linklater Center in New York City and Orin, Scotland. The contemporary issue voice and dialect coaches must face is how to balance the gate-keeping privilege of VASTA with the unregulated generation of online voice and dialect coaching that borrows its business model from influencer-style online promotion. Inherent in both models of access to this discipline is still the prevalence of implicit and explicit biases that contribute to the enforcement of negative Linguistic stereotypes seen through entertainment. The tension

between access and professional training informs many of the best practices that I discuss in the conclusion of this dissertation.

5.0 Towards a cognitive conception of voice training

While the bulk of this chapter has explained the arguments and assumptions practitioners use in training, I turn my attention now to how theatrical audiences factor into this work. My focus for the following chapter will be the second side of intelligibility, which is the perception of how understandable speakers are to the average audience member. Intelligibility begins with listening, both for the actor and crucially in the audience. Both of these types of listening are subject to normative language ideology, especially given the privileged arena of theatrical speech. Voice trainers have conceived of theatrical audiences as arenas that require extra perceptual expertise and have historically fashioned their work around this expectation. Voice trainer Marian Hampton speaks of clarity and intelligibility in her opening essay on standard language,

As teachers, we must guide students in listening astutely to the speech of others so that they may adopt those characteristics which will contribute to the establishment of character, yet choose carefully what will help in this process without destroying an audience's ability to understand the text of the play... We've all seen and heard productions in which the accent, albeit accurate beyond question, is so broad as to render the play unintelligible (15).

Hampton uses her conception of the near universal experience of perception of the audience to establish the knowledge base upon which she judges accents for actors, and simultaneously pits accuracy—not authenticity—against the needs of the audience. Intelligibility to Hampton is placed squarely on the execution of an “accurate yet broad”

rendition of an accent and leaves little room for audience autonomy in the interpretation of said accents. This refrain is the foundation of the profession of voice and dialect training after nearly one hundred years of philosophy and work of successive generations of voice professionals. The history of this profession necessarily informs the modern conception of intelligibility in its colloquial use by this profession. Understanding intelligibility in this way, however, eliminates the autonomy of individual audience members and their experiences. Honoring the autonomy of the experiences of audience members is the intervention of the cognitive conception of intelligibility. The second half of this dissertation will center individual experiences of audience members from a cognitive perspective, thereby creating room for contemporary voice and dialect training to embrace a more diverse and deeper understanding of intelligibility that will result in the inclusion of a more diverse group of performers and audiences.

The emphasis on actor training in this profession misses the audience's role in meaning-making in production. Through the use of "audience reception," Susan Bennet weakens the audience's role in meaning-making processes when experiencing performance, by reducing audience members to passive receptacles of meaning in the theatre (4). Reception implies a passive almost literary role for each audience member, which limits each audience member's agency as a participant in the theatrical event. Bennet uses reception in part in response to other burgeoning theories of the time, most famously the reader-reception theory, thereby extending the metaphor of literacy or "reading" to a theatrical event (6). To counteract this passivity, I propose shifting from reception to *perception*, borrowing from psychological use of the term. Perception recognizes a person's participation in *consciously* or *subconsciously* recognizing and

organizing sensory information due to a number of ecological and psychological factors (Michel). This shift in terminology both gives agency back to the audience member, but also acknowledges the complicated and precise cognitive processes that are activated when an audience member creates meaning for the performance they are witnessing in a way where the voice profession has flattened individual experiences into the professional's expectation of audience experience.

Audience spectatorship is not necessarily a uniquely cognitive activity but uses human cognitive faculties available to other modes of perception in a unique configuration for theatrical spectating. In Dr. Thalia R. Goldstein's article, "Questions of Realness," where she debates the role of cognition in Realism, she claims "Theatre is obviously artifice," yet audiences have come to expect this artifice as a stand-in for authenticity. More so, theatre uses real humans in real-time; even the most experimental of forms still often include humans on stage. Several questions about the tensions related to audience experience of authenticity and artifice may unlock the secret to creating a balance between artificial dialects and accents with accurately portraying accents represented onstage. How does the audience parse what is the artifice of theatre in the form of performance, versus the real automatic cognitive reactions to witnessing human behavior in real time? How are we balancing these imagined scenarios with the reality of our automatic responses that have been shaped by our experience in the rest of society? The question of representation is particularly pressing when audiences can perceive that certain accents are representing certain and often derogatory character traits. The ultimate ethical responsibility of presenting these accents lies with the voice professionals and production teams, which becomes fraught with the historical resonances of so many

harmful practices of the past.

Examining the profession of voice and dialect may not be enough to untangle expectations around authenticity and intelligibility. According to Bruce McConachie, contemporary approaches to acting in general, even while they owe their origins to Stanislavski, heavily employ the actor-as-container metaphor, and envisions actor bodies as empty vessels ready to be filled with emotion and the psychological means to access character (44). The words spoken onstage amount to what Stanislavski called “verbal action” and ought to be considered an integral part of any acting approach but more importantly, a reflection of the environment in which practitioners and audience members find themselves (Moore 69). Instead of asking what an actor can fill themselves with, cognitive approaches to acting conceive the actor as a permeable part of a larger system, asking, “what does it mean to build characters from the ecosystem up, rather than a more psychologically focused method of character assessment?” (Cook 117). This approach necessarily considers the contexts in which audiences and performers find themselves, which lends a sharper vocabulary and tools to confront the overarching raciolinguistic ideologies imposed by a society with normative listening ideologies. Conceiving actor and voice training as an ecological event, inextricably connected to the context in which theatre is created leads to a larger access to empathy, and new pathways for meaning-making for both artists and audiences alike.

By reconsidering both spectating and training as embodied processes that cannot be separated from ecological and social conditions, theatre producers can uncover new and surprising ways to make meaning on the stage, and by extension can conceive of new ways to approach voice training. These surprising ways are bolstered by evidence from

several neurological studies. Italian scientists made an exciting discovery in the early 2000s when they observed the motor neurons (and not just the sense neurons) in monkeys' brains light up when they watched their handlers perform actions with their hands (McConachie 70). The use of motor neurons points toward how empathy might be built in the brain, seeing an emotion could mean the perceiver is activating and feeling that same emotion in their own brain. In his book, Bruce McConachie says, "visuomotor representations... provide spectators with the ability to 'read the minds' of actor/characters, to intuit their beliefs, intentions, and emotions by watching their motor actions...Empathy is not an emotion, but it readily leads viewers to emotional engagements" (65). Conceiving of empathy as an automatic cognitive process in theatrical audiences, theatre producers can stop wondering about the necessary and sufficient conditions to create empathy in audiences and instead consider audience members as an integral interactive part of the process of theatrical creation. In this way, producers and audience members alike not only access new modes of meaning making but can justify in a very real way the role that performance plays in our social fabric (Dissanayake 89).

Activating empathy and subsequent emotional experience drives the interactive cognitive model of audience perception, deconstructing the rational model of audience experience that divides subjects from objects and assumes a separate rational world where meaning is made. The antidote to this objective approach can be found in Lakoff and Johnson's concept of embodied realism. They write in *Philosophy of the Flesh*,

The alternative we propose, embodied realism, relies on the fact that we are coupled to the work through our embodied interactions...what disembodied

realism misses...is that, as embodied, imaginative creatures, *we never were separated or divorced from reality in the first place* (emphasis author's own, 93). Crucially, this rational objectivism destroys embodied experience as a mode of meaning-making. The assumption that objects lie "out there" and subjectivity lies within the audience destroys an opportunity to conceptualize meaning creation as a connective ecological event, dependent upon the exact conditions and contexts of each performance.

The kernel at the center of the voice and dialect profession is the very use of voice, which practitioners can also approach through embodied realism. In very concrete terms, the voice is the result of the physical configuration of an individual's vocal tract and the subsequent effect that configuration has on how air travels through that system. The essentials of having a voice requires the vocal tract, but also usually the movement of air—which is usually provided by the lungs or colloquially the breath. The vocal tract has some features that can be consciously manipulated, while components of voice are much harder or impossible to change. For example, physiological features, like vocal tract length, and medical issues like a deviated septum are nearly immutable, while placement of the lips, tongue and jaw can be changed rapidly. The combination of these mutable and conservative features creates an individual sound or voice. Honoring both the physiological circumstances and the context in which speakers and listeners find themselves completes the picture of understanding the role of intelligibility in voice training. Embodied realism even supports the metaphorical use of voice—actors and even playwrights are encouraged to "find their voice" when performing in theatre. The very experience of using your voice in the theatre lends itself to metaphorical ways of conceiving of theatrical practice. In *Metaphors We Live By* George Lakoff and Mark

Johnson argue that nearly all metaphorical language that we use comes from a deeply embodied and very much non-metaphorical lived experience. From there, lived experience ought to be centered when creating meaning-filled work such as performance.

By examining the assumptions and underlying philosophies of these trainers through the lens of cognitive audience studies, I can bring the discipline of voice training into a more contemporary conversation with actor training and cognitive understandings of how humans make meaning out of the world around them. The historical conception of voice training contributes to contemporary understanding of knowledge transmission from each generation to the next, and demonstrates where subsequent practitioners transmitted ideas while others resisted ideas shows clues to how the voice profession is situated in the larger ecosystem of theatre production. In the tradition of pushing back against prior generations, I advocate for a system that meets actors and audience members from where they linguistically hail, explicitly honoring the diversity of experience that has led them to inhabit a small dark room together to experience a specialized form of communication. Yet, the process of the audience's use of intelligibility in their experience of performance remains a large question as part of this cognitive conception of voice. To probe deeper into the role of intelligibility and meaning-making for individual audience members, I will introduce my own linguistics cognitive studies in the following chapter. My empirical investigation created around these very questions coupled with cognitive humanities studies will demonstrate how audiences can subjectively construct intelligibility onstage and how intelligibility can no longer act as a reliable objective measure for voice professionals and actors to use in their work. Finally, I will follow that chapter by picking up where this critical history leaves

off, with an eye towards the future of voice discipline. I will discuss the needs of contemporary theatre makers for voice training, along with highlighting some practitioners who I believe are on the right path to account for historical biases and constructions of authority in this discipline. I will use results from the linguistics chapter to further explore our cognitive understanding of how individual audience members use context and their prior experience with authority to construct intelligibility of what they see and hear onstage.

CHAPTER III

CONSTRUCTING AUDIENCE INTELLIGIBILITY USING EMPIRICAL INQUIRY

“. . . [O]nce most people really come to understand what an embodied conception of mind entails, they are going to be upset about it. Much of what they hold dear is at stake – their view of mind, meaning, thought, knowledge, science, morality, religion, and politics.” -Mark Johnson, *The Meaning of the Body*, 15

1. Rationale for empirical approach

In the previous chapter, I examined the assumptions and ways that dialect and voice trainers construct their authority and knowledge in their field, along with their understanding of how audiences perceive performed speech through an in-depth look at how the profession is constructed. This chapter will challenge the ideas that established the profession by asking specific empirical questions about audience understanding, or intelligibility, which appears to be the yardstick by which voice and dialect coaches measure the effectiveness of their training. Many voice trainers measure the success of actor voice and dialect training through their perception of how intelligible the actor sounds on stage. Cognitive perception and meaning making by audience members are highly context dependent and built over time according to the experience of the listener, trainer, and actor. In this chapter, I take aim specifically at intelligibility as a socially constructed phenomenon that is the direct result of speaker (performer), listener (audience member), and the specific context (e.g., performance space, context of the story, previous historical encounters with voice).

Contemporary research of speech perception supports the ideas of context dependent constructions of intelligibility explored in the last chapter. I will demonstrate this by constructing my own empirical research that tests the question of specific social contexts and experiences in audience members. To test the specific influence of performance context in this construction of intelligibility, I have devised two empirical

experiments that manipulate the listener's belief—asking whether explicit knowledge that what they are about to hear is performed speech versus spontaneous affects how they judge the voices they hear. What happens to a listener when they are expecting a context with maximum intelligibility; how are they constructing the voices they hear in the context of performed speech? In this chapter, I use the idea of expectation of performance as a direct stand-in for expectation of intelligibility, and will simply use the term “expectation” throughout these experiments as a shorthand phrase.

The best way to approach the gap in voice practice is to use a field of inquiry that specializes in understanding the mechanisms of linguistic perception, drawing useful information from primary studies, and creating a theory from whence practitioners can work. Combining these two fields takes a careful approach because the vocabulary in one field can overlap with the vocabulary in the other field, while having two different meanings. For example, while researchers in linguistics operationalize intelligibility as accuracy of speech transcription, voice practitioners use intelligibility as a general measure of audience understanding. In this case when I refer to intelligibility, I must be careful to either highlight the colloquial usage by theatre practitioners or use intelligibility as a linguist. In some situations, intelligibility for both hold similar meanings; when practitioner Dudley Knight uses intelligibility, he carefully defines his usage as the “amount of linguistic information a listener can gather” (Knight 140). In contrast, intelligibility is defined by Derwing and Munro as, “the extent to which the native speaker understands the intended message” and is specifically measured by recall of key terms in subsequent experiments (2). The difference lies in the expectations of knowledge recall of the listener. Untangling this distinction between linguistic usage and

colloquial usage will be part of the work of this chapter, which is essential in the task of truly understanding how audiences perceive language onstage. Going into depth about the different uses of this term specifically will illuminate the gaps in knowledge that practitioners have been carrying despite their years of embodied subjective knowledge. A fresh new perspective on the terms any profession is using ought to be welcome at any stage in training.

Using research from an adjacent empirical field is the practice of interdisciplinary research of the cognitive humanities. I am continuing this tradition with this research, with one difference. Often cognitive humanities make use of research studies as the primary sources of theorizing. The research in this chapter offers a unique intervention by featuring a custom designed study, which means this chapter will consist of literature review of relevant linguistic studies, a lab report of the experiments I conducted, and then a subsequent discussion of implications of the findings in the report that incorporates theories from cognitive humanities. The literature review is thematically organized. The lab report is often the primary resource of cognitive humanities research and therefore summarized without the raw analysis; this dissertation offers the opportunity to examine the research in the form that cognitive humanities refer to but does not often display. In the future, I hope that more theorists in this field choose to work with linguists and scientists to present their primary findings in an accessible way for humanities researchers and practitioners alike. Combining interdisciplinary research into a format that is accessible to both disciplines has always been a goal of mine, and I will use this chapter as a blueprint for future research.

Numerous linguistic research studies support the idea that the objective-sounding

measure of intelligibility is susceptible to social context and standard language ideology, through the various social experiences of different listeners or audience members. Despite this, in his discussion of standard accents and intelligibility, voice practitioner Dudley Knight off-handedly laments, “it appears that little if any research into intelligibility has been done up to now” (70). Knight is mistaken; since the foundational 1960 Wallace Lambert et al. article “Evaluational reactions to spoken languages,” thousands of articles on speech intelligibility have been published in the field of linguistics and our understanding of the social effects of language continues to grow. This means, contrary to Knight’s claims, that objective measures of language perception such as intelligibility are subject to language standards, especially in voice and dialect training where the environment explicitly judges speakers on their perceived intelligibility.

Linguistics provides the tools and vocabulary necessary to explicitly examine how listeners use intelligibility to construct what they are hearing on stage and provide evidence for the embodied realism approach to knowledge construction of Lakoff and Johnson (4). Listeners are exposed to significant variation in speech, including different accents and dialects, and subsequently they hold variable beliefs about how ideal language should sound. Regardless of the variability encountered in speech, listeners can not only parse information from speech, but also associate this speech signal with perceived physical and social qualities of the speaker²⁵ (Agheyisi and Fishman 146). Generally, accent perception and, by extension, meaning construction can be construed as

²⁵ Agheyisi and Fishman summarize the use of attitudinal matched guise studies investigating these qualities including various languages, regional dialects, races, socioeconomic status, religion, and gender.

a combination of relying upon two different types of cognitive processes, bottom-up or subconscious processes, and top-down or conscious processes. Bottom-up processes relate to how automatic cognitive processes decipher the acoustic signal that strikes the ear drums in the listener (Rauss and Pourtois 276). Top-down processes which help the language users predict patterns in the signals they are hearing mold the perception of these acoustic signals (Rauss and Pourtois 276). This investigation closely examines the effect of using social context in top-down processing in accent perception and meaning construction in performance. These top-down processes are affected by conscious and subconscious socially ingrained ideas about language, gathered through a lifetime of being a language user. Listeners of spoken languages often judge accents and dialects and by extension the speakers of these accents and dialects, and comprehension of the speech signal can suffer as a result of these judgments (Gluszek and Dovidio 215). These socially ingrained language attitudes are close in concept to Lippi-Green's "standard language ideology" (27). Decades of research on non-native accent perception have demonstrated that listeners carry specific language attitudes towards non-native speech (Moyer 114). Simultaneously, several factors impact perception of non-native speech, all of which carry the ability to affect audience perception of intelligibility of the performer on stage.

This chapter offers a succinct voice practitioner-friendly summary of the relevant literature in cognitive linguistics about non-native accent perception and other linguistic phenomena that are important to the practice of voice and dialect. Most importantly, accented speech perception is a composite of different factors that listeners weigh when hearing speech. To complicate matters, linguists use the term intelligibility in a very

narrow sense; they operationalize intelligibility as a numerical measure of the amount of information that the listener receives and subsequently can reproduce (Munro and Derwing 287). Often, intelligibility is measured using the proportion of correct words recalled by listeners in various listening environments. Significant prior research has examined the constellation of related factors that affect accented speech perception that are closely related to intelligibility: particularly comprehensibility, and accentedness (Derwing and Munro 1, Flege, Munro and MacKay 3129). Accentedness is a subjective measure that researchers define as how strong or heavy a listener believes an accent to be. Comprehensibility is also a subjective measure, which asks the listener how easily they can understand the speaker. Many social contexts often determine these scores, including contexts intrinsic to the speaker, intrinsic to the listener, or related to the environment in which the language is perceived (Moyer 144). Performance, in this case, can be construed as a particularly specialized social context in which a listener is encountering the speaker.. The limitation to this research is that I must simplify this interaction to how listeners are reacting to the voice of the performer before introducing visuals to the working model. Clarifying the role of the actor's voice can still lead to clues to how an audience member might behave; even with the impoverished signal of voice only, the listener can do a lot of work to fill in social expectation and qualities for the speaker.

This chapter demonstrates that the very environment of theatre or performance might contribute to skewing that seeming-objective measure of intelligibility, or other ways to judge communication onstage. Intelligibility, like other terms used in the theatre like 'authenticity' or 'clarity' is in fact a privileged form of judgment that listeners and practitioners use as shorthand, affected by the individual experiences of listeners. In other

words, these terms or qualities are continuously constructed by their users, and subsequently these terms are affected by every instance of use. To believe these are objective measures is to lead audience and performers astray towards the belief that their perception of objective fact is the correct approach to this work. To examine this socially constructed idea of intelligibility, I explore factors over and above accentedness, comprehensibility and intelligibility that a listener may use to construct their perceptions of the voice and the actor they are encountering, including subjective factors that are often used in discussion of an actor's voice, including discussions of 'clarity,' 'effort,' and 'authenticity.'

The final section of this chapter considers how my research might respond to voice and dialect practitioners in their calls for more interdisciplinary investigation into these issues. I discuss the results with an explicit eye to how these results could be interpreted considering assumptions made by the voice and dialect profession, which will carry into the final chapter where I consider my position as a white cisgender practitioner myself. In the discussion, I seek to establish questions of how standard language ideology may be navigated in this craft, which will be considered in my conclusion. The discussion of this thread will lead directly into the final chapter, which will summarize the contemporary issues and challenges the voice and dialect industry faces and will provide best practices and considerations of those who seek to incorporate dialects into their productions.

1.2 Language perception: General mechanisms, several models

Before addressing the social aspect of speech perception, a deeper understanding of speech perception is necessary to describe some of the more general speech perception

theories and models that influence linguistic research today. Within this section, I will highlight how these models might serve as points of access for voice trainers into the field of speech perception. First, in language perception more generally, several puzzles or issues of perception exist that speech models must address. These puzzles, like social context surround speech, may at first blush appear to have simple answers, but linguistics research will reveal that these puzzles are difficult principles to unravel. Models must address the most common issues that includes linearity (tracking the order in which the speech signal is received), segmentation (being able to perceive discrete meaningful units of language), speaker normalization (accounting for speech differences in different speakers), and the basic unit of speech perception (Ferrand 394).

Often, speech practitioners begin instructing voice students with the most popular proposed basic unit of speech, the phoneme. A phoneme is a unit of sound that is perceptually distinct and can help distinguish one word from another and are often taught by voice practitioners to their students through use of the International Phonetic Alphabet or their own proprietary writing system (Blumenfeld 12). For example, /p/ and /b/²⁶ are phonemes because in English, these sounds distinguish between the words “pad” and “bad” (Catford 184). Phonemes are an important key to understanding basic speech perception theories, as the other principles of speech perception are built from these units. For instance, the linearity principle and segmentation principles use phonemes to refer to the idea that a specific sound in different words corresponds to the same specific phoneme (Ferrand 393). As a practical example, this principle might assert that the /k/

²⁶Linguistic notation often places phonemes in forward slashes, which is a convention I adopt for this dissertation.

sound in “cat” is the same sound at the end of “back.” The segmentation principle asserts that the speech signal can be divided into discrete units that correspond to specific phonemes. Therefore, according to these two principles, the /k/ sound in “cat” and “back” not only should be the same sound, but they should be easily discernible in the speech signal, and both sounds ought to be identified as the phoneme /k/. Overwhelming evidence, however, has established that this is not the case. The exact acoustic sound characteristics of the /k/ in “cat” and “back” vary because of differing characteristics of the contexts in which this sound is produced. Voice practitioners can often explore these differences with their students; I often have my students explore the physical difference in the back of the mouth placement for the initial /k/ sound in “cup” versus “key.” The placement for “key” is closer to the front of the mouth than with “cup,” because of the mouth placement of the following vowel. This phenomenon is referred to as coarticulation, where muscular preparation for one sound affects the production of the immediately surrounding sounds. The mystery remains that there is no clear-cut one-to-one mapping of the acoustic signal to discrete sounds; we perceive speech as a series of separate and distinct phonemes and words even as the acoustic boundaries between phonemes are blurred and highly variable within one speaker, much less between speakers of the same language (Ferrand 394).

The issue of speech perception compounds when listeners perceive speech not from one speaker but from many different speakers throughout the day. Theories of speech perception try to account for differences in speakers. Different factors (e.g., age, gender, language background) lead to wide variations in speech, that includes pitch, loudness, stress, and rate of speech. However, theories posit that listeners can account for

these differences by somehow attuning to the ranges that speakers produce in their acoustic signals. Somehow, listeners can ignore irrelevant differences between productions of a given sound, while focusing on the acoustic features that indicate differences between meaningful units of speech (Ferrand 394). These units of speech also produce a linguistic quandary, do language users store and perceive speech as acoustic-phonetic features, abstract sound categories like phonemes, or larger units like syllables or small word units? This question might also have a different answer depending upon the age of the perceiver, children may process auditory information using larger units and later shift into adult-like behavior where they may depend upon smaller units like phonemes (Nittrouer 280). Units might also be sensitive to environment; a person may be able to attune to smaller units of sounds in a quiet situation than in a noisy situation, where they might rely upon context to predict the linguistic sounds they are hearing. Clearly, these four issues—linearity, segmentation, speaker normalization, and units of perception—present unique challenges to creating models of speech perception that account for the wide variety signals that a listener hears and of which they must make sense. The leading theories of speech perception grapple with these challenges that establish a basis for understanding the social implications of speech perception and ultimately how a listener might construct the idea of intelligibility to aid them in their perceptual journey.

One influential speech perception model is the Motor Theory of perception. This theory stresses the link between knowledge of production of speech and perception (Liberman and Mattingly 2). At its most basic form, the theory posits that a listener can perceive speech because they produce speech. Listeners are aware on some level of the

relationship between the theoretical sounds in their speech, and the articulatory gestures they produce to get there. Listeners are taught to perceive in terms of different types of mouth gestures but does not track the actual movements, instead they are tracking abstract articulatory plans that results is a perfect production of the utterance (Hawkins 127). Acoustic Invariance Theory assumes a similar abstract articulatory plan for each sound found in a specific language. This theory focuses on core acoustic properties, however, and can be conceptualized as a template against which the listener compares the incoming sound (Stevens and Blumstein 1358). The listener is still working with abstract representations of speech. In both theories, listeners are abstracting essential features of the incoming acoustic signal and subsequently make a decision about its identity by checking against a theoretical list of features.

Developed in the 1980s, Direct Realism²⁷ pushes back against notions of specialized abstract representation of speech sounds (Ferrand 399). This theory posits that direct knowledge of speech perception does not only stem from the acoustic signal itself, but also from prior experience of the listener perceiving the speech signal. Integrating direct experience more explicitly into further explanation of how prior experience shapes the speech signal, the TRACE model reflects a connectionist approach that integrates parallel processing across multiple sources of information in speech perception (McClelland and Elman 41). In other words, listeners are processing sounds across different levels simultaneously, including phonetic features, phonemes, words, and vitally social contexts of the speech. Units of perception can be as small as phonemes, or as large as words (e.g., a logogen or another type of unit associated with words in a

²⁷ Not related to the theatre movement of Realism.

listener's vocabulary). Each experience of a unit is tagged with useful information, such as perceived qualities of the speaker, to help the listener recall these units more efficiently in the future. These models reflect processes that researchers were using in computing, and these models became more sophisticated as knowledge about computers advanced.

Native Language Magnet Theory has been an influential model over the beginning decades of the twenty-first century and begins to take seriously the notion that language perception is a weighted collection of individual experiences for the listener (Frieda et al. 130). The critical element of this theory is that phonetic categories are organized in terms of prototypes (like theories that came before), but these prototypes function as perceptual magnets that assimilate variations in production towards these categories. Distinctions within the same category that are close to the prototype are reduced (e.g., the /k/ in both "key" and "cup"), but perceptual distinctions between category boundaries become even more distinct, so the boundaries between sound categories are clearly divided (Kuhl et al. 684). Thus, listeners in this model can account for a range of differences while maintaining a sound system that supports specific language perception and production. These general perceptual theories lay the groundwork for language perception in an even more variable environment, which is language produced by a speaker who is speaking in their second, or even third language.

A model that is specifically concerned with perception of non-native contrasts in language is the Perceptual Assimilation Model (Best and McRoberts 193). This model can be used by voice practitioners to help predict the relative ease or difficulty an actor may have in acquiring a new performed dialect. Importantly, the Perceptual Assimilation

Model can help account for difficulties in perceiving and acquiring a target performed dialect that is close to the actor's own idiolect. This model incorporates perceptual attunement to the physical consequences of articulatory gestures that signal contrasts between speech sounds, incorporating speech gesture into a statistical experiential model. Degrees of distinctions between phonological categories are what perceivers attend to; listeners can also learn to tune out phonetic sequences that do not signal a change of meaning. The listener can attune to a hierarchy of phonological differences and assigns weight to these degrees which are incorporated into the model (i.e., sound differences within category receive little weight, while differences that signal meaning change have higher weights). The model PAM predicts that discrimination performance on non-native contrasts will vary from poor to excellent depending on how the contrasting non-native phones are assimilated (according to the weights assigned) to native phonological categories (Farris, Best and Tyler EL1). If non-native sounds are not fully incorporated into a listener's collective experience of sounds, they are *categorized* as examples of phonological sounds with ratings from excellent (e.g., sounds native-like) to poor. If the features of the non-native sound are not consistent with any one native sound category, then it is *uncategorized*, and if it is not heard as speech (such as a lateral click from the Xhosa language), then it is non-assimilable into the listener's sound inventory. PAM accounts for the counter-intuitive notion that sounds close to an actor's own phonology may be more difficult to acquire precisely because their own native categories overlap with the sounds they are trying to acquire as part of an accent due to their inability to perceive these sounds. These categories can predict how well a listener can understand a non-native speaker and can help explain why some sounds are more salient than others.

The next section takes the idea of *categorization* and expands the idea that listeners are tracking more than just the sounds they hear, but the contexts in which they find themselves as well.

1.3 Top-down processes help organize the speech signal

The work of a voice practitioner is a continual balancing act between expert judgment and fine phonetic work with their actors and understanding empirical findings of this field about how these judgments work will strengthen their work. Theories of cognitive processes in the previous section describe the automatic and subconscious ways that listeners use to perceive speech, and introduces just how variable the acoustic signal can be, which presents the first challenge to voice practitioners. In this section, I describe the processes that listeners use to assign meaning to that acoustic signal, focusing specifically on those processes that live closer to the surface of the listener's consciousness and experience of their world, which adds another layer to the complicated story of how listening and perception works in performance. Contexts in which people encounter speech affects the way that that speech is processed, as demonstrated by countless experiments that measure listeners' language attitudes. Social context even affects the very notions of "dialect" and "accent," which turns out to be a less stable notion than voice and dialect practitioners would like to admit. It appears that all these conscious and socially constructed processes come at a cognitive cost; listeners who judge speakers as less easily understood in turn are less likely to absorb information or content in the speech signal, creating a feedback loop where perception creates the results of that perception (Rubin 522). The following section explores the causes and consequences of these social top-down expectations and how they may relate to the work

the precarious position voice and dialect practitioners hold by navigating these social contexts.

1.3.1 What is a dialect anyway? Social construction of an accent

To answer the question “what is a dialect?” requires consideration of how power intersects with the lived linguistic lives of everyday people. This power manifests itself by way of individuals' standard language ideology, but these individual ideologies are shaped by authorities and the broader society in which they live imposing an indexical order on speakers who are accessing macro-sociological categories as individual values (Silverstein 193). Answering this question points to who society at large believes who has an “accent” or a “dialect” or non-standard variation of a standard language. These speakers are already marked by the community as deviant from the norm or accepted language usage. Pragmatic knowledge perceived as individual value judgments attached to linguistic forms varies depending on the availability, accuracy, detail and control that speakers in a community have (Preston 188). Michael Silverstein (1981) refers to these dimensions in his work on indexicality of the linguistic forms in the minds of speakers and listeners. The term indexicality is the notion that “attitudes towards and folk beliefs about languages are not isolated instances, but reflect patterned and structured ideologies within cultures and speech communities” (Silverstein, qtd in Preston p.182). Voice professionals must account for their own and their students’ systematic beliefs, the beliefs speakers they are trying to emulate, and the audience who will hear these speakers in specific instances, which creates a rich ecosystem of ideology to unpack as part of the rehearsal process. This accounting begins with the overarching idea of who even has an accent in the first place, as Preston (2016) notes:

When the folk say that someone has an accent, there are at least two important differences. First, for linguists, if the word “accent” is a technical term at all, it refers exclusively to the phonetic/phonological level. Folk respondents very often refer to the entire linguistic system with this word. Second, and more importantly, linguists know that everyone speaks some regional variety, even those heavily invested in removing such matters from their speech. Folk comment abounds, however, with the idea that somewhere there is “accent-free” speech; in the United States, for example, many respondents identify the Upper Midwest as “accent-free,” perhaps particularly those from the area itself. (182)

Given these attitudes, explicit standardization—the codification of pronunciation, grammar, lexicon, or spelling for a given language variety—often interacts with explicit political structures (Moyer 85). Sociological indexicality affects power structures on the personal and institutional level. Often the type of political power involved with standardization includes the right to determine what may count as a “language” and what may count as a “dialect.”

As illustration of the interrelating power structures of language ideologies, I highlight the particular issue of labeling dialects within English, which could be an aid for voice practitioners in approaching indexicality of their chosen accent or dialect. Even within regional and international dialects of English, power structures dictate “inner” and “outer” circle dialects to legitimize certain speakers of English over others. The use of inner and outer delineates a type of privilege that has been awarded to historically white English speakers, while those on the “outside” have been relegated to a second or lower

class of acceptable English usage. The Inner circle consists of the United Kingdom²⁸, Ireland, US, Canada, Australia, and New Zealand, and represents about 3–5% of English use in the world. The remaining 97–95% of usage, English is a second language or heavily influenced by local languages and used in official governmental context (Moyer 91). Using this delineation disrupts the colloquially accepted idea of *nativeness*, since some English dialects have indeed undergone language change independent of the trajectory of English in the inner circle (Mollin 170). Dialect coach Jerry Blunt implicitly uses this inner/outer orientation towards his collection of audio dialects, where he sought non-native dialects, but not in the context of the speakers living in the locations where each language variety is found. Instead, he sought speakers from “outer” countries living in “inner” locations like the UK or the United States to collect audio samples. Given these complications, in these experiments, I will refer to non-native dialects and mean that these are English dialects where the speaker has learned English as their second language in a location outside of the United States.

Dialects standardized explicitly by prestigious institutions such as formal education or voice professionals are direct reflections of ideas about what accent is right or standard in any given culture. How these standardized accents are treated and reified can reveal elements of indexicality of the traits that are important to key authorities who control access to these artificial dialects. Though different dialects do not innately indicate characteristics of the speaker, the perception remains that dialect or accent can indeed reveal qualities about the character of the speaker and their status in wider society.

²⁸ Within the United Kingdom some varieties of English are not as widely accepted as others, depending upon markers of class and geographical location. Perceptions of outsider versus insider status also vary with identification of the label “United Kingdom.”

These associations are often instilled in language users at a young age by various forms of entertainment; one study examining different Disney characters and accents found that often villains were portrayed with regional or non-native accents²⁹, while heroes often spoke in some form of unmarked or General American accent (Lippi-Green 92). The same phonemic trait can appear as both a marker of high and working class, and the dialect coach's job is to aid the actor or student in disentangling phonetic realities from the social expectation of character traits. Another example examines the use of rhoticity, defined as the appearance or disappearance of the phoneme /r/ in certain accents and dialects, especially in syllable-final position. The mere presence or absence of rhoticity in a dialect does not inherently point towards character traits good or bad. Dialects of different prestige and mainstream acceptability employ r-lessness in various degrees. These accents are often seen as indicative of higher-class speakers specifically because of the perception of Received Pronunciation and other higher-class British accents as accents that were taught in schools and to performers and used in middle class pursuits such as theatre. However, this type of r-lessness also applies to working class dialects found in some neighborhood of Boston, such as the Southie neighborhood, as popularized by figures who have actively cultivated a working-class image or persona such as Whitey Bulger, and Mark Wahlberg (Ulin).

The decision to use accents is influenced by stereotyped representations, where representations of accent and dialect in media are used as cognitive shortcuts for characterization of the characters (Bakanic 14). In the Disney study, foreign-accented

²⁹ Since this 1997 study, the pattern has continued. In *Lion King* (2019), Scar (Jeremy Irons and Chiwetel Ejiofor) and Zazu (Rowan Atkinson and John Oliver) are the only British-sounding characters in *Lion King*, with the former the villain and the latter the sort of busybody killjoy.

characters in particular were more often seen as poor, uneducated, and as the bad guys, enforcing stereotypes that foreign-sounding speakers of English are not to be trusted (Lippi-Green 93). Further evidence points towards performed speech as a reflection of the standard or idealized speech of the dominant time in which the media is produced (Elliott 120). Elliott investigates the predominance of certain speech varieties by tracking the change in rhoticity, or the use of /r/ at the end of syllables, throughout a century's worth of movies and correlating that decline with the general decline in rhoticity in English in the United States over the same period of time. Though this claim could be attributed to language change in general and the fact that actors are also part of a language community, a tantalizing theory exists that the speech of actors (especially explicitly trained in dialect or voice for the stage and film) can and does represent an ideal or standard style of speech, especially as it relates to Skinner and her Transatlantic speech. Listeners may use expectation of idealized social norms for accents to assign character traits to accents, and may be using this same ideology to judge authenticity and other social factors to the accents they are hearing.

Using a performed accent (especially with actors who are trained to acquire a new accent in a short period of time) in this dissertation may offer access into the processes by which expectation of a standard accent affects perception. In this case, a listener may perceive a non-native accent through their conception of authenticity or intelligibility. German-speaking listeners were able to identify the origin of different imitated non-native accents (e.g., French, American, Italian) better than authentic non-native accents (Neuhauser & Simpson 1805). However, they were less accurate at judging the authenticity of the presented accents. That is, listeners' expectation of authenticity does

not translate to ability to judge authenticity of the accent. This may be because listeners are identifying stereotypical traits in the imitated speech they are hearing as evidence of authenticity, while these stereotypical cues are missing from the authentic accents that they are hearing. The research in this chapter will further examine the effects of expectation on imitated and natural accents by examining other social factors that may be susceptible to these types of effects. Using expectations in this way will reveal the various layers of indexicality that listeners place upon their conception of dialect the role dialects play in creating meaning for performance.

1.3.2 Social construction of an accent affects perception

Standard language expectations have also been demonstrated to have consequences in educational environments, demonstrating that comprehension is affected by listeners' attitudes toward a speaker. In a foundational study on perceived accentedness in 1990, Donald Rubin and Kim Smith investigated lecturer ethnicity and lecture topic as factors in undergraduates' attitudes towards International Teaching Assistants (337). They measured comprehensibility ratings after playing 4-minute lessons either in a 'moderate' or 'strong' accent for 92 undergraduate students while projecting one of two lecturers, indicated by photograph of a white or an East Asian instructor. Degree of accent correlated negatively to perceptions of teaching competence. In 1992, Rubin followed-up this study by demonstrating that college students' language perception and comprehension can be influenced by perceived race. Even when using a standard American accent as the audio signal, students who saw a picture of an East Asian woman while listening to a lecture performed more poorly on the content exams in both the science and humanities post-tests (Rubin 516). In later research, Rubin named this

phenomenon “reverse linguistic stereotyping,” demonstrating that listeners’ perceptions are sensitive even to the suggestion of racial context (Kang and Rubin 441). Kevin McGowan explored the reverse of this effect in 2015, demonstrating that foreign accented English paired with a picture of a person of a different perceived race resulted in similar detrimental effects on the listener (e.g., Chinese-accented English paired with a picture of a white woman). That is, listener expectation runs both ways; if the listener hears foreign-accented speech, they expect the image of the person to match the signal to which they are listening. In other words, listeners carry standard language expectations for more than their own language, and are poised to carry standard expectations for most of the accents they encounter in their lives. Voice practitioners and dialect coaches especially may use of standard expectations or stereotypes to affect how audiences perceive the speakers of these accents. For example, if actors are using dialect in a surprising non-stereotypical way, dialect coaches can account for the adjustment that audiences must make when they encounter these accents on stage for the first time.

To test the audience’s perceptions of stereotypical accents in performance context, I will use foreign or non-native dialect as the tool for inquiry into this idea of intelligibility as a specific example of a context in which audiences are creating meaning using the voice in performance. For the purposes of these experiments, I needed a target dialect or accent from which to work that could be controlled in the lab setting for my experiments. Even with controls in place, asking a voice trainer to help an actor sound their best or most intelligible runs the risk of introducing many different variables. Beginning with a target dialect of Russian-accented English to be trained for an American English-speaking actor at least creates a target than can offer insight into

listeners' ideas of stereotype. I acknowledge that reducing the lived linguistic experiences of speakers is quite near the opposite of what I have been arguing in the dissertation up until this chapter. However, I do need an entry point into the world of context for intelligibility, and I need an entry point that will elicit reactions about that particular accent from listeners. Often, listeners are more likely to give explicit judgments or ratings when they are listening to non-native accents (Wester and Mayo). Starting so specifically with an accent like this means that the patterns and phenomena captured in the subsequent experiments in this chapter may not be generalizable to every moment (as I have argued so far in this dissertation), however, they may serve as a baseline for further inquiry into perception of intelligibility. Scientific knowledge, after all, starts with as many variables controlled as possible and introduces more variables as the model becomes more complex. The next section contains the description of factors I will use the results of two experiments designed to tease apart the social measures behind intelligibility.

1.3.3 Measures of factors affecting accented speech perception

In order to test audience expectations, I must be able to measure some kind of experience the listeners are having. In order to do this, I will be using language attitudes that listeners can use to label their perceptual experience. I am interested in attitudes intrinsic to the listener that affect the factors of accentedness, comprehensibility, and intelligibility, like in previous research, and the context in which the listeners perceive speech (Munro and Derwing 285). For performance-specific questions related to voice and dialect, I turn to other listener-intrinsic qualities assigned to accented speech that can be measured on scales like those used to judge accentedness and comprehensibility. For

example, when judging qualities of speech, listeners used adjectives such as “appealing”, “clear,” “pleasant,” “intelligent,” and “sophisticated” in different amounts on a five-point scale while listening to different regional accents in North America (LaMonica). These different qualities can reveal specific attitudes about different accents and dialects. In LaMonica’s study, listeners rated Southern dialects as more appealing, yet not as sophisticated as accents found in the Midwestern United States, demonstrating that, while these scales are often aligned, there is some independence in descriptions of accents. The factors affecting perception of accented speech are not fixed within the speaker or even the listener, as these factors can be influenced by the context in which the speech is being perceived, including expectations of the listener (Kang & Rubin 450). In the present study, we ask specifically about the effect of these contexts on the factors of accentedness, comprehensibility and intelligibility, along with other qualities with which listeners may be associating in particular with non-native accents both within and outside of the context of performance.

To test the influence of expectation of standard accents on social factors in perceiving accented speech and ultimately intelligibility, I employ a modified matched guise experimental design (Lambert et al. 44), using different instructions to different groups of listeners to make them believe they are in different scenarios. Experiments utilizing matched guise attempt to hold as many variables as possible constant and employ a single speaker speaking the same material over a series of accents or over a series of contexts (Giles & Coupland 34). For example, in one experiment employing matched guise, participants listened to the same stimulus while being assigned to one of two different listening contexts; being told whether the speaker is a native speaker of

Cantonese or an American speaker (Hu and Lindemann 254). Listeners who were told they were listening to a Cantonese speaker gave higher accentedness ratings than those who believed they were listening to an American speaker.

I employ a matched guise paradigm in a similar way by introducing different listening contexts to different groups of listeners by using stimuli from a trained actor imitating a dialect and from natural speakers of that dialect. While this series of experiments captures a mere fraction of the rich environment an audience member would encounter while experiencing theatre, I first start with the voice or audio as a way to directly compare this work to work that I have just reviewed.³⁰ I want to test if the mere suggestion of expecting performed or imitated speech affected audience perception. If perception is affected with the mere suggestion, I would expect future work to demonstrate that the entire audience experience affects speech perception in profound ways that are yet to be documented. To support these experiments, I conducted a preliminary study using this matched guise paradigm, to determine whether listeners are sensitive to the differences between an actor imitating an accent and natural speakers of the target accent, and whether patterns of description exist for listeners. Data from this pilot have helped me identify social factors that may be associated with a socially idealized or maximally intelligible accent. Below, I review this study, using conventions from linguistic inquiry. Implications of these data for the field of cognitive humanities follow the description of the experiment.

³⁰ Another issue I must address is that all of the experiments conducted for this dissertation was during the COVID-19 pandemic and I only had access to my participants via the internet, so I decided the best way to control for differences in technology and access was to focus on the audio portion of experiencing voice onstage. Adding video or a picture of these speakers will be an excellent future direction for further study of this effect.

2. Preliminary study

2.1 Participants

I used Mechanical Turk³¹ to gather results from 108 participants in an experiment that took an average of 2.1 minutes to complete. These participants were in the United States, at least 18 years of age, and had indicated that they had completed some high school. Importantly, these participants were from locations outside of the University of Oregon community, so I was able to gather a larger variety of listeners.

2.2 Stimuli

Four recordings of one randomly selected sentence from the Hearing in Noise Task (HINT) sentences (Nilsson, Soli and Sullivan) from the Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings (ALLSSTAR) corpus (Bradlow, Kim and Blasingame) were used as stimuli for the experiment. Three recordings of three Russian-accented English speakers were selected from the corpus. The fourth speaker was a university student who was trained in a Russian accent through a Voice and Dialect theatre class offered at the University of Oregon. This actor was subsequently privately coached specifically on all the sentences for 4 hours (2 two-hour sessions, one for the first 60 sentences and another for second set of 60 sentences) and then recorded the sentences. Acceptability was determined by the dialect coach³²; recording would continue until the dialect coach was confident each sentence was successfully produced in the target Russian accent. The student was coached to read the sentences, and to not act out

³¹ Mechanical Turk is an online service provided by Amazon that employs HITs, or Human Intelligence Tasks that rewards workers a small amount of money when they complete a HIT. In this case, I paid for this HIT at an equivalent rate of \$15/hour.

³² Much appreciation to Dr. Tricia Rodley for her contribution as dialect coach, and to Christian Mitchell as the actor in this experiment

the sentences, as they might normally do in a theatre class. Russian-accented English was chosen as the target accent partially because of the availability of this accent through the theatre class, but also because this Russian-accented English has been described by the research of Stephanie Lindemann as “correct but not pleasant,” and occupies an intersection of intelligibility and accentedness, where the speakers are perceived as heavily accented yet still intelligible (204).

2.3 Listening groups

Participants were divided into four listening groups in a 2X2 design that examines speaker type (trained vs. untrained) and expectation (no expectation vs. expectation), as described in Table 1 below. This design allows for making multiple types of comparisons of the data by comparing either rows or columns to one another, or comparing all four groups to each other. By separating by expectation (columns), we can examine the effect the listener has on intelligibility and other factors, while separating by training (rows), we can examine the specific effect of speaker training on intelligibility and other factors. A listener heard either the two real Russian-Accented English speakers (untrained group), or a mixture of one real Russian-accented English speaker and an actor (trained group). No participant heard a different combination of the real Russian-English Accented speakers, nor did they hear any other speaker compared to the actor. These two groups of speakers were crossed with two listening conditions. In one condition, the listeners were explicitly told that there is an actor in the group of two speakers (expectation). In the other, listeners were not explicitly told there is an actor in the group (no expectation). Listeners were randomly assigned to one of four listening groups.

Table 1. Four different listening groups in the experiment.

	Expectation	No Expectation
Trained	Group 1	Group 2
Un-trained	Group 3	Group 4

2.4 Procedure

Participants were instructed to complete the experiment using headphones. In the expectation condition, participants were first informed that they would hear two voices and that one of those two voices were an actor who was been trained to perform an accent. In the no expectation condition, they were informed that they would hear two speakers, but were given no other information about those speakers. They were then presented with the audio of the sentence spoken by each of two speakers and could listen to each audio clip as many times as they wished. Regardless of condition, participants advanced to the next screen where they were asked to select which audio clip contained who they thought was an actor in a two alternative forced choice task. Note that only half the participants were told in advance that there would be an actor in producing speech and in only half the conditions was an actor actually included in the sound files. After their selection, an attention question was asked, “how did you listen to the audio samples today?” Finally, participants were asked to explain their choice of actor using a free response text box.

2.5 Results (two alternative forced choice task)

For the trained condition in both expectation and no expectation, participants selected the actor more often than the other speaker. In the untrained condition in both

expectation and no expectation, participants selected speaker 142 more often over speaker 140. Exact percentages are found in Table 2.

Table 2. Results of two alternative forced choice task

	Expectation		No Expectation	
Trained	<i>Actor</i>	<i>Speaker 144</i>	<i>Actor</i>	<i>Speaker 144</i>
	59%	41%	65%	35%
Untrained	<i>Speaker 142</i>	<i>Speaker 140</i>	<i>Speaker 142</i>	<i>Speaker 140</i>
	53%	47%	65%	35%

Independent t-tests show that all these percentages are not significantly different from chance (all p-values >.05), probably due in part to the small number of participants in each square. T-tests also show that proportions within condition (expectation versus no expectation) and within listening groups (trained versus untrained) are not significantly different from each other, while demonstrating a trend in the direction towards the first option that they were given in the experiment. Non-significance in this case means there is no detectable difference between conditions. That is, we do not have evidence that listeners are sensitive to the presence of an imitated accent. However,, like in all scientific experiments, the null result should be interpreted cautiously, as a null result neither proves nor disproves hypotheses that are established as part of the experiment. The trend in the task demonstrates enough possibilities that I adopted this procedure for the main experiments below.

2.6 Results (Free response question)

Participants were able to type free responses to the question “Why did you pick that speaker as the actor?” Responses were coded with a type-token count by tagging each response with a keyword (sometimes multiple keywords). A type-token count refers to how many (tokens) of each type of keyword was found in the responses. Responses containing the word “unnatural” accounts for over 20 individual responses of 108 participants in both conditions and listening groups. The term “unnatural” was followed in order by “clear,” “forced,” “fake,” “exaggerated,” “recognizable,” “natural,” and “not authentic.” These descriptors appear to have positive and negative connotations in their use, with negative terms “unnatural” and “not authentic” being the most transparent negations to their counterparts “natural” and “authentic”.

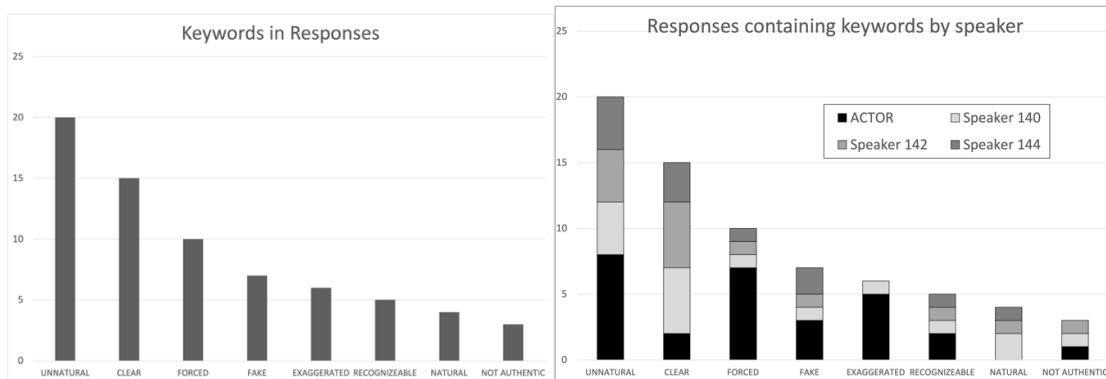


Figure 1. Histogram of responses by keyword, and again with keywords by speaker.

Dividing each key term by the choice that was made by the listener reveals that different terms used by listeners pattern differently in each instance of these keyword. For example, 70% of the times the participants used “forced” to describe their choice of actor, they correctly chose the actor, compared to only 13% of the times they used “clear” to describe their choice when they correctly chose the actor. In particular, the use of the descriptor “clear” seems to pattern with selection of spontaneous accents, since most

listeners who chose the descriptor of “clear” also chose one of the three speakers of the real speech samples of Russian-Accented English as their selection for the actor. These selections show that listeners use different factors when selecting for an actor when they hear an imitated accent—the terms where the highest proportion of listeners select the actor include “unnatural”, “forced”, and “exaggerated”. Factors that listeners use in selection of natural speakers as the “actor” include, “clear”, “fake”, and “natural”. Different proportions in keywords used to describe their choices points to at least some kind of sensitivity to the difference between imitated and natural accents.

When examining the keywords that are used by expectation group (comparing the columns of the experiment design), another pattern arises in the responses. One key word “recognizable” is only used when the listener is explicitly expecting to hear an actor. With explicit expectation, listeners were sensitive to this voice as imitation with one participant describing the actor’s voice as imitating a famous actor. Another key word “natural” is nearly used exclusively while explicitly expecting to hear an actor. The fact that keywords appear in different proportions in expectation and no expectation listening conditions points to a possible difference in factors that listeners are using to judge imitated versus natural accents, regardless of their ability to accurately detect imitated accents versus natural accents.

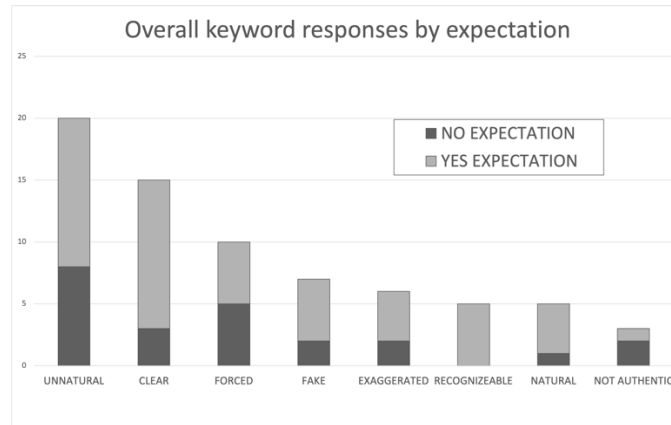


Figure 2. Histogram of responses by keyword, by expectation condition

2.7 Interim Discussion

Data from the initial pilot study reveals the exciting possibility of patterns in how listeners conceptualize the voices they are hearing when there is an explicit expectation of intelligible speech (e.g., performed speech from an actor). Because of the trends and promising results, these findings lead to the experiments conducted specifically for this dissertation. I will employ specifically the keywords from the free response that were coded and analyzed. Above these terms, these experiments ask how subjective measures such as accentedness, comprehensibility, and other measures of language are viewed when the listener expects maximum intelligibility. Further, these experiments dissect just what listeners mean when they are listening for intelligible speech. Experiment One examines what kind of social qualities listeners are assigning to each of the four speakers (the actor and the three Russian-English accented speakers) before the notion of expectation is introduced at the end of the experiment. From the findings of the pilot, the possibility that listeners are perceiving idealized forms of language as intelligible speech when asked to listen for an actor is not immediately clear, so I explicitly ask listeners to select who they believe to be the stereotypical accent, which helps to approach the idea of intelligibility through a more top-down processes and conscious activation of social

standard dialect expectations.

To answer the other research question, which of the social factors are listeners using to subjectively determine if they are hearing idealized or standard forms of language, the free response keywords will be used. In Experiment Two, listeners employ descriptions from the keywords of the pilot on Likert scales (scales of 1 to 9, a standard practice in surveys and social experimentation) to see how expectation of performance affects these descriptions. Using these descriptors, I can approach the social construction of intelligibility and explore what types of qualities listeners use while perceiving speakers they expect are performing for them. Determining which of these special qualities that listeners are using can clue dialect and voice coaches towards new goals in vocal (and more specifically dialect) production. What if, while voice professionals were using intelligibility as a benchmark, we could use a different quality as a goal instead? If audiences are indeed sensitive to social context and expectation, why achieve “authenticity,” when authenticity is constructed out of expectation of stereotype and not experience of reality for these accents? These experiments aim to answer these questions in service of expanding the notion of who has an acceptable performing voice.

Copious research evidence exists that points toward social expectation shaping how speakers are perceived by listeners. The hypothesis is that this special case of expectation—that performance is a social context that triggers a change in listeners’ perceptions of language—is not any different. Therefore, if listeners are using performance as a special social context when we test different groups of listeners (those in the Expectation condition and those in the No Expectation condition), we should see changes between the conditions. We should see that expectation affects how listeners

score their listeners on dimensions such as accentedness, comprehensibility, and the adjectives that were found in the pilot experiment. This would demonstrate, in part, that objective-seeming adjectives are constructed through the social context that listeners have in the scenario. If there is no change between the Expectation and No Expectation groups, then performed speech is not a factor that listeners account for in their perception of speech.

3. Experiment 1

3.1 Method: Participants, stimuli, procedure

I recruited forty-five (45) participants whose first language is English and had no prior experience with Russian from the human subjects pool at the University of Oregon. In addition to collecting language background information from each participant, I also asked about their experience in performing and attending live performance (acting, improv, role playing, and any other type of performance). Stimuli are the same speakers as in the pilot study. Three Russian-accented English speakers from the ALLSSTAR Corpus and the actor are the same speakers as in the pilot study. Forty-two HINT sentences were selected (see appendix). Each participant heard ten different sentences from each speaker. Two additional sentences were selected, where the participant heard all four speakers. In Qualtrics³³, participants were randomly assigned to one of two expectation conditions. In one condition (expectation), listeners were explicitly informed that one of the four speakers they are about to hear is an actor, and that they will be asked to choose who they believe the actor is. In the second condition (no expectation), participants were not informed that one of the four speakers they are about to hear is an

³³ Qualtrics is an online survey platform.

actor. Each participant heard each sentence selected for the experiment.

After each sentence, the participant transcribed the sentence. These transcriptions were scored for accuracy using AutoScore³⁴ (Borrie, Barrett and Yoho). After transcribing each sentence, participants then rated each sentence on a 9-point scale for both accentedness (i.e., “how accented is this sentence?”) and comprehensibility (i.e., “how easy is it to understand is this sentence?”) (Derwing and Munro). After participants responded to each of the forty sentences, they were presented with the audio of the *same* sentence spoken by each of the speakers and could listen to each audio clip as many times as they wish. They were asked to select from the four voices which person they believe is the actor in a four-alternative forced choice task. To test explicit language attitudes about stereotype and authenticity, a second sentence was played with all four speakers and the listener was asked, “which is the closest to a stereotypical Russian accent?” The results of these procedures follow, first examining how accentedness, comprehensibility and intelligibility vary by speaker (either the native Russian speakers or the actor), and then how these attributes vary by listener condition (expectation versus no expectation).

3.2 Results

3.2.1 Intelligibility (Accuracy of recall)

Since the work of this dissertation directly challenges the notion of intelligibility, I am electing here to rename intelligibility to what was being measured functionally from the experiment. Intelligibility, in this case, is the proportion of correct words recalled to the number of words in the sentence. These intelligibility scores are at ceiling or nearly

³⁴ AutoScore is a program that compares the ideal sentence to a response sentence given by a listener, and automatically counts the number of words correct in the sentence, automating the process of analysis.

100%, precisely because this experiment was designed to maximize this score to examine how accentedness and comprehensibility behave with maximum perceived intelligibility. What follows is a table that shows the proportions of words correct for each speaker. Because the measure is a proportion, the closer the number to one (1), the more accurate listeners were in their transcription of the sentences that they heard. Standard deviation is used to indicate the extent of deviation for the group as a whole. In other words, this measures how different each of the individual scores for each of the sentences are from one another. This means that a smaller number indicates scores that are all very similar, while a larger number shows a larger variation in scores.

Table 3. Accuracy of transcription of sentences for each speaker.

Speaker	Mean accuracy	Standard Deviation
<i>Actor</i>	.9574	.1621
<i>Speaker 140</i>	.9504	.1325
<i>Speaker 142</i>	.9338	.1791
<i>Speaker 144</i>	.9386	.1308

The Actor overall has a higher accuracy of recall ratings than the three other native Russian-English speakers. However, when compared for significance using t-tests, none of these accuracy ratings are significantly different from each other. The t-test between the Actor and the speaker with the lowest rating (Speaker 142) does not reveal significant differences between the two ($t(677)=1.7981$, $p=0.0726$). The following table reveals accuracy of recall by expectation condition, combining all four speakers in both categories. These data are the same responses that make up the above Table 3, but are divided in a different way that might help reveal the role of social expectation in accuracy

of recall. Listeners in the Expectation condition showed a higher accuracy than No Expectation in their transcription of the sentences. However, a t-test reveals that these results are not necessarily significantly different from each other ($t(1356)=1.2303$ $p>.05$). Overall, these results show a slight trend towards more accuracy for the actor, and more accuracy for the Expectation condition.

Table 4. Accuracy of transcription of sentences by expectation by listening condition

Condition	Mean Accuracy	Standard Deviation
<i>Expectation</i>	.9505	.1281
<i>No Expectation</i>	.9403	.1714

3.2.2 Accentedness and Comprehensibility

Listeners scored each sentence on Likert scales from 1 to 9 for both accentedness and comprehensibility. Listeners gave a score of 1 for “not at all accented” and 9 for “extremely accented.” For comprehensibility, listeners gave a score of 1 for “easiest to understand” and a score of 9 for “extremely difficult to understand.” In other words, the higher the score, the more accented and less comprehensible each sentence is judged to be. Results are presented in box plots, which show a summary of a set of data. Each box represents the first and third quartile³⁵ of the data set, while the horizontal line represents the median. The ends of the whisker—or the lines above and below the box—represent the minimum and maximum values in each set of data. In this case, the minimum and maximum are always 1 and 9 for any set of data. Below is the box and whisker plot that directly compares the scores for the Actor and the three Russian-Accented English

³⁵ A quartile is the median of the data below and above the median of the entire data set.

speakers. Each data point in this plot is the individual sentences that each speaker produced for this experiment. Speakers 140 and 142 are judged to have the similar accentedness ratings and received over all very similar scores. The mean accentedness rating is 5.7 (s.d.³⁶=1.91) for Speaker 140 and 5.6 (s.d.=2.0) for Speaker 142. The actor received a mean score of 5.1 (s.d.=1.87), meaning he was judged as less accented than speakers 140 and 142. Speaker 144 received a mean score of 4.6 (s.d. = 2.08), which means they were judged the least accented of the four speakers, while simultaneously demonstrating the widest variation in scores.

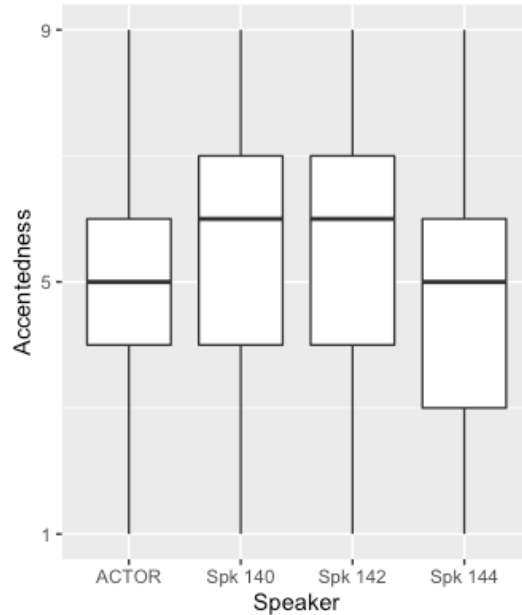


Figure 3. Box and whisker plot that shows the median accentedness scores for all four speakers.

Compared to accentedness ratings, each speaker has a lower overall mean comprehensibility rating. Again, for in Figure 3, the higher the number, the less comprehensible the speaker sounds. The actor, while demonstrating accentedness ratings that are similar to Speakers 140 and 142, appears more comprehensible than these

³⁶ s.d. stands for standard deviation.

speakers, patterning this time with Speaker 144. The mean comprehensibility rating for Speaker 140 is 3.94 (s.d. = 2.21) and for Speaker 142 is 4.1 (s.d.=2.27). The mean comprehensibility rating for the Actor is 3.1 (s.d.=2.09), while the mean comprehensibility rating for Speaker 144 is 3.0 (s.d.=2.21). While the literature posits that accentedness and comprehensibility ratings can be independent of each other, these results clearly demonstrate that these scores do not always correlate with one another. Listeners are indeed constructing different understandings of accentedness and comprehensibility for each of these speakers, and appear sensitive enough to the differences between each speaker to rate them differently from one another. The intelligibility scores for each of these speakers do not differ significantly, yet speakers still receive different scores for their accentedness and comprehensibility of their speech. These differences demonstrate that there are different ways to construct maximum intelligibility, regardless of expectation. The next section teases apart these results further by examining listening by expectation.

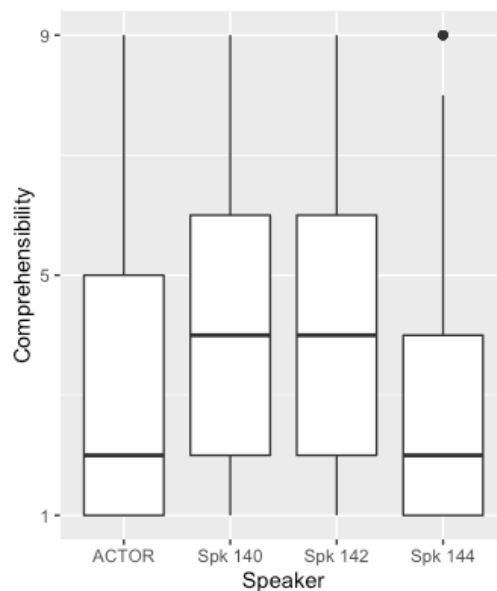


Figure 4. Box and whisker plot that shows the median comprehensibility scores for all four speakers.

3.2.3 Accentedness and comprehensibility by expectation

Results and these next two figures further divide the data of each speaker into two listening conditions: expectation versus no expectation. The same data from the first analysis of accentedness and comprehensibility appear again, this time further divided into expectation;. Recall the prediction that, if listeners are sensitive to social context, they would adjust their ratings in a different direction, thus demonstrating that intelligibility is further constructed out of different subjective judgments of voice and is highly sensitive to the context in which the listeners and speakers find themselves.

In the first box plot, the red boxes represent listeners who were explicitly told to listen for an actor in the experiment, while green represents listeners who were not informed they were listening to an actor until asked to determine who the actor is in the experiment. The following box plot demonstrates that accentedness is not necessarily sensitive to the social context of expectation. Because the boxes overlap and look similar between the two conditions for each speaker, it appears that there are little to no differences in this chart between the conditions, with the exception that for speaker 144, the median is different. As a comparison, the results of the comprehensibility question show different patterns between the Expectation and No Expectation conditions.

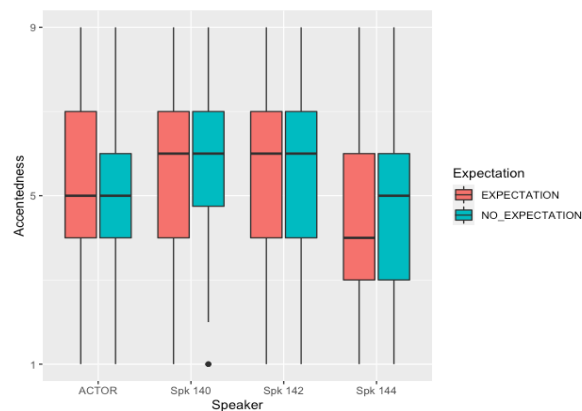


Figure 5. Box and whisker plot for accentedness with expectation conditions.

In this next plot, similar patterns for Speakers 140, 142 and 144 show no differences between the two expectation conditions. The Actor shows that in the No Expectation condition, listeners were more consistent in their scores concentrating around a score of 2 (where the median is, right at the bottom of the box). While in the Expectation condition, answers varied widely, with the first quartile at 1, and the third quartile ending at a score of 5. Changes in the variability of scores may indicate a change in behavior as a result of expecting a certain type of voice or performance context while listening to speakers.

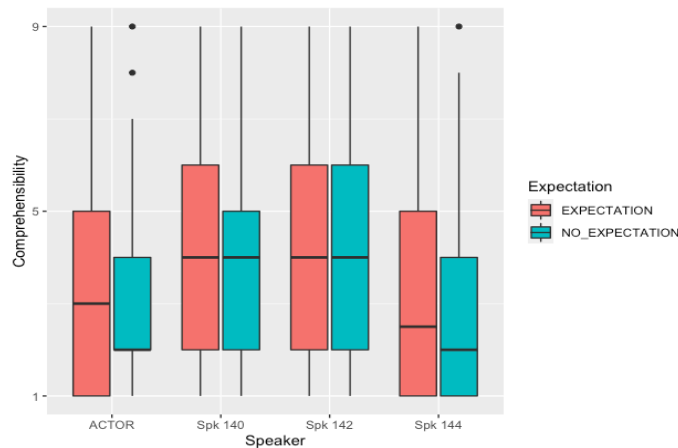


Figure 6. Box and whisker plot for comprehensibility with expectation conditions.

3.2.4 Who is the actor?

The second half of this experiment asked each listener to determine who they thought was the actor after listening to a block of forty sentences that contained ten sentences each for all four speakers in the experiment. Remember, each listener gave ten intelligibility, accentedness and comprehensibility ratings per speaker, but only gave one actor selection per experiment, so the number of data points for this question is a lot lower than the above figures that contain hundreds of data points. Though the overall number of data points is lower, valuable insights about listener behavior can still be

found. In the following figures, the y-axis is the proportion of times that a listener selected a particular speaker in the task. While each of the four speakers demonstrated different scores for the above three characteristics (i.e., intelligibility, accentedness and comprehensibility), listeners chose the actor the most in both expectation conditions. As shown in Figure 7, social Expectation of performance helped listeners choose the Actor more overwhelmingly than in the No Expectation condition.

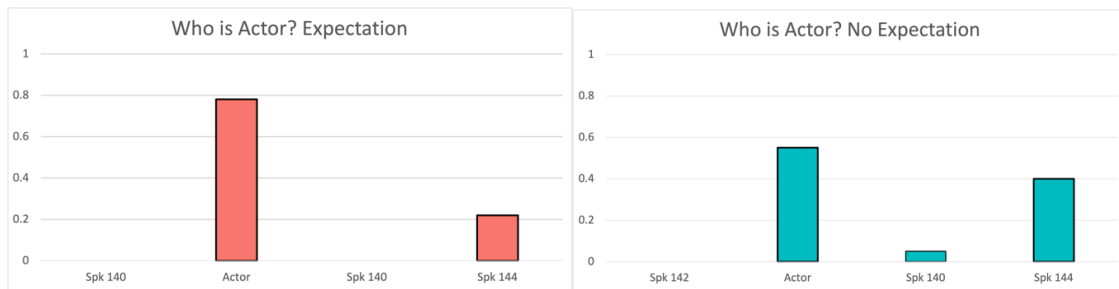


Figure 7. Comparison of expectation conditions for selection of speaker most likely to be the actor.

3.2.5 Who has the most stereotypical accent?

The second question in the experiment, about the listener’s perception of stereotypical accents shows a different pattern. Immediately following the first actor selection question is a second question, asking the listener to select who they think is the most stereotypical Russian-accented English speaker. This question was asked to try to explicitly access the top-down social judgments that listeners might be using while they are listening to speakers of an accent that they can recognize. In fig. 8, selections appear as a wider range of possible acceptable answers. Unlike the first question that was asked of listeners, the idea of a stereotypical accent is up to the interpretation of the listener and there is no “right” answer in the experiment. These preferences subsequently appear in the form of a wider range of answers. For example, Speaker 142 and Speaker 140 were selected in this question and given that they had higher accentedness ratings than the

other two speakers, this answer seems perfectly acceptable. These selections are interesting given the individual profiles of intelligibility, accentedness, and comprehensibility with no clear correlation between the scores these speakers received and the selections listeners made at the end of the experiment.

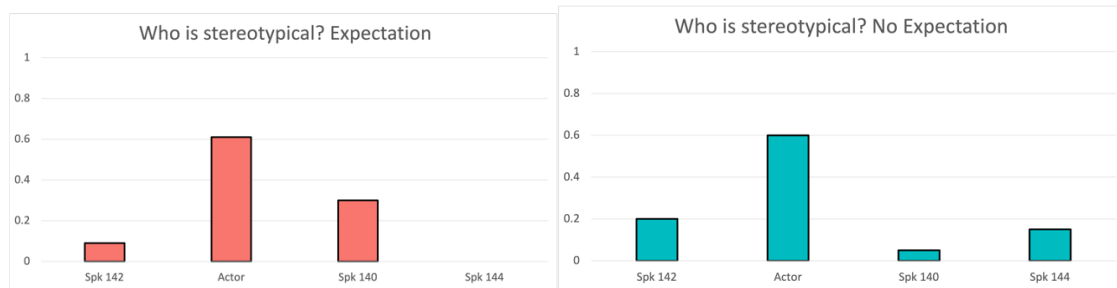


Figure 8. Comparison of expectation conditions for selection of speaker most likely to be the most stereotypical speaker of Russian-accented English.

4. Experiment 2

In addition to the performance specific questions being asked for this dissertation, I designed a second experiment in an attempt to find what kinds of judgments listeners are making while in different social contexts of listening. Managing to capture any kind of differences between Expectation and No Expectation conditions demonstrates that listeners may be sensitive to different listening contexts and adjust their parameters of accent judgment in response to these contexts. The second experiment is designed to probe more closely into the different types of subjective judgments a listener might use in a context of performance, guided by the pilot project and keywords that appear when training voice and dialect. The hypothesis is that these adjectives are more sensitive to changes in social context since they are more associated with the idea of performance as demonstrated in the free answer of the preliminary study.

4.1 Method: Participants, stimuli, procedure

Again, I recruited forty-five (45) additional participants whose first language is

English and had no prior experience with Russian and had not participated in the first experiment. In addition to collecting language background information from each participant, we also asked about their experience in performing and attending live performance (acting, improv, role playing, and any other type of performance). Stimuli are the same sentences used in Experiment 1. Three Russian-accented English speakers from the ALLSSTAR Corpus (Speaker 140, 142 and 144) and the actor are the same speakers as in the first experiment (see appendix for sentences that the listeners heard).

Each participant was randomly assigned to one of two expectation conditions like in Experiment 1. Then each participant listened to each sentence one at a time in a random order. In the Expectation condition, participants were told that they are listening to sentences from four different speakers and one of those speakers is an actor. There was no such explicit instruction in the no expectation condition. After each sentence, participants were asked to judge each speaker on nine-point judgment scales for the five most frequent responses in the pilot study. The scales were established so that 1 stood for “extremely natural,” “extremely authentic,” “not forced,” “extremely clear,” and “not exaggerated.” A score of 9 represented the opposite of these adjectives (shown in Table 5 below). Scales were aligned so that speech that sounded spontaneous would receive a lower score, while the higher scores are attributes that indicated performed speech, according to the adjectives gathered in the pilot study. The exception to this rule was the adjective “clear”, where 1 represents careful, planned or otherwise performed speech. This is the mistake in my experiment design because I assumed that spontaneous speech would be clearer than performed speech, despite the extensive literature that points to the opposite condition.

Table 5. Adjective alignment on the Likert scales for experiment 2.

Score of 1 (spontaneous sounding)	Score of 9 (performance sounding)
Extremely natural	Not Natural at all
Extremely authentic	Not Authentic
Not forced	Extremely forced
*Extremely clear	*Not Clear at all or UNclear
Not exaggerated	Extremely exaggerated

4.2 Results

4.2.1 Adjectives by Speaker

The five box plots in Figure 9 below show the results for all five adjectives. A few patterns are immediately apparent; for example, Speakers 140 and 142 have similar patterns, appearing with nearly identical profiles in all adjective cases. However, in some cases, the scores for the Actor behaves like Speakers 140 and 142, and in some cases the Actor behaves like Speaker 144. The different behaviors for the scores of the Actor may indicate that different vocal qualities are valued in performed voices than others. Even more apparent from this part of the experiment, is that listeners often assigned different values of each quality to speakers, which indicates there is a range of acceptable qualities for these speakers while maintaining the same basic level of intelligibility, as indicated in the first experiment.

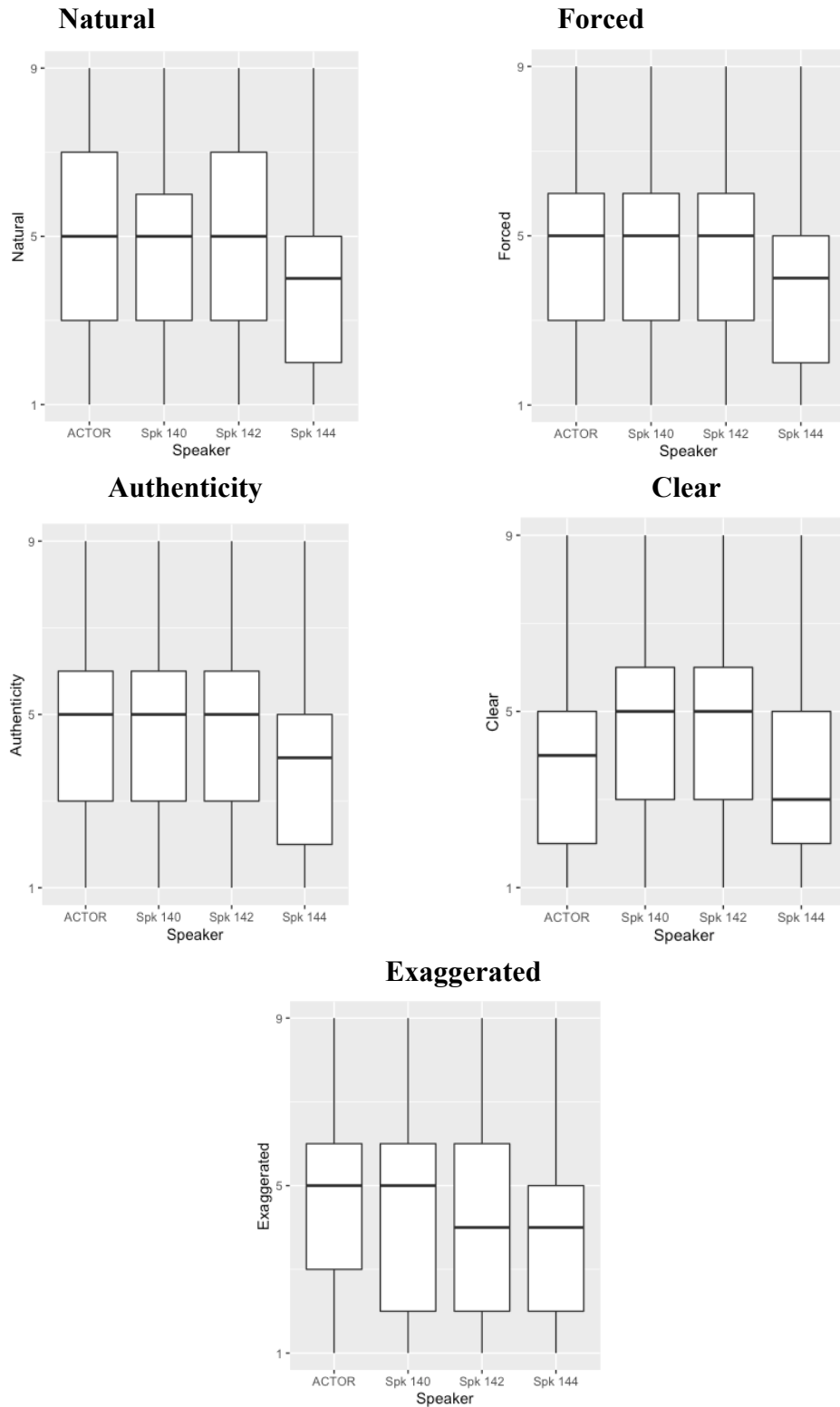


Figure 9. Box and whisker plots of the results from the Likert rating of all five adjectives, by speaker.

For comparison's sake, I have also provided a table of the mean score and standard deviation of each adjective with each speaker. Note in Table 6, with the exception of the adjective "clear," the Actor averaged higher ratings in all other adjectives. Recall that the Likert scale for "clear" was flipped as a result of the experimental design, so having a lower average means they still sound the most performative in that adjective category. In fact, the Actor scored as the most performative sounding out of all the speakers regardless of adjective or listening condition.

These scores indicate that each speaker exhibits a unique set of attributes compared to the other speakers. Each of the three Russian-accented English speakers is different from the other two speakers and have their own unique scores. In the case of Speaker 144, he scores lower than the other two native Russian English speakers on all attributes, indicating that speaker 144 could be perceived as a more "performative" spontaneous speaker *at the same time speaking more clearly than the other two native speakers*. With their similar intelligibility scores from the first experiment, these results indicate speakers can exhibit different combinations of these adjectives and still be sufficiently intelligible for performance. As a speaker, the Actor shares some attributes with Speakers 140 and 142 (scoring similarly in natural, forced, and authentic) and shares other attributes with speaker 144 (scoring similarly in the clear category). These attributes for the actor point to a possible special attenuation to how they are speaking, indicating there might be unique acoustic properties that the actor exhibits that may clue a listener into performance outside of social expectation.

Table 6. Mean and Standard Deviation for all five adjectives, by speaker.

		Actor	Speaker 140	Speaker 142	Speaker 144
Natural	<i>Mean</i>	4.97	4.82	4.76	3.81
	<i>S.D.</i>	2.23	2.10	2.25	1.92
Authentic	<i>Mean</i>	4.91	4.54	4.51	3.86
	<i>S.D.</i>	2.10	1.97	2.10	1.94
Forced	<i>Mean</i>	4.83	4.79	4.48	3.85
	<i>S.D.</i>	2.22	2.12	2.22	1.99
Clear	<i>Mean</i>	3.72	4.68	4.75	3.55
	<i>S.D.</i>	1.91	2.12	2.22	2.97
Exaggerated	<i>Mean</i>	4.53	4.32	4.01	3.75
	<i>S.D.</i>	2.72	2.15	2.16	2.01

4.2.2 Adjectives by Expectation

The following box and whisker plots further divide the data seen above into two listening conditions—Expectation and No Expectation. If social context is a factor in judging these speakers, we should see a difference between the two conditions. How that difference manifests is the key factor in these results and common sense may dictate that expectation of performance may affect a listener’s perceptions more towards indicating that voices are more performance like. For this hypothesis to be true, a listener who is explicitly expecting an actor will score speakers higher on the Likert scales (with the

exception of clear) which indicates the listener believes that the speakers are exhibiting performative behaviors in their voice. In fact, trends in the data show the opposite of this prediction. Figure 10 shows a slight pattern in the *opposite* direction of this prediction. Again, note that red (or the boxes on the right side of each column) indicates listeners were explicitly expecting to hear the voice of an actor.

The overall pattern does not show a lot of change between the No Expectation and the Expectation conditions, meaning the boxes in the conditions for each adjective almost completely overlap and share descriptive attributes in both listening conditions. However, examining a few combinations of speaker and adjectives leads to an unexpected result. Listeners score Speakers 140 and 142 higher in natural in the No Expectation condition than in the Expectation condition, which is the opposite effect than what the hypothesis of using social expectations predicts. Further, Speaker 142 also exhibits similar behavior with the adjective 'clear.' Social expectation of performance has resulted in listeners scoring Speaker 142 as clearer (with a lower score) than when listeners are not expecting performance. This possibly indicates that a speaker might be adjusting their social expectations for speech towards a more generous mode of assessment if they know they are expecting performed speech. This effect seems to become more pronounced for speakers that would otherwise be judged as less intelligible or comprehensible, (e.g., Speaker 142 had the lowest intelligibility mean score of .933, (s.d. = .179)).

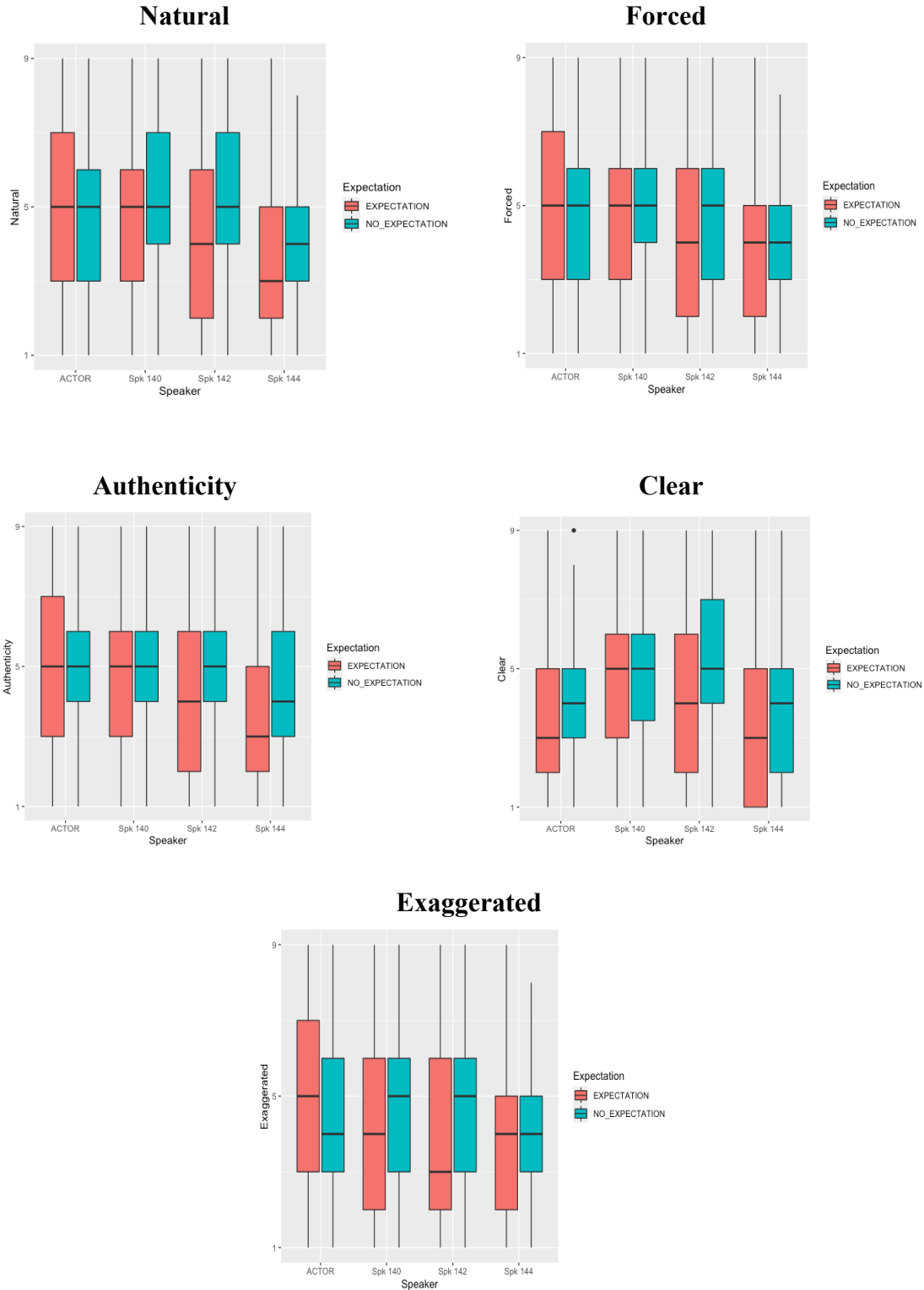


Figure 10. Box and whisker plots for the five adjectives that compare the two listening conditions.

4.2.3 Who is the actor?

The results from the actor selection question pattern differently than those who were asked in the first experiment. Figure 11 shows the comparison between the two expectation questions. The y-axis is the proportion of answers that were given for that speaker. When listeners were expecting to hear an actor, they overwhelmingly chose the Actor as the correct choice almost 60% of the time. However, when listeners were not expecting to hear an actor, they selected the Actor and Speaker 144 around 40% of the time. While strictly better than a chance guess (25%), it does demonstrate that Speaker 144 and the Actor may share characteristics in common. This even split between selecting Speaker 144 and the Actor appears in the first experiment, as well (see fig. 7). Without social expectation of consciously listening for the traits of performed speech, listeners may be tapping other social expectations to determine their choice. In many aspects, Speaker 144 and the Actor scored quite similarly in the rating experiments.

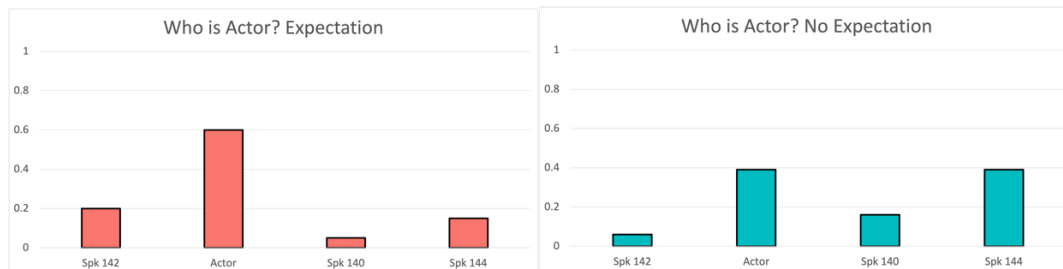


Figure 11. Comparison of expectation conditions for selection of speaker most likely to be the actor in experiment 2.

4.2.4 Who has the most stereotypical accent?

Listeners answered the question about stereotype immediately after selecting their choice for who they believe was an actor in the experiment. Figure 12. compares the two listening conditions, giving proportions of answers that listeners selected. Answers were spread more evenly between the four speakers when listeners were expecting to hear an actor, selecting the Actor approximately half of the time. However, listeners selected the

Actor more than any other speaker when they were not primed to expect to hear an actor in the voices. This curious pattern reflects some aspects of the choices that listeners made in the voices. This curious pattern reflects some aspects of the choices that listeners made in the first experiment. Listeners' expectations of performance might have shifted in the second experiment, since they had just completed fifteen minutes of an experiment that asked for them to listen for "authentic" and "natural" voices, indicating that one or all of these speakers might not be authentic spontaneous speakers of Russian-Accented English. This result demonstrates opinions about "stereotype" can fluctuate from listener to listener. This question about stereotype does not necessarily have one right answer in the way that asking who they believe the actor is has one correct answer.

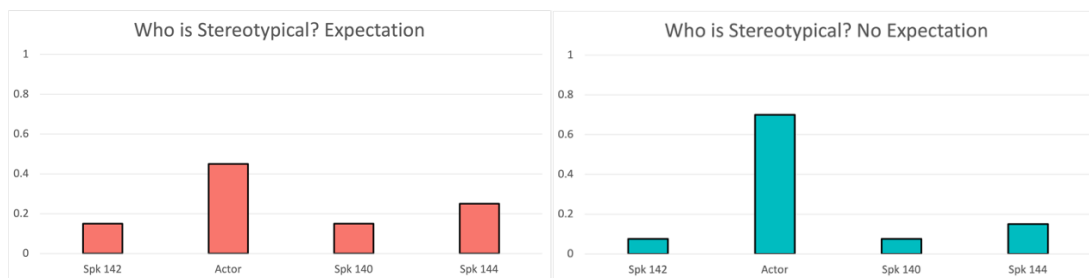


Figure 12. Comparison of expectation conditions for selection of speaker most likely to be a stereotypical speaker of Russian-accented English in experiment 2.

5. Discussion: What can practitioners take from this chapter?

Overall, the results of these experiments show that there are multiple ways to construct maximally intelligible speech. All four speakers were not statistically different from one another in their intelligibility scores, and yet exhibited different levels of different attributes such as accentedness, comprehensibility or different levels of the five adjectives featured in Experiment two. If practitioners of voice and dialect were strictly using measures of intelligibility such that Dudley Knight's advocates, practitioners would be happy to accept all four voices in this experiment as acceptable voices for onstage performance. However, subjectively, differences between the four different speakers exist

and practitioners might feel the need to work with the speakers to make them more ideal for performance contexts. Instead of affecting intelligibility (which is already at maximum, according to these experiments), speakers would modify different aspects of their voices to affect the other aspects that listeners are using to judge these accents like “clear” or “natural” or “authentic.” This work with these voices would not be necessary, since they are already perceived as maximally intelligible.

Using these parameters to perceive these voices becomes even more complicated, because these experiments have demonstrated that the very context of expecting performed speech affects a listener’s ability to judge. Just by mentioning a context change, the goal posts for “authenticity” move when listeners expect performed speech in an unexpected direction. Listeners are more generous in their observations (i.e., willing to accept performed speech as spontaneous-like) and are more likely to be generous in their ratings and rate inauthentic speech as more authentic than contexts in which they would hear the same type of speech outside of performance. This phenomenon could indicate a privileging of performed speech by listeners, which also means that authenticity could be an easier target for voice and dialect coaches than previously imagined.

Another possible explanation of these results is that listeners are only using the word “authenticity” in a particular context that requires some doubt of the veracity of the speech they are encountering, and therefore expecting performed speech more readily primes a listener to use this term. This phenomenon is similar to when listeners describe marginalized voices as “articulate,” implying that the default expectation of the marginalized speaker is that they cannot achieve a certain level of articulation that is acceptable to the listener. This type of compliment is a backhanded way to show surprise

that also can be explained by privileging the white listening apparatus and attendant expectations. To highlight the inequality of the use of this adjective, language advocates have criticized some users of this adjective. For example, when President Joe Biden is quoted as saying that former president Barack Obama is an “articulate and bright and clean and a nice-looking guy,” he is implicitly arguing that he could not expect a Black man in Obama’s position to speak and look the way he does (Alim and Smitherman 10). Results from these experiments show that authenticity might be used in a similar sense by listeners, especially when primed to expect a fake or performed accent, or in a scenario where “authenticity” might be questioned by a listener.

One further aspect of these results worth noting demonstrates that the Actor as a speaker of imitated Russian-accented English is not always the most stereotypical speaker. These results, hand-in-hand with the relative struggle of listeners to actually identify the actor in the experiment, point towards different criteria used by listeners to determine performed speech other than use of stereotype to identify performance. Other research has demonstrated that listeners have some ability at identifying different types of accents from one another (such as German and French accents), but cannot reliably determine if the accent they are hearing is authentic or imitated (Neuhauser and Simpson 1805). Due to the variety in answers in the experiments, listeners may not be relying on their stereotypical representations of this accent to help guide their judgments of performed speech and stereotype. The relative spread of the answers that listeners provided to both the actor and the stereotype question demonstrates that pinpointing authenticity and performance are still relatively difficult for listeners to do and they are not reliable in this task. The expectation from their experiences of performed speech and

stereotypical accents in performance did not reliably provide listeners with enough examples to accurately determine which speaker was the performing imitator. Voice and dialect practitioners should take these results with a responsibility towards their listening audiences; the listener will not always be accurate in telling imitated accents from authentic ones.

On the other hand, the inability of listeners to tell spontaneous accents from performed accents can work in the favor of marginalized actors who do not speak a mainstream version of English who wish to perform onstage. Recall that the intelligibility measure for all four speakers was essentially the same, even between listening conditions. This means that listeners, even while they vary in their judgments of the accents they are hearing, are still receiving information and meaning from every speaker in this experiment. This means that even non-native speakers in their own accents can easily be understood and should be treated as such as performers in their own right, and voice practitioners ought to take care of how they approach intelligibility with their acting students. Specifically in the United States, students and actors who have non-native English accents or regionally accented English can therefore participate with their own unique voice and not necessarily fear that they are unintelligible—in both the colloquial and the linguistic sense—to audience members in performance. However, the accent or dialect they are speaking will still trigger normative language attitudes and can contribute to subjective meaning-making when the individual audience members combine their ideas about how these speakers sound and what they are hearing as part of the script. Social expectation of performance can even push the boundaries of acceptable or performance voices more since listeners are more willing to judge accents to be more

spontaneous-like in their delivery, as demonstrated by experiment 2. The results of these experiments show that practitioners have more flexibility than once thought because listeners can accept a wider range of what it means to be an intelligible voice onstage. Simply expecting performed speech boosts these voices towards a more generous interpretation for a host of attributes that listeners use in the audience.

Future directions include creating a better design to tease apart expectation of performance and expectation of stereotype, since these experiments have not made clear the link between the types of judgments listeners were asked to make and their selection for stereotype, as evident in the wide spread of answers for the stereotype question. In the future, the order in the procedure of the experiment can change, where I ask listeners their stereotype judgments before they participate in the rating task. Then, after the rating tasks, I can ask listeners who they think is the most stereotypical speaker. If listeners change their answer and subsequently change their scores for these attributes, I can gain more insights into the types of judgments listeners make when considering stereotypical voices. I can then compare first impressions of listeners with listeners after they have had fifteen minutes to adapt to these voices. I can also create a scenario where the speech is masked (i.e., less intelligible in the empirical sense), which might result in a wider array of judgment scores for all four speakers and help to reveal any significant differences that social expectation makes for listeners, especially in a scenario where they must rely more heavily on the social expectation of performance. Another challenge is present in the form of Rubin and Kang's "reverse linguistic stereotyping." Marginalized actors in bodies that read as non-white will also face a challenge with respect to social expectations of how they are expected to sound onstage. A future direction in this

research can explore that space between expectation of accent from marginalized bodies and adaptation to these voices when these expectations are not initially met. How can we exploit the gap between expectation and reality to train audience members to expand their initial expectations when interacting with new people?

The main results that these experiments demonstrate for voice professionals fall into three themes. The first theme is that the context of performance will necessarily change how listeners are perceiving the language they encounter. The second theme is that there are many different ways to create a voice that has maximum intelligibility; all of the voices in this experiment could convey linguistic information to listeners. This second theme means that voices that have historically been excluded in performance are not excluded because they were unintelligible; prior prejudices and catering to raciolinguistic ideas of the white listener have excluded these performers in the past. The third theme is the ambiguous relationship between authenticity, imitation, and stereotype of which voice practitioners can take advantage. Listeners cannot reliably discern spontaneous accents from imitated or performed accents, which means there exists a space where voice practitioners can use this fact to highlight real voices onstage next to trained accents and dialects and manipulate audience expectations towards the benefit of marginalized or non-standard dialects. These three themes will influence the practical steps this profession can take that I highlight in the next chapter. The results combined with the alternative approach of conceiving voice training in general will offer a foundation for the best practices for voice training—and listening!—in the future. I will then tie these steps into other work that is being done to expand the idea of theatre in the final chapter and conclusion to this dissertation.

CHAPTER IV

TOWARDS A NEW TRAINING PARADIGM FOR VOICE PROFESSIONALS

“We celebrate it when white actors nail an accent...We celebrate it! Until we can offer that same detail and attention to all linguistic identities, and to a myriad of accents, we’re still going to be erasing the humanity of those stories and those characters.” Cynthia Santos DeCure, “How Should Black People Sound?” *New York Times*

“One of the most repressed things for Black people in this country has been our voice. Right now, we’re seeing if we can really find our voice, at this time, and this specific moment, to specifically tell this story — this beautiful thing — the way the team wants it to be told.” Tre Cotten “How Should Black People Sound?” *New York Times*

1. Enacting the expansive imagination in voice

I begin with the above quotes from the article “How Should Black People Sound?” by Reid Singer, which ran in the *New York Times* on October 28, 2020 as a vision for the future of the profession of voice and dialect, explicitly correcting the white supremacist foundations upon which this profession was built. My research offers a critical look at these supremacist foundations, examining the assumptions upon which the profession is built, and then interrogating those assumptions in a way that reveals the social construction of what it means to experience a voice onstage. I have used two critical lenses on the profession and practice of voice professionals to examine how the profession might grow more equitably into the remainder of the 21st century and beyond. One lens, a critical philosophical examination of the assumptions of voice practitioners throughout the history of the practice, demonstrates the role of the profession in cultivating and enforcing explicit racial and cultural stereotypes. The assumptions of these professionals asked two important questions. The first question, a question of unlearning, seeks to eradicate any trace of what Micha Espinosa refers to as “cultural voice” or marker of class, gender, or race (75). The second question of previous voice

professionals seeks to implant a sanitized version of the voice, whether through generalized standard accents, or through stereotypical foreign and regional dialects.

These two questions appeared in the name of an objective measure of clarity or intelligibility; the idea that the audience needs a sanitized version of authenticity governs the previous choices of voice professionals. Paradoxically, voice professionals assume the audiences' hunger or awareness for authenticity on the stage but historically have not offered a system with which authenticity of various lived linguistic experiences may be honored. In this case, voice trainers' perceptions of audience need drives the profession, and not necessarily the lived experiences of actors nor the audience. Intelligibility, a driving force in the working epistemology of voice, has been elevated as an objective and separate gold standard by which voice professionals ought to measure the success of their work. By treating intelligibility as an objective measurement, voice professionals again denigrate the autonomy and lived experiences of actors and audience alike, by excluding any differences in subjective experience in society.

A second lens, and the key intervention of this dissertation, takes seriously the notion of intelligibility as a socially constructed judgment that has a real-world effect on perception and affects individuals differently when they use intelligibility to perceive their world. This lens is from the cognitive linguistics field and offers a different explanation of how speech perception may work in performance, which opens an avenue to respect different lived linguistic experiences of actors and audience members alike. This lens uses cognitive research into speech perception to establish an experiment of intelligibility in performance, demonstrating that intelligibility may not be as objective or easy to measure as voice professionals may desire. Interrogating this assumption of

intelligibility is necessary to support a contemporary approach to voice training that actively resists the racist underpinnings of the profession. While these findings complicate the picture of providing professional voice services in some ways, this also demonstrates that audiences may be more willing to meet the actor in their own lived linguistic experience and may not be relying upon “authenticity” as a measure as common sense might imply. Further, the use of “authenticity” only arises when listeners are using their expectations of stereotype to guide their listening. This linguistic experiment demonstrates that listeners are willing to change their perceptual patterns *only with the mere suggestion* of performance or pretend. In other words, empirical evidence suggests that audiences are more likely to receive marginalized voices in the realm of performance and theatre.

With a critical look into the past of these practices and a cognitive look into the present-day reality of actual audience behavior, I can now offer my own approach as a white theatre maker and cognitive scholar that incorporates a more just and equitable voice practice for the future. To do this, I look to colleagues in language education, and will incorporate what April Baker-Bell calls an “Anti-Racist Black Language Pedagogy” that explicitly acknowledges and actively works against the white supremacist structure that the practice of which performed language is a part (11). By acknowledging the problematic foundations, introducing the working assumptions of the profession, and offering evidence backed ways to push against these assumptions, I can contribute to a profession that can honor healthy holistic approaches to the voice and lived linguistic experiences of those who have been excluded from theatre, film, and entertainment. With enough practice, this approach can actively push back against the larger societal

phenomenon of linguistic discrimination by extending the audience and performers' imaginations in honoring actor's own lived experiences and introducing accents and dialects in dynamic and surprising ways.

Creating an anti-racist approach to dialect training that does not rely upon pre-existing notions of standard language ideology or imperialism offers the chance for practitioners and audience members alike to experience a different form of empathy by enacting their imaginations towards a world where people of all different lived linguistic experiences have a right to their own stories in performance. Anti-racist approaches to dialect and voice training require an explicit examination of the assumptions that drive this behavior, and the choices that continue to be made regarding using accented language onstage. This type of empathy does not necessarily always mean that the audience will be comfortable, as confronting detrimental biases sometimes require sacrifice of comfort and complacency in the moment. As theatre trainer and anti-racist activist Nicole Brewer says, "It's okay to explore another's lived experience with empathy and curiosity, and deep profound listening, and it's okay not feeling comfortable" ("Training with a Difference"). The ethical imperative of vocal professionals is to provide a carefully scaffolded cultural framework to introduce these conversations into the production process, from pre-production through actor training. The benefits include creating a linguistically diverse soundscape onstage which means that more actors and producers have access to American theatre-making.

According to my preliminary linguistic data, audiences might be ready and eager to incorporate this soundscape into their meaning-making while experiencing the theatre. It is the ethical responsibility of the voice practitioner to take the steps necessary to

respectfully represent linguistic lives on stage; and this dissertation has offered another critical intervention to ensure the respectful representation of dialect by considering audience experience as a direct factor in perceiving these dialects. At the same time, Micha Espinosa still explains the core goals of a voice practitioner,

I do believe that a voice should be free of those pesky glottal attacks and/or have the ability to sustain throughout a run of a show, but it was at that moment that I became aware of the cultural voice. A voice that has endured the dirt and struggle of constantly crossing borders might not be as aesthetically pleasing to some, but it was a lot more interesting to me (78).

Balancing audience expectations between “interesting” and intelligible need not be tied to racial, gender and class expectations or stereotypes. Just as listeners construct the notion of intelligibility in everyday speech interactions, we, the performers, and practitioners, can construct the idea of intelligibility to create an inclusive approach to voice. In other words, the choices we make as artists and practitioners hold real-world consequences, who we represent linguistically on stage reflects the values we hold while making our art. We must interrogate our own assumptions that guide our art creation, and to do so means we ought to be reaching for new tools and lines of inquiry. This is the lesson above all that I hope practitioners and scholars alike learn: theatre-making itself can be a critical tool in combating the detrimental belief system of a society that believes that how a person sounds is related directly to their worth.

This dissertation has sought to analyze the profession of voice and dialect through the dual methodologies of cognitive philosophy and linguistics in order to interrogate the assumptions driving this profession. In the preceding chapters, I detail the historical

material circumstances and the prevailing assumptions about voice to establish the challenges facing contemporary vocal trainers. I created my own empirical inquiry to demonstrate that the notion of intelligibility, one of the driving qualities behind much vocal professional work, is in fact a subjective measure that is susceptible to implicit biases of the listener—information that should be a central concern to theatre practitioners and all cultural workers. These normative assumptions are baked into the theatre-creating apparatus and have contributed to a negative feedback that reinscribes and damages marginalized practitioners. The question before us as theatre practitioners generally is: how will we respond to the present call for a reckoning with the normative raciolinguistic listening practices of professional and regional theatre in the United States? This dissertation is my attempt as a white theatre maker and voice coach with specific linguistic training to wrestle with the central question of equity in approaches to vocal training. I will not be able to correct all the historical violence that this profession has wrought on marginalized performers, but I offer an approach, using my own experience, that can begin to crack open anti-racist practice of voice training.

The previous chapters have established a complicated story of the interchange between audience expectations and the expectations of theatrical practitioners to build the idea of linguistic intelligibility on stage. In what follows, I will use the three themes that appeared in my experimental work to highlight my approach to voice and dialect practice. The three themes—listening is context dependent, maximally intelligible voices have a wide array of attributes and qualities, and the complicated relationship between authenticity, stereotype, and imitation—will appear in my past examples of work as I grapple with the question presented above. I will also use the historical lessons and

assumptions that previous voice professionals have infused into this profession. I will push back against the two core goals: unmaking voices in their grit and lived linguistic experience and replacing these voices with sanitized or simplified versions of voice.

I am gearing this last chapter towards dialect coaching specifically for two reasons. The first reason is that the practice of training dialects has a highly fraught and explicit history in enforcing stereotypes, therefore will be a fruitful site for anti-racist intervention. As linguist Vijay Ramjattan once remarked, “language practices are racialized, and language practices racialize as well” (Twitter). The language practices of theatre are not immune from the racializing practices of assigning accents and dialects to characters with unsavory attributes, linking how people sound with innate negative stereotypes. The demand for the practice of dialect coaching is only growing and this is a prudent time to create an ethical framework from which to work (Singer). Secondly, I have personal experience in dialect training throughout my career as a dialect coach and have taken concepts from this research in this dissertation and applied them to project based work. The three examples I will draw most heavily upon are my work in *Pilgrims Sheri and Musa in the New World* by Youssef El Guindi, *Good People* by David Lindsay-Abaire, and *The Language Archive* by Julia Cho. To accompany my practical work, I will draw upon theoretical examples from several other plays.

This work of undoing years of explicitly racist structures in the profession mirrors work by other practitioners in other areas of representation on stage. My own practices run in conversation with practices offered by Nicole Brewer and guidance set forth by the group *WE SEE YOU WHITE AMERICAN THEATRE*, which advocate for centering marginalized voices in all aspects of production, from onstage to offstage to the front of

house and artistic management staff (“Training with a Difference”). Some voice work in this vein is gaining recognition in professional circles, from practitioner Daron Oram’s 2019 article “De-Colonizing Listening: Toward an Equitable Approach to Speech Training for the Actor” to featuring the story of dialect coach Tre Cotten as a dialect coach in the New York Times article that began this chapter. I borrow not only from theatre practitioners, but also language teaching advocates like April Baker-Bell’s “Anti-Racist Black Language Pedagogy” in her 2020 book *Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy*. Many different scholars and artists from marginalized communities have already begun this work and I will highlight some of this work in a section that prefaces my own practices by featuring the work of Latnix theatre makers. All of these examples of alternate approaches to theatre production serve to liberate and celebrate historically marginalized communities, offering an alternate vision of how theatre can shake loose the shackles of linguistic white supremacy.

This chapter serves examples of liberatory linguistic practices that already exist within theatrical production, and the findings in my own linguistic research and dialect coaching support these practices. I seek not to reinvent the wheel, but to support and reaffirm the work already being done. Like the rest of my research and work, the method by which I conduct my work as a dialect coach sits at an intersection of use of my background as a linguist and as a theatre practitioner, utilizing the subjective knowledge I have accumulated as vocal coach and scientific knowledge gathered from my peers. Some of these practices have already been employed and some of these practices are steps I would recommend for the field. What has become abundantly clear throughout this research is that simple and robust solutions to equity, diversity and cultural

competency in voice practices do not exist. The thorny history of white supremacy, raciolinguistic practices, and the material circumstances of theatrical creation has limited historic approaches to voice.

2. Case Study: contemporary linguistic needs for Latinx Actors and Directors

Contemporary voice practitioners are already creating space for voices that do not necessarily fit the mold of accepted performance, and I would like to discuss an approach that one facet of this community is using to push back against the dominant white mode of theatre production. Borrowing from the scholarship of Gloria Anzaldúa, Micha Espinosa, a voice trainer and professor at Arizona State University, uses the term “cultural voice” to describe voices of actors who do not fit the dominant mode of acting, “Cultural voice is described as the self-constructed, emotionally bound, non-dominant performer’s identity and identification with the social-historical values and principles of one or more cultures” (75). Espinosa enumerates the difficulties of working as a marginalized voice in a white-dominated space by connecting her own struggles as a voice practitioner with Anzaldúa’s 1987 essay “How to Tame a Wild Tongue” by connecting her work as a voice practitioner in higher education with the linguistic terrorism of the “authentic wild tongue” (75). While in the white-dominated space, Espinosa feels, “to succeed, Latinos and Mexican students have to negotiate an identity with the psychological and physical realities they have been given. Both students and teachers often find themselves working with unexamined and opposing sets of external and internalized beliefs” (78). In this way, Anzaldúa’s concept of border identity is reflected in Espinosa’s personal experience, in a way that evokes physical location and environment as stand-in for the linguistic realities of students:

I have been straddling that *tejas*-Mexican border, and others, all my life. It's not a comfortable territory to live in, this place of contradictions. Hatred, anger, and exploitation are the prominent features of this landscape. However, there have been compensations for this *mestiza*, and certain joys. Living on borders and in margins, keeping intact one's shifting and multiple identity and integrity, is like trying to swim in a new element, an "alien" element (18).

I have demonstrated that the desire to "Tame a Wild Tongue" and its detrimental effects can be seen throughout the history of the voice profession, and yet this desire still remains in contemporary practices of training in both theatre and film. The discomfort of marginalized speakers in this new element of catering to the white listener is the focus of Espinosa's training and her commitment to honoring the voices that "constantly endure the struggles of constantly crossing borders" as rightful participants in the theatre making apparatus of higher education theatre. Oftentimes, the desire in terms of voice is for a mildly Hispanic accent that can read as "spicy" or "foreign" as a cognitive shortcut to character or entertainment, but the accent or dialect does not read as coming from a distinct culture or geographical area. The psychological damage done to the actor is described succinctly by Espinosa, "When we unconsciously continue to cast that one Latino student as a spirit, or 'other,' we again propagate Eurocentric dominance and the student's social marginalization" (81).

To get a deeper read on the contemporary linguistic issues facing Latinx actors and theatre producers, I would like to highlight a recent conversation I had with fellow scholar Dr. Olga Sanchez Saltveit, who is a director and educator with over twenty years

of experience bringing Latinx stories to the stage.³⁷ Sanchez Saltveit shared her frustrations with the entertainment industry. Specifically, Sanchez Saltveit points to historic assumptions of directors and producers that, “any Latinx can be any kind of Latinx” (Personal Communication). This expectation results in actors creating a type of accent or dialect that is not reflective of any one geographical area as part of fulfilling the expected role. These stereotypical accents often accompany stereotypical roles, reinforcing harmful social and racial stereotypes. Further, these accents erase any type of Latinx indigeneity; speakers of minority languages such as Quechua (spoken in some South American countries) are not represented with these generic Spanish-dominant accents. Additionally, Sanchez Saltveit points out that actors are sometimes asked to translate from Spanish to English on the fly, thereby performing a type of free labor for bilingual productions. Sanchez Saltveit’s own personal experience is reflected in larger industry patterns, and Mexican-born film producer and director Batán Silva remarks in a recent *New York Times* article, “There’s nothing worse than a Mexican character who sounds like an Argentinian or a Spaniard...Or actors who say seven things in Spanish, and then miraculously switch to English” (qtd. in Singer). He follows this statement by remarking that production studios are beginning the work to diversify their production staff truly and deeply to more accurately reflect the specific cultures required to fulfill the script’s demands, actors and voice professionals included. Sanchez Saltveit has seen

³⁷ Thanks to Dr. Olga Sanchez Saltveit for permission and input on this section that summarizes this conversation. The original scope this dissertation included a systematic survey and interviews with practitioners, coaches, and directors, but that plan had to be altered in the wake of COVID-19. What follows is a summary of a conversation that could serve as a source of future inquiry with other theatre professionals.

promising signs of change for the industry as well, as professional theatre casting has begun to reflect each script's specific cultural demands more accurately.

Balancing these pressing issues of representation are the audibility requirements and needs of the production, namely that actors still must be heard by their audience. In Higher Education, where a fair number of Latinx productions are produced, Sanchez Saltveit warns against using actor training to reinforce its own type of standard language expectations. She says, "Asking Latinx actors for Latino accents ultimately enforces codeswitching and appropriate times to be 'Latinx' or not" (Personal communication). The use of the appropriateness paradigm in language teaching again assumes the primacy of the typical white listener as the target for language (Rosa and Flores). While she is sensitive to how alienating university and professional theatrical production can be for marginalized students, Sanchez Saltveit agrees that actors must deliver their text with clarity. Whatever accent they may bring to the process, she asks all actors to be "clear, strong and vocally present" (Personal Communication). When I asked her further about the definition of clarity, she responds that she defines clarity as,

The ability to understand people's words, what they are saying. I know my hearing of what people are trying to say is broader because I have more experience with so-called accents. I grew up listening to people speak accented English. *Clarity has to do with the receiver, and the ability for your audience to discern what the actor is saying.* (Personal Communication, emphasis my own).

The professional instincts of Sanchez Saltveit in this one conversation confirm what the linguistic experiments of the previous chapter demonstrate—that judgments of intelligibility and clarity are in the minds of the audience. With a little practice, listeners

can also increase their hearing range to understand a wider breadth of voices onstage. Further linguistic evidence of practice and accommodation can be found in the work of Melissa Baese-Berk, Ann Bradlow and Beverly Wright, where listeners are able to adapt to novel accented speech after training on a variety of other types of accented speech (EL177). Change happens by two avenues in both production and perceptions—including historically marginalized theatre professionals in the larger production apparatus and exposing audiences to a wider variety of lived linguistic experiences.

This brief conversation offers a very specific viewpoint of linguistic issues in a particular identity group and does not begin to address the nuance and myriad of approaches in on theatre producing community. Other marginalized theatre creators run into similar obstacles while creating theatre in the larger theatrical apparatus in white-dominated field. The rest of this chapter is dedicated to imagining a world and creating an ethical framework that not only minimizes the harms of centuries of stereotyping in entertainment, but also offers a way to actively undo harms that have been created by the dominating raciolinguistic assumptions of larger society. As executive producer Lang Fisher says, “We don’t want caricatures, and so it’s important not to have actors just winging the accent” (qtd. in Singer). The rest of this chapter is aimed at encouraging thoughtful and diligent work from pre-production through audience interface.

3. Pragmatic answers to utopian questions

These practices are especially fraught when the production still requires training an accent or using dialect work as part of building the world onstage. This scenario involves a production team—casting, directing, and vocal coach—who all attempt to not only create an artistic statement satisfying to those who are part of production, but also

desire to cast this show as ethically as possible. The old approach to this issue would be to hire a dialect coach and have them teach a pre-created accent or dialect that the production requires, either from available materials, or custom crafted by the dialect coach to fit the desires of the production team. Often times this accent or dialect is made of different elements or “ingredients” of the voice. with “a recipe for every person’s voice” without much regard as to the source of these ingredients (Barton, qtd. in Sakland 30). This approach, with the two explicit goals of voice practice of un-creating the actor’s voice and the re-creation of a pared down version of an accent from pre-existing ingredients, is a reification of raciolinguistic ideals aimed with a white listener in mind. I have established through the work of this dissertation that these practices are problematic at best, and actively harmful at worst. To counteract this issue, production teams have several options. The most extreme of these options is to forego the accent or dialect entirely. This option risks flattening everyone’s lived experience to a “neutral” or General American accent that disqualifies the vast majority of actors who do not have a middle-class white cisgender background. Another option is to have actors use their home dialect or accent when in production. This option honors the lived linguistic experiences of individual actors, but might not serve the story as intended, and runs the risk of again reinscribing harmful stereotypes (if, say, an actor plays an angry character and happens to speak in a dialect from a marginalized group). Neither of these options serve to honor the cultural specificity that this era of entertainment and theatre production deserves.

Casting considerations may also reveal yet a third option for using dialects and accents in production. Responsible casting in theatrical production requires casting actors who share the basic traits of the characters they are to portray. For example, the

documentary *Disclosure* advocates heavily for roles about trans characters to be given to trans actors (Feder 0:05:16). Production teams might advocate for casting actors who already speak the accent or dialect in question for production. At first blush, this advocacy avoids the traps set forth by the first two options for our production team, and third option reveals the desire for authenticity in representation onstage, which is a term that has been heavily problematized in this dissertation. As this desire for absolute authenticity is often practically impossible, I would caution against this enthusiasm. Brian Herrera addresses this issue regarding Latinx casting in production,

A more rigorous advocacy for culturally competent presentations of plays engaging Latinx racial, ethnic, and gender diversity need not solely rely upon demands for *authenticity*, indeed the hunger for authenticity—often rooted in some combination of fear and fantasy—can risk fetishization as readily as if promised the reward of cultural validation (33).

The use actors with authentic dialects and accents risks reducing the nuances of the lived experiences of actors to a token that is ultimately used as a stand-in for the accent, for which, as my experiments demonstrate, audiences will be carrying a stereotype regardless. Both Herrera and one of the takeaway points from my experiments agree, “the appearance of authenticity always lay in the eye of the beholder...the priority of presenting...requires a more reliable and more rigorous protocol than authenticity” (33).

The average production team is put into an impossible position; they cannot approach the idea of accent and dialect through appealing to authenticity or without any consideration or regard to the effects of how different voices present onstage. The following section explores yet more options to approach dialects and accents onstage, a

place where dialect training can still be practiced, but is practiced with a heightened awareness and care towards the crushing mechanisms of white-dominated cultural spaces. The keys to approaching linguistic casting for the stage includes research with respect to the sources of this research, collaboration, and honoring the lived linguistic experiences of both the actor and the obligations of character as written or envisioned by the production team. The approach requires interrogation of the playwright, direction, voice training, and ultimately the audience to truly determine the role that accent and dialect training play in theatre making. There is hope yet for voice and dialect training!

4. My own practice recommendations for dialect

A strong vision of voice and dialect work within production situates the dialect or voice coach in conversation with the usual players—playwrights, actors, and directors—and incorporates their positionality as theatre makers within the white-dominated space that theatrical production has historically been a participant. We can expand that linguistic positionality awareness to those often excluded from voice practice and include dramaturgs into the practice of accent and dialect. To keep my own practice as ethical as possible, as adopted from Chelsea Pace and Laura Rickard in their book *Staging Sex: Staging Sex Best Practices, Tools, and Techniques for Theatrical Intimacy*, I aim to be the first to name power structures in the room, both in production and working with actors to begin to demystify the myriad structures that uphold the practice of theatre (16). This looks like naming my point of view in the world as a white cisgender femme queer theatre maker³⁸ who is (sometimes) paid for my expertise to teach and train with actors under my care. These practices I offer are from my limited and privileged point of view

³⁸ I aim to use adjective that describe myself from most visible label to least visible.

as a white theatre maker who aims to become an anti-racist co-conspirator in breaking the white-dominated structures that still govern theatre today (McIntosh 6).

In plays that require dialect, the two positions of playwright and dramaturg contribute to interrogating the necessity of the dialect or accent in the first place. Because theatre makers strive to create stories that feature characters from many different backgrounds and perspectives, these characters will require the vocal obligation of cultural competency, promoting a space for respectful vocal training that acknowledges the necessity of trained dialects. What follows is a guide that roughly divides the production process and the ideal roles that a vocal professional may play throughout the life of a production, beginning with season/play selection and ending with public engagement about language in the play. Some of these practices spring from my own experiences as a coach, while others are suggestions that will enhance approaching dialect with a mind towards ethical and responsible considerations of voice.

Following each subsection concerning guidance for each segment of targeted audience, I offer questions as guideposts for voice professionals in pre-production, working with the actor, and working with the audience. These questions build upon the suggestions of Bonnie Raphael in her 2000 article “Dancing on Shifting Ground” and Kim James Bey in her 2014 article “Speech Stereotypes: good vs. evil.” These questions will bear quite a resemblance to Elinor Fuchs’ “Visit to a Small Planet.” The final section of this conclusion will be thoughts that speculates on the future of the profession itself and how we can put into place practices for a more ethical profession going forward.

4.1 Who let this dialect: Pre-production and season selection

In contrast to contemporary approaches to dialect coaching, where coaches may

be selected after the director has determined they want a dialect in their production, the dialect or vocal coach ought to be invited to the initial conversations around accent or dialect needs for an entire season. In effect, the dialect coach serves to remind producers to consider ethical issues of voice and representation from the beginning. Historically, inclusion of dialect coaches in season selection has not been the case; in a report that surveys productions in the 2018-2019 season for member theatres of the Theatre Communications Group, Melissa Tanning-Kollwitz, Joe Hetterly, and I found that the overwhelming majority of determining need for dialect work lay with either the artistic director of a theatre company at season selection, or with the director when they are initially assigned their production (5).

Dialect selection practices would benefit greatly from the keen expertise of dialect coaches participating in initial conversations with production teams. Further, this position ought to be adequately compensated so that the dialect coach may not be tempted to advocate for dialect work in the production so that they may be tempted to pay themselves for the work within the season if they recommend dialect work for a particular production. The model of inclusion of a dialect or voice professional reflects the movement to de-gigify or create steady creative work in theatrical production, working back towards the models of theatrical creation in the 1950s and 1960s where professionals are hired for entire seasons or as permanent staff at professional theatre companies (Zazzali 48). Some regional theatre companies, like Oregon Shakespeare Company, do employ full-time voice professionals as company members. More positions like those at the larger theatre companies ought to be the norm, with a retooling of season selection to include the expertise of voice and dialect coaches. This recommendation

reflects a call towards more stable employment in general in the sector of theatre, put forth by Brian Bell and Sam Hunter in *American Theatre*, arguing for a vast expansion of federal and state funding for live performance (“How U.S. States Could Fund Repertory Resident Theatres”).

Another consideration in dialect coach selection is fitting the lived linguistic background of the dialect coach with the anticipated needs of the production or season. In this sense, actors who come from different backgrounds are more likely to see their lived experience affirmed through the professional conduct and creative decisions of the dialect coach. Micha Espinosa describes this affirming choice through her first experience of studying at a Patsy Rodenburg voice intensive under the direction of David Carrey:

I had never discussed the aesthetics of voice. I had adopted my Anglo teacher’s aesthetic. The voice teachers all agreed on the benefits of a clear tone and a healthy instrument. But one of the voice teachers, a non-native English speaker, liked a voice with a little dirt in it. A voice that sounded like it had life. Maybe that life was hard? Maybe that voice had imperfections? (78)

By training and employing voice professionals of different backgrounds, the profession can already begin to deconstruct the assumptions behind the chosen “aesthetics of voice” that has dominated the practice. In this case, the perceived voice aesthetic of the profession Espinosa was entering did not have metaphorical dirt by matching the listening expectations as determined by the voice professionals that were instructing Espinosa in this workshop. Productions that employ marginalized or non-standard versions of different language varieties ought to endeavor to find and employ voice professionals of similar backgrounds. Luckily, resources are emerging that help support

this recommendation for best practices. VASTA, as the leading professional organization that tracks voice professionals, has started to include search terms in their search service that honor different experiences. These terms include “cultural identity”, “equity/diversity/inclusion,” and “social justice” (“Find a Pro”).

The ethical responsibility of the dialect coach is to ask loudly and often if the accent is indeed necessary for the production. For example, if the playwright desires a Southern American regional dialect for a side character that is often teased for being stupid or dumb, the dialect coach may question that choice by asking if the playwright desires to play into stereotypes and/or might question the use of that accent as a short cut within the play or performance to signal to the audience about the character’s intelligence. From my expertise, the moral duty of the dialect coach is to remind the playwright that accent is not indicative of intelligence and will ask the playwright to deeply consider their own biases and to make a new choice. In some instances, the choice of using a perceived non-accent or General American English³⁹ accent is also worth consideration as part of the meaning-making process for the audience. The dialect coach must interrogate this choice, since the choice to or desire to “do away with accent altogether” privileges the idea of general or neutral accents as maximally intelligible, which is an idea that this dissertation works diligently to interrogate. In entertainment, this route to eschew expectations of matching character background to accent has been used to great success in HBO’s television series *Chernobyl*. In an interview, the show’s

³⁹ The definition of General American English that I am using is from Tonning-Kollwitz and Hetterly 2018 and defined as, “a dialect of North American English that is free from regional characteristics” (295). The specific phonology is available as online supplemental material for this article.

creator Craig Mazin explains, “We didn’t want to fall into the “Boris and Natasha” cliched accent [from The Rocky and Bullwinkle Show] because the Russian accent can turn comic very easily” (Freeth). Unbridled from the concerns of authenticity, choosing not to match accent or dialect with immediate expectations may provide for a fruitful avenue of theatre creation.

In this position, a voice professional must account for ethical considerations in this position at the time of season selection including include a deep consideration of which accent or dialect is a) required by the playwright b) desired by the director or production team and c) appropriate for the actor who must use this dialect. The first source of requirement ought to be the intentions and effects of accents used by the playwright in their writing. In new works development, a voice professional ought to act in the way of dramaturg and must ask the playwright two very important questions if the playwright desires to use a specific foreign or regional dialect for their characters. The first question, familiar to new works dramaturgy is, “What is the work that you think the dialect is doing for creating meaning for the play?” and the second question is, “what is the work that the dialect is actually doing for the play?” Oftentimes playwrights may desire the use of dialect as a sort of cognitive shortcut or stand-in for certain traits of their characters, which may cut down on exposition. However, these same cognitive shortcuts are very close to societal stereotypes and may in fact be reinforcing biases and stereotypes in ways unintended for the playwright. The answers to the two above questions may begin to disentangle intent of the playwright in a new work with the potential impact of the new work. To illustrate how dialects may be used in new work development, I will use two different examples from established plays and playwrights to

highlight the need to carefully think through the use of dialect as character trait.

Often, to achieve their desire for a certain dialect or accent, playwrights might include instructions on how they prefer various characters sound. These instructions vary between vague instructions in the stage directions to spelling changes to indicate phonetic differences between characters. The dialect coach's responsibility in all cases is to interpret the intentions of the playwright and the work the dialect is doing for creating meaning on the stage. To demonstrate how to explore the ethical ramifications of dialect and accent requirements in playwriting, I will contrast two approaches to dialect that are baked into the writing of the play which will include the background of the playwright and the work that the addition of an accent or dialect might be doing in meaning-construction for the audience.

In *Pilgrims Musa and Sheri in the New World*, playwright Youssef El Guindi uses only sparse instructions for dialect work in his stage directions, which indicates that the stage directions that are included are pointing towards pertinent details that must be included in the production. The opening character instructions read, "Musa (*Offstage; accent*)" followed immediately by "Sheri (*Offstage*)" (66). El Guindi indicates his desire for one of his main characters to have an accent—thereby also implying that the other character does not have an accent or speaks with a General American English accent that is read by the audience as neutral or accent-less. In this case, El Guindi wants to vocally separate Musa, who is a recent immigrant to the United States of America, from Sheri, who is a character native to the United States of America. Throughout the play, Musa reveals his desire to assimilate to American culture; for him to be vocally marked and "Othered" to prevent total assimilation reminds Musa that he cannot ever achieve his

desire. In this case, a dialect coach or vocal professional considering this play may use these clues to conclude an accent or dialect is required for this play. Other clues of requirements such as differences in orthography written into the dialogue of the text present a different challenge for the dialect coach as these clues do not immediately lend themselves to ease of identification for the type of dialect that is required by the playwright. In these cases, careful consideration of the backgrounds and social statuses of the characters within the play and of the playwright is required to determine the need for dialects.

Other instances of dialect desires might not be as clear cut, and the dialect coach must carefully weigh the desires of the playwright with the potential to actively harm marginalized groups further. Often this decision becomes more difficult when the production in question is intended as comedy or satire. For example, in *Avenue Q* (2003), book writer Jeff Whitty writes the desire for dialect implicitly into the Christmas Eve character's grammar of her lines, "He a pervert. You no spending time with him" (14). Christmas Eve as a character is written to be a smart East Asian girlfriend of Brian, and she often laments that people cannot see or hear her brilliance due to her accent. In some ways, this pastiche East Asian accent, often drawn from stereotypical examples of accents in popular culture (e.g., Mickey Rooney in *Breakfast at Tiffany's*), reinforces audience expectations by once again tying the audio experience with Christmas Eve's character traits. Perhaps, Christmas Eve was written as a smart satire, but Jeff Whitty does not do anything with her character that implies subversion, and in fact draws upon Herman and Herman's exact examples in their *Foreign Dialects* book from the 1950s.

A dialect coach may see the attempts at satire and is left with a decision. The next question ought to be, is the comedy “punching up” or cleverly uplifting a historically marginalized group? Another way to examine this question; would a dialect coach or vocal professional, doing their due diligence and involving community input, be proud to present this character and dialect concept to a member of the community this character is trying to represent? Accent is often used as a cognitive shortcut to lead the audience to a conclusion about the character in question, and the dialect coach ought to examine every facet of this conclusion. Another option in this sticky instance is to investigate the implications of not using the desired accent. In this instance, if Christmas Eve used a General American Accent, or even location-specific New York accent, the audience would find themselves recreating the Kang and Rubin experiment of “reverse linguistic stereotyping,” where the reason Christmas Eve cannot find clients in “Sucks to be Me” is the audience’s *expectation* of her accent (442). Given the implications of either decision, the dialect coach has a lot to weigh and ought to be given enough power and respect to make the most ethical decision.

An unlikely ally of the voice professional ought to be the dramaturg, who together can determine deeply situated expertise that can guide the process of the production. Early incorporation and respect of the dialect or vocal coach helps to shape the production in the way that best combats racist, classist, and sexist linguistic stereotypes, an arm of the basic cultural competency responsibilities of the entire production team. The dialect coach, much like the dramaturg, can offer expertise about the linguistic lives of the characters in a way that shapes the overall meaning-making in production. For new works, the dialect coach’s number one question is, “what is this accent doing to enhance

or detract from meaning-making for the audience?” This production team integration ought to combine with individual access to actors with ample time to integrate the linguistic requirements for the role. Practically, in my own experiences, I am often asked to coach past the time for thoughtful pre-production integration and am often left with little time to work with actors on a desired accent, yet alone be able to discuss with the director about the motivations behind the desired accent.

4.1.1 A note on casting

A discussion of pre-production would not be complete without addressing the multi-faceted issue of casting and representation, which is a topic that has been addressed extensively in other venues of research and deserves its own dissertation’s worth of exploration. However, this guide would be incomplete without consideration of the effect of casting on dialect and accent coaching. Directors and dialect coaches have the responsibility to explicitly account for the power structures in society and the barriers of marginalized actors from entering the profession. In their quest for equitable representation, directors and dialect coaches must also be wary of fetishizing or tokenizing individuals as representatives of their race, ethnicity, gender, or socioeconomic status.

For dialect coaches to stand the best chance of ethically doing the work they are hired to do, the explicit responsibility of representation of actors onstage ought to rest on the artistic directors and director’s shoulders. A dialect coach’s job of ethical representation and training of lived linguistic experiences stand the greatest chance of succeeding if ethical approaches to casting are employed from the very outset of the production process. Cultural competency is the responsibility of all on the production

team, with the director's immediate responsibility to consider the implications of the types of bodies onstage that they choose to represent the characters in their productions. Acknowledging the racist power structures that exist in theatrical creation means that in production, roles that are created specifically for Black, Indigenous, and People of Color ought to be filled with individuals who fit the demographics of the character description as closely as possible or through thoughtful consideration of coalition casting (Herrera 33). A director ought to also consider casting in the opposite tenor, extending roles historically created for white actors to Black, Indigenous, and actors of color. With this foundation, dialect coaches can approach training these actors with as much specificity and care as in the casting process, tailoring the approach for each actor. Part of this process is creating space in training to explicitly acknowledge how society treats the lived linguistic experiences of not only the target dialect, but the lived linguistic experiences of the actor, as well. I will detail honoring such lived linguistic experiences in the following section about approaching the encounter with actors in training.

Pre-production work is vital for the success of the work that the dialect coach must do with the actors in the next phase of production. The following ethical questions and guidelines for pre-production work are not limited to only the dialect or voice coach. These questions can become the responsibility of the director, dramaturg and ultimately the artistic director, especially if the production company has an eye towards equity and representation. Much work has been done for authentic representation on stage and the time is now to further extend that work to linguistic representation, which is an area that has been neglected. Production team members can work together towards ethical representation onstage while making room for the expansive possibilities exceeding the

use of authenticity in production. Dialects and voice onstage can start from grounded research in the lived linguistic experiences of speakers, but used thoughtfully against expectation, can lead to new ways to make meaning in the stories we present onstage. The work of the dialect coach can also fall in line with new and important positions that are being created for production that ultimately respect the bodily autonomy and lived experience of the actor, in a similar vein to how Theatrical Intimacy Educators approach completing work while respecting the bodily autonomy of actors and producers alike (Pace). Like simulated intimate acts, this approach to voice and dialect coaching aims to respect the actor and provide vocabulary to provide a simulation of the dialect or accent that respects all parties. Cultural competency and equity ought to be on the minds of every practitioners in production. Both professions build upon an exchange between the student or actor and their coach, in a configuration that can be physically, emotionally, and psychologically intimate and can be susceptible to abuse of power dynamics.

4.1.2 Questions to ask a play (and production team)

<p>What are the dialect <i>requirements</i> of the play/playwright?</p>	<ul style="list-style-type: none"> ○ What character or personality traits do the accents point towards? Is the accent supposed to “stand in” for any character trait? ○ What does the use of a particular dialect reveal about the dramaturgical life of the characters? ○ What dramaturgically does a dialect contribute to the play? ○ How does the positionality or background of the playwright interact with the desired dialects or accents in the script?
<p>What stereotypes/expectations (racist, classist, sexist) are the dialects participating/perpetuating?</p>	<ul style="list-style-type: none"> ○ In what ways is dialect <i>enforcing</i> these stereotypes/expectations? ○ In what ways is dialect use <i>subverting</i> these stereotypes/expectations? (Are you sure?)

What are the dialect <i>requirements</i> of the production company and director?	<ul style="list-style-type: none"> ○ Are there expectations of a unifying dialect?
What are the dialect <i>desires</i> of the play/playwright?	<ul style="list-style-type: none"> ○ Are there ways to include dialect/accent in a manner that <i>subverts</i> stereotype/expectations of the audience? How does this enhance the production? ○ How does dialect work connect into community outreach and audience education?

4.2. The heart of the work: one-on-one with the actor

In the time between completing pre-production activities with the production team, answering the above questions, and getting a crystal-clear picture of the desires and ethical responsibilities of the vocal professional, the next phase is extensive research on the dialect or accent, working with real speakers of the desired dialect. This type of research that a vocal coach must do includes specific research into the different intersecting identities of the characters (e.g., socioeconomic status, gender, race/ethnicity) with the aim to become as specific as possible when constructing the lived linguistic realities of these characters. Sources for academic research can vary from specific linguistic descriptions of languages (in the case of second language speakers), linguistic atlases (useful for pinpointing direct regional dialects), to dialect materials⁴⁰ that already exist for performance. An important source can be direct recordings from speakers, which are available oftentimes on YouTube, podcasts, and other archives of material. One overlooked source of language attitudes towards particular accents or dialects can be popular linguistics videos and articles, and general folk linguistic articles that include non-experts' opinions about the accent in question. These types of resources can point an

⁴⁰ See the bibliography of *The Dialect Handbook* (2003) by Ginny Kopf for a comprehensive collection of dialect instructional resources collected over the twentieth century. A future project of mine will be to collect resources created since this guide appeared.

actor towards how their character might feel about how they sound, which gives actors access to more nuanced choices about the level of character self-confidence in any given scene and might affect their choices about how a character ought to sound when they are feeling an extreme emotion. For example, an actor might make creative linguistic choices maintaining control over their character's more middle-class accent that they acquired when they were older when their character is being teased about their lower-class upbringing. I have found that compiling materials about language attitudes towards speakers of the target accent or dialect can create access to a new form of dialogue that feeds into an actor's autonomy over their sound on stage.

A voice professional who aims to be an anti-racist co-conspirator ought to consider the level of formal linguistic education actors might have along with time constraints of training sessions with actors and cater the material accordingly. The voice professional can compile a resource similar to dramaturgical research that includes the reasons behind the accent or dialect that was selected, and the details of the target accent. Often, I include audio and video materials for actors as they desire an audio example from which to work, which activates the perceptual system, but can interfere with production of target sounds in unexpected ways (Kato and Baese-Berk 7). Regarding sanitized or theatrical examples of popular dialects, I will sometimes provide actors with materials from established dialect coaches (e.g., Blumenfeld, Singer, or even Knight and Thompson) with explicit discussion of the constructed and standardized nature of these dialects. I do not use pre-written dialects from countries outside of the United States and Great Britain. I will also always provide audio samples of real speakers. One popular dialect that I often use with pre-created material is Received pronunciation, a popular

dialect that often represents British characters on stages in the United States, which consists of a series of worksheets and information adapted from colleague Dr. Tricia Rodley and Robert Blumenfeld (29). With careful framing of how to approach standardized dialects, I include a forward to the materials that acknowledges power structures and a discussion of how these accents became standardized in the first place. I invite discussions with directors and actors alike to approach the desire for dialect in the first place.

Another approach to creating dialect work is to create a dialect or accent from scratch with which actors may work. Dialect creation has historically been a tool for voice professionals to impose their dialect ideologies upon speakers who enter their elocution classrooms and rehearsal rooms. There is room, however, for this type of dialect creation to be used in performance, especially when the characters' backgrounds are fictionalized to the point of being from made-up locations. In film and television, extreme examples of dialect creation include creating entire separate languages. These constructed languages figure heavily in science fiction, from Mark Okrand's creation of Klingon from Star Trek to high fantasy with various languages in J.R.R. Tolkien's *Lord of The Rings* series, and languages created for *Game of Thrones*. (Klingon Language Institute). These constructed languages often borrow their phonology and sound systems from languages that exist around the world. In a similar vein, constructing a dialect or accent would again consider the questions posed in the prior section. When real-world sources are selected for these types of characters, utmost care must be taken in order to ethically represent these fictionalized linguistic lives so that they may not reflect linguistic stereotypes that exist in the real world.

As an example of my own work in this area, I constructed the dialect for two characters Alta and Restin in *The Language Archive* by Julia Cho for my undergraduate student production in October 2013 for University of New Mexico. Alta and Restin are the last two speakers of a language called Ell-o-wa, a language that the main character George has a vested interest in preserving (Cho 20). One of the few linguistic cues the play offers as to how these two speakers sound is in a monologue from Restin wherein he describes a “golden fleek” (38). This demonstrates an important sound substitution at the end of words—and a few lines of dialogue in this playwright created language between Alta and Restin. The production team determined that Restin and Alta are from some area where Slavic speakers live, so I, wearing many hats as part of a student production of dramaturg, linguist, and director then turned my attention to the specific phonetic and phonemic categories of several major Slavic languages (e.g., Polish, Moscow Russian-accented English, and Czech) and selected several target consonant and vowel sound substitutions that would be targets for the actors. Part of the justification for targeting this part of the world was an explicit discussion of power dynamics that overlap with the bodies of the actors who would be featured speaking this dialect. While these dialects do not carry the overt prestige of western European dialects, they did carry a relatively perceived neutral ethnic prestige in which speakers from other racialized parts of the world do not participate.

In the play, Alta and Restin would sound like they were hailing from a foreign country, but the origin would be hard to pin down for the average audience member. This was important, because the two actors who were cast as these characters were of different ethnicities—one actor identified as white and the other actor self-identified as non-white

Hispanic. I wanted an accent that audience members would mark as “foreign” without necessarily marking one actor with more negative stereotypes. What resulted was an accent that gestured towards a certain part of the world but could live in the mouths of the two fictional Ellowan characters believably. Subsequent discussions with actors showed where each sound decision was made and how they related to real-world linguistic experiences. We could therefore approximate what could be considered the homeland of Restin and Alta without the need to fall back upon stereotypes of sounding stereotypically Eastern European. Did the audience catch these nuances? The answer, according to the results of my linguistic research in chapter 3, is mostly like not, but the important aspects of training and respect for the lived linguistic realities of speakers of minority languages remained a fruitful means of discussion and character creation.

4.2.1 Finally meeting the actor

The initial meeting between coach and actor is a crucial moment that sets the stage and approaches to the work. It is in this initial meeting that misconceptions and unconsciously held biases are explored in a safe space one-on-one with the actor and discussed, along with the practical work of actors becoming intimately acquainted with their vocal apparatus. This is the first opportunity to create a space of mutual respect that acknowledges the power structures at play. The first lesson I impart during this meeting is that we all carry within us a voice that has been shaped and created by the location we grew up in, those who we called peers, and every linguistic interaction since acquiring language from a young age. The second lesson, directly from the lessons the study of linguistics imparts, is we also live in a society where everyone carries with them standard language ideologies for language varieties, and these ideologies govern which varieties of

language are more societally acceptable than others. Expanding approaches to dialect and voice work past the individual to examine the socially built structures that govern language usage connects nicely to other work in character building and audience reception. Instead of approaching the individual as a psychological container unto themselves, this approach to dialect work foregrounds the idea that language is a tool for community use, and therefore is necessarily shaped by the ecology in which a community finds itself. Again, I invoke Amy Cook's question about character building, "what does it mean to build characters from the ecosystem up, rather than a more psychologically focused method of character assessment?" (117). I take the same approach to dialect coaching and center language usage as an integrated part of community and ecology over and above any individual's psychological experience.

After discussing general language societal language attitudes with the actor, I begin by introducing the ideas of raciolinguistics, and how theatre is set up to cater towards an assumed white listener. In this discussion, the idea that accents and dialects are not inherently connected to character traits is introduced, but that society has an overwhelming tendency to assign character traits to how people sound when they speak. The framework we use will incorporate societal expectations (i.e., the average audience member) into the work while explicitly working against assigning character traits to how the actor will sound. This discussion is accompanied by any stereotypes we might be pushing back against with the production. Part of this work is a series of questions that we begin as a discussion point (see "4.2.2 Questions to ask the actor") and unpack how the actor thinks they sound before creating work on a new dialect. We then turn to discuss intelligibility as an attribute and feature of the listener. I have not had the pleasure

of sharing the results of my research in dialect coaching (partially because the pandemic has all but eliminated production for the time being), but I look forward to the day of showing actors how listeners react to performed accents and demonstrating that listeners might be more generous in their interpretations of the accents they will hear onstage. We then discuss the term authenticity and all the implications around the word, including its strange and uneasy relationship with how theatre uses representation to create meaning in the audience. This opens discussion for how we approach dialects for production; we will be targeting a few key sounds and rhythms to create a dialect that the audience will read as “authentic” while honoring the source of the dialect. What follows is an information session about the historical, ethnic, and socioeconomic circumstances that surround the target dialect (even if the dialect is constructed, such as Received Pronunciation). This discussion is often the bulk of the initial session, with a small introduction to language notation and an initial approach to the work.

After conversations about the theoretical framework and a small warm up, I introduce the actor to the idea of linguistic notation and offer a few brief exercises in thinking about how words sound (phonemes) as opposed to how they are spelled (orthography). Often, I incorporate a small introduction to the International Phonetic Alphabet with a historic framing about the origin of this alphabet. We begin to connect these symbols and how they are arranged on paper with the lived reality of the actor’s vocal apparatus. We then calibrate how we approach the work by exploring how the actor responds to commands to “make this sound harder/softer/wider” for future instruction and physiologically attune them to what their vocal apparatus is doing when they make certain linguistic sounds (Colaianni). At this point in the instruction, I also introduce a

sense of serious playfulness to the work, emphasizing the fact that it took six to seven years for the actor to master their first sound system (Statler, Heracleous and Jacobs). This type of work requires the playfulness of toddlers exploring their sound systems combined with the targeted work that an actor must do to achieve desired linguistic results. Opening the session for play helps to relax the actor and remind them that acquiring new skills will inevitably include making mistakes in the process. After initial discussion and exploration, I give actors training sheets that include practice of the IPA, target sounds of the dialect, and links to audio sources for practice.

Intense one-on-one sessions vary depending on the availability and schedule of the production and actors. Subsequent sessions include questions and answer sessions from initial discussion of framework, specific questions about target sounds and target lines in the production, and targeted practice with notes. Like Bonnie Raphael, I aspire to give notes, “always stated in vocal rather than in acting terms” (168). I also give several physiological options for actors to access the sound they desire. When I provide sociolinguistic background to actors, I remind them that they may have sociolinguistic tools to make decisions about their dialect and accent usage but that they ultimately have autonomy to make acting choices with their dialects onstage. I will also participate in rehearsals at least once a week, and more often near the end of rehearsal when there are full runs of the production to give specific notes. While speakers are often variable⁴¹ in their speech, I am listening for moments and target sounds that are attempted in dialect in production that do not quite make their targets. I am also gathering feedback from the

⁴¹ For example, Huspeck demonstrates that almost every American English speaker is variable in their pronunciation of -ing at the end of words like “running” (152).

director, stage manager and any assistant stage managers for their overall impression to ensure actors sound like they are of the same linguistic world, even if individual actors have different accent or dialect targets.

4.2.2 Questions to ask an actor

<p>Initial questions to teach accent awareness and discuss lived linguistic experience with the actor before even approaching the target dialect:</p>	<ul style="list-style-type: none"> ○ Do you have an accent? ○ How do you feel when asked if you have an accent? ○ Have you been told, or do you think that you have an accent? ○ Where did you grow up? ○ Do the people you grew up with have an accent? ○ When was the last time someone pointed out to you that you had an accent? ○ How did that feel? Was it a positive experience?
<p>Approaches to language notation, accompanied by an exercise that explores sounds in the mouth as they are arranged on the International Phonetic Alphabet:</p>	<ul style="list-style-type: none"> ○ Have you noticed that spelling does not always match pronunciation? Which sounds do you notice this discrepancy the most? ○ How will you consistently notate sound changes in your practice?
<p>Finally, we combine language attitudes with practice and their attitudes towards their character's accent or dialect. I ask the actor as their character the initial questions from above.</p>	

4.3 Audience outreach: Working within the community with the dramaturg(e)

Best practices for audience and community outreach are to build long-term, lasting relationships with audiences and the community far and above single productions, or even single seasons focused on marginalized groups. Like representation issues in casting, much research has been conducted in building community more generally with

groups that have historically been excluded from the practice of theatre (Lacko 21). The focus of this section is on particular roles that dialect coaches and voice professionals may fill over the course of production that lead to audience and community interfacing. These suggestions call for a tighter relationship between the dialect coach and dramaturg and whatever apparatus the production company has set up for community interface. In the absence of these roles, the dialect or vocal coach may fulfill these community obligations themselves.

Community outreach and coalition building may start within the pre-production when a dialect coach can partner with consultants to provide insights and audio material for dialect construction. This is an important part of the research process; access to audio for reference is an important part of the training available to actors. Access to collaborators also means the possibility of audio references that are custom catered for the production. For example, in the 2019 production of *Pilgrims Musa and Sheri in the New World*, I was able to create custom audio tracks that included pronunciation of the Arabic that was used in the show, thanks to access to a willing Arabic speaker⁴². Not only were the exact phrases available, but I was able to work with my consultant to create slower instructional tracks that included pronunciation one phoneme at a time. In this case, the production was a community production, and the consultant was not monetary compensated, but they were invited to a preview night and thanked in the program. Ideally, dialect coaches in a production ought to be allocated funding from the production budget to compensate consultants or pay to access other audio materials that are not readily available.

⁴²Deepest thanks to Ryan Sayegh as consultant for this production.

Part of the dialect coach's community outreach can be creating opportunities to educate the audience about their own linguistic biases as part of the dramaturgical process. The careful nuanced work that happens with actors must somehow extend to audience members, since it is within each audience member that meaning-making and perceptual processes happen. Usual venues for audience outreach include materials in the programs, interpretative displays in the lobby, and interactive events such as talkbacks and meet-the-artist type events. When dialect is considered core to the story or character, the dialect coach ought to be afforded the opportunity to provide materials and feedback on the dramaturg's displays and notes for the program. Even short descriptions of accent or dialect choice for a given production can help give audience members insight into the complicated processes of how language attitudes and accent perception govern audience members' everyday lives. Interpretative displays can include discussions of accents and actor training. Another source of audience interaction can take the form of talkbacks with linguistic experts in the field. For the 2013 production of *The Language Archive* where the play was about a linguist performing field work on a language with only two speakers remaining, I arranged an evening of talkbacks with two Linguistics PhD students who conduct research similar to the main character of the play. This accomplished two objectives—one was connecting audiences to a different way of conceptualizing language and language use, and another was introducing an entirely different audience to theatrical production, as the house was packed with linguistics professors and students who had never received the explicit invitation to a theatrical production by those who would be featured in the programming. Talk backs can be structured in a similar fashion, with options for guests that include leaders from the speech community featured in the show,

linguistic experts, or even the dialect coach themselves. This real time interfacing can be incredibly valuable to all involved participants.

However, due to historical mistrust and exclusion from theatre creation, some members of communities may resist inclusion in production. Voice professionals, even in their good intentions and research plus community outreach, ought to be prepared for a non-response, or even negative response from the communities they research and ultimately bear responsibility for representation in production. An example of this rejection can be as compassionate as how Brian Andrew Cheslik describes in his article “ASL and Theatre: Here’s what not to do” where he opines the best intentions from directors who are trying to incorporate his native language of American Sign Language (because not all lived linguistic realities are auditory) into theatrical productions. Cheslik’s first request is casting Deaf characters with Deaf actors, to bring authenticity, representation, and work to these actors. Cheslik follows with, “While I appreciate that you want to share your student’s hard work learning sign language with my students, we are not interested in coming to see new signers butcher our language” (Cheslik). This point is an excellent reminder that even the most carefully researched, rehearsed, and practiced dialect for production will not in any meaningful way truly capture the depth of a person’s lived linguistic experience. This sentiment matters because it is directly connected to how theatres develop new audiences; without the ability to experience linguistic representation onstage, communities historically exploited for their culture will not trust these endeavors. Speakers of the group that will be featured have the right to not experience poor simulated approaches to their lives, even with the most careful approach from pre-production through community engagement.

While a still a socially constructed term, authenticity still has a large role to play in community representation and outreach. Cheslik cites “authenticity” as the ultimate reason for his hesitance for engaging hearing productions that feature ASL but also often points to poor planning and incorporation at all stages of production. “There has to be a reason why you have decided to do the show in ASL and English. Basically, there should be Deaf actors in the show. Do not just do a bilingual show with ASL and English without there being a Deaf performer involved” (Cheslik). While authenticity can be one avenue for community coalitions, Cheslik leaves the door open for Herrera’s coalitional casting approach to producing theatre with marginalized culture. Including the marginalized group about whom the piece of theatre is about (at various steps in the production process) is the ethical path for production.

While included at the end of this discussion, to approach community coalition building and outreach at the end of the process is a mistake for production. These relationships need to build over time and in coalition with one another. These approaches help to avoid an exploitative relationship between producing company and community around which these plays and productions revolve. Such exploitative relationships directly contribute to the harm and stereotype creation that has governed this profession since its inception with elocution teachers. There has to be a balancing act between the tensions of authentic *representation* and authentic *recreation* on stage.

4.3.1 Questions to Ask a Dramaturg(e)/Community Outreach

How soon can we begin building community outreach?	<ul style="list-style-type: none"> ○ When and how are we tracking community interaction and outreach?
What clubs or organizations can we reach out to collaborate?	<ul style="list-style-type: none"> ○ What is the budget to compensate community consultants for their contributions to dialect? ○ How else can we support a

	community or organization that historically has been omitted from theatre production?
How can we search for actors and consultants who are of the same background (overlap with casting considerations of the production team)?	<ul style="list-style-type: none"> ○ What is our interaction with the casting process?
What do the dramaturgy materials look like?	<ul style="list-style-type: none"> ○ Will there be interface via program/lobby display? ○ Will there be talk backs or other interactive elements? ○ How do we create materials and community outreach that is as accessible as possible?
After the production, how can we critically engage with new audience members?	<ul style="list-style-type: none"> ○ What does continued support and collaboration look like for communities that we have already engaged?

5. Challenges that Remain, Where Do We Go from Here?

This ecological approach to building the profession of dialect coaching complements existing anti-racist and decolonialization efforts in the theatre. I struggle against the prevailing tide of voice training as a singular white theatre maker, and I look to practitioners and scholars in actor training for answers. One such practitioner is Nicole Brewer, who is an advocate for decolonializing acting training. In her article “Training with a Difference,” Brewer advocates against equity, diversity, and inclusion in that the policies that are often created are not strong enough to address the deepest underlying issues. She, like myself, argues for an anti-racist framework in approaching theatre production and training, "It’s a problem to not say racism. You have to turn to face racism. Lacking clarity around an anti-racist policy allows it to persist, despite your intention” (Brewer). Because dialect and voice training has not addressed its own explicitly racist and white supremacist foundations, this profession continues to

reproduce and promote harmful practices, and I have a difficult time with my individual activism through one-on-one training with actors in project-based approaches to voice.

Dismantling biases using a discipline that has been explicitly built upon standard language which is driven by white-dominated structures does not necessarily inspire confidence in the success of the venture. After all, as Audre Lorde has said, “the master’s tools will never dismantle the master’s house” and therefore this enterprise may be doomed from the start (27). Lorde, however, qualifies this statement, “They may allow us temporarily to beat him at his own game, but they will never enable us to bring about genuine change. And this fact is only threatening to those women who still define the master’s house as their only source of support” (27). I must wrestle with the perceived threat of the master’s house—the institution of theatre in which I have been trained—as an ineffective at best and violent at worst approach to performance and theatrical creation. The methods by which these metaphorical tools in voice training have been employed uphold the structures of white supremacy and linguistic racism. By conceiving of these tools as just tools, we can begin to conceive of voice training in a way that can combat these structures. We must walk the knife’s edge of using voice training techniques and materials established as part of this structure and acknowledge their histories, while allowing room to do the work of linguistic transformation as production calls for it. I must examine the privilege that my position as dialect coach comes with an assumed institutional authority on what is said can be deemed acceptable to appear on the stage. As discussed in the second thread, this reverence for authority reflects the standard language practices that have been in place since early elocutionists at the turn of the twentieth century (Knight, 2000). In these past years, I have come to grow uncomfortable

with replicating this societal structure while still serving as a source of confidence building and joy for actors. Thus, my own deepest desire is to create a system that empowers actors to become their own linguistic expert; I wish to become a kind of “guide on the side” that helps actors understand that the language they are already producing contains years of embodied knowledge of environment, and that knowledge is a powerful lens in which they may enter the lived linguistic experiences of others.

Yet another struggle that I come up against dialect coaching is usually project-based, and thus is part of a larger production team. A dialect coach is given an accent in a particular production and the directive to make an actor sound clear and accented in the performance. It is the coach’s responsibility to interpret “clarity” from many different production team members (most importantly the director) and balance the respect for the lived linguistic lives of actors on stage. Thus, a large portion of my approach is determining the best ways to deal with that balance and offer tools for producers and directors guided by the research in this dissertation. This type of work challenges the preconceived notion of accent as character trait by asking every participant in production from the actor to the producer about the role of trained accent or dialect in performance onstage. After that question is addressed, I offer other steps in training and production that continue to keep the fundamental idea of accent and dialect not as superfluous accessory to be added to an actor’s repertoire, but as a deeply embodied part of the dramaturgical and practical work of inhabiting a character. The practice of using accent and dialect will remain a part of performance and entertainment; I offer my perspective as a white theatre maker what equates to harm reduction to portray accent in a way that challenges the prevailing attitudes about who gets to sound like whom.

Often, when presenting work outside of specific dialects and in voice training in general, my goal is not to only introduce the new dialect, but again to empower participants to fully accept their own embodied voice, by exploring a brief text that is meaningful to them outside of the project at hand. This may be reminiscent of other approaches by voice practitioners like in Kristin Linklater's *Freeing the Natural Voice*) However, in my work, I always introduce to participants the explicit societal structures that reinforce standard language biases, which grounds this practice into real-world pressures we all face as language users. As discussed in the introduction, other practitioners like most notably Patsy Rodenburg in her *Right to Speak* also address these issues implicitly, but my argument is that work of this nature *must* begin with this framework to even attempt a respectful representation of voices onstage. Since the publication of her book, Rodenburg has acknowledged that linguistic discrimination is its own form of violence towards marginalized actors and has since adjusted how she approaches Received Pronunciation in her own classroom.

What I believe is that if you teach an accent that has painful historic resonances you must teach that accent with grace and sensitivity. You must also understand that the student has a right not to master or even speak that accent without the fear of failing a course. Of course, not speaking certain accents will affect an actor's potential casting—Received Pronunciation is still very important for British actors' careers—and that fact has to be very clearly communicated to the student. Most of my students, who have emotional problems with Received Pronunciation, when given the above option and having their pain honored, do learn and own Received Pronunciation. (qtd. in Espinosa 82)

Actors and students of voice deserve this type of sensitivity to individual experiences with sociolinguistic or raciolinguistic discrimination at all stages of their training and professional lives. At the heart of the work, the actor's autonomy over their own instrument must be of paramount importance.

These structural patterns of white-dominated structures in theatre still persist, yet there are examples of contemporary practitioners who are resisting the dominant ideas and attitudes of how voice training has been established. Teachers who come from marginalized communities themselves are fighting for this type of freedom to acknowledge and work with students in this fashion. However, voice teachers are only one part of the theatrical and film production and their roles can be limited. For example, when Micha Espinosa worked with one well-meaning director, the director suggested,

To record the one African-American student I was working with at the beginning of the semester, and he expected that by the end of the semester this student would speak in a Standard American dialect. When I rejected his advice, the director felt that I was not teaching this student the skills he needed. He had no knowledge of the emotional carnage that following his advice would have inflicted (81).

This type of behavior is driven by the raciolinguistic device of the idealized white listener, and it is obvious that standard language ideology can and does carry through other production roles in professional settings as well as higher education. To ethically practice voice and dialect, voice trainers ought to have not only the responsibility of saying "no" to other production team members, but the authority to do so as well. My ethical responsibility as a white theatre maker to ensure this fight for equity and equal

linguistic representation does not fall upon the shoulder of my marginalized colleagues; equity and cultural competency is everyone's responsibility.

Another challenge that the voice profession will face in the coming years is the question of licensing and qualification, especially given the growing trend of voice coaches advertising their services on the internet. VASTA, the professional organization for voice trainers, is not a licensing body or board that regulates the industry or proliferation of online dialect and accent coaches. Licensing does however often appear in the form of certifications from individual voice trainers and their schools. For example, the *Linklater Center* in New York City and Orkney, Scotland offers certification to students in higher education or participants who can pay the high fees associated with this official voice training program. This tension of who has access or the right to study their voice lies at the heart of my own criticism of the material circumstances that have arisen around this industry and this model may be changing. The COVID-19 pandemic has forced established voice institutions to inhabit digital space in an unprecedented way. This increases access to teaching and classes, but often presents issues for work that often requires in-person intimate contact. The internet has facilitated both a rise in access to this type of work and an influx of digital influencer-styled voice and dialect coaches.

The larger question remains is how to balance the gate-keeping privilege of VASTA with the unregulated generation of online voice and dialect coaching that borrows its business model from influencer-style online promotion. Inherent in both models of access to this discipline is still the prevalence of implicit and explicit biases that contribute to the enforcement of negative Linguistic stereotypes seen through entertainment. The desire to quantify qualifications into easy-to-read but hard-to-attain

lines in one's curriculum vitae or resumé has existed for a long time in theatre production at large. One solution can be garnered from the emerging field of intimacy coordination, where leading intimacy coordinators have cautioned against the idea of a certification system. Chelsea Pace writes in a newsletter from June 2021, "The existence of 'certification' leverages systems of power that promote inequality, exclusion, and the dynamics of deeply problematic master-teacher models to capitalize, financially or otherwise, on gatekeeping access to knowledge and opportunity" (2). Access to certification also disproportionately affects Black, Indigenous and People of Color who do not have the same resources and can be overlooked for the knowledge and skillset they have. This issue of certification and access will continue to be a problem into the future of this profession but will necessarily reflect how willing this field wants to change to correct historic harms and exclusions. Without a reimagining of the very structure of this profession, no progress can be made towards the ethical responsibility of harm reduction and inclusion.

At the heart of this work both in ethical approaches to voice training is the idea that every person communicates differently due to their lived experiences, and every human deserves dignity and respect for how they sound. Everyone's voice

...comes from where we come from, but then every single one of us gets influenced in ways that are both conscious and unconscious through our entire life: who we dated, what we liked to watch when we were younger, a formative iconic figure for us during the era that we were growing up, what our age is, who we wanted to hang out with, where else we've lived in the world. There is a conscious and unconscious way in which our voice tells a story of who we are

(Bay, qtd. in Feller).

The future of this voice profession is navigating a way that honors every language user's unique experience, recognizes the power and rich resources of that unique experience, and makes room for serious and respectful linguistic play. A piece of examining socially constructed values is to critically interface with audience expectations to analyze how these social constructions arise through empirical inquiry. By recognizing the harms of historic voice practices that erased individual experiences and endeavored to replace these experiences with a stereotypical depiction, we can begin to correct and mitigate harm by recognizing voice practice as a tool where we can ascribe our own social meaning in pursuit of an ethical practice of voice.

APPENDIX
STIMULI MATERIALS USED

Hearing In Noise Test (HINT) sentences (Nilsson, Soli and Sullivan) used as stimuli in Experiments 1 and 2 from Chapter 3.

HINT 1

1. A boy fell from the window.
2. The wife helped her husband.
3. Big dogs can be dangerous.
4. The shoes were very dirty.
5. The player lost a shoe.
6. Somebody stole the money.
7. The fire was very hot.
8. She's drinking from her own cup.
9. The picture came from a book.
10. The car is going too fast.
11. The paint dripped on the ground.
12. The towel fell on the floor.
13. The family likes fish.
14. The bananas are too ripe.
15. He grew lots of vegetables.
16. She argues with her sister.
17. The kitchen window was clean.
18. He hung up his raincoat.
19. The mailman brought a letter.
20. The mother heard the baby.
21. She found her purse in the trash.
22. The table has three legs.
23. The children waved at the train.
24. Her coat is on a chair.
25. The girl is fixing her dress.
26. It's time to go to bed.
27. Mother read the instructions.
28. The dog is eating some meat.
29. Father forgot the bread.
30. The road goes up a hill.
31. The painter uses a brush.
32. The family bought a house.
33. Swimmers can hold their breath.
34. She cut the steak with her knife.
35. They're pushing an old car.
36. The food is expensive.
37. The children are walking home.
38. They had two empty bottles.
39. Milk comes in a carton.
40. The dog sleeps in a basket.
41. The house had nine bedrooms.
42. They're shopping for school clothes.
43. They're playing in the park.
44. Rain is good for trees.
45. They sat on a wooden bench.
46. The child drank some fresh milk.
47. The baby slept all night.
48. The salt shaker is empty.
49. The policeman knows the way.
50. The buckets fill up quickly.
51. The boy is running away.
52. A towel is near the sink.
53. Flowers can grow in the pot.
54. He's skating with his friend.
55. The janitor swept the floor.
56. The lady washed the shirt.
57. She took off her fur coat.
58. The match boxes are empty.
59. The man is painting a sign.
60. The dog came home at last.

HINT 2

1. They heard a funny noise.
2. They found his brother hiding.
3. The dog played with a stick.
4. The book tells a story.
5. The matches are on a shelf.
6. The milk was by the front door.
7. The broom was in the corner.
8. The new road is on the map.
9. She lost her credit card.
10. The team is playing well.
11. The boy did a handstand.
12. They took some food outside.
13. The young people are dancing.
14. They waited for an hour.
15. The shirts are in the closet.
16. They watched the scary movie.
17. The milk is in a pitcher.
18. The truck drove up the road.
19. The tall man tied his shoes.
20. A letter fell on the floor.
21. The ball bounced very high.
22. Mother cut the birthday cake.
23. The football game is over.
24. She stood near the window.
25. The kitchen clock was wrong.
26. The children helped their teacher.
27. They carried some shopping bags.
28. Someone is crossing the road.
29. She uses her spoon to eat.
30. The cat lay on the bed.
31. They're running past the house.
32. He's washing his face with soap.
33. The dog is chasing the cat.
34. The milkman drives a small truck.
35. The bus leaves before the train.
36. The baby has blue eyes.
37. The bag fell off the shelf.
38. They are coming for dinner.
39. They wanted some potatoes.
40. They knocked on the window.
41. School got out early today.
42. The football hit the goalpost.
43. The boy ran away from school.
44. Sugar is very sweet.
45. The two children are laughing.
46. The firetruck is coming.
47. Mother got a sauce pan.
48. The baby wants his bottle.
49. The ball broke the window.
50. There was a bad train wreck.
51. The waiter brought the cream.
52. The teapot is very hot.
53. The apple pie is good.
54. The jelly jar was full.
55. The girl is washing her hair.
56. The girl played with the baby.
57. The cow is milked every day.
58. They called an ambulance.
59. They are drinking coffee.
60. He climbed up the ladder.

REFERENCES CITED

- “About.” *Joy of Phonetics*, n.d. <https://www.joyofphonetics.com>
- “About Phonetic Pillows.” *Joy of Phonetics*, n.d. <https://www.joyofphonetics.com/phonetic-pillows-2/phonetic-pillows/>.
- “About the Work.” *Knight-Thompson Speechwork*, n.d. <https://ktspeechwork.org/about-the-work/>
- Agheysi, Rebecca, and Joshua A. Fishman. “Language Attitude Studies: A Brief Survey of Methodological Approaches.” *Anthropological Linguistics*, vol. 12, no. 5, 1970, pp. 137–57.
- Alim, H. Samy, and Geneva Smitherman. *Articulate while Black: Barack Obama, language, and race in the US*. Oxford University Press, 2012.
- Anzaldúa, Gloria. “How To Tame a Wild Tongue.” *Borderlands / La Frontera: The New Mestiza*, Aunt Lute Books, 1999.
- Austin, John Langshaw. *How to do things with words*. Oxford university press, 1975.
- Bakanic, Von. *Prejudice: Attitudes about race, class, and gender*. Prentice Hall, 2009.
- Bartoskova, Michaela. “The Role of the Psoas Major Muscle in Speaking and Singing.” *Voice and Speech Review*, vol. 15, no. 2, Mar. 2021, pp. 1–11.
- Bell, Brian, and Sam Hunter. “How U.S. States Could Fund Repertory Resident Theatres.” *AMERICAN THEATRE*, 21 June 2021, www.americantheatre.org/2021/06/21/how-u-s-states-could-fund-repertory-resident-theatres/.
- Benedetti, Jean. *Stanislavski: An introduction*. Taylor & Francis, 2004.
- Bennett, Susan. *Theatre Audiences: A theory of production and reception*. Psychology Press, 1997.
- Berkely, Anne. “Changing Views of Knowledge and the Struggle for Undergraduate Theatre Curriculum, 1900-1980.” Fliotsos and Medford. 7-30.
- Blumenfeld, Robert. *Accents: A manual for actors- revised and expanded edition*. Limelight Editions, 2002.
- Borrie, Stephanie A., Tyson S. Barrett, and Sarah E. Yoho. "Autoscore: An open-source automated tool for scoring listener perception of speech." *The Journal of the Acoustical Society of America* 145.1 (2019): 392-399.

- Bradlow, Ann R., and Tessa Bent. "Perceptual adaptation to non-native speech." *Cognition* 106.2 (2008): 707-729.
- Bradlow, Ann R., Midam Kim, and Michael Blasingame. "Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate." *The Journal of the Acoustical Society of America* 141.2 (2017): 886-899.
- Brewer, Nicole. "Training With a Difference." *AMERICAN THEATRE*, 4 Jan. 2018, www.americantheatre.org/2018/01/04/training-with-a-difference/.
- Brown, Bruce L., Howard Giles, and Jitendra N. Thakerar. "Speaker evaluations as a function of speech rate, accent and context." *Language & Communication* (1985).
- Brown, Stan. "Column the cultural voice." *Voice and Speech Review* 1.1 (2000): 17-18.
- Brunstetter, Bekah. *The Cake*. Samuel French, 2019.
- Calta, Louis "26 Stage Troupes Form League to Bargain with Actors Equity" *New York Times*, 4 April 1966, pg. 26.
- Catford, John Cunnison. *A practical introduction to phonetics*. Oxford: Clarendon Press, 1988.
- Cho, Julia. *The Language Archive*. Dramatists Play Service, 2012.
- Colaiani, Louis. *The Joy of Phonetics and Accents*. Joy Press, 1995.
- Cook, Amy. *Building Character: The art and science of casting*. University of Michigan Press, 2018.
- Cotter, Kelley. "Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram." *New Media & Society* 21.4 (2019): 895-913.
- Dal Vera, Rocco, and Voice and Speech Trainers Association. *Standard Speech: Essays on voice and speech*. Applause Books, 2000.
- Derwing, Tracey M., and Murray J. Munro. "ACCENT, INTELLIGIBILITY, AND COMPREHENSIBILITY: Evidence from Four L1s." *Studies in Second Language Acquisition*, vol. 19, no. 1, Mar. 1997, pp. 1-16.
- DeWitt, Marguerite E. *Euphon English in America*. EP Dutton & Company, 1924.
- Diderot, Denis. *The Paradox of Acting Translated with Annotations from Diderot's 'Paradoxe sur le Comedien'*. Trans. by W. H. Pollock. London: Strangeway and Sons, 1883. *Openlibrary*. Web. 2 April 2020..

- Disclosure: Trans Lives on Screen*. Directed by Sam Feder, Netflix Movies, 2020.
- Dissanayake, Ellen. *Homo Aestheticus: Where art comes from and why*. University of Washington Press, 2001.
- “Dr Geoff Lindsey • speech coach.” *YouTube Channel*, n.d.
<https://www.youtube.com/user/englishspeechservice>
- El Guindi, Youssef. “Playscript: Pilgrims Musa and Sheri in the New World.” *American Theatre*, September 2012: 63-80.
- Elliott, Nancy C. "Peer-reviewed Article Rhoticity in the Accents of American Film Actors: A Sociolinguistic Study." *Voice and Speech Review* 1.1 (2000): 103-130.
- Espinosa, Micha. "A Call to Action: Embracing the Cultural Voice or Taming the Wild Tongue." *Voice and Speech Review* 7.1 (2011): 75-86.
- “Find a Voice Pro.” *VASTA*. n.d.
https://www.vasta.org/content.aspx?page_id=154&club_id=516524
- Ferrand, Carole T. *Speech Science: An integrated approach to theory and clinical practice*. Pearson Education, 2013.
- Flege, James E. "Second language speech learning: Theory, findings, and problems." *Speech Perception and Linguistic Experience: Issues in cross-language research* 92 (1995): 233-277.
- Flege, James E., and Ocke-Schwen Bohn. "The Revised Speech Learning Model." *Unpublished Preprint*, 20 August 2020.
https://www.researchgate.net/publication/342923241_The_revised_Speech_Learning_Model
- Flege, James Emil, et al. “Factors Affecting Strength of Perceived Foreign Accent in a Second Language.” *The Journal of the Acoustical Society of America*, vol. 97, no. 5, May 1995, pp. 3125–34.
- Freeth, Becky. “Chernobyl Creators Explain Why the Characters Don't Have Russian Accents.” *Metro*, Metro.co.uk, 7 June 2019, metro.co.uk/2019/06/07/chernobyl-cast-asked-not-put-russian-accents-make-emotion-authentic-9854176/?ito=cbshare.
- Frieda, Elaina M., et al. "Adults' perception and production of the English vowel /i/." *Journal of Speech, Language, and Hearing Research* 43.1 (2000): 129-143.
- Frumkin, Lara. "Influences of accent and ethnic background on perceptions of eyewitness testimony." *Psychology, Crime & Law* 13.3 (2007): 317-331.

- Fuchs, Elinor. "EF's visit to a small planet: some questions to ask a play." *Theater* 34.2 (2004): 4-9.
- Giles, Howard, and Nikolas Coupland. *Language: Contexts and consequences*. Thomson Brooks/Cole Publishing Co, 1991.
- Gluszek, Agata, and John F. Dovidio. "The Way They Speak: A Social Psychological Perspective on the Stigma of Nonnative Accents in Communication." *Personality and Social Psychology Review*, vol. 14, no. 2, May 2010, pp. 214–37. *SAGE Journals*,
- Goldstein, Thalia. "Questions of Realness." *The Junkyard*, The Junkyard, 6 Apr. 2018, junkyardofthemind.com/blog/2017/8/14/questions-of-realness.
- Graham, David A. "A Short History of Whether Obama Is Black Enough, Featuring Rupert Murdoch." *The Atlantic*, Atlantic Media Company, 8 Oct. 2015, www.theatlantic.com/politics/archive/2015/10/a-short-history-of-whether-obama-is-black-enough-featuring-rupert-murdoch/409642/.
- Grover, Purva, et al. "Perceived usefulness, ease of use and user acceptance of blockchain technology for digital transactions—insights from user-generated content on Twitter." *Enterprise Information Systems* 13.6 (2019): 771-800.
- Hampton, Marian. "Editorial Column The Golden Rule for Standard Speech." *Voice and Speech Review* 1.1 (2000): 13-16.
- Hammond, David. "Peer-reviewed Article Sidebar 'Good Speech in Classic Plays': The Historical Perspective." *Voice and Speech Review* 1.1 (2000): 143-147.
- Hampton, Marian. "Editorial Column The Golden Rule for Standard Speech." *Voice and Speech Review* 1.1 (2000): 13-16.
- Hawkins, Eric W. "Foreign language study and language awareness." *Language awareness* 8.3-4 (1999): 124-142.
- Herman, Lewis, and Marguerite Shalett Herman. *Foreign Dialects: A Manual for Actors, Directors, and Writers*. Routledge, 1997.
- Hay, Jennifer, and Katie Drager. "Stuffed toys and speech perception." *Linguistics* 48.4 (2010): 865-892.
- Herman, Lewis and Marguerit Shalett Herman *American Dialects: A manual for Actors, Directors, and Writers*. Routledge 1947.

- Herrera, Brian Eugenio. "'But Do We Have the Actors for That?': Some Principles of Practice for Staging Latinx Plays in a University Theatre Context." *Theatre Topics* 27.1 (2017): 23-35.
- Hobbs, Robert L. *Teach Yourself Transatlantic: Theatre speech for actors*. Mayfield Publishing Company, 1986.
- Hodge, Robert, and Gunther Kress. "Social semiotics, style and ideology." *Sociolinguistics*. Palgrave, London, 1997. 49-54.
- Hodge, Robert, et al. *Social semiotics*. Cornell University Press, 1988.
- Hope, Donna. *American English Pronunciation: It's No Good Unless You're Understood*. Cold Wind Press, 2006.
- Hu, Guiling, and Stephanie Lindemann. "Stereotypes of Cantonese English, apparent native/non-native status, and their effect on non-native English speakers' perception." *Journal of multilingual and multicultural development* 30.3 (2009): 253-269
- Huspek, Michael R. "Linguistic variation, context, and meaning: A case of-ing/in'variation in North American workers' speech." *Language in Society* 15.2 (1986): 149-163.
- "Immigrants in the Progressive Era : Progressive Era to New Era, 1900-1929 ." *The Library of Congress*, n.d. www.loc.gov/classroom-materials/united-states-history-primary-source-timeline/progressive-era-to-new-era-1900-1929/immigrants-in-progressive-era/.
- Johnson, Mark. "The meaning of the body." *Developmental perspectives on embodiment and consciousness*. Psychology Press, 2007. 35-60.
- Kato, Misaki, and Melissa Michaud Baese-Berk. "The effect of input prompts on the relationship between perception and production of non-native sounds." *Journal of Phonetics* 79 (2020): 1-20.
- Kang, Okim, and Donald L. Rubin. "Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation." *Journal of Language and Social Psychology* 28.4 (2009): 441-456.
- Kendi, Ibram X. *Stamped from the Beginning: The definitive history of racist ideas in America*. Hachette UK, 2016.
- Knight, Dudley. "Peer Reviewed Article: Standards". *Voice and Speech Review* 1.1 (2000): 61-88.
- Knight, Dudley. *Speaking with Skill: A Skills Based Approach to Speech Training*. Methuen Drama, 2012.

- Knight, Dudley. *Speaking with skill: An introduction to Knight-Thompson speech work*. A&C Black, 2013.
- Knight, Dudley. "Reprint Standard Speech: The Ongoing Debate." *Voice and Speech Review*, vol. 1, no. 1, Jan. 2000, pp. 31–54.
- Knight, Dudley, and Philip Thompson. "About the Work." *Ktspeechwork.org*, ktspeechwork.org/about-the-work/. Accessed 19 Dec. 2020.
- Knowles, Richard Paul. "SHAKESPEARE, VOICE, AND IDEOLOGY: Interrogating the Natural Voice." *SHAKESPEARE, THEORY, AND PERFORMANCE*, edited by James C Bulman, Taylor & Francis, 1996, pp. 95-116.
- Kuhl, Patricia K., et al. "Cross-language analysis of phonetic units in language addressed to infants." *Science* 277.5326 (1997): 684-686.
- Lacko, Ivan. "Imaginative Communities: The Role, Practice and Outreach of Community-Based Theatre." *Ars Aeterna* 6.2 (2014): 21-27.
- Lakoff, George, and Mark Johnson. *Philosophy in the Flesh: The cognitive unconscious and the embodied mind: How the embodied mind creates philosophy*. Basic Books, 1999.
- Lambert, Wallace E., et al. "Evaluational reactions to spoken languages." *The Journal of Abnormal and Social Psychology* 60.1 (1960): 44.
- LaMonica, C. (2019). "Factors in an acoustical-attitudinal account of dialect perception." *New Ways of Analyzing Variation* 48. Poster (2019).
- "Language Opens Worlds." *Klingon Language Institute*, www.kli.org/.
- Liberman, Alvin M., and Ignatius G. Mattingly. "The motor theory of speech perception revised." *Cognition* 21.1 (1985): 1-36.
- Lindemann, Stephanie. "Who speaks "broken English"? US undergraduates' perceptions of non-native English 1." *International Journal of Applied Linguistics* 15.2 (2005): 187-212.
- Lindsay-Abaire, David. *Good People*. Theatre Communications Group, 2011.
- "Linguistix Pronunciation." *YouTube Channel*, n.d. <https://www.youtube.com/channel/UC3wikulG1obZp9k1sH2cH3Q>,
- Linklater, Kristin, Phyllis Epp, and William Snow. *Freeing the Natural Voice*. Drama Book Publishers, 1976.

- Linklater, Kristin. *Freeing Shakespeare's Voice: The actor's guide to talking the text*. Theatre Communications Group, 1992.
- Lippi-Green, Rosina. *English with an Accent: Language, ideology and discrimination in the United States*. Routledge, 2012.
- Lippi-Green, Rosina. "The Standard Language Myth" *Voice and Speech Review* 1.1 (2000): 23-30.
- Lopez, Robert, et al. *Avenue Q, the Musical: The complete book and lyrics of the Broadway musical*. Applause Theatre & Cinema Books, 2003.
- Lorde, Audre. "The master's tools will never dismantle the master's house." *Feminist Postcolonial Theory: A reader* 25 (2003): 25-40.
- McClelland, James L., and Jeffrey L. Elman. "The TRACE model of speech perception." *Cognitive Psychology* 18.1 (1986): 1-86.
- McConachie, Bruce. *Engaging Audiences: A cognitive approach to spectating in the theatre*. Springer, 2008.
- McConachie, Bruce. *Evolution, Cognition, and Performance*. Cambridge University Press, 2015.
- McConachie, Bruce. "Metaphors we act by: Kinesthetics, cognitive psychology, and historical structures." *Journal of Dramatic Theory and Criticism* (1993): 25-46.
- McGowan, Kevin B. "Social expectation improves speech perception in noise." *Language and Speech* 58.4 (2015): 502-521.
- McIntosh, Peggy. *White Privilege: Unpacking the invisible knapsack*. January 1990, convention.myacpa.org/houston2018/wp-content/uploads/2017/11/UnpackingTheKnapsack.pdf.
- Michel, Alexandra. "Cognition and Perception: Is There Really a Distinction?" *Association for Psychological Science - APS*, 29 Jan. 2020, www.psychologicalscience.org/observer/cognition-and-perception-is-there-really-a-distinction.
- Modiano, Marko. "Euro-English from a 'deficit linguistics' perspective." *World Englishes* 26.4 (2007): 525-533.
- Mollin, Sandra. "New variety or learner English?: Criteria for variety status and the case of Euro-English." *English World-Wide* 28.2 (2007): 167-185.

- Moore, Adrienne. "The History of the Voice and Speech Trainers Association (VASTA)." *Voice and Speech Review*, vol. 13, no. 1, Jan. 2019, pp. 97–105.
- Moore, Sonia. *The Stanislavski System: The professional training of an actor*. Penguin, 1984.
- Moyer, Alene. *Foreign Accent: The phenomenon of non-native speech*. Cambridge University Press, 2013.
- Mudd, Derek. *Staging the Voice: Towards a critical vocal performance pedagogy*. 2014. Louisiana State University, PhD
- Munro, Murray J., and Tracey M. Derwing. "Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners." *Language Learning*, vol. 49, Jan. 1999, pp. 285–310.
- Neuhauser, Sara, and Adrian P. Simpson. "Imitated or authentic? Listeners' judgements of foreign accents." *Proceedings of the 16th international congress of phonetic sciences*. 2007. 1805-1808
- Niedzielski, Nancy. "The effect of social information on the perception of sociolinguistic variables." *Journal of language and social psychology* 18.1 (1999): 62-85.
- Nilsson, Michael, Sigfrid D. Soli, and Jean A. Sullivan. "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise." *The Journal of the Acoustical Society of America* 95.2 (1994): 1085-1099.
- Nittrouer, Susan. "Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds." *Journal of Speech, Language, and Hearing Research* 39.2 (1996): 278-297.
- Nolan, Ian T., et al. "The role of voice therapy and phonosurgery in transgender vocal feminization." *Journal of Craniofacial Surgery* 30.5 (2019): 1368-1375.
- Oram, Daron. "De-Colonizing Listening: Toward an Equitable Approach to Speech Training for the Actor." *Voice and Speech Review* 13.3 (2019): 279-297.
- Pace, Chelsea. "The Certification Question." *The Journal for Consent-Based Performance*, 1 June 2021, www.journalcbp.com/the-certification-question.
- Pace, Chelsea, and Laura Rikard. *Staging Sex: Staging Sex Best Practices, Tools, and Techniques for Theatrical Intimacy*. Routledge, 2020.
- Pao, Angela Chia-yi. "False accents: Embodied dialects and the characterization of ethnicity and nationality." *Theatre Topics* 14.1 (2004): 353-372.

- Preston, Dennis R. "Whaddayaknow now." *Awareness and Control in Sociolinguistic Research* (2016): 177-199.
- Ramjattan, Vijay. "language practices are racialized, and language practices racialize as well" *Twitter*. 29 June 2021.
- Raphael, Bonnie N. "Peer-reviewed Article Dancing on Shifting Ground." *Voice and Speech Review* 1.1 (2000): 165-170.
- Reinares-Lara, Eva, Josefa D. Martín-Santana, and Clara Muela-Molina. "The effects of accent, differentiation, and stigmatization on spokesperson credibility in radio advertising." *Journal of Global Marketing* 29.1 (2016): 15-28.
- Robbins, Sanford. "Essay Edith Warman Skinner: A Former Student's Recollection and Appreciation." *Voice and Speech Review* 1.1 (2000): 55-60.
- Roach, Joseph. *The Player's Passion: Studies in the Science of Acting*. 1985. Ann Arbor: U of Michigan P, 1993.
- Rodenburg, Patsy. *The Right to Speak: Working with the voice*. Bloomsbury Publishing, 1993.
- Royde-Smith, John Graham. "World War II." *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., 15 May 2021, www.britannica.com/event/World-War-II.
- Rubin, Donald L. "Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants." *Research in Higher education* 33.4 (1992): 511-531.
- Rubin, Donald L., and Kim A. Smith. "Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants." *International journal of intercultural relations* 14.3 (1990): 337-353.
- Sabia, Joe. "Movie Accent Expert Breaks Down 32 Actors' Accents." *Wired YouTube Channel*, n.d. <https://www.youtube.com/watch?v=NvDvESEXcgE>
- Sabia, Joe. "Movie Accent Expert Breaks down 31 Actors Playing Real People." *Wired YouTube Channel*, n.d. <https://www.youtube.com/watch?v=lZSCGZphjq0>
- Sakland, Nancy. *Voice and Speech Training in the New Millennium: Conversations with Master Teachers*. Applause Theatre & Cinema Books, 2011.
- Sakland, Nancy. "Robert Barton." *Voice and speech training in the new millennium: Conversations with master teachers*. Applause Theatre & Cinema Books, 2011, pp. 29-38.

- Sansom, Rockford. "The unspoken voice and speech debate [or] the sacred cow in the conservatory." *Voice and Speech Review* 10.2-3 (2016): 157-168.
- Sedgman, Kirsty. "Challenges of cultural industry knowledge exchange in live performance audience research." *Cultural Trends* 28.2-3 (2019): 103-117.
- Siegel, Jeff. *Second Dialect Acquisition*. Cambridge University Press, 2010.
- Silverstein, Michael. "Indexical order and the dialectics of sociolinguistic life." *Language & Communication* 23.3-4 (2003): 193-229.
- Silverstein, Michael. "The Limits of Awareness. Sociolinguistic Working Paper Number 84." (1981).
- Singer, Reid. "How Should Black People Sound?" *The New York Times*, The New York Times, 28 Oct. 2020, www.nytimes.com/2020/10/28/style/hollywood-accent-coaches.html.
- Skinner, Edith, et al. *Speak with Distinction*. Applause Theatre Book Publishers, 1990
- Smiljanić, Rajka, and Ann R. Bradlow. "Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness." *The Journal of the Acoustical Society of America* 130.6 (2011): 4020-4031.
- Staff, ABQJournal News. "Student Production Explores the Connection between Linguistics & Love." *Albuquerque Journal*, 25 Oct. 2013, www.abqjournal.com/287877/albuquerque-theater-4.html.
- Statler, Matt, Loizos Heracleous, and Claus D. Jacobs. "Serious play as a practice of paradox." *The Journal of Applied Behavioral Science* 47.2 (2011): 236-256.
- Stevens, Kenneth N., and Sheila E. Blumstein. "Invariant cues for place of articulation in stop consonants." *The Journal of the Acoustical Society of America* 64.5 (1978): 1358-1368.
- Stoller, Amy, et al. "Speech stereotypes: good vs. evil." *Voice and Speech Review* 8.1 (2014): 78-92.
- Sullivan, Gail. "John Cho of 'Selfie': 'I Experienced Racism!'" *The Washington Post*, WP Company, 2 May 2019, www.washingtonpost.com/news/morning-mix/wp/2014/10/09/john-cho-of-selfie-wants-roles-outside-any-asian-stereotype-2/.
- Sumner, Meghan, et al. "The socially weighted encoding of spoken words: A dual-route approach to speech perception." *Frontiers in Psychology* 4 (2014): 1015.

“Tim Monich.” *Imdb*, n.d. <https://www.imdb.com/name/nm0598106/>.

Tonning-Kollwitz, Melissa, and Joe Hetterly. "The Current Use of Standard Dialects in Speech Practice and Pedagogy: A Mixed Method Study Examining the VASTA Community in the United States." *Voice and Speech Review* 12.3 (2018): 295-315.

Tonning-Kollwitz, Melissa, Joe Hetterly, and Ellen Kress. "The Current Use of Standard Dialects in the United States Theatre Industry." *Voice and Speech Review* (2021): 1-15.

Trudgill, Peter, and Jean Hannah. *International English: A guide to the varieties of standard English*. Routledge, 2013.

Ulin, David L. "Whitey Bulger' Digs Deep into a Gangster's Tale." *Los Angeles Times*, Los Angeles Times, 1 Mar. 2013, www.latimes.com/books/la-xpm-2013-mar-01-la-ca-jc-whitey-bulger-20130303-story.html.

Wester, Mirjam, and Cassie Mayo. "Accent Rating by Native and Non-Native Listeners." 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014.

We See You W.A.T., n.d., www.weseeyouwat.com/.

Wilkinson, Alec. "Talk This Way." *The New Yorker*, 11 Nov. 2009, www.newyorker.com/magazine/2009/11/09/talk-this-way.

"Who Is Kristin Linklater." *Linklater Voice*, n.d. www.linklatervoice.com/linklater-voice/who-is-kristin-linklater,

Zazzali, Peter. *Acting in the Academy: The history of professional actor training in US higher education*. Routledge, 2016.