UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Speech Stereotypes of Female Sexuality

Permalink
https://escholarship.org/uc/item/9c5407bt

Author
Barron-Lutzross, Auburn

Publication Date
2018

Peer reviewed|Thesis/dissertation

Speech Stereotypes of Female Sexuality

by

Auburn Lupine Barron-Lutzross

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Linguistics

and the Designated Emphasis

in

Women, Gender and Sexuality

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Keith Johnson, Co-Chair
Professor Susan Lin, Co-Chair
Professor Justin Davidson
Professor Mel Chen

Spring 2018

Abstract

Speech Stereotypes of Female Sexuality

by

Auburn Lupine Barron-Lutzross

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Keith Johnson, Co-Chair

Professor Susan Lin, Co-Chair

At its core, my dissertation addresses one primary question: What does it mean to sound like a lesbian? On the surface, this may seem a relatively simple question, but my work takes a broad perspective, approaching this single question from a multitude of perspectives. To do so I carried out a combination of experiments, interpreting the results through the Attention Weighted Schema Abstraction model that I developed.

Following the introduction Chapter 2 lays out the AWSA model in the context of previous literature on stereotype conception and speech and sexuality. Chapter 3 presents the production experiment, which recorded speakers reading a series of single words and sentences and interviews discussing stereotypes of sexuality. Phonetic analysis showed that though speech did not vary categorically by sexual orientation, familiarity with Queer culture played a significant role in variation of speech rate and mean pitch. This pattern was only seen for straight and bisexual speakers, suggesting that lesbian stereotypes are used to present an affinity with Queer culture, which was further supported by the decrease or loss of their significance in interview speech.

Chapter 4 describes three perception experiments. First, listeners judged speakers on attributes including likelihood of being a lesbian. Despite production results showing no variation by sexual orientation, listeners correctly identified lesbian speakers as more likely to be lesbians. Though not accurately predicting self-identification, four phonetic features contributed directly to lesbian ratings. A Principal Components Analysis of the ratings identified four recognizable social personae, including an 'angry butch lesbian' persona, showing a complex lesbian speech stereotype. The second experiment tested how listeners would judge sexuality without prompting. Listeners provided five adjectives to describe each speaker. Of 3480 adjectives, only one referenced sexual orientation. Finally, listeners judged only sexual orientation. Listeners were highly consistent in lesbian and straight ratings, but not at all for bisexual. In interpreting my results through the AWSA model, I posit a process of perception of lesbian speech, in which, rather than cueing orientation directly, phonetic features prompt perception of gender expression. Non-linguistic stereotypes then link these to sexual orientation.

Table of Contents

List of Figures

List of Tables

Acknowledgments

No dissertation is truly the creation of a single person, and mine is certainly no exception. More people have contributed than I can possibly mention here – more probably than I even realize – but here are the most important.

Keith Johnson and Susan Lin have been excellent advisors throughout this process. When I was first applying to grad school, my conversation with Keith convinced me that I should pursue a graduate education at Berkeley. He has been an amazing advisor in so many ways. Every meeting leaves me more inspired. My ideas become more clear, and the bridges between them become more solid. Susan supports me in all my research endeavors, helping me to approach the same question in different ways. With both advisors I am always challenged to ask the bigger questions, "How?", "Why?", and "Who cares?", which has made me a stronger scientist and more confident in my work.

Interactions with each committee member, Keith Johnson, Susan Lin, Justin Davidson, and Mel Chen, have been productive individually, but meetings with the whole group are a blast! Somehow I have managed to bring together four brilliant minds to talk excitedly and insightfully about my own work. Seeing them walk out of meetings with such enthusiasm encourages me to believe that my research might continue to inspire others.

I want to thank the many other linguists who have inspired me with their work and encouragement over the years. Ben Munson, Kathryn Campbell-Kibler, Katie Drager, Meghan Sumner, Molly Babel, Rob Podesva, and countless others. Molly and Rob in particular have been especially impactful. Many people have motivated me to incorporate psychology into my research in sociolinguistics, and Molly Babel has been top among them, through her research, conversations with her, and her impact on the department and our shared advisor. Rob Podesva helped me with the project that would eventually become this dissertation before it was even my Division III at Hampshire college, and he has continues to provide me with valuable feedback throughout the years.

The overwhelming effort involved in this research, especially the more tedious tasks like transcription and alignment, has been made more manageable and enjoyable with the help of all my Linguistic Research Assistants: Lucien Cady, Akshay Aitha, Katrina Olsen, Edie Sussman, Beth Martyn, and Q Cerabino-Hess. You have all inspired me with your work ethic and enthusiasm. Edie and Beth especially, in addition to contributing great ideas along with your time and effort, you were great reminders of why I got so excited about this topic in the first place.

I also want to thank all the members of the Berkeley PhonLab for their feedback and inspirational work. Talk about amazing linguistics! I especially want to thank my collaborators on the projects of Power and Accommodation and the LSA symposium on Sociolinguistic Cogntition: Alice Shen, Emily Remirez, Eric Wilbanks, and Andrew Cheng. Thank you also to everyone who ever participated in SPREAD, especially my co-founder Jevon Heath. The format of the group morphs and shifts every semester, but the

goals of collaborative learning and shared inspiration have remained constant. I hope it keeps this focused and organic nature for years to come.

Belen Flores, Paula Floro, and Ron Sprouse have helped me throughout this process in ways no one else could. From what I can tell, they are all-knowing. Belen, especially seems to always have answers to my questions before I even knew to ask them.

Thank you also to my many office mates, past and present: Jonathan Manker, Sarah Bakst, Matt Faytak, Meg Cychoz, Yevgeniy Melguy, Andrew Cheng, Emily Remirez, Jevon Heath, and Alice Shen. I know Dwinelle 50 can get pretty cramped at times, but being in a workplace with such fantastic people always reminds me how fun linguistics is.

I need to acknowledge the kindness of my other Ling 100 GSIs and instructor this past semester, Alice Shen, Eric Wilbanks, Lev Michael, and Emily Remirez for being so kind and understanding and making the last few weeks of writing my dissertation bearable, especially you Emily; I am awed by the hard work you put in for your friends and students.

Alice, thank you so much for doing your best to make me less of a hermit these past few years and adding joy to my life with your cynicism.

I have been so lucky to end up at grad school so close to my very loving parents and sister. Thank you Jim for all the lunches and chats. I'm very grateful for all the times you ask how my work is going, for dropping the topic when it's clear I don't want to talk about it, or engaging excitedly when I do. Thanks Mom, for being a true mom-a-lot and everything that that entails, from dinners, to medical advice, to random gifts, to rides to the airport at ungodly hours in the morning. I truly lucked out to get my amazing sister, Morgen Lutzross. You always challenge me to be my best and to never take that task too seriously.

Darjeeling Barron-Lutzross joined our team the summer before I started at Berkeley. As a fellow student with a passion for ornithology, she has always done a great job of keeping me sane, reminding me to stay curious and critical about the world..

Finally, my wife has been an incredible support for the last nine years. Michelle, you inspire me every day. Your respect and enthusiasm push me to do my very best in everything and give me the courage to believe in who I am and who I can be.

1 Introduction

"I think ... stereotypically they would have some sort of masculine like quality to them as well. So like that would mean in terms of like their voice they'd be like deeper or more like aggressive or like dominant. Like like when I speak in like the middle of like my sorority sisters, you hear my voice booming over people, and like like during our little ceremony we go off like in little groups of five, then have to like read the same verses, and you just hear mine infinitely over everyone 'cause it's like a solid octave lower and it's just like very loud and like 'uhn!' ... I also like am known for my voice. My identifiers are like oh like super blonde like you know whatever like like shaved body and then like deep voice and so it's like they don't expect me to talk the way I do and to have like the personality I have because I have very like stereotypical like sorority girl like 'woo!' like face and like appearance."

- taken from a sociolinguistic interview with a self-identified bisexual woman

1.1 Stereotypes and expectation

Most people aren't aware of a stereotype about the way lesbians sound, and even if they are, they tend to have difficulty articulating its distinctive features. However, as this speaker pointed out, these features made her voice odd and unexpected to others observing her appearance. These expectations clue directly to an underlying stereotype about lesbian style and how it connects to other social attributes such as the masculinity, aggression, or dominance mentioned in this quote. The comment points to a number of questions that I will address in this dissertation.

- What precisely are these stereotypes about speech and female sexuality?
- To what degree do these stereotypes reflect reality?
- In what ways do the stereotypes of lesbian speech intersect with other social stereotypes and personae?
- How are these and other stereotypes represented cognitively and how does this fit into sociolinguistic models of representation?

1.2 Conceptual grounds and categories of reference

As in any work investigating aspects of gender and sexuality, it is necessary to understand from the beginning what is meant here by the terms *gender* and *sexuality*, and why. For the most part in my research I define terms according to self-identification. That is, the subjects in my study who I refer to as *lesbian* chose this term to describe themselves. This is likewise true for *bisexual* and *straight* speakers. Any reference to sexual orientation that I make will be to these groups as subjects chose to define them. My default definition for each is self-identified women who are attracted romantically

and sexually to exclusively or primarily other women (lesbian), self-identified women who are attracted to men and women (bisexual), and self-identified women who are attracted to exclusively or primarily to men. Speakers and listeners when asked for their sexual orientation were asked to choose one of those three, and listeners when asked for their gender were asked to choose between *male* and *female*; I will refer to them as such for the rest of this dissertation

Although this selection goes beyond the binary limitations of sexual orientation in any previous research, given the range of possible gender identities and sexual orientations, this selection is still quite limited. The decision to set these limitations was made with careful consideration of the goals and scope of this particular study. I decided that, given the time and resources available, a statistical analysis would be most meaningful if gender were not variable and sexual orientation as a variable were set to three conditions. This allows me to interpret the results of phonetic analysis with reasonable confidence as to the influence of sexual orientation and as it interacts with other social and personal attributes of the speakers and listeners.

The recruitment flyer for the production experiment (Appendix A) explicitly called for women of any sexual orientation. It was expected that this would encourage only self-identified women to participate. This was successful with the exception of one individual who identified as gender queer. They were assigned female at birth but currently did not identify as female and so for the reasons given above, they were not included as a speaker in this study.

The findings of this study, as well as the few that have come before it, have indicated pitch or fundamental frequency[1] and first and second formants as stereotypes of lesbian speech. This reflects an important concept in how physiology contributes to gender versus sex. Although they can be and are regularly manipulated, these phonetic features are also influenced by the length of the vocal tract (for a review of this influence, see Simpson, 2009). This is related in many ways to sex and gender, particularly the positive correlation between height and perceived masculinity (Munson, 2007). I will lay out in the final chapter a more detailed model of the social and cognitive psychology underlying this correlation as well as how it relates to my larger cognitive model of stereotype abstraction.

1.3 Structure of the dissertation

The primary goal of this dissertation is to address these questions from a broad perspective. As a sociolinguist I approach the nature of linguistic variation as one which is both accessible to listeners as a means of understanding others and to speakers as a

---

[1] Which of these is used and how precisely it is interpreted varies by study. I measure of pitch in Equivalent Rectangular Bandwidth, described in Section 3.2.4.)

means of presenting some element of ones identity. In exploring the underlying nature of sociolinguistic stereotyping is clear that these concepts extend in a broader sense beyond sociolinguistics to incorporate elements of psycholinguistics as well as social and cognitive psychology. In my analysis of sociolinguistic stereotyping, I developed a model of cognition which I term the Attention Weighted Schema Abstraction. Based at its core on the models of Nosofsky (1986) and Hintzman (1986), AWSA models stereotype construction, perception, and access as a single process in which exemplars are stored and activated. Abstraction across these exemplars is not a separate system existing at some higher level, but rather occurs through the process of activation. The second chapter of this dissertation presents previous research on stereotype theory and speech and sexuality and provides a detailed description of AWSA.

The third chapter is a discussion of a speech production experiment, in which I recorded and analyzed the speech of self-identified lesbian, straight, and bisexual women. In this experiment I collected data about the speakers, including their sexual orientation and their self-evaluated familiarity with Queer culture. Speakers were recorded reading a series of individual words and sentences, as well as in casual interviews discussing stereotypes of gay men and lesbians in speech and other social stereotypes. I found that self-identified sexual orientation was not independently significant in predicting phonetic variation. However, the familiarity a speaker reported to have with Queer culture (however she interpreted that) was a significant feature. In read speech, speakers' familiarity with Queer culture correlated positively with a lower mean pitch and faster speech rate. Most notably, when broken down by sexual orientation, only straight and bisexual participants showed this pattern. It would appear that this style of speech represented some element of affiliation or ally identity which the self-identified lesbian speakers did not have reason to access. This interpretation is further supported by the decrease in significance of familiarity in interview speech. The use of this speech stereotype in the 'how' speech mode rather than 'what' speech indicates that it does carry some element of performance, and familiarity with Queer culture is not likely to predict speech patterns when sexuality is not salient.

In the fourth chapter I discuss a set of three perception experiments testing from various perspectives whether and in what ways sexual orientation is perceived in female speech. All three use recordings from the speech production experiment. The first, Experiment 2a, collected judgments of listeners on a set of social and personal attributes (including likelihood of being a lesbian) and analyzed how these were predicted by a wide range of phonetic and acoustic variables. As part of this experiment, I ran a Principal Components Analysis resulting in a set of principal components resembling common social personae. The second, Experiment 2b, used free elicitation to collect descriptions of speakers. In this way I could compare what judgments people made when prompted to rate specific attributes to the ways people stereotyped others without prompting. The resulting list of adjectives were analyzed using Latent Dirichlet Analysis combining them into a smaller set of shared personality types, similar to but distinct from the personae of Experiment 2a and involving almost no reference to sexual orientation. The third, Experiment 2c, took

the opposite approach from 2b, asking listeners to rate sexual orientation exclusively. Bisexual ratings were apparently random, indicating that the social issue of bisexual invisibility extends to inaudibility as well. However, straight and lesbian ratings were highly consistent. The results indicate a clear lesbian speech stereotype, but also provide insight into the variation in findings of previous research. People recognize stereotypically lesbian speech, but only when they are explicitly asked to do so. The conclusion summarizes the results of Experiments 1 and 2 in relation to each other and provides a broader interpretation through the lens of the AWSA model.

2 Previous literature and the AWSA model

2.1 Introduction

2.2 The psychology of stereotypes

   The discussion of stereotype formation has often been addressed from one of two perspectives. One perspective is of individuals or groups projecting stereotypes onto others, whereas the other is of individuals self-stereotyping according to their in-group. Due to a general lack of communication between researchers and theorists of these perspectives, two areas which would logically be closely related are viewed quite differently with different explanations, theories, and models of how stereotypes are formed. In this paper, I give a review of literature on both topics, and preset a cognitive model of how the formation and application of stereotypes to others and the self are part of an identical process. The model is exemplar-based, linking the MINERVA 2 model presented by Hintzman (1986), which incorporates episodic memory in schema abstraction, and Nosofsky's (1986) model of identification and categorization, which includes the influence of selective attention. This integrative model of Attention Weighted Schema Abstraction (AWSA) is then applied to a number of theories and areas of research, which, though they have not been connected before this, can be incorporated as support for this model, demonstrating that a single process predicts the various results of a wide range of research areas.

2.2.1 Exemplar models and abstraction

2.2.1.1. Schema abstraction

   In 1986, Hintzman presented the MINERVA2 model, which posits that episodic memory allows for the creation of abstract schemata through a process of trace activation. Hintzman distinguishes between the Primary Memory (PM), which is the representation of the current experience, and the Secondary Memory (SM) which is composed of exemplars from previous experiences. The abstract representations are not acquired but are produced as a response to a *retrieval cue*. This retrieval cue is an active representation of an experience in PM, also referred to by Hintzman as a *probe*. In Hintzman's model, each experience is stored in the memory as a unique *trace*. Even if experiences are very similar (e.g.: encounters with the same individual) they will be stored separately. When a new probe is encountered, this activates all traces stored in SM and creates what Hintzman terms an *echo*. This echo is the sum of all traces, but each is activated with different strength depending on similarity to the probe. Hintzman describes similarity between two traces as "related roughly to the number of properties they share." Traces that are particularly similar to the probe will be more influential to the *content*, the pattern of properties, of that echo, and the more traces exist that are similar to the probe, the greater the *intensity* of the echo will be.

2.2.1.2. Selective attention

Also in 1986, Nososky presented a model of how stimuli are identified and categorized according to similarity to other exemplars, rather than abstract group prototypes. In this model, as in MINERVA2, exemplars are similar based on the overlap of their properties, which Nosofsky refers to as *dimensions*. One of the central aspects of the model is the concept of *attention weights*. Nosofsky explains this as the assumption that "subjects will distribute attention among the component dimensions so as to optimize performance in a given categorization paradigm" (p 41). In this way, different dimensions will receive degrees of attention, not only in comparison with other dimensions, but also compared to the same dimension in other contexts with a different categorization paradigm. In testing the identification and categorization patterns of two individuals, Nosofsky demonstrated that people do, in fact, attend to different dimensions differently. This was manipulated through comparisons with categories that differ according to certain dimensions. When the participants identified stimuli, these dimensions influenced their judgments more strongly than others which did not provide information about groups. Similarity between exemplars is not an objective sum of their shared properties but is subjective and depends on the context in which the exemplars were formed.

Smith and Zarate (1992) applied Nosofsky's model to social judgments, where the stored exemplars are of individuals. They demonstrated through a review of literature on social judgments, that selective attention is seen in perceiving individuals, just as it is in perceiving the visual object stimuli in Nosofsky's experiment. Studies have shown that similarity between perceived individuals is judged differently according to context, such as the task or goal, comparisons to different groups, and repeated exposure. They have also demonstrated that individuals often have dimensions that they tend to pay more attention to, and that social defaults exist, so that when an individual deviates from this default, attention will be paid to the dimension or dimensions in which they vary. For example, a 1990 study by Smith and Zarate found that in a categorization task, subjects categorized individuals according to the dimension in which they deviated from the default of White male. Black men were categorized as Black, rather than male, and White women were categorized as women rather than White. The literature on stereotype formation and self-stereotyping that I present also demonstrate the influence of such factors.

In Smith and Zarate's (1992) model of social identification and categorization, stereotypes were considered to be applied to individuals after they are categorized. Their focus was not on the formation of stereotypes, so how they were used or applied was not a concern for the authors. However, the idea that stereotypes are developed in the memory and stored separately doesn't fit well into an exemplar model such as theirs. Hintzman's model of abstraction from exemplars explains how stereotypes are formed, and that they are created each time a new experience is stored in memory. However,

studies such as the ones cited by Smith and Zarate, as well as ones discussed later in this paper, indicate that selective attention is extremely influential in perceived similarity between traces and probes, and thus, the content of stereotypes.

2.2.1.3. AWSA and stereotype formation

When the MINERVA2 model is applied to social stereotyping, the content of a stereotype can be thought of as the abstract schema of the echo that is activated when the probe is an individual, a group, attributes of the group, or even a simple request for a stereotype. For example, when people in a study are asked to list stereotypes of a lesbian speech style, this will activate an echo of all memory traces simultaneously, but ones that are identified as being part of the lesbian category will have the strongest influence on the content of the echo.

The mechanics of how this functions in AWSA are broken down into four steps.[2]

1) An individual experiences a probe in Primary Memory (PM).
   - Depending on context, different dimensions of the probe with receive more attention compared to other probes and other dimensions of the same probe.
2) Probe stored as a trace in Secondary Memory
   - Dimensions receiving more attention will have greater weight (**Attention Weighted**)
3) Probe activates all traces from Secondary Memory, forming an abstract echo (**Schema Abstraction**)
   - Traces with overlapping dimensions are activated with more strength and more influential on the content of the echo
4) Echo also stored as a trace

For example: a listener hears a single utterance from a female speaker and, based to context, attends more strongly to the gender of the speaker than sexual orientation or other dimensions. This probe activates all other traces as an echo. Traces with an overlapping dimension of gender, especially where gender in those existing traces received greater attention, contribute more to the content of the echo. This echo represents an abstraction of speech from female speakers; dimensions of the acoustics which were attended to with considerable weight will contribute to a stereotype of female speech but not to speech according to sexual orientation.

2.2.1.4. Selective exemplar storage (Garrett & Johnson, 2012)

Garrett and Johnson (2012) present a simulation of sound change using an exemplar model which follows the Labov (1994, 586-587) suggestion that

---

[2] These are not necessarily chronological, as multiple steps overlap. I make no claims without direct testing regarding the time course of the process

'misunderstood tokens may never form part of the pool of tokens that are used.' Any token (in this case a word) that is not understood will not contribute to the average value for that cloud of exemplars. The process of attention weighting supports this directly. Where a token is not understood to be a particular lexeme, it cannot be stored as such, even if the utterance may be stored as an experience in some form. For instance, an utterance in a language not spoken by that listener may still be stored as an example of that language, but not as an expression of the meaning intended by the speaker. Based on research in speech imitation and accommodation, Garrett and Johnson also argue that listeners may selectively store exemplars based on social biases such as group status. As I will discuss later in this chapter, multiple social factors may contribute to this. Exemplars may be stored within a subgroup which does not affect the mean value of the larger set. They may be stored as outside the set entirely as being produced by out-group speaker, or finally tokens may be too socially 'insignificant' to listeners to receive any attention weighting.

2.2.1.5. Trait aggregation model (Can & Sanver, 2009)

According to the AWSA model, all traces are equal in weight; stereotypes are not determined by which individuals form stronger traces, but by how they compare in their various dimensions to the probe. This contrasts with models such as the mathematical formalization of the attribution hypothesis by Can and Sanver (2009), where stereotypes are perceptions of groups that a person forms as an aggregation of common traits. Through a number of proofs, the authors propose a model of stereotype formation in contrast to the very simple idea that people apply to a larger group the traits that they witness in the majority. Their model includes the influence of p, the subjective weight an individual assigns to the visible members of the group, ω, the weight distribution based on an individual's perception of the representativeness of subgroups, and q, members or majority that are included or excluded from the overall representation due to an individuals' existing prejudices. The authors provide a clear formalization of a proposed process but do not attempt to apply it to existing research.

The model in many ways reflects the concepts behind an exemplar model, primarily in the concept that traits associated with a group will be based on observations of members of the group. It contrasts distinctly, however, in the concepts that individuals make judgments of how representative the observed group members are and that they make abstractions based on the knowledge that other group members exist. According to most exemplar models, judgments are based only on stored memory traces. Echoes of these traces create abstract prototypes, which are in turn stored as traces; these are not influenced by abstract awareness of "invisible" members of the group. In fact, in a model of memory trace abstraction, awareness of invisible members (even just from being told they exist) makes them inherently visible. Any conceived members, whether directly observed or not, exist as traces in memory, and inferred knowledge of these members,

regardless of accuracy, will influence the echo of a probe for that group. In fact, when one of the only dimensions of this trace is membership in a given group, other dimensions will strongly influence the content of the echo, more so than many traces formed from direct interactions. The strength of a trace's effect on the content of the echo is determined by its similarity to the probe. The influence of a trace is not determined in the process of forming that trace from an experience, but in how it is activated based on similarity to a probe. When a new individual is encountered, information is stored and will contribute to a stereotype probe that is sufficiently similar. If a given trait shows up in most traces (and has attention weight), it will be part of the content of the echo. Thus, while there is not prototypical representation of that group, echoes of traces will contribute to the concept of a stereotype (based on a stereotype probe).

2.2.2 Stereotyping others

This model of schema abstraction in a multiple-trace memory has not been previously applied to stereotyping. Most existing models of stereotype formation include stereotypes as a more abstract concept, often a kind of prototype of the group that is formed in various ways. An earlier example of this is the integrative model of stereotype formation of Bar-Tal (1997), which combined a variety of existing models in social psychology and cognitive stereotype theory. Existing models tended to focus on one area or the other, either illustrating how various aspects of social psychology contribute to societal development of stereotypes, or how cognitive processes affect the ways in which individuals perceive and stereotype others. This theory was developed to integrate these different perspectives to not only explain the formation and change of stereotypes (specifically ethnic and national), but also to account for their varying intensity and extensiveness. The model identifies three overall types of variables that can influence stereotypes. These include the *background variables*, such as historical interaction or political climate, that determine the relationship between groups, *transmitting variables*, such as family relations or direct contact with out-group members, that determine how stereotypes are transmitted between individuals and generations, and *personal mediating variables*, such as individual motivations or cognitive style, that influence how an individual processes and perceives stereotypes. Through analyzing and understanding these different variables, Bar-Tal argued that patterns in the formation and change of stereotypes could be predicted, although this was only applied to the larger established groups of ethnicity and nationality.

Bar-Tal's model included a cognitive perspective, but with the AWSA model, the social variables that Bar-Tal incorporated can also be viewed from a cognitive perspective. As he explains, background variables like political climate are important in stereotype formation in that they influence intergroup attitudes and ideas of what are important aspects of that group (p 498). In an exemplar model like AWSA these can be argued as important for the same reason that Bar-Tal argued, but the exemplar model

goes into more depth by explaining the mechanism of the variables' influence. Background variables influence what dimensions of an experience receive greater attention. For example, when there is extreme or salient conflict between two nations or cultures, members of the out-group will be stored as memory traces with stronger attention paid to their group status and dimensions that are associated with that group. Bar-Tal's transmitting variables act in a similar way, influencing not only selective attention, but in the case of contact with out-group members, will result in a greater number and more varied memory traces of members of that group. Personal mediating variables included in Bar-Tal's model are ones which are more clearly related (as he stated) to the cognitive processes involved in stereotype formation. These are specifically referencing patterns common to that individual. As I will discuss in more detail later, individual variation in cognitive stereotype formation includes differences in whether an individual is inclined to attend to more or fewer dimensions in creating memory traces, or if they have certain dimensions which they default to regardless of context.

In AWSA, both *subtyping* and *subgrouping* and involved in the cognitive process of stereotype formation. Subtyping is the process whereby group members who are apparently distinct from or at odds with the stereotypes and standards of the larger group are clustered together and set aside as exceptions, dissimilar from the group as a whole. Although individuals may perceive these exceptions as being members of the group in that this membership is one of their features, they will not consider the group as including them in that they do not affect the overall perception of the group. Subgrouping is a similar process where individuals form cognitive clusters of members of a group based on similarities between members of the subgroups and differences from other members of the larger group. In the case of subgrouping, unlike subtyping, smaller subgroups influence the overall perception of the group.

Most earlier research focused on one of these two processes, but did not compare the two. In 1995, Maurer, Park, and Rothbart combined research into a single model of cognitive processing in stereotype formation and change. Individuals are clearly able to perceive larger social groups while still maintaining subgroups within them. However, the ways in which subgrouping occurs is not always the same, and in this paper, the authors presented a model incorporating the two concepts of subtyping and subgrouping.

Maurer and colleagues conducted two studies to test their model and their theory that whether an individual is subtyped or subgrouped depends on how that group member is represented in relation to the superordinate group. Subjects in the first study were given descriptions of 16 fictional members of a certain group along with a general description of the group at large. The descriptions were presented so that members could be easily divided into five subgroups, with one subgroup containing only members who disconfirmed the general group stereotypes, the subtype. Subjects were divided into three conditions, 1) nonsort condition – they were only told to form an impression of the group, 2) subgroup condition – they were told to cluster together members who were similar in

some way to form as many groups as they like, and 3) subtype condition – they were told to sort members into individuals who did or did not fit the group as a whole. Following this step, subjects completed a number of judgments about the group and the different members, including the degree to which the group had stereotypic traits, how heterogeneous the group was, and how typical each group member was.

The results confirmed the hypothesis that subjects in the subtyping condition would rate the disconfirming members as more atypical and perceive the superordinate group to be more stereotypical and homogeneous, because they were not including the disconfirming members as part of the superordinate group. In fact, the more atypical the disconfirming members were considered to be, the smaller the range was judged. The authors stated that these perceptions are likely not causal but do interact. Subjects in the subgrouping condition judged the members and group homogeneity in the opposite direction, perceiving disconfirmers as less atypical and the group overall to be less homogeneous. The nonsort subjects gave ratings in between subjects in the other two conditions. Contrary to their hypothesis, the authors found that the subjects in both subtyping and subgrouping estimated large ranges in group members. They argued that the disconfirming members were especially striking for subtyping subjects (as evidenced by their ability to recall details in the recall task), so, even though these group members were atypical and did not contribute to the stereotypes of the group at large, they were still recognized as group members, and were even, in some cases, more noticeable.

The second study applied the model to real groups with existing stereotypes to test whether, without direct instruction, subjects would naturally have different strategies and different tendencies towards including or excluding disconfirming members. The target stereotype of the second study was gay men. Subjects were presented with descriptions of 15 fictional gay men. Twelve of them confirmed the stereotypes of gay men as effeminate, removed from the social mainstream, not involved in contact sports, and involved in short-term or casual relationships. The other three members disconfirmed all four stereotypes. Similar to the first study, the authors found that the more atypical the subjects rated disconfirming group members, the stronger their stereotype of the group. They also found that there was natural variation among individuals, indicating that some subjects were more inclined to subtype disconfirming individuals while others subgrouped them. The results of these studies provide the first direct comparison between the strategies of subgrouping and subtyping. They support the theory that both are ways of processing members of a group and that the same information can be processed differently according to the instructions or the general tendencies of the individual.

These studies provide strong support both of the key aspects of the AWSA model, the role of similarity in trace activation, and the existence of selective attention in determining what is perceived as most similar. As these studies indicate, similarity plays an important role determining which disconfirmers are subtyped and which

are subgrouped. Probes will more strongly activate traces that are more similar. While they are still given the same group label, the disconfirmers which are sufficiently distinct from the others will not be strongly activated with them and will not contribute to the content of the echo. Similarity is not objective but is dependent on the amount of attention certain dimensions receive. Allocation of attention to different dimensions can depend on a number of factors, but one of these is the context in which the memory trace is formed. Evidence for this is shown in the findings of Maurer et al. that the same disconfirmers may be subtyped or subgrouped depending on the context in which they were encountered and which dimensions were salient in comparisons to other members of the group. When disconfirmers are different from other members in regards to these dimensions, the difference will receive more attention and they will not be perceived as similar to most traces of group members. Because of this, the label of the group will be applied, but they will not be as strongly activated by new probes such as the request for an overall stereotype. They will not contribute to the echo from these new probes in the same way that group members disconfirming other dimensions (ones that are perceived with less attention) will. Because subjects still recognize these disconfirmers as members of the group, they rate diversity of the group equally high or higher, but the general stereotype is not affected.

Approaching lesbians as a social group with different subgroups can provide insight into how the processes of subtyping and subgrouping are applied outside of the lab. Studies have shown that 1) subjects of studies across cultures identify similar subgroups of lesbians, and 2) there appears to be a default set of traits associated with lesbians as a group at large, though it is sometimes vague. The first of these can be seen in a number of studies on lesbian stereotypes. Some studies investigated stereotypes of lesbians held by non-lesbians (Geiger et al., 2006; Brambilla et al., 2011), while others collected information from interviews with lesbians and bisexual women about ideas of lesbian identity within the lesbian community (Eves, 2004; Farquhar, 2000; and Huxley et al., 2013; Kennedy and Davis, 1993).

Geiger et al. (2006) collected stereotypes of lesbians from straight college students in the US using two tasks, first eliciting descriptions of lesbians with who subjects were familiar, and later having other subjects sort the previously provided traits into specific types of lesbians. While this method may have been flawed in that it forced subjects in the second experiment to incorporate traits that they may not have included in any of their own stereotypes of lesbians, it nonetheless found a number of lesbian subgroups that were shared by most subjects. These included 'lipstick lesbian', 'career-oriented feminist', 'soft-butch', 'free-spirit', 'hypersexual', 'sexually confused', 'sexually deviant', and 'angry butch'. The study also does not indicate which of these types of lesbians were the most commonly identified, but the descriptions of the 'lipstick lesbian' and 'angry butch' align with descriptions of the 'femme' and 'butch' stereotypes that are identified by lesbians and non-lesbians in other studies.

Another study looking at straight men and women was conducted Brambilla et al. (2011). The authors collected perceived lesbian subgroups from college students in Italy. The first part of the study identified four commonly perceived lesbian subgroups, 'feminine', 'butch', 'outed', and 'closeted'. The second part had subjects make social judgments about either the group as a whole, or the four different subgroups. Following the frame of Stereotype Content Theory (SCT), subgroups or the supergroup were judged on scales of warmth and competence. The results showed that the lesbian supergroup had generally neutral ratings on both axes, but, demonstrated by the range positive and negative ratings for the subgroups, this was likely a result of the fact that 'lesbian' is not as significant a category for people as the individual subgroups. In both of these studies, a number of stereotypes are applied to lesbians, clustering traits and creating subgroups. In the first there was no test for a general concept of lesbians, only for more specific groups, and these are shown to range widely in extremes of different dimensions. The second study demonstrates that there may not be a single one of these groups that is a default stereotype; when subjects are asked for only one description of lesbians, it is vague, falling somewhere between the ratings given for the different subgroups. Subjects appear to be applying an average, where memory traces are activated more or less equally.

The question of subtyping versus subgrouping has also been explored in female sexuality from the perspective of in-group self-stereotyping. Kennedy and Davis (1993) discuss a set of quite distinct types of lesbians, in this case specifically as perceived by members of a lesbian community in Buffalo, NY, which is said to be representative of many other communities of this sort. The individuals interviewed all cite the difference between the more masculine "butch" lesbian, and the more feminine "femme" lesbian. Whether these subgroups are representative of individual identity, relationship role, or performance depends on the individual, but the two subgroups are clearly reflected in interviews with lesbians and bisexual women in subsequent research by Farquhar (2000), Eves (2004), and Huxley et al. (2013).

Interviews by Huxley et al. focused on norms of style and appearance. Interviewees indicated a split between these masculine and feminine subgroups, but they also show a clear default, with most using the term "butch" in describing lesbian style. The interviews were conducted partly as a way of understanding whether this default stereotype is changing or being dissolved, but the results demonstrate that this is only true to a small degree.

Other studies where subjects are not encouraged to consider more than one type of lesbian demonstrate the existence of a default stereotype, or at least, stronger activation of masculine exemplars. This can be seen in the research by Wright and Canetto (2009) looking at stereotypes of older lesbians and gay men, which found that gender-related stereotypes of lesbians were similar to those of straight men. A study by Eliason et al. (1992), showed that nursing students' stereotypes of lesbians include a "male aura". In my own research, interviews with participants found similarly that

stereotypes of lesbian speech and appearance included aspects generally perceived as masculine, with the term "masculine" being used by many participants to describe the speech style, as in the interview quote given in the introduction. In the perception study, voices that were judged as sounding like a lesbian were significantly more likely to be judged as masculine and uncaring (an attribute commonly associated with masculinity).

Smith and Zarate discuss this concept of a cultural default. In their discussion they focus on default groups which are generally not considered subgroups of any larger groups. They cite the common cultural default of the young, White, straight, male, without disabilities. However, there is no reason to believe this same phenomenon will not exist for subgroups. If this is the case, we may consider the "butch" subgroup to be a cultural default, while the "femme" lesbian is a subtyped group that is separated out and does not significantly affect the stereotype. The stereotype that people present when only giving one, generally trends toward a butch role, but different studies show that this can be vague. This vagueness indicates that individuals 1) have different degrees of strength of an overall concept of lesbians with no actual subgroups, which I will discuss later, or 2) vary in whether they subgroup or subtype disconfirming (non-butch) lesbian exemplars.

The idea of a cultural default questions to what extent individuals form stereotypes by comparing groups to this particular default rather than in comparison to other existing groups. In AWSA, comparison to other exemplars is a primary factor in the role of attention weighting and which dimensions receive this attention. Wyer and colleagues (2002) presented research investigating the significance of intergroup comparison in stereotype formation and stereotype change. This study built off of theories of subtyping, where individuals are not likely to change their stereotypes of a group if they subtype, rather than subgroup the disconfirming group members. The two experiments tested the theory that context in stereotype formation can affect the way stereotypes change later when new and disconfirming group members are encountered.

As with the earlier study by Maurer et al. but contrary to some other researchers, the authors argued that the method of inclusion used is not necessarily based on the features of the group member. While they agree with the general assumption that individuals will subtype disconfirmers if they are sufficiently different from existing group members, they argue that the perceived degree of difference between members depends on the comparative context in which the group stereotypes were originally presented and thus the salience.

Two experiments were conducted to test this theory. In both experiments subjects were presented with information about members of a focal group, specifically learning about stereotypic behaviors along the dimensions of 'academically motivated' and 'politically liberal'. Subjects formed impressions of the focal group in contrast to one of two comparative groups. One group had members who showed behaviors identified as academically unmotivated and artistic, while the other group showed behaviors that were

14

politically conservative and campus leaders. Each comparative group contrasted with the focal group in one dimension but not the other. Subjects were then presented with descriptions of additional members of the focal group who disconfirmed both stereotypes. The authors hypothesized that 1) subjects would form stronger stereotypes about the attribute of the group that contrasted with the comparative group and 2) with the addition of disconfirming members, subjects' stereotypes would change more for the attribute that did not serve to contrast the two groups despite the fact that they disconfirmed both stereotypes. Both of these hypotheses were confirmed, supporting the subtyping argument that group members are more likely to be subtyped if they disconfirm the comparative stereotype.

Though the authors do not make this connection, it could be argued that the unequal significance of the dimension that differentiates a group is an example of selective attention in social categorization. The contrast effect as Wyer et al. describe it, is one aspect of context which shifts attention to another dimension. When categories are shown to differ according to a certain dimension, that dimension will receive more attention in categorization tasks. The shift, as with shifts for other reasons, causes exemplars that would be perceived as similar in other contexts to be perceived as different. The results provide support for the model in directly demonstrating that similarity is subjective and depends on how experiences are perceived and thus how they are stored as traces.

This contrast effect can become complex when one group is compared to multiple others that may not all differ along the same dimension. One example of this is the perception and stereotyping of lesbians. Lesbians are most often compared to straight women because the contrasting dimension can be said to only be sexual orientation. All are women, and this dimension of gender is arguably more visible (both in the literal and non-literal sense) than the dimension of sexual orientation. This comparison and visibility could be part of the reason for the masculine stereotype of lesbians in general. Because gender is such a salient and visible dimension, it is often applied to dimensions that are not directly related. As a way of distinguishing them(selves) from women of a perceptually opposite sexual orientation, individuals, including lesbians, apply contrasts in the dimension of gender to indicate differences in sexual orientation. This is only one explanation and is likely only one of many components contributing to the "butch" stereotype.

Another possible explanation is comparison to straight men, who, though they differ on the two primary dimensions of gender and sexual orientation, are similar to lesbians in the gender of their sexual and/or romantic partners. When attention in a context is given to an individual's sexuality (which can often be the context in which a woman's sexual orientation becomes known), similar traces will be ones where attraction is towards women. Finally, this effect of contrast may provide insight into why sexual orientations other than homosexual and heterosexual, so often lack clear stereotypes

(Eves, 2004; Farquhar, 2000; and Huxley et al., 2013) – they are not being contrasted with a single group.

2.2.2.1. Familiarity and effort (Smith et al., 2006)

It is often assumed in social and cognitive psychology that repeated exposure to an individual leads to familiarity and perception of them as more unique, and often to increased liking. This is shown to generally be the case in real life and is an argument for increased diversity as a means of changing stereotypes and decreasing discrimination. However, it has also been shown that with no motivation for thoughtful processing, people often put in less effort as a way to reduce use of limited processing capacity. Analytical and effortful processing is reserved for novel objects and situations.

Smith and colleagues (2006) ran two experiments to test the theory that in a context in which subjects have no social motivation to process individuals analytically, increased familiarity will lead to decreased individuation and increased stereotyping based on group membership. Subjects were presented in the first experiment with photos of individuals and then later with photos and descriptions, including occupation, of individuals. Half of the individuals in the second set were repeated from the first set. Individuals in the second set were given descriptions that contradicted common stereotypes of their given occupation. Subjects were then asked to rate the people on five traits (two fillers and three fitting occupational stereotypes), as well as rating their likeability. As the authors hypothesized, likeability did not change with repeated exposure; subjects rated people more highly on stereotypical traits if they had already viewed them in the first set, demonstrating that repeated exposure to individuals can lead to increased stereotyping.

The second experiment followed the same general design, but added, for half of the subjects, a question about the subject's familiarity with each person after the judgment task. Previous studies have indicated that awareness or repeated exposure can negate the effect of other judgments. Supporting their hypothesis, the authors found that repeated exposure increased feelings of familiarity in addition to increasing stereotyping of individuals. While the authors note that in most cases of increased familiarity, there is also increased closeness, which usually leads to greater liking and individuation. However, in cases when exposure is repeated without direct interaction (such as with someone on a bus or a waiter in a restaurant), individuation will actually decrease, and people will stereotype more. Smith et al. based the argument for their findings on the concept of minimal effort; when people encounter something again, they will use less effort than when encountering a new item. However, the exemplar-based model that I present supports their findings equally well and gives an explanation for why stereotyping decreases in most real-life encounters. The issue is not a matter of familiarity, but of similarity in traces. In cases such as these two experiments, where the

exemplars are almost identical, an activation of the memory traces will contribute to the content of the echo as if they were a single exemplar, but their influence will be stronger simply because of their increased number. When, as in most experiences in life, exemplars of the same individual vary in many ways and for many reasons (primarily context), each of these traces will contribute to the echo in different ways, producing not only a more complex representation, but one that is based less on abstractions from other members of the group and changes further with each new encounter.

When this concept is applied to stereotypes of sexual orientation, and specifically a "Lesbian Speech Style", it is important to consider the two options for familiarity, one being repeated exposure with more information revealed in each exemplar (usually direct contact) the other being repeated exposure with no/little new information each time (usually indirect contact). Here different genders have different degrees of accessibility of stereotyped features. For lesbians there usually has to be some amount of direct contact in order for the knowledge of their sexuality to become known (due to a lower level of visibility), so repeated exposure means more information and less stereotyping. For gay men, this may be less true, especially if they use more stereotypically gay male speech patterns. This provides an example of how linguistic stereotypes can become self-perpetuating. Using stereotypical speech patterns allows for group identity to be known to others without direct interaction; with this group identity stored as part of the memory trace, sociolinguistic variables are linked to groups without any interaction that would reveal more information. This creates more similar traces and stronger stereotypes. When a sociolinguistic stereotype is not as strong for a group, the process will be much slower, if it occurs at all.

To give real-life example, at a café I go to regularly, one of the baristas speaks with features of stereotypical gay male speech. The style is recognizable to myself and even though I have never engaged in a conversation beyond ordering a cup of coffee, I am fairly confident of his sexual orientation. Even though I don't know that he is gay, traces of those experiences store him with that label. In this way, more nearly identical traces of a gay man with a stereotypical gay speech style are added to my memory and influence the intensity and the content of an echo when activated by a probe that has similar phonetic or social features. Any similar experiences I have with lesbians are rarer and thus produce a weaker echo.

2.2.2.2. Stereotype directionality (Cox & Devine, 2015)
In most stereotype research, both from a cognitive perspective and in the field of social psychology, stereotypes are conceived of as being traits attributed to a particular group of people. However, it is sometimes the case that, rather than an attribute being activated cognitively by the activation of a group, a group can be activated by the activation of an attribute.

In 2015, Cox and Devine presented a model, based on a connectionist approach, in which cognitive activation can occur in different directions depending on the attribute and the group. These can be Group ⇒ Attribute, Attribute ⇒ Group, or the bidirectional Group ⇔ Attribute. The authors argue that all stereotypes are actually bidirectional to some degree, but most are primarily one direction or the other. The examples generally given in the paper are Black ⇒ Athletic versus Fashionable ⇒ Gay. Cox and Devine noted that they are specifically referring to men in both of these examples, though they continued to refer to them as a single attribute (rather than male and athletic) and a superordinate group (rather than Black male). They also conceded that groups could be considered to be a specific attribute, in that gay could be a group or an attribute, but they do not discuss this further, and I have not found it discussed in any other work. Through a series of large-scale studies, the authors test and confirm this theory of heterogeneous directionality.

In the first study, subjects were asked to list the first five things that come to mind "when you picture a man who _____". Each subject was given only a single stimulus item. The sentences were completed with one of the following, 20 stereotypic attributes of Black men, 20 stereotypic attributes of gay men, 17 attributes that are stereotypic of both groups, and the name of the 2 groups in question. In the results analysis, the authors computed a d-score for each link which they used as a measure of directionality, a positive score corresponding to a Group ⇒ Attribute direction, a negative score corresponding to an Attribute ⇒ Group direction, and a score close to zero being bidirectional. As hypothesized, the authors found that most of the links between attributes and Black men were stronger in the direction of Black ⇒Attribute, and most links between attributes and gay men were stronger in the other direction, even the same attributes. However, there were cases for both groups where the attribute and group were linked in the opposite direction than expected. This confirmed the hypothesis that stereotypes are not always in the direction of Group ⇒Attribute, and that this direction can depend both on the group and the attribute. A second experiment tested this with other groups and found the same patterns. Certain attributes activated the associated group more than that group activated the attribute, and vice versa. They also found that, similar to the difference between Black men and gay men, certain groups in general activated attributes more than they were activated by attributes (Black men) and vice versa (gay men). One of these latter groups was lesbians. The results indicated a number of attributes that were closely associated with lesbians, but most of these worked in an Attribute ⇒ Group direction.

In Minerva terms, a variety of probes are possible and may be more or less selective of a stereotype. The idea of directionality is essentially a question of what is functioning as a probe. Unlike many other stereotype experiments, those conducted by Cox and Devine do not ask subjects explicitly to provide a stereotype. In this case, they may in some ways to be accessing stereotypes (as defined earlier) more directly than

others. They are getting the connections people make, rather than individuals' explicit ideas of the stereotype. As the authors note, the gay men have more associations in an Attribute $\Rightarrow$ Group form than Black men, likely because ethnicity is generally more visible than sexual orientation. In an exemplar model this is explained simply by the fact that echoes of a probe will only be influenced by traces in the memory that are most similar to the probe. Traces of a gay man will not affect the echo if his homosexuality is not known (or assumed). This will be true of Black men as well, but a larger percentage of exemplars of Black men will be stored as such simply due to the fact that the ethnicity is visible in more instances than sexual orientation. Sexual orientation in women is generally accepted as less visible than in men. This explains Cox and Devine's findings and findings from other experiments and interviews (including my own), that stereotypes of lesbians are less strong and less clear than those of gay men. As discussed in the context of repeated exposure to exemplars, this process can become cyclical when visibility of a group leads to more homogenous perception and stronger stereotypes, which can then be used by members of that group for greater visibility.

Despite obvious overlap, the concepts of self-stereotyping and formation of out-group stereotype are rarely analyzed simultaneously as part of the same process. With the AWSA model, these two phenomena are simply variations of the process of activation of memory traces. Other exemplar-based models only address the perception of others, where traces are formed of other individuals. Despite the fact they most theories define memory traces as cognitive representations of experiences, very few discussions of exemplar theory have included the self as part of that experience. In describing his model of speech perception based on activation of exemplars, Johnson (1997) mentions that, "some speech experiences are of one's own speech, which presumably code not only auditory properties and categorical labels, but also articulatory properties. Therefore, an exemplar model can, in principle, also be used to give an account of the production-perception link." (p. 153). In addition to perceiving and storing memory traces of others, people perceive and store memory traces of themselves. This has not been discussed in any detail in existing literature, and no model has been presented, but there is no reason to believe that it differs from the model of perceiving others, or that the self-traces should be any different beyond being stored with the information that they are memories of oneself. Questions and research about self-stereotyping can be addressed through this model, and it allows for a direct comparison between the process of applying stereotypes to others and applying them to oneself.

2.2.3 Self-stereotyping

In order to incorporate self-stereotyping into a cognitive model it is important to first understand the theoretical base for most social psychological and cognitive models of how humans perceive and categorize themselves, the Social Identity Theory (SIT). Research on the topic of self-stereotyping focused initially on understanding the

mechanisms underlying intergroup relations and conflicts. In the 1970s and 1980s Tajfel and Turner developed SIT, which introduced the concept of a *social identity* to explain intergroup behaviors based on group status, discrimination, and social categorization. Social identity is defined as the identification individuals derive from being members of a social group, usually in comparison to another group.

This theory, though developed from a perspective of social psychology is based on concepts of self-perception and the role of context and comparison. These concepts are particularly critical in understanding the detailed functioning of an exemplar-based model, and further studies based on SIT have shown evidence for my own model in multiple ways.

The Social Identity Theory was developed as an account for in-group favoritism, and a means of predicting intergroup behavior based on perceived group status. It did not address the concepts of human social selves and how individuals identify and categorize themselves. The Self-Categorization Theory was developed by John Turner and colleagues, building off of ideas of SIT to create a general theory of how individuals act as a group, incorporating SIT as a single derivation.

The idea was introduced by Turner et al. in the 1987 publication *Rediscovering the Social Group: A Self-Categorization Theory*. Here they define *self-categorizations* as "cognitive groupings of oneself and some class of stimuli as the same (identical, similar, equivalent, interchangeable, and so on) in contrast to some other class of stimuli." (p. 44) These self-categorizations exist in a hierarchical structure, with three distinct structures, developed as equivalent to the concept of superordinate, intermediate, and subordinate levels of comparison. The superordinate self-categorization is that of the *human identity*, in which people compare themselves as humans to other forms of life. Below this, at the intermediate level, is the *social identity*, in which people define themselves as a member of a particular social group and not others. As part of the social identity, individuals often perceive themselves as non-distinct and interchangeable members of that group. The subordinate self-categorization is the *personal self*, in which people see themselves as unique individuals with a particular personality, and different from other group members. Turner et al. present the concept *depersonalization*, where the person's sense of self shifts from personal identity to social identity, and features shared with other members of a group or group stereotypes become more salient, leading to self-stereotyping. This theory as it is discussed in *Rediscovering the Social Group: A Self-Categorization Theory* forms the basis for a great deal of following research, including newer models and theories of self-stereotyping.

The AWSA model predicts how the basic concepts of SCT, personal vs. social selves, depersonalization, and self-categorization, result from perception of self and others through exemplars. When memory traces of the self are created, different dimensions receive more or less attention depending on context. This is seen clearly in

cases where aspects of the situation result in greater salience of dimensions that the individual shares with exemplars of other members of the in-group. A retrieval cue with more attention to these dimensions will activate a larger number of in-group members, including traces of abstract stereotypes. The resulting echo is an example of a social identity, whereas cues with greater attention weight on other, more individualizing dimensions will not activate other traces, resulting in a personal identity. This process of depersonalization allows for an infinite number of social identities, as well as personal identities, based on context and the social salience of particular dimensions.

The patterns seen in studies of lesbian stereotypes and lesbian identities indicate that depersonalization may be generally less common for members of the overall lesbian in-group. For example, an overview of historical perspectives of lesbian identity from within the community demonstrates disagreement over what it means to be a lesbian. The "butch" role or identity has been rejected by many as mimicry of heterosexual gender norms, but it is clearly adopted by many lesbians as an identity or temporary role (Newton, 1984; Kennedy and Davis, 1993; Farquhar, 2000; Eves, 2004; Huxley et al., 2013). Whether this identity should be defined as a social identity or a personal one depends in part on whether "butch" is defined as a subgroup of lesbians or an individual attribute. The most likely explanation is that this difference depends on situation and individual perspective. "Butch" (as well as "femme" and the relationship between these roles) is an excellent example of whether group labels can function as a dimension or attributes in the same way that vowel formants or hairstyle do, and how this affects issues of stereotyping both self and others.

One situational element determining the salience of an attribute and its potential role as an identity is the size and status of the associated group. This aspect of self-stereotyping was tested in two laboratory experiments by Simon and Hamilton (1994). The first experiment tested the effect of relative in-group size, that is, whether the subject is in the minority group or majority group. Subjects were led to believe that their ratings of a set of paintings placed them either in the majority (80% of all people prefer that painter) or the minority (10% prefer that painter). They were also told that liking that particular painter was related to being either an introvert or an extrovert. To avoid status effects, subjects were explicitly told that neither introvert nor extrovert is better. Subjects were then asked to rate themselves for three adjectives associated with introversion ('calm', 'careful', and 'rigid') and three associated with extroversion ('lively', 'social', and 'restless'). After this, subjects provided perceived similarity between themselves and other members of the group (people who preferred that painter) in those same six adjectives. Subjects then indicated their estimated homogeneity of the group, the extent to which they saw themselves as members of the group, *groupness* (the extent to which they saw the in-group (and out-group) as a psychological group), and then confirmed the manipulations saying that the in-group was more introverted (or more extroverted) than the out-group and that the various attributes were more typical of the intended introversion or extroversion. Results for the first experiment showed generally that

members of the minority group stereotyped themselves more. It was clear that this was not done simply as a means of presenting themselves in more positive terms, because subjects did not endorse negative out-group attributes. They also did not reject negative in-group attributes, which would have gone against self-stereotyping. Instead, by rejecting negative out-group attributes, and endorsing positive in-group attributes, members of the minority group presented themselves positively, while maintaining similarity with the group at large.

The second experiment added group status as an independent variable in a variation of the earlier experiment. In-status was manipulated by telling subjects that the painter was either "very distinguished" and highly regarded by experts in the field or "not very distinguished" and not highly regarded by experts in the field. The statement that neither introverts nor extroverts are better was removed. The authors predicted that high status of the in-group would increase self-stereotyping for minority groups members, but not members of the majority. This was confirmed in the results of the experiment. Increased self-stereotyping was found in high status minority group members for both positive and negative attributes. Self-categorizing (viewing themselves as members of the group) was also stronger with a higher status for minority group members but not for majority group members. This finding confirmed the hypothesized relationship between self-categorization and self-stereotyping, where "self-categorization underlies self-stereotyping, which in turn reinforces the former" (p. 709)

To expand on the research on self-stereotyping and self-categorizing, Verkuyten and Nekuee (1999) conducted a study that investigated this phenomenon outside of laboratory. Their experiment surveyed Iranians living in the Netherlands to determine how self-stereotyping, group identification, and group threat interact, and what affect they have on in-group bias. They measured each of these variables separately with questions that targeted how participants perceived threats to the group (group threat), whether they considered themselves to be similar to other members of the group (self-stereotyping), how positively they felt about being a member of the group (identification), and their overall evaluation of the group (in-group bias). Like Spears et al., the authors found that members who identified strongly with the group also self-stereotyped more, but only when they had a higher level of perceived group threat, though the reverse was not true (low identifiers did not distance themselves with greater group threat). They also found that the only predictor of in-group bias was self-stereotyping, but only for high identifiers.

Both of these studies demonstrate that self-stereotyping is distinct from, but not independent of, other variables such as identification and group threat. These findings support the notion of selective attention in forming memory traces and how different aspects can influence what dimensions of an experience are socially and personally salient. Level of identification and perceived threat (to status and distinctiveness, and likely other elements) will shift attention when creating memory traces of the self. When

individuals identify more strongly with a group, this overlap is shown to be particularly important when the group is under threat. Memory traces of the self will be stored with more attention to the dimensions that are salient to the group (particularly in inter-group relations). This will result in the self-stereotyping I described earlier, and predicts the patterns shown in research on the issue of group threat and group identification.

As this phenomenon applies to groups of sexual orientation, it provides insight into the differences in self-stereotyping between lesbians and gay men. Lesbians, according to accounts from Kennedy and Davis (1993), have less self-perceived homogeneity and a less self-stereotyping than is usually observed in gay men. Within the AWSA model, this reflects differences in perceived group threat and identification with the group. As a group, lesbians have historically been under less social threat than gay men. This is not a result of greater status or out-group acceptance so much as general invisibility. For various reasons, the group as a whole was generally ignored. One example of this is that laws about homosexuality tended to only reference men. Gay men, on the other hand, show generally more self-stereotyping, but many gay men also distance themselves by self-stereotyping less; a division that is shown in studies to result from greater perceived threat. The flexibility and confusability between social and personal selves demonstrates a complexity that is not completely captured by the original conception of SCT.

Andersen and Chen in 2002 expand on the concept of two levels of social identity, proposing a representation of personality existing at the level of relational identity, based directly on individuals' relationships with significant others. This term is defined by the authors as "any individual who is or has been deeply influential in one's life and in whom one is or once was emotionally invested. Relationships and attitudes towards significant others form the relational identity and affect interactions with new individuals through the phenomenon of *transference*. Research by these authors has shown that many aspects of subjects' interactions with new people seem to be influenced by the new person's similarity to certain significant others. For example, subjects were given descriptions of individuals that were designed to be similar to subjects' descriptions of significant others in their own lives. Descriptions were created to be neutral in valence, but attitudes towards the individuals were more positive if the descriptions matched ones of significant others towards whom the subjects had positive attitudes, than if they had negative attitudes towards the particular significant other. Another study by Andersen tested the effect of role congruence or discrepancy in transference. Subjects described a positively regarded significant other who was a role model or authority figure to them. Subjects then received a description of someone they were told was either an expert in a task to be done together, or a novice in the task. When the description matched the significant other and was introduced as an expert, the roles were congruent and subjects reported more positive moods. When the description matched the significant other and was a novice, there was discrepancy in the roles, and subjects experience a negative, depressed mood. When the description did not match the significant other, the level of expertise had no

23

positive or negative effect on the subject's mood. Andersen and Chen argue that an individual has multiple separate selves, each one a link between a significant other and the self, and that these identities are applied in an if-then model of personality. The ifs are situations in which a new person is encountered and, due to his or her similarity to a significant other, activates a representation of that significant other through transference. The thens are the responses to the new person based on the relational identity associated with that significant other.

The idea of a relational identity, just like a social identity can be understood as the activation of traces of others that are similar to the retrieval cue of the self. As in the process of stereotyping others, as argued by Wyer et al. (2002), this similarity is mediated by selective attention. Attention is determined by contrast to other exemplars. In the case of social identity, where individuals define themselves based on group membership, similar traces are ones created of other in-group members or the self in the group context; for the relational identity presented by Andersen and Chen, similar traces are ones of the self created in the context of a relationship with a significant other. The significant others that Andersen and Chen discuss are ones that received particular attention in the memory trace creation, or ones that are encountered so often that more traces of them exist in the memory than of most other individuals. For this reason, they will be activated with more strength and will have a larger influence on the context of the echo. Newly encountered individuals will activate traces of these significant others that are similar according to context and aspects of the new individual as they relate to the self. This breakdown of self-stereotyping as a relational identity has been incorporated more recently as involving the concept of self-anchoring

In 2016. van Veelen and collaborators presented a theory of overlap between self and in-group that integrates the concepts of self-stereotyping and self-anchoring as two separate cognitive pathways. Their Integrative Model of Social Identification (IMSI) differs from SCT models of social identification as a top-down process in which individuals assimilate to the stereotypes of their in-group. In many cases, the in-group does not have an established stereotype. If individuals cannot use these to identify as a member of the group, it raises the question of how they establish a social identification. The IMSI places the personal self in a predominant role and incorporates self-anchoring as a primary means of establishing self-group links, combining the two as parallel and overlapping processes. As the authors note, social psychologists disagree on which process is the default. Proponents of SIT and SCT place self-stereotyping as the default based are on the self-as-distinct model in which the personal self is what separates an individual from others, and that depersonalization is what connects an individual to the group. Proponents of self-anchoring as default base the argument on the concept of the personal self as "the locus of experience and thus a basic source of inferences about others" (p. 14). Van Veleen et al. reconcile these conflicting theories by demonstrating that the default depends on context and discuss the different group contexts that moderate which process is more likely to be used.

From the perspective of an exemplar model like AWSA, the results can be analyzed in the same way as those of Otten and Epstude (2006); when there are more or more similar traces of the self stored as a member of the group than other group member traces, an abstraction of that group will be primarily composed of traits of the self (self-anchoring), and self-stereotyping occurs when the opposite is true. Unlike earlier models, the IMSI allows for simultaneous self-anchoring and self-stereotyping, which is equally allowed in an exemplar model if one considers that the number of memory traces of the self in relation to traces of others is not binary but exists on a continuum. Likewise, context influences the probe, which determines which traces are activated most strongly in the echo. Thus, the AWSA model makes the same prediction as the IMSI as to when and why self-anchoring or self-stereotyping occur; they can occur simultaneously, and the strength of each is determined by awareness of the group (existing memory traces) and the context (cognitive probe).

Self-stereotyping and stereotyping others are at their core the same process, an activation of memory traces involving out- or in-groups and the individuals involved. This process of activation would imply that individuals are only aware of their own stereotypes and do not have access to the stereotypes that others hold. Most people would agree that this is untrue, including many of my interviewees asked to give stereotypes of lesbian speech. They often made statements to the effect of, "I don't think this, but there's a general stereotype that lesbians sound…". This idea of a "general stereotype" and individuals' awareness of it is referred to as *stereotype consensus*.

2.2.4 Stereotype consensus (Haslam & Wilson, 2000)

In their 1999 paper, Haslam and Wilson combined concepts of self-stereotyping with the application of stereotypes to others. The authors made use of the Self-Categorization Theory to explain how members of an in-group arrive at stereotype consensus. Previous research demonstrated that group members shared strong stereotypes of the in-group, but earlier theories explained stereotype consensus as developing from intergroup friction, with stereotypes being generally derogatory and applied by members of conflicting groups. The authors focused particularly on the SCT concept of fit, "the degree to which a social categorization matches reality" (p. 810). Self-stereotyping, was argued by the authors to enable in-group stereotype consensualization in three ways: 1) enhancing perceived in-group homogeneity, 2) establishing expected agreement between group members on issues of particular relevance to that group, and 3) pressuring members to reach a consensus of belief or attitude regarding those issues. Haslam et al. presented an experiment testing the assertion that people are more likely to develop a shared in-group stereotype if they identify and interact more as sharing a common social identity. The experiment was based on Australians as the social group, testing differences when subjects were asked to consider aspects of their identity either as individuals, or as members of that group. The results showed that with increased salience of social identity,

the stereotypes of the in-group were more similar to those of other members and generally more positive. This effect was enhanced when subjects were asked to complete the task in groups

In 2000, Haslam and Wilson tested the nature of the long-assumed relationship between prejudice and stereotypes and to what degree these two social judgments are personal or shared. The experiment involved three independent conditions in which Australian high school students used a checklist to indicate elements of their own or cultural stereotypes and beliefs about Aborigines. In all conditions, subjects checked all of the traits they believed to be part of the "cultural stereotype" of Aborigines and highlighted the five most central to the cultural stereotype. In both the Group and Group Interaction conditions, subjects stated whether they identified with a pro-Aboriginal or anti-Aboriginal group. Subjects in the Group condition then completed the same checklist according to the stereotypes which they believed other members of the pro- or anti-Aboriginal group held about Aborigines. In the Group Interaction condition, subjects with the same identification were clustered into groups of three or four. After a discussion between group members, subjects were asked independently to identify the beliefs of the group. Subjects in all conditions completed a checklist indicating their individual prejudice. As predicted, the results showed that in all cases, prejudice was more influenced by personal beliefs than cultural stereotypes. When they were elicited in the abstract, personal beliefs were not better predictors of prejudice than shared group stereotypes were but were better when they were elicited in the context of group membership. This study demonstrates that individuals, even as young as 15 years old, are aware of the idea of cultural stereotypes and that they may differ, perhaps drastically from their own personal beliefs. Although the authors do not mention this, the results also demonstrate that individual prejudice is not based on cultural stereotypes but rather on personal beliefs, particularly when they are shared by members of the relevant in-group. The in-groups that they chose were based on the issue being discussed. The personal beliefs shared by members of other in-groups not based entirely on the issue of prejudice may have relevance in different ways, such as the significance of their peer group with whom they likely making an effort to agree.

From the perspective that stereotypes are activations of existing memory traces, it must be the case that information is stored when the trace is created that allows individuals to recognize it as influencing a stereotype that does not overlap entirely with their own. This can be understood as perceived communicative intent, awareness that the "performances" people give, or ideas they express, express the stereotype. Cultural stereotypes, usually ones that are most accessible to individuals (other than their own), are reflected in explicit or implicit statements of shared stereotypes. They can also be reflected in performances that are imitations or exaggerated representations of members of a given group. Who shares/holds these stereotypes can be inferred from the context of intended audience. This allows people to give different responses, aware of who holds what stereotypes. People can have their own response to a probe based off of their

own memory traces and how they are weighted, but they can also have a response to a probe based off of exemplars shared by others, with potentially different weights. For example, they may have weak stereotypes about how lesbians speak because they have few exemplars, or the exemplars vary too much, but know that others (eg: society at large) hold stronger stereotypes because they have heard these stated, and/or less variation is shown in media, and exemplars that are shown exaggerate those features. In my own interviews, many people were able to give examples of stereotypes held by most people, or by society at large, but most said they did not share these stereotypes because the lesbians they knew spoke in very different ways and didn't use that style. Within the AWSA model the dimensions of a probe are determined by attention which can in turn be influenced by the many factors described above, among any number of others. A common dimension in stored exemplars within sociolinguistics is the question of gender and sexuality. In the following section I provide a background of research on lesbian speech situated in the more extensive area of gay speech research.

2.3 Research on language and sexuality

Research on language and sexuality has focused primarily on male speech (for a recent review see Lawson 2015). With the exception of a study conducted in 1983 (which I discuss below), female sexual orientation has only been investigated empirically as a sociolinguistic variable in the last two decades. While the production and perception of gay male speech is useful in understanding aspects of sociolinguistics overall, it potentially functions very differently from speech and sexuality in women. The "gay accent" is a very salient and recognizable speech style (see introductions of e.g, Campbell-Kibler 2011 and Mack and Munson 2012). Lesbian speech styles, on the other hand, are much less salient. Many people are not aware that they exist and even those who might think someone "sounds like a lesbian" would be hard-pressed to explain why they think so. Nonetheless, a few studies have investigated aspects of speech variation based on sexual orientation in women, and have found evidence of an existing stereotype about lesbian speech. The purpose of this study is to identify the patterns that exemplify this stereotype and better understand why such patterns occur, looking at sexual orientation as a complex identity that intersects with other personal attributes and experiences. Based on theories of intersectionality (Levon 2015), and incorporating exemplar based models of speech production and perception (Johnson 2006), I test the influence of such attributes and experiences through an in-depth analysis of individuals' sexual orientations and levels of familiarity with Queer culture in both production and perception of sexuality-based stereotypes.

The following sections provide an overview of previous research on speech and sexuality in women, and elaborate on the differences between production and perception research and their findings.

2.3.1 Production of lesbian speech

Existing research on phonetic variation in lesbians is described below. Studies have so far been limited to elements of pitch or fundamental frequency (f0) (Camp, 2009; Levon, 2011; Moonwomon-Baird, 1997; Van Borsel et al., 2013; Waksler, 2001) and the first two formants (F1 and F2) (Munson et al., 2006a; Pierrehumbert et al., 2001). Studies have ranged in speaker number from four (Moonwomon-Baird, 1997) to 102 (Van Borsel et al., 2013), and have collected a variety of speech types. They have also looked at speakers of multiple unrelated languages. The findings are inconclusive, which is perhaps not surprising given the variability in experiment design.

Speaker pitch (measured as fundamental frequency or a related perceptual scale) is the most frequently analyzed of any phonetic variable, included in the earliest as well as the most recent of the published lesbian speech studies. Published in 1997, Moonwomon-Baird's study is composed of two experiments, one production and one perception. Originally conducted in 1983 and 1984, the work looked at conversational speech in American English. In the production experiment, the author compared the speech of two lesbian speakers and two straight speakers in casual conversation. Compared to the straight speakers, lesbian speakers showed significantly lower pitch peaks and a general trend for smaller ranges which did not reach significance. The speech was collected primarily for use in the following perception study, and the subject pool of only four speakers makes it difficult to draw strong conclusions from the study, but the pattern supports a general perception of lesbians having lower voices. Also looking at pitch production in American speakers, Waksler (2001) compared twelve lesbian and twelve straight speakers retelling the Wizard of Oz. Waksler found no significant differences between the two groups, but there was a general trend for pitch range to be larger among lesbian speakers.

Three researchers have studied pitch in lesbian speakers of languages other than American English. Camp's 2009 dissertation on Japanese lesbian speech explored the average pitch height and width as well as the maximum and minimum pitch measured in semitones. Camp collected speech from interviews with twelve lesbian or bisexual women and seven straight women, and compared their speech based on average pitch height and width as well as the maximum and minimum pitch measured in semitones. She found that the average, minimum, and maximum pitch were all significantly lower in lesbian/bisexual speakers. The average width for speaker pitch was also smaller for lesbian/bisexual women, but this was not statistically significant. Levon (2011) examined pitch mean and range in interview speech as part of a larger project on language and sexuality in Israel. In this paper, rather than comparing the speech patterns of lesbians to straight speakers, he tested differences between two different groups of Israeli lesbians. The two groups were contrastive across the political spectrum and in their beliefs about the place of sexuality in the public sphere. The 'mainstream' group supports

incorporating homosexuality as part of current Israeli culture, while 'radicals' call for a reconfiguring of Israeli culture to successfully incorporate homosexuality. Levon found that speakers in both groups used a significantly higher mean pitch when talking about topics relating to sexual orientation than when talking about unrelated topics. Van Borsel, Vandael, and Corthals (2013) looked at read speech from 68 straight and 34 lesbian speakers of Brazilian Portuguese. The study took into account aspects of subjects beyond sexual orientation such as their job, whether they were in a committed relationship, etc., including the age of "coming out" for lesbian participants. The authors measured mean pitch, as well as the variation, calculated as the sum of absolute value of all F0 changes divided by the duration of the utterance. They found that lesbian subjects used both significantly lower pitch averages and less pitch variation compared to straight subjects. Age was the only other attribute that correlated with these features, but it was independent of sexual orientation. Though not all research on pitch variables has shown consistent variation, those that have, generally show an overall trend for straight speakers to use a higher average pitch and wider pitch range compared to lesbian speakers.

Production studies concerning vowel formants are fewer in number than those that focus on pitch and show a general pattern of lower formants in lesbian speakers. In 2004, Pierrehumbert, Bent, Munson, Bradlow, and Bailey recorded a total of 102 men and women in the U.S. divided into sexual orientations of either LGB or straight. Subjects read sentences which were then compared for F1 and F2 differences. The authors found significant differences in specific vowels. Lesbians used significantly lower F1 only in /u/ and /ɑ/, and lower F2 only in /ɑ/. In a study of 11 LB and 11 straight women, Munson, McDonald, DeBoe, and White (2006), using single words rather than sentences, found a similar pattern but with different vowels. The results showed that lesbians used significantly lower F1 in /ɛ/ and lower F2 in /oʊ/. In 2008, Rendall, Vasey, and McKenzie looked at read speech from 33 straight women and 29 lesbians reading three types of stimuli: isolated vowels, single-syllable words, and short sentences. The authors found differences between the two groups, also only in certain vowels. Five of the vowels (/i/, /ʌ/, /ə/, /ɪ/, /oʊ/) were lower in one or a combination of F1, F2, or F4, with only one (/æ/) having a higher F1.

2.3.2 Perception of lesbian speech

Very little can be drawn from an overview of previous literature, primarily because each study differs from the others in multiple ways, from language, to number of speakers, to the type of speech recorded. What is clear is that lesbians do not differ from straight speakers in reliable ways. In all studies there is an overall trend for lower pitch and lower formants, but the fact that formants differ only for specific vowels suggests that these patterns apply differently for different groups and for different reasons, in line with findings throughout sociolinguistics of different linguistic styles showing variation

in specific phonemes, such as greater /t/-release in high school "nerd girls" (Bucholtz 2001).

In addition to looking at speech production according to sexual orientation in women, researchers have investigated how phonetic variation influences listener judgments of speaker sexual orientation. Some of these studies were carried out in tandem with the production, so there is considerable overlap in the phonetic variation under scrutiny. These perception studies focus on pitch and vowel formants, with one additional study looking at center of gravity in /s/.

In the perception experiment of Moonwomon-Baird's study, 21 listeners heard a series of recordings of six lesbians and six straight women (different from the four speakers in the production study), and filled out a questionnaire eliciting judgments about personal attributes of each speaker: class, age, educational background, region of upbringing, ethnicity, and sexual preference as well as phonetic variables: rate of speech, pitch, amplitude, and "forcefulness". The author found no correlations between the perceived sexual orientation and perceived acoustic attributes of the voice or the perceptions of other social aspects of the speaker, with the exception of a positive correlation between listener judgments of "lesbian" and "Jewish". She did find that speakers were much less likely to be judged as lesbian than as straight. This, in addition to comments by listeners, indicate an apparent discomfort making a determination about a speaker's sexual preference and specifically in labeling a speaker as lesbian. Waksler (2001) examined perception as well but in a way that is rare among studies of sexual orientation. The study analyzed pitch in recordings of four film actresses portraying both a lesbian character and a straight character within the same two years. In comparing the pitch ranges for each actor in her lesbian role and her straight role, Waksler found that for all four women, the straight character had a wider range, though there was a wide diversity among speakers overall. Waksler concluded from this that the stereotype of lesbians having a smaller pitch range is shown in portrayals of lesbians on film. Camp's 2009 study of speech and sexuality in Japan also included a matched guise perception experiment. Fifteen participants listened to recordings of read speech from one male and one female. Pitch in recordings was manipulated to create three different heights: normal, high, and low, and three different widths: natural, wide, and narrow. Listeners rated speakers on five seven-point scales: "less attractive/more attractive", "calm/emotional", "masculine/feminine", "unassertive/assertive", and "heterosexual/homosexual". For the female voice, Camp found that both height and width were significant, a lower pitch and narrower range being perceived as more homosexual.

Two studies have looked at perception as it relates to phonetic variables other than pitch. The 2006 study of formant variation by Munson et al. looked at how F1 and F2 influenced listeners' perceptions of sexual orientation. 40 listeners rated speakers' perceived sexual orientation on a five-point scale. Participants listened to recordings of the 44 speakers recorded in the production experiment (22 of each gender divided evenly

between LGB and straight). Listeners rated each speaker four times based on four types of words which differed in the presence of particular phonemes. Overall, listener ratings corresponded accurately with the sexual orientation of the speakers. The greatest predictors of listener ratings were vowel formant frequencies. For low-front vowels, a high F1 frequency correlated with a higher straight rating for women. Also in 2006, Munson, Jefferson, and McDonald investigated the possibility that the perceived sexual orientation of a speaker can influence how the listener perceives the sounds. It built off of research by Strand and Johnson (1996) who found that listeners identified ambiguous fricative sounds differently according to the gender of the speaker. Munson et al. created a nine-step sibilant continuum from a clear /s/ to a clear /ʃ/. They used tokens of /æk/ and /ɪp/ taken from the same 44 speakers recorded in the earlier study to create stimuli of *sack – shack* and *sip – ship* continua for each speaker. Ten listeners then heard in a series of two /s/ – /ʃ/ identification tasks and identified each token as either *sip* or *ship* or *sack* or *shack*. Based on previous ratings, the authors grouped the speakers as sounding more or less homosexual and more or less feminine/masculine. By analyzing the sibilant perception according to the grouping of each speaker, Munson et al. found that there were significant effects of speaker gender, perceived sexual orientation, and acoustics of the following vowel. The results showed that female speakers sounding homosexual or less feminine elicited more /s/ percepts than straight or feminine women.

The results of perception studies have mirrored those of the production studies but with somewhat more consistency. Listeners judge speakers according to the same stereotypes: lower average pitch and smaller pitch range, as well as lower vowel formants (for certain vowels) made a speaker more likely to be judged as sounding like a lesbian. Though this was not studied in production, a lower center of gravity for /s/ was also associated with sounding like a lesbian. There is clearly some general stereotype about the way lesbians sound, and this stereotype is used (in some situations) by lesbians more than by straight women. The details of this are rather unclear, and my research aims to probe the reasons for the use of this speech stereotype to project and perceive a lesbian identity, as well as the relationship between this stereotype and others.

2.4 Conclusion
In this chapter I have presented a model of stereotype construction and access in which abstract concepts ranging from phonemes to social stereotypes are all underlying the same format: the set of related exemplars that are activated in the Primary Memory by a probe. With this model, I analyzed previous findings from social and cognitive psychology on self-identification and group stereotyping, and have provided background on the existing research regarding speech and sexuality in women.

Past studies on lesbian speech have covered a range of languages, speech modes, and number of participants. However, the small number of studies published so far shows

an unclear picture of the relation between sexual orientation and speech in women. Some show that pitch is significant, but others indicate the difference is only a general tendency. The first and second formants show statistical significance, but only for certain vowels, and these vowels differ by study. Only one study has explored phonetic variation in fricatives. The paucity of data does not allow for any conclusive picture of what a lesbian speech style is, whether and to what extent listener stereotypes accurately reflect patterns in speech, and what other variables such as perceived femininity or self-identified bisexuality of speakers contribute to a general stereotype of sexual orientation and speech.

In the following chapter I present an experiment, Experiment 1, in which I recorded and analyzed the speech of lesbian, bisexual, and straight women. This experiment tests a broader set of phonetic features than have so far been studied and includes in the analysis multiple personal attributes of the speakers beyond sexual orientation.

3 Speaker stereotype use: How sexual orientation is performed through speech

3.1 Research questions and hypotheses
This study expands upon methods from earlier research presented in Chapter 2 in three important ways, 1) by analyzing a wide range of previously unexplored phonetic features, 2) by testing sexual orientation beyond the binary of straight and lesbian, and 3) by incorporating intersections with personal experience and attitude as well as demographic information.

The results of earlier research on phonetic variation and sexual orientation in women are varied but point to some overall patterns of lesbian speech, specifically a lower f0 as well as lower first and second formants overall or within specific vowels.

The results of previous research interpreted through the AWSA model lead to three primary hypotheses:

- Hypothesis 1: Speaker variation will not pattern exclusively according to self-identified sexual orientation.

- Hypothesis 2: Familiarity with Queer culture will be one of these other personal and social attributes contributing to phonetic variation.

- Hypothesis 3: To the extent that phonetic variation is predicted by sexual orientation, bisexual speakers will not pattern with lesbian speakers.

3.2 Methodology

To test these hypotheses, I collected read and interview speech from self-identified lesbian, bisexual, and straight women. I performed acoustic analyses on these recordings in search of patterns of phonetic variation relating to personal and social attributes.

3.2.1 Participants

Recordings were made of 64 self-identified women from Berkeley, CA and the surrounding area. Due to technical issues and experimenter error, only 54 of these recordings were usable. Participants were solicited through various forms of advertisements and announcements (around the UC Berkeley campus and surrounding area, online, in class, and by word-of-mouth). A copy of the recruitment flyer is given in Appendix A. Previous studies have focused on sexual orientation as a binary either between straight and lesbian or straight and lesbian/bi.[3] This binary is particularly problematic for women as demographic research has shown that more women identify as bisexual than lesbian (Gates, 2011). Moreover, though previous studies incorporating bisexual speakers group them with lesbians, interviewees in my own research generally agree that bisexual women are stereotyped as sounding more similar to straight women than lesbians. Therefore I chose to expand the subject pool to include bisexual speakers and analyzed them as a unique group. Participants self-identified as either lesbian (N=12), bisexual (N=18), or straight (N=24).

Participant demographics were collected in a questionnaire asking for their sexual orientation, the number of their family members or friends who identify as homosexual or bisexual, and their self-reported familiarity with Queer culture, as well as their age, ethnicity, city of origin, and city of current residence. The cities were categorized into dialect regions according to Labov et al. (2006). Ethnic identities were given as options taken from the racial/ethnic categories in the US Census. Familiarity with Queer culture was given a 5-point scale from 1 = "not at all familiar with Queer culture" to 5 = "extremely familiar with Queer culture"; neither *familiarity* nor *Queer culture* were defined for speaker and were instead left open for interpretation. Possible definitions and it's effect are described in Section 3.3.1.3. All of these answers were considered as potentially influencing phonetic variation. I refer to these answers (e.g., sexual orientation) from here on out as a speaker's *social and personal attributes*.

Participants' ages ranged from 18 to 54, but since most participants were university students the numbers were skewed towards ages between 18 and 21, with only six participants above the age of 30. Participants self-identified as Asian, Black, Hispanic, and White, as well as combinations of any of those four. Speaker recruitment was primarily aimed at recording as equal an amount of speakers for each orientation as

---

[3] The exception to this is Levon's (2011) study of speech variation in two different groups of Israeli lesbians.

possible. As a result, the distribution of other attributes, such as ethnic identity and region of origin was skewed across sexual orientation. This distribution can be seen in Tables 3.1 and 3.2 respectively.

Ethnic identity	Bisexual	Straight	Lesbian
White	12	6	7
Asian	3	14	1
Hispanic/White	1	1	3
Hispanic	0	2	0
Black	1	1	0
Asian/White	0	0	1
Black/White	1	0	0

Table 3.1: Speaker sexual orientation according to ethnic identity

Region of origin	Bisexual	Straight	Lesbian
Canada	1	0	0
Inland North	1	0	0
Mid-Atlantic	0	0	1
New York City	0	1	1
St. Louis Corridor	0	1	1
Texas South	0	0	1
The South	1	0	0
The West	13	10	19
West New England	2	0	0

Table 3.2: Speaker sexual orientation according to dialect region of origin

3.2.2 Stimuli

The recording procedure was split into two reading blocks and an interview. The first was a word reading block using a set of 100 words. The entire set of 200 words was

presented in random order. These words, taken from Tracy et al. (2015) were 1-3 syllable English words chosen for phonetic balance. The entire list is given in Appendix B. They provide multiple examples of a range of phonemes, consonant clusters, and a variety of spoken word frequencies (all frequent enough to be familiar to speakers). The second block of reading showed sentences and questions in random order, also twice each. Sentences were ten simple structured utterances ranging from seven to nine words in length. Questions were five polar (yes/no) and five content questions (wh-) ranging from seven to ten words. Examples of these are shown in 1-3 and the entire list is given in Appendix C.

1)   After the dance they went directly home.

2)   How do ducks fly North with no compass?

3)   Is that medicine better than the one I have?

3.2.3 Procedure

Based on the hypothesis that a speaker's background and familiarity with Queer culture influence ways in which certain speech stereotypes are used, this experiment was designed to cue participants in on the topic of speech and sexuality in women, where the influence of familiarity and sexuality would most likely be found. In order to best understand how sexual orientation and familiarity influence speech style, the purpose of the study was intentionally made clear to participants. Intra-speaker variation is expected, and most theories of sociolinguistics hold that variation within individuals is based on multiple factors, including interlocutor, stance, and topic. In the case of this study, while the "topic" of the speech was limited to a list of words and sentences, the topic of the interaction at large involved questions of sexuality. The recruitment flyer for the experiment advertised it as a study of speech and sexuality in women, and the questionnaire including the self-rating of familiarity with Queer culture, as well as sexual orientation and other related questions was filled out by participants before recording began.

After completing the questionnaire, participants sat at a desk in a sound attenuated booth in front a screen with word and sentence prompts being displayed using OpenSesame (Mathôt, Schreij, and Theeuwes 2012). Audio recording were conducted using an AKG 535 EB microphone positioned approximately 8 inches from the speaker, slightly to the right to avoid aspiration 'pops'. Recordings were made at a 48Hz sampling rate and digitized with a Steinberg UR22 USB audio interface. A Canon FX100 video camera, positioned behind the screen of the teleprompter simultaneously recorded video and audio with an internal microphone for use in future video-based research. Audio files were stored on a Dell Optiplex desktop PC. In the first two blocks, speakers were given the choice of being joined in the sound booth by the researcher. Most expressed no

preference, in which case the researcher remained in the booth in case any issues or questions came up. The optional third part of the experiment was a discussion/interview regarding views of a lesbian speech style. A total of 37 subjects participated in the interview. This conversation with the researcher elicited a more natural speech style as well as whether participants were aware of lesbian speech stereotypes and what they considered those to be.

3.2.4 Measurements and statistical analysis

Transcripts were made of the recorded speech using the order of words and sentences saved in the experiment program OpenSesame. Transcripts were aligned with recordings using an implementation of the Penn forced aligner (Yuan and Liberman 2009). The automatic alignments were visually checked in Praat (Boersma 2001), and major misalignments were hand corrected. Measurements were made of phonetic features with more detail extracted in R. These phonetic features were chosen based on findings from previous research described in Section 2.3 and below, as well as answers from participants in interviews. They were mean pitch, pitch range, means of F1 and F2, formant values for each phoneme, vowel space, rate of speech, fricative spectral balance, /t/ release, and creak, and will henceforth be referred to as *phonetic features*, variation of which will be referred to as *phonetic variation*.

Pitch and vowel formants have been the primary focus of previous research on speech and sexuality in women. Pitch in particular, both average and range, has been shown in multiple studies to vary according to self-identified and/or perceived speaker sexual orientation. F1 and F2 have also been found to vary, though this has only been researched in perception experiments. The fundamental frequency and the first two formants of each vowel token were measured at the midpoint using the Inverse Filter Control method (Watenabe 2001). Rather than using the Hertz values, which do not reflect the auditory representation of vowel formants, measurements were converted to the Equivalent Rectangular Bandwidth (ERB) rate using the function:

$$ERB = 21.4*\log_{10}(4.37*(f/1000)+1)$$

This auditory frequency scale is similar to Bark or other functions that allow researchers to measure a variable more similar to how vowels are processed through the auditory system. The means of the pitch and formants were calculated for the primary and secondary stressed vowels separately. The median was taken for each vowel quality (/i/, /ɪ/, /e/, /ɛ/, /ə/, /ɚ/, /æ/, /ɑ/, /o/, and /u/) from which the overall mean was calculated. Diphthongs /aʊ/, /ɔɪ/, and /aɪ/ were not included for this calculation. Pitch range was measured as the interquartile range for all values. This, rather than the total range, was used to avoid extreme values not representing speakers' normal ranges. The vowel space for each speaker was calculated using the function "convexhull.xy" in the R package "spatstat" (Baddeley, Rubak, and Turner 2015). Given a set of points with x and y values

(F1 and F2) the function calculates the convex hull, providing a measure of the overall size of the vowel space. Outlying speakers with median pitch values below 100ERB were not included in the analysis of mean pitch and pitch range, removing pitch-related data from three speakers in this experiment.

Speech rate has not been studied to any great degree as a sociolinguistic variable, but it was noted by multiple participants in Experiment 1b to be an element of stereotypical lesbian speech. In work that has studied speech rate as a sociolinguistic variable, men were shown to use faster speech (Salmons, Jacewicz, and Fox 2008; Yuan, Liberman, and Cieri 2006). Based on its demonstrated association with gender and its perceived association with lesbian speech, the rate of speech was included as a variable to be analyzed and was measured as a function of vowel length. For each vowel, the time of the vowel onset was subtracted from the time of the vowel offset. The median length was measured for each vowel quality and the mean of the medians indicated the each speaker's overall rate of speech.

Variation in /s/ production has also been included in studies of language and sexuality. Though only one study has investigated variation based on sexuality in women, it was shown to reflect perceived sexual orientation. It is a prominent element of the stereotype of a "gay lisp" in men (e.g., Campbell-Kibler 2011; Mack 2011; Mack and Munson 2012; Munson 2007; Munson, McDonald, DeBoe, and White 2006a; Tracy, Bainter, and Satariano 2015), and though it is not expected that women will necessarily pattern in the opposite way, previous findings from Munson et al. (2006a) indicate that it may feature in a lesbian speech stereotype as well. For this reason, as well as its association with gender, variation of /s/ was included in the analysis. The fricative spectral balance was calculated as the ratio of the summed amplitude in the Mel frequency spectrum above 300 Hz and below 2,400 Hz relative to the summed amplitude in the spectrum above 2,400 Hz. The Mel spectrum was calculated at the midpoint of fricatives using the 'melspec' function in the ESPS speech signal processing system (Sprouse 2014). This is a relatively robust way to measure spectral tilt (Forrest, Weismer, Milenkovic, and Dougall 1988).

Though it has never featured in studies of language and sexuality in women, /t/ release has been shown in multiple studies to have associations with gender and style and with sexual orientation of men (for a review, see Eckert 2008). Release of word final /t/ was measured by extracting all words with final /t/ - based on the Praat TextGrid alignments. The words (including 50 ms beyond the marked end of the word) were then played and judged by ear to either be released or unreleased. The data entered into the analysis is the proportion of released /t/s for each speaker.

Finally, creakiness of speakers was also examined. Creak is a relatively new variable in sociophonetic research, particularly in American English (for a review, see Podesva 2013). It has been found to be more prevalent in women, particularly young

women. It is also shown to be strongly associated with toughness and an authoritative stance, both of which are generally considered masculine traits and were mentioned by multiple participants to feature as part of a lesbian stereotype. This simultaneous association with masculine and/or lesbian traits while being used primarily by women, indicated a potential link with female sexual orientation. The amount of creaky voice for each speaker was measured using the artificial neural network described by Drugman, et al. (2012). Their matlab script calculates a number of acoustic features and is trained on a database of speech that has been tagged for creaky-voice.

3.3 Results and discussion

3.3.1 Read speech

Speech variation was modeled to understand which phonetic features were salient as sociolinguistic variation, as well as which personal attributes these variables expressed. These were run as linear regression models for each acoustic feature, with the feature as the dependent variable and the various personal attributes from the initial questionnaire as independent variables. Tables 3.3 – 3.5 show the output of such a model using familiarity with Queer culture, sexual orientation, and ethnic identities as potentially determining the mean speech rate either independently or in interaction with one another. Of the many personal attributes, familiarity with Queer culture was the only one to independently predict acoustic variation.

3.3.1.1. Ethnic identity and region of origin

Before I turn to familiarity and orientation effects. I want to address the issue of race. As mentioned in the section on methodology, speakers were not equally distributed across ethnic identity and region of origin. It was of particular concern that despite there being a large number of Asian speakers, the vast majority of them were straight and over half of the lesbian speakers were white. There were no models in which the acoustic variable was best predicted by ethnic identity alone, but there were three acoustic variables for which there was a significant or near significant interaction between ethnic identity, familiarity, and sexual orientation. The following tables shows the results of ANOVAs for speech rate (Table 3.3), mean F1 (Table 3.4), and proportion of /t/ release (Table 3.5). I also ran a small experiment testing the perception of race, which demonstrated that variation by ethnicity was not significant to listeners. This experiment is discussed in Appendix D.

anova(lm(**vdur ~ S.familiarity * Sexuality * Ethnic.ident.,** data = all.df))					
Source	df	SS	MS	F	P
Familiarity	3	0.0025211	0.00084036	3.1298	0.04137 *

Sexuality	2	0.0000450	0.00002248	0.0837	0.91991
Ethnicity	6	0.0009768	0.00016280	0.6063	0.72303
Familiarity:Sexuality	6	0.0013066	0.00021777	0.8111	0.57023
Familiarity:Ethnicity	4	0.0007130	0.00017825	0.6639	0.62233
Sexuality:Ethnicity	2	0.0000069	0.00000344	0.0128	0.98727
Familiarity:Sexuality:Ethnicity	*1*	*0.0009346*	*0.00093460*	*3.4808*	*0.07259 .*

*Table 3.3: ANOVA of speech rate ~ familiarity * sexual orientation * ethnic identity*

anova(lm(**menf1 ~ S.familiarity * Sexuality * Ethnic.ident.,** data = all.df))					
Source	df	SS	MS	F	P
Familiarity	3	2304	767.9	0.4659	0.7084
Sexuality	2	912	456.1	0.2767	0.7603
Ethnicity	6	7358	1226.4	0.7440	0.6190
Familiarity:Sexuality	6	10809	1801.4	1.0928	0.3909
Familiarity:Ethnicity	4	6503	1625.7	0.9863	0.4311
Sexuality:Ethnicity	2	3620	1809.8	1.0979	0.3475
Familiarity:Sexuality:Ethnicity	1	*7510*	*7509.8*	*4.5559*	*0.0417 **

*Table 3.4: ANOVA of mean F1 ~ familiarity * sexual orientation * ethnic identity*

anova(lm(**tprop ~ Sexuality * S.familiarity * Ethnic.ident.,** data = all.df))					
Source	df	SS	MS	F	P
Familiarity	2	0.09948	0.049741	1.2080	0.3139
Ethnicity	3	0.05707	0.019025	0.4620	0.7110
Sexuality	6	0.20550	0.034250	0.8318	0.5555
Sexuality:	6	0.19920	0.033200	0.8063	0.5736
Sexuality:Ethnicity	4	0.03508	0.008770	0.2130	0.9290
Familiarity:Ethnicity	2	0.00572	0.002862	0.0695	0.9330
Sexuality:Familiarity:Ethnicity	*1*	*0.14692*	*0.146922*	*3.5682*	*0.0693 .*

*Table 3.5: ANOVA of proportion of /t/-release ~ familiarity * sexual orientation * ethnic identity*

3.3.1.2. Sexual orientation

As hypothesized, no phonetic features measured in the study varied exclusively with speaker sexual orientation. Even average pitch and pitch range, which have been found in some previous studies to be associated with a lesbian identity (Van Borsel et al. 2013; Moonwomon-Baird 1996; Camp 2009; Waksler 2001), did not vary according to sexual orientation of speakers in this study. With the exception of familiarity with Queer culture, other social and personal attributes, such as ethnic identity of area of origin, also showed no direct correlation with any acoustic feature.

3.3.1.3. Familiarity with Queer culture

Participants were asked to rate their "familiarity with Queer culture" on a scale of 1 (Very Low) to 5 (Very High). The interpretation of this was left open to speakers. Possible definitions from participants and other researchers include:

- Ability to recognize in-group references, jokes, and stereotypes
- Knowledge of LGBTQ magazines, podcasts, and other media
- Knowledge of/membership in LGBTQ-related groups
- Participation in pride and LGBTQ-related events
- Having many or primarily self-identified LGBTQ friends

Of the various information collected from the speaker, this self-stated familiarity with Queer culture was the only one to show direct correlation with phonetic variation of the individual's speech. A higher familiarity with Queer culture is correlated with a lower mean pitch ($F(1,51) = 9.499$, $p = 0.004$), as well as a faster rate of speech ($F(1,51) = 9.354$, $p = 0.004$). These are illustrated in Figures 3.1 and 3.3.2 respectively.
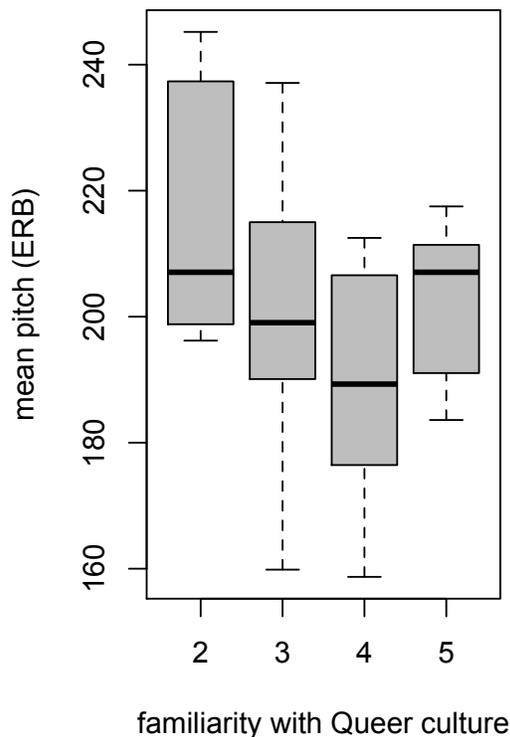
Figure 3.1: Familiarity and mean pitch
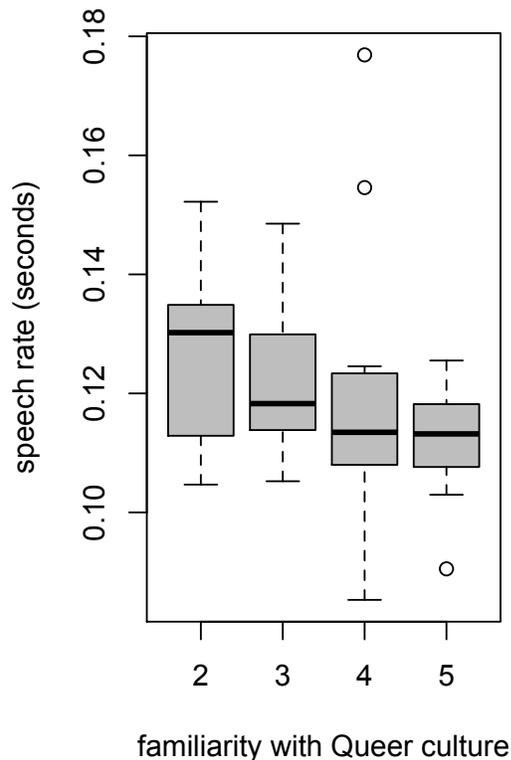
Figure 3.2: Familiarity and speech rate

While there was no direct correlation between sexual orientation and any phonetic variation, there was an interaction between sexual orientation and familiarity with Queer culture for both mean pitch and speech rate. As reported above, the mean pitch is lower and duration of vowels is shorter as the familiarity of the speaker increases. The pattern is somewhat different when speakers are analyzed separately by sexual orientation. The straight and bisexual speakers follow the same overall pattern, while the lesbian speakers show no correlation between phonetic variables and familiarity with Queer culture. Table 3.6 shows correlations between familiarity and selected acoustic features. This includes all speakers combined as well as analyzed separately according to sexual orientation. Figures 3.3 and 3.4 illustrate the different correlations between familiarity and mean pitch (Figure 3.3) and speech rate (Figure 3.4) when divided by sexual orientation. In Experiment 1a) familiarity showed a strong positive correlation with mean F2 in lesbian speakers. As shown in Table 3.7, this was also true in 1b) and expanded to influence F1. The correlation, particularly the direction, is unexpected and introduces new questions. However further investigation is beyond the scope of this dissertation.

Pearson's Correlations – R Between Speaker Familiarity and Phonetic Variables in Read Speech				
	Speaker Familiarity			
	Familiarity	Familiarity straight	Familiarity bi	Familiarity lesbian
mean pitch	-0.3962542**	-0.5459264**	-0.2385528	-0.2344672
pitch range	0.05819061	-0.05472472	0.300981	-0.01127752
mean F1	0.005771557	-0.1295928	-0.1464746	0.2785556
mean F2	0.069333	-0.3142536	0.03312854	0.7850122 **
speech rate	-0.396741**	-0.5011549 *	-0.4238789 .	-0.04331858
/t/ release	-0.04263515	0.02442732	0.002079029	-0.5169945 .

Table 3.6: Correlations of familiarity with Queer culture with acoustic variables in Exp 1a
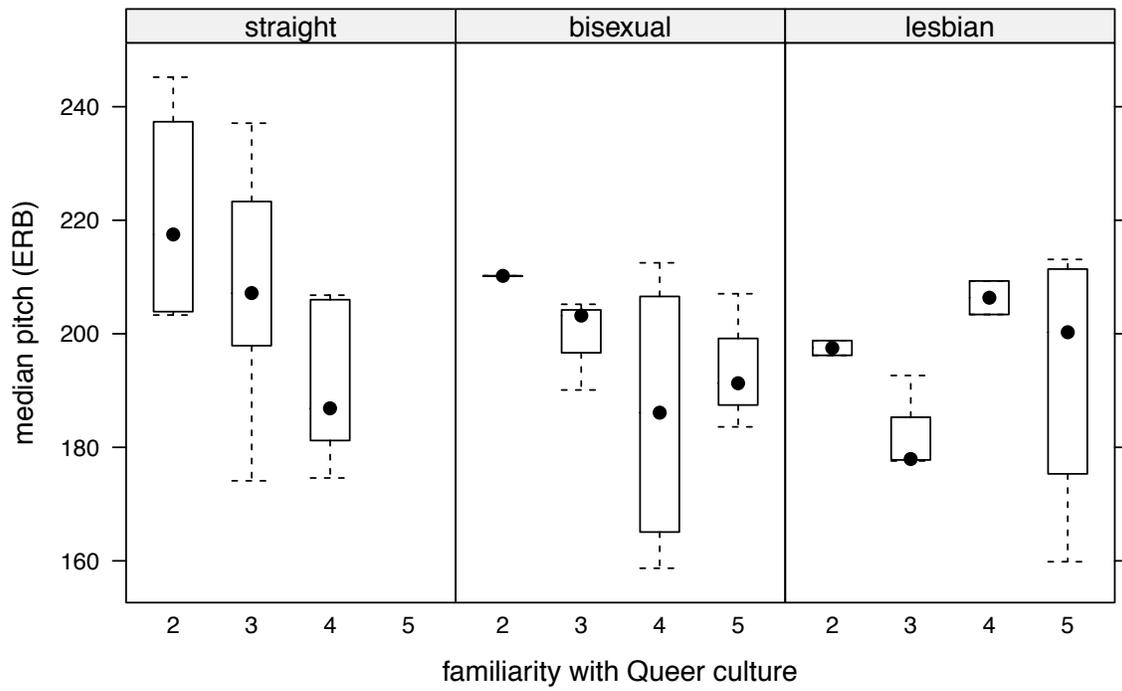
Figure 3.3: Familiarity with Queer culture and mean pitch divided by sexual orientation
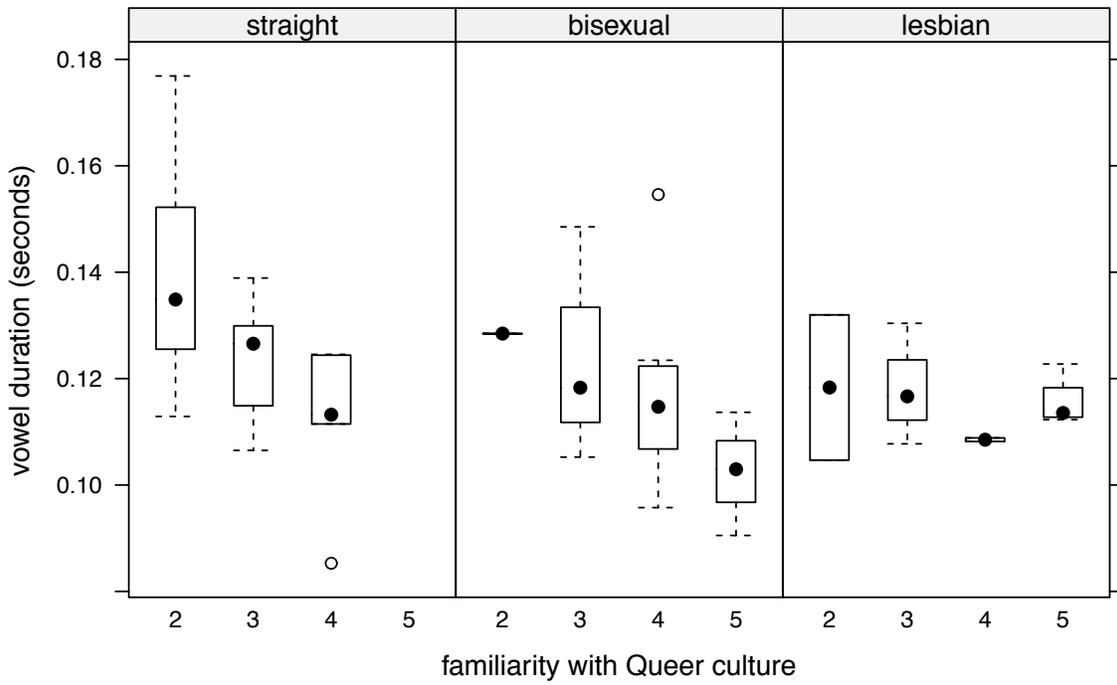

Figure 3.4: Familiarity with Queer culture and speech rate divided by sexual orientation

3.3.2 Interview speech

Differences between speakers were also analyzed for speech patterns in interviews. Phonetic variables measured in read speech were also measured in interview speech. The process of analysis for Experiment 1b) were identical to that of Experiment 1a).

3.3.2.1. Sexual orientation

Interview speech followed many of the same patterns as read speech. As with read speech, the self-identified sexual orientation of speakers did not function independently as a determining factor in phonetic variation between speakers.

3.3.2.2. Familiarity with Queer culture

Speech in interviews did not vary as much according to familiarity. Though it did not reach significance, there were minimal correlations between familiarity and mean pitch ($F(1,32) = 3.542$, $p = 0.069$) as well as mean F1 ($F(1,32) = 3.698$, $p = 0.063$). For straight speakers, the effect of familiarity on acoustic variation either decreased in significance, as with mean pitch ($F(1,14) = 3.352$, $p = 0.089$) or disappeared entirely, as was the case for speech rate. This, of course meant a change in effect for the speakers as a whole.

Pearson's Correlations – R Between Speaker Familiarity and Phonetic Variables in InterviewSpeech				
	Speaker Familiarity			
	Familiarity	Familiarity straight	Familiarity bisexual	Familiarity lesbian
mean pitch	-0.3156675 .	-0.439531 .	-0.406357	-0.2299402
pitch range	-0.1982	-0.01225115	-0.6166226 .	0.3182643
mean F1	0.2830724 .	-0.01370501	0.4224063	0.7259502 *
mean F2	0.0763164	-0.4836136 .	-0.131408	0.9266372***
speech rate	-0.2330613	0.04995079	-0.4119806	-0.4040777

Table 3.7: Correlations of familiarity with Queer culture with acoustic variables in Exp 1b

These results further support the hypothesis that individuals are using speech stereotypes as a means of asserting affiliation with but not identity within a minority culture. Had lesbian and bisexual speakers, for example, been using a different style of speech from straight speakers, we would expect the effects of this to be greater in less monitored speech, as is so often the case with regional, ethnic, or socio-economic linguistic variation with similar experimental methodology (e.g., Labov 1967 et alia). In this case, subjects are using this style more consistently in read speech, when they are more conscious of the way they sound. In interviews, when subjects are more concerned

with what they're saying than how they're saying it, the performance of an ally identity through speech is no longer important.

3.4 Conclusions

Variation in the phonetic features that were measured did not correlate directly with the self-identified sexual orientation of the speakers. It was not the case, as in some previous studies, that lesbians as a group tended to speak with a lower pitch or use different F1 or F2, nor was there any correlation with phonetic features not previously studied. However, although sexual orientation does not directly predict variations in speech production, it does, as hypothesized, interact with other social attributes. The main difference between members of different sexual orientations is not the way they sound, but the way their familiarity with Queer culture influences how they sound. In an interaction targeting questions of sexuality, it would seem that sounding familiar with Queer culture is less of a concern for lesbian speakers. Straight and bisexual speakers who are more familiar with Queer culture appear to be using phonetic variables to express this through speech. These features and their patterns are ones that have been shown in past research and/or expressed by participants during interviews to be features of a lesbian speech stereotype. While "familiarity with Queer culture" is a general and abstract concept, open to interpretation by participants, it is clearly meaningful, and almost certainly involves affinity with that larger group and interactions with people who identify as Queer. For people who do not identify as Queer but associate with that group, taking on aspects of the speech stereotype may be a useful way of expressing that association. The loss or reduction of this effect in interview speech further supports this argument. In the more monitored speech of reading words and sentences, speakers will attend more to style; in interviews, the attention is instead on the content of their answers. Finally, it should be noted that the few studies which have included bisexuals have grouped them with lesbians as part of an LB group, while, as hypothesized, bisexuals in this experiment patterned more closely with straight speakers than lesbians.

4 Listener stereotype use: How sexual orientation is perceived through speech

This chapter discusses three complementary experiments, which I will refer to as Experiment 2a, 2b, and 2c. The results of the production research presented in Chapter 3 demonstrate that, as hypothesized, women are not using phonetic features in speech to directly project or perform their sexual orientation. Rather, it is used - only by straight and bisexual speakers - to indicate familiarity or affiliation with Queer culture. The specific phonetic variables that indicate a higher familiarity (lower mean pitch and faster speech rate) are ones which have been shown by earlier studies to be associated with a lesbian identity (Camp, 2009; Moonwomon-Baird, 1997; Munson et al., 2006a; Munson et al., 2006b; Munson, 2007; Pierrehumbert et al., 2004; Van Borsel et al., 2013). Straight and bisexual women seem to be using existing stereotypes about the way lesbians sound to project association with that group. The overall purpose of the three perception experiments described in this chapter was to understand in more detail what the phonetic stereotypes associated with sexual orientation in women are, and how they fit with other sociophonetic stereotypes of female speech.

4.1 Experiment 2a: Perceiving sexual orientation as a personal attribute

The first of the three perception experiment was designed to test for which acoustic variables contribute to a perceptual stereotype of lesbian speech, and how this stereotype intersects with other social stereotypes such as femininity, shyness, and compassion.

4.1.1 Experiment 2a – Research questions and hypotheses

Subjects listened to clips from the speech collected in Experiment 1 and rated speakers on a set of social and personal attributes, including the likelihood that the speaker is a lesbian. These ratings were analyzed for correlations between each other and with the phonetic variables measured in the production analysis. Analyses also incorporated listener attributes such as sexual orientation and familiarity with Queer culture, to test for ways in which these may influence how they judge individuals based on speech.

The results of production analysis in combination with results of previous related research lead to three primary hypotheses:

- Hypothesis 1: Judgments of sexuality will pattern according to elements of pitch and vowel formants as well as potentially speech rate, sibilant COG creak, and word final /t/-release.

- Hypothesis 2: Speakers who are judged to sound more like a lesbian will also be judged to sound more masculine as well as potentially more educated, casual, uncaring, and outgoing.
- Hypothesis 3: Listeners' judgments and perceptions will vary in kind with their own sexual orientation and familiarity with Queer culture.

4.1.2 Experiment 2a – Methodology

4.1.2.1. Participants

All perception experiments were run on Amazon Mechanical Turk (MTurk), an online platform used for distributing web-based tasks including sociophonetic experiments like the perception experiments of this dissertation. Participants were limited to MTurk "masters", users that have been rated highly for their participation in other experiments. A total of 58 individuals participated in this experiment. They ranged in age from 21 to 69 years with an average of 37.4 years, and included 23 women and 35 men. All were self reported native speakers of English, coming from a wide range of locations around the US, with one from British Columbia, Canada. As with participants in the production experiment, region of origin was skewed across sexual orientation as well as gender.

4.1.2.2. Stimuli

Stimuli for this experiment were taken from recordings of words in isolation, read by the 54 speakers in Experiment 1a. These were *absent*, *dose*, *locate*, *museum*, and *popular*. These particular words were chosen to get a range of phonetic features, particularly targeting features that have been found to be relevant in previous research on topics of sexuality and gender expression. Specifically, this choice of words highlights differences in the phonetic characteristics of /s/, word-final /t/, and the vowels /i/, /ʌ/, /æ/, /e/, /o/, /u/, /a/, /ɚ/, discussed in Chapter 3 as potential cues to sexual orientation.

4.1.2.3. Procedure

This experiment was divided into two nearly identical MTurk Human Intelligence Tasks (HITs). Participants in each HIT heard recordings from the 54 speakers analyzed in the first part of the study. Each recording had the same five words, in alphabetic order, with one speaker per recording. Rather than being played separately, words were combined into sets of five to reduce participant fatigue. For each set of words, participants were asked to rate the speaker on a scale from 1 and 5 for three social variables. These variables targeted aspects of perceived speaker personality, background, presentation, and orientation. The experiment was broken into two HITs so that each listener only judged three of the six variables. In one HIT (HIT 1), listeners rated subjects

on three five-point scales of *educated – uneducated*, *formal – casual*, and *masculine – feminine*. In the other HIT (HIT 2), listeners rated speakers on five-point scales of *shy – outgoing*, *compassionate – uncaring,* and likelihood of being a lesbian. This division was done for two reasons, 1) to shorten the total time of the experiment, avoiding tedium for the participants and 2) to separate judgments for *feminine* and *lesbian*, the expected correlation of which would potentially become stronger if every listener were asked to judge both.

4.1.2.4. Measurements and statistical analysis

Thirty individuals participated in HIT 1 and 28 in HIT 2. The ratings generated by these participants were averaged for each speaker, resulting in a single score for each variable for each speaker. This number was used in analyzing the correlation with ratings and the various phonetic variables examined in Experiment 2. Speakers mean ratings were also used in conducting a principal components analysis of the social variables.

The acoustic data analyzed in this experiment are a subset of the data collected in Experiment 1 – only the five words that listeners heard when making their judgments for the Experiment 2a. These showed the same phonetic variation seen in the complete recordings for each speaker, and following similar testing used in Experiment 1 analysis, the results showed the same patterns: no significant correlations with sexual orientation, and correlations between a high familiarity with Queer culture and lower average pitch as well as familiarity and faster rate of speech.

A principal components analysis (PCA) was run comparing the six different social variables, scaled to unit variance, using the *prcomp* function in the 'stats' package of R (R Core Team, 2013). Strong correlations between variables indicated redundancy in the informativity of those measurements. A PCA of social ratings highlights certain perceived personae such as straight women who are very feminine, or highly educated women who use very formal speech. The model I ran used a dataframe with the score for each speaker for each of the six attributes.

    prcomp(attributes.df)

As with the speakers who participated in the production experiment, listeners in these experiments were asked for their sexual orientation, the number of their family members or friends who identify as homosexual or bisexual, and their self-reported familiarity with Queer culture, as well as their gender, age, ethnicity, state or province of origin, and state or province of current residence. All of these variables were considered as potentially affecting perception of sociolinguistic variation. The sexual orientation of the listeners and their familiarity with Queer culture were not evenly distributed due to the nature of collecting subjects in MechanicalTurk, as can be seen in Table 4.2. To create more even distribution, listeners with "very low" and "low" familiarity were

combined as a group of lower level familiarity. Likewise, listeners with "very high" and "high" familiarity were combined as higher familiarity speakers.

Familiarity	Women			Men		
	Bisexual	Straight	Lesbian	Bisexual	Straight	Gay
Very High	0	0	1	0	0	0
High	0	4	0	1	1	0
Medium	2	9	0	0	8	1
Low	0	7	0	0	19	0
Very Low	0	1	0	0	1	3

Table 4.2: Listener familiarity with Queer culture and sexual orientation in Exp 2a

Three sets of tests were run to determine which of these variables influenced how listeners rated female speech. Stepwise regression tests were conducted in R for each rating (casual, educated, feminine, lesbian, outgoing, and uncaring) with attributes of the speaker as predictor variables, and were compared for best fit. The equivalent was done testing whether attributes of the listener influenced how they rated speakers. Finally, the same was done with phonetic features as predictor variables for each rating. Examples of each are shown below.

lm(feminine ~ speaker familiarity + speaker orientation + speaker ethnicity)

lm(feminine ~ mean pitch + mean F2 + speech rate)

lm(feminine ~ listener familiarity + listener orientation + listener gender)

4.1.3 Experiment 2a – Results

4.1.3.1. Speaker attributes

Analysis showed that listeners are generally accurate in rating a speaker's self-identified sexual orientation. Straight speakers were significantly less likely to be judged as sounding like a lesbian than lesbian speakers ($F_{(2,50)} = 4.694$, $p = 0.0036$). Bisexuality was not a variable which listeners were given to judge; the variable was a range from 1 = "very likely to be a lesbian" to 5 = "not at all likely to be a lesbian". Listeners' lesbian ratings of bisexual speakers patterned more closely to listener ratings of lesbian speakers, but the difference between straight and bisexual speaker ratings did not reach significance ($F_{(2,50)} = 4.694$, $p = 0.081$). Sexual orientation of speakers was also correlated with the feminine-masculine rating and the compassionate-uncaring rating. Straight women were significantly more likely to be rated as being more feminine than lesbians ($F_{(2,50)} = 5.19$, $p = 0.0092$), and bisexuals were somewhat more so ($F_{(2,50)} = 5.19$, $p = 0.0597$). Lesbian speakers were also more likely to be rated as sounding

uncaring than straight speakers (F(2,50) = 4.825, p = 0.0047) and bisexual speakers (F(2,50) = 4.825, p = 0.057).

4.1.3.2. Phonetic variation

Average lesbian ratings by listeners correlated with a number of phonetic variables in speakers' productions. These were mean pitch, pitch range, median F2, and creak, as shown in Figures 4.1, 4.2, 4.3, and 4.4, respectively. Mean pitch shows a strong negative correlation with lesbian ratings. Speakers with a higher mean pitch were rated as less likely to be a lesbian (F(1,51) = 9.167, p = 0.00386). Like median pitch, the range of a speaker's pitch was correlated with her average lesbian rating (F(1,51) = 6.965, p = 0.011). Women with wider pitch ranges tended to be rated by listeners as more likely to sound like a lesbian. The second formant patterns in the same way as pitch; a higher F2 is correlated with a lower lesbian rating (F(1,51) = 11.67, p = 0.00125). Speakers with a higher lesbian rating also tended to have a higher proportion of creaky voice (F(1,51) = 8.734, p = 0.00472).
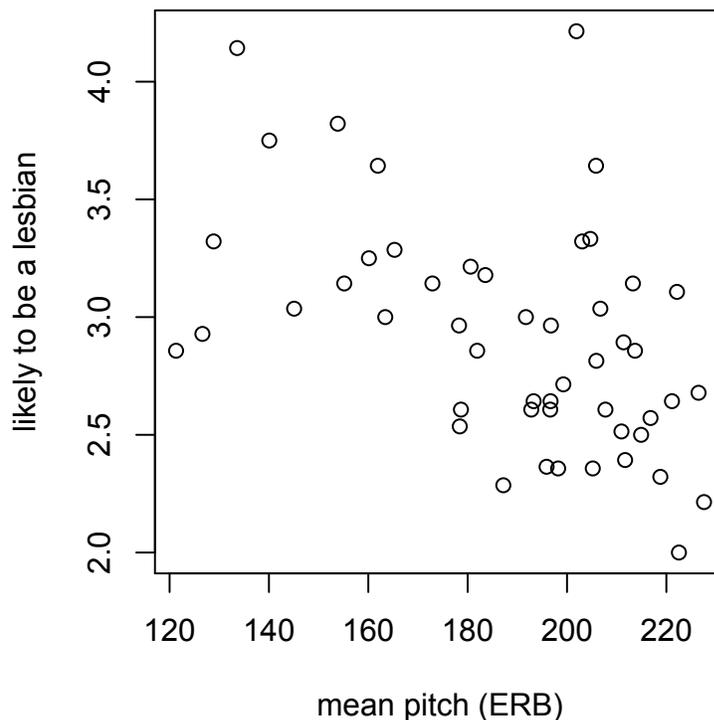


Figure 4.1: Mean pitch and lesbian likelihood rating (1 = "very likely to be a lesbian", 5 = "not at all likely to be a lesbian")
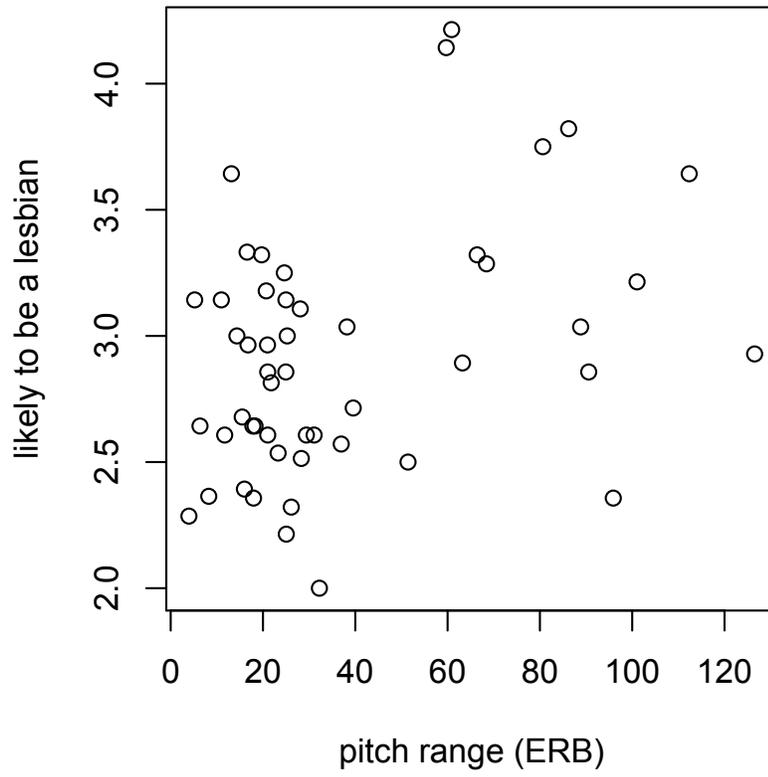
Figure 4.2: Pitch range and lesbian likelihood rating (1 = "very likely to be a lesbian", 5 = "not at all likely to be a lesbian")
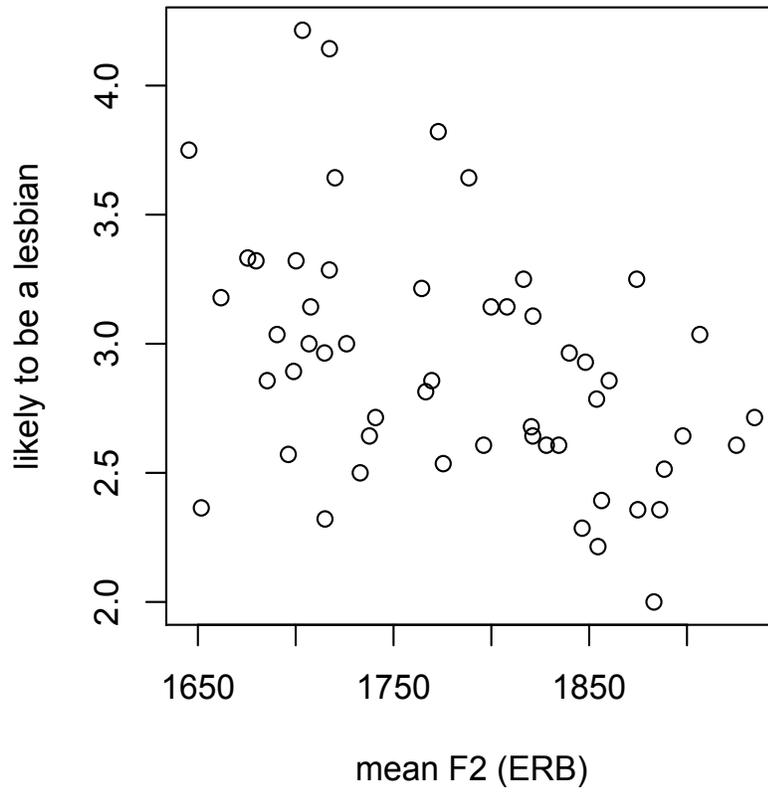
Figure 4.3: Mean F2 and lesbian likelihood rating (1 = "very likely to be a lesbian", 5 = "not at all likely to be a lesbian")
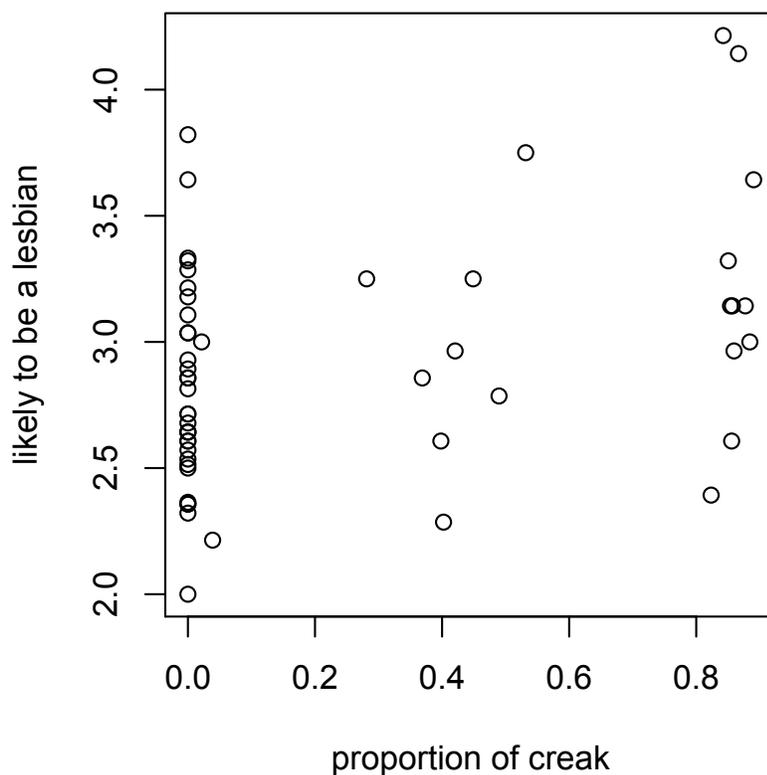
*Figure 4.4: Proportion of creak and lesbian likelihood rating (1 = "very likely to be a lesbian", 5 = "not at all likely to be a lesbian")*Listener attributes

Listeners were also influenced differently by specific phonetic variables according to their familiarity with Queer culture. Regardless of gender, listeners with higher familiarity showed more sensitivity to the median pitch ($F(3,1449) = 19.46$, $p < 0.0001$), median F2 ($F(3,1421) = 24.08$, $p = 0.001$), and proportion of creak ($F(3,1421) = 8.626$, $p = 0.045$) in lesbian ratings. This can be seen respectively in figures 4.5, 4.6, and 4.7. Listener familiarity had no effect on accuracy of their ratings of sexual orientation.
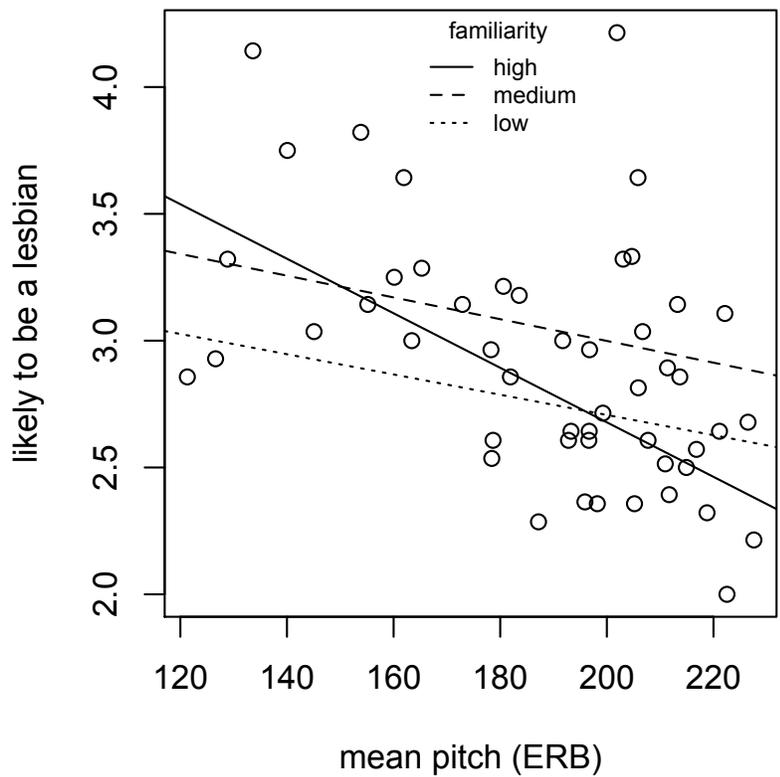
Figure 4.5: Mean pitch and lesbian likelihood rating (1 = "very likely to be a lesbian", 5 = "not at all likely to be a lesbian") showing listener familiarity with Queer culture
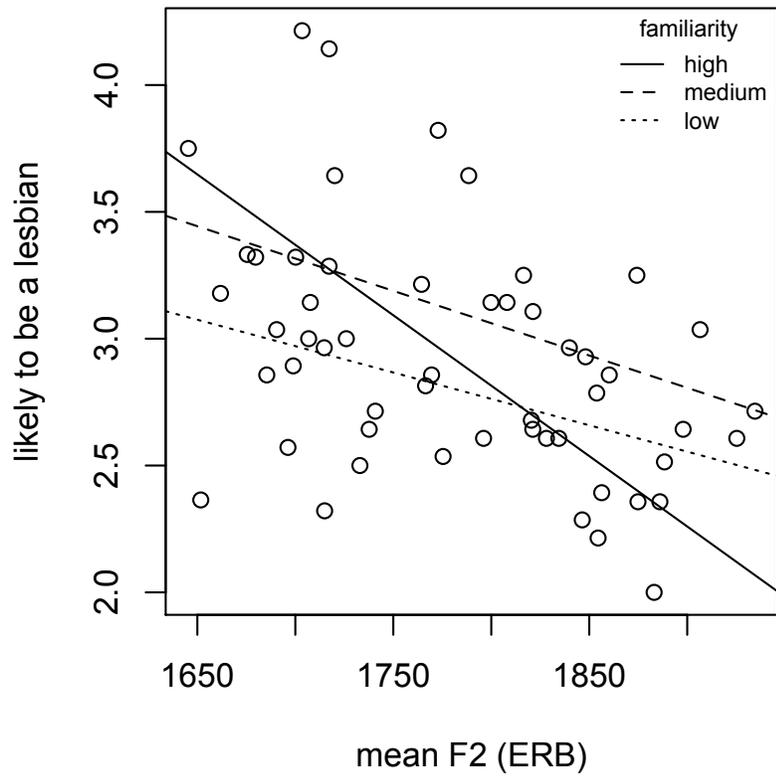
Figure 4.6: Mean F2 and lesbian likelihood rating (1 = "very likely to be a lesbian", 5 = "not at all likely to be a lesbian")showing listener familiarity with Queer culture
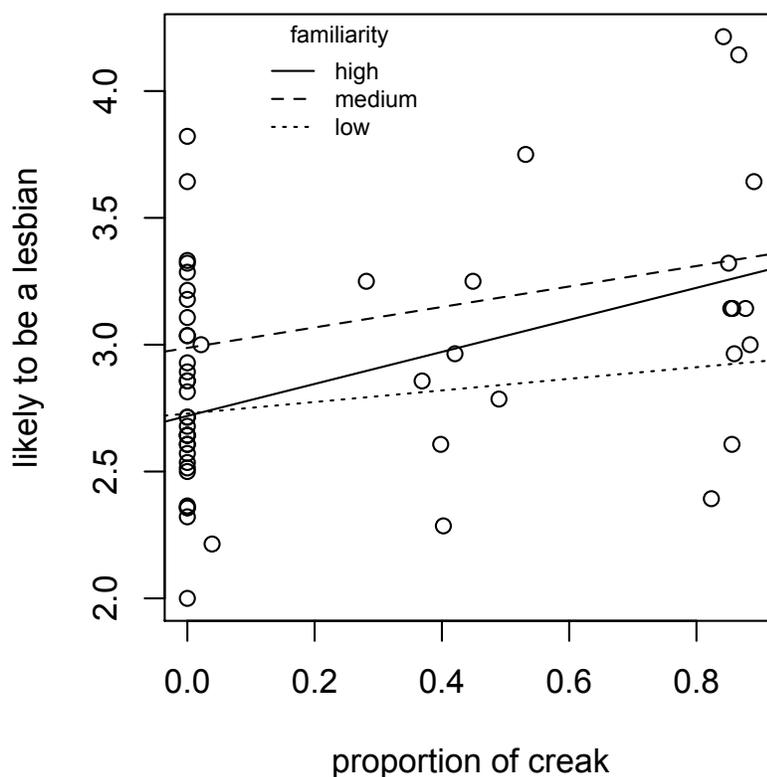
Figure 4.7: Proportion of creak and lesbian likelihood rating (1 = "very likely to be a lesbian", 5 = "not at all likely to be a lesbian") showing listener familiarity with Queer culture

4.1.3.3. Personae

The first four principal components of the PCA described earlier account for over 95% of the variance. These PCs can be interpreted as representing social personae. I define persona as a cluster of social and linguistic features indexing a character type such as a "diva" persona (Podesva, 2006) or the Beijing "smooth operator" persona (Zhang, 2008). Personae are more complex and give a more complete and realistic understanding of social variables than the six individual variables that were given individually to listeners to rate. The first principal component or persona – most notably highly feminine, highly compassionate, and not at all likely to be a lesbian – showed a positive correlation with mean F2 ($F(1,51) = 15.4$, $p = 0.000261$) and mean pitch ($F(1,51) = 5.602$, $p = 0.0218$) and a negative correlation with speech rate ($F(1,51) = 5.788$, $p = 0.0198$). These are shown in Figures 4.8, 4.9, and 4.10 respectively. The second persona, defined primarily by high perceived levels of education, formalness, and extroversion, correlated positively with /t/ release ($F(1,51) = 5.278$, $p = 0.0257$) and mean pitch ($F(1,51) = 3.307$, $p = 0.0749$). The third is described primarily as very outgoing, and also somewhat uneducated, and correlated positively with mean F1 ($F(1,51) = 26.26$, $p < 0.0001$) and pitch range ($F(1,51) = 9.34$, $p = 0.00356$). The fourth persona, which accounts for only a further 5.15% of the variance (perhaps not strong enough to be considered a unique persona) is defined primarily as likely to be a lesbian and compassionate and showed

only a positive correlation with word-final /t/ release ($F(1,51) = 5.109$, $p = 0.0281$). Rotations, standard deviation, percent of variance, and cumulative percept variance are shown in Table

	Principal components			
	PC 1	PC 2	PC 3	PC 4
lesbian	-0.49394740	0.2767793	0.2398844	0.65011123
outgoing	-0.01168662	0.4486815	0.7935182	-0.36681453
uncaring	-0.54398315	0.1668289	-0.2718564	-0.47862877
casual	-0.29874637	-0.5726306	0.2339781	-0.38204387
educated	0.24999585	0.5856799	-0.3505309	-0.25997183
feminine	0.55517061	-0.1527075	0.2475093	0.01319522
Standard deviation	1.6113	1.3860	0.9481	0.55627
% of Variance	0.4327	0.3202	0.1498	0.05157
Cumulative %	0.4327	0.7529	0.9027	0.95424

Table 8: Rotations for the first four principal components

	Principal components			
	PC 1	PC 2	PC 3	PC 4
median pitch	0.31*	-0.25 .	0.004	0.07
pitch range	-0.26 .	0.1	0.39**	-0.1
mean F1	0.18	0.15	0.58***	-0.02
mean F2	0.48***	-0.1	0.08	-0.07
creak	-0.26 .	0.14	-0.14	0.14
speech rate	-0.32*	-0.13	0.09	-0.0003

| /t/ release | 0.1 | 0.31* | -0.19 | 0.3* |

'***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.08

Table 9: Correlations of personae and acoustic variables
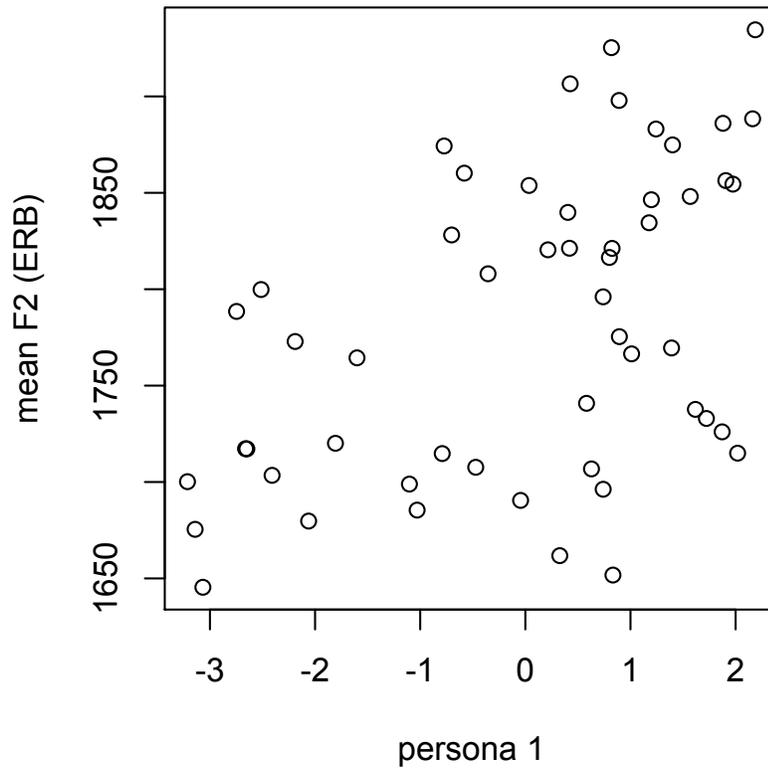
Figure 4.8: Persona 1 score and mean F2
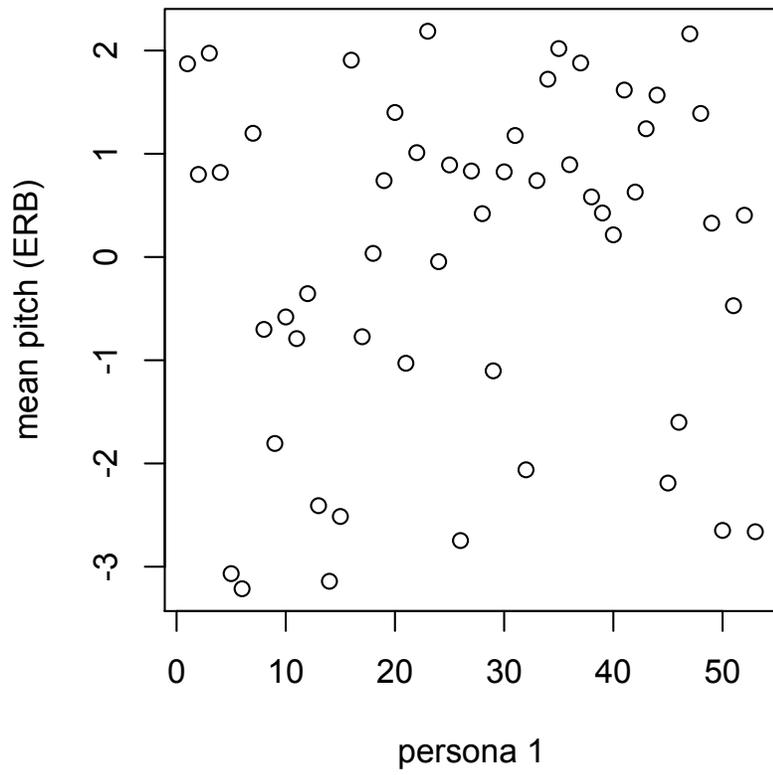
Figure 4.9: Persona 1 score and mean pitch
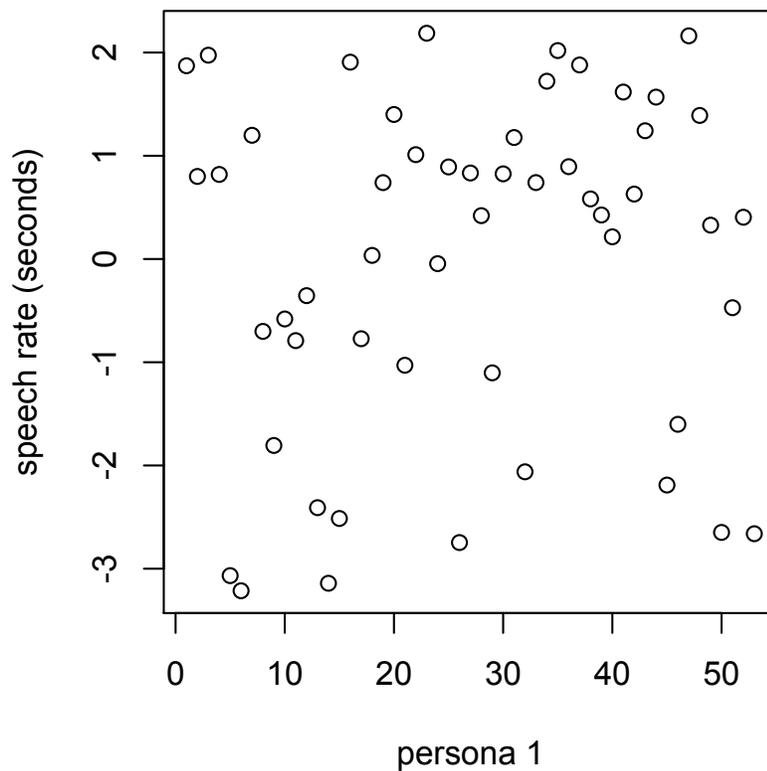
Figure 4.10: Persona 1 score and speech rate

4.1.4 Experiment 2a – Discussion

4.1.4.1. Lesbian rating accuracy

Ratings were not only consistent across listeners, demonstrating that listeners were cuing into a common stereotype of lesbian speech, but they were also strongly correlated with the actual sexual orientation of the speaker, indicating that this lesbian speech style is in fact used more often by the lesbian speakers in Experiment 1a. As discussed in Experiment 1, the self-identified sexual orientation of the speakers did not correlate directly with any individual phonetic variable that was measured, so how is it that listeners are able to accurately identify the sexual orientation of lesbian speakers? Two possible (and potentially overlapping) situations could explain these results. The variables may be interacting in more complex ways than were originally tested, or the listeners may be using phonetic variables that were not measured in this study but are part of a real lesbian speech style. These cues could be other phonetic variables, such as /s/ duration or a different measurement of speech rate, but they could also be differences in *when* variables are used by speakers. Even though limiting the stimuli to only five words restricts time variation, it is still possible for speakers to vary in where in the word they

use creak, for example, or how they realize different intonational contours. Intonation is complex and naturally variable, even with single words; there could be differences in timing, for example. This study measured a wide range of segmental and suprasegmental variables, but their analysis did not capture finer distinctions in timing and intonation which listeners may be able pick up on.

This study has shown that listeners are able to rate a speaker's sexual orientation with a significant degree of accuracy. It has also demonstrated a number of variables that compose a perceived lesbian speech stereotype that is not used consistently by lesbian speakers. However, although this study has analyzed more phonetic variables than any previous work on lesbian speech combined, it has not uncovered the variables in the speech of self-identified lesbian speakers which allow listeners to accurately predict speaker sexual orientation.

4.1.4.2. Effects of listener attributes

Only a handful of previous perception studies have examined the interaction between gender and/or sexual orientation of listeners and their responses. Smyth, Jacobs, and Rogers (2003) found an interaction between gender and sexual orientation of the listeners in which gay male listeners were more likely to judge speakers as gay. Camp (2009) found an influence of listener's gender on how they rated voices on a number of different personality variables such as emotionality and femininity. Levon (2014) found that listener attitudes influenced the role of pitch in their judgments of the masculinity and sexual orientation of speakers.

No previous studies have investigated the influence of listener familiarity with Queer culture or the equivalent. This may be due to the assumption that this influence would be captured by looking at listener sexual orientation, but this implies that the only people with high familiarity with Queer culture are ones who identify as homosexual, and that all self-identified homosexuals have a high familiarity with Queer culture. The results of my study are somewhat complicated by the fact that in the listener population, all the homosexual or bisexual speakers did report having a high or very high familiarity with Queer culture. However, this direct correlation was not true of the speakers, and there were a number of listeners in Experiment 2a who were straight with a high familiarity. Regardless of sexual orientation, familiarity with Queer culture was influential in how listeners judged the sexual orientation of others through speech.

A listener's familiarity with Queer culture influenced the relative predictive power of different phonetic variables. It wasn't an overall trend to judge people as more lesbian-sounding, but a clear sensitivity to lesbian speech stereotypes, with ratings more directly based on specific phonetic variables. This connection indicates that stereotypes of particular groups or identities are formed and strengthened through greater familiarity

with those groups, an important point that is salient for aspects of social psychology within and beyond linguistics.

Despite their greater use of specific phonetic features of the lesbian stereotype, listeners with higher Queer culture familiarity were not more successful in their ratings. There was no interaction between listener familiarity and speaker sexual orientation in their ratings of how likely a speaker was to be a lesbian. Evidently, more familiarity leads to greater use of linguistic stereotypes, but not a more accurate application of those stereotypes in categorizing others based on their speech.

4.1.4.3. Sociolinguistic personae

In most studies investigating the perception of gay and/or lesbian speech, listeners are asked to judge only the sexual orientation of the speaker, using either a binary choice of straight or gay or a 1-dimensional scale between the two. Some studies have also included the variable of femininity or masculinity, recognizing that the perception of sexual orientation is often associated closely with gender expression. This study included judgments of variables beyond sexual orientation and masculinity/femininity for multiple reasons. The primary reason was to understand what other associations listeners have with female sexual orientation and how strong those associations are. For example, it was expressed in some interviews from Experiment 1b that lesbians tend to use more casual speech. Additionally, these ratings were included as a distraction for listeners to prevent them from focusing exclusively on sexual orientation and gender expression which may result in their overthinking their judgments.

As expected, I found that a number of these attributes showed strong correlations with one another. To provide a more complete picture of listeners' perceptions, I combined the attributes in a principal component analysis. The results of the PCA revealed a number of principal components which can be interpreted as social personae. These include a compassionate, straight, and feminine persona (P1), an educated, formal, and outgoing persona (P2), and a very outgoing and somewhat uneducated persona (P3). The fourth (P4) is less recognizable, a compassionate lesbian persona. While it may be that this is a representation of the 'femme' lesbian in contrast to the 'butch' lesbian represented by the P1, there is no correlation with femininity. Note that these personae do not show overall clustering of social variables; when using formal speech, a woman sounding outgoing is considered to be more educated, as demonstrated by P2. However, without this formal speech, that same outgoing attitude is associated with a lower level of education, as shown in P3.

With PCA, each speaker is given a score on the scale persona. This score makes it possible to see which phonetic variables are associated with each persona. In the data from this study, for instance, I found that the "angry butch lesbian" persona (masculine, uncaring, and likely to be a lesbian) is marked by a lower mean pitch, lower mean F2,

and faster speech rate. Sociophonetic representations of these multidimensional perception gives us, as linguists, a better understanding of what listeners are really judging when they hear specific phonetic features, such as a lower average pitch or smaller pitch range. I contend that listeners are less likely to perceive a given speaker on a scale for separate social attributes and are more likely to classify them into a certain socially established persona.

4.2 Experiment 2b – Listener-elicited descriptions

4.2.1 Experiment 2b – Research questions and hypotheses

Experiment 1 directly tested for phonetic stereotypes regarding female sexuality and the interactions with other social and personal attributes. The results demonstrated a clear phonetic stereotype of lesbian speech, as well as clusters of attributes suggesting that listeners perceived personae in which sexual orientation was an attribute, rather than directly perceiving sexual orientation. However, these attributes were sampled from previous studies of speech perception and comments on sexual orientation from Experiment 1b interviews; they may not reflect the automatic associations listeners make when hearing speech. Experiment 2b was designed to allow subjects to provide their own descriptors, without prompting, and to study how these relate acoustically or socially to the judgments made in Experiment 2a. The results from the production study and Experiment 2a suggest three hypotheses for Experiment 2c.

- Hypothesis 1: Unprompted associations between personal and social attributes will overlap with the attributes listeners were given to rate in Experiment 1, but will align more closely with the personae resulting from the PCA, which were interpreted to be more accurate representations of sociolinguistic perceptions.
- Hypothesis 2: Sexual orientation will not be elicited as an unprompted association.
- Hypothesis 3: Phonetic variation will follow patterns according personae, and correlations will be stronger between the freely elicited associations and phonetic features than with predetermined adjectives.

4.2.2 Experiment 2b – Methodology

4.2.2.1. Participants

As with Experiment 1, Experiment 2 was administered to MTurk "masters". A total of 58 individuals participated in this experiment. They ranged in age from 23 to 63 years with an average of 39.4, and included 29 women and 29 men. All were native speakers of English, coming from a wide range of locations around the US and two from areas in India. Note that the distribution is highly skewed towards straight for both

genders, especially women. Similarly, the vast majority of listeners reported having a familiarity level of medium or lower; only 1 of 58 having a "very high" familiarity.

Familiarity	Women			Men		
	Straight	Bi	Lesbian	Straight	Bi	Gay
Very Low	4	0	0	5	0	0
Low	6	0	0	10	1	1
Medium	15	0	0	8	0	1
High	2	1	0	2	0	1
Very High	1	0	0	0	0	0

4.2.2.2. Stimuli

As with those of Experiment 1, the stimuli of Experiment 2 were recordings from the production part of the study. However, for Experiment 2, these stimuli were limited to 48 speakers, and were a full sentence rather than individual words from the list. The single sentence used for all speakers was "After the dance they went directly home." This sentence was chosen for best quality across speakers (avoiding false starts, speech errors, etc.), and to target certain phonetic features (/s/, /ae/N vs. /ae/, etc.).

4.2.2.3. Procedure

Each HIT contained recordings from 12 speakers, with sexual orientations of speakers in each HIT proportional to the complete set. Participants each heard twelve recordings of different speakers reading the same sentence and were asked to write five adjectives that best describe the speaker. A total of 58 people participated in the experiment, 25 participants for the first HIT and 11 participants each for the other three.

4.2.2.4. Measurements and statistical analysis

Acoustic measurements were taken of the same phonetic variables as in Experiment 1. Values were taken from each speaker's sentence, and not the speaker's full data. I hand-corrected obvious misspellings in listeners' responses so that all entries of the same word would be recognized as such. Using Latent Dirichlet allocation (Blei, Ng & Jordan, 2003), I fitted a topic model of 6 topics. This method is used primarily for analysis of documents; in this case speakers were analyzed as documents composed of the list of adjectives they received. In this method, each topic is a ranking of the prevalence of a given word as it is applied to each of a set of documents (or speakers). Each speaker is treated as a combination of topics with a score for the fit of each topic. An example of such a topic would be one ranked highest for "confident", "loud", and

"happy". These scores were used in testing correlations with attribute ratings and personae scores from Experiment 2a.

4.2.3 Experiment 2b – Results

4.2.3.1. Personality types

Table 4.3 below shows the top four terms for each of the six topics. With the exception of "feminine" and "intelligent" most terms refer to attitude (e.g., friendly or shy) or sentiment (e.g., happy or tired), rather than many of the social attributes used in Experiment 1. For this reason, I refer to them as personality types and consider them distinct from the social personae of the PCA in Experiment 1.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
young	smart	confident	calm	bored	quiet
feminine	intelligent	loud	friendly	tired	shy
fun	calm	happy	soft	sad	timid
nice	serious	friendly	annoyed	slow	reserved

Table 4.3: Top four terms for each topic

4.2.3.2. Phonetic variation

Table 4.4 shows the Pearson's correlations between the personality types and all measured variables with a p-value of less than 0.08. Unlike the given attributes in Experiment 1, or the resulting personae, the topics in Experiment 2 were much less likely to show correlations with phonetic variation. As shown below, Topics 1 and 3 show no clear correlation with any of these variables. The only personality type with phonetic correlations that reached significance was Topic 5; speakers with a higher score in Topic 5 were significantly more likely to have a higher F1 ($F(1,43) = 4.984$, $p = 0.0308$) and higher F3 ($F(1,43) = 4.358$, $p = 0.0428$).

Listener ratings of social variables					
Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**	**Topic 6**

mean pitch	-0.075	-0.27.	0.155	-0.015	0.272.	-0.114
mean F1	-0.212	-0.052	0.181	0.067	0.322*	-0.197
mean F2	-0.148	-0.29.	0.181	0.045	0.229	-0.045
mean F3	-0.082	0.074	0.231	-0.202	0.303*	-0.253.
speech rate	-0.041	-0.029	0.169	-0.252.	0.037	0.034

'***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.08

Table 4.4: Correlations of topics and acoustic variables

4.2.3.3. Compared to results from Experiment 2a

Topic 5 was also the only personality type which showed correlation with any of the attributes rated in Experiment 2a, and similarly the only one correlating to the personae resulting from the PCA. The personality type marked by sounding "bored", "tired", "sad", and "slow" correlates negatively with ratings of "uncaring" ($F(1,43) = 4.333$, $p = 0.0434$) and high PC3 scores ($F(1,43) = 6.5$, $p = 0.0144$). Tables 4.5 and 4.6 show correlations and significance between topics and ratings and topics and personae respectively.

Adjective Topic Scores						
	Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**	**Topic 6**
casual	-0.169	0.015	-0.074	-0.091	0.024	0.198
educated	0.16	0.112	-0.035	0.091	-0.098	-0.131
feminine	-0.21	0.067	0.063	-0.115	0.239	-0.035
lesbian	0.232	0.062	-0.169	0.081	-0.12	-0.056
outgoing	-0.095	-0.057	-0.092	0.081	0.245	-0.064

uncaring	0.189	0.128	-0.092	0.049	-0.303*	0.050

'***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.08$

Table 4.5: Correlations of topics with attribute ratings from Exp 2a

	Adjective Topic Scores					
	Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**	**Topic 6**
PC1	-0.143	-0.023	0.112	-0.047	-0.195	-0.072
PC2	0.197	0.043	-0.060	0.135	-0.058	-0.161
PC3	-0.235	-0.090	-0.081	-0.015	0.362*	0.013
PC4	0.23	-0.073	0.022	0.019	-0.005	-0.164

'***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.08$

Table 4.6: Correlations of topics with personae from Exp 1a

4.2.4 Experiment 2b – Discussion

The second perception experiment was designed to determine what unprompted associations listeners make when hearing female speech. Free-elicitation adjectives were collected and, using topic modeling, I uncovered six primary personality types. Each speaker then had a score for each of the six personality types, and I was able to uncover how these aligned with their scores on the social attributes ratings and social personae scores from Experiment 2a. I was also able to test for how phonetic variation determined which associations were made for a given speaker.

4.2.4.1. Compared to results from Experiment 2a

Hypothesis 1 stated that topics would align somewhat with certain predetermined attributes, but would show stronger correlations with social personae, as they were interpreted to be more accurate sociolinguistic perceptions. The only topic which correlated significantly with any of the adjectives or personae was Topic 5, which only

showed correlation with PC3 and "uncaring". The correlation was stronger with the persona than the adjective, but only minimally so.

Evidently, the adjectives used in Experiment 2a were not generally representative of the social categories or attributes that most listeners immediately draw upon, and because the adjectives listeners used were so varied, the resulting PCs also did not capture these social categories with any clarity. Hypothesis 2, that sexual orientation would not feature as an automatic association, was also confirmed. Of the 3480 adjectives elicited in Experiment 2b, only one of them ("gay") referenced sexual orientation. Listeners are generally very consistent in rating sexual orientation when asked to do so, but without direction, this is not a salient social attribute for these voices. Gender expression, on the other hand, were quite common, especially the term "feminine", but also terms like "tomboy", "girly", or "mannish".

I propose a cognitive representation directly linking acoustic input from female speech and perceived gender expression, but not necessarily to sexual orientation. Sexual orientation is instead linked to the gender expression. When asked to judge female sexual orientation through speech, listeners are able to do so consistently, but the process involves first making reference to aspects like femininity. This would account for the very strong correlation between the attributes of "likely to be a lesbian" and "masculine", as well as the difficulty interviewees had in describing lesbian speech stereotypes. It was much easier for participants in the production experiment to describe speech stereotypes about gay men, as well as lesbian stereotypes involving non-speech features (such as appearance or habits). This brings up two questions that require further testing. First, does this direct perceptual link between speech and sexual orientation exist for men while it is absent in women? Second, does there exist a direct link between female sexual orientation and non-speech features such as clothing or hairstyle, which was suggested in responses in speaker interviews from Experiment 1b?

Experiment 2b was designed to elicit the social associations listeners make automatically from speech. The results are interpreted as set of personality types (or topics) that speakers are perceived as belonging to. However, the dearth of phonetic markers that correlate significantly with any of the topics suggests that these personality types are not representative of individuals' perceptions; automatic associations are likely less common across speakers than originally expected. Listeners appear able to access common social stereotypes (e.g., those represented in popular media), but each individual may perceive social attributes and phonetic features differently – what is salient for one individual may not be for another. Previous research has shown that listeners use different cues to judge speakers, but so far these have all been based on predetermined attributes and do not test how automatic associations differ.

4.2.4.2. Phonetic variation

Hypothesis 3 was also confirmed by the results of this experiment – because judgments were based exclusively on speech, correlations between different judgments should be reflected in the acoustics. However, only two significant correlations between judgments from Experiments 2a and 2b were found. Topic 5 and *uncaring* ratings were significantly correlated and both were marked by a higher F2. Topic 5 also correlated positively with PC3, both of which were marked by a higher F1 (though this effect did not reach significance for Topic 5).

4.3 Experiment 2c: Orientation and listener confidence

4.3.1 Experiment 2c – Research questions and hypotheses

While the results of Experiment 2b suggest that sexual orientation was not a salient social feature for listeners, Experiment 2a results demonstrate that stereotypes about this feature exist nonetheless. Experiment 2a only investigated sexual orientation as a single dimension between lesbian and straight. This is common among sociolinguistic and social psychology research, even when acknowledging that sexual orientation is not binary. Furthermore, demographic studies have shown that more women identify as bisexual than lesbian (Gates, 2011).

This then brings up the question of whether a speech stereotype of bisexual women exists and how it compares to the stereotypes uncovered in Experiments 2a and 2b. I designed a third experiment to answer this question. When asked for stereotypes about bisexuals, interviewee responses generally reflected social attitudes of bisexual invisibility. These included comments that there aren't any stereotypes, that you can't tell if a woman is bisexual, or that bisexual women look/sound just like straight women. Based on these responses as well as research in female sexuality and psychology, I developed the following hypotheses.

- Hypothesis 1: straight ratings opposite to lesbian ratings
- 2: Fewer speakers would be judged as sounding like a bisexual than straight or lesbian.
- Hypothesis 3: Listeners would have lower confidence when judging speakers to be bisexual.
- Hypothesis 4: Speakers judged as sounding bisexual would exhibit similar phonetic variation as speakers judged as sounding straight.

4.3.2 Experiment 2c – Methodology

4.3.2.1. Participants

As with Experiments 2a and 2b, Experiment 2c was run on Mturk. A total of 60 people participated, but three participants' data were excluded for incompleteness. The same requirements for participants were applied. Information such as listener sexual orientation and familiarity with Queer culture were collected with the same questions as in Experiments 2a and 2b, and as with Experiment 2b, numbers of friends and family members who identified as bisexual or homosexual were limited to sets to avoid outliers and answers like "all" or "200" which were given by some participants in Experiment 2b. Participant ages ranged from 23 to 63 years old, with an average age of 39.35.

4.3.2.2. Stimuli

The stimuli for Experiment 2c was identical to that of Experiment 2b. The same 48 subjects' recordings of the sentence, "After the dance they went directly home" were played to listeners. Participants heard to each recording, then chose what they felt was the most likely sexual orientation of the speaker in the recording and indicated on a scale of 0-100 their confidence in that choice.

4.3.2.3. Measurements and statistical analysis

Each speaker received a separate score for bisexual, lesbian, and straight which were calculated as the sum of listener confidence ratings for each orientation. Using these, ANOVAs tested predictions of sexual orientation ratings by other social attributes from Experiments 2a and 2b, as well as phonetic variation between speakers. Experiment 2c used the same stimuli as Experiment 2b, and so the same phonetic variables were analyzed.

4.3.3 Experiment 2c – Results

4.3.3.1. Phonetic variation

Table 4.8 below shows the correlations between scores of each sexual orientation and the phonetic features that were significant or approached significance. Not surprisingly, the scores for sexual orientation show similar acoustic correlations to the ratings for PC1 and its component attributes (femininity, lesbian likelihood, and compassion). The apparent role of mean pitch, which was significant for PC1, lesbian likelihood, and compassion, is decreased when subjects are asked to pick a sexual orientation. The mean F1 and F2, the spectral balance of /s/ and the standard deviation of vowel duration showed the reverse and were more significant in this direct judgment.

There was no clear stereotype for what made a speaker sound bisexual; no phonetic feature showed a significant correlation. In interviews, many speakers noted a lack of bisexual speech stereotype, and most said that bisexuals are generally perceived as sounding straight. The results of this experiment suggest otherwise – bisexual sounding speakers don't sound straight. If anything, they appear to sound more like lesbians, but it seems that the total lack of stereotype means that judgments of bisexuality were essentially random. Ratings of lesbian versus straight on the other hand are almost completely opposed. Every phonetic feature that makes a woman sound like a lesbian will make her sound straight if reversed. Figure 4.11 shows the negative correlation between straight and lesbian ratings.

	Listener ratings of social variables		
	bisexual	lesbian	straight
median pitch	-0.205	-0.258.	0.271.
mean F1	-0.226	-0.397**	0.441**
mean F2	-0.249.	-0.561***	0.517***
mean F3	0.066	-0.385*	0.339*
/s/ spectral balance	0.03	0.374*	-0.27.
vowel dur standard dev.	0.117	0.328*	-0.337*

'***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$

Table 4.8: Correlations with perceived orientation and acoustic variables
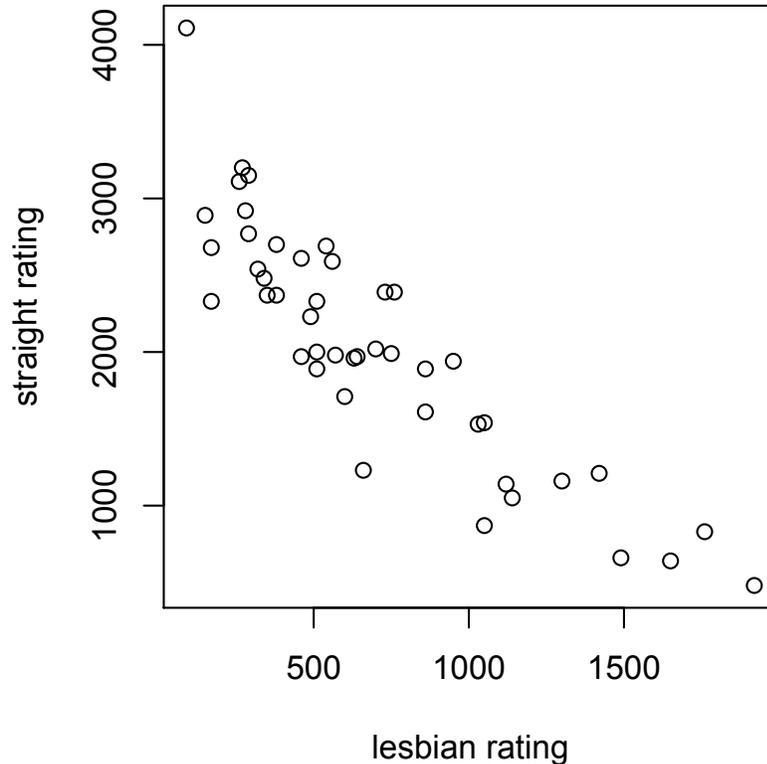
Figure 4.11: Lesbian ratings and straight ratings

4.3.3.2. Compared to results from Experiment 2a

This same pattern can be seen equally well in the correlations with ratings of social attributes and persona values. Lesbian- and straight-sounding speakers pattern in nearly opposite ways. Lesbian-likelihood ratings correlated positively with lesbian ratings in Experiment 2c $(F(1,43) = 13.58$, $p = 0.00064)$ and negatively with straight ratings $(F(1,43) = 8.168$, $p = 0.0066)$. Femininity showed a negative correlation with lesbian ratings $(F(1,43) = 24.66$, $p < 0.0001)$ and a positive correlation with straight ratings $(F(1,43) = 17.57$, $p = 0.00014)$. Uncaring ratings were positively correlated with lesbian ratings $(F(1,43) = 33.9$, $p < 0.0001)$ and negatively correlated with straight ratings $(F(1,43) = 32.33$, $p < 0.0001)$. The apparent lack of a bisexual stereotype can be seen here as well. There is no significant correlation between sexual orientation category scores and any attribute or persona from Experiment 2a. With social attribute ratings, as with phonetic variables, the variable with the strongest correlation for straight and lesbian (uncaring) approached significance for bisexual, but it was likewise in the same direction

as the lesbian rating as opposed to reported stereotypes of bisexuals sounding straight. Table 4.9 shows the correlations for all social attribute ratings.

	bisexual	lesbian	straight
casual	-0.065	0.249.	-0.199
educated	0.154	-0.136	0.09
feminine	-0.024	-0.604***	0.539***
lesbian	0.063	0.49***	-0.4**
outgoing	0.052	-0.019	0.159
uncaring	0.281.	0.664***	-0.655***

'***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, '.' $p < 0.1$

Table 4.10: Correlations of perceived sexual orientation with attribute ratings from Exp 1a

Persona correlations show an unexpected difference and may give insight into the nature of the confusing PC4 (compassionate lesbian). The only persona that showed a clear correlation with bisexual scores was one which was not significant for either of the other sexual orientations. Table 4.10 shows the correlations for all four personae.

	bisexual	lesbian	straight
PC1	-0.083	-0.643***	0.574***
PC2	0.156	0.006	-0.033
PC3	-0.102	-0.26.	0.267.

PC4	-0.257.	-0.014	0.126

'***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1

Table 4.11: Correlations of perceived sexual orientation and personae from Exp 2a

4.3.3.3. Compared to results from Experiment 2b

As with Experiment 2a, Experiment 2c required participants to rate predetermined adjectives or attributes. The topics of Experiment 2b were instead volunteered by participants themselves. These are the social features that listeners hear without prompting, and are thus interpreted as a more accurate representation of casual sociolinguistic perception. The fact that none of these shows a significant correlation with any of the sexual orientation scores further confirms the inference that sexual orientation is not a notable feature of female speech. Furthermore, the only personality type which approaches significance is the one which includes gender expression as a top concept. This also adds support for the theory that, although not linked unconsciously to phonetic variation, judgment of sexual orientation based on speech is consistent when solicited because it is connected through perceptions of gender expression. Table 4.11 shows the correlations for all six topics.

	bisexual	lesbian	straight
Topic 1	0.065	0.251.	-0.227
Topic 2	0.227	0.132	-0.141
Topic 3	-0.07	-0.127	0.128
Topic 4	0.093	-0.034	-0.082
Topic 5	0.038	-0.166	0.137
Topic 6	-0.197	-0.075	0.106

'***' p < 0.001, '**' p < 0.01, '*' p < 0.05, '.' p < 0.1

Table 4.12: Correlations of perceived sexual orientation and topics from Exp 2b

4.3.4 Experiment 2c – Discussion

Contrary to my hypothesis, bisexual ratings did not pattern similarly to straight ratings. While straight and lesbian ratings were generally found on opposite ends of most variables tested, both in terms of who sounded more straight or lesbian and thus in the phonetic features that predicted the ratings, bisexual ratings followed no such pattern. Based on the significance of these correlations, the ratings of bisexuality seemed to be more or less random, certainly much less predictable than straight or lesbian. Listeners were perfectly willing to choose bisexual as a rating, and were no less confident in their ratings than for lesbian ratings (distribution of confidence ratings shown in Figure 4.12), but bisexual ratings were inconsistent.
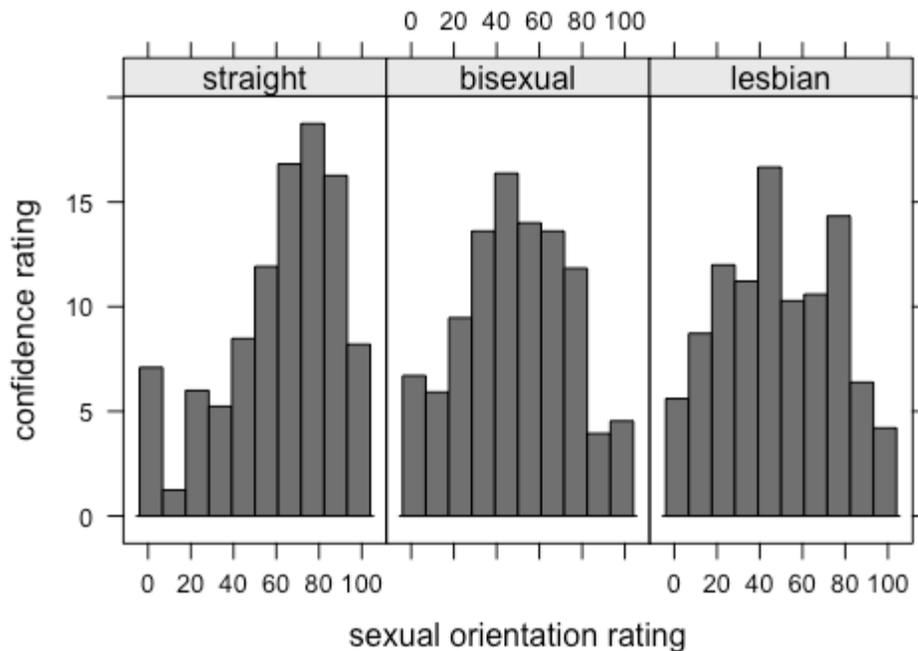
Figure 4.12: Confidence scores according to perceived sexual orientation

It would seem that the inclination of researchers to ignore bisexuality as a sexual orientation, is not mirrored by participants (hypothesis 1 and 2 were not confirmed). However, it is clear that, unlike straight or lesbian speech, there is no consistent social stereotype about bisexual speech in women. Whether this is true of bisexual men or non-speech features of bisexual women has yet to be tested.

4.4 Conclusions

4.4.1 Stereotype perception through speech

This chapter described three experiments designed to together test the perception of sexual orientation in women's speech, based on sentence or word readings. The first used Likert scale judgments of a set of personal and social attributes, including likelihood that the speaker is a lesbian. The second targeted individual listeners' associations through open-ended adjective elicitation. The third compared isolated judgments of sexual orientation as non-binary, including bisexuality as a possible choice.

Combined, the studies help to illuminate the cognitive processes of stereotype perception. When asked to rate certain social attributes, listeners show consistency both within their own judgments of multiple speakers, and across listeners. Certain attributes cluster together, indicating that individuals do not judge them independently, but as parts of larger social stereotypes. However, while these social stereotypes are common across listeners and marked by the same phonetic variables, listeners with higher levels of familiarity with Queer culture are more consistent in their use of those variables (though they are no more accurate in their classification of women as being lesbian or bisexual).

Listeners are able to reference these common social stereotypes, as shown in the personae of Experiment 2a, and can easily judge stereotypical lesbian or straight speech, though not bisexual speech, as shown in Experiment 2c. However, Experiment 2b results demonstrate that these are not generally salient features for listeners when they are not elicited explicitly. Finally, this conclusion brings up a number of other questions. Do these automatic associations vary according to social attributes of the listener? Previous research has demonstrated that attributes like gender and attitude, such as the MRAS of Levon's (2011) study make a difference. How do associations for female speech differ from (a) aspects of female appearance or non-speech behavior and (b) associations for male speech? Are the same phonetic cues marking the same associations, or do they have different meanings for different speakers? How are these associations learned (going back to question 1) and how are they processed cognitively? According to the AWSA model, such associations are the result of different listeners attending to different social and phonetic features, storing connections between them with different weight, or even making different connections.

In the next and final chapter, I discuss the results of the production study and perception studies together, addressing the larger questions of how production and perception are linked in the presentation and perception of sexuality in speech and language, and how both are an important key to understanding the processes underlying sociolinguistic stereotyping.

5 Conclusion

5.1 Summary of dissertation goals

At its core, my dissertation addresses a single primary question: What does it mean to sound like a lesbian? On the surface, this may seem a relatively simple question, something to the effect of, 'In what ways does lesbian speech differ from that of other women?' However, my work takes a broader perspective, approaching this single question from a multitude of perspectives. Such perspective include the following questions:

1. What do lesbian speakers sound like? How does this relate to straight or bisexual speakers?
2. Why might speakers not identified as lesbians make use of lesbian stereotypes in speech?
3. Are self-identified lesbians perceived accurately as such through their speech?
4. When listeners are asked to judge sexual orientation, are there patterns of speech that they consistently judge as sounding lesbian-like (regardless of accuracy)?
5. Is there even such a thing as a "lesbian speech stereotype"? How does it interact with other social stereotypes?
6. How meaningful or salient is the concept of a lesbian speech stereotype when participants are not asked to judge it?
7. How are other sexual orientations perceived through speech? Is there a bisexual speech stereotype and how does it relate to sounding like a lesbian?
8. What is the cognitive process of perceiving female sexual orientation through speech?

To answer these questions, I carried out a combination of production and perception experiments and interpreted through the exemplar-based Attention Weighted Schema Abstraction.

5.2 What speech production reveals about stereotypes

To answer the first two of the above questions, I recorded a total of 54 women who self-identified as bisexual, lesbian, or straight reading a series of single words and sentences (Experiment 1a) and collected information about their ages, ethnic identities, dialectal regions or origin, and aspects of their interactions with Queer individuals and Queer culture. The read speech recordings were optionally followed by a sociolinguistic interview discussing different stereotypes of male and female sexuality, including possible lesbian speech stereotypes (Experiment 1b). Input from the interviews helped to guide interpretations of the research and were analyzed primarily for acoustic variation. Variation was measured for a wide range of phonetic features including vowel formants,

pitch, speech rate, sibilant spectral balance, final /t/ release, and creak. Following previous research and answers from participant interviews, I hypothesized that speaker sexual orientation would not be directly predicted by phonetic variation. This hypothesis was confirmed. In answer to Question 1, lesbians evidently don't vary consistently in their style of speech in read speech or sociolinguistic interviews. On the other hand, variation in two acoustic features followed speakers' self-ratings of their familiarity with Queer culture. Interpretation of this concept was left open to speakers, but however they interpreted it, the effect was consistent. Speakers with higher familiarity had lower mean pitch and faster speech rate, both of which were identified by participants as stereotypically lesbian-like. This variation by familiarity was only true for straight and bisexual speakers, indicating that the use of this speech style may be a means for speakers to express their affinity with Queer or lesbian culture.

In answer to Question 2, this style of speech can be used, not to mark a queer or lesbian sexual orientation, but to show affinity with that culture. This would be of less value to participants who had already identified themselves as lesbians as a part of the study. That the use of this style is more a performance than speakers' unmarked speech was further supported by its decrease or loss of significance in interview speech. Speakers monitor speech more in reading and are more likely to attend to stylistic variation than in sociolinguistic interviews when they attend instead to the content of their answers.

5.3 What speech perception reveals about stereotypes

Experiment 2 was composed of three related studies of sociolinguistic perception, with the aim of answering the remaining questions. In the first of these, listeners heard a clip of five words from the speakers in Experiment 1a. For each speaker, listeners judged them on three out of six social or personal attributes, including likelihood of being a lesbian. The results showed that, despite acoustic analysis of Experiment 1 failing to show variation according to sexual orientation, listeners were able, with only five words per speakers, to identify lesbians as more likely to be lesbians with greater than chance accuracy. The answer to Question 3 is a clear 'yes'. How listeners are able to do this is still unanswered.

In answer to Question 4, there were four acoustic variables that contributed directly to the perception of lesbian likelihood, though these variables did not accurately predict speaker self-identification. Speakers with lower mean pitch, wider pitch range, lower mean F2, and higher proportion of creak were perceived as generally more likely to be lesbians. Lesbian likelihood was also positively correlated with ratings of *masculine* and *uncaring*. Through the use of a principal components analysis, I was able to identify four recognizable social personae by which listeners were judging speakers. The first of these was an 'angry butch lesbian', providing insight to Question 5. There is a lesbian

speech stereotype (at least when listeners are asked to judge sexual orientation), and its social meaning is simultaneously more specific and more broad than attraction to the same sex.

The second perception experiment (Experiment 2b) was designed to directly address Question 6 and test whether and how listeners would judge sexual orientation if they were not prompted to do so. Listeners heard a single sentence from a total of 12 speakers (four nearly identical experiments allowed analysis of 48 speakers overall). For each speaker they provided five adjectives they felt described that speaker. These ranged from adjectives like 'shy', 'simple', or 'serene', but of a complete list of 3480 adjectives, only one of them ('gay') referenced sexual orientation. While plenty of adjectives referenced elements of gender expression (e.g., 'girly' or 'mannish') the salience of sexual orientation in the context of speech perception (in an online experiment) is almost nonexistent.

The third perception experiment was essentially the opposite of Experiment 2b; testing what patterns listeners show when asked explicitly and exclusively to judge sexual orientation. This experiment included 'bisexual' as a possible sexual orientation in addition to 'straight' and 'lesbian'. Listeners were highly consistent in their judgments of lesbian and straight orientations, and the acoustic variations were almost exactly reversed. Any variable that corresponded positively with lesbian ratings corresponded negatively with straight, and vice versa. This further confirms the answer to Question 4 provided by Experiment 2a. In answer to Question 7, the listeners were highly inconsistent in ratings of bisexuality. It was not the case that no speakers received judgments of 'bisexual', but which speakers received these judgments varied across listeners and showed no pattern of acoustic variation. Evidently there is no common speech stereotype for bisexual speakers, which makes sense when sexual orientation is perceived linguistically as a linear scale from lesbian to straight.

In interpreting my results as a whole through the AWSA model I developed, I posit a cognitive process of lesbian speech perception which takes an indirect route from acoustic input to judgment of sexual orientation as illustrated in Figure 4.13. Rather than cueing sexual orientation directly, phonetic features like lower pitch or higher proportion of creak prompt perception of gender expression like masculinity or boyishness. Stereotypes beyond language link these gender expressions to sexuality and sexual orientation.

Recall the example given in Section 2.2.1.3.: a listener hears a single utterance from a female speaker and, based to context, attends more strongly to the gender of the speaker than sexual orientation or other dimensions. This probe activates all other traces as an echo. Traces with an overlapping dimension of gender, especially where gender in those existing traces received greater attention, contribute more to the content of the echo. This echo represents an abstraction of speech from female speakers; dimensions of the

acoustics which were attended to with considerable weight will contribute to a stereotype of female speech but not to speech according to sexual orientation.

In this 2-step model of the perception of lesbian speech, most listeners have very few if any traces stored with strong attention weighting of both the sexual orientation of the speaker as well as phonetic features of her speech. On the other hand, most listeners have traces where the dimension of gender expression is strongly weighted with phonetic features. As mentioned in the first chapter of this dissertation, pitch and the first two formants have been shown to have a strong connection perceptually to gender expression. These phonetic features are culturally and physiologically associated with greater height and masculinity. Listeners likewise have traces of women where gender and sexual orientation are strongly weighted. Without reference in the task to sexual orientation, female speech as a probe will create an echo that does not contain that particular dimension. On the other hand, a direct request for sexual orientation judgment, activates traces in which sexual orientation was particularly strongly weighted.

The model as it was tested in the three perception experiments is illustrated in Figure 4.13 below. In all three experiments, the probe that listeners experienced was purely phonetic input (no visual stimuli and variation in syntax, morphology, etc.) In Experiment 2a, the phonetic input activated exemplars with weighting on the scales subjects were asked to rate. This activation is illustrated in the figure with orange dashed lines. This activation of specific features consequently creates echoes of personae like the angry butch/compassionate femme scale that is represented in the PC1 score. In the figure, this personae is illustrated as a circle encompassing femininity and compassion. Listeners were asked to judge the sexual orientation of the speakers, and they were able to do so by accessing traces that had were weighted for sexual orientation as well as other elements of this persona. In Experiment 2b, the phonetic probe was meant to activate most strongly the associations listeners had with phonetic features of the input. These activations are represented with purple dashed lines. Some of these concepts or dimensions were given that had already been used in Experiment 2a (aspects of *shyness*, *femininity*, and *compassion*), others like *boredom*, and *happiness* had not. The lack of activation of any concept of sexual orientation is represented by the lack of a connection between phonetic as well as any other dimensions. Finally, Experiment 2c explicitly asked listeners to judge sexual orientation, but the 2-step model of perception of lesbian speech means that these judgments must be made by first accessing traces with heavily weighted femininity (gender expression) and compassion. This process is represented with the green solid lines.

It is important to note that this model is not necessarily true for all listeners. The figure illustrates what appears to be the cognitive process for most listeners, but there was one, for example, who listed 'gay' as an adjective. Speaking from my own experience, I have at this point stored a countless number of probes in which attention is paid to both the phonetic features of the speech and the (actual or perceived) sexual orientation of the

speaker. Those exemplars mean that an illustration of my personal cognitive processing would have direct connections between phonetics and sexual orientation.
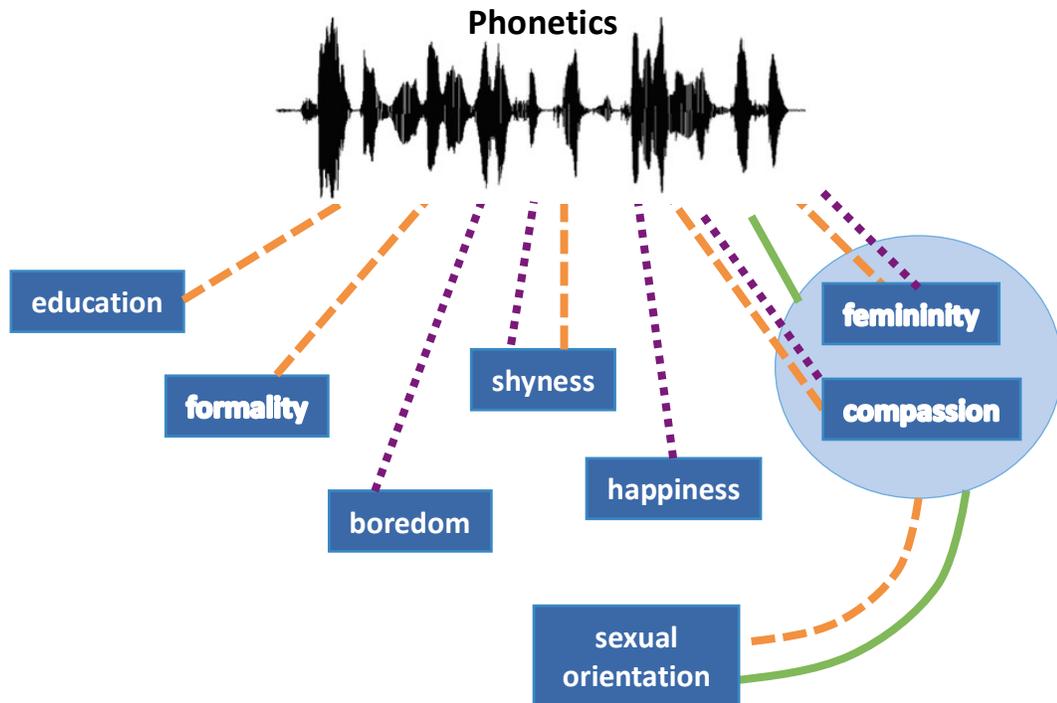

Figure 4.13: two-step model of sexual orientation perceived through speech

The 2-step model of most listeners allows them to use common acoustic cues to judge sexual orientation, but leaves them unlikely to do so without direct prompting. Though interviewees had difficulty conjuring lesbian speech stereotypes consciously, interview responses were rife with references to gender expression, as can be seen in the quote given in the introduction of this dissertation.

References

Andersen, S. M., & Chen, S. (2002). The relational self: an interpersonal social-cognitive theory. *Psychological Review*, *109*(4), 619–645.

Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. CRC Press.

Baeck, H., Corthals, P., & Van Borsel, J. (2011). Pitch Characteristics of Homosexual Males. *Journal of Voice 25*(5), 211-214. doi:10.1016/j.jvoice.2010.10.019

Bar-Tal, D. (1997). Formation and change of ethnic and national stereotypes: An integrative model. *International Journal of Intercultural Relations*, *21*(4), 491–523.

Blashill, A. J., & Powlishta, K. K. (2009). Gay stereotypes: The use of sexual orientation as a cue for gender-related attributes. *Sex Roles*, *61*(11-12), 783–793.

Blei, D.M.; Ng, A.Y.; Jordan, M.I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research, 3*(4–5), 993–1022.

Brambilla, M., Carnaghi, A., & Ravenna, M. (2011). Status and cooperation shape lesbian stereotypes. *Social Psychology*, *42*(2), 101–110

Brewer, M. B., & Gardner, W. (1996). Who Is This "we"? Levels of Collective Identity and Self Representations. *Journal of Personality and Social Psychology*, *71*(1), 83–93.

Bucholtz, M. (2001). The Whiteness of Nerds: Superstandard English and Racial Markedness. *Journal of Linguistic Anthropology 11*(1), 84-100. doi: 10.1525/jlin.2001.11.1.84.

Cadinu, M., Galdi, S., & Maass, A. (2013). Chameleonic social identities: Context induces shifts in homosexuals' self-stereotyping and self-categorization. *European Journal of Social Psychology*, *43*(6), 471–481.

Cameron, D. (2005). Language, Gender, and Sexuality: Current Issues and New Directions. *Applied Linguistics 26*(4), 482-502. doi: 10.1093/applin/ami027.

Cameron, D., & Kulick, D. (2003). *Language and Sexuality*. Cambridge University Press.

Camp, M. (2009). Japanese Lesbian Speech: Sexuality, Gender Identity, and Language. Ph.D. diss., Department of East Asian Studies. The University of Arizona.

Campbell-Kibler, K. (2011). Intersecting Variables and Perceived Sexual Orientation in Men. *American Speech 86*(1), 52-68. doi: 10.1215/00031283-1277510.

Can, B. & Sanver, M. R. (2009). Stereotype formation as trait aggregation. *Mathematical Social Sciences*, *58*(2), 226–237.

Cartei, V. & Reby, D. (2012). Acting gay: Male actors shift the frequency components of their voices towards female values when playing homosexual characters. Journal of Nonverbal Behavior, *36*, 79-93.

Cox, W. T. L. & Devine, P. G. (2015) Stereotypes Possess Heterogeneous Directionality: A Theoretical and Empirical Exploration of Stereotype Structure and Content. *PLOS ONE*.

Doucherty, G. J., & Foulkes, P. (2014). An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua*, *142*, 42–56.

Drugman, T., Kane, J., & Gobl, C. (2012). Resonator-based creaky voice detection. Paper presented at *the International Conference on Spoken Language Processing (Interspeech 2012)*. (Software accessed from https://github.com/jckane/Voice_Analysis_Toolkit/tree/master/creak_fcns.)

Eckert, P. (2010). Three waves of variation study: The emergence of meaning in the study of variation. <web.stanford.edu/~eckert/PDF/ThreeWavesofVariation.pdf> (accessed 9/25/17).

Eliason, M., Donelan, C., & Randall, C. (1992). Lesbian stereotypes. *Health Care for Women International*, *13*(March), 131–144.

Ellemers, N., Korenkaas, P., & Ouwerkerk, J. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, *29*(October 1997), 371–389.

Eves, A. (2004). Queer Theory, Butch/Femme Identities and Lesbian Space. *Sexualities*, *7*(4), 480–496.

Farquhar, C. (2000). `Lesbian' in a post-lesbian world? Policing identity, sex and image. *Sexualities*, *3*(2), 219–236.

Fiedler, K. (2014). From intrapsychic to ecological theories in social psychology: Outlines of a functional theory approach. *European Journal of Social Psychology*, 44, 657–670.

Fingerhut, A. W., Peplau, L. A., & Ghavami, N. (2005). A dual-identity framework for understanding lesbian experience. *Psychology of Women Quarterly*, *29*(2), 129–139.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, *84*(1), 115-123. doi: 10.1121/1.396977.

Geiger, W., Hardwood, J., & Hummert, M. L. (2006). College Students' Multiple Stereotypes of Lesbians. *Journal of Homosexuality*, *51*(3), 165–182.

Haslam, S. A., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social Identity Salience and the Emergence of Stereotype Consensus. *Personality and Social Psychology Bulletin*, *25*(7), 809–818.

Haslam, S. A., & Wilson, A. (2000). In what sense are prejudicial beliefs personal ? The importance of an in-group's shared stereotypes, (June 2015), 45–63.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428.

Huxley, C., Clarke, V., & Halliwell, E. (2013). Resisting and conforming to the "Lesbian Look": The importance of appearance norms for lesbian and bisexual women. *Journal of Community and Applied Social Psychology*.

Johnson, K. (1997). Speech perception without speaker normalization: an exemplar model. *Talker Variation in Speech Processing*.

Johnson, K. (2005). Speaker normalization in speech perception. *The Handbook of Speech Perception, 363-389*

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, *34*(4), 485-499. doi: 10.1016/j.wocn.2005.08.004.

Kennedy, E.L. and Davis, M. (1993). *Boots of Leather, Slippers of Gold*. Routledge.

Kristiansen, G. (2001). Social and linguistic stereotyping: A cognitive approach to accents. *Estudios Ingleses de la Universidad Complutense*, 9, 129-145.

Kite, M. E., & Bryant-Lees, K. B. (2016). Historical and contemporary attitudes toward homosexuality. *Teaching of Psychology*, *43*(2).

Lawson, R. (2015). Sexuality and sociophonetics. *The International Encyclopedia of Human Sexuality*. 1115-1354. doi: 10.1002/9781118896877.wbiehs469.

Levon, E. (2011). Teasing apart to bring together: gender and sexuality in variationist research. *American Speech 86*(1), 69-84. doi: 10.1215/00031283-1277519.

Levon, E. (2014). Categories, stereotypes, and the linguistic perception of sexuality. *Language and Society 43*(1), 539-566. doi: 10.1017/S0047404514000554.

Levon, E. (2015). Integrating intersectionality in language, gender, and sexuality research. *Language and Linguistics Compass*, *9*(7), 295-308. doi: 10.1111/lnc3.12147.

Mack, S. (2010). Perception and identity: Stereotypes of speech and sexual orientation in Puerto Rican Spanish. In *Selected Proceedings of the 12th Hispanic Linguistics Symposium*, ed. Claudia Borgonovo et al., 136-147.

Mack, S. (2011). A sociophonetic analysis of /s/ variation in Puerto Rican Spanish. In *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, edited by Luis A. Ortiz-Lopez, 81-93.

Mack, S. & Munson, B. (2012). The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp' stereotype. *Journal of Phonetics, 40*(1), 198-212. doi: 10.1016/j.wocn.2011.10.002.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods 44*(2), 314-324. doi:10.3758/s13428-011-0168-7

Maurer, K. L., Park, B., & Rothbart, M. (1995). Subtyping versus subgrouping processes in stereotype representation. *Journal of Personality and Social Psychology*, *69*(5), 812–824.

Moonwoman-Baird, B. (1997). Toward the study of lesbian speech. In *Queerly Phrased*, edited by Anna Livia and Kira Hall, 202–13. Oxford: Oxford University Press.

Munson, B. (2007). The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech, 50*(1), 125-142. doi: 10.1177/00238309070500010601.

Munson, B. & Babel, M. (2007). Loose lips and silver tongues, or, projecting sexual orientation through speech. *Language and Linguistics Compass, 1*(5), 416-449. doi: 10.1111/j.1749-818X.2007.00028.x.

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The Influence of Perceived Sexual Orientation on Fricative Identification. *The Journal of the Acoustical Society of America, 119*(4), 2427-2437. doi: 10.1121/1.2173521.

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The Acoustic and Perceptual Bases of Judgments of Women and Men's Sexual Orientation from Read Speech. *Journal of Phonetics, 32*(1), 202-240. doi: 10.1016/j.wocn.2005.05.003.

Newton, E. (1984). The mythic mannish lesbian: Radclyffe Hall and the new woman. *Signs*, *9*, 557–575.

Nosofsky, R. M. (1986). Attention , Similarity , and the Identification-Categorization Relationship, *M*(1), 39–57.

Otten, S., & Epstude, K. (2006). Overlapping Mental Representations of Self, Ingroup, and Outgroup: Unraveling Self-Stereotyping and Self-Anchoring. *Personality and Social Psychology Bulletin*, *32*(7July), 957–969.

Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., & Bailey, J. M. (2004). The influence of sexual orientation on vowel production (L). *The Journal of the Acoustical Society of America, 116*(4), 1905-1908. doi: 10.1121/1.1788729.

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

Podesva, R. J. (2011). The California vowel shift and gay identity. *American Speech 86*(1), 32-51. doi: 10.1215/00031283-1277501.

Podesva, R. J. (2011). Salience and the Social Meaning of Declarative Contours: Three Case Studies of Gay Professionals. *Journal of English Linguistics*, 9(3), 233-264.

Podesva, R. J. (2013). Gender and the Social Meaning of Non-modal Phonation Types. *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society,* 427-448. doi: 10.3765/bls.v37i1.832.

Queen, R. (2007). Sociolinguistic horizons: Language and sexuality. *Language and Linguistics Compass* 1(4), 314-330.

Rendall, D., Vasey, P. L., & McKenzie, J. (2008). The Queen's English: An alternative, biosocial hypothesis for the distinctive features of 'gay speech'. *Archives of Sexual Behavior* 37(1), 188-204. doi: 10.1007/s10508-007-9269-x

Roccas, S., & Brewer, M. B. (2002). Social Identity Complexity. *Personality and Social Psychology Review*, *6*(2), 88–106.

Salmons, J., Jacewicz, E., & Fox, R. A. (2008). Fast talkers versus slow talkers: Speech rate across dialect, generation, and gender. Paper presented at *the American Dialect Society Annual Meeting, Chicago*.

Simon, B., & Hamilton, D. L. (1994). Self-stereotyping and social context: The effects of relative in-group size and in-group status. *Journal of Personality and Social Psychology*, *66*(4), 699–711.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistic Compass, 3*(2), 621-640.

Spears, R., Doosje, B., & Ellemers, N. (1997). Self-stereotyping in the face of threats to group status and distinctiveness: The role of group identification. *Personality and Social Psychology Bulletin*, *23*, 538–553.

Smith, E. R. & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*, 3-21.

Smith, E. R., Miller, D. A., Maitner, A. T., Crump, S. A., Garcia-Marques, T., & Mackie, D. M. (2006). Familiarity can increase stereotyping. *Journal of Experimental Social Psychology*, *42*(4), 471–478.

Smyth, R., Jacobs, G., & Rogers, H. (2003). Male Voices and Perceived Sexual Orientation: An Experimental and Theoretical Approach. *Language in Society, 32*(1), 329-350. doi: 10.1017/S0047404503323024.

Sprouse, R. (2014). ESPSFree. <https://github.com/rsprouse/espsfree> (accessed June, 2017.)

Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology*, *18*, 86-99.

Strand, E. A. & Johnson, K. (1996). Gradient and Visual Speaker Normalization in the Perception of Fricatives. In *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference, Bielfeld, October 1996*, edited by Dafydd Gibbon, 14-26.

Tajfel, H. and Turner, J. C. (1986). The social identity theory in intergroup behavior. In W. G. Austin and S. Worchel (eds.), *Psychology of Intergroup Relations*, 7-24. Chicago: Nelson-Hall.

Tracy, E. C., Bainter, S. A., & Satariano, N P. (2015). Judgments of self-identified gay and heterosexual male speakers: Which phonemes are most salient in determining sexual orientation? *Journal of Phonetics*, *52*(1), 13-25. doi: 10.1016/j.wocn.2015.04.001.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*, 454-463.

Ueda, Y., Hamakawa, T., Sakata, T., Hario, S., & Watanabe, A. (2007). A real-time formant tracker based on the inverse filter control method. *Acoustical Science and Technology*, *28*(4), 271-274. doi: 10.1250/ast.28.271.

Van Borsel, J., De Bruyn, E., Lefebvre, E., Sokoloff, A., De Ley, S., & Baudonck, N. (2009). The prevalence of lisping in gay men. *Journal of Communication Disorders, 42*(2), 100-106.

Van Borsel, J., Vandaele, J., & Corthals, P. (2013). Pitch and pitch variation in lesbian women. *Journal of Voice, 27*(5), 656.e13-656.e16. doi: 10.1016/j.jvoice.2013.04.008.

van Veelen, R.; Otten, S.; Cadinu, M. and Hansen, N. (2016). An integrative model of social identification: Self-stereotyping and self-anchoring as two cognitive pathways. *Personality and Social Psychology Review 20*(1), 3-26.

Verkuyten, M., & Nekuee, S. (1999). Ingroup bias: The effect of self-stereotyping, identification, and group threat. *European Journal of Social Psychology*, *29*, 411-418.

Waksler, R. (2001). Pitch range and women's sexual orientation. *Word*, *52*, 69-77.

Wells, J. C. (1982). *Accents of English* (Vol. 1). Cambridge: Cambridge University Press.

Wright, S. L., & Sara Canetto, S. (2009). Stereotypes of Older Lesbians and Gay Men. *Educational Gerontology*, *35*(5), 424–452.

Wyer, N. A., Sadler, M. S., & Judd, C. M. (2002). Contrast effects in stereotype formation and change: The role of comparative context. *Journal of Experimental Social Psychology*, *38*(5), 443–458.

Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation. Paper presented at *the International Conference on Spoken Language Processing (Interspeech 2006)*, Pittsburgh.

Yuan, J. & Liberman, M. (2009). Investigating /l/ Variation in English through Forced Alignment. *Tenth Annual Conference of the International Speech Communication Association*.

Zhang, Q. (2008). Rhoticization and the 'Beijing Smooth Operator': The social meaning of a linguistic variable. *Journal of Sociolinguistics, 12*, 201-222.

APPENDIX A

EXPERIMENT 1 RECRUITMENT FLYER TEXT

Paid linguistics study on speech and sexuality

The University of California Berkeley's Phonology Lab is collecting information on speech patterns of homosexual and bisexual women. The work aims to understand the stereotypes and reality of a "lesbian accent".

If you are a native speaker of English, you can participate by being recorded reading words and short sentences. Women of any sexual orientation are welcome, though we are looking especially for those who identify as non-heterosexual.

The experiment takes place on the UC Berkeley campus in Dwinelle Hall, room 52. You will be given $5 for 20-30 minutes of participation.

If you are interested, please visit the following site and sign up for a time.

APPENDIX B

EXPERIMENT 1A WORDLIST

above	engineer	introduce	perhaps	soon
absent	enough	Japan	personnel	support
afternoon	equipment	jet	pipe	then
age	evidence	knife	popular	title
authentic	famous	lavender	previous	today
back	felt	locate	quality	tone
barn	food	loose	quiet	typical
became	forever	machine	recommend	under
calm	forget	mass	register	upset
case	garage	meal	represent	usual
cigarette	guess	motor	robbery	vein
concern	guitar	museum	royal	victory
contribute	have	mystery	sad	volume
deal	help	navy	safe	volunteer
defeat	hotel	negative	sample	western
defense	however	niece	seed	wet
division	image	not	selection	wonderful
dose	include	open	sell	worry
early	incomplete	opinion	soap	yard
eight	internal	organize	sock	yellow

APPENDIX C

EXPERIMENT 1A SENTENCES

1. A blue crane is a tall wading bird.
2. A fresh start will work such wonders.
3. After the dance they went directly home.
4. Can you make ice by freezing water?
5. Do both brothers wear the same size?
6. Even the worst will beat his lowest score.
7. Float the soap on top of the bath water.
8. He wrote his last novel there in the inn.
9. How do ducks fly north with no compass?
10. Is that medicine better than the one I have?
11. Should I serve tea from the blue jug?
12. The brown house was on fire to the attic.
13. The club rented the rink for the fifth night.
14. The hostess taught the new maid to serve.
15. The lure is used to catch trout and flounder.
16. What will you eat if I don't make dinner?
17. When did the rush for funds reach its peak?
18. Who slashed the gold cloth to tiny ribbons?
19. Why did they order his head chopped off?
20. Will the play begin as soon as we sit down?

APPENDIX D

ETHNICITY FOLLOW-UP EXPERIMENT

1. Intro

This follow-up experiment was carried out to test the potential influence of ethnicity on acoustic variation in speakers of Experiment 1. There was a worry that skewed distribution might mask the role of ethnicity in variation which was otherwise shown to result from an interaction of sexual orientation and familiarity with Queer culture. The experiment was designed with the idea that were the role of ethnicity significant, listener judgments of ethnicity should correlate in some way with speaker familiarity or sexual orientation. This was not the case. Just as acoustic variation was not successfully predictive of speaker ethnicity, neither were listener ratings, supporting the conclusion that (in the case of these speakers) speech varied according to familiarity and orientation, and not ethnic identity.

2. Methods

There were a total of 30 listeners (19 male and 11 female). All but three self-identified as straight; one male participant and one female participant identified as bisexual and one female participant identified as homosexual.

Listeners heard the same single sentence, "After the dance they went directly home." from each of the 54 original speakers. The experiment was run online using Amazon Mechanical Turk. For each recording, listeners judged the speaker's ethnic identity and rated their confidence in that judgment.

Speaker ethnic ratings were calculated as the total number of ratings each for the ethnic identities chosen by listeners. The effect of perceived ethnicity was tested in an Analysis of Variance with speaker familiarity as the independent variable, speaker ethnic rating as a fixed effect and speaker and listener as random effects.

3. Results

Perceived ethnicity did not successfully predict speaker familiarity with Queer culture in this analysis. It did, however, approach significance ($F_{(18,1469)} = 1.517$, $p = 0.07528$).

4. Discussion

The results of the original production experiment (1a) indicate that familiarity with Queer culture is a primary factor in the variation of speech in the context of this study. However, the skewed distribution of ethnic identity across levels of familiarity made it possible that ethnicity be the true determining factor of acoustic variation among these speakers.

To most accurately test this hypothesis, the ideal subject pool would be limited to a single ethnic group. There have been three studies so far on the sociophonetics of sexual orientation in women which have limited subjects according to nationality and ethnicity, focusing exclusively on Israeli (Levon, 2015), Brazilian (Van Borsel et al., 2013) or Japanese (Camp, 2009) speakers. Other researchers have not reported the ethnic diversity of speakers, so it is difficult to evaluate how the results of Experiment 1a and 1b compare to the influence of ethnicity in those studies, or indeed whether ethnicity varied in those studies at all. The results of this follow-up experiment show that ethnicity was not audibly predictive of familiarity with Queer culture – the primary personal attribute contributing to acoustic variation among speakers in this study.

Based on the results from Experiment 1a and 1b, I argue that speakers' awareness of the focus of the study contributed to the role of familiarity with Queer culture in acoustic variation. Following this argument, I would anticipate variation according to ethnic identity or some related attribute (the equivalent of familiarity with Queer culture) if speakers were under the impression that the focus of this study were on ethnic speech stereotypes. This theory is beyond the direct scope of my research, but certainly merits further investigation, as does the intersectionality of race and sexuality in phonetics in general.