UP:13/09/2013-04:57:53 WM:13/09/2013-04:57:55 M:LG981-7-FY A:12a1 R:1203860 C:D69D6E1B2F6F61C1A01EAD644BF54439FEFF805C

Comprehension & Sociolinguistic Expectations:

Social indexicality of AAVE and MUSE and its effects on speech perception

The University of Essex

2013

A thesis submitted in fulfillment of the requirement for the Master's degree
in Sociolinguistics at The University of Essex by

E. Lenn

Abstract

This dissertation focuses on speech perceptions of African American Vernacular English (AAVE) and Mainstream United States English (MUSE). The following study offers a sociolinguistic investigation of speech perceptions consisting of 62 participants from various locations in Michigan. The use of social indexical cues, including ethnicity, and the perception of AAVE and MUSE speech samples is explored. Participant expectations regarding the visual appearance and sound of speakers is also examined as a potential contributing factor to the perception of speech. Additionally the study examines the role of age and sex in the comprehension of these two American language varieties.

The results of this study show that while an overwhelming amount of participants asserted that they did not use photos, collected data proves otherwise. Social indexical cues were used by participants in transcribing both AAVE and MUSE speech. Linguistic and visual factors of perception of ethnicity were also important. Counterintuitively, speech samples were more accurately transcribed when perceived ethnicity and dialect did not match. Initially, age was not found to be significant, but redefining age boundaries showed that the youngest and oldest participants were less accurate overall than the middle group. Lastly, it is found that female participants were more accurate in transcription than their male counterparts.

Acknowledgments

Thanks are due first and foremost to my supervisor, Dr. Vineeta Chand. Despite the fact that I set part of your cutting board and kitchen counter on fire, you still found it in your heart to encourage and support me throughout this dissertation. I would also like to extend my gratitude to the faculty and staff of the Department of Language and Linguistics at the University of Essex for their help, especially my program coordinator Professor Peter Patrick.

A thanks to one of my best friends, David Bogojevich for being there for me. Without you, I doubt I would have been able to complete this dissertation. Further thanks are also due to my father, who has been a continual source of support throughout my life and for that I am incredibly grateful. Thank you for putting up with my constant calls home to check in on you and Katze.

To the friends and colleagues that I have come to know during my time at Essex-thank you. Though it would be impossible to mention you all by name, know that you have played a significant role in my life and have made it one filled with laughter and fond memories. Remember that you will always be welcome wherever I am.

To all my friends back home in the States that kept in touch with me while I was abroad- I am so lucky to have you and know that I can always rely on you guys, even if there is an ocean separating us.

Finally, the greatest thanks of all go to the individuals who took part in the study, both as the speakers on which the study was built and as the participants themselves. Without your willingness to donate your time and energy to this project it could have never come to be what it is today.

List of Figures

Figure 1: Example of matched guise 16
Figure 2: Map of the United States with the State of Michigan Circled 17
Figure 3: Counties of Michigan where research was conducted 18
Figure 4: Images used in the McGowan 2012 experiment 20
Figure 5: Visual stimuli used in this experiment 21
Figure 6: Chi-test formula taken from Light (2008) 26
Figure 7: Example of audio sound file on Praat 42
Figure 8: Female Silhouette 44

List of Tables

Table 1: Related terms to AAVE taken from Green 2002 2

Table 2: Altered sentences 22

Table 3: Male & female transcription accuracy concerning photos 27

Table 4: Female transcription accuracy concerning photos 28

Table 5: Male transcription accuracy concerning photos 28

Table 6: Male & female participants' questionnaire response of photo usage 29

Table 7: Questionnaire response regarding photo usage separated by sex 29

Table 8: Transcription accuracy of AAVE 31

Table 9: Transcription accuracy of MUSE 31

Table 10: Male & female transcription accuracy of AAVE vs. MUSE 32

Table 11: Female transcription accuracy of AAVE vs. MUSE 32

Table 12: Male transcription accuracy of AAVE vs. MUSE 32

Table 13: Male & female transcription concerning photos & dialect 33

Table 14: Female transcription concerning photos & dialect 34

Table 15: Male transcription concerning photos & dialect 34

Table 16: Male & female transcription of Black & AAVE vs. White & MUSE 35

Table 17: Female transcription of Black & AAVE vs. White & MUSE 35

Table 18: Male & transcription of Black & AAVE vs. White & MUSE 35

Table 19: Transcription accuracy of 19-24 & 40-60 year olds 36

Table 20: Redefined age boundaries & transcription accuracy 37

Table 21: Female transcription accuracy divided by age 38

Table 22: Male transcription accuracy divided by age 38

Table 23: Male & female transcription accuracy 38

Table 24: AAVE speakers & participants' transcription accuracy 39

Table 25: AAVE speakers & female participants' transcription accuracy 39

Table 26: AAVE speakers & male participants' transcription accuracy 40

Table 27: MUSE speakers & participants' transcription accuracy 40

Table 28: MUSE speakers & female participants' transcription accuracy 40

Table 29: MUSE speakers & male participants' transcription accuracy 41

Table 30: AAVE speech rate (syllables per second) 42

Table 31: MUSE speech rate (syllables per second) 42

Table of Contents

Abstract ii
Acknowledgements iii
List of Figures iv
List of Tables v

1. Introduction 1
1.1 Terminology 1
2 Brief Overview of AAVE 3
2.1 AAVE & its Relationship to MUSE 3
2.12 Origins of AAVE 3
2.13 Current State of AAVE: Convergence or Divergence 4
3. Importance of Studying AAVE 5
3.1 Linguistic Discrimination & Linguistic Profiling 5
3.2. AAVE & the Education System 6
3.21 The 1979 Ann Arbor Case 7
3.22 The 1996 Oakland School Board Controversy 7
3.3 Communicative Breakdown 8
4. Language Attitudes & Speech Perception 9
4.1 Matched guise 10
4.2 Speech Perception & Sociolinguistics 11
4.3 Language Attitudes & AAVE 12
5. Research Questions 13
6. Predictions 14
7. Methodology 15
7.1 Targeted Participants 19
7.2 Visual Stimuli 20
7.3 Sentences 21
7.4 Speakers 21
7.41 AAVE Speakers 22
7.42 MUSE Speakers 22
7.5 White noise 23
7.6 Experimental Environment & Directions 25
8. Analysis of Data 25
9. Results & Discussion 27
9.1 Research Question 1a.Does the use of social indexical cues affect participants' accuracy in transcribing speech? 27
9.12 Research Question 1b. Does perceived ethnicity affect transcription accuracy? 30
9.2 Research Question 2: Do people transcribe more accurately when the dialect matches the given visual stimuli? 33
9.3 Research Question 3: 3. Are people ages 40-60 able to transcribe more accurately than those who are 19-24 years old? 36
9.4 Research question 4. Are females more accurate in transcription than males? 38
9.5 Are participants responding to idiosyncratic results for speakers used within the experiment? 39
10. Brief Discussion of Methodological Issues 44
11. Conclusion 45
References 48
Appendix 55

1. Introduction

For this dissertation, I conducted a study manipulating audio and visual stimuli in order to see how participants are able to perceive both African American Vernacular English (henceforth referred to as AAVE) and Mainstream United States English (henceforth referred to as MUSE). The first question addresses whether people are using social indexical cues and whether perceived ethnicity is a factor. It is then asked if participants' expectations of what a speaker should look and sound like affect their perception of what is being said. Age and sex are also explored, as they may play role in AAVE and MUSE comprehension. AAVE has been central to sociolinguistic studies for over five decades, yet there remains a shortage on AAVE in reference to perception and attitudes. It is hoped that this dissertation closes some of the gap, as not much can be found in regards to this particular area; most research on AAVE concerns syntactic (e.g. Labov 1969, 1972, Wolfram 1969) and more recently phonetic and phonological characteristics of AAVE (e.g. Green 2002, Thomas 2007).

In the first section, I will go over the terminology of both AAVE and MUSE. This is then followed by a brief overview of AAVE. The third section explains the importance of this study. It is in section four where the connection between language attitudes and speech perception is explored. Section five introduces the research questions at hand. The research questions are then succeeded by the predictions of this experiment in section six. Methodology comes next in section seven and data collection and experiment set-up is explained in detail. Section eight describes how the collected data was coded and analyzed. Results and discussion make up section nine. After the results and discussion, section ten gives a brief overview of the methodological issues that need to be addressed. Section eleven concludes the overall findings of this study.

1.1 Terminology

Sociolinguistics largely revolves around dialects, which are widely considered as sub-varieties of a single language (Meyerhoff 2006). This brings forth the concept of "descriptivism" and "prescriptivism." Simply put, descriptivism focuses on how and what people say, while prescriptivism involves what people *should* say-this is usually regulated by grammar books and/or other recognized language "authority" sources (McWhorter 1998, Wolfram and Schilling-Estes 2006). Prescriptivism is often associated with what people consider the "standard" way of speaking, which brings up the concept of "ethnolect." An ethnolect is defined as "[a] variety of language which is strongly associated with a particular ethnic population" (Wolfram and Schilling-Estes 2006: 393) and are generally considered non-standard. However, this is not to say that ethnicity is a determining factor in regards to certain speech practices, but it can demonstrate an individual's co-construction and membership to a particular community. As Fought (2006) posits, ethnic

identity is something that is negotiated in a social contingency whereas the acknowledgement of others belonging in a certain community is crucial.

This dissertation will focus on the ethnolect known as AAVE. AAVE has various labels; some commonly used terms are listed below:

Common Labels for AAVE
Negro dialect
Nonstandard Negro English
American Negro speech
Black communications
Black dialect
Black folk speech
Black street speech
Black English
Black English Vernacular (BEV)
Afro American English
African American English (AAE)
African American Language
African American Vernacular English (AAVE)

Table 1: Related terms to AAVE taken from Green 2002

Nonetheless, due to changing times and to give way to more "neutral references," most linguists commonly refer to this particular dialect as AAVE or AAE (Wolfram and Schilling-Estes 2006: 21). Green elucidates that the latter three from the above table are more commonly used (Green 2002: 6).

A brief, but important issue to address is that MUSE is also commonly known as Standard American English (SAE). While many people are led to believe that there is a "standard" way to speak within the US, this is a misnomer, rather it is more often defined "…by what it is not, than what it is" (Cargile et al. 2008). Wolfram and Schilling-Estes (2006: 283) state that MUSE is a variety of English where it perceived as "devoid of both general and local socially stigmatized features, as well as regionally conspicuous phonological and grammatical features." For this reason, I shall not be using the term SAE.

2. Brief Overview of AAVE

Years of linguistic research have shown that language change and the resulting variation is a normal process and because of this, language differences are bound to arise (e.g. Milroy and Gordon 2003, Tagliamonte 2012). Green points out that AAVE has been one of the most studied American English varieties for nearly 50 years (Green 2004: 76). Despite the wealth of research on AAVE, many questions are still left unanswered. Nevertheless, since there has been so much literature on AAVE, it is difficult to cover all areas of research, but a brief overview of AAVE in the US is necessary in order to explain the complex history and the reason as to why AAVE is still a contentious topic. Wolfram and Schilling-Estes point out that three general issues have been brought forth concerning AAVE: the relationship of AAVE to that of MUSE, the development and origins of AAVE, and the current state of AAVE (Wolfram and Schilling-Estes 2006: 212). The latter part addresses as to whether AAVE is converging or diverging towards MUSE. These three general issues concerning AAVE will be discussed below.

2.1 AAVE & Relationship to MUSE

Due to the fact that AAVE shares many linguistic features with that of Southern White Vernacular English (SWVE), much sociolinguistic literature has provided many studies that have been devoted to the comparison of the two (e.g. Bailey 2001, Cukor-Avila 1995, 2001, Fasold 1981, Mufwene 2003, Thomas 2007). This is well documented in the work conducted by Bailey who notes that many shared linguistic features can be attributed to the sociohistorical contexts (e.g. historical impact of slavery in the South) of both AAVE and SWVE (Bailey 2001). Nevertheless, as Thomas denotes, the majority of AAVE research has focused on morphological and syntactic variables (Thomas 2007: 450).

Academics often debate the relationship between AAVE and MUSE, as much of this can depend upon regional differences. For instance, most research on AAVE has been conducted in major urban cities (e.g. Labov 1972, Labov et al. 1968,Wolfram 1969, Anderson 2002) It should also be noted that not all AAVE speakers, similar to speakers of other dialects, use the same range of linguistic features, rather individual AAVE usage lies on a continuum (Green 2002, Wolfram and Schilling-Estes 2006).

2.12 Origins of AAVE

Trying to determine the origin of AAVE is based off comparative data taken from other varieties of non-standard English, such as English in the African diaspora (Poplack and Tagliamonte 1991) and creoles found in the Caribbean (Green 2002: 8). Other sources taken from hoodoo/voodoo texts[1] (Ewers 1995) and ex-slave narratives have expanded views upon AAVE origins, nevertheless, there has yet to be an overall

[1] A large collection of on interviews revolving around witchcraft.
[2] Although Lippi-Green asserts that the idea of a communicative breakdown causes people to perceive

consensus amongst linguists as to how AAVE has emerged. For instance, the Substratist hypothesis argues that West African or other substrate languages influenced the sentences and phonological features of AAVE in reference to English (Green 2002: 8). Other linguists take the Creolist view, which posits that AAVE began as a creole (e.g., Jamaican Creole and Gullah) and attribute this to the fact that many linguistic patterns found in AAVE can be found in creole varieties of English along with other English dialects (Green 2002: 9).

However, due to limited data and different strategies used by researchers to draw conclusions on whether AAVE was initially a creole, it is difficult to determine which view is correct, as both are supported in the literature. The Anglicist view provides another hypothesis on the origins of AAVE relying on the fact that many characteristic patterns of AAVE are found in similar varieties of English. Another hypothesis comes from Mufwene (2000), who proposes that Africans coming into the American colonies would adapt to the local norms of the area (in this case, English), rather than establish their own. Others suggest a mixture of both the Creolist view and that by Mufwene-believing that AAVE was never a creole, but created by African Slaves due to contact between Europeans and Africans in the southern region of the United States during the seventeenth century. This briefly demonstrates that even academics cannot readily agree upon the origins and relationship of AAVE in comparison to MUSE, which is why more research is needed in regards to AAVE.

2.13 Current State of AAVE: Convergence or Divergence

The question of whether AAVE is converging or diverging in regards to MUSE was largely addressed in the 1980s. It began with Labov and Harris (1986) who presented the idea that AAVE and MUSE were diverging in Philadelphia, this was closely followed by the Bailey and Maynor (1987) paper discovering black and white vernaculars were also diverging in the southern region of the US (Ewers 1995: 92). What was compelling is that both studies suggesting divergence were done independently of each other using different research methods, research teams, involving both rural and urban participants, with results applicable to over 130 years (Bailey and Maynor 1989: 12). This led to a heated debate and criticism, as the media became more involved, and was the focus of the NWAVE 1987 panel. A large critique of Labov came from Vaughn-Cooke who believed that Labov's claims of divergence gave rise to "alarming questions about the language-learning capabilities of black people" (Vaughn-Cooke 1987: 13).

While views on the listed issues are still debated upon and there is still no unified consensus on issues brought forth in AAVE research, this leaves most academics calling for more data. For instance, Rickford and Wolfram advocate for more research on perceptions of African Americans that not only focus on linguistic features, but also ideological stances and attitudes towards AAVE (Rickford 1992, Wolfram 1987, 1990). AAVE studies have focused on only certain types of features, chiefly morphological and syntactic features and some phonological features, though there is still a large void within the literature on AAVE perceptions and language attitudes.

3. Importance of Studying AAVE

Moving beyond terms and the brief overview, it is important to surmise as to why the study of perceptions and attitudes toward AAVE are still pertinent. This shall start with what is called "linguistic discrimination" or "linguistic profiling" and move onto the relationship between AAVE and the educational system in the US. Lastly, what has been referred to as a "communicative burden" or "communicative breakdown" will be explored. Another important reason for studying AAVE perceptions and attitudes comes from a recent post from the VAR-list (a public emailing group largely targeted at sociolinguists). John Rickford brought up the recent case involving the testimony of Rachel Jeantel-a 19-year old prosecution witness in the Florida trial of George Zimmerman. Zimmerman was on trial for shooting and killing Trayvon Martin-an unarmed black 17-year old. As Rickford wrote on both the Language Log (a blog consisting of posts created by linguists), and through the VAR-list, Jeantel's use of AAVE during her testimony was found difficult to understand by jurors (Rickford 1 Aug. 2013) and that afterwards, talk shows and various online social media sites, "castigated" Jeantel on her improper grammar and AAVE usage (Rickford 10 July 2013). Not only did this hinder comprehension, it was found that jurors were unable to relate to her or find her credible (Rickford 1 Aug. 2013)-thus serving a prime example of a "communicative breakdown," which is discussed in section 3.3. Rickford notes that most AAVE research revolves around individual grammar or lexical features. He asked for any other research specifically focusing on the intelligibility of AAVE to non-AAVE speakers and/or "…on intelligibility among speakers of other ethnic, social, and regional varieties, especially of English, but of other languages/dialects as well." (Rickford 1 Aug. 2013).

3.1 Linguistic Discrimination & Linguistic Profiling

Many studies within sociolinguistics concerning ethnolects gave rise to the concepts of linguistic discrimination and linguistic profiling. It is noted that language prejudices appear to be more resistant to change than other types of prejudice and that dialect prejudices are some of the last to go (Adger et al 2007: 26). Research on linguistic discrimination and linguistic profiling stems from the work of John Baugh and his associates. Baugh was the first to coin the term "linguistic profiling"-which is when people are appointed a particular ethnicity or ethnicity based on their use of prosody, phonetics, phonology, and dialect (Makoni et al. 2003), particularly when the only information about a person that is made available is their voice; for example, an interviewer may linguistically profile (consciously or subconsciously) a job applicant when conducting phone interviews. Keeping this in mind, research by Purnell et al. (1999) explores the issue of linguistic discrimination and profiling through four experiments in regards to the housing in the United States. Preliminary findings indicate that dialect-based discrimination indeed takes place and that ethnic group affiliation can be easily retrieved from speech (Purnell et al. 1999).

For one experiment, Baugh conducted telephone interviews; it is important to note that Baugh is a tridialectal speaker and able to speak AAVE, MUSE, and Chicano English-which is also known as ChE a variety of Latino English spoken in the US (Lippi-Green 2012: 261). Here, he discovered that there was a pattern between potential discrimination and dialects. To see if participants were able to identify speaker ethnicity at the macro level, they were presented with the sentence, "Hello, I'm calling to see about the apartment you have advertised in the paper", the three target dialects were randomized and it was found that participants were able to distinguish the targeted ethnolect. To see if this held true at the micro level, the utterance of "hello" was recorded by Baugh; in total, it was found that listeners were able to successfully identify tokens more than 70% of the time (Purnell et al. 1990). Therefore, following the findings on the experiments conducted by Purnell et al. (1990), it is seen that people are easily able to identify accents. This is an important aspect of this dissertation, as there are two accents (AAVE and MUSE) that are presented in this study.

3.2 AAVE & the Educational System

Within the US educational system, it has been publicly acknowledged that stratification has separated African Americans from other situated groups. Starting in the 1960s, AAVE has attracted much attention from academics because there was a large interest in minority-group education and attention towards low-income minority-group people and because of the prominent history of African Americans in the US; because of this, language differences between lower-class African American students in these programs in comparison to their middle-class correspondents became ascertainable (Wolfram and Fasold 1974: 33). On top of this, speakers of nonstandard varieties of English have demonstrated lower scores on standardized assessments of English (Hudley and Mallinson 2011: 3), which could explain why there are negative connotations attached to AAVE. Though regional differences are often interpreted by Americans as "matters of quaint curiosity," social status differences play a much more significant role (Wolfram 1991: 91), and unfortunately, in the case of AAVE, those who use AAVE are generally "downgraded on status-related traits" (Cargile et al. 2008). Research on social evaluations in regards to language have been extensively studied and critiqued, and stigmas attached to differing linguistic varieties are often seen as "less intelligent, less wealthy, less powerful and even shorter and less good-looking" (Campbell-Kibler 2010: 379).

Public education serves as a manner in which children are exposed to certain values, attitudes, and beliefs about language and other language varieties; when the public educational system does not comply with the community from which children are coming, conflicts will and do arise (Wolfram and Fasold 1974: 177). These issue within the educational system must be addressed in order to close the gap between groups of "historically underprivileged students" in regards to factors such as gender, social class, and ethnic background to name a few (Hudley and Mallinson 2011: 4). Nonetheless, concerning the issues of

dialects of English, and in this case AAVE, these are rarely ever based solely on linguistic features, but rather extend to socio-political affairs (Green 2002: 217).

3.21 The 1979 Ann Arbor Case

The Ann Arbor case is often referred to as the Black English Trial (Baugh 1998: 282), however is more formally known as *Martin Luther King Junior Elementary School Children et. al., vs. Ann Arbor School District Board* (Smitherman 2006, Rickford and Rickford 2000). Ann Arbor is a city situated on the east side of the state of Michigan. It was here that eleven African-American students were placed in remedial special education classes for failing to adhere to MUSE (Baugh 1998: 282) and given labels such as "…learning disabled, behavior problems, emotionally disturbed" (Labov 1982). In this trial, linguists such as Labov and Smitherman were called to testify and argue that AAVE is in itself systematic and rule-governed. The claims made by plaintiffs were that the school district had failed to take into consideration the students' 'racial and linguistic backgrounds' (Labov 1982: 168). Two years later, Federal Judge Charles W. Joiner, ruled in favor of the *King* children (Smitherman 2006: 12). Although this case brought many linguists together and prompted new research on AAVE (Labov 1982: 195) -as seen in the following section on the Oakland School Board case-this issue was yet to be formally resolved.

3.22 The 1996 Oakland School Board Controversy

This latter case was adopted by the Oakland School Board in late 1996 and early 1997 in the state of California. Oakland is located in the East Bay in the San Francisco region (Lippi-Green: 2012: 306). Rather than being a single school (like the Black English Trial), this case involved the whole school district. A few conditions on Oakland in 1996 were that African American children were not doing academically well in comparison to other children and this was due to the fact many came to school using AAVE (Green 2012: 308). A report from the Oakland School District stated that 71% of students enrolled in special education classes were African American (Vaughn-Cooke 2007: 254). In order for teachers and the educational system to address these issues, and to make education both relevant and rewarding for these children, more funding was needed (Green 2012: 308).

It was decided upon by school officials that in order for African American children to succeed, they needed to be engaged and have access to more funds; at this time, it was recognized that money was available for teaching English to those who spoke a different language. So in order to do this, they proposed MUSE as being the "home language as a conduit" (the presumed first language) and AAVE as its own separate language (Green 2012). In response to the decision of the Oakland School board, many linguists came to the defense. In 1997, the Linguistic Society of America (LSA) released the one of the following statements:

> *The variety known as "Ebonics," "African American Vernacular English" (AAVE), and "Vernacular Black English" and by other names is systematic and rule-governed like all natural speech varieties. In fact, all human linguistic systems-spoken, signed, and written are fundamentally regular. The systematic and expressive nature of the grammar and pronunciation patterns of the African American vernacular has been established by numerous scientific studies over the past thirty years. Characterizations of Ebonics as "slang," mutant," lazy," defective," "ungrammatical," or "broken English" are incorrect and demeaning* (excerpt taken from Vaughn-Cooke 2007: 272*).*

It was not until January 15, 1997 that the Oakland educators present an amended version of their resolution, however, despite their revisions and the statements presented by the LSA, these modifications were and still found to be controversial (Baugh 2000: 47).

3.3 Communicative Breakdown

Why then are the above issues on linguistic discrimination and AAVE within the educational system important? The means by which a hearer responds to linguistic and paralinguistic cues in messages is central to the "language-communication intersection" (Cargile et al. 1994: 211). Because language is observed as a source of social power, it can be molded by people with either unintentional or intentional ramifications (Cargile et al 1994: 2005). Lippi-Green states that upon first sight, communication is seemingly simple-one person talks and another person listens. However, like most things, it is much more complex than this and she notes that rarely ever is communication completely neutral (Lippi-Green 2012: 71). Research has shown that perceived ethnicity or regionality of a speaker has an effect on language attitudes and can lead to what is called a "communicative breakdown"; this is seen in the following study (Cargile et al. 1994: 216).

The Rubin (1992) study explored how expectations of undergraduate students of foreign instructors affected their overall attitudes and learning experiences in the classroom. Expectations and ethnicity were used as variables that affected student perceptions and performance. Students would listen to a short reading passage on an introductory topic (science or humanities). As they were listening, they were shown a picture of what was meant to represent the instructor that was speaking (Caucasian woman, or Asian woman); the recordings used were from the same speaker and readings were of the same passages (this is a prime example of what is called a matched guise and will be explained in further detail in the next section). After listening to the test, they were asked to compare the person speaking to themselves in terms of similarity or differences. The more similar they were to the teacher, the more positive the rating. This was then followed up with a questionnaire asking students to rate the accent, ethnicity, and teaching quality. It was found that students had a higher comprehension level when presented with a picture of a Caucasian female, and lower when the picture was of an Asian female; students also perceived an accent when presented with the photo of an Asian woman, and the higher the level of perceived accent, the lower the

teacher rating. This proves that the visual stimuli and linguistic expectations are significant in the way in which students evaluated language and that the idea of there being a communicative breakdown/burden is truly possible. Lippi-Green believes that these "preconceptions and fear" are strong enough forces to make it possible for students to perceive "imaginary accents, and fictional communicative breakdowns" (Lippi-Green 2012: 94).[2]

Another example of communicative breakdown is a personal one in which I was conducting research for a graduate seminar. The assignment involved asking strangers what they thought of me (e.g. what I studied, where was I from, etc). Important background information in this instance is that I am Korean-American and this particular example occurred in England. The first person I asked to participate in this assignment stated that I was from China and studying accounting or business. After I informed him where I was from, he said he should have known that I was from US because of the way I talked. A similar situation occurred when I was talking to a British woman at a bus stop. Though she was too polite to say where she thought I was from and what I did, she did admit that she should have known where I was from by the way I talked. Countless times both within the US and in England, I would be complimented on my English-Speaking abilities, despite the fact that English is my first language. This demonstrates that despite what people hear, what they see visually can override what is really heard.

4. Language Attitudes & Speech Perception

Speech perception experiments are prevalent in many branches of linguistics, particularly to spoken language processing and the relationship between the acoustic properties of speech, and underlying phonological/lexical features (Nygaard 2005: 391). Campbell-Kibler asserts that the most "prolific" area that concerns social perceptions of languages is that of language attitudes-which explores both emotions and beliefs about certain language varieties and behaviors (Campbell-Kibler 2010: 378). This is shown in Campbell-Kibler's work on the sociolinguistic variable (ing), which is discussed in detail in section 4.2. Speech perception and sociolinguistics intersect, showing that linguistic variation between talkers on the grounds of regional and ethnic difference is an important aspect of spoken language (Clopper and Pisoni 2005: 315). In order to understand how language is related to social domains, language studies should not only focus on how it is produced, but how it is perceived (Campbell-Kibler 2010: 377).

While linguists assert that every language and dialect is equal in terms of communication, there are "subjective differences" on what is deemed as acceptable and this is also dependent on who and what the situational circumstances are (Scott 1999: 1); however, bringing these subjective differences to light can prove to be somewhat difficult. This is not to say that what is considered acceptable in terms of communication is done randomly, rather it can be thought of something that deals with power. Language

[2] Although Lippi-Green asserts that the idea of a communicative breakdown causes people to perceive "imaginary accents" and are "fictional," these are real-life occurrences as many times what is trying to be communicated by a person is misunderstood.

and power is brought up by Bourdieu and his concept of the linguistic marketplace (Bourdieu 1991). The idea of the linguistic marketplace states that certain language practices and behaviors used by those in power are ideologically labeled as being "good," "correct," or the "standard," while other language behaviors and practices used by those who have less power are seen as being "incorrect," "non-standard" and "ungrammatical; in doing so, this allows the indirect study of language and captures the relationship between power and certain social groups (Bourdieu 1991). While there are a multitude of ways in which to gather data on language attitudes and ideologies, a common method used is the matched guise. This method was chosen because it is a sufficient way to collect implicit data on attitudes and ideologies, rather than rely on explicit notions-especially since the notion of AAVE is so disputed within and outside of academia.

4.1 Matched guise

The matched guise was introduced by Lambert et al. in 1960. This technique uses various languages, dialects, and visual stimuli in order to elicit stereotypical or biased acumens within a certain group of individuals (Lambert 1967: 93). The matched guise involves a given stimulus and participants are asked to answer certain questions relating to the perception of one stimulus as influenced by the other; this is done in order to elicit covert ideologies about groups of speakers. Lambert et al. (1960) specifically used the matched guise to determine how language played a part to listeners when it came to evaluating reactions to English and French; it found that English speakers reacted more favorably to English guises than French guises (Lambert et al. 1960). However, it is important to note that not only has the matched guise been used for foreign languages and foreign accents, but it has also been applied to different dialects. For instance, Luhman (1990), employed this technique to compare attitudes towards Appalachian English and MUSE; here, it was established that speakers of Appalachian English are partially accepting of low status evaluations of their speech community, yet renounce other negative stereotypes associated with their speech community in regards to integrity and social attractiveness (Luhman 1990). As seen from these studies, difficult information to obtain, such as status and solidarity can be revealed.

Another well-known study making use of the matched guise was conducted Rubin and Smith (1990). It explored ethnicity, subject matter, and level of accentedness in a series of three studies. Targeted participants were undergraduates and their perception of NNSTAs (nonnative English speaking teaching assistants). The first study is of interest to my experiment, as it too makes use of a matched guise in by manipulating ethnicity via photos. The major results from the Rubin and Smith (1990) study indicated that students had a higher comprehension when presented with a photograph of a Caucasian instructor versus a photo with an Asian instructor; students also reported a higher level of accent when the photograph they were given was of an Asian instructor (Rubin and Smith 1990). This demonstrates that the perceived ethnicity of a speaker affects listener comprehension and the perception of accentedness.

4.2 Speech Perception & Sociolinguistics

Firstly, it is important to define what "perception" entails. In sociolinguistics and other similar areas of study, Campbell-Kibler (2010) elucidates that "perception" is often used to "refer to beliefs or ideologies that people hold on a given topic" (Campbell-Kibler 2010: 378). Language attitudes are pertinent to sociolinguistic studies, as they can enhance models of both social and linguistic behavior-which can then be extended to speakers from other areas of the world (Campbell-Kibler 2010: 378). It is often found that these models of social and linguistic behavior demonstrate the relationship between speaker comprehension and linguistic form. A pioneer of speech perception and language attitudes comes from Preston who conducted research on how speakers in the US viewed themselves and their own speech in comparison to people from other regional areas (Preston 2003). The objective of these studies was to investigate the underlying beliefs, stereotypes, and presuppositions that support the existence of language attitudes, and how these beliefs relate to certain actions and behaviors. These types of experiments have been extended to studies within other countries (Preston 1999, Long-Preston 2002). It is therefore unsurprising that attitudes towards certain languages and varieties are related to attitudes towards groups of people (Preston 2003: 40) and that there are real life repercussions to using stigmatized linguistics forms-such as earlier discussed on Jeantel's testimony in the Zimmerman case.

To demonstrate the connection between speech perception and sociolinguistics Niedzielski's (1999) did such and such and found study demonstrating that people are using social indexical cues in the perception of certain sociolinguistic variables. Previous work in Detroit show that Detroit residents hold certain stereotypes about Canadian English (e.g. Canadian Raising-in which the dipthong /aw/ is raised and fronted); on top of this, research has shown that Detroiters believe the dialect of White, middle-class residents of Detroit speak MUSE (Niedzielski 1999: 63). Nevertheless, research stemming largely from Labov (1994) shows that Detroiters do not speak MUSE and are involved in the Northern City Chain Shift (NCS or NCCS)- this is the dialect generally spoken in the state of Michigan[3].

Niedzielski then asks why Detroiters do not recognize Canadian Raising and NCS-vowels in their fellow Detroiter's speech and if their expectations of what Detroit residents *should* sound like has a greater impact than received acoustic information. It is asked that if led to believe that a speaker is from a particular area, would Detroiters be more likely to notice certain stereotypical linguistic features. Niedzielski conducted a series of perceptual tests on resynthesized vowels and presented it to Detroit residents who were then asked to identify vowels that best fit the speech of a fellow Detroiter. Half of the participants were told that the speaker was from Canada, and the other half were told that the speaker was from Detroit. When given words containing the stereotyped /aw/, it was found listeners heard the stereotyped feature moreso when the speakers were labeled as Canadian, however, when labeled from Michigan-participants were less likely to "hear" the raised variant. Since the only item that was changed in

[3] See Gordon 2001 and Labov et al. 2006 for a more extensive explanation on the NCS.

this experiment was the label of nationality, this suggests that listeners do indeed use social indexical cues when it comes to perceiving speech. Niedzielski's use of labeling speakers from a certain region is called priming, which is when participants have a heightened sensitivity to certain stimulus due to prior experience (Cherry). The concept of priming is of importance to this study because I too use this technique, but rather than using labels denoting speaker nationality, I use photos to suggest ethnicity.

To further connect speech perceptions and sociolinguistics, take the work conducted by Campbell-Kibler (2005, 2007, 2011) concerning the sociolinguistic variable of (ING). In her 2005 dissertation, the matched guise was used to understand the structure of sociolinguistic variation and how information is conveyed to listeners; it was found that (ING) is involved in a "network of social meanings" (this includes education, articulateness, formality, region, and rural/urban divide"). Campbell-Kibler (2007) finds that the alveolar form *–in* and Southern accents were perceived as a lack of education, the country (rural), and associated with the term "redneck." The –ing and gay accent represented lower masculinity, the city (urban), and the associated with the term "metrosexual." In her 2011 paper, Campbell-Kibler also used matched guises involving *–in, -ing*, and a neutral guise with no audible (ING) tokens. Here, *-ing* was rated as more intelligent/educated, articulate, and less likely to be said by a student than *–ing* and neutral guises-which behaved similarly in which they did not differ significantly; *–in* guises were judged as less formal and less likely to be gay than the *–ing* and neutral guises. All of these studies prove that listener perception can also be affected by linguistic variation and that these linguistic variations can carry certain social meanings-demonstrating that social indexical cues can be both linguistically and visually oriented.

4.3 Language Attitudes & AAVE

Lippi-Green (2012) posits that in language-focused discrimination, it is not necessarily language that is important, but rather how an individual's beliefs about language affect institutional practices (Lippi-Green 2012: 67). This brings language ideologies and attitudes to the forefront of applied sociolinguistics. As previously mentioned, there is a lack of studies concerning language attitudes towards AAVE, but there are a few studies that are worth mentioning in regards to AAVE language attitudes and perceptions.

Teacher attitudes towards AAVE are something that has also been studied throughout the years. Calloway (1989) found that secondary teachers believed that the decline of AAVE would have a positive effect on society and that there was a correlation between "speech output" and "academic achievement." A survey conducted by Taylor (1973) of 422 teachers' attitudes toward AAVE found that forty percent of the respondents reacted negatively towards the "structure" and "usefulness" of AAVE, forty percent were positive on this topic, and the remaining twenty percent responded neutrally. Although a fair amount of teachers had a positive outlook on AAVE, it is important to note that there was an equal amount that responded negatively, implying that more research and educational instruction needs to take place in regards to AAVE.

In 1996, the dissertations of both Boyd and Campbell discovered that elementary teachers had negative attitudes towards AAVE (Boyd 1996, Campbell 1996). These two dissertations highlight the possibility that these particularly negative attitudes could adversely affect teacher interactions with AAVE speaking children. Ervin (2005) was a bit more complex and looked at faculty reactions to AAVE from 2 year colleges and 4 year universities; here it was found that female faculty had more favorable attitudes towards AAVE in regards to the structure, inherent usefulness of AAVE, and the philosophical principles concerning the use and acceptance of AAVE in the educational setting than their male counterparts. It is also found that attitudes of faculty members at the 2-year colleges demonstrated more acceptance of AAVE than faculty at 4-year universities. So as seen from the latter example, negative attitudes towards AAVE can be found in the elementary schools and at 2 and 4-year colleges.

It is also of importance to see how other African Americans react and what their views are towards AAVE. Rahman (2008) specifically examined middle-class African Americans. It was found via a variety of methodologies (including online survey and interviews), that there was a correlation between the perception of African American identities and judgments; while participants appreciated AAVE as their "heritage language," they saw MUSE as the variety where they could "meet all demands of all environments" (Rahman 2008: 141). In a recent dissertation written by Oliver (2012), it was found that African American preschool teachers' attitudes towards the use of AAVE by students were neutral. As seen from both of these studies, there are variations on language attitudes and perceptions of AAVE from within the African-American community.

5. Research Questions

The purpose of this study was to see how people from the state of Michigan and from two different age groups react to the dialects of AAVE and MUSE that are primed with visual stimuli. This is done in order to explore how these stimuli affect people's perception of what they hear. This dissertation is broken into a series of five research questions and investigates the following:

1a. Does the use of social indexical cues affect participants' accuracy in transcribing speech?
1b. Does perceived ethnicity affect transcription accuracy?
2. Do people transcribe more accurately when the dialect matches the given visual stimuli?
3. Are people ages 40-60 able to transcribe more accurately than those who are 19-24 years old?
4. Are females more accurate in transcription than males?

Though there were initially four questions, one last question presented itself while coding and analyzing the data. The question asks as to whether participants were reacting to idiosyncratic results to the speakers that were used in this experiment. This is explored after research question four.

6. Predictions

Firstly, it is predicted that people are using social indexical cues when transcribing speech and this will affect how accurately they transcribe. The idea of people using social indexical cues is not something new in sociolinguistics (e.g. Campbell-Kibler 2005, 2007, 2010, 2011, Niedzielski 1990, Labov 1963, Rahman 2008, Staum 2008). For instance, in the classic sociolinguistic study conducted by Labov in Martha's Vineyard on the dipthongs /au/ and /ai/, a major finding was that fisherman centralized both /au/ and /ai/ in order to show their social identity as islanders and to separate themselves from islander outsiders; this non-standard dialect involving the centralization of /au/ and /ai/ represented a linguistic divide and desirable values that were often associated with the fishermen (Labov 1963).

The sub-question of the first research question asks as to whether the perception of ethnicity affects transcription accuracy. Of course social indexical cues are not only used to demonstrate solidarity, but also serve as a division between groups of people. Ethnicity has also been shown in psychological and other similar areas of study to play a role in how people are categorized and judged (e.g. Purnell et al 1999. It is then predicted that participants will use social indexical information revolving around ethnicity (both visually and linguistically) and this will in turn affect how accurately they perceive the given stimulus.

The second research question addresses both visual and audio stimuli concurrently. Keeping research question 1a. and 1b. in mind, it is hypothesized that if the given dialect and photo match (in terms of ethnicity), then participants will be able to transcribe more accurately than if they did not. Thirdly, it is asked as to whether the 40-60 year olds are more accurate in transcription than those who are 19-24. Assuming that participants are using social indexical cues, it is believed that age may be a factor. In the study conducted by Tan (2012), results suggested that young Singaporeans are "deaf" to ethnic variations and it was shown that the older participants were more accurate in identifying the ethnicity of speakers than their younger counterparts. Tan (2012) attributed this largely to the fact that those that were older had more exposure and experience to other languages. For these reasons, it is predicted that those who are 40-60 will be more accurate in their transcriptions than those 19-24.

Lastly, gender is also an important area to address in regards to this experiment. Loudermilk's (submitted) study on the sociolinguistic variable (ing) found that females were more perceptive to the non-canonical form (in'). On top of this, it has been shown repeatedly across sociolinguistic literature that women are largely responsible for being the linguistic innovators (e.g. Arana 2013, Bohn and Mastumoto 2008, Labov 1990, Haeri 1994). Therefore, as linguistic innovators, it is also plausible that females more linguistically aware. Due to females' heightened linguistic perception, it is believed that they will be more accurate in their transcriptions than males.

7. Methodology

As shown the previous studies cited above, there are many ways in which to elicit language attitudes. While useful, Preston's map-drawing task was not appropriate for what I wanted to study. Therefore I chose the matched guise to elicit language attitudes. To further correlate the matched guise, a questionnaire was also included at the end of the experiment. Nevertheless, these two techniques reveal explicit information, rather than implicit information. For gathering implicit information, it is possible to use what is called an "Implicit Association Test (IAT). This test was developed by Greenwald et al. (1998) and is a timed reaction test in which participants sort positive and negative adjectives into good or bad categories tied to people belonging to a certain group, such as ethnic groups. This test is often ideal for studies with racial attitude implications, as many times participants are unwilling to share what is on their minds publically in regards to such contentious topics. However, had I used an IAT, participants' perceptions of MUSE and AAVE dialects would have been tied to the semantic content of what they were saying.

Another way to elicit language attitudes is through the sociolinguistic interviews. Interviews can prove to be quite useful when exploring language attitudes and perceptions as "they can be a way of exploring relationships between different aspects of a situation" and "interviewing is a powerful way of helping people make explicit things that have hitherto been implicit" (Arksey and Knight 1999: 32). Despite the usefulness of this method, it was not logical for this study due to large sample size and time constraints, which would not have allowed me to go thoroughly through every interview.

However, unlike the traditional matched guise, the matched guise in this experiment includes visual and audio stimuli that are matched and mismatched. Cargile et al. (2008) notes that most speaker evaluation studies of AAVE have only used male speakers, so in this experiment, I decided to use female speakers and photos. For reference, Figure 1 is a diagram that briefly demonstrates what is meant by "match" and "mismatch." The green boxes are representative of when the visual stimuli (photo) matched the audio stimuli (dialect; AAVE or MUSE), the blue boxes are when they did not match. The last two photos in the diagram below represent the control part of the experiment (silhouette), in which the Black and White photos were compared against

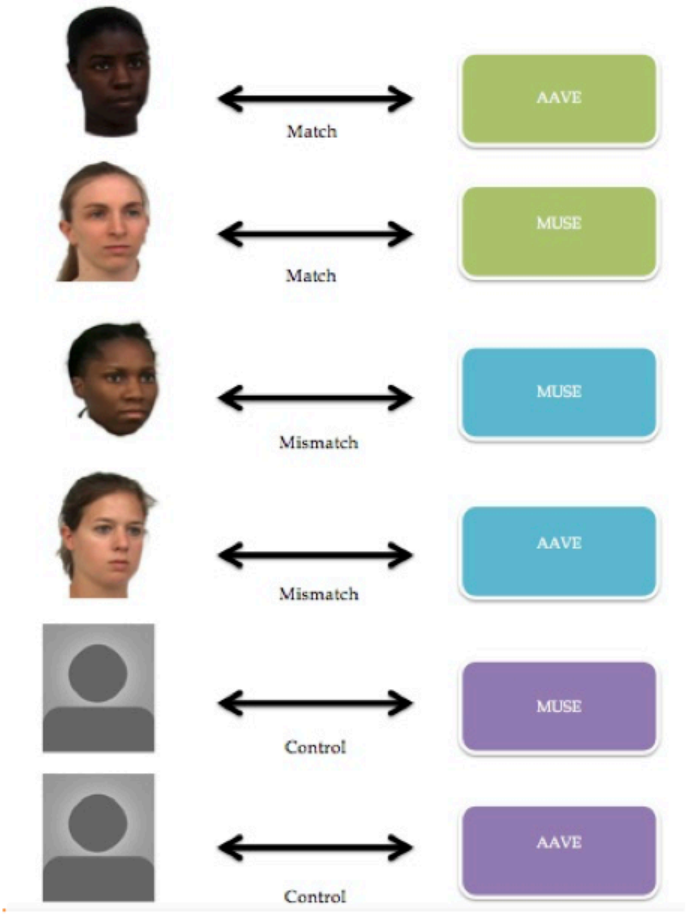

Figure 1: Example of matched guise

Six speakers were used as the linguistic input in this experiment; three were AAVE speakers and three were MUSE speakers and each speaker was recorded saying five sentences. These sentences were taken from a previous experiment conducted by Bradlow and Alexander (2007) and are discussed in more detail in section 7.3. In total, participants were presented thirty sentences with added white noise in the background; fifteen sentences were of AAVE and fifteen were of MUSE. These thirty files were put into a PowerPoint.[4] First participants were presented with a photo and the next slide depicted the same photo with audio. Pictures and sentences were quasi-randomly placed in order to prevent the same picture and audio spoken by the same speaker being one after another. At the start of the experiment, participants are asked to sign a consent form[5] and then given a worksheet[6] that gave them space to write down their transcriptions and at the end a brief questionnaire[7] is also given. The questionnaire consisted of both open and closed questions, and asked general background questions (e.g. age, sex, where participants grew up). A

[4]Although there were over 70 slides presented in the PowerPoint, in the Appendix A is a brief sample of what was shown and how slides were set up.
[5] Consent form found in Appendix B.
[6] A brief sample of the worksheet given to participants is shown in Appendix C.
[7] The questionnaire given to participants is shown in Appendix D.

questionnaire seemed appropriate as they are often used to obtain attitudes and perceptions of certain languages or dialects (Rasinger 2010).

All the data for this dissertation were collected in various regions of Michigan, located in the Midwest region of the United States. After all the data was collected, it was coded using the program IBM SPSS 19. Right and wrong answers were evaluated by looking at the target word, which was the word at the end of each sentence. For example, with the sentence "A book tells a story," the target word in this sentence would be "story."

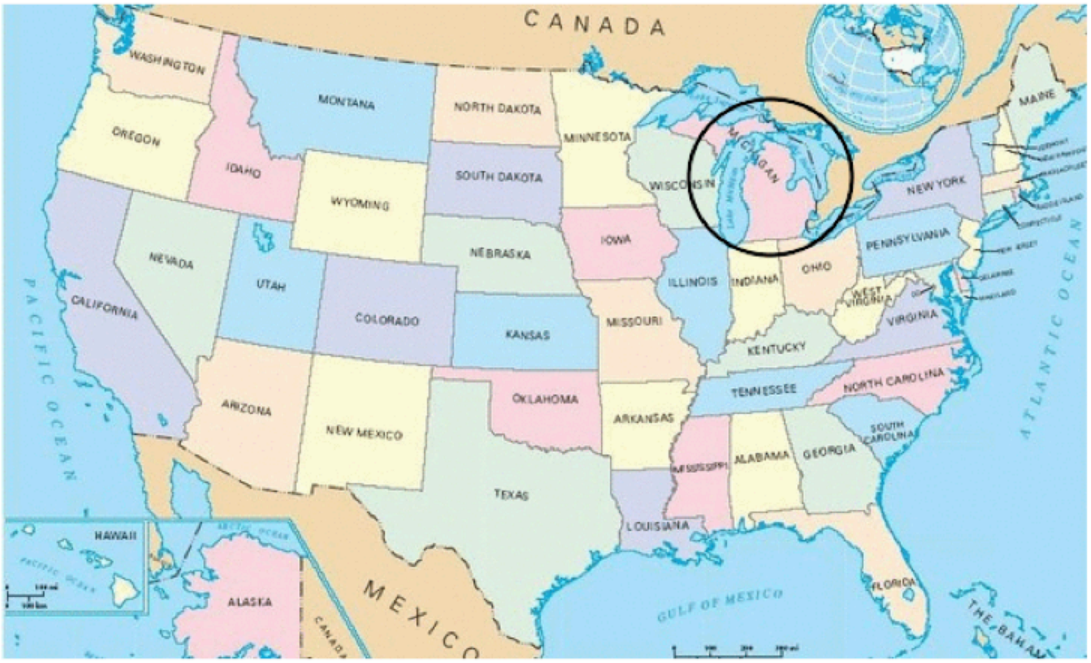

Figure 2: Map of the United States with the State of Michigan Circled[8]

Data was collected from Montcalm County-located in mid-Michigan; East Lansing (Ingham County)-a city located adjacent to the state capital of Michigan; Mt. Pleasant (Isabella County)-located in the northern-central part of Michigan; Ann Arbor (Washtenaw County) in the southeast region; and on the east side of Michigan in Bloomfield Hills (Oakland County)-a suburb of Detroit. The majority of this data was collected at libraries from three major Michigan universities-Michigan State University, the University of Michigan, and Central Michigan University.

[8] Original map found here: http://www.newyorkstatesearch.com/United_States_Maps/United_States_Map/United_States_map.jpg

Figure 3: Counties of Michigan where research was conducted [9]

Michigan was chosen because this is the state I most readily had access to. Linguistically, Michigan has been has largely been studied in regards to the NCS (e.g. Gordon 2001, Labov 1994, 2006), however, has also been used in a well-known speech perception study conducted by Preston (2004). Here, language attitudes and perceptions are captured via a map-drawing task[10] and it is revealed that Michigan residents consider themselves as speaking the most "correct" English out of all the states in the US To linguists, this is also known as "linguistic security," which is another reason as to why Michigan is an interesting region to conduct this study.

---

[9] Original image found here: http://www.econsultant.com/usa-state-maps-with-county-lines-county-names/mi-3-state-map-with-county-lines-county-names.gif

[10] Participants are given a blank map and asked to label dialect regions according to their perceptions of a certain region.

7.1 Targeted Participants

Similarly, like Rubin and Smith (1990) and McGowan (2012), my targeted participants were college-aged (19-24 years old); however, I also collected data from those that are older (40-60 years old). While McGowan (2012) looked at "experienced listeners" and "inexperienced listeners," it was deemed too difficult to separate listeners into "experienced" and "inexperienced" groups because-as with many other non-standard speakers found across the globe-many speakers are aware of social judgments associated with certain linguistic practices and will not acknowledge that they speak a "different" or "marked." Determining those that do have regular access or interactions with AAVE speakers is also difficult, as the concept of AAVE is an ongoing controversial and heated debate within the US-both within academia (e.g. the origin of AAVE or whether diverging or converging towards MUSE) and outside of academia (as seen from the Black English trial and Oakland school Board Case). Nonetheless, the majority of the public tends to take a prescriptive stance towards AAVE). Nonetheless, this does not conceal the fact that "experience" or exposure to AAVE may be a factor and affect results. In order to reconcile this problem, I had a questionnaire attached at the end of the experiment that asked if people spoke AAVE, had a lot of contact with people who used AAVE, if they were familiar with AAVE, and if they were familiar with hip-hop and rap culture. These are all questions that help determine whether "experience" or "inexperience" is a factor when it comes to transcribing, and will be discussed in the results and discussion section of this dissertation.

Both Rubin (1992) and McGowan (2012) gathered their data from major universities and all participants were undergraduates. Similarly, to keep the sample of participants relatively homogenous, I decided to make sure that the participants I was targeting also had some form of higher education. I wanted to ensure that my participants did not come in knowing that I study language, as this knowledge could possibly change the outcome of the data. Lastly, I excluded those with hearing disabilities, as this would also skew results. While participants from McGowan's study were recruited via undergraduate classes and were either given partial course credit or paid $15.00 for partaking in the experiment, my circumstances did not allow for either of these incentives, so finding participants willing to take the time to complete my task within the given timeframe proved challenging.

Although, data was collected from 75 participants, thirteen were excluded; three did not fit the targeted age range, one did not meet the educational requirements, another was discarded because the participant knew that I was studying linguistics, six had a hearing disability, one was excluded because they were from another country, and the last one was excluded because they did not follow the directions and listened to sentences multiple times. This leaves 62 participants who will be considered in the data analysis, including 16 males and 15 females between 19-24 years old and 13 males and 18 females between the ages of 40-60.

7.2 Visual Stimuli

Originally, I intended to follow McGowan's method for providing visual stimuli. An example of which is provided below:

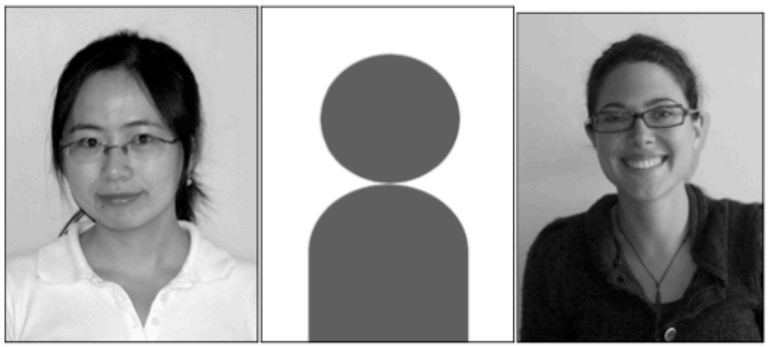

Figure 4: Images used in the McGowan 2012 experiment

The images that McGowan used were license-free portraits that were found online via web search. Then an informal survey of several linguistic graduate students was used to determine which photos to use. In the present study, I intended on taking a series of photos of personal contacts and use an informal survey to help me choose which photos to use in this experiment. However, in doing so, I found that it was difficult to get people who had similar features. Instead, I decided to find photos that were more consistent and did not allow for other biases to emerge based on how the person was dressed, how they smiled, etc. To do so, I took my photos from the Face Place, which is a face database originally created by the Tarrlab at Brown University, but is currently hosted at Carnegie Mellon; it consists of images of over 200 individuals and of different ethnicities under the same lighting, hosting different emotions and disguises, and taken from different camera angles (Tarr). No article of clothing was shown, as this too could potentially cause some bias among participants.

In total, seven photos were used; six from the Face Place and one of a geometric silhouette found through a web search. Similar to McGowan's study, the purpose of the silhouette image was to serve as the controlled variable. Three images of African American females and three images of White women were used, each of these images were repeated four times throughout the experiment. There were also six instances where the silhouette was used. These were then placed in a PowerPoint in a pseudo-random manner that ensured that the same photos were not placed one right after another.

Figure 5: Visual stimuli used in this experiment

Overall, participants were exposed to thirty photos, each accompanied by a spoken sentence; there were fifteen sentences in each dialect (AAVE and MUSE) presented. Every participant was shown six pictures of a silhouette, twelve of Black and White photos; out of the twelve Black and White photos, six photos matched the perceived speaker dialect and photo and six would not match the dialect and photo. For example, for mismatched photos, the White photo would have the AAVE dialect and vice versa. The experiment took about fifteen minutes to complete and a questionnaire was given to participants after they completed the experiment.

7.3 Sentences

In the work conducted by McGowan (2012), sixty sentences were used; thirty were high predictability sentences and thirty were deemed low predictability sentences. All of these sentences were taken from an experiment done by Bradlow and Alexander (2007). The high predictability sentences were easily able to be completed by groups of non-native and native speakers of English; for instance: "Elephants are big _____" (Elephants are big animals). The low predictability sentences were sentences that were not easily recognizable or completed. For the purpose of this study, I only used the low predictability sentences because all of my targeted participants were native English speakers and I wanted to measure what they heard and did not want to risk them correctly guessing the target word because of semantic recognition due to the surrounding words or their best guesses. [11]

[11] See Appendix E for the list of the low predictability sentences used in this experiment.

Three sentences were changed at request of one of my AAVE speakers. She noted that in AAVE, she would change the sentences by deleting the word "is." Table 2 shows what was changed:

Original Sentences	Altered Sentences
This is her favorite sport.	This ⌀her favorite sport.
This is her favorite week.	This ⌀ her favorite week.
This is her favorite time.	This ⌀ her favorite time.

Table 2: Altered sentences

This linguistic feature of AAVE is also called zero copula deletion and has been well documented in sociolinguistic literature (e.g. Labov 1969). In layman terms, zero copula deletion is a systematic and rule-governed feature in which there is an absence of the copular verb *to be*. Rickford and Rickford (2000) elucidate that zero copula deletion is a "distinctive and identity-affirming" characteristic of AAVE (Rickford and Rickford 2000: 125); Wardhaugh notes that this linguistic feature is rarely found in the speech spoken by white or lower-income Southern white speakers, though it is important to remember that not all African Americans use this feature (Wardhaugh 2010: 365). Nevertheless, because zero copula deletion is a distinguishing feature of AAVE, I decided to keep these altered sentences and distribute them equally amongst the photos, although recent work from Thomas and Reaser (2004) demonstrate that a high degree of Americans can distinguish African American voices based on phonetic features and with the absence of stereotypical syntactic and morphological features characteristic of AAVE.

7.4 Speakers

In total, six female speakers were recorded using the Olympus VN-712PC voice recorder and sound files were saved as .mp3 files and then converted to .wav formats. All were native English speakers and grew up in Michigan.

7.41 AAVE Speakers

There were a total of three AAVE speakers. The first speaker is a bidialectal (AAVE and MUSE) 23-year old female who grew up in Detroit. The second AAVE speaker is also bidialectal (AAVE and MUSE). She is 21 years of age and grew up in Lansing, Michigan. The third AAVE speaker used was someone I had not previously known. This speaker was 18 years old and had also grown up in Detroit. I had met her in one of the dormitories at MSU and recognized that she was an AAVE speaker due to the manner in which she spoke to other people. Much linguistic research has demonstrated that people who are talking to strangers or people they do not know usually resort to a more formal register or shift styles to a more socially accepted style of speech; Labov introduces what is known as "The Observer's Paradox,"-which occurs

when an informant in an interview is aware that they are being monitored, therefore causing them to speak differently (Labov 1972). Although this is in reference to what is known as a sociolinguistic interview, this can be extended to my experiment. Labov's research (1966) shows that those participants that talk to people who they are unfamiliar with or that are being recorded tend to resort to what is referred to as "careful speech." Although the Observer's Paradox is framed as problematic in capturing vernacular speech, I asked this last AAVE speaker if she would let me record her reading five sentences. After recording her, I observed that she continued in the same manner and kept using AAVE. So therefore, this shows she was not influenced by the Observer's Paradox.

7.42 MUSE Speakers

The three speakers that I chose to read sentences in MUSE were unknown to me before I recorded them. I met them in one of the dormitories that housed students for the summer semester at Michigan State University. The reasoning for recording strangers was to ensure that people were using what they thought is a "standard" accent. Within the United States, it is noted that there really is no "standard" or "prestige," but there is a notion of what constitutes as such. In other words, a certain language or dialect is designated as "normative" in comparison to others; Wolfram and Fasold generalize this notion of "standard language" as being "a codified set of language norms which are considered socially acceptable to the most prestigious social class in society" (Wolfram and Fasold 1974: 18). Following Labov's idea of the Observer's Paradox and its ability to capture what people believe to be standard speech or the prestige, this is why I believe it was best to record strangers to get my MUSE recordings.

7.5 White noise

White noise was added to the background of the sound files via a Gaussian equation in order to prevent participants from getting all of their answers correct in order to avoid "ceiling effect." This was done through the programs of Praat version 5.3.48 (Boersma and Weenick 2013) and Audacity 2.0.3 on a MacBook Pro™ OSX version 10.8.3. The Gaussian equation goes as follows:

$$1/2 * \sin(2*\pi*377*x) + \text{randomGauss}(0,0.1)$$

To alternate the amount of white noise, the 0.1 part of the equation can be changed; the higher the number, the more white noise is added to the sound file. It was hoped that the added white noise would allow enough transcription errors that the ceiling effect would be avoided and data would yield interesting results. Added noise also made sense in the fact that rarely ever does a listener not have some type of added background noise when conversing with other people. It should be noted that the McGowan (2012) study

used multi-talker babble over other types of noise and had done this to match the plausibility of stimuli for participants. Rather than doing this, I decided to add white noise because with present day situations, noise is not just coming from people, but from electronics and other sources. On top of this, it was thought that multi-talker babble could potentially interfere with the audio stimulus that was being presented.

A large issue with this experiment and what proved to be the most time consuming was gauging the appropriate amount of white noise to be added to the background. In speech perception experiments, there is little research that specifically tells how the researcher determines how much added noise is enough. For instance, in McGowan's (2012) study, he wrote that they conducted a series of pilots using the full set of experimental sentences with no noise and then gave an informal listening task to several researchers who were unfamiliar with the content of the sentences. The researchers then suggested mixing the noise with a certain signal so that the result would allow a "sufficient" amount of transcription errors for the purpose of the experiment. Though it is difficult to determine what is a "sufficient" amount of noise to add to the sound files that would allow for the desired proportion of transcription errors, this is why three total pilot studies were conducted.

Firstly, white noise via a Gaussian equation on Praat was added to the audio files; the highest began at 0.4 and went down in increments of 0.05. In total, there were audio files with noise of 0.4, 0.35, 0.3, 0.25. 0.2, and 0.15. The audio files were then sent to three people who did not know what the experiment was and asked them to start listening to the files starting with the highest amount of white noise (0.4) and then pick the sound file where they could somewhat accurately hear what was being said, but not too easily. The overall consensus was at 0.15. From thereon, a PowerPoint was created and added the sound files to the pictures and ran another informal pilot was conducted. Results from the latter pilot showed that there was still too much noise. Because of this, one last informal pilot study was conducted where it was found that 0.10 made the transcription task too easy and a ceiling effect was shown. So to average this out, 0.12 was chosen as the set amount of noise that was to be added to the sound files.

The difficulty in adding noise to sound files is that once noise was added, it cannot be adjusted. Another issue that arose is that given the file format, the files could not be opened via Praat. I was able to rectify this issue by opening the sound files using Audacity. Then I would export the file to my desktop saved as a .wav file. In doing so, this allowed me to open the sound files through Praat and add noise to them. In no manner does this degrade the audio quality.

7.6 Experimental Environment & Directions

Participants were seated in a quiet area and used a pair of noise cancelling BeatsAudio™ executive headphones to listen to the audio stimuli. They were also provided with a worksheet that allowed them to write down the transcribed sentences as well as complete attached questionnaire. Prior to the start of the task, participants were given a clip of white noise lasting for about five seconds order for them to get accustomed to the noise. This was then followed by a practice clip with the sentence "Elephants are big animals" taken from one of Bradlow and Alexander's (2007) high predictability sentences. The directions that appeared to the participants were as followed:

Transcription Task

- You will not be timed.
- **There is no right or wrong answer.**
- This task will take about 10 minutes to complete.
- It includes 1 practice sentence, 30 short sentences, and a brief questionnaire.
- Please write down your answers as **neatly/legibly** as possible.

Directions:

- A photo is given.
- To move to the next slide, push the → arrow.
- A photo appears. Push the à arrow again and an audio icon will appear.
- Push the → arrow to listen to the audio.
- After listening to the audio, **write down what you heard to the best of your ability**
- You may only **listen to an audio one time**.
- If you are unsure, try your **best guess** of what you think you heard.

Please note:

- Audio files will have white noise added to the background and volume may change because there are different speakers who are speaking, so listen carefully.

White noise:

- This is what the white noise will sound like.
- Adjust the volume accordingly.

8. Analysis of Data

Both the data from the experiment and questionnaire were coded using the statistical program IBM SPSS 19. IBM SPSS was used to perform a chi-squared test, excluding research questions one and two. IBM SPSS 19 was inappropriate to use for these research questions because it did not take into consideration the fact that the control population for the photo response questions was based on the visual stimuli of the silhouette-rather it included the silhouette, black photos, and white photos in the statistical analysis. Therefore, instead of using the chi-squared function via IBM SPSS 19, an online chi-square calculator (Ryan 2013) was used because numbers could be added manually.

The Pearson chi-squared test was deemed appropriate, as enough data was gathered to prove useful; the Fisher's exact test was also considered, but this is generally done when presented with smaller amounts of data (Gorman and Johnson 2013: 214). This chi-squared test is also referred to as a goodness-of-fit test and is used to see the "...proportional distribution of categoric variables" (Garner 2005: 235). There were some instances where participants got the majority of transcriptions correct and some instances where most participants transcribed the sentence incorrectly. However, none of this data was excluded due to the fact that it is advised that when performing a chi-squared test, nothing should be eliminated (Light 2008), and because as a researcher and sociolinguist, all variation should be accounted for. Excluding certain sentences due to participants' performance could therein possibly skew the statistical outcomes.

Perhaps one of the most well known sociolinguistic study employing the chi-squared test was Labov's 1966 study on the stratification of post-vocalic *r* in New York City. Here, he elicited tokens from the phrase "fourth floor" from store employees from three Manhattan department stores (Kleins vs. Macy's vs. Saks) representative of three different social classes (lower class, middle class, upper class) in New York (Labov 2006). To ensure that the probability of the use of post-vocalic *r* was not due to chance, the chi-squared test was used; though his use of the chi-squared test was not published in the original 1966 paper, it is discussed in further detail in other works specifically about different methodologies and approaches used within sociolinguistics (e.g. Gorman and Johnson 2013).

To uncover as to whether the data demonstrated some correlation with the given input, the chi-squared test was administered. The formula for a chi-squared test goes as follows:

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Figure 6: Chi-test formula taken from Light (2008)

X^2 is the sum of the chi-square, *O* represents the observed frequencies found in the data, and *E* is the expected outcome. The chi-squared test essentially checks as to whether observed frequencies, or what was found in the data, is close to the frequencies of what was expected -this is known as the null hypothesis (Garner 2005: 194). The alpha value set prior to this experiment is 0.05, this means that there is less than a five percent chance that the differences of the expected outcomes were due to chance. If the probability value (hereon referred to as the "p-value") exceeds the alpha value of 0.05, results are determined to not be significant and the null hypothesis must be accepted. If p is less than 0.05, then the null hypothesis is rejected. To get the p-value from the chi-squared result (X^2), a probability table with degrees of freedom (df) must be consulted. The df indicate how many variables in the table are truly independent, and simply put, "are equal to the number of rows minus one times the number of columns minus one" (Light 2008).

The dependent variable was participants' accuracy, and the independent variables that were being manipulated were the visual and audio stimulus. The data is also divided into male and female to see in closer detail how it was dispersed; Salkind (2004) suggests that the parametric statistics must fulfill the assumption that a given sample is representative of a population and that to fulfill this assumption, data must be collected from a sampling size of about 30 people (Salkind 2004: 285). However, in the case of dividing up males and females with my data, it came to there being 29 males and 33 females. Therefore, a small amount of caution must be taken in regards to results referring to only male participants.

9. Results & Discussion

The following section will introduce and attempt to answer the research questions that are asked in this dissertation, pulling from data collected from the experiment and questionnaire.

9.1 Research question 1a: Does the use of social indexical cues affect participants' accuracy in transcribing speech?

My first research question asks if the use of social indexical cues affect participants' accuracy in transcribing speech. Data from the experiment indicates that photos were indeed a significant factor. Results showed that participants transcribed most accurately when presented with the silhouette photo than with a Black or White photo. Since the control variable is the silhouette photo, this in turn will be compared against the Black and White photos. The chi-squared for Black photos (df=1) was significant at 67.7456897, resulting in a p-value < 0.001, and the null hypothesis is rejected, which means that Black photos affected participants' transcription accuracy; out of the three types of stimuli presented, participants transcribed audio stimuli with Black photos less accurately than the White and silhouetted photos. For White photos, the chi-squared was 13.2535022 (df = 1) with a p-value < 0.001- also significant- so therefore the null hypothesis stating that White photos did not affect transcription accuracy is also rejected. The following table in percentages and raw numbers (N) show how the data was distributed for all participants:

Photo (total)	Correct (N)
Black (744)	54.8% (408)
White (744)	62.6% (466)
Silhouette (372)	68.8% (256)

Table 3: Male & female transcription accuracy with varied photo types

For female participants, the analysis on Black photos yielded a chi-squared of 32.6418729 (df=1) and a p-value < 0.001, while with White photos, the chi-squared result was 1.20716986 (df=1), in which $0.1 < p < 0.5$. The latter finding is not significant; so in other words, females viewing White photos behaved similarly to that of the control (silhouette), yet had significantly different transcription accuracy when presented with the Black photos regardless of the linguistic input. Table 4 shows how female participants performed on the transcription task concerning photos:

Photo (total)	Correct (N)
Black (396)	57.1% (226)
White (396)	67.7% (268)
Silhouette (198)	70.2% (139)

Table 4: Female transcription accuracy with varied photo types

However, for males, both the Black and White photos were significant. Black photos for males came back with a chi-squared value of 35.2748538 (df=1), while White photos resulted in a chi-squared sum of 16.9068826 (df=1); both p-values < 0.001- therefore in these two instances, the null hypothesis of male participants not being affected by Black and White photos is rejected. The following depicts male accuracy based performance in reference to photos:

Photo (total)	Correct (N)
Black (348)	52.3% (182)
White (348)	56.9% (198)
Silhouette (198)	67.2% (117)

Table 5: Male transcription accuracy with varied photo types

Overall, while the White and Black photos were significant for the males, only the black photos were significant for women. This raises the question as to why both types of photos affected males, but only one for women. No explanation was readily available as to why this has occurred, so this is an area that should be addressed with future research. Seeing as there were only 29 male participants in comparison to 33 females, more information may be able to shed light on this perplexing finding.

A questionnaire was also given to participants at the end of the experiment in order to gain a better insight into the collected data. To see whether people were truly using the photos to aid them in transcription, the questionnaire asked whether the photos were helpful in the transcription exercise; in this case, 82.3% of the participants answered "no," 16.1% said "yes," and 1.6% were unsure. The table below shows both female and male participants' responses combined:

Did you use the photos to help transcribe?	Response (N)
Yes	16.1% (10)
No	82.3% (51)
Unsure	1.6% (1)

Table 6: Questionnaire response regarding photo usage

When separated, it is seen that females and males answered similarly:

Did you use the photos to help transcribe?	Female	Male
Yes	9.7% (6)	6.5% (4)
No	41.9% (26)	40.3% (25)
Unsure	1.6% (1)	0.0% (0)

Table 7: Questionnaire response regarding photo usage separated by sex

Nonetheless, it should be noted that despite the assertion that many participants stated they did not use the photos, participants were consistent and many of them wrote similar statements claiming that they chose not to look at photos because they were too distracting, which can explain as to why photos proved to be a significant factor in the statistical analysis. Participants wrote that they purposely did not look at the photos because they did not want to think of stereotypes, that the photos were purposely trying to "prejudice" them, or that they purposely did not look at the pictures because it would cause them to think of "other things." Interestingly enough, of the 82.3% (51) of the participants who stated that they did not use the photos, 39.2% (20) stated that they did not use the photos because it was easier to focus and transcribe the audio when they were not paying attention to the photos; I personally noted that many participants closed their eyes when taking my experiment or purposely did not look at the computer screen, presumably to hear more clearly-as expressed multiple times within the questionnaire answers.

Looking at the questionnaire, if one were to combine the participants who said that they did not look at photos because they needed to focus and the photos would be distracting plus the answers where people said they did not look at the photos because the mouths did not move, then this accounts for 45% (23) out of the 82.3%. Another relevant component to consider is that while the majority wrote that they did not look at the photos, six wrote the blank clips (silhouette) were easiest, three mentioned the white woman was easier, and eight mentioned the black woman was harder to understand. So despite participant's claims to not use photos, it shows that in reality, they really were using them.

9.12 Research Question 1b: Does perceived ethnicity affect transcription accuracy?

As seen from the first part of this question, participants did use social indexical information, so now it is asked as to whether the concept of ethnicity played a role in transcription accuracy. Question 1a demonstrated that both male and female participants were more accurate in transcribing the silhouette and White photos, than the Black photos. However to explore this in greater detail, responses from the questionnaire will be discussed. Since there were open-ended questions in the questionnaire, it is not possible to categorically organize all the data from the questionnaire and present exact percentages given that this form of data is qualitative in nature. Instead, answers were grouped together according to similarities and counted by the frequency in which they were mentioned. It was noted by respondents six times that it was easiest to transcribe when presented with the silhouette. It was also written in six instances that the Black photos and voices were difficult to understand, however, participants could have meant three different things; it could have been the case that the black photos were hard to understand, that they audio of AAVE was difficult to understand, or response were ambiguous and could have meant the photo, the audio, or both. One person clearly stated that the photo with the "black woman" was hard to understand, one stated that that those who used AAVE were difficult to understand, while four participants were ambiguous in what they wrote and it could have been interpreted that either AAVE audio, the picture, or both were difficult to understand.

Three participants additionally wrote that the white photo was easier to understand, and two stated that the Black photo was easier to understand. It was mentioned eleven times that people could not figure out a pattern within the photos and that they did not seem to be relevant because the photos did not fit with their expectations, while three participants noted that because the mouths did not move, they did not bother to look at the photos. The latter statement further supports the idea people were assuming that the visual stimuli would match with audio.

On the questionnaire, only two people wrote that the African American woman was easier to understand: one was a male in the 40-60 year old age group and the other was a female in the 19-24 age group. The male wrote that he spoke AAVE, so this could be the reason as to why he perceived the African American woman as easier to understand-research on paralinguistic cues have shown that listeners prefer

speech similar to their own; Campbell-Kibler (2010) applies this to what is known as the Accommodation theory -which "...proposes that interactants who feel positively toward one another tend to align their speech patterns with one another on a variety of dimensions" (Campbell-Kibler 2010: 379). Results from the following table show the average of how people transcribed AAVE, and as shown, both sexes and age groups performed similarly:

Sex (age)	Total sentences of AAVE	Correct
Females (19-24)	225	59.1% (133)
Females (40-60)	270	60.4% (163)
Males (19-24)	240	51.2% (123)
Males (40-60)	195	54.9% (107)

Table 8: Transcription accuracy of AAVE

Sex (age)	Total sentences of MUSE	Correct
Females (19-24)	225	67.6% (152)
Females (40-60)	270	68.5% (185)
Males (19-24)	240	61.7% (148)
Males (40-60)	195	61.0% (119)

Table 9: Transcription accuracy of MUSE

Individual results from the male participant (in the 40-60 age group) who wrote that he spoke AAVE and that he thought the "colored lady" was easier to understand showed that he did about the same as the average of males in his age group; transcribing AAVE sentences at 53.3% (the average for male participants ages 40-60 in AAVE transcription was 54.9%). Unlike the male participant, the female participant (in the 19-24 age group) that wrote that the African American woman was easier to understand did perform better than the average with 73.3%; this is quite high when compared to the average of her age group, which had 59.6% average of correctness. It is difficult to explain as to why she did so much better than the average, considering that she does not speak AAVE, have contact with AAVE speakers, and is unfamiliar with AAVE and hip-hop/rap culture. So in this case, it is possible that because of her preconceived notions that the Black photo was easier to understand, this may have ultimately helped her in transcribing AAVE.

Moving on from photos and questionnaire responses, just dialect was looked at and AAVE was compared to MUSE. The null hypothesis was rejected as the chi-squared result was 13.718 (df=1) and the p-value < 0.001. Fifteen sentences of both dialects (AAVE and MUSE) were given to each participant, giving way to 930 sentences for AAVE and 930 sentences of MUSE total.

Females and Males combined:

Dialect (total sentences)	Correct (N)
AAVE (930)	56.6% (526)
MUSE (930)	64.9% (604)

Table 10: Transcription accuracy of AAVE vs. MUSE

Individually for females, the chi-squared result came back with a value of 7.364 (df=1) and a p-value was 0.007. Table 10 shows the distribution of data for females:

Dialect (total sentences)	Correct (N)
AAVE (495)	59.8% (296)
MUSE (495)	68.1% (337)

Table 11: Female transcription accuracy of AAVE vs. MUSE

Similarly, the chi-squared for males was 6.425(df=1) with a p-value of 0.011. Table 11 shows how male participants performed in transcription accuracy according to dialect.

Dialect (total sentences)	Correct (N)
AAVE (435)	52.9% (230)
MUSE (435)	61.4% (267)

Table 12: Male transcription accuracy of AAVE vs. MUSE

Overall, participants were better at transcribing MUSE than AAVE, so dialect was a significant factor in the transcription process in which MUSE was more accurately transcribed at 64.9% than AAVE at 56.6% Furthermore, when separated, it is interesting to note that both males and females differed at about 9% (8.3% for females and 8.5% for males) between their accuracy for MUSE and AAVE. A chi-squared test was then performed to see if participants did statistically significant in regards to correct and incorrect answers for MUSE and AAVE. The chi-square test come back with a value of 8.33424658 (df=1) and a $p< 0.01$, which proves there was a statistical difference between how people performed in AAVE vs. MUSE. This demonstrates that participants did indeed do significantly worse in transcribing AAVE than MUSE, which is something that was predicted.

Taking the questionnaire data and the statistical information gathered from the first part of the research question showing that Black photos were significant for both females and males and both Black and White photos were significant for males. Research question 1a. found that participants were more accurate in transcription when presented with MUSE than AAVE. Taken together, it is safe to say that perceived ethnicity from the visual stimuli and dialect was indeed a factor in this experiment.

9.2 Research question 2: Do people transcribe more accurately when the dialect matches the given photo?

The next question involves the analysis of both visual and audio stimuli together, rather than looking at them separately. The question at hand asks whether participants were able to transcribe more accurately when the dialect matches the given visual stimuli. The chi-squared test proves that this was significant with a value of 85.7406385 ($df = 1$), $p < 0.001$ for matching photos; those that were mismatched resulted in a chi-squared value of 6.82091864 (df=1) and $p < 0.01$. The table below depicts the overall results for both male and female participants, and shows that overall, participants transcribed more accurately when the visual and audio stimuli did not match.

Photos & Dialect (total)	Correct
Silhouette (372)	68.8% (256)
Match (744)	53.1% (395)
Mismatch (744)	64.4% (479)

Table 13: Male & female matched guise results regarding transcription accuracy

The data is further separated in order to see how females performed, the chi-squared results for females was 32.6418729 (df=1) with $p < 0.001$. Therefore, the null hypothesis stating that photos and dialect were insignificant is rejected. Below shows the exact percentages and raw numbers for females:

Photos & Dialect (total)	Correct (N)
Silhouette (198)	70.2% (139)
Match (396)	57.1% (226)
Mismatch (396)	67.7% (268)

Table 14: Female matched guise results regarding transcription accuracy

Male responses are looked at next and the following table shows how their data was distributed. The chi-squared result for photos and dialect that matched was 55.1169591 (df=1)-significant and resulting in a $p < 0.001$. The same went for mismatched photos, but was a little less significant with a chi-squared of 6.90103464 (df=1) and $p < 0.001$.

Photos & Dialect (total)	Correct (N)
Silhouette (174)	67.2% (117)
Match (348)	48.6% (169)
Mismatch (348)	60.6% (211)

Table 15: Male matched guise results regarding transcription accuracy

The prediction that participants would transcribe more accurately when the photo matched the dialect is proven to be incorrect, as there were a higher percentage of participants who correctly transcribed when the photos and dialect did not match. This is a quite puzzling result and there is no explanation as to why this occurred. Regardless as to whether or not they used the photos (41.9% of females and 40.3% of males said "no"), when the data is separated into males and females, it appears that females did better when the photos and dialect matched. Many behavioral studies have shown that facial recognition is dependent upon the gender of the observed face and that of the participant (Ino et al. 2010), and it is repeatedly found that females have a higher level of recognition for female faces than males (Cross et al. 1971; Lewin and Herlitz 2002; Shaw and Skolnick 1994, 1999; Wright and Sladden 2003). Therefore, because females are more adept to facial recognition and because the photos that were used in this experiment were all females, it is plausible that they were reacting to this gender bias.

However, to gain a further insight on data, it is then divided into MUSE and White photos versus AAVE and Black photos. Table 16 how both male and female participants performed when the dialect and photo matched:

Photos & Dialect (total)	Correct (N)
Black & AAVE (372)	41.4% (154)
White & MUSE (372)	64.8% (241)

Table 16: Male & female transcription accuracy of Black & AAVE vs. White & MUSE

The chi-squared result was 40.850 (df=1) and a $p < 0.001$ for all participants. So in re-evaluating the data, it is seen that the null hypothesis is rejected and that the matching stimuli did have an effect on participants. As seen from above, both male and female participants did significantly worse in transcription accuracy when presented with a Black photo and AAVE, than with a White photo and MUSE. The above data of matching photos and dialect is further broken down into female and male participants. As seen below, females were more accurate in transcribing White photos with MUSE, than Black photos and AAVE. This is proven significant with a chi-squared of 27.870 (df=1) and $p < 0.001$.

Photos & Dialect (total)	Correct (N)
Black & AAVE (198)	43.9% (87)
White & MUSE (198)	70.2% (139)

Table 17: Female transcription accuracy of Black & AAVE vs. White & MUSE

Similar to female participants, male participants transcribed more accurately when presented White photos with MUSE than Black photos and AAVE. This too was significant with a chi-squared value of 14.092 (df=1) and $p < 0.001$.

Photos & Dialect (total)	Correct (N)
Black & AAVE (174)	38.5% (67)
White & MUSE (174)	58.6% (102)

Table 18: Male & transcription accuracy of Black & AAVE vs. White & MUSE

So it is observed that when the data is redefined, matching audio and visual stimuli proves significant and both male and female participants transcribed more accurately when presented with a White photo and MUSE.

9.3 Research question 3: Are participants ages 40-60 able to transcribe more accurately than those who are 19-24 years old?

It was further hypothesized that age may have been a factor in this experiment. Initially, from the table below, it appears that age is slightly significant:

Age	Correct (N)
19-24 (930)	59.8% (556)
40-60 (930)	61.7% (574)

Table 19: Transcription accuracy of 19-24 & 40-60 year olds

However, the chi-squared came back with a value of 0.731(df=1) and a p-value of 0.393. Since this number is above the alpha value (0.05), the null hypothesis is not rejected and age proves to not be a factor in this instance. When the data was broken down even further, it is still seen as being insignificant for both females and males. The chi-square value for females was 0.131 (df=1) and a p-value of 0.717, while males had a chi-squared value of 0.195 (df=1) and p-value of 0.659. Since both of these are quite above the alpha of 0.05, the null hypothesis must be accepted. Although it was initially predicted that those that were ages 40-60 would be less accurate in their transcriptions than those that were ages 19-24, results from the experiment show that this prediction was incorrect, and there could be a couple different reasons as to why age did not factor in the statistical analysis.

Out of the 75 participants, thirteen were excluded, and six of these were discarded because they had a hearing disability. All six participants with hearing disabilities were in the 40-60 age group. While it would have been possible to include these people in the analysis to see if they would result in a different outcome, all of those that I tested had varying levels of hearing disabilities that were quite vast; some were completely deaf in one ear, others had trouble hearing different pitches and one had a hearing aid. Because I was purposely targeting participants who did not have a hearing disability, a statistical analysis would be inappropriate as this is not representative of a population. Another reason that could explain why age was not a factor is possibly due to the manner in which I divided my age boundaries.

Rather than have 40-60 year olds, I could divide the ages 40-49 and 50-60. After recoding the data and redefining the age boundaries into three groups (19-24, 40-49, 50-60), it turns out that age is significant as the chi-squared value of 8.710 (df=2) resulted in a p value of 0.013, which is less than the alpha value of 0.05. Therefore, in this case, the null hypothesis must be rejected.

Age	Correct (N)
19-24 (930)	59.8% (556)
40-49 (452)	66.4% (300)
50-60 (478)	57.3% (274)

Table 20: Redefined age boundaries & transcription accuracy

Interestingly enough, for some reason, the 40-49 year olds were more accurate in transcribing than the 19-24 year olds. It is possible that since Anderson Gosselin & Gagné (2011) found that adults expend more listening effort than those who are younger, this could explain as to why they performed better. However, just because adults put forth more listening effort, this does not mean that they would be able to transcribe more accurately nor does this explain as to why the 50-60 year olds did worse, but perhaps after a certain age and combined age-related hearing loss (Moller 2006), listening effort proves to not be as helpful when it comes to transcription accuracy. Life experience is also an important factor to consider, as Tan (2012) found, adults were better at recognizing ethnic variations than those who were younger and this was largely attributed to exposure and experience. It is therefore possible that those ages 40-49 had more experience than the younger age group and those that were 50-60 years old.

Furthermore, Clarke (2003) found through a series of experiments that listeners altered their expectations for phone categories from information about the speakers given from prior exposure to the speaker and that there was a perceptual benefit of being exposed to the non-native speech before taking an experiment where they were exposed to a foreign-accented speech. It was also found that the control groups were able to adapt to the accented voice at the end of the experiment (Clarke 2003: 11), which thus shows that experience (e.g. age) and exposure (e.g. the professional worksphere, media) could play a role in results and that perhaps those that were 40-49 had more experience and exposure than the 19-24 year olds and the 50-60 year old groups. Overall, my prediction that those in the older age group were more accurate in transcription than younger is partly correct, but only when the age boundaries are redefined into three categories.

9.4 Research question 4: Are female participants more accurate in transcription than their male counterparts?

Lastly, the question as to whether one sex was more accurate in transcription than the other. The statistical analysis denotes that sex indeed played a significant role in transcription accuracy; the chi-squared came back with a value of 9.014 (df=1) and a p-value of 0.003. Below, table 21 and 22 show that females and males of both age groups performed similarly.

Sex (age)	Correct (N)
Female (19-24)	63.3% (285)
Female (40-60)	64.4% (348)

Table 21: Female transcription accuracy divided by age

Sex (age)	Correct (N)
Male (19-24)	56.5% (271)
Male (40-60)	57.9% (226)

Table 22: Male transcription accuracy divided by age

The overall results for both males and females were as followed:

Sex (total sentences)	Correct (N)
Female (990)	63.9% (633)
Male (870)	57.1% (497)

Table 23: Male & female transcription accuracy

From the above results, it is seen that on average, females are slightly better in transcription accuracy than males. This is what I initially predicted due to the fact that much sociolinguistic literature asserts that women are at the forefront in regards to language variation and change, and because of their higher use of innovative vernacular forms occurring below the level of consciousness than males (Labov 1990), therefore they may also be more aware to dialect variations. However, similar to what was hypothesized in the second research question, this could also possibly due to another case of gender bias. Perhaps females are more accurate because the visual and audio stimuli that I used were female. Unfortunately, as previously noted, I was not able to locate previous literature regarding transcription accuracy in regards to sex; this would be an interesting area for future research.

9.5 Are participants responding to idiosyncratic results for speakers used within the experiment?

Additionally, another question presented itself while analyzing the data. This question relates to whether participants were responding to idiosyncratic results for speakers that were used within this experiment. In the case of females and males, they were all able to transcribe the speech of AAVE Speakers 1 and 2, however, both groups proved to have more difficulty with AAVE Speaker 3. These are marked in the tables below with an asterisk. When looking at the MUSE speakers, similar to AAVE results, participants appeared to have trouble with the third speaker. Concerning the results from the speakers, the chi-square value was 125.106 (df=5) and the p-value < 0.001 for both males and females, meaning the null hypothesis assuming participants were unaffected by speakers is false.

Females and males combined:

Dialect and Speaker (total sentences)	Correct (N)
AAVE Speaker 1 (310)	69.4% (215)
AAVE Speaker 2 (310)	56.1% (174)
*AAVE Speaker 3 (310)	44.2% (137)

Table 24: AAVE speakers & participants' transcription accuracy

Female response to AAVE speakers:

Dialect and Speaker (total sentences)	Correct (N)
AAVE Speaker 1 (165)	72.1% (119)
AAVE Speaker 2 (165)	60.0 (99)
*AAVE Speaker 3 (165)	47.3% (78)

Table 25: AAVE speakers & female participants' transcription accuracy

Male response to AAVE speakers:

Dialect and Speaker (total sentences)	Correct (N)
AAVE Speaker 1 (145)	66.2% (96)
AAVE Speaker 2 (145)	51.7% (75)
*AAVE Speaker 3 (145)	40.7% (59)

Table 26: AAVE speakers & male participants' transcription accuracy

The following tables concern the MUSE dialect. Again, the asterisk shows where participants generally had more trouble in transcribing accurately.

Females and males combined:

Dialect and Speakers (total sentences)	Correct (N)
MUSE Speaker 1 (310)	67.4% (209)
MUSE Speaker 2 (310)	80.0% (248)
*MUSE Speaker 3 (310)	47.4% (147)

Table 27: MUSE speakers & participants' transcription accuracy

Female response to MUSE speakers:

Dialect and Speaker (total sentences)	Correct (N)
MUSE Speaker 1 (165)	69.1% (114)
MUSE Speaker 2 (165)	83.6 %(138)
*MUSE Speaker 3 (165)	51.5% (85)

Table 28: MUSE speakers & female participants' transcription accuracy

Male response to MUSE speakers:

Dialect and Speaker (total sentences)	Correct (N)
MUSE Speaker 1 (145)	65.5% (95)
MUSE Speaker 2 (145)	75.9% (110)
*MUSE Speaker 3 (145)	42.8% (62)

Table 29: MUSE speakers & male participants' transcription accuracy

Because participants had trouble with the third speakers of both AAVE and MUSE, I decided to re-examine my data and there is an explanation as to why this may be. I noted that AAVE Speaker 3 was recorded in a different environment than all the other speakers (both AAVE and MUSE). I had recorded this speaker in her apartment at her request, while the others were recorded in the same dormitory at Michigan State University. Therefore, it is possible that the environment in which the recordings took place affected the overall quality of the audio.

It is also possible that AAVE Speaker 3 had a "thicker" AAVE accent. As noted from the brief overview, not all AAVE speakers use the same linguistic features equally, as regionality, and individual styles/registers can vary (as found with other dialects); rather it is best to think of AAVE lying on something of a continuum. Keeping this in mind, AAVE Speaker 3 was the AAVE speaker that asked if she could use copula deletion in three of her sentences, therefore it is entirely possible that AAVE Speaker 3 was more progressive in AAVE than the other AAVE Speakers used in this experiment.

It is unknown why participants had more trouble with MUSE Speaker 3. It was thought that speech rate may have been a factor. To prove this quantitatively, I initially tried to measure speech rate via a Praat script [12]. This script was supposed to label syllables and mark parts of the sound files when there were pauses or silence, however, when I inspected some of the sentences manually, I found that many times syllables were labeled incorrectly (e.g. "Mom" was counted as three syllables rather than one). After discovering this, it was decided that sound files were to be measured manually in syllables per second. For all sound files, I used Praat to label syllable boundaries and then divided by the total time of the sound file. Figure 7 is an example of one of my sound files:

[12] This Praat script is written about extensively by De Jong and Wepe 2009, which also includes where the script can be found.

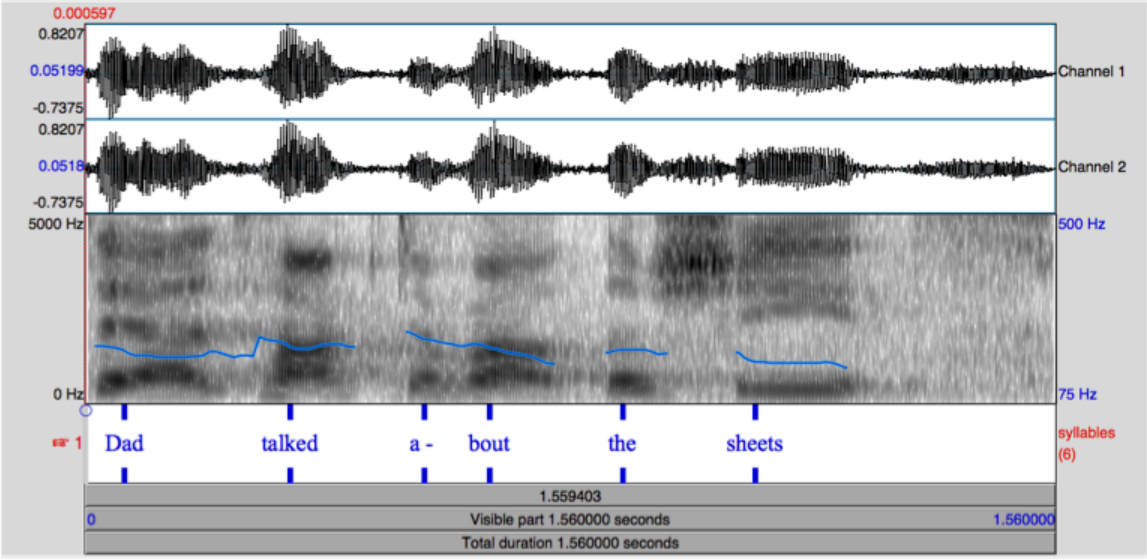

Figure 7: Example of audio sound file on Praat

Syllables were labeled as demonstrated above, but the syllable count is also listed on the bottom right hand corner in red. "Total duration" (found at the bottom center of the picture; e.g. 1.560000 seconds) was also taken into account. To calculate syllables per second, the syllables were taken from each sound file the speakers produced and added up; then total duration of the sound files were also added up. Then the total number of syllables were divided by the sum of the total durations. This was done for all five sentences each speaker produced and were rounded two decimal places. Tables 30 and 31 show the speech rate for AAVE and MUSE speakers.

Speaker	Syllables per second
AAVE Speaker 1	4.85
AAVE Speaker 2	5.25
AAVE Speaker 3	4.28
Total Average	4.79

Table 30: AAVE speech rate (syllables per second)

Speaker	Syllables per second
MUSE Speaker 1	3.41
MUSE Speaker 2	3.67
MUSE Speaker 3	4.17
Total Average	3.75

Table 31: MUSE speech rate (syllables per second)

As seen from the above tables, all AAVE Speakers spoke at a faster speech rate than the MUSE speakers and this may possibly explain as to why participants transcribed AAVE less accurately than MUSE, yet this does not explain the reason participants did worse on Black photos than White photos (as I had mixed MUSE and AAVE with different photos). For future reference, it is recommended that perhaps speaker speech rate should be more controlled. However, I do not believe these factors would have adversely affected the data, as speaker variability is found in everyday communication and varies from person to person.

The questionnaire also addressed transcription difficulty. Here, 20 participants wrote within the questionnaire that speaker speech rate was a factor. Interestingly enough, one female participant believed the "fast talkers" to be using AAVE, while another wrote that it was difficult to understand people who use AAVE when they talked "too quickly." From the calculations done on speech rate, they were indeed correct, as all AAVE speakers spoke faster than MUSE speakers. However, rather than controlling speaker speech rate, it was decided that the variability found in the speaker's speech rate would be left alone. Despite research on speech rate showing that people prefer mid-range speeds and over fast and slow talkers (Street and Brady 1982), other research stemming from Brady et al. (1983) indicate that individual listeners react more positively to speakers with perceived speech rates similar to their own, so since participants' speech rate varied, the variation found in the speakers used in the experiment were kept as is.

This still does not explain as to why participants had trouble with MUSE Speaker 3. Of all the informants that I used in this experiment, it was MUSE Speaker 3 that was most reluctant to take part. As noted in the methodology, all MUSE speakers I used in this experiment where complete strangers to me because I thought this was the best manner in which to elicit "standard" pronunciations. Perhaps emotional tone could have affected results; Mullenix et al. (2002) found through a series of experiments that variation of emotional tone of voice had a "detrimental" effect on the perceptual processing of speech. However, in their experiments, they compared "angry voice," "surprised voice," and "commanding voice" (both males and females) using a singular word utterance (Tom and Todd). While this could have been useful, they did not measure a neutral voiced utterance to see how people reacted when there was no emotion to be found. Nonetheless, neutral voice was explored in the experiments conducted by Nygaard and Queen (2008); they suggest that the emotional tone of voice influences the time involved in lexical processing. In other words, it is proposed that information on emotional tone in regards to linguistic processing can affect semantic recognition and the naming of spoken words in an "emotion-congruent manner." Here, they too used single word utterances said by both males and females. However, results showed that participants had a lower latency when presented with stimuli with emotional tone (in this case, happy or sad). So while neither of the discussed articles on emotional tone support my idea that emotional tone may have adversely affected participants' transcription accuracy, neither used the emotional tone of reluctance or impatient, so perhaps future research on this topic can be expanded to see if this type of tone can be detected by participants and if latency is affected.

10. Brief Discussion of Methodological Issues

A large issue in this experiment was speaker speech rate. While every attempt was made to control speed, it was apparent after recording that speech rate varied amongst the speakers. However, I decided to leave the sentences as they were because speaker variation is to be expected in everyday speech, and since my participants would speak at different speech rates themselves, it made sense to keep them as they were. If to be replicated again, it would be more ideal to include sentences taken from spontaneous speech, as it is shown that listeners are able to discern between read passages and those taken from spontaneous speech and this in turn can affect outcomes of the data (Campbell-Kibler 2010).

Another issue presented in this experiment concerns White noise, which was talked about extensively in the methodology section of this dissertation. All the sentences except for one had the same amount of noise added to them. Ideally, it would be best to have all sentences have the same amount of white noise added. The problematic sentence was said by AAVE Speaker 3, but the overall average that participants got this sentence correct was 61.3%, which was close to the other overall averages. Therefore, this sentence was kept in the statistical data analysis; if it were the case that almost all participants missed this sentence, then this would have been dealt with differently and would have possibly resulted in this sentence being excluded. Another concern involving white noise is that there is a chance that the ethnolects were masked by the noise. Although the added white noise was necessary in order to avoid a ceiling effect, the possibility that the white noise masks ethnolects is something to consider in future studies.

The controlling of visual stimuli was also done, in which photos that were chosen were of a neutral face, the same lighting, and revealed no clothing, background, and hairstyle. Nevertheless, from the results, it was shown that perhaps females did slightly better than males when it came to matching the dialect and photos. Research has shown that there is the possibility of gender bias. My control photo of the silhouette I used was genderless, so this may contribute to how participants used it. It is suggested that when doing a similar experiment, depending on what sex the visual stimuli is, if using a silhouetted photo; it would perhaps be best to use a silhouette that may suggest the same sex as the visual stimuli. For instance, since all the photos and audio recordings were female, perhaps this silhouette would have been better suited.

Figure 8: Female Silhouette [13]

[13] Image found here: http://www.clker.com/cliparts/z/r/A/H/X/g/silhouette-female-grey-md.png

As previously discussed, my data was unbalanced in some areas, especially in regards to experience, contact, and familiarity with AAVE. Since these are all factors that could affect the data, it would be recommended to target an even amount of participants that were familiar and unfamiliar with AAVE for future reference, in order to see if experience and familiarity/exposure played a role in transcription accuracy.

Lastly, it is important to address the manner in which I arranged the photos and audio files. As explained in the methodology section, stimuli were quasi-randomly placed so that the same speaker and the same photo did not follow one another. It could be possible that the hardest files to transcribe were placed with a black photo and the easiest with those of the white. While it is felt that this is unlikely possibility, it is still an issue that could affect the results, so future research should consider this concern.

11. Conclusion

In summary, it is my hope that this dissertation answered some questions in regards to AAVE speech perception and attitudes. From the data obtained from both the experiment and questionnaire, it is seen that all research questions addressed in this dissertation were of importance.

Question 1a asked if the use of social indexical cues affect participants' transcription accuracy. Both male and female participants transcribed more accurately when the photo was of the silhouette and devoid of any cues that would suggest race or gender. When Black and White photos were compared against the silhouette, it was discovered that males and females were affected by the Black photos, yet only the White photos affected male participants.

When examining the questionnaire as it relates to this question, 82.3% of participants stated that they did not use photos, however, out of the 82.3% that said they did not use photos, 39.2% said they did not look at the photos because it was easier for them to transcribe the audio stimulus when they did not look at the photos. Furthermore, when combining the participants who said they did not look at photos because they needed to focus and because the mouths of the speakers did not move, this accounts for 45% of the 82.3% who said they did not look at photos. Overall, both the experiment results and questionnaire results show that participants were using social indexical cues in transcribing and that participants transcribed more accurately with the silhouette, followed by the White photos and then Black photos.

Question 1b. builds on the idea from research question 1a. (where participants were affected by the photos; male participants were affected by Black and White photos, while females were solely affected by Black photos). Overall, participants were more accurate in transcribing the silhouette and White photos than Black photos. When just looking at AAVE and MUSE, it was found that participants transcribed more

accurately with MUSE than AAVE. So taking these findings together, it is believed that perceived ethnicity, both visually and linguistically, proved to affect transcription accuracy.

The second research question addresses both the audio and visual stimuli collectively. Interestingly enough, results found that male and female participants transcribed more accurately when the dialect and photo did not match. However, when separated more, it is found that male and female participants were significantly affected by photos and dialect. Overall, females did better when the photos and dialect matched; this result could be possibly due to gender bias, as female speakers and photos were used in this experiment.

To gain more insight on the experiment results, data was divided into MUSE and White photos and AAVE and Black photos, leaving out the silhouette (control) photo. When compared against each other, it was found that male and female participants were significantly affected by the matching dialects and photos; both sexes did significantly better in transcription accuracy with White photos and MUSE, than Black photos and AAVE.

In the third research question, it was hypothesized that participants who were 40-60 would transcribe more accurately than their younger counterparts who were 19-24. However, results indicated that age was not significant. It is not until the age boundaries are redefined into three categories (19-24, 40-49, 50-60) that age proves to be significant. Interestingly enough, those who were 40-49 transcribed more accurately than those 19-24 and 50-60. It is then discussed that this finding may be due to a combination of prior experience, exposure, and age-related hearing loss, although it is suggested more research should be conducted to see as to whether this holds true.

It is then asked as to whether female participants transcribed more accurately than male participants. As discussed previously, findings from Loudermilk (submitted) found that females were more perceptive to non-canonical forms. Females are also largely accredited as being linguistic innovators, therefore it is predicted females have a heightened linguistic sense. Taking these into account, it was predicted that female participants would transcribe more accurately and this is found to be true.

Finally, an additional question presented itself, regarding whether participants were reacting to idiosyncratic results from speakers used within this experiment, and this was found to be accurate. The idiosyncratic results did have an effect on participants. It was found that participants had trouble with MUSE and AAVE Speaker 3. The issue of speech rate arose; although this does not explain why participants performed worse when it came to MUSE and AAVE Speaker 3, it does present the possibility as to why participants transcribed AAVE less accurately than MUSE, as it was shown that all three AAVE speakers had a faster speech rate than MUSE speakers.

Although age was significant when the age boundaries were redefined, the finding that those who were 40-49 did better than the younger (19-24) and older (50-60) age groups is quite interesting. A replication of this experiment without the criterion that participants do not have a hearing disability could possibly help shed some light on this finding. This could be beneficial not only in the linguistic sense, but for future research concerning auditory processes and how age affects hearing abilities throughout a person's life. It was also found that there is a lack of research in regards to transcription accuracy, gender and speech perception, and studies on speech perception throughout the lifespan. Hopefully, more research within sociolinguistics and other related fields will expand on this shortage of information.

Although more research is necessary, it is hoped that this dissertation has provided a somewhat clearer picture regarding language attitudes and speech perception experiments as they relate to AAVE. This experiment was unable to address how previous experiences, exposure, or prior knowledge affect speaker perception, however this too can be fixed through future studies. Taking information from the questionnaire and experiment into consideration, it is evident that a large portion of participants perceived the Black photo/speaker to be more difficult to understand than the White photo/speaker and this is shown in transcription accuracy results.

References

Adger, Carolyn Temple, Walt Wolfram, and Donna Christian. *Dialects in Schools and Communities*. Mawah, New Jersey: Lawrence Erlbaum Associates, Inc., 2007. Print.

Anderson, Bridget L. "Dialect leveling and /ai/ monophthongization among African American Detroiters." *Journal of Sociolinguistics* 6/1 (2002): 86-98. Print.

Anderson, Penny Gosselin and Jean-Pierre Gagné. "Older Adults Expend More Listening Effort than Young Adults Recognizing Speech in Noise." *Journal of Speech, Language, and Hearing Research* Vol. 54 (2011): 944-958. Print.

Arana, Gabriel. "Creaky Voice: Yet Another Example of Young Women's Linguistic Ingenuity." *The Atlantic.* 10 Jan. 2013. Web. Accessed on 24 Aug. 2013.

Arksey, Hilary and Peter Knight. *Interviewing for Social Scientists*. London: Sage, 1999. Print.

Bradlow, Ann R. and Jennifer A. Alexander. "Semantic and phonetic enhancements for speech-in- noise recognition by native and non-native listeners." *Acoustical Society of America.* Vol. 121, No. 4 (2007): 2339-2349. Print.

Bailey, Guy and Natalie Manor. "Decreolization?" *Language in Society* Vol. 16 No. 4 (1987): 449-473. Print.

Bailey, Guy and Natalie Maynor. "The Divergence Controversy." American Speech Vol. 64 No. 1 (1989): 12-39. Print.

Bailey, Guy. "The relationship between African American and White Vernaculars in the American South." *Sociocultural and Historical Contexts of African American English.* Ed. Sonja L. Lanehart. Philadelphia, Pennsylvania: John Benjamins Publishing Company, 2001. 53-91. Print.

Baugh, John. "Linguistics, Education, and the Law: Educational Reform for African-American Language Minority Students." *African American English Structure, history and use.* Eds. Mufwene, Salinkoko S. et al. New York, New York: Routledge, 1998. 282-301. Print.

Baugh, John. *Beyond Ebonics: Linguistic Pride and Racial Prejudice.* New York, Oxford University Press, 2000. Print.

Bohn, Mariko T. and Yoshio Matsumoto. "Young women in the Meiji period as linguistic trendsetters." *Gender and Language* Vol. 2 No. 1 (2008). Print.

Bourdieu, Pierre. *Language and Symbolic Power.* Malden, Massachusetts: Polity Press, 1991. Print.

Boyd, Gwendolyn Aline. "Teacher attitudes toward African-American Vernacular English: The Relationship to Students' Perceptions of Classroom Climate." Diss. University of Memphis, 1996.

Calloway, Sandra Louise. "Contemporary Secondary Teachers' Perceptions of African American Students and Their Language." Diss. University of San Francisco, 1989.

Campbell-Kibler, Kathyrn. "Listener Perceptions of Sociolinguistic Variables: The Case of (ing)." Diss. Stanford Univeristy, 2005.

Campbell-Kibler, Kathryn. "Accent, (ing), and the social logic of listener perceptions." *American Speech* Vol. 82, No. 1 (2007): 32-64. Print.

Campbell-Kibler, Kathryn. "Sociolinguistics and Perception." *Language and Linguistics Compass.* Vol. 4 No.6 (2010): 377-389. Print.

Campbell-Kibler, Kathryn. "The sociolinguistic variant as a carrier of social meaning." *Language Variation and Change* Vol. 22 (2011): 423-441. Print.

Campbell, Murphysteen. "The Relationship Between Teacher Attitudes Towards African American Vernacular English and Teacher-Student Classroom Interaction." Diss. University of Memphis, 1996.

Cargile, Aaron C. et al. "Language Attitudes as a Social Process: A Conceptual Model and New Directions." *Language and Communication* Vol. 14, No. 3 (1994): 211-236. Print.

Cherry, Kendra. "What Is Priming." *About.com* Web. Accessed on 9 Sept. 2013.

Clarke, Margaret. "Processing time effects of short-term exposure to foreign-accented English." Diss. University of Arizona, 2003.

Clopper, Cynthia G. and David B. Pisoni. "Perception of Dialect Variation." *The Handbook of Speech Perception.* Eds. Pisoni, David B. and Robert E. Remez. Oxford, United Kingdom: Blackwell Publishing Ltd, 2005. 313-337. Print.

Cross, John F., Cross, Jane, and James Daly. "Sex, race, age and beauty as factors in recognition of faces." *Perception & Psychophysics* Vol. 10 (1971): 393–396. Print.

Cukor-Avila, Patricia. *The evolution of AAVE in a rural Texas community: An ethnolinguistic study.* Diss. University of Michigan, 1995.

Cukor-Avila, Patricia. "Co-existing grammars: The relationship between the evolution of African American and White Vernacular English in the South." *Sociocultural and Historical Contexts of African American English.* Ed. Sonja L. Lanehart. Philadelphia, Pennsylvania: John Benjamins Publishing Company, 2001. 93-127. Print.

Cukor-Avila, Patricia. "The complex grammatical history of African-American and white vernaculars in the South." *English in the Southern United States*. Ed. Nagle, Steven J. and Sara L. Sanders. Cambridge, United Kingdom: Cambridge University Press, 2003. 82-105. Print.

De Jong, N.H. and Wempe, T. "Praat script to detect syllable nuclei and measure speech rate automatically." *Behavior research methods*, Vol. 41 No. 2 (2009), 385 - 390. Print.

Eckert, Penelope. "Variation and the indexical field." *Journal of Sociolinguistics.* Vol. 12, No. 4 (2008): 453-476. Print.

Ervin, Staussa C. "The Attitudes of Higher Education Faculty Toward African American Vernacular English." Diss. Texas A&M University, 2005.

Ewers, Traute. *The Origins of American Black English.* New York: Mouton de Gruyter, 1996. Print.

Fasold, Ralph. "The relationship between Black and White speech in the South." *American Speech* Vol. 56, No. 2 (1981): 163-189. Print.

Fought, Carmen. *Language and Ethnicity.* Cambridge, United Kingdom: Cambridge University Press, 2006. Print.

Garner, Roberta. *The Joy of Stats: A Short Guide to Introductory Statistics in the Social Sciences.* United Kingdom: University of Toronto Press Incorporated, 2010. Print.

Gordon, Matthew J. "Phonological Correlates of Ethnic Identity: Evidence of Divergence?" *American Speech* Vol. 75 No. 2 (2000): 115-136. Print.

Gordon, Matthew J. *Small-town Values and Big-city Vowels: A Study of the Northern Cities Shift in Michigan*. Durham, North Carolina, 2001: Duke University Press. Print.

Gorman, Kyle and Daniel Ezra Johnson. "Quantitative Analysis." *The Oxford Handbook of Sociolinguistics.* Ed. Bayley, Robert, and Richard Cameron, and Ceil Lucas. New York, New York: Oxford University Press, 2013. 214-240. Print.

Green, Lisa J. *African American English A Linguistic Introduction.* Cambridge, United Kingdom: Cambridge University Press, 2002. Print.

Green, Lisa. "Research on African American English since 1998 Origins, Description, Theory, and Practice. *Journal of English Linguistics* Vol. 32 No. 3 (2004): 210-229. Print.

Green, Lisa. "African-American vernacular English." Eds. Finegan, Edward and John R. Rickford. *Language in the USA*. New York: Cambridge University Press, 2004. Print.

Haeri, Niloofar. "A Linguistic Innovation of Women in Cairo." *Language Variation and Change*, Vol. 6 No. 1 (1994): 87-112. Print.

Hudley, Anne H. Charity and Christine Mallinson. *Understanding English Language Variation in U.S. Schools.* New York, New York: Teachers College Press, 2011. Print.

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

Ino, Tadashi et al. "Gender Differences in Brain Activation During Encoding and Recognition of Male and Female Face." *Brain Imaging and Behavior* Vol. 4 (2010): 55-67. Print.

Lambert, W.E. et al. "Evaluational Reactions to Spoken Languages." *Journal of Abnormal and Social Psychology* Vol. 60, No. 1 (1960): 44-51. Print.

Lambert, Wallace E. "A Social Psychology of Bilingualism" *Journal of Social Issues* Vol XXIII, No. 2 (1967) 91-109. Print.

Labov, William. "The social motivation of a sound change." *Word* Vol. 19 (1963): 273-309. Print.

Labov, William, Paul Cohen Clarence Robbins, and John Lewis. *A Study of the Non-Standard English of Negro and Puerto-Rican Speakers in New York City. Final Report, Cooperative Research Project 3228*. Vols. I and II (1968). Philadelphia: US Regional Survey. Print.

Labov, William. "Contraction, Deletion, and Inherent Variability of the English Copula." *Language* Vol. 45, No. 4 (1969): 715-762. Print.

Labov, William. *Sociolinguistic Patterns*. Oxford: Blackwell, 1972. Print.

Labov, William. "Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor." *Language in Society* Vol. 11, No. 2 (1982): 165-201. Print.

Labov, William. "Are Black and White Vernaculars Diverging? Papers from the NWAVE XIV Panel Discussion." *American Speech* Vol. 62 No.1 (1987): 5-12. Print.

Labov, William. "The intersection of sex and social class in the course of linguistic change." *Language Variation and Change* Vol. 2 (1990), 205-254. Print.

Labov, William. *Principles of Linguistic Change Vol. 1: Internal factors.* Oxford: Blackwell Publishers, 1994. Print.

Labov, William. *The Social Stratification of English in New York City.* Cambridge: Cambridge University Press, 2006. Print.

Labov, William and Sharon Ash and Charles Boberg. *The Atlas of North American English*. Berlin: Mouton-de Gruyter, 2006. Print.

Labov, William, and Wendell A. Harris. 1986. "De facto segregation of Black and White vernaculars." David Sankoff, ed., *Diversity and Diachrony*. Amsterdam: John Benjamins, 1-24. Print.

Lewin, C. and Herlitz, A. "Sex differences in face recognition: women's faces make the difference." *Brain and Cognition*. Vol. 50 (2002): 121– 128. Print.

Light, Caitlin. "Tutorial: Pearson's Chi-square Test for Independence." *University of Pennsylvania*. Web. Accessed on 22 June 2013.

Lippi-Green, Rosina. *English with an Accent.* New York, New York: Routledge, 2012. Print.

Long, Daniel and Denis R. Preston eds. *Handbook of Perceptual Dialectology* Vol. 1.Amsterdam: John Benjamins Publishing Company, 2002. Print.

Loudermilk, B., Gutierrez, E., & Corina, D. (submitted). "What are you talkin' about? Perception of sociolinguistic variation: Evidence from event-related potentials." *Journal of Cognitive Neuroscience.* PDF.

Luhman, Reid. "Appalachian English stereotypes: Language attitudes in Kentucky." *Language in Society* Vol. 19 (1990): 331-348. Print.

Makoni, S., Smitherman, G., Ball, A., and Spears, A. *Black Linguistics Language, Society, and Politics in Africa and the Americas.* New York: NY. Routledge, 2003. Print.

Map of the United States. Digital image. *New York Search State Search*. New York Search Engine, n.d. Web. <http://www.newyorkstatesearch.com/United_States_Maps/United_States_Map/United_States_map.jpg>

McGowan, Kevin B. "Socioindexical expectation and speech perception in noise: experienced and inexperienced listeners." 2012. 1-33. PDF.

McWhorter, John. *Word on the Street Debunking the Myth of a "Pure" Standard English.* United States: Basic Books, 1998. Print.

Meyerhoff, Miriam. *Introducing Sociolinguistics.* New York, New York: Routledge, 2006. Print.

Milroy, Leslie and Matthew Gordon. *Sociolinguistics Method and Interpretation.* Oxford, United Kingdom: Blackwell Publishing, 2003. Print.

Moller, Aage R. *Hearing: Anatomy, Physiology, and Disorders of the Auditory System, 2nd Edition.* Burlington, Massachusetts: Academic Press, 2006. Print.

Mufwene, Salikoko. "The shared ancestry of African-American and American White Southern Englishes: Some speculations dictated by history." *English in the Southern United States*. Ed.

Nagle, Steven J. and Sara L. Sanders. Cambridge, United Kingdom: Cambridge University Press, 2003. 64-81. Print.

Mufwene, Salikoko. "Creolization is a social not a structural process." In Ingrid Neumann-Holzschuh & Edgar Schneider, eds., *Degrees of restructuring in Creole languages*. Philadelphia: John Benjamins, 2000. 65-84. Print.

Mullenix, J. W.et al. " Effects of Variation in Emotional Tone of Voice on Speech Perception." *Language and Speech* Vol. 45 No. 3 (2002): 255-283. Print.

Niedzielski, Nancy. "The Effect of Social Information on the Perception of Sociolinguistic Variables." *Journal of Language and Social Psychology* Vol. 18, No.1 (1999): 62-85. Print.

Nygaard, Lynne C. "Perceptual Integration of Linguistic and Nonlinguistic Properties of Speech." *The Handbook of Speech Perception.* Eds. Pisoni, David B. and Robert E. Remez. Oxford, United Kingdom: Blackwell Publishing Ltd, 2005. 390-414. Print.

Nygaard, Lynne C. and Jennifer S. Queen. "Communicating emotion: Linking affective prosody and word meaning." *Journal of Experimental Psychology: Human Perception and Performance,* Vol 34 No. 4 (2008): 1017-1030. Print.

Oliver, Wendy Seay. "Attitudes Towards African American Vernacular English in African American Preschool Teachers." Diss. Capella University, 2012.

Perry, Theresa and Lisa Delpit. *The Real Ebonics Debate: Power, Language, and the Education of African American Children.* Boston, Massachusetts: Beacon Press, 1998. Print.

Poplack, Shana and Sali Tagliamonte. *African American English in the Diaspora.* Oxford, United Kingdom: Blackwell Publishers, 2001. Print.

Boersma, Paul & Weenink, David (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.55, retrieved October 2013 from http://www.praat.org/

Preston, Dennis R., ed. *Handbook of Perceptual Dialectology* Vol. 2. Amsterdam: John Benjamins Publishing Company, 1999. Print.

Preston, Dennis R. "Language with an Attitude. "The Handbook of Language Variation and Change. Ed. Chambers, J.K., Trudgill, Peter and Natalie Schilling-Estes. Oxford: Blackwell Publishing, 2004. Print.

Purnell, Thomas and William Idsardi and John Bough. "Perceptual and Phonetic Experiments on American Dialect Identification." *Journal of Language and Social Psychology.* Vol. 18, No. 1 (1999): 10-30. Print.

Rahman, Jacquelyn. "Middle-Class African American Reactions and Attitudes toward African American English." *American Speech.* Vol. 83, No. 2 (2008): 141-176. Print.

Rasinger, Sebastian M. "Quantitative Methods: Concepts, Frameworks and Issues." *Research Methods in Linguistics.* Ed. Liatosseliti. New York, New York: continuum, 2010. 49-67. Print.

Rickford, John R. *African American Vernacular English.* Oxford, United Kingdom: Blackwell Publishers Inc., 1999. Print.

Rickford, John R. "Grammatical variation and divergence in Vernacular Black English." *Internal and external factors in syntactic change.* Eds. Gerritsen, Marinel and Dieter Stein. New York: Mouton de Gruyter, 1992. 174-200. Print.

Rickford, John. "Rachel Jeantel's language in the Zimmerman trial." *Language Log.* Blog. 10 July 2013. Web. Accessed on 1 Aug. 2013.

Rickford, John "Intelligibility of AAVE and other dialects by speakers of other varieties." Email to VAR-L@jiscmail.ac.uk. 1 Aug. 2013. Web. Accessed on 1 Aug. 2013.

Rickford, John R. and Russell John Rickford. *Spoken Soul The Story of Black English.* New York: John Riley & Sons, Inc, 2000. Print.

Rubin, Donald L. and Kim A. Smith. "Effects of Accent, Ethnicity, and Lecture Topic on Undergraduates' Perceptions of Nonnative English-Speaking Teaching Assistants." *International Journal of Intercultural Relations* Vol. 14 (1990): 337-353. Print.

Rubin, Donald L. "Nonlanguage Factors Affecting Undergraduates' Judgments of Nonnative English-Speaking Teaching Assistants." *Research on Higher Education* Vol. 33, No 4 (1992): 511-531. Print.

Ryan, James M. "The Chi Square Statistic." *Hobart and William Smith Colleges.* Web. Accessed on 5 July 2013.

Scott, Barbara Ann. "Differential Listener Perceptions of African-American English." Diss. University of Illinois at Urbana-Champaign, 1999.

Salkind, Neil J. *Statistics For People Who (Think They) Hate Statistics.* London, United Kingdoom: Sage, 2004. Print.

Shaw, J. I., & Skolnick, P. "Sex-differences, weapon focus, and eyewitness reliability." *Journal of Social Psychology.* Vol. 134 (1994): 413–420. Print.

Shaw, J. I., & Skolnick, P. "Weapon focus and gender differences in eyewitness accuracy: arousal versus salience." *Journal of Applied Social Psychology.* Vol. 29 (1999): 2328–2341. Print.

Silhouette Female Grey. Digital image. *The Online Royalty Free Public Domain Clip Art*. N.p., n.d. Web. <http://www.clker.com/cliparts/z/r/A/H/X/g/silhouette-female-grey-md.png>

Smitherman, Geneva. *Word from the Mother Language and African Americans*. New York, New York: Routledge, 2006. Print.

State Map of Michigan and Michigan Counties. Digital image. *Ask the EConsultant*. EConsultant, n.d. Web. <http://www.econsultant.com/usa-state-maps-with-county-lines-county-names/mi-3-state-map-with-county-lines-county-names.gif>.

Staum, Laura. "Experimental Investigations of Sociolinguistic Knowledge." Diss. Stanford University, 2008.

Street, R. L., Jr., and Brady, R. M. "Speech rate acceptance range as a function of evaluative domain, listener speech rate, and communication context." *Communication Monographs* Vol. 49 (1982): 290-308. Print.

Street, R. L., Jr., Brady, R. M. and W. Outman. "The influence of speech rate stereotypes and rate similarity on listerners' evaluations of speakers." *Journal of Language and Social Psychology* Vol. 2 (1983): 37-56. Print.

Tagliamonte, Sali A. *Variationist Sociolinguistics Change, Observation, Interpretation.* Oxford, United Kingdom:Wiley-Blackwell, 2012. Print.

Tan, Ying Ying. "Age as a factor in ethnic accent identification in Singapore." *Journal of Multilingual and Multicultural Development* Vol. 33, No. 6 (2012): 569-587. Print.

Tarr, Michael J. *Center for the Neural Basis of Cognition and Department of Psychology,* Carnegie Mellon University, http://www.tarrlab.org/. Funding provided by NSF award.

Taylor, Orlando L. and Dorian L. Lee. "Teachers' attitudes toward black and nonstandard English as measured by the language attitude scale." *Language Attitudes: Current Trends and Prospects.* Eds *Press* Roger W. Shuy and Ralph W. Fasold. Washington, DC: Georgetown University, 1973. 174-201. Print.

Thomas, Erik R. "Phonological and Phonetic Characteristics of African American Vernacular English." *Language and Linguistics Compass* 1/5 (2007): 450-475. Print.

Thomas, Erik R. & Jeffrey Reaser. "Delimiting perceptual cues used for the ethnic labeling of African American and European American voices." *Journal of Sociolinguistics* Vol. 8 No. 1 (2004): 54–87. Print.

Vaughn-Cooke, Fay Boyd. "Are Black and White Vernaculars Diverging? Papers from the NWAVE XIV Panel Discussion." *American Speech* Vol. 62 No.1 (1987): 12-32. Print.

Vaughn-Cooke, Fay Boyd. "Lessons learned from the Ebonic controversy: implications for language assessment." *Sociolinguistic Variation.* Eds. Bayley, Robert and Ceil Lucas. Cambrdige, United Kingdom: Cambridge University Press, 2007. 254-275. Print.

Wardhough, Ronald. *An Introduction to Sociolinguistics Sixth Edition.* United Kingdom: Wiley-Blackwell, 2010. Print.

Wolfram, W. A. *A Sociolinguistic Description of Detroit Negro Speech*. Center for Applied Linguistics, 1969. *ProQuest.* 11 Sept. 2013. Web.

Wolfram, Walt. "Are Black and White Vernaculars Diverging? Papers from the NWAVE XIV Panel Discussion." *American Speech* Vol. 62 No.1 (1987): 40-48. Print.

Wolfram, Walt. "Re-Examining Vernacular Black English." Linguistic Society of America Vol. 66 No. 1 (1990): 121-133. Print.

Wolfram, Walt. *Dialects and American English.* London, United Kingdom: Prentince-Hall, 1991. Print.

Wolfram, Walt and Natalie Schilling-Estes. *American English.* Malden: Blackwell Publishing, 2006. Print.

Wolfram, Walt and Ralph W. Fasold. *The study of Social Dialects in American English.* Englewood Cliffs, New Jersey: Prentice-Hall, 1974. Print.

Wright, D. B., and Sladden, B. "An own gender bias and the importance of hair in face recognition." *Acta Psychologica.* Vol. 114 (2003): 101–114. Print.

APPENDIX

Appendix A – PowerPoint Example

Transcription Task

- You will not be timed.
- **There is no right or wrong answer.**
- This task will take about 10 minutes to complete.
- It includes 1 practice sentence, 30 short sentences, and a brief questionnaire.
- Please write down your answers as **neatly/legibly** as possible.

Directions:

- A photo is given.
- To move to the next slide, push the → arrow.
- A photo appears. Push the → arrow again and an audio icon will appear.
- Push the → arrow to listen to the audio.
- After listening to the audio, **write down what you heard to the best of your ability**
- You may only **listen to an audio one time**.
- If you are unsure, try your **best guess** of what you think you heard.

Please note:

- Audio files will have white noise added to the background and volume may change because there are different speakers who are speaking, so listen carefully.

White Noise

- This is what the white noise will sound like.
- Adjust the volume accordingly.

Practice #1

Practice #1

End of practice.

Do you have any questions?

1.

1.

2.

2.

3.

3.

4.

4.

5.

5.

6.

6.

7.

7.

8.

8.

Appendix B – Consent Form

CONSENT FORM
University of Essex
Dept. of Language & Linguistics
Colchester CO4 3SQ

Task

I am a current graduate student studying at the University of Essex. The task that you are about to take will be done in order to help me complete my master's degree thesis; you will be asked to write down 30 short sentences that will be played aloud. All the data I will receive from you will be saved in an archive and only used for research purposes.

I, Elisa Lenneman, promise:

- To not publish real names/addresses in any project reports, or give them out to the public. I'll protect, to the best of my ability, the confidentiality of all participants that partake in this task.
- The materials made as part of the research will only be used for educational/scholarly purposes (not for profit).
- No copies of these materials will be made, and nothing from them will be published without the consent of my program supervisor or my supervisor. If you have questions, please feel free to contact them at: patrickp@essex.ac.uk or vineeta@essex.ac.uk. If you have any questions you'd like to ask me, you can email at: elenne@essex.ac.uk

Experiment conductor: Elisa Lenneman — Program Coordinator: Professor Peter Patrick

Date: ___/___/___ — Supervisor: Dr. Vineeta Chand

The person(s) taking this task agrees:

☐ I've had a chance to ask questions. I know I can withdraw at any time from the task, with no penalty and without giving any reasons.
☐ It's okay with me if some excerpts from the task I take are published, in print or on the web, for research purposes – as long as they protect names, addresses & identifying personal information.
☐ I understand the researcher isn't doing this for the money, and I'm not asking to be paid either.
☐ My choice to participate in this task is voluntary.

☐ If I have other restrictions on use of this information I'll make them clear now, so that we can agree on them now; if we can't, I have the right to see all materials destroyed.

Name(s) of person(s) participating:

_____

Signed: _____ _____

Date: _____ _____

Additional Comments:

Appendix C – Worksheet

Transcription Task

CONTACT INFORMATION:

Elisa Lenneman: elenne@essex.ac.uk
PhD Vineeta Chand (advisor): vineeta@essex.ac.uk
PhD Peter Patrick (program coordinator): patrickp@essex.ac.uk

Directions:
This task will take about 10 minutes to complete. You will be listening to 30 short sentences accompanied by a photo. To move to the next slide, push the → arrow. First a photo will be shown. When you go to the next slide, the same photo will be shown and a sentence will automatically play. There is one practice sentence so you can adjust the volume accordingly. Every sentence will have added white noise in the background. After listening to a sentence, please write down what you were able to hear. You will only be able to listen to each sentence once. Please write as neatly as you can, and try your best to write down these sentences as accurately as possible. If you are unsure, write down what you think you heard. *There is no right or wrong answer, you will not be timed, and all responses will be anonymous.* If you have any questions, you can ask me at any time before, during, or after the task. After this is completed, I will tell you what I am studying and why I used this task. You are free to stop this task at any time you wish and without giving any reasons.

PRACTICE:
#1.

End of practice…

1.

2.

3.

4.

5.

6.

7.

Appendix D – Questionnaire

Age:	
Gender:	
Education Level:	
Where did you grow up? (City, State)	
Do you have a hearing disability?	

1. How did you feel about this task?

2. Were there parts that were easier for you to transcribe? Which ones and why?

3. Were there parts that were more difficult for you to transcribe? Which ones and why?

4. Did the photos help you with this transcription task? Why or why not?

5. Do you speak African American Vernacular English (AAVE), otherwise known as "ebonics"?

YES or NO

6. Do you have a lot of contact with people who speak AAVE?

YES or NO

7. Please circle how familiar you are with AAVE:

1 = I am unfamiliar with AAVE

2 = I am familiar with AAVE

3 = I am very familiar with AAVE

8. Please circle how familiar you with hip-hop/rap culture?

1 = I am unfamiliar with hip-hop/rap culture

2 = I am familiar with hip-hop/rap culture

3 = I am very familiar with hip-hop/rap culture

9. Is there anything else you noticed?

10. Do you have any other comments you would like to add?

Appendix E – Sentences

Low Predictability Sentences
He pointed at the *animals*.
We pointed at the *bird*.
Dad talked about the *bomb*.
He pointed at the *cents*.
She looked at the *clock*.
We read about the *coach*.
Mom pointed at the *coffee*.
There are many *days*.
He talked about the *dinner*.
We read about the *family*.
She thinks that it is *fast*.
Mom pointed at his *father*.
Mom looked at her *feet*.
Dad pointed at the *grass*.
She pointed at her *head*.
Mom looked at the *juice*.
She talked about the *leaves*.
She talked about their *necks*.
Mom talked about the *pie*.
Dad talked about the *sheets*.
Dad read about the *sky*.
He looked at the *sleeves*.
This is her favorite *sport*.
We looked at the *story*.
This is her favorite *time*.
He read about the *trees*.
We talked about the *water*.
This is her favorite *week*.
He looked at her wrist.
Mom thinks that it is *yellow*.

Low predictability sentences used in this experiment from McGowan (2012), Originally from Bradlow and Alexander (2007). Target words are italicized.