

“Sounding Black”: The Legal Implications of Linguistic Profiling

Abbie MacNeal, Katherine Fiallo, Alexander Jones, Shaughnessy Jones, Samantha Laureano, Matthew Monjarrez, Yian Xu

Introduction

Linguistic profiling occurs when a listener uses auditory cues to identify social characteristics, such as race, gender, sexual orientation, or geographic origin. Linguistic profiling is a natural and automatic psychological process. Though it is not itself inherently discriminatory, it can contribute to racial profiling, which is inherently discriminatory. Evidence of discrimination resulting from linguistic profiling is shown in research by John Baugh, Jeffrey Grogger, and others. Linguistic profiling interacts with the law in a variety of ways, both discriminatory and non-discriminatory. Witnesses have based their testimonies on linguistic profiling that occurred during the crime, when they overheard the suspect. Linguistic profiling has also led to race-based discrimination in the housing market, and several cases have been brought to court. In other cases, linguistic profiling has led to key witness testimony being thrown out because of discrimination attached to a witness’s speech. The goal of this article is to highlight the kinds of linguistic profiling that bear on legal issues.

1 Linguistic Profiling

1.1 Some Studies

After confirming several appointments to view an apartment over the phone, linguist John Baugh was turned away when he arrived in person to see them. Baugh, an African American man who speaks “regular,” or Standard, English, suspected he was being turned away because he was African American and wondered if he would have even been able to book the same appointments if he “sounded black” on the phone. This personal experience led Baugh to his research on linguistic profiling.

Baugh defined “linguistic profiling” as “racial identification based on speech.” His first experiment¹ with Thomas Purnell and William Idsardi addressed whether housing developers would use linguistic profiling simply based on audio cues. Would applicants calling about apartments be treated differently based on their speech? Baugh, Purnell, and Idsardi theorized that they would; they suspected that callers using dialects of English which were not “standard” would confirm less appointments to view apartments.

In this experiment, Baugh called landlords who had apartments advertised for rent in five different geographic areas varying in racial and socioeconomic demographics. He used three distinct dialects to call each landlord: Standard American English (SAE), African American English (AAE), and Chicano English (ChE). Standard American English is accepted as the “mainstream” dialect of American English and is typically spoken by white Americans. African American English is the dialect of English spoken mainly by working and middle class African Americans. AAE is also sometimes referred to as African American Vernacular English, Black English, or Ebonics, though AAE is the term considered to be standard among linguists. Chicano English is spoken mainly by Mexican Americans. These dialects sound distinctly different and

¹ Thomas Purnell, William Idsardi & John Baugh, *Perceptual and Phonetic Experiments on American English Dialect Identification*, 18 *Journal of Language and Social Psychology* 10–30 (1999).

have different grammatical rules. Baugh, Purnell, and Idsardi hypothesized that SAE would be preferred over AAE and ChE.

The results supported their hypothesis and revealed some other telling patterns. The percentage of calls using SAE that confirmed appointments to view the apartments was relatively similar across the sample areas, regardless of the area's white population, which demonstrated a general acceptance and preference for the dialect. The percentage of AAE and ChE calls that confirmed appointments in each area paralleled the black and Hispanic populations of each area, respectively. In one area with a high white population and lower black populations, 70.1% of SAE calls confirmed an appointment, whereas only 28.7% of AAE calls confirmed an appointment. In another area, this one with a low white population and higher black population, 68.7% of white calls confirmed appointments and 72.0% of AAE calls confirmed appointments. Despite these areas large differences in white populations, the percentage of SAE calls which received appointments were similar, whereas the percentage of AAE calls that confirmed appointments increased as the black population in the area increased.

The authors' claim that the differences in appointment acceptance rates reflects dialect differences is supported by a series of earlier, related studies. These earlier studies found that dialects could be easily identified by listeners. The first of these studies found that listeners could identify the dialect used in the sentence "Hello, I am calling to see about the apartment you have advertised in the paper." The second study found that listeners could reliably identify dialect just from hearing the word "hello." These results indicate that these differences in accepted appointments are very likely to be due to differences in dialect.

Jeffrey Grogger² found evidence that dialect differences were responsible for wage discrimination. He asked listeners to try to identify a speaker's race from an audio recording. He then compared the speakers perceived race with their total earnings. He found that listeners could accurately identify a speaker's race—with 83.6% accuracy for white speakers and 77.1% accuracy for black speakers—but could not accurately identify their education level or region of origin (less than 50% accuracy). Grogger also found a correlation between a speaker's perceived race and their total earnings: African American speakers who were identified as black based on their speech earned 12% less than the African American workers who were not identified as black. Those African American speakers who were not identified as black (i.e. who 'sounded white') earned as much as white workers.

1.2 Linguistic Profiling vs. Racial Profiling

Linguistic profiling is often conflated with racial profiling, but as Baugh emphasizes, they are not the same. Linguistic profiling occurs when a listener uses auditory cues to identify social characteristics, such as race, gender, sexual orientation, or geographic origin, while racial profiling relies on all information about a person's race, including but not limited to audio and visual cues. Racial profiling always refers to a discriminatory situation, by which we mean unjust or prejudicial treatment based on race, which can be illegal. Linguistic profiling, however, is used only to describe the use of auditory cues to determine a speaker's race. While it can lead to discrimination on the part of the listener, it is not, in itself, a discriminatory act. Some cases of linguistic profiling, such as the housing discrimination shown in Baugh's experiments, are clear examples of discrimination and racial profiling whereas some other instances are not as obvious. Other instances lead to no discrimination at all. For instance, linguistic profiling occurs when a person recognizes someone else as coming from the Southern United States based on their

² Jeffrey Grogger, *Speech Patterns and Racial Wage Inequality*, 46 *Journal of Human Resources* 1–25 (2010).

speech. This normally does not lead to prejudice, though a listener can attach prejudice to it (e.g. determining that someone is from the South based on linguistic profiling and then applying discriminatory stereotypes, such as assuming the speaker is dumb because they have a “backwater” accent).

1.3 An Outline of This Paper

The next section will discuss the psychological basis for linguistic profiling – the natural human tendency to categorize people by their language. The Section 3 demonstrates the linguistic facts that show that no dialect or language is more or less intelligent, inherently better or worse than any other. Section 4 turns to several examples that show how linguistic profiling can interact with the law. This section includes examples of criminal cases that involve linguistic profiling as a way for witnesses to identify likely characteristics of suspects, an example of discriminatory linguistic profiling in the housing market which lead to legal ramifications, and an example of linguistic profiling in the courtroom causing key testimony to be thrown out.

2 The Psychological Perspective: The Multi-Faceted Nature of Linguistic Profiling

2.1 What Underlies Linguistic Profiling?

In order to better understand linguistic profiling as a type of human behavior, we will devote this section to a psychological discussion on the nature of linguistic profiling, including the fundamental reasons for it. The definition of linguistic profiling – identifying individuals based on their speech – relates to cognitive processes. In the context of legal cases it is important to ask whether linguistic profiling should be classified as a “motivated act.”

Cognitive and social psychologists have spent decades exploring how people see others and make sense of the world. We live in a complicated environment where the sheer number of objects we encounter every day is simply beyond the scope of our mental capacity to track on an individual basis. It is simply impossible for humans to keep track of every single detail of each entity. However, having some general knowledge about the objects in our environment is essential for survival. We need to know which fruits are edible and which are poisonous, just as which groups of people to welcome and which to avoid. How do we resolve the tension between our limited cognitive resources and the vital demand on these resources to understand the environment?

The answer is that we rely on categories. Categories are the basis for how people represent and reason about the empirical world. They allow us to group together entities that share common features (based on perceptual, functional, or other cues) and form a common mental representation to store all the critical information. Depending on the specific domain and objects, sometimes those category features are grounded on provable facts: for example, it is evolutionarily beneficial to recognize that tigers attack humans, but this statement may not equally apply to every single tiger. Other times, a generic statement about a certain group of entities may simply be wrong – for example, the stereotypical claim that all Asians are good at math. It is widely accepted in the scientific community that individual variation within a singular racial group sufficiently exceeds the variation across racial groups.³

³ Dar-Nimrod, I., & Heine, S. (2011). Genetic Essentialism: On the Deceptive Determinism of DNA. *Psychological Bulletin*, 137(5), 800-818.

2.2 Cognitive Categories: An Automatic and Neutral Psychological Mechanism

Regardless of the factual accuracy of recognized category features, category concepts serve an important storage function in inductive reasoning. In other words, these concepts allow people to apply what they already know about the world to make predictions about what they do not know. This is true even for children as young as four. In a 2002 study by Jaswal and Markman, children relied on category labels provided by the adult experimenters rather than perceptual similarities among the stimuli to make judgments on features that a novel stimulus would exhibit.⁴ These toddlers seemed to perceive that category membership, denoted by a shared underlying factor expressed in a label, determines the important features of individual members. In another study, preschool-aged participants were given a scenario in which a rabbit was raised by a group of monkeys (the “Switched-at-Birth” paradigm). Experimenters asked the children whether they thought the rabbit would show the same physical and/or behavioral traits as either rabbits or monkeys.⁵ The preschoolers claimed that the rabbit would still have long ears, just like other rabbits, and would like to eat carrots. They demonstrated that even though they knew the rabbit grew up with monkeys, it still belonged to the “rabbit” category and shared characteristics with other rabbits. Taken together, evidence from a body of developmental research suggests an innate and universal tendency for humans to:

1. rely on category concepts and labels as an inductive basis to make predictions about novel instances, and
2. attribute outward individual traits to a deep, underlying cause.

Cognitive psychologists use the term ‘psychological essentialism’ to describe the belief system that an underlying cause determines category membership and gives rise to individual features.^{5,6} It is important to note that this theory describes the structure of mental representations but does not claim that these are objectively true or supported by scientific proof. Under this framework of categorization, linguistic profiling could be understood as a particular manifestation of how people rely on accessible cues. In this case, speech patterns are used to form meaningful categories of social groups. As we sort others into social groups by speech pattern, we infer other traits based on our definition of that social group.

Distinct speech patterns feed into another category, one based on race, since we find that certain dialects – African American English, for example – are predominantly spoken by members of a certain race – African Americans. And since our cognitive predisposition to sort entities into categories applies to all kinds of categorization,⁷ when we sort people by their speech patterns, linguistic profiling is the result. So to be clear, linguistic profiling is the result of a fundamental human cognitive process.

⁴ Jaswal, V. K., & Markman, E. M. (2002). Children’s Acceptance and Use of Unexpected Category Labels to Draw Non-Obvious Inferences. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 24.

⁵ Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, 38(3), 213-244.

⁶ Gelman, S. A. (2003). *The essential child: origins of essentialism in everyday thought*. New York: Oxford University Press.

⁷ Medin, D. L., & Aguilar, C. M. (1999). Categorization. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 104-106). Cambridge, MA: MIT Press.

2.3 Do People Use Speech Patterns to Form Categories?

A number of studies have investigated the extent to which people essentialize a variety of different social cues, such as race, gender, socio-economic status, nationality, and so forth. Empirical evidence across cultural samples suggests that people categorize social groups the same way that they categorize other biological species without realizing that many of the features are socially constructed.⁸ Nevertheless, people do not assign the same weights to different social cues. According to Haslam and colleagues,⁹ the underlying essence of a social category includes two aspects:

1. a social category is naturally defined instead of artificially created (naturalness)
2. members of the same category share uniform features (entitativity)

Therefore, knowing that someone belongs to a category provides information about that person. Evidence suggests that people vary along these two sub-scales of essentialist thinking both within and between social categories.¹⁰ For instance, in Haslam's original study, politically affiliated categories were highly essentialized on the entitativity but not the naturalness dimension and the category of "male" was highly essentialized on the naturalness but not the entitativity dimension.

But now the question arises: Do people use speech patterns to define social categories? The answer is yes – and from a young age. In one experiment by Hirschfeld and Gelman, researchers presented a hypothetical scenario to a group of five-year-olds, in which they were told about a child born to a family that speaks English, but was raised by another family that speaks Portuguese (and vice versa in another subject group). They were asked to determine whether the child would speak the same language as their birth parents or adoptive parents. By five years of age, most children favored birth parents over adoptive parents, indicating that they believe language is something inheritable and fixed at birth.¹¹ Even though the adoptive family speaks another language, children reasoned that language was something tied directly to that child's essence and would always remain.

Another study, by Kinzler and Dautel, explicitly tested children's essentialist reasoning about race and language by pitting these two elements directly against each other.¹² In this study, a group of five- to six-year-old European American children were given a trial task. Each trial presented a sample face labeled with their race and the language they spoke, and then two options. One option matched the race of the sample face but spoke a different language; the other option matched the language of the sample face but was of a different race. The subjects were asked to judge whether the sample face would grow up to be similar to the one who spoke the same language or the one that was the same race. Intriguingly, these five- to six-year-old children favored the language-match option, suggesting that they view one's language as even more stable and immutable than one's race. These two studies taken together provide important

⁸ Haslam, Nick. "More Human than Others? A Critique of Cypryańska Et Al. (2017)." *The Journal of Social Psychology* 157, no. 2 (2017): 143-47.

⁹ Haslam, N., Rothschild, L. and Ernst, D. (2000), Essentialist beliefs about social categories. *British Journal of Social Psychology*, 39: 113–127.

¹⁰ *Id.*

¹¹ Hirschfeld, L. A., & Gelman, S. A. (1997). What young children think about the relationship between language variation and social difference. *Cognitive Development*, 12, 213-238.

¹² Kinzler, K. D. and Dautel, J. B. (2012), Children's essentialist reasoning about language and race. *Developmental Science*, 15: 131–138.

developmental evidence suggesting that spoken language, although not yet specified at the level of speech pattern within a given language, constitutes a powerful social category.

2.4 Linguistic Profiling: Is it Conscious or Unconscious?

Now that we see that (1) linguistic profiling can be understood as a form of mental categorization and (2) spoken language constitutes a crucial category for people to use when reasoning about others, to what extent should we consider linguistic profiling to be a conscious, motivated act? Category concepts can be activated automatically and subconsciously, along with their related stereotypes. In one classic experiment, Bargh, Chen, and Burrows primed participants with the category of “elderly” using an implicit measure, a scrambled sentence task.¹³ Participants were told that they were just completing a verbal test. However, in those sentences, they were not processing random words. Instead, half of the participants saw words that were stereotypically relevant to the category of “elderly,” such as “retired” or “Florida.” Once participants unscrambled the sentences and completed some filler questions, they were told they were done and instructed to leave. Surprisingly, without being aware of it, those primed with the stereotypical “elderly” concept walked out more slowly than those who had not seen “elderly” related words. Similar evidence was collected from other studies, demonstrating that category concepts could be easily activated below the conscious level. Furthermore, this also revealed that the automatic activation of category concepts can have an effect on outward behaviors.

Much of the previous research has discussed the dissociation between explicit and implicit (or subconscious, uncontrollable) attitudes towards social groups and suggested that these may underlie two distinct cognitive processes. So far, we have been discussing implicit attitudes towards social groups and the cognitive processes underlying them. But much research investigates our conscious, explicit attitudes towards social groups, and has suggested that these are driven by a distinct cognitive process. One example is overcompensation of prejudice, where a 2013 study by Mendes & Koslov has shown that Caucasian participants exhibited more positive behaviors towards Black targets than Caucasian targets, and that such overcompensation is even more pronounced in those with higher anti-black bias measured by an implicit association task.¹⁴ Whereas implicit attitudes are automatic, and do not take cognitive effort, this study showed that overcompensation requires cognitive resources, and such overcompensation was reduced when participants were under cognitive load (in other words, when their mental resources were being occupied by other tasks). A similarly-structured study focusing on anti-gay bias has suggested that heterosexual participants overcompensated their anti-gay bias by behaving overly positively towards gay targets

Piecing the evidence together, it is clear that linguistic profiling is activated automatically, without awareness and motivation. Of course, this is not to deny that humans also make conscious categorizations, such as when a listener wants to compensate for his or her unwanted unconscious biases. But linguistic profiling is a clear example of an automatic human categorization process.

¹³Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2), 230-244.

¹⁴ Mendes, W. B., & Koslov, K. (2013). Brittle smiles: Positive biases toward stigmatized and outgroup targets. *Journal of Experimental Psychology: General*, 142(3), 923-933.

2.5 Summary

To sum up, we discussed the nature of linguistic profiling by asking three questions. First, what is the cognitive basis for linguistic profiling? Second, since profiling requires social categorization, to what extent do people form social categories using language or speech patterns? Third, is linguistic profiling an explicit, conscious act? By reviewing experimental evidence from social, cognitive and developmental psychology research, we conclude that:

1. linguistic profiling reflects the central cognitive capacity of categorization, which helps us to construct our mental reality and inductively reason through novel instances;
2. people rely on many different social cues and intuitively endorse the belief that members of the same social category share an underlying factor in common, which determines individual features. In particular, language constitutes a crucial social category, which may – falsely – be considered inheritable and fixed at birth, and in some cases, even more immutable or predictable than race
3. linguistic profiling, just like other category concepts, can be activated automatically without the engagement of attentive awareness.
4. It is important to emphasize that linguistic profiling does not involve explicit processes. As such, linguistic profiling is an automatic and neutral cognitive process of social categorization.

3. The Grammar of African American English (AAE)

In 1983, John Baugh referenced “black street speech” to identify the dialect of English now referred to as African American English (AAE) or African American Vernacular English (AAVE). Contrary to Baugh’s phrase (and contrary to popular conceptions), this dialect is not spoken only in poor, urban communities; many of its speakers live outside of urban areas and come from a range of social backgrounds. Nevertheless, like other dialects typically associated with people of low socioeconomic status (for example, Appalachian English or Cockney English), the dialect has been devalued and is often seen as “incorrect” or “simplified” English. While there are idiosyncrasies to AAE, the same is true of all dialects, including those that are more highly valued in our society.

Like every other documented dialect, AAE follows strict grammatical rules.¹⁵ When linguists talk about grammatical rules, they don’t mean the prescribed conventions for formal English taught in school, like “Use *whom* for objects (*Whom did you invite?*) and *who* for subjects (*Who invited you?*). Rather, they are rules or formulas that express the patterns that the dialect follows. Most of the rules in AAE overlap with the rules of Standard American English (for instance, every complete sentence contains a verb). But some of them are particular to AAE alone and make the dialect distinctive.

3.1 AAE Syntax

Consider the rules for the verb “be”.¹⁶ When the goal is to express the idea that an action is “habitual,” Standard American English uses the present tense verb form, as in “Lily **jogs** after school.” The meaning is that Lily generally or usually (habitually) jogs after school. In African American English, the same meaning is expressed with the bare verb “be” followed by the verb

¹⁵ William Labov, Unendangered Dialect, *Endangered People: The Case Of African American Vernacular English*, 18 *Transforming Anthropology* 15–27 (2010).

¹⁶ Lisa J. Green, *African American English: a linguistic introduction* (2009).

in its progressive form (the form ending in *-ing*), such as in “Lily **be jogging** after school.” These rules apply consistently, so the same habitual interpretation applies whenever these constructions occur: “We **play** basketball on Saturday mornings” and “The doctor **is** here on Thursdays” in SAE are expressed as and “We **be playing** basketball on Saturday mornings” and “The doctor **be** here on Thursdays” in AAE.

AAE also has a unique rule when the use of “be” is not habitual, but only “momentary”. Consider the exchanges in (1):

(1) A: I can't find Lily.
 B: AAE She \emptyset playing basketball. I think she \emptyset in the schoolyard.

In the two sentences spoken by B, “is,” a form of “be,” is missing. In momentary contexts (she is playing basketball at this moment), “be” forms can be deleted in exactly the same places where it can be contracted in SAE, as shown in (2). (Note that this AAE rule is optional rule; be can also be present or contracted as in SAE.)

(2) A: I can't find Lily.
 B: SAE She is playing basketball. I think she is in the schoolyard.
 or, She's playing basketball. I think she's in the schoolyard.

But where “be” is NOT contractible in SAE, it is not deletable in AAE (* means not possible, ungrammatical in that dialect):

(3) A: I can't find Lily.
 B: I don't know where **she is**
 SAE *I don't know where **she's**
 AAE *I don't know where **she \emptyset**

Notice that deleted momentary “be” is a separate construction from habitual “be,” governed by a second set of rules of the AAE dialect.

3.2 AAE Morphology: Agreement

The AAE agreement system also differs systematically from SAE. It contains three rules that delete agreement markers. One rule deletes the possessive 's marker from a noun. The possessive is still part of the AAE dialect, as is clear from possessive pronouns like *her* and *his*, which don't use the marker:

(4)

SAE	AAE
daddy's house	daddy \emptyset house
his daddy's fiancée's house	his daddy \emptyset fiancée \emptyset house
his daddy	his daddy
her house	her house

Two other deletion rules apply to agreement markers. Regular plural forms in SAE, which are marked with -s and third person singular verbs, also marked with -s, are both unmarked in AAE, as shown in (5):

(5)

	SAE	AAE	
plural		a couple seconds later	a couple second \emptyset later
3rd person sg.		he says	he say \emptyset

3.3 AAE Phonology: Rules of the Sound System

The grammatical rules unique to AAE also include rules of the sound system, or phonology. One well-known rule turns the “th” sound, which is absent in AAE to a related sound, either t, d, f, or v, depending on the context. In SAE, there are two versions of “th”, the unvoiced version (as in “thy”) and the voiced version (as in “thigh”), which are written in the phonetic alphabet as [θ] and [ð] respectively. The sounds that substitute for the two versions of th are their phonetic neighbors. “t” and “d” are produced in the same place in the mouth and only differ in being “stops” (which stop the air completely before releasing it, not “fricatives” that don’t). “f” and “v”, like [θ] and [ð] are fricatives, but they are produced slightly forward in the mouth. The substitutions always match in voicing: unvoiced sounds are replaced by unvoiced sounds and voiced sounds by voiced sounds. The substitutions are shown in this chart, adapted from Green’s African American English: A Linguistic Introduction:

	SAE	SAE Phonetic Transcription	AAE	AAE Phonetic Transcription
1	thing	[θɪŋ]	thing	[θɪŋ]
	think	[θɪŋk]	think	[θɪŋk]
2	these	[ðiz]	dese	[diz]
	that	[ðæt]	dat	[dæt]
3	bath	[bætθ]	baf	[bæf]
4	with	[wɪθ]	wif	[wɪf]
			wit	[wɪt]
5	month,	[mʌntθ]	mont monf	[mʌnt] [mʌnf]
6	bathe	[beð]	bave	[bev]
7	smooth	[smuð]	smoove	[smuv]
8	Bethlehem	[beθləhem]	Beflehem	[befləhim]
9	mother	[mʌðə]	muva	[mʌvə]

At the beginning of a word, (rows 1 and 2) “th” is not replaced when it is unvoiced: [θ] (**th**ing stays [θɪŋ]), but is always replaced when it is voiced [ð]: **th**ese becomes [diz]; **th**at becomes [dæt]. At the end of a word, (rows 3-7) “th” is always replaced by a sound that matches in voicing: unvoiced θ (**ba**th, **mo**nth) is replaced either by “t” [mʌnt] or “f” ([bæf], [mʌnf]); voiced ð (**ba**the, **smoo**the) by “v” [bev], [smuv]. In the middle of a word (rows 8-9) Unvoiced θ (Bethlehem) is replaced by “f” [bɛfləhim]; voiced ð (mother) is replaced by “v” [mʌvə]. Using these rules, we can accurately predict the possible - and impossible - AAE pronunciations of words with a “th” sound.

Grammatical rules, like those that dictate AAE’s syntactic patterns and its phonology (its sound system) govern all dialects of all languages. However, speakers of other dialects of English do not generally recognize the systematicity of AAE. The dialect is often devalued in American society, a result of the devaluing of its speakers. A thorough linguistic analysis shows that the value of a dialect is tied to socioeconomic and racial factors in a given culture, not to any problems with its linguistic validity or correctness.

4. Overview of Linguistic Profiling and the Law

Linguistic profiling is often referred to as the linguistic equivalent of racial profiling because people can often easily determine a speaker’s ethnicity, race, or country of origin based on their speech.¹⁷ Linguistic profiling can often be relevant evidence in court cases. For example, if the victim of a crime could not see the perpetrator but heard his voice, that evidence can be invaluable to proving the perpetrator’s identity at trial. On the other hand, courts are tasked with enforcing anti-discrimination laws. If a person discriminates against someone else based on how that person speaks, then they have most likely violated state or federal anti-discrimination laws. These are the two main ways that linguistic profiling appears in courts. Courts generally allow the use of linguistic profiling in criminal cases to help identify the perpetrator of a crime, and also generally accept that discriminating against someone because of linguistic profiling violates anti-discrimination laws. In this section, we examine these two appearances of linguistic profiling in criminal and civil cases.

4.1 Criminal Cases

In all criminal cases, the prosecution must prove that the defendant before the court is the person who committed the offense charged. This is known as the “issue of identity,” and it is an essential element of all crimes. In order to establish identity, generally, courts in criminal cases allow linguistic profiling; that is, they usually allow witnesses to testify that a person “sounded” like they belonged to a particular race or ethnic group. The most common objection to the use of linguistic profiling in criminal cases is that a witness is not an expert in linguistics and cannot use linguistic profiling correctly. However, courts usually reject this argument, and hold that witnesses can testify about an accent or dialect as long as the witness has sufficient experience with it. But how much experience is “sufficient”? *People v. Sanchez* provides a good example.¹⁸

In *Sanchez*, the defendant, Wilberto Sanchez, was on trial for second-degree murder.¹⁹ During the police investigation, Mr. Sanchez, a Puerto Rican, told police that a man from the Dominican Republic who looked like him actually committed the crime.²⁰ This placed the issue

¹⁷ Dawn L. Smalls, *Linguistic Profiling and the Law*, 15 STAN. L. & POL’Y REV. 579, 579 (2004).

¹⁸ 492 N.Y.S.2d 683 (N. Y. Sup. Ct., Bronx C’nty 1985).

¹⁹ *Id.* at 683.

²⁰ *Id.* at 683–84.

of identity into contention. To counter this, the prosecution called a witness named Israel Torres who testified that he had heard the defendant arguing with the victim in Spanish just before the victim was shot.²¹ The prosecution asked Mr. Torres if the man arguing with the victim had a Puerto Rican or Dominican accent.²² The defense attorney objected to this question, arguing that Mr. Torres was not an expert in linguistics, and therefore could not opine to the court about the speaker's accent.²³

This objection raised the issue of the difference between lay opinion testimony and expert opinion testimony, which are governed by each individual jurisdiction's rules of evidence. All states and the federal judiciary have rules governing whether certain types of evidence and testimony can be admitted as evidence in court. These are usually called the "rules of evidence." If a statement or document is acceptable under these rules, it is called "admissible." If a statement or document is unacceptable, for whatever reason, and cannot come into evidence in court, it is called "inadmissible." Sometimes rules of evidence are published as a code of rules by statute (such as in the federal system), while in other jurisdictions the rules of evidence are not compiled on one code (such as in Massachusetts, which uses case law to develop rules of evidence, and therefore has never passed a full code of rules via statute). Under the rules of evidence applicable to the *Sanchez* court in 1985, a lay witness (one who is not an expert) could not give his or her opinion on a set of facts unless: "(1) the facts which constitute the opinion are incapable of description; (2) the subject matter does not require expert knowledge; and (3) the witness is qualified to give his opinion."²⁴

Under the New York rule applicable in *Sanchez*, a witness could give his or her opinion about several things. Witnesses could give their opinions on "matters of taste, smell, touch, color, weight, size, quantity, velocity, heat, cold, sickness, health, excitement, intoxication and disposition."²⁵ The *Sanchez* court went on to note that no decision appeared to address the issue of whether a lay witness can opine about the accent or dialect of a speaker.²⁶ However, the court noted that a person with proper experience could easily tell the difference between different types of accents or dialects.²⁷ The court found that an opinion about the accent or dialect of a speaker was a matter that was both incapable of description and did not require expert knowledge. The court also found that Mr. Torres was qualified to give his opinion about the speaker's accent because Mr. Torres had spoken Spanish his entire life with both Dominicans and Puerto Ricans and testified that the two groups have distinct dialects that are distinguishable from one another.²⁸ Based on this, the *Sanchez* court allowed Mr. Torres to testify as to whether the speaker "sounded" Puerto Rican or Dominican.²⁹

That same year, a panel of the Court of Appeals of Washington decided *State v. Kinard* in a similar manner.³⁰ Mr. Kinard, who was black, was convicted of a burglary in which the victim, Mrs. Barbara Cardell, was robbed in her home by two men, who covered her face with a pillow,

²¹ *Id.* at 683.

²² *Id.* at 683–84.

²³ *Id.* at 684.

²⁴ *Id.*

²⁵ *Id.*

²⁶ *Id.*

²⁷ *Id.*

²⁸ *Id.*

²⁹ *Id.* at 685.

³⁰ 696 P.2d 603 (Wash. App. 1985).

making it impossible for her to see them.^{31 32} However, she was able to hear them, and the trial judge allowed her to testify that the man who jumped on her bed “sounded black to me.”³³ The issue on appeal in *Kinard* was whether the trial judge was correct in allowing Mrs. Cardell to say her assailant “sounded black.”³⁴

In determining whether Mrs. Cardell’s statement was admissible, the *Kinard* court turned to the applicable Washington Rule of Evidence, Rule 701,³⁵ which governs the state’s admissibility of opinions from lay witnesses (those who are not experts). According to Rule 701, “A lay witness may give an opinion, so long as it is rationally based on her perceptions and helpful to the jury.”³⁶ Mrs. Cardell testified based on the voices she heard the night she was attacked, as well as her years interacting with black people when she lived in the southern United States.³⁷

The court noted that trial judges have broad discretion pertaining to the admissibility of lay opinions,³⁸ citing other state courts in North Carolina and Missouri that had allowed lay witnesses to testify about whether people they could not see “sounded black.”³⁹ As for the opinion of Mrs. Cardell the court who was born in New Zealand, the court found that she had sufficient experience in determining whether someone “sounded black” based on her years of experience interacting with African Americans in the American South. And though she never specifically identified Mr. Kinard, the defendant, as her attacker, her testimony was found by the court to be helpful for the jury because it helped establish the attacker’s identity.^{40 41}

In *Clifford v. Commonwealth*, which was decided by the Kentucky Supreme Court,⁴² the defendant, a black man, Mr. Charles Clifford, was convicted of drug trafficking.⁴³ Mr. Clifford was arrested in connection to a drug purchase that was set up by the police, with an undercover police officer wearing a wire throughout the transaction.⁴⁴ Later, Officer Darin Smith listened to the conversation of the people involved in the transaction.⁴⁵ At trial, Officer Smith testified that, based on his thirteen years as a police officer interacting with black people, he heard someone who sounded like a “male black” offer the undercover officer crack cocaine.^{46 47 48}

Upon his conviction, Mr. Clifford appealed to the Kentucky Supreme Court, arguing that the trial judge should not have allowed Officer Smith to give his opinion about the race of the speaker he overheard.⁴⁹ The Kentucky Supreme Court began its analysis on this question with

³¹ *Id.* at 604.

³² *Id.*

³³ *Id.*

³⁴ *Id.*

³⁵ *Id.* at 605.

³⁶ *Id.* citing Washington Rule of Evidence 701.

³⁷ *Id.* at 604

³⁸ *Id.* at 605

³⁹ *Id.* citing *State v. Phillips*, 212 S.E.2d 172 (N.C. 1975) and *State v. McDaniel*, 392 S.W.2d 310 (Mo. 1965).

⁴⁰ *Id.* at 604–05.

⁴¹ *Id.*

⁴² 7 S.W.3d 371 (Ky. 1999).

⁴³ *Id.* at 373.

⁴⁴ *Id.*

⁴⁵ *Id.*

⁴⁶ *Id.*

⁴⁷ *Id.*

⁴⁸ *Id.*

⁴⁹ *Id.* at 374.

Kentucky Rule of 701. Like Washington Rule of Evidence 701,⁵⁰ Kentucky Rule of Evidence 701 allows a lay witness to “express an opinion which is rationally based on the perception of the witness and helpful to a determination of a fact in issue.”⁵¹ The court also noted a corollary to this rule, the “collective facts rule,” “which permits a lay witness to resort to a conclusion or an opinion to describe an observed phenomenon where there exists no other feasible alternative by which to communicate that observation.”⁵²

While the Kentucky Supreme Court acknowledged that linguistic profiling was a new issue in Kentucky, it noted that many other state court decisions—including *Kinard*, *Phillips*, and *McDaniel*—had allowed witnesses to testify about the perceived race of speakers they heard but could not necessarily see.⁵³ The *Clifford* court cited *Sanchez*, which found that witnesses can testify about dialects and accents they are familiar with.⁵⁴ Ultimately deciding that the requirements of Kentucky Rule of Evidence 701 were met because Officer Smith’s thirteen years of experience with black people qualified him to identify an African American dialect.⁵⁵

Finally, in *People v. Poole*, the defendant, Mr. Cedrick Poole, a black man, was convicted in California of crimes against a woman, “N.Y.”⁵⁶ N.Y. was a 31-year-old college student studying linguistics, a Japanese national, a resident of the United States for only two years before the attack, and legally blind.⁵⁷ N.Y. was approached one evening by a man she later identified by voice as Mr. Poole, who groped her, threw her to the ground, and punched her.⁵⁸

Upon his conviction, Poole appealed to the California Court of Appeal, arguing that N.Y.’s identification of was unreliable,⁵⁹ because, he claimed, N.Y. did not have sufficient experience with African American dialects to be able to say whether a person sounded like an African American.⁶⁰ The Court of Appeal was unconvinced.⁶¹ N.Y. was a linguistics student, had been blind since birth and was therefore better at identifying voices, and also had an African-American boyfriend with whom she interacted daily.⁶² For the California Court of Appeal, this was sufficient experience for a non-native English speaker and foreign national to be able to say that someone “sounded” black.

Courts’ willingness to allow witnesses with sufficient experience to testify that someone “sounded black,” is distinct from allowing testimony that says that a person “acted black,” a distinction made in a federal case, *United States v. Card*.⁶³ Witnesses wished to testify that the robbers in the case “talked like” and “acted like” African Americans.⁶⁴ Though the *Card* court

⁵⁰ *Id.* citing Kentucky Rule of Evidence 701. Many states follow examples of the Federal Rules of Evidence, which also has a Rule 701, which is the basis for the Washington and Kentucky rules.

⁵¹ *Id.*

⁵² *Id.*

⁵³ *Id.*

⁵⁴ *Id.*

⁵⁵ *Id.* at 376.

⁵⁶ 2001 WL 1526968, *1 (Cal. App. 2001) (unpublished).

⁵⁷ *Id.* at *1, 3.

⁵⁸ *Id.* at *1.

⁵⁹ *Id.* *3.

⁶⁰ *Id.* at *3–4.

⁶¹ *Id.* at *3.

⁶² *Id.*

⁶³ 86 F.Supp.2d 1115 (D. Utah 2000).

⁶⁴ *Id.*

was not definitive about testimony that the robbers “acted like” African Americans, they accepted that witnesses can, with proper foundation, testify about the dialect of a speaker.⁶⁵

4.2 Civil Cases

In civil cases linguistic profiling is accepted; courts in civil anti-discrimination lawsuits accept a person can tell the race of another person based on speech. However, they consider it illegal to discriminate based on linguistic profiling, a violation of either federal or state anti-discrimination laws. This is illustrated in the district court pretrial opinion in *United States v. Habersham Properties*, a civil lawsuit brought by the federal government to enforce the Fair Housing Act in the suburbs of Atlanta.⁶⁶

In early 2001, in Decatur, Georgia, a black woman named Lynda Osbourne telephoned to get an apartment at Crescent Court Apartments, which was run by Habersham,⁶⁷ speaking over the phone in a British accent.⁶⁸ Though she was told there was an apartment,⁶⁹ when she arrived, she was told there was no apartment available.⁷⁰ Osbourne complained to the federal government, which used both black and white testers, who called and then visited to view apartments.⁷¹ In most cases, the white callers and visitors were told that apartments were available while the black callers and visitors were told the opposite.⁷² The federal government sued, alleging violations of the federal Fair Housing Act,⁷³ which prohibits a landlord from engaging in a “pattern” of racial discrimination in leasing apartments.⁷⁴ Though it left the question about whether there was a pattern of discrimination for a jury, and the case settled before the trial went on trial,⁷⁵ the federal district court in *Habersham* did find enough evidence to show discrimination, accepting that discrimination based on speech was a kind of racial discrimination. Recall our discussion in Section 1: this case illustrates John Baugh’s experiments playing out in a court of law.

4.3 When Linguistic Profiling Affects the Court

In *State of Florida v. George Zimmerman*, defendant George Zimmerman was charged with the murder of a 17-year-old black male, Trayvon Martin. Zimmerman would later be acquitted on these charges due to insufficient evidence, a result that was met with backlash from across the nation, especially from those who believed that Martin’s killing was racially motivated.

In addition to being widely publicized, the case is remarkable because it highlights an instance of linguistic profiling occurring inside the courtroom itself. During the trial, the defense produced their star witness: Rachel Jeantel. Jeantel was on the phone with Martin during his encounter with George Zimmerman, and her testimony, which went on for nearly six hours, longer than any other single witness at the trial, may have held the key to discovering what

⁶⁵ *Id.* at 1118.

⁶⁶ 319 F. Supp. 2d 1366 (N.D. Geo. 2003).

⁶⁷ *Id.* at 1369.

⁶⁸ *Id.*

⁶⁹ *Id.*

⁷⁰ *Id.*

⁷¹ *Id.* at 1370.

⁷² *Id.*

⁷³ *Id.* at 1371. *See also* 42 U.S.C. § 3604, *and* Int’l Bhd. of Teamsters v. United States, 431 U.S. 324, 336 (1977).

⁷⁴ *Id.* at 1372.

Id. at 1373–75.

See Consent Decr

⁷⁵ *Id.*

transpired the night of the attack. However, according to one juror, ('no one mentioned Jeantel in [16+ hour] jury deliberations. Her testimony played no role whatsoever in their decision' (Juror Maddy, as reported in Rickford & King, citing Bloom 2014:148)).

Below is an excerpt of her testimony.⁷⁶ As discussed above in section 3.2.1, in AAE, certain forms of the verb "be" and agreement markers on plural and possessive nouns and verbs are regularly deleted. These deletion sites are marked by "∅" in the excerpt of Jeantel's testimony in (6), below.

- (6) Excerpt from the Courtroom Testimony of Rachel Jeantel (RJ), Day 1
Prosecutor Bernie de la Rionda (BR) questioning, as recorded by the court reporter (CR)
and annotated by us
[= Rickford and King]

∅ = deletions

There were four types of deletions: (a) **is/are** verb forms; (b) **'s** possessive noun marker; (c) **-s** noun plural marker; and (d) **-s** third person singular present tense marker

RJ: He said he ∅ from—he—I asked him where he ∅ at. An he told me he ∅ at the back of his daddy∅ fiancée∅ house, like in the area where his daddy fiancée—BY his daddy∅ fiancée∅ house. Like—I said, 'Oh, you better keep running.' He said, naw, he lost him.

BR: Okay. Let me stop you a second. This—this lady [the Court Reporter] has got to take everything down, so you make sure you're—Okay. So after he said he lost him, what happened then?

RJ: And he say he—he ∅by—um—the area that his daddy∅ house is, his daddy∅ fiancée∅ house is, and I told him 'Keep running.' He—and he said, 'Naw,' he'll just walk faster. I'm like, 'Oh oh.' And I—I ain't complain, 'cause he was breathing hard, so I understand why. Soo

BR: What—what happened after that?

RJ: And then, second∅ later—ah—Trayvon come and say 'Oh, shit!'

CR: [Unintelligible—requesting clarification] 'Second later?'

RJ: A couple second∅ later, Trayvon come and say, 'Oh, shit!'

BR: Okay. Let me interrupt you a second. When you say the words, 'Oh, shit,' pardon my language, who said that?

RJ: Trayvon.

BR: He said it to YOU?

RJ: Yes.

BR: Okay. And after he used, pardon my language, he said, 'Oh, shit,' what happened then?

RJ: The nigga ∅ behind me.

CR: I'm sorry, what? (22:7–23:7)

RJ: [Slowly, deliberately] The nigga's behind—the nigga ∅ behind me.

BR: Okay. He used the N word again and said the nigger is behind me?

⁷⁶ John Rickford and Sharese King (2016) "Language and Linguistics on Trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond" *Language* 92(4), 958-988.

Because of the way she spoke, Jeantel's speech was often misinterpreted, and her credibility was attacked.⁷⁷ As such, the courtroom failed to understand or respect Jeantel's speech, which was not a butchering of the English language, but a systematic, rule governed, way of speaking shared by many people in the black community.

Lisa Bloom, a lawyer who witnessed Jeantel's testimony asserted "her grammar and diction could be hard to interpret... in addition, Jeantel spoke using urban teenaged lingo that was an alien tongue to most of the white, suburban, middle-aged jurors." Jeantel, as a result, was ridiculed outside of the courtroom.

Jury composition matters.⁷⁸ Rickford notes that "African Americans on the jury – especially fluent AAE speakers – would have understood Jeantel." He also points out that any juror who spoke AAE could ensure that the other jurors understood what she was saying. The jury pool in the case was homogenous, composed of white jurors. In the state of Florida, all-white jury pools convict black defendants 16% more often than their counterparts, a gap that disappears with the inclusion of a single black jury member.

Beyond the jurors' comprehension problems were misunderstandings of some of her expressions.⁷⁹ For example, Jeantel referred to Zimmerman as a "creepy-ass cracka", and frequently used the word "nigga." But Jeantel explained that her use of "nigga" was not to be interpreted as describing a black person, but instead to describe any person, regardless of race or ethnicity. But the jury ignored her remarks and found this language off-putting; Maddy, a Puerto Rican juror, stated that "All the other jurors... were offended by 'creepy-ass cracka' and they were done with Jeantel once they heard that."

When jurors commented on how the jurors reacted to Jeantel's testimony during their deliberations, juror B37, said that over the course of the jury deliberations, Jeantel was not mentioned even once. Because people attached preconceived prejudices to Jeantel's dialect, her testimony - which should have been a key factor in the jurors' deliberations - was completely ignored. Had her testimony been accepted by the courtroom, its contents could have changed the result of the trial.

5. Discussion

This article has investigated "linguistic profiling" from a variety of perspectives. We began by distinguishing it from the more familiar term, "racial profiling," showing that "linguistic profiling" is not inherently discriminatory. Thinking that someone "sounds like" they come from a French-speaking country doesn't entail making a negative judgement about that speaker. Nevertheless, as we have seen, our tendency to engage in linguistic profiling - to use auditory cues to identify social characteristics, such as race, gender, sexual orientation, or geographic origin - can *lead to* discrimination. We have seen legal cases - both civil and criminal - that have engaged linguistic profiling in various ways.

Our exploration of the psychological research showed that the human brain automatically sifts through and categorizes all the input we get every day. Linguistic profiling is one result of this process. As a cognitive process, these human mental "shortcuts" to organize and classify our

⁷⁷ John Rickford and Sharese King (2016) "Language and Linguistics on Trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond" *Language* 92(4), 958-988.

⁷⁸ Shamera Anwar, Patrick Bayer, & Randi Hjalmarsson, The Impact of Jury Race in Criminal Trials. *Quarterly Journal of Economics*, 127(2), 1017-1055 (2010).

⁷⁹ Grace Sullivan, Problematizing Minority Voices: Intertextuality and Ideology in the Court Reporter's Representation of Rachel Jeantel's Voice in the State of Florida v. George Zimmerman Murder Trial (2017).

surroundings are not inherently good or bad but useful tools. They only become problematic when we infuse these linguistic categories this process creates with biases.

As we explored the interaction of linguistic profiling with legal cases, we have seen both civil and criminal cases where it has played a role, in a variety of ways. First, it became clear to us that in criminal cases, our brain's tendency to categorize speech can help establish a suspect's identity. In many cases, the court has ruled that testimony based on linguistic profiling does not require expert testimony but that it does require sufficient experience to testify that someone "sounded black" or "sounded Puerto Rican." Various cases have defined "sufficient experience" differently. In *People v. Sanchez*, the witness had spoken Spanish with Dominicans and Puerto Ricans his whole life and was a member of the speech community. The court deemed that sufficient to identify speakers as Puerto Rican or Dominican. In *Clifford v. Commonwealth*, the witness had interacted with AAE speakers for his ten-year career as a police officer and the court ruled that this was sufficient to identify a speaker as black. While the individual judge may be responsible for deciding what constitutes "sufficient experience," it seems clear that discrimination is not allowed when using linguistic profiling. In *United States v. Card*, a witness was able to describe the speaker as "sounding black" but could not testify that someone "acted black" or that someone who "sounded black" was more likely to have committed a crime (implying that they were more likely to have committed the crime because they were black). In the civil case we looked at, *United States v. Habersham Properties*, we saw that linguistic profiling could fall under purview of the law when it led to discrimination, such as when Habersham Properties would not book appointments to view apartments to people who "sounded black" over the phone. In another criminal trial, *Florida v. Zimmerman*, we saw the effect of linguistic profiling in the courtroom. Rachel Jeantel, a young African American woman speaking with an AAE accent, had her testimony disregarded because people attached negative stereotypes to the way she spoke.

What does this mean for the legal system? Linguistic profiling can be used as a tool in courts. As many courts have already recognized, testimony based on linguistic profiling is helpful in establishing the issue of identity. Of course, courts must be careful to discourage discrimination in this realm: the way someone speaks and their race do not make them more likely to have committed a crime. Courts also need to recognize that linguistic profiling leads to discrimination but is not discrimination. Linguistic profiling has real and illegal impact on people's lives when it comes to housing, employment, and other areas where discrimination is punishable by law. Finally, courts must be careful to keep discriminatory linguistic profiling from affecting jury decisions and the weight of witness testimony. A key witness testimony cannot be allowed to be disregarded because the dialect used by the witness is non-standard and a jury or judge attaches a prejudice to that dialect. As we have seen, our tendency to categorize makes linguistic profiling a natural process but all dialects are rule-driven and no language is inherently better or worse than another language. Linguistic profiling plays a role in court cases all the time and it is imperative that the courts be aware in order to curtail discrimination based on linguistic profiling.