2022

AMBIGUOUS APPALACHIANNESS: A LINGUISTIC AND PERCEPTUAL INVESTIGATION INTO ARC-LABELED PENNSYLVANIA COUNTIES

Crissandra J. George
University of Kentucky, crissandrajg@gmail.com
Author ORCID Identifier:
https://orcid.org/0000-0001-5223-0156
Digital Object Identifier: https://doi.org/10.13023/etd.2022.217

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Crissandra J. George, Student

Dr. Allison Burkette, Major Professor

Dr. Allison Burkette, Director of Graduate Studies

AMBIGUOUS APPALACHIANNESS: A LINGUISTIC AND PERCEPTUAL INVESTIGATION INTO ARC-LABELED PENNSYLVANIA COUNTIES

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Arts in the College of Arts and Sciences
at the University of Kentucky

By
Crissandra George
Lexington, Kentucky
Director: Dr. Allison Burkette, Professor of Linguistics
Lexington, Kentucky
2022

https://orcid.org/0000-0001-5223-0156

ABSTRACT OF THESIS

AMBIGUOUS APPALACHIANNESS: A LINGUISTIC AND PERCEPTUAL INVESTIGATION INTO ARC-LABELED PENNSYLVANIA COUNTIES

The Appalachian Regional Commission (2022) designates 52 of Pennsylvania's 67 counties as Appalachia, excluding only the southeast portion of the state. Matthew Ferrence, in *Appalachia North*, states that his "home is sometimes called Appalachia, sometimes Rust Belt, other times Midwest, even though very few who live there would accept any of those labels as correct" (xi). This ambiguous and fluid identity is due to the shaping, forming, and changing of Pennsylvania's role within society from a founding colony to a thriving state with industry, unselfishly spoiling others, to the grounds of converging identities (Ferrence xi). This ambiguous identity makes the voice of Northern Appalachian speakers difficult to capture. Watt and Llamas (2017: 193) note that place is not just a location, but rather "states of mind, stances, attitudes, and the status that individuals hold within their social networks and society at large." Historically, and even currently, stereotyping and defining these Appalachian regions has come from "outsiders" or "spectators" within society that continue add dynamic and fluid definitions that vary depending on a multitude of contexts (Ulack and Raitz 1982). Both language use and language perception play a big part "in how territories bounded by borders with their neighbors are defined" (Watt and Llamas 2017:191). By looking at language and perceptual excerpts from the Linguistic Atlas Project and present-day interviews with Northern Appalachian speakers themselves, one can compare these linguistic patterns with other patterns studied in Appalachian Englishes and investigate the identities of these speakers to understand where Pennsylvania fits into the region that is Appalachia, giving writers, researchers, and society voices and identities to capture.

KEYWORDS: Appalachian Englishes, Northern Appalachia, Linguistic Atlas Project (LAP), place-based identity, perceptual dialectology, sociolinguistics

Crissandra George

04/15/2022

AMBIGUOUS APPALACHIANNESS: A LINGUISTIC AND PERCEPTUAL INVESTIGATION INTO ARC-LABELED PENNSYLVANIA COUNTIES

By
Crissandra George

Dr. Allison Burkette
Director of Thesis

Dr. Allison Burkette
Director of Graduate Studies

04/15/2022
Date

DEDICATION

To the mountains, hills, valleys, ridges, hollers, rivers, creeks, cricks, streams, and runs of Pennsylvania that brought me countless memories, taught me life's greatest lessons, and continue to inspire me.

ACKNOWLEDGMENTS

Without the insight and direction of several people, this thesis would have only been merely an idea, never acted upon. First, my Thesis Chair, Dr. Allison Burkette, exemplifies the high-quality scholarship and passion for language to which I aspire and has always pushed me to achieve anything I put my mind to with endless support. Dr. Allison Burkette and The Linguistic Atlas Project have always fostered a personal environment of critical inquiry and academic exploration for which I am grateful to have been exposed to and given the honor to work with for several years. In addition, Dr. Jennifer Cramer, a member of my Thesis committee, continued to motivate my work through instructive comments, timely evaluation, and supportive feedback within the thesis and in my academic career. Dr. Josef Fruehwald, another member of my Thesis committee, provided insights that challenged my thinking and improved my research abilities. Additionally, I wish to thank Dr. Mark Lauersdorf, who believed in my research and abilities more than I even did myself. Lastly, I am grateful for the countless Northern Appalachian researchers that I shared my research with and reaffirmed the rewarding nature of this field and inspiring scholars that occupy it.

In addition to the technical and instrumental assistance above, I received equally important assistance from family, friends, and colleagues. My mother, father, and grandfather provided on-going support throughout the thesis process. I am eternally grateful for their love and support and without them, I would not be who I am today. Finally, I wish to thank the respondents of my study (who remain anonymous for confidentiality purposes).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 Appalachian Pennsylvania

Small creeks flow into larger streams and rivers creating an interconnected, interdependent network of individual elements resulting in something much bigger than the sum of its parts. Histories and cultures follow this pattern. Historical events flow into bigger trends and movements all shaping a region's identity, within which a network of individual identities is embedded. The linguistic identity of Pennsylvania (PA) is shaped and formed similarly. Once a founding colony to a thriving state with industry, unselfishly spoiling others with natural resources, to the grounds of converging identities. Matthew Ferrence (2019), in *Appalachia North*, states that his "home is sometimes called Appalachia, sometimes Rust Belt, other times Midwest, even though very few who live there would accept any of those labels as correct" (xi). This ambiguous “fluid” identity is due to the shaping, forming, and changing of Pennsylvania’s role within society (Ferrence, 2019, xi). This ambiguous identity makes the voice of Northern Appalachian speakers difficult to capture. This investigation aims at exploring this lesser known and studied voices and identities in a deeper setting and seeing the convergence of language and place within Appalachian Regional Commission (ARC) counties in PA and how these variables effect the perception of PA speakers within their own communities.

This study will review previous research conducted within this region as well as the various definitions placed in Appalachian regions and PA regions, giving sociocultural context to the geographic locations being examined. Along with giving a sociocultural background, this investigation will also cover these regions through a

linguistic and perceptual lens prior to giving description of the methodology used to conduct this comparative investigation.

1.2 Overarching Goal of the Investigation

The overarching goal of this investigation is to compare linguistic patterns, commonly indexed as Appalachian and present in historical survey-style interviews within PA, from the Linguistic Atlas Project (LAP) to current survey-style interviews from PA speakers. The LAP data will serve as a benchmark of comparison for production patterns within these regions. Along with investigating production (phonetic, morpho-syntactic, lexical) patterns, I will be comparing perceptions of regional identity across present-day speakers. This historical and present-day production and perception data will bring light to the ambiguous and fluid regional identities of ARC counties in PA, giving all languages and identities to authentically capture.

1.3 Research Questions

1. Do LAP and present-day PA speakers, in ARC labeled counties, follow linguistic production patterns commonly indexed in Appalachian Englishes?
2. Do present-day speakers, in ARC labeled counties, regionally identify with Appalachia?
3. Based on this data, where does PA fit into the region of Appalachia?

1.4 Overview of the Investigation

Through this methodology and research questions, one will see the presence of Appalachian Englishes features present in both historical and current PA speakers. In addition to the presence of these linguistic features, regional identities of the modern-day Northern Appalachians vary depending on a multitude of contexts related to geography, sociocultural variables, language, and perceptions of the speakers themselves as well as outsider perceptions related to stereotyping. Through the combination of the linguistic and perceptual data, speakers use detailed discursive mapping to create boundaries and identify themselves regionally. Some of these regional identities include Appalachia and some include Appalachia in addition to another regional label while others do not include Appalachia at all with speakers struggling to navigate the ambiguity associated with their regional identity. Language and perception come together within this investigation to show these ambiguous and fluid identities that further bring insight into who and where this ambiguity is attached to whether it be the region, the speakers, or if this ambiguity is a regional identity within itself, allowing a further discussion into whether this ambiguity is due to multiple overlapping identities or an absence of identities. While this study is not extensive and large enough to fully answer these complex questions surrounding regional identity, it will demonstrate the diverse, dynamic, and heterogenous entity that is Northern Appalachia. Through analyzing the language and regional perceptions of historical and current speakers in PA Appalachian counties, one will take part in the beginning steps of researching and understanding these understudied regions of Appalachia.

CHAPTER 2. LITERATURE REVIEW

2.1 Definitions of Appalachia

Appalachia is a contested term due to the history of varying definitions. Prior to the twentieth century, scholarship in Appalachia was primarily focused on ethnography and folklore (Drake, 2001, p.129). Geographers, geologists, and natural science researchers had long studied the region, but there was little investigation into the culture and lifestyle besides nineteenth century fictional literary depictions of the people and region. These literary depictions brought a definition of *Appalachia* that was based on stereotypes and caricatures, even though they were intended for entertainment (Lewis, 1999, p. 21). When studies then began exploring Appalachian culture, many new definitions of the term continued to be invented and discovered. The history of the term itself, in everyday life, is still fairly new and there is very little formal history of the term used within academia. Due to this, Appalachia gets different definitions based on many different criteria and contexts and "it seems to really depend on who you ask" (Hasty, 2020, p.6). Some regions lack these varying definitions and are often "difficult to delimit" while definitions of Appalachia are "are quite vivid and have been much studied yet lack definitive boundaries" (Ulack and Raitz, 1981, p. 40). With this lack of definitive boundaries, comes many interpretations of what Appalachia is depending on if you are "talking about Appalachia geographically, politically, socially, perceptually, or linguistically" (Hasty, 2020, p.7). Through these varying definitions, the regional term, *Appalachia*, and meanings attached to it are diverse and have a variety of contexts and connotations embedded within them that require analysis.

Geographically, Appalachia is the area within the boundaries from Northern Alabama to Newfoundland, but this was later adjusted to only include the "portion south of the Hudson and Mohawk valleys in New York" (Ulack and Raitz: 1981, p. 40). This physiographical abbreviation was later adjusted again that excluded "Appalachia north of the Maryland-Pennsylvania border (Mason-Dixon Line)" (Ulack and Raitz: 1981, p. 41). While geographic definitions of this region have shifted, these definitions still only show landforms and geospatial information which lacks the information to investigate social and cultural variables that further connect geography to society.

Along with changing geographic definitions, Appalachia has also received many definitions based on socio-economic, political, and educational values by the Appalachian Regional Commission (ARC). The ARC was created to address and improve socioeconomic issues that persistently flooded this region. From the ARC, sociopolitical mappings were created to implement economic development programs within Appalachian regions. These ARC designated regions within the United States were said to be "abundant in natural resources and rich in potential" but lag "behind the rest of the nation in its economic growth and that its people have not shared properly in the Nation's prosperity" (House of Representatives, 2007, p. 187). The ARC definition of Appalachia, and labeling of Appalachian sub-regions, has gone through multiple revisions. The most recent definition of the region, based on socioeconomic variables, expands across 13 state governments including 423 counties ranging from New York to Mississippi (Appalachian Regional Commission, 2021). Figure 1 contains the most recent ARC map defining the subregions: Northern, North Central, Central, South Central, and

Southern. While the ARC map includes societal variables, it is still subject to continued adjustment and lacks in certain contextual elements.

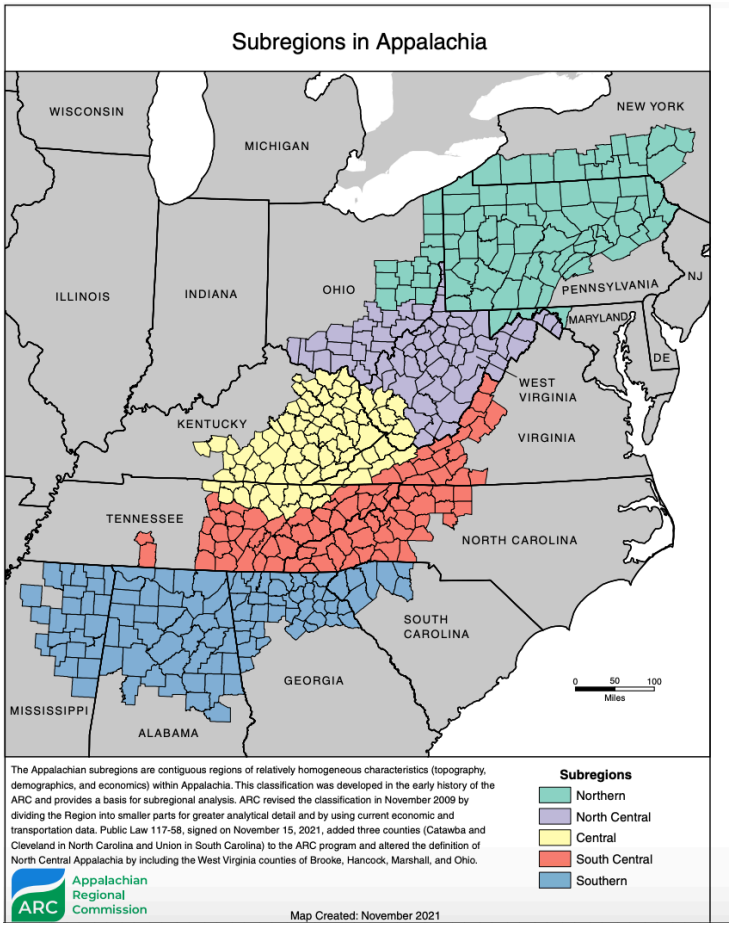

Figure 1. Subregions of Appalachia (Appalachian Regional Commission, 2021).

The ARC map, as well as many other mappings of Appalachia, do not showcase the "dramatic contrasts in landforms, land use, and resource potential" along with the diversity in culture and economy (Ulack and Raitz, 1982, p. 727). McCann (1998) brings to light that socioeconomic mapping of regions, such as the ARC, have their limitations and "breed stereotypes that come back to haunt local communities" (p. 91). This is due to the quick snapshot nature of these visuals and that they are created for utility and a specific purpose, not for accuracy in all contexts.

A solution to approaching a representation of Appalachia for accuracy in all contexts was thought to be cognitive mapping, but this has its limitations as well. Cognitive mapping is categorized cognitive information put into the form of a graphic and this methodology can show the relationship between people and places (Raitz and Ulack, 1981, p. 201). The categorized cognitive information is derived from participants' perceptions of the region and perceptions of residents within the region that is compiled and then visualized in the form of geo-spatial map. Previous studies have included cognitive mapping methods based on "insiders", "outsiders", and "cognitive outsiders" to the region of Appalachia (Ulack and Raitz, 1981, p. 750), which bring representations of the region and its people from three different perspectives. Insiders are participants within the study who reside in Appalachia and are aware that they live in this ARC designated region. Outsiders are participants who live outside of Appalachia. Cognitive outsiders are participants who live in Appalachia, but who are not aware that they reside within the ARC boundaries. These participant groups represent diverse perceptions of the region from the viewpoint of three distinct regional identities. This study and Ulack and Raitz (1982) further show a variety of participants' perceptions that differ in viewing Appalachia as a social place (their stances and opinions surrounding society and culture in Appalachia) and a physical place (the boundaries that respondents place and label as a part of the region). Within their study, Ulack and Raitz's respondents mapped only 1/3 of area designated by the ARC boundaries and only 10% of the respondents agreed with the ARC boundaries (Ulack and Raitz, 1982, p. 51). One distinction noted within this study is that outsiders expanded the south-central region only a bit wider than insiders, demonstrating the multitude of outsider and insider views on the perceived location of the

region (Ulack and Raitz, 1982, p. 51). Through this investigation, there is a contested definition on the location of Appalachia that varies between different participant groups, but this methodology, at the very least, includes the perceptions of Appalachians themselves. While cognitive mapping includes the voices and perceptions of residents and non-residents, they still are affected and influenced by stereotypes and portrayal of the region, but this does not take away the vitality of understanding socio-cultural perceptions when regionalizing Appalachia. Through geographical, socioeconomic, and cognitive mapping, Appalachia has been represented in a variety of ways creating an image of Appalachians that "is confused and largely negative" (Ulack and Raitz, 1982, p.727), but this clearly popularized image of a region still lacks clearly defined boundaries.

Boundary making is inherent to humans and society. They help people make sense of society and the world around them. Specific perceptions get associated with boundaries, resulting in attitudes, ideologies, and perceptions tied to these boundaries, creating a regional identity. These perceptions and identities can shift and change depending on people and society, which seems to be the case with Appalachia. Boundaries are loose in some contexts, but then rigid and restricted in others. This is exemplified within regional boundaries where counties, cities, and states all have differing views of the locations around them and even themselves. Through this boundary making, identities and representations are created, which constructs an individual's and even larger group's mapping of their reality. These constructed realities, whether they be from insiders or outsiders, can be the source of various characterizations, representations, and stereotypes.

2.2 Appalachia through a Stereotypical Lens

Affecting these representations seen in various mapping approaches is the perceived image of Appalachia and the region's residents. Graphical and mental representations have not been the only methods that influence the perceived image of Appalachia. Many media and textual representations have resulted in extreme characterizations and negative stereotypes of the region that historically have been present within society in the United States. Ferrence (2012) states that "Appalachia, in particular, cannot find its way out of punchlines" (p. 114) and highlights that "individuals living within the region are offered little agency; their lives are assumed to be lived in accordance with the expected stereotypes" (p. 119). These expected stereotypes depict low socioeconomic status, poor education, and a backwards way of living. These depictions are further extended through literature, media, and popular culture creating this idea of Appalachia and Appalachians as a homogenous identity, but this in no way encompasses the reality of this diverse and complex region.

2.3 Appalachia through a Linguistic Lens

The perceived homogenous identity of the region extends into the perception of the language in Appalachia, but again it is not the reality. While many believe the language of Appalachia is old-fashioned, indicative of low education, and informal, Montgomery (2013) states that Appalachia "does not have just a single dialect" (p.29), but rather it is linguistically diverse containing a variety of linguistic patterns. This lack of homogeneity is what brings to light the term "Appalachian Englishes" (Hazen and Fluharty, 2004) and is meant to convey the complex nature of the language varieties itself and reinforces that it

is not a "single entity" (Cramer, 2018, p.46). Appalachian Englishes are highly stigmatized, just as the region and its people are, and some features in the language remain marked for social variables of inferiority, low education status, low social class, and lack of pleasantness. These linguistic perceptions are seen in multiple previous studies showcasing the stereotyping and influence of stereotypes associated with these varieties (see, for example, Luhman 1990 and Cramer 2012, 2013, 2014, 2018). The perceptions of these nonstandard marked features range "from comical and uneducated to demeaning and isolated" (Cramer, 2018, p. 57). Just as the region and people are heavily stereotyped, so are the varieties of Appalachian Englishes. While not all features seen in Appalachian Englishes (see Table 1) are markers for these harsh stereotypes, the features still are highly marked indicating perceived group membership.

Table 1. Appalachian Englishes Phonological (Reed 2020) and Morphosyntactic Features (Hazen 2020)

Phonological Features	**Examples**
/w/ /ʍ/ distinction	Difference between “which” and “witch
/l/ /ɬ/ distinction	Sounds like [beli] instead of [beɬi] for the term “belly”
/z/ changes to /d/ in word middle and word-final	business to “bidness”
intial “h” sound	“ain’t” to “hain’t”
consonant deletions	“there” to “ ’ere”
/ai/ ungliding	“bite” and “price” is monopthongized
rotation of front vowels	[bɪt] for "bit" [bɪt] for "beet"
fronting of back vowels	[but] sounds like [biewt]
vowel breaking	[bæd] sounds like [bæId]
pin/pen merger before nasals	“pin” and “pen” sound the same
cot/caught merger	“cot” and “caught” sound the same
rooted rise in pitch without cause/emphasis	“MIKE left EARLY”
Morphosyntactic Features	**Example**
a-prefixing	"a-fishing"
multiple negation	"he didn't have no common sense"
was-leveling	"The rumors was starting down there"
demonstrative them	"Them dogs are running"
reflexive regularization	"hisself" and "theirselves"
Second person plural	"y'all" "you guys" "you ones" "yinz"
[-ing] variation	"goin" (less formal) and "going" (more formal)
Quotative “be like”	"The dog is like a baby"
non-standard past tense	"I knowed about what they were doing."
double modals	"I might could do that" and "we are fixin to go to the mall"
verb construction	"The cats need fed" or "the clothes need washed"

2.4 Appalachia through the Lens of Regional Identity Construction

Due to Appalachian Englishes being heavily marked as non-standard varieties, the linguistic features index an immediately identifiable group membership. This indexicality of Appalachian Englishes comes from the context dependent associated information (identity, gender, region, socio-cultural variables, and more) attached to language utterances, such as regional varieties being connected to the identity of the speaker (Hanks, 1999, p.124). Indexing people or even a region through pronunciations, grammatical structures, words, and writing systems is language being "used as a proxy for place" (Watts and Llamas, 2017, p. 194) which shows the embedded meaning within language. This also shows how language is closely related to identity and how "linguistic variation is a meaningful and significant symbol of group membership" (Tamasi and Antieau, 2014, p. 20). Through variation, these meaningful connections create complex and fluid identities. Language variation, indexed speech patterns, and identity are not singular within individual speakers, which is why it is important to "recognize that speakers have not one, but many identities that be reflected within their speech" (Tamasi and Antieau, 2014, p. 20). Society, literature, and popular culture continually perpetuate that the Appalachian identity is singular and do not acknowledge the heterogenous entity that is Appalachia and the speakers within the region. Monolithic Appalachian stereotypes, perpetuated through outsiders looking into the region have "been the object of curiosity and even romanticization, perhaps mostly because of its old-fashioned flavor and its colorful and seemingly quaint usages" (Montgomery, 2007, p.42). Some of these romanticized features and sociocultural portrayals in literature then further create an image of Appalachia constructed by writers that are "outsiders" and create a construction

of the area that to others may seem accurate, even when it is not. Ferrence (2019) comments on this romanticization of Appalachia in literature and states that much known "about Appalachia is built from writers" and that Appalachia is "as much a literary construction as it is anything else" (p.179). While some of these writers' intentions were probably not ill-willed, they continue to create misconceptions that the languages, and even culture, of Appalachia are stagnant, when in fact these languages and cultures are "exhibiting features in various stages of change" (Montgomery, 2007, p. 43). These literary misconceptions further increase the stereotypes and misrepresentations of Appalachia portraying the region, people, and languages as a singular stagnant entity with no regard to the diversity of identities and perceptions that can lie in a region comprised of many people.

Appalachia and the speakers in the region, containing a variety individual languages, identities, and perceptions, create an established connection between language and place. This is exemplified with the features in Appalachian Englishes that "give social and regional identity and cultural cohesion" (Montgomery, 2007, p. 43). Due to Appalachian Englishes being heavily indexed for regional and social cohesion, speakers recognize this and can indicate this close natured connection as seen in Cramer (2020), which states that "Appalachians recognize that their speech reveals an attachment to culture, heritage, home, and family, and they perceive their speech to be pleasant and beautiful" (p. 80). Appalachians acknowledge and are aware of the negative stereotypes, seen in society and literature, showing that place is not just a location, but rather "states of mind, stances, attitudes, and the status that individuals hold within their social networks and society at large" (Watt and Llamas, 2017, p. 193). This acknowledgment

also shows the deeply embedded connection between language and place that plays a big part "in how territories bounded by borders with their neighbors are defined" (Watt and Llamas, 2017, p. 191). This metalinguistic and perceptual information is vital to understanding this region and the subregions within it. Without looking into how Appalachians perceive their languages and identities as well as the knowledge they have of outside perceptions, one will not achieve a deepened understanding the region itself and the varying perceptions, ideologies, and attitudes within it.

2.5 ARC Counties in PA

Just as *Appalachia* has varying definitions, contexts, perceptions, and connotations that change based on geography, socio-cultural variables, and outsider perceptions, so do the subregions of Appalachia. PA is part of the subregion labeled by the ARC as Northern Appalachia, but these does not include the whole state (see Figure 2). Northern Appalachia is only 52 out of 67 counties within PA (Appalachian Regional Commission, 2021). How do these diverse definitions, languages, and perceptions of Appalachia apply to these ARC counties in PA, a state only partially within this region? By looking into Appalachian PA through different lenses, such as geographically, culturally, linguistically, and perceptually, one can achieve a deepened understanding of this Northern Appalachian region.

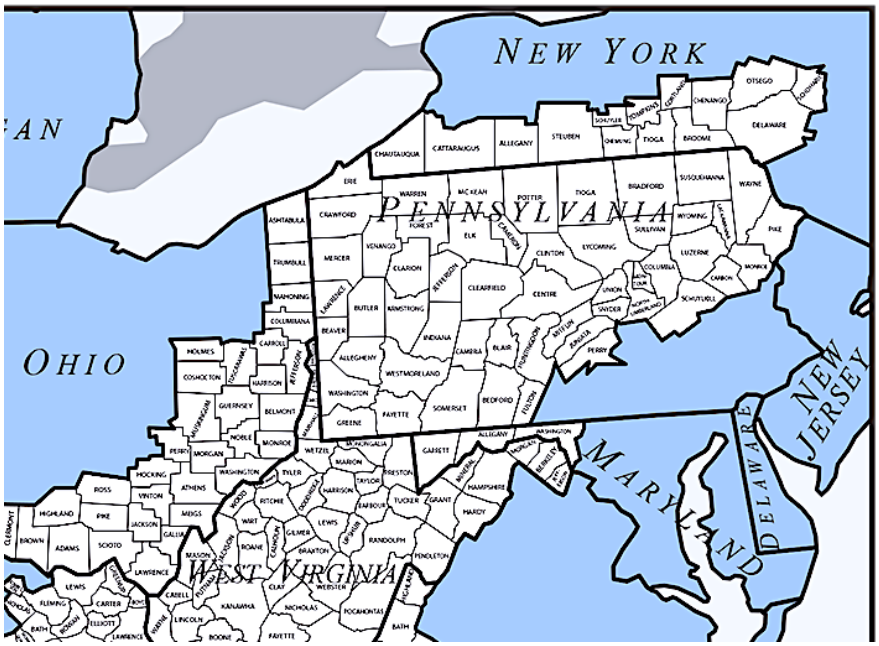

Figure 2. ARC Counties in Pennsylvania (52 in total) provided by the Appalachian Regional Commission

2.6 PA-ARC Counties through a Geographical Lens

Physiographically, there are six different geographic provinces within this PA (see Figure 3), which refer to physical features and processes of landforms and each of these provinces is categorized based on similar geology and landscape (Barnes, 2014, p. 31). These provinces are: Appalachian Plateaus, Atlantic Coastal Plain, Central Lowlands, New England, Piedmont, and Ridge and Valley (Barnes, 2014, p. 31-34). The mineral wealth made way for the coal, oil, natural gas, and non-fuel resources that created great industry within the state. These industry booms are what many people see and hear about in the sociocultural history of Appalachia. Within the state of PA, industry was greatly connected to the identities and lifestyles of residents further establishing a reality that was constructed around the natural resources.

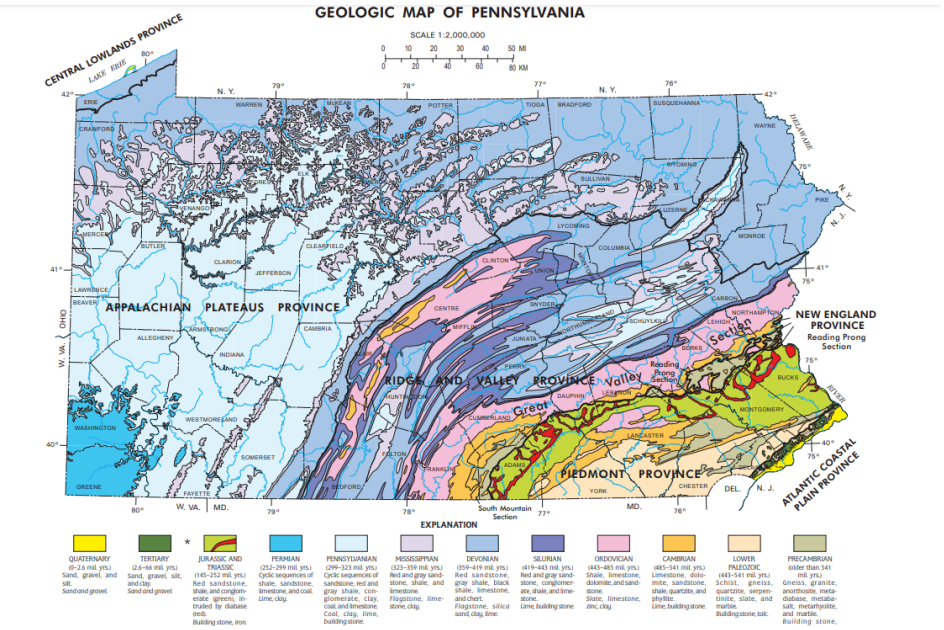

Figure 3. Geologic Map of PA (Pennsylvania Geological Survey, 2007)

2.7 PA-ARC Counties through a Sociocultural Lens

Geography and society have always been tightly connected within PA bleeding into the identities and lifestyles of people within these ARC counties. The abundance of many natural resources in conjunction with the extraction of these natural resources created many cycles of industry. Ferrence (2019) explains this cycle of industry in PA as "First, the land was home. Next, the mountains were obstacles. Then they were conquered, offering a triumph narrative for America. And then there was coal, gas, and the ripping of these things from the ground" (p.55). These PA counties' importance and contribution to the nation became defined by the resources beneath the ground. From this, the realities of residents were constructed around the geography and natural resources that were valuable to the region and valuable to society within the United States.

The culture of these areas was affected by many different shifts of industry. Prior to the coal and oil industry, PA was first known for the logging industry which began to

grow in the 1850's, after PA towns "fell behind statewide averages in capitalization and mechanization" (Sandow, 2010, p. 272). By the 1860's, these counties in PA were leading producers of lumber, but this thriving industry caught the attention of outside developers and national politicians, which began a long process of exploitation of the natural resources and residents of these areas. This exploitation led to wide-spread deforestation, which resulted in a downfall of the logging industry. Inhabitants of these areas were left devastated economically and grew a great disdain for "outsiders" and federal authorities (Sandow, 2010, p. 275). This devastation and disdain were perceived in many negative ways by outsiders, leading to stereotyping and the view of these people as "ignorant and easily imposed upon by designing politicians" (Sandow, 2010, p. 275). With lumber operations migrating out of PA and into West Virginia, the coal industry began to bring hope to these devastated areas, but quickly after brought "extreme poverty" (Sandow, 2010, p. 278) and continued exploitation of the natural resources. These continued struggles led to protests and rebellions which further painted the area and its people as "the worst class of human beings, both native and foreign, to be found in this country" (Sandow, 2010, p. 281). The circumstances surrounding social and environmental exploitation further perpetuated negative stereotypes by outsiders about the region and its people as poor, inferior, hopeless, and easily coerced. Through this industrial history and perpetuation of negative stereotypes, one can see the culture and lives of residents that still affect the region to this day.

Stereotyping and the perception of this area as an industrial, inferior, poverty-struck region still affects residents today. In many of these areas the mining industry no longer exists as a form of employment, but "it still subjects longtime inhabitants to the

physical/embodied and place-related elements of a mining region" (Meade, 2019, p. 102). The aftermath of environmental and social exploitation still exists today in the pollution of waterways and continuous economic struggles. This failed environmental and economic landscape is what brought these counties into the ARC in 1965. Cleaning of polluted waterways from mine runoff, reclamation of abandoned mines and improvement of various toxic waste centers continue to be a priority within the state government and ARC (Appalachian Regional Commission, 2015). Along with these environmental priorities, economic strategies and initiatives to replace previous industries that have left continue to be pressing issues that require attention from local, state, and federal governments (Appalachian Regional Commission, 2015). While these struggles are a reality for many, they continue to be romanticized within literature further perpetuating stereotypes and negative attitudes of the area, like other regions within Appalachia.

The romanticizing of these counties in PA continues these stereotypes furthering the negative associations with the region. Ferrence (2019) notes this romanticization in multiple novels stating that these PA counties are "considered as curious, strange, backwards, and somehow malevolent" (p. 180). Some of these authors that Ferrence (2019) mentions also use these qualities to promote themselves and the region while furthering negative attitudes and stereotypes, such as the self-written biography from Tawni O'Dell who describes her town in Western, PA like so; "the rolling hills are pitted with dead gray mining towns like cigarettes on a green carpet" (p.186) and that she "is half Pennsylvania redneck and half southern white trash" (p. 186). Literature and authors of the area continue to paint this region, and the residents within it, in a dim light that further invokes negativity and stereotyping, similar to other Appalachian literature. This

one literary example by O'Dell does not encompass other identities of speakers within the region that have a voice, but their regional identities are less clear to researchers and possibly even themselves.

2.8 PA-ARC Counties through a Linguistics Lens

When looking at PA, many people initially think of the two main cities: Pittsburgh and Philadelphia. These two main cities create a dynamic that instills rivalry across the state, insisting you choose which side you are on, and this even applies outside of sports. Only Pittsburgh is included within the ARC boundaries therefore, this study will have a heavier focus on this city.

Pittsburgh, PA, "The Steel City," was a part of these industrial and economic struggles. This city is still dealing with the economic aftermath of the once booming steel industry, but just as Appalachia is not a homogenous entity, neither are the ARC counties of PA. While most ARC counties in PA are predominately rural, Allegheny County, where Pittsburgh is located, is still considered urban based on population density. In fact, Hasty (2020) labels Allegheny County as "the only county is Appalachia that could officially be considered as urbanized area" (p. 16) and Pittsburgh has been mentioned before as the "Paris of Appalachia" (O'Neill, 2009). Even local newspapers such as the *Pittsburgh Post-Gazette* have published articles demonstrating their inclusion in the region of Appalachia such as "Yes, we and yinz are part of Appalachia" (O'Neill, 2011). Through Pittsburgh being the largest city within Appalachian boundaries, the language variety of the Pittsburgh residents is much more known to people outside of Appalachia.

Pittsburgh, linguistically, is fairly well-known compared to linguistic features in other PA counties. Johnstone (2006) explains how Pittsburgh is located at the convergence of many different language variety boundaries resulting in a "large number of sounds, words, and structures that sound nonstandard" which resulted in the American English variety that is Pittsburghese (p. 86). Common features of Pittsburghese are the cot/caught merger, "yinz" for plural you, merger of /i/ and /ɪ/ before /l/, /au/ monopthongization, /θ/ to /s/ assimilation, "jagoff" (an epithet for an annoying idiot), "gumband" for rubberband, and "red up" for tidying up, just to name a few (Johnstone and Pollak, 2016). These linguistic features are "widespread in central and western Pennsylvania, if not throughout the United States, and some of the lexical and morphosyntactic feature thought of as local can be heard throughout the Ohio Valley or the Midland, Southern, and/or Appalachian dialect areas" (Johnstone, 2006, p.87). Pittsburghese's presense within PA and other regions continues to experience perceptual shifts (from insiders and outsiders) linked to language, place, and identity. Pittsburghese became linked with "working class identity, incorrectness and/or lack of education" (Johnstone, 2009, p.163) further following the perceptions of PA during industrialization eras. Post-industrialization, perceptions of the language variety changed to index local identity, "pride and nostalgia, even among people who do not identify themselves as working-class or as speakers of a nonstandard variety" (Johnstone, 2009, p. 163). These features connected to language became a part of the regional identity indicating a deeper social meaning locally and nationally through enregisterment and commodification (Johnstone, 2009, p. 157). Pittsburghese and the connection to the city itself lead to the creation of this language variety through displaying, standardizing, and creating social

meaning to these linguistic features. The linguistic features can be seen in every area of the city from clothing, signs, memorabilia, and more. This intense enregisterment, commodification, and standardization created a dynamic that PA speakers inside and even outside of Pittsburgh also associated their identity through these features.

Through enregisterment and commodification, this language variety became representative of the city and other widespread surrounding counties. Other counties within Appalachian PA, show enregisterment and commodification, but on a much smaller scale still resulting in local identity and local representation of the language variety. In the PA coal regions, specifically Schuylkill County, shirts and memorabilia contain phrases like “I speak 2 languages: English and Coal Region” (see Figure 4) and many other shirts contains terms specific to the region such as “bolio” (a typical alcoholic drink during Christmastime in these regions sometimes referred to as the “champagne of the PA coal region) and “ho butt” (used to address people) can be seen on mugs and other items for sale. Within the realm of other PA counties, there are also items of clothing and memorabilia that indicate insider regional perceptions, such as clothing in Erie, PA stating that “It’s okay to love Erie” among other phrases (see Figure 5). While neither of these examples are on the same scale as Pittsburgh’s enregisterment and commodification, it does begin to show linguistic variants that have diffused into many areas of PA and how linguistic perceptions of the region have embodied the history and cultural identity of the state.

Figure 4. Enregisterment and Commodification in the PA Coal Region in local Historical Society Gift Shop (Shuylkill County Historical Society, 2021)

Figure 5. "It's Okay to Love Erie" Apparel (Erie Apparel, 2022)

Language varieties in PA have influence from many other dialect regions due to a variety of historical and cultural factors. Due to east to west fanning migration, the North and Midland dialect boundary runs right across the state of Pennsylvania (Wolfram, 2016, p. 29). In addition to this dialect boundary, there is German influence seen in syntactic structure and lexical terms such as “Are you going with?” and “stollen” the term for a kind of fruit cake (Wolfram, 2016, p. 32) along with influence from Pennsylvania

Dutch. Along with these features, sporadic sound changes are noted in Southern parts of PA such as “pa’lor” and “lib’ary” (Wolfram, 2016, p. 54). Other areas of these PA Appalachian counties that have been studied, such as Erie, PA, present features aligned with Midland vowel systems such as short-a raising before nasals and the cot/caught merger, but there is a divergence of the Midland boundary, influenced by Northern Cities, from a lack of participating in strong fronting of /ow/ and unrounded, lower open-o compared to Pittsburghese (Evanini, 2008). These small pockets of overlapping linguistic boundaries begin to show the various linguistic patterns in the area influenced by language contact. Additionally, Western Pennsylvania speakers also use verbal constructions, such as the use of “punctual whenever”, as noted in Table 2, which show the similarities to patterns in Appalachian Englishes. This connection to Appalachian Englishes is also seen in the cot/caught merger. With this influx of contact from varying linguistic boundaries, PA Englishes is just as diverse as the region and people but lacks investigation into areas outside of the two main cities, especially the ARC designated counties.

Table 2. Common Features from PA-ARC Counties and Appalachia (Montgomery, 2003, Hasty 2020, 2021, Reed 2020, Hazen 2020)

Features in Both PA and Appalachia	Examples
"leave" for "let	"Leave him go"
want + preposition with elliptical infinitive	"want in"
term for noisy mock wedding-night celebration (PA, WV)	"belling"
term for small stream	"run"
a-prefixing	"a-fishing"
term for brittle	brickle
formation of nouns and pronouns with the addition of 'un	"you'uns" "young'uns"
term for "nowadays"	"anymore"
directional adverb indicating "distance"	"yonder"
term for a sack	"poke"
term for low place between two mountains	"holler"
double modals	"might should"
completive done	"I done told you once"
term for "going to"	"fixin to"
rhotic insertion	"warsh"
nonstandard past tense	"knowed"
Theta deletion to [d]	"dis"
cot/caught merger	"cot" and "caught" sound the same
quotative "be like"	"the dog is like a baby"
demonstrative them	"Them dogs are running"
[-ing] variation	"going" more formal; "goin'" less formal
Punctual whenever	“whenever I was..”

While specific ARC counties in PA are not as studied as Pittsburgh, with only one study looking at specifically Erie County (Evanini, 2008), there is a convergence of features common to language within PA. A few previous studies that investigate the linguistic perception of these areas demonstrating that speakers in these ARC Northern Appalachian counties are “much less willing to admit using these features” of Appalachian Englishes (Hasty, 2021, p. 85), but they “appear to hear these features being used around them at apparently similar levels” (Hasty, 2021, p. 76). With a connection

between ARC counties in PA and Appalachian Englishes, one can see similar linguistic patterns, but showing a link between language, place, and identity will develop a deeper understanding of this region.

2.9 PA-ARC Counties through the Lens of Regional Identity Construction

Just as Appalachia and these Appalachian counties in PA are not a singular entity, their regional identity is not either. By only looking at singular details, one will not be able to capture or even fully understand these diverse and fluid regional identities. When combining multiple variables such as language production, perception, and geography, one can achieve a deepened insight into these ARC counties. While there is not an extensive amount of research looking at PA as a part of Appalachia, there is an emerging area of investigation into language and how to connects to the fluidity of Northern Appalachian identities. Anderson (2014) explains that people in ARC counties that live in the broader regions of Appalachia "may not understand what the label 'Appalachian English' means, but they call themselves hillbillies and hill-jacks and play hill-hop and hick-hop music" (p. 5) resulting in speakers in this area living in a state of ambiguous regional identity. Even educational and pedagogical studies note this such as Hayes (2017) that states "defining Appalachian identity is further complicated because it is largely a matter of self-definition, a self-definition that may or may not incorporate the term Appalachian" (p. 74). This self-identification and language used to communicate this Appalachian identity is further ambiguous when speakers are hesitant to associate with a highly stigmatized area and language. This linguistic hesitancy, born from

stereotypes and negative perceptions, may not be the only source complicating these Appalachian identities.

Possibly, this regional identity in PA is indicative of the many different perceptual factors that shape regional identity such as convergence of many industrial, linguistic, and cultural boundaries. Many labels can come to mind when discussing this area, from the standpoint of outsiders and insiders in these counties. In PA, there is an overlap between ARC designated boundaries and the "Rust Belt region" investigated in Autumn (2016) showing how economic growth increased and then declined, leaving these Appalachian Rust Belt areas devasted with many more consequences economically and culturally (p. 176). *Rust Belt* is a commonly heard term used in these areas to define the people and the region, but also *Midwestern* is another term used by insiders to describe themselves along with the term *Appalachian*. This confluence of labels could be due to the geographic proximity of all these areas. The proximity of these geographic boundaries and how they influence society are crucial when investigating regional identity and regional perceptions. Ferrence (2019) describes his own personal struggle in defining a regional identity and states "My home is sometimes called Appalachia, sometimes Rust Belt, other times the Midwest, even though very few would accept these labels as correct" (xi). The ambiguous regional identities in these PA-ARC counties, as noted in Ferrance (2019), indicate that there are multiple interpretations on which label is correct and that there seems to be no alternative choices that people within this area believe fit perfectly, whether it be for the identity of themselves or for the region. Studying the ambiguous identity of these people and their region is of the utmost importance because people are "very strongly predisposed towards partitioning and

demarcating our spatial surroundings" (Watt and Llamas, 2017, p.191) and "it seems almost self-evident that language plays a crucial part in how territories bounded by borders with their neighbors are defined" (Watt and Llamas, 2017, p.191). The language usage as well as language used to describe these regions play a huge role in understanding the construction of regional identity, or rather regional identities. Once we combine language and perception of PA ARC counties from speakers themselves, then we can truly begin to understand their regional identity

CHAPTER 3. METHODOLOGY

3.1 Historical Data Overview: Linguistic Atlas Project

The Linguistic Atlas Project (LAP) is a set of regional survey-based research projects that began in 1929, under Hans Kurath, that contain a tremendous amount of language variation of American English as well as cultural data on the different regions of the United States. The Linguistic Atlas Project covers over 800 linguistic features, which are referred to as *targets*, in every region. Through these surveys and topics, one sees the vast phonetic, morphological, syntactic, and lexical variation in a much wider context that can be applied to further continue research on regional variation (Linguistic Atlas Project, 2022).

Within these regional-based survey-style interviews, fieldworkers collected as much data as they could in the form of manual handwritten narrow phonetic transcriptions and some of the regional based projects have accompanying audio files recording the whole interview. For each interview, the fieldworker is equipped with a bound workbook with a copy of the worksheets (containing the specific targets) and a notebook with blank worksheets with left hand margins marked with numbers 1-8. This numbering system is used for the fieldworker to document the linguistic features in IPA in an organized manner since each page within the LAP has 8 targets. While documenting the narrow phonetic transcriptions, the fieldworkers must elicit responses to get these targets during the interview in a productive way that mimics conversation. In the *Handbook of the Linguistic Geography of New England* by Hans Kurath (1939), he lays out instructions for the fieldworkers on how to elicit accurate and representative responses without pressuring the respondents (p.49-50). The instructions and procedures in the first LAP project, the

Linguistic Atlas of New England (LANE), were then passed down and used as a standard for many more LAP projects to come.

3.2 Historical Dataset: Linguistic Atlas of the Middle and South Atlantic States

In the summer of 1939, the Linguistic Atlas of the Middle and South Atlantic States (LAMSAS) began doing fieldwork and interviewing residents in New York, New Jersey, Pennsylvania, West Virginia, Delaware, Maryland, District of Columbia, Virginia, North Carolina, South Carolina, Georgia, and Florida, which includes three dialectal regions: Northern, Midland, and Southern (Kretzschmar, 1994, p. 2). LAMSAS was a project that look much longer than LANE. This was due to the historical social events occurring within the United States. Kretzschmar (1994) states that the Great Depression hit the United States just as the project was being organized and "universities and foundations found their resources drastically curtailed, and in much of the Atlantic Seaboard such resources did not exist on the scale they did in New England" (p.1). Due to this decrease in resources, "funds were barely enough to keep one investigator in the field till 1941; with the involvement of the United States in World War II field work became impossible and the federal student support that provided a number of competent editorial assistants came to an end" (Kretzschmar, 1994, p. 1). With these difficulties came a much longer time frame over which the interviews were conducted and then organized to be published, but this does not take away the immense language and sociocultural information within LAMSAS.

LAMSAS used the same general interview and transcription procedure as used in LANE and "used a finely graded alphabet based on the International Phonetic Alphabet"

(Kretzschmar, 1994, p. 113), see Figure 6, 7, and 8 to see the vowel, diacritics, and consonant charts for LAMSAS.

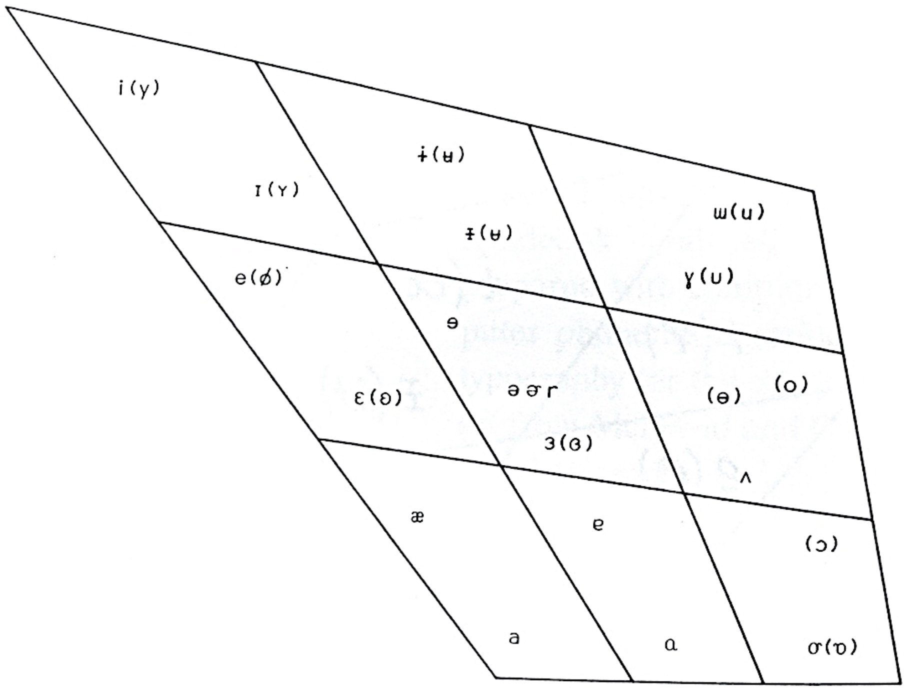

Figure 6. LAMSAS Vowel Chart (McDavid and O'Cain 1980)

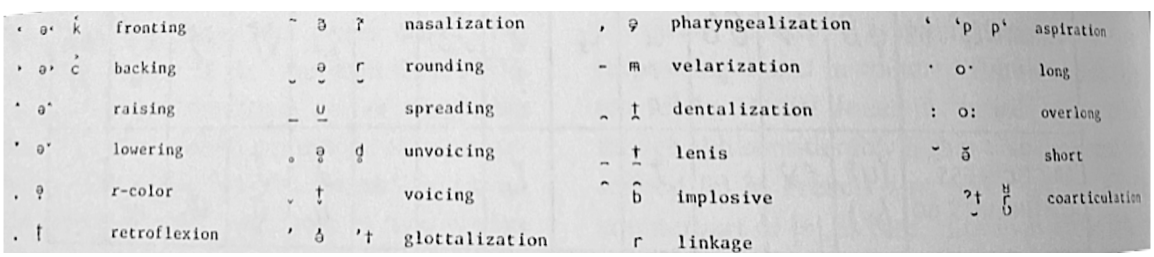

Figure 7. LAMSAS Modifying Diacritics (McDavid and O'Cain 1980)

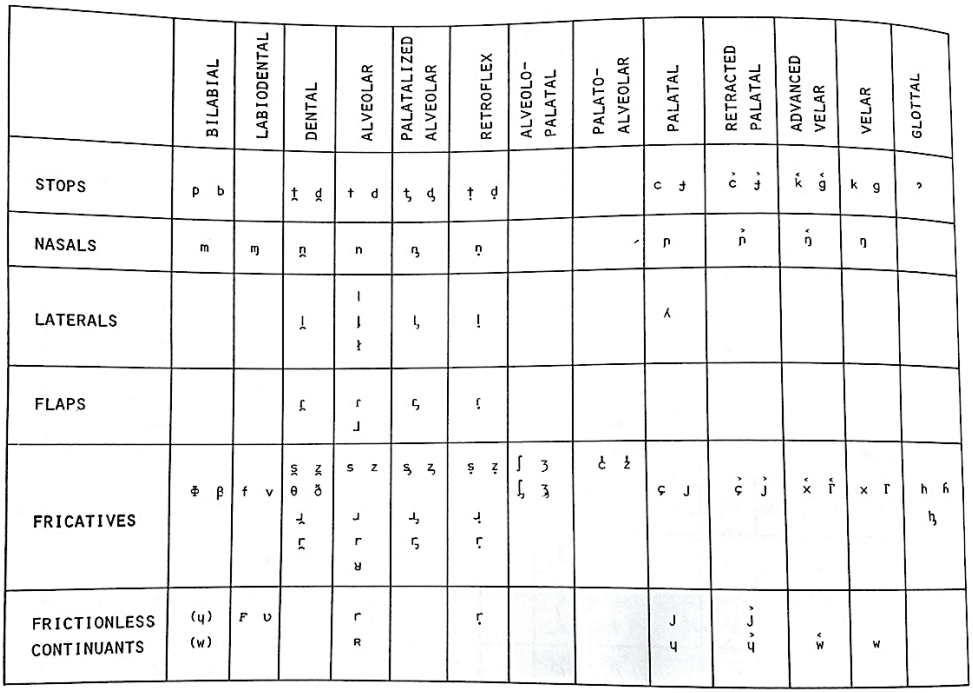

Figure 8. LAMSAS Consonant Chart (McDavid and O'Cain 1980)

Through these procedures and standards, fieldworkers conducted interviews with specific phonetic, morphosyntactic, and lexical targets in mind such as terms around the house, farm, and community including numerals, expressions of time, verb forms, adverbs, and terms for topography and roads, to name a few. The fieldworkers also made comments on sociocultural and perceptual information from the informant during the interview. For example, if the informant said many linguistic features that are indexed as Southern features, then the fieldworker may make a point to mark this down. These perceptual comments look at linguistic perception of the informant from the viewpoint of the fieldworker. This linguistic perception data is derived from the informant biographies (production) and perceptual information stated by the informant themselves (metalinguistic), which could be invaluable for studies in highly indexed and marked language varieties, such as Appalachian Englishes. Within LAMSAS, large parts of Appalachia are included within this interview data along with linguistic and perceptual

information of Appalachian Englishes. Northern Appalachia, by ARC standards, is included in LAMSAS as well as parts of Central and Southern Appalachia. This can also be observed through the topographical map of the LAMSAS region, see Figure 8.

Using the LAMSAS map (Figure 8) in comparison to the ARC map (Figure 1), one can see the similarities of land area covered. Included in these maps, are the ARC counties within PA. In LAMSAS, there were informants in all 67 counties totaling to 158 PA informants. Out of those 158 PA informants, 110 of them resided in ARC counties making this a valuable set of historical, linguistic, and sociocultural data for examining Northern Appalachian speakers.

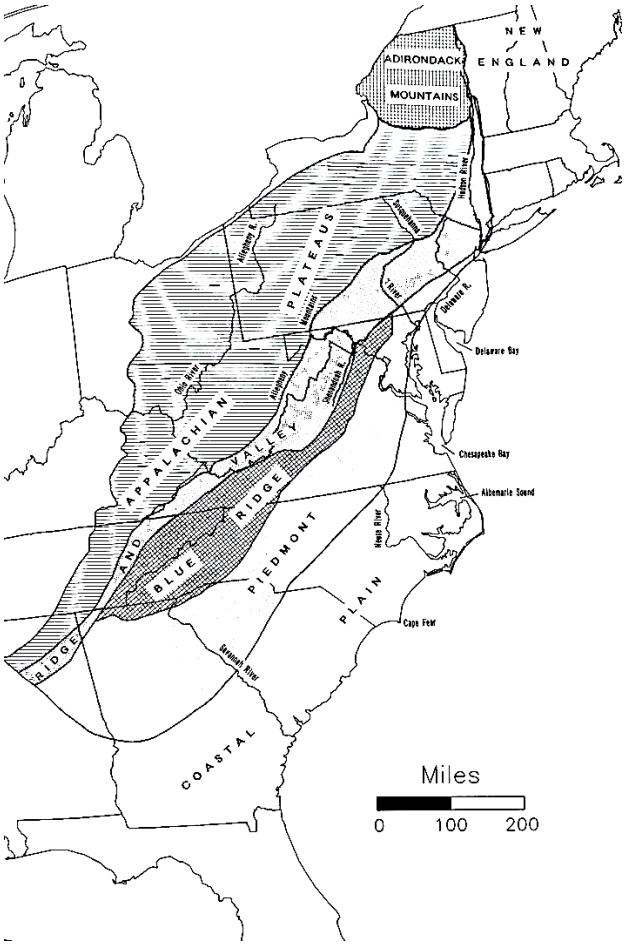

Figure 9. Topographical Map of LAMSAS region (Kretzschmar 1994)

These 110 historical informants provide crucial insight into these regions linguistically. Where LAMSAS tends to lack a bit is in the perceptual information from the viewpoint of the speakers (and their respective perceptions of their region and regional identity), but for this study, these informants will be used as a benchmark to compare linguistic production in this region to see if these features are still present within modern-day speakers. Phonetic and morphosyntactic features will be quantified based on their frequency of usage during the interviews, including normalized percentages that account for the amount an informant used these features compared to the overall amount that they "could" have been used through specific target elicitations from the fieldworkers. Lexical features will only be quantified based on their presence throughout the duration of the interview. Through this methodology, one will be able to see the presence of Appalachian Englishes features across historical PA speakers as means to investigate present day speakers.

3.3 Methodology of the Current Interviews: Production and Perception

To be able to obtain current data from these speakers, current in-person interviews were conducted with approval from the University of Kentucky Institutional Review Board (IRB). These interviews were conducted in a manner similar to the LAP survey-style interviews, but on a much smaller scale in quantity of informants and length of interviews. Another distinct factor between these current interviews and the LAP interviews is that they were structured into four parts: Brief Autobiographical Oral History, Language Production, Perception, and Demographic Information Questionnaire. All of these materials can be found in the appendices. I began the interview by asking the speaker to

tell some of their favorite childhood memories during the Summer, Winter, Spring, and Fall in PA. This was intended to create a comfortable and relaxing environment to elicit natural speech while being recorded. For the language production section, specific targets were selected in these interviews based on LAP targets and prototypical Appalachian Englishes features and PA Englishes features. The phonetic, morphosyntactic, and lexical features were then elicited through prepared questions, similar to the questions LAP fieldworkers would ask during survey-style interviews. After this, I proceeded to ask perceptual questions related to regional boundaries within the state of PA (Appalachia, Midwest, Great Lakes, and Rust Belt) such as "Where is Appalachia?", "Why did you choose this area as Appalachia?", and "What are some things that come to mind when you think of this region?". Lastly, after the interviewer stopped the recording, for confidentiality concerns, the speakers were asked to fill out a brief demographic questionnaire that asked about social variables such as age, occupation, current residing location (city and state), and their hometown in PA. All these factors are documented within the interview metadata along with fieldworker notes that describe the fieldworker's perception of the informant, like the procedures conducted within the LAP. Various structures within the present-day interviews were established to mimic the procedures within the LAP while including a designated section for perceptual data. This will ensure a valid and consistent means of comparison when analyzing the benchmark LAP data to the current interview data.

3.4 Scope and Content of the Current Interviews

The current interviews were comprised of speakers who were from the ARC-labeled counties in PA. To be eligible for this study, speakers had to be at least 18 years old, from the one of the 52 ARC-labeled counties in PA, and willing participate in an interview where audio would be recorded, per IRB established procedures. These current interviews include 22 informants from 7 of the ARC-labeled counties in PA (see Figure 10). These counties are primarily located in Western and Central PA. 12 of the 22 informants were male and 10 of the informants were female. The average age of the informants was 38 years old, keeping in mind that speakers were not able to participate in this study unless they were 18 years of age or older. The youngest informant was 18 years old, and the oldest informant was 82 years old. Education and occupation varied within the speaker set ranging from a high school education to degrees in higher education.

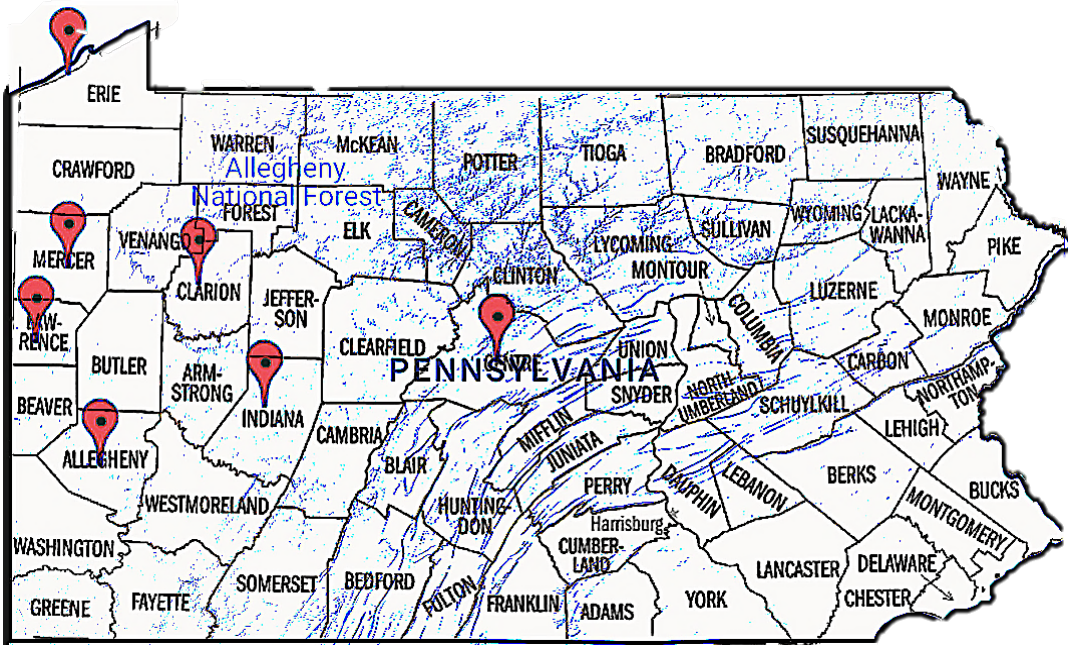

Figure 10. Counties Represented in the Current Data

The interviews in total lasted between 1-2 hours, depending on how talkative the informant was and how many people were being interviewed at once. Only one informant did not complete the perceptual part of the interview due to having to leave the interview early, so in total only 21 informants participated in the perceptual section of the interview. The same fieldworker expectations and procedures that were used within LAMSAS were replicated within these current interviews. In total, these current interviews contain 50 targets relating to linguistic features and 15 targets relating to perceptual information that were categorized into 5 groups: location and perception of Appalachia, location and perception of Midwest, location and perception of Great Lakes Region, location and perception of Rust Belt, and their own place-based regional identity. Each of these 5 groups, included 3 target questions that asked the speakers where they believed this region was, why they believe that this is the location of the regional label, and what their perceptions of the region are. These perceptions were then recorded as qualitative discursive comments and later grouped into six different groupings to visualize these perceptions easier (see Table 3). The speakers' perceptions were categorized into groups based on their perceptual comments surrounding their own regional identity and their perceived location of Appalachia. Through these groupings, the informants can be further categorized based on their perceptions to investigate if there are larger connections across the informants and similar perceptions.

Table 3. Perceptual Groupings

Regional Identification Groupings
Appalachia
Appalachia and other regions
Rust Belt only
Midwest only
Great Lakes only
Part of PA only; not Appalachia at all
Locating Appalachia Groupings
mark PA as a part of Appalachia only
mark PA as a part of Appalachia and other regions
Rust Belt only
Midwest only
Great Lakes only
Not Appalachia at all

CHAPTER 4. RESULTS

4.1 Historical Data: LAMSAS Linguistic Production Data

Prior to looking at the linguistic data, I made sure to construct a metadata worksheet with each informant's demographic information to specifically gather data from informants in ARC counties in PA. The demographic information recorded was gender, age, education, occupation, city, county, and fieldworker notes (see appendices for detailed information). From these 110 informants, 7 were female and 103 were male. The average age of the informants was 61 years old with the oldest being 90 years old and the youngest being 36 years old. After this, I began to collect and compile the linguistic production data. To collect the LAMSAS data, I referred to prior digitized transcriptions of the field records or the field records themselves, depending on if they have been digitized or not at the LAP. After collecting the specific targets, I organized them into a spreadsheet with informant number, target, response, and comments. The targets that I collected can be seen in Table 4. These targets were selected based on the criteria that they are commonly indexed as features of Appalachian Englishes and accessible through the LAP within a timely manner. Each of these targets and their corresponding descriptions are structured based on the LAMSAS worksheets and targeted elicitations used by the fieldworkers to document the speaker's use of each specific feature. For example, rhotic insertion will only be analyzed through the term "wash" because that is the available target that would record rhotic insertion that was accessible and available. While rhotic insertion could be seen in many other parts of a standard LAP interview, all these opportunities that the speaker "could have" used this feature cannot be captured through this study due to the inaccessibility of the full LAMSAS interviews along with the fact that not all field record documents are

digitized. Through this methodology, the total number of tokens per feature are based on the total responses within these specific targets and this methodology is consistent in the historical and current interviews examined within this investigation to create a comparable method of comparison.

Table 4. LAMSAS Targets Collected (Linguistic Atlas Project, 2022).

Phonetic Targets	**Morphosyntactic targets**	**Lexical Targets**
[ʃ] or [t͡ʃ] *("Appalachia")*	Gerund Variation (Switching between *"going" and "goin'"* in the same sentence*)*	terms for various roads *("road" "lane" "side street" "turnpike" "highway" "backroad")*
Offglide reduction *("m[ə]" for "my")*	Use of *"fixin'* to" instead of *"going to"*	terms for bodies of water *("river" "creek" "crick" "stream" "run" "marsh" "swimmin' hole" "gravel pit")*
Diphthongization *("one" and "two"- "waən" and "tʊu")*	Am not contraction *("ain't done")*	terms for different geographic features *("holler" "hollow" "mountain" "hill" "valley" "ridge" "cliff" "gulley")*
Monophthongization (*"five")*	A-prefixing *("a- singing and a-laughing")*	plural you *("yinz" "y'all" "you guys" "you" "you ones")*
Initial voiced theta deletion *("It's 'em")*	Double Modals *("might could")*	terms for outdoorsman/backwoodsman *("backwoodsman" "outdoorsman" "granola" "sportsman")*
[ɛ] [ɪ] substitution *("forgit")*	Punctual whenever *("Whenever I was cooking, a dog jumped over the couch.")*	terms for chipmunk *("grinnie" "chipmunk" "ground squirrel" "whistlepig" "chippy")*
Theta deletion [ð]to [d] *("over [d]ere")*	Past tense forms of climb *("clum" or "climbed")*	terms for spoiled milk *("turned" "spoiled" "bad" "lobbered" "rotten")*
Velar Nasal Fronting [ŋ] to [n] *("goin'")*	Past tense forms of dive *("dived" or "dove")*	Nearby states *("Maryland" "West Virginia" "Ohio" "New York" "New Jersey")*
Rhotic insertion *("warsh" "warshing" "warshed")*	Other non-standard past tense forms found in other targets *("spoked" "spoilt")*	Nearby cities *("Pittsburgh" "Philadelphia" "Erie" "Harrisburg" "Allentown" "Lancaster")*

From these targets, I then organized the informants and their corresponding responses into a spreadsheet and conducted a variety of counting methods to analyze the data. Token counts and observed frequencies were documented within the data. Token counts of the number of individual responses were used to create observed frequencies of the occurrences of each feature. The observed frequencies are the total number of times the informants used these Appalachian Englishes features (as documented by the LAMSAS fieldworkers within their field records) divided by the total number of responses to the specific target elicitation across all informants. While these frequencies do not show the total and exhaustive number of opportunities the informant had to use these features throughout the full interview, due to the full LAMSAS interviews not being accessible, it still begins to show the usage of these features historically (see Table 5 and 6).

Table 5. LAMSAS Phonetic Counts and Frequencies

Phonetic Targets	**Counts**	**Observed Frequency** *Feature occurred ÷ total target responses *100*
[ʃ] or [t͡ʃ] *("Appalachia")*	0	0
Offglide reduction *("m[ə]" for "my")*	42	35.80%
Diphthongization *("one" and "two")*	75	64.10%
Monophthongization *("five")*	95	86.30%
Initial voiced theta deletion *("It's them")*	32	8.62%
[ɛ] [ɪ] substitution *("forgit")*	0	0
Theta deletion [ð]to [d] *("over dere")*	65	17.52%
Velar Nasal Fronting [ŋ] to [n] *("goin'")*	58	10.10%
Rhotic insertion *("warsh" "warshing" "warshed")*	30	24.79%

Table 6. LAMSAS Morphosyntactic Counts and Frequencies

Morphosyntactic Targets	**Counts**	**Observed Frequency** *Feature occurred ÷ total target responses *100*
Gerund Variation (Switching between *"going" and "goin'"* in the same sentence	201	36.81%%
Use of *"fixin'"* to instead of *"going to"*	7	26.92%%
Am not contraction *("ain't done")*	126	11.00%
A-prefixing *("a- singing and a-laughing")*	144	26.37%%
Double Modals *("might could")*	173	22.12%
Punctual whenever *("Whenever I was cooking, a dog jumped over the couch.")*	0	0.00%
Past tense forms of climb *("clum" or "climbed")*	60	34.48%
Past tense forms of dive *("dived" or "dove")*	72	46.75%

From these tables, we begin to see the presence of traditional Appalachian Englishes features throughout all the informants. Along with this, there is also the occurrence of terms common to PA Englishes. Based on the historical LAMSAS data, there were no occurrences of the phonetic variation in the term "Appalachia" in total, so these token counts were zero along with the observed frequency. Speakers also did not show any production of [ɛ] to [ɪ] substitution of "forg[ɪ]t" indicating that this traditionally Appalachian Englishes feature is not present or common across all the speakers within these ARC PA counties in these specific targets. Diphthongization (in the targets "one" and "two") were produced frequently along with monophthongization (of "5"), rhotic insertions (of "warsh, warshing, and warshed"), and offglide reduction when using the

possessive pronoun “my.” All these features had the highest observed frequencies within the collected phonetic data compared to the number of targeted responses. There are also occurrences of velar nasal fronting (10% frequency) and morphosyntactic gerund variation (36.81%) within the same sentence showing the variants of these features and their presence within the data. Within this dataset, there is also the presence of a-prefixing occurring 114 times creating a 26.37% frequency of the production of this feature compared to the total responses the speakers produced within the collected target. There is a high frequency of double modals demonstrating, that out of 100 instances to use double modals, speakers produced this feature 22% of the time. Non-standard past tense forms of “climb” occurred 34% within the target and non-standard past-tense forms of “dive” occured 46% of the time within these targets across all informants. The phonetic and morphosyntactic features present are highly frequent, but there seems to be lesser lexical variation across these informants. This could be due to the rigid nature of the lexical elicitations or lack of additional terms offered by the speakers during the interviews. The lexical items with the highest lexical variation across all the speakers were terms for outdoorsman/backwoodsman and terms for plural you. Some examples of the different terms for outdoorsman/backwoodsman were “backwoodser,” “backwoodsman,” “mossback,” “mountaineer,” and “hick”. Some terms used by informants for plural you were “yinz,” “you,” “you people,” “yous,” “you’ns,” “you’m,” “you folks,” and “you folkses.” These LAMSAS observed frequencies will serve as a means of comparison and benchmark to analyze the current interview data from historical interviews.

4.2 Current Interview Data: Linguistic Production Data

For the current interviews, similar methods from the LAMSAS data were conducted. A metadata spreadsheet was compiled with demographic information (see Scope and Content). Along with this, the same targets were analyzed with only a few major differences. There were two groups of targets added looking at lexical features, which were asking informants to name words they felt were unique to PA and to name various regions within the United States. In addition to 2 new lexical targets, a couple of phonetic and morphosyntactic targets of PA and Appalachian Englishes features, such as dropping of the infinitive after "need" (*"needs washed")* were added to observe this morphosyntactic feature in addition to the phonetic feature of rhotic insertion and other non-standard past tense forms that occurred throughout the interview. Along with this, specific perceptual targets were asked, as mentioned earlier within the methodology. This data was collected through audio recorded interviews that were then processed and orthographically transcribed similar to the methodology of LAP interviews, except the step of creating field records was omitted and IPA transcriptions were only done on specific phonetic features. Table 8 shows these targets and Table 9 shows the perceptual targets added to these present-day interviews.

Table 7. Targets for the Current Interviews

Phonetic Targets	**Morphosyntactic targets**	**Lexical Targets**
[ʃ] or [t͡ʃ] *("Appalachia")*	Gerund Variation (Switching between *"going" and "goin'"* in the same sentence*)*	terms for various roads *("road" "lane" "side street" "turnpike" "highway" "backroad")*
Offglide reduction *("m[ə]" for "my")*	Non-standard past tense forms of climb, and dive, *("clum" or "climbed" "dove" "dived")*	terms for bodies of water *("river" "creek" "crick" "stream" "run" "marsh" "swimmin' hole" "gravel pit")*
Diphthongization *("one" and "two"- "waən" and "tʊu")*	Non-standard past tense occurrences *("spoked")*	terms for spoiled milk *("turned" "spoiled" "bad" "lobbered" "rotten")*
Monophthongization (*"five")*	Use of *"fixin'* to" instead of *"going to"*	terms for different geographic features *("holler" "hollow" "mountain" "hill" "valley" "ridge" "cliff" "gulley")*
Initial voiced theta deletion *("It's 'em")*	Am not contraction *("ain't done")*	plural you *("yinz" "y'all" "you guys" "you" "you ones")*
[ɛ] [ɪ] substitution *("forgit")*	A-prefixing *("a- singing and a-laughing")*	terms for chipmunk *("grinnie" "chipmunk" "ground squirrel" "whistlepig" "chippy")*
Theta deletion [ð]to [d] *("Over dere")*	Double Modals *("might could")*	terms for outdoorsman/backwoodsman *("backwoodsman" "outdoorsman" "granola" "sportsman")*
Velar Nasal Fronting [ŋ] to [n] *("goin'")*	Punctual whenever *("Whenever I was cooking, a dog jumped over the couch.")*	States that border PA *("Maryland" "West Virginia" "Ohio" "New York" "New Jersey")*
Rhotic insertion *("warsh" "warshing" "warshed")*	Dropping of Infinitive *("needs washed")*	Regions of the US *("Midwest" "Appalachia" "South")*
		Major cities in PA *("Pittsburgh" "Philadelphia" "Erie" "Harrisburg" "Allentown" "Lancaster")*
		Unique words to PA *("pop" "Ohi[a]" "yinz" "gumband")*

Table 8. Perceptual Targets from Current Interviews

Perceptual Targets
Location and Perception of Appalachia
Location and Perception of the Great Lakes Region
Location and Perception of the Midwest
Location and Perception of the Rust Belt
Perception of PA and how they personally label their region

From these targets, data was gathered to look at the presence and production of Appalachian Englishes features and perceptions of several regional boundaries within PA. The token counts and observed frequencies of these features can be seen in Table 10 and 11. These token counts and observed frequencies were calculated and conducted in the same manner as stated above in the LAMSAS data. This was kept consistent to create a valid means of comparison across the historical and current data.

Table 9. Current Interview Phonetic Counts and Frequencies

Phonetic Targets	**Tokens**	**Observed Frequency** *Feature occurred ÷ total target responses *100*
[ʃ] or [t͡ʃ] *("Appalachia")*	16	69.56%
Offglide reduction *("m[ə]" for "my")*	9	14.75%
Diphthongization *("one" and "two")*	18	29.50%
Monophthongization (*"five")*	32	52.45%
Initial voiced theta deletion *("It's 'em")*	12	46.15%
[ɛ] [ɪ] substitution *("forgit")*	1	5%
Theta deletion [ð]to [d] *("over dere")*	10	38.46%
Velar Nasal Fronting [ŋ] to [n] *("goin'")*	57	32.75%
Rhotic insertion *("warsh" "warshing" "warshed")*	13	22.03%

Table 10. Current Interview Morphosyntactic Counts and Frequencies

Morphosyntactic Targets	**Counts**	**Observed Frequency** *Feature occurred ÷ total target responses *100*
non-standard past tense occurrences	21	22.82%
Gerund Variation (Switching between *"going" and "goin'"* in the same sentence*)*	65	40.88%
Use of *"fixin'"* to instead of *"going to"*	1	2.38%
Am not contraction *("ain't done")*	0	0.00%
A-prefixing *("a- singing and a-laughing")*	1	0.57%
Double Modals *("might could")*	1	2.12%
Punctual whenever *("Whenever I was cooking, a dog jumped over the couch.")*	20	64.51%
Non-standard past tense of climb *("climbed" "clum")*	2	8.00%
non-standard past tense of dive *("dived" "dove")*	12	57.14%
Dropping of Infinitive *("needs washed")*	22	37.29%

From these tables, one can see the presence of features in Appalachian Englishes in present-day speakers. Some features seem to be occurring just as frequent as the observed frequencies in the historical data, while others are not as present, if at all. Phonetically, there is a strong presence of diphthongization and monophthongization occurring at a frequency of 29.5% and 52.45% out of the total responses within these targets. Along with this, there is a large occurrence of velar nasal fronting with 57 occurrences and a frequency of 32.75%, highlighting the production across all the speakers. In addition to this, there is frequent production of off-glide reduction of the possessive

pronoun “my” (14.75% frequency) and rhotic insertion (22.03%). Many of these phonetic features of Appalachian Englishes are present within the data, along with the variation of the pronunciation of “Appalachia,” which was not able to be documented within the LAMSAS data. Appalachia was pronounced as /æ.pəˈleɪ.ʃə/ at a frequency of 69.56% across the speakers within these targets. While this may not show a traditional Appalachian Englishes feature, it does indicate a variation in pronunciation of the region that is not present within LAMSAS data.

In terms of morphosyntactic features, the biggest differences are the lack of a-prefixing and double modals with only 1 occurrence of them each within the current interviews. Each of these occurrences are from the same informant. The informant (PA011) that produced the double modal “might could” commented that this term is not heard as frequently anymore and explained that it is more prevalent in Lancaster, PA. This same informant (PA011) produced the only instance of a-prefixing when he was telling a story about his childhood stating that his father “came a-runnin’” but this feature was not used again for the rest of the interview. There is also less occurrences of non-standard past tense forms of “climb, “occurring at a frequency of 8%, but when looking at the non-standard past tense forms of “dive” there is much more variation, occurring at 37.29%. Many informants commented on not knowing if “dove,” “dived,” or “doved” was the standard past tense form when being asked to elicit this target, which may have added to the frequency of this feature compared to the past tense forms of climb, which no informants commented confusion when responding to this target. The verbal construction of “needs washed” is prevalent within the data and occurs 22 times with 3 different types of the features across the speakers. The types of this morphosyntactic construction were “needs

washed," "needs to be washed," and "need washed." One informant (PA019) indicated that the verbal construction of "needs to be" is only used in Philadelphia. Other speakers interviewed also stated that they know the "proper" way to use this verbal construction is "need to be" and commented that people have corrected them on this morphosyntactic feature.

With there being variation in the tokens and frequencies, it is of the utmost important to compare these features and frequencies in a normalized manner, which can be seen in Table 11 and 12. The normalized frequencies were calculated to a desired size of how often the feature was produced "per 100 words" to accurately compare the two interviews that had varying lengths. The token count was multiplied by the desired normalization size, in this study it is 100, and then divided by the total number of target responses that feature in that specific target elicitation. This was done for each feature in each context of the elicitations of each interview type (historical and current).Through this normalization, the frequencies represent the number of instances the feature occurred every 100 words relative to the specific elicitations of the historical and current data across all the informants. For example, in LAMSAS, the feature of offglide reduction was used 42 times across all LAMSAS speakers combined, resulting in 42 total tokens, this number would be multiplied by 100 and then divided by 178 (the total number of responses within the targets collected) resulting in a normalized frequency of 23.59 out of 100 (23.59%). This indicates that offglide reduction, when using the possessive pronoun "my," was used 23.59% of the time out of 100 responses to the target elicitations asked by the fieldworker. This is a way to look at frequencies of features across two different length interviews in a

comparable manner to further understand the usage and production of these features historically and currently.

Table 11. Normalized Frequencies in LAMSAS and Current Interview Phonetic Data

Phonetic Targets	LAMSAS Counts	LAMSAS Normalized Frequencies	Current Interview Counts	Current Interview Normalized Frequencies
[ʃ] or [t͡ʃ] *("Appalachia")*	0	0%	16	69.57%
Offglide reduction *("my")*	42	23.59%	9	14.75%
Diphthongization *("one" and "two"- "waən" and "tʊu")*	75	64.10%	18	29.51%
Monophthongization (*"five")*	95	86.36%	32	52.46%
Initial voiced theta deletion *("It's 'em")*	32	8.63%	12	46.15%
[ɛ] [ɪ] substitution *("forgit")*	0	0%	1	5%
Theta deletion [ð]to [d] *("over [d]ere")*	65	17.52%	10	38.46%
Velar Nasal Fronting [ŋ] to [n] *("goin'")*	58	10.10%	57	32.76%
Rhotic insertion *("warsh" "warshing" "warshed")*	30	24.79%	13	22.03%

From these feature-specific normalized frequencies per 100 instances of production, offglide reduction is more frequent in the LAP than the current data, showing decreased production of this feature in current the data. This decrease in use is also noted in monophthongization and diphthongization possibly hinting at a decrease in usage of this

feature. While some features are not as prevalent in the current interviews, theta deletion to a dental plosive and initial theta deletion are occurring more frequently (38.46% and 46.15% respectively) in the current interviews versus the historical interviews. Along with this, velar nasal fronting seems to be occurring more frequently within the current interviews at a frequency of 32.76% in the current interviews compared to 10.10% in the historical interviews. Overall, there seems to be a decrease in a majority of phonetic Appalachian Englishes features in current ARC PA speakers. While this study will not investigate in detail historical language change and variation in these PA counties, this decrease in frequency of features could be attributed to a variety of sociocultural factors such as industry, history, (LAMSAS was conducted in 1939 when the coal, oil, and steel industries in PA were still very prominent whereas these industries are in decline or non-existent now in these areas of PA), and society within PA during the time of the historical interviews. A reasoning, seen within the current data, that could contribute to this analysis is the informant that used these lesser frequent phonetic features such as diphthongization ("one" and "two") and monophthongization ("five" and "nine") was the oldest speaker (82 years old) in the current interviews (PA011). This is the same speaker who used the morphosyntactic features of a-prefixing and double modals. While this is only a small piece of evidence to give context to the change in frequency of phonetic features, it could be a beginning point of looking into language change within these areas. Due to language and speakers being dynamic and everchanging, this historical variation could be occurring based on a multitude of reasons that are better investigated with a larger and more representative modern-day interview dataset.

Table 12. Normalized Frequencies of LAMSAS and Current Interview Morphosyntactic Data

Morphosyntactic Targets	LAMSAS Counts	LAMSAS Normalized Frequencies	Current Interview Counts	Current Interview Normalized Frequencies
Gerund Variation (Switching between "going" and "goin'" in the same sentence)	201	36.81%	65	40.88%
Use of "fixin' to" instead of "going to"	7	26.92%	1	2.38%
Am not contraction ("ain't done")	126	11%	0	0%
A-prefixing ("a-singing and a-laughing")	144	26.37%	1	0.57%
Double Modals ("might could")	173	22.12%	1	2.13%
Punctual whenever ("Whenever I was cooking, a dog jumped over the couch.")	0	0%	20	64.52%
Non-standard past tense of climb ("clum")	60	34.48%	2	8%
Non-standard past tense of dive ("dived" "doved")	72	46.75%	12	57.14%

As some phonetic features are not as present in the current data, there seems to be some historical variation across the morphosyntactic and lexical features of Appalachian Englishes in these areas. LAMSAS informants were using double modals, a-prefixing, and "fixin' to" (22.12%, 26.37%, and 26.92% respectively) at a much higher frequency than

current speakers which showed very little occurrences of these features (2.13%, .57%, and 2.38% respectively). The one informant (PA003) who did use “fixin’ to” in the current interviews made the comment, after responding to the target elicitation, that “he has not heard that term in PA for a long time and now many people refer to it as a southern term” even though he has heard it in PA before. Along with this, as mentioned above, the only informant (PA011) who used double modals and a-prefixing commented that “might could” is not heard as frequently anymore and produced the only instance of a-prefixing when he was telling a story about his childhood. The non-standard past tense use of “climb” was still much higher in the historical interviews compared to the current interviews with the LAMSAS interview showing a frequency of 34.48% and the current interviews showing a frequency of 8%. While the non-standard past tense form of “climb” showed a decrease in frequency, the non-standard past tense form of “dive” occurred 57.14% out of the current data compared to 46.75% in the historical data. In the current data, informants commented on the confusion on what was the “proper” past tense of “dive,” but did not have this same confusion with “climb.” While this contextual commentary is not as easily accessible in the historical data, it would be intriguing to see if this confusion and uncertainty is accounted for by the historical speakers in LAMSAS as well. Along with this, there is an increase in frequency of gerund variation in current data speakers (40.88%) showing usage of “-ing” and “-in” within the same sentence. Some of variation in the presence of these features could be due to sociocultural variables or even the context through which the speakers are using these features. For instance, PA011 used a-prefixing through the context of storytelling, which could indicate more natural unconscious speech or maybe the storytelling warranted the use of a-prefixing. Along with this, the occurrence

of non-standard past tense forms of “climb” and “dive” vary across the current and historical speakers. From the current speakers, there seems to be uncertainty on which form is grammatically correct, which warranted more frequency, but this cannot be confirmed or denied through the contextual commentary on behalf of the historical informants. These are only a few considerations that need to be taken into consideration when comparing the frequencies and counts within these datasets that vary in contextual information and interview methodology. These considerations are crucial when looking into all language patterns and features.

Lexically, both sets of data show the presence of stereotypical Appalachian terms historically and currently. The term “run” as a small body of water was seen in both LAMSAS data and in the current interviews. Within the current interviews, many informants noted that sometimes “run” is used as a proper noun while to some speakers it was a common noun. Additionally current speakers also noted a semantic difference between “creek” and “crick” (PA008 & PA009) stating that these distinct terms mean two different bodies of water. Along with this, many informants discussed the differences between the terms “holler”, “hollow”, and “valley”. Some informants stated semantic differences based on the presence of a road, body of water, and town. A few informants commented that “holler” can be a proper noun for a specific part of town such as “pigs holler” (PA011). While there seemed to be less lexical variation in the LAMSAS data across bodies of water and geographic features, there was a lot of variation across different terms for plural you such as “yinz,” “you’ns,” “you’m,” and “yous.” These variants were seen within the current interviews as well with the additional presense of “y’all” and “you guys.” Both the historical and current data noted many terms for

"outdoorsman/backwoodsman" such as "hillbilly", "redneck", "hick" and "mossback". Within the current interviews, these terms were documented, but many current speakers have specific terms dependent on the outdoor activity such as "hunter," "camper," "hiker," and "backpacker." The current speakers also made distinctions based on the type of "outdoorsman" such as "granola" or "crunchy" (for an outdoorsy person who is a nature-lover) along with a distinction on the socioeconomic factors such as a "sportsman" is someone who likes to be outdoors but will spend a lot of money on expensive outdoor equipment. Lexical data in the current interviews show many more terms for surrounding states such as "West Virginny" (PA002) and "Merland" (PA010). One lexical term that both datasets showed a large amount of variation was the pronunciation of Ohio as "Ohia." Within the current data, many informants note that people from PA "always say Ohia" (PA008 and PA010) showing different variants that indicate a speaker is from PA. Specifically, the current interview target asking speakers about unique words in PA shows diverse terms that indicate or index speakers within the region. While this target shows large lexical variation due to the nature of this question and freedom of the speaker to discuss any terms that come to mind, it begins to show some of the unique indexed terms within the region. Speakers repeatedly noted terms such as "red up" (to tidy an area), "still" (for steel), "Stillers" (for the NFL team in PA), "sweeper" for a vacuum, "poke" (for a bag), "mups" (for tourists from Pittsburgh who come visit), "pop" (for a carbonated beverage), "yinz" (for plural you), "gumband" (for rubberband), and pronouncing "window" and "Ohio" with an [a] at the end ("winda" and "Ohia"). Overall, due to the circumstances of the LAMSAS, there would seem to be less lexical variation in the historical dataset, but this could be indicative of various methodological constraints.

Differences between the lexical variation could be due to many factors embedded within the methodology of both historical and current interview procedures. LAMSAS informants may not have noted as many different words for each target. Also, the LAMSAS elicitation questions asked by fieldworkers were specific and geared towards specific responses on targets due to the longer interview length and ability to ask more specific questions. Along with this, the list of targets for the LAMSAS interviews are much longer than the list of targets within the current-day study. In the current interviews, the questions were more open-ended aiming to get as much language production data in a 1-2 hour time frame. Without taking into consideration the circumstantial context and methodological information surrounding both the historical and current interviews, these lexical features could propose there is more lexical variation currently in these counties across speakers than before, but this would be a rather large conclusion to make and would need more expansive research to be conducted to confirm or deny this statement. While this is a large conclusion for this study to make, one can see, through these tables, the variation in frequency of Appalachian Englishes features historically and currently showing varying occurrences between the two interview sets, but still indicating the presence of commonly indexed Appalachian Englishes features.

4.3 Current Interview Data: Perceptual Information

After processing, compiling, and analyzing the production data, the perceptual data was then converted into categorical groupings to be able to conduct data analysis and various visualizations derived from this data. The groupings are based on specific locations that the speakers designated as Appalachia and was categorized based on if they considered

PA as a part of Appalachia (see Table 13). From this table, we can see that 18 informants considered PA to be a part of Appalachia and 6 (of these 18) believe that PA is a part of Appalachia in addition to another regional boundary. The 3 informants who did not perceive PA as Appalachia, noted that PA was part of other regions, specifically the Rust Belt, Great Lakes, and the Midwest.

Table 13. Inclusion of PA in Appalachia

Informant	County	Perception
PA001	Allegheny	Appalachia
PA002	Lawrence	Appalachia
PA003	Clarion	Appalachia
PA004	Erie	Appalachia
PA005	Clarion	Appalachia
PA006	Erie	Appalachia
PA007	Erie	Appalachia and other
PA008	Lawrence	Appalachia
PA009	Mercer	Appalachia
PA010	Erie	Appalachia and other
PA011	Clarion	Appalachia
PA013	Erie	Rust Belt only
PA014	Erie	Appalachia and other
PA015	Erie	Appalachia and other
PA016	Erie	Appalachia and other
PA017	Erie	Appalachia and other
PA018	Erie	Great Lakes only
PA019	Centre	Midwest only
PA020	Erie	Appalachia and other
PA021	Erie	Appalachia and other
PA022	Indiana	Appalachia

During the perceptual part of the interview, a map without any labels of the United States (see appendices) was provided in case any informants wanted to use it to point and explain their perception of where Appalachia, the Midwest, the Great Lakes region, and

the Rust Belt is located. Some informants had varying perceived locations of Appalachia based on geography, sociocultural variables, and perceptions of the area. PA010 described how the region of Appalachia changes "based on the mountains compared to history and culture" and that "their boundaries are different for each" when locating Appalachia, but both versions of Appalachia included PA. PA002 commented that PA was included in Appalachia and only a little bit of Ohio because "we don't wanna give them that much credit." PA002 also commented that "purists would disagree with this location of Appalachia (including PA) because people feel it is more Kentucky and West Virginia." This perception that others may not agree with their inclusion of PA was stated by many speakers within the interview including PA003 who felt their "idea of Appalachia is more northern, and some people would not agree with that." Other informants would only include parts of PA as a part of Appalachia and commented that there are "cut off counties" such as PA009 stating that "Clarion County is Appalachia, but Crawford County and Erie County are not." Within some of the group interviews, informants commented on each other's inclusion of PA in Appalachia such as PA008 commenting to PA009, "you don't think Westmoreland County is a part of Appalachia?!" showing a bit of concern that PA009 did not include this county as a part of the region. Along with this, other informants noted as they were describing the location of the region of Appalachia that "if I was being generous, I would maybe include up to central PA, but definitely not Pittsburgh"(PA021). Some informants commented that Appalachia is only located in the region "where the Appalachian Trail runs through and that is what I was taught in school" (PA011) or that "central PA is called Pennsyltucky, so that is Appalachia" (PA015) choosing their interpretation of Appalachia based on geography or states perceived as Appalachian. Along

with this, society and pop culture also determined for informants their perceived location of Appalachia. PA019 discussed that they watched a documentary recently informing them that Appalachia is "only where southern states touch West Virginia, so I know the answer to this for sure!" Along with this, many informants had varying explanations and perceptions as to why this was the location of Appalachia showing a diversity of opinions surrounding the location of this region that 18 out of 21 times of the times in total included PA, even if it was only specific parts of PA that overlapped with other boundaries such as the Midwest, Rust Belt or Great Lakes region. These results may seem definitive, showing that PA is Appalachian, but the variation lies within the informant's perception and self-identification of Appalachian. This is precisely what the other categorical grouping investigated (see Table 14).

Table 14. Regional Identification of Speakers

Informant	County	Perception
PA001	Allegheny	Pittsburgh
PA002	Lawrence	Appalachia
PA003	Clarion	Appalachia
PA004	Erie	Appalachia and other
PA005	Clarion	Appalachia and other
PA006	Erie	Great Lakes only
PA007	Erie	Great Lakes only
PA008	Lawrence	Appalachia and other
PA009	Mercer	Northern PA
PA010	Erie	Rust Belt only
PA011	Clarion	Western PA
PA013	Erie	Appalachia and Great Lakes
PA014	Erie	Northwestern PA only
PA015	Erie	Appalachia
PA016	Erie	Great Lakes only
PA017	Erie	Appalachia and other
PA018	Erie	Appalachia and other
PA019	Centre	Midwest only
PA020	Erie	Appalachia and other
PA021	Erie	Appalachia and other
PA022	Indiana	Appalachia

From Table 14, we see that only 4 speakers would identify as only "Appalachian" while 8 identified as Appalachian and another region within the US ranging from Rust Belt, Midwest, Great Lakes, to even just specific parts of PA. Here we begin to truly see these ambiguous and fluid identities come to the surface even though a majority of informants are locating their state and even hometown as a part of Appalachia, but when asked if it is their identity, some seem to deny.

The 4 informants that identified as "Appalachian only" had similar statements describing their regional identity. Many stated that their identity was unique and

uncommon to others. PA022 stated that "I am a weird Appalachian because I consider myself in a box that is a different type of Appalachian because PA is a different subset of Appalachia." PA015, another informant who identified as Appalachian discussed how "everywhere except for Philly and Pittsburgh is Appalachia, so that is why when people ask where I am from, I say Northern Appalachia." While this identification seems based on location and geography, some informants identified their Appalachianness through societal values connected to the region, such as PA003, who identified as Appalachian stating that "I was born in Appalachia and I am Appalachian, but I did not truly realize what this meant until I moved and saw how other regions did not have the same sense of home, hardworking mentality, and prideful demeanor as I grew up with." The last informant (PA002), that described themselves as Appalachian, hinted at the stereotypes of the region to further describe their regional identity by saying that "people are playin' banjos on porches in my hometown too, but I am in the boonies. Many people would say I do not count as Appalachian because I am outside the 'heart' of Appalachia." While these informants were clear in their regional identities, there were some that were contemplating and discussing the many regional identities that shape their individual regional identity.

The informants who described their regional identity as "Appalachian and other" frequently described why they could not just pick one based on many reasons. This is noted in PA004, who stated that "I am from Appalachia, but also the Great Lakes due to Lake Erie being central to just about everything here. Also, the industrial aspect of the Rust Belt has shaped who I am, but I am not midwestern. You see corn around here, but not that much." PA004 and PA013 have similar ways of identifying themselves and PA013 described how in Erie County "you have aspects of the culture of Appalachia, but you also

have a lake that is connected to just about everything around here" showing how the Great Lakes region and Appalachia interact to create this ambiguous or rather multiple identities within one individual. PA008 has similar comments on their regional identity explaining that they "would use all these regional terms to identify" themselves. PA017 and PA018 identified based on the fact that their region and values are "less city-focused, a little mix of redneck and rural" (PA017) and that "if I had to choose one, I do not think I could" (PA018). Showing that choosing only one of these terms was complicated and almost impossible to do, hinting at the ambiguity of this identity. All these informants commented that their identity is tied to many things crucial to society and the place they call home, showing that just "Appalachian," is not enough or indicative of their true regional identity. While many informants regionally identified as Appalachian or Ambiguous Appalachians, others described themselves as "midwestern" (PA019), "Rust Belt" (PA010), "Great Lakes region" (PA006, PA007), and others described themselves based on the area of PA they are from (PA001, PA009, PA014).

The informants that described their identity outside the regional identity of Appalachian, discussed their varying perceptions on why that is the identity they possess. PA019 stated that "I always tell people I am from the Midwest because I do not think the Great Lakes or Rust Belt is a regional identity, so Midwest is always easiest." This shows that while these other boundaries are prevalent to the state, they sometimes do not seem to be actual regional identities, but more just locations within the state. PA010 stated that "I am a part of the Rust Belt region. I mean I cannot deny the effect that Cleveland, Buffalo, and Erie have on me. I identify with those values" showing how their identity is defined by the influence of history and culture of their area and surrounding locations shape their

own identity. Informants that described their regional identity as part of PA varied in their individual descriptions from “Northern PA” (PA009), “Western PA” (PA011), to even more specific descriptions such as “Northwestern PA” (PA014). PA009 stated that “these regional terms we used in this interview do not really apply to me. I would not define myself based on any of these. If someone asked me where I am from, I would say Northern PA.” The ambiguity surrounding which label would define the informants was common across the informants that identified based on their region of PA which is also seen by PA011 who said, “I guess I would just say Western PA because I have never defined a regional label for where I am from.” This informant hints at the idea that this area does not seem to have a regional label outside of just Pennsylvania. PA014 has these same sentiments by discussing that “if I had to pick a label then maybe Great Lakes, but I always just say Northwestern PA.” This begins to show that this region seems to have an unclear regional label that applies to the individual identities within the state. Further, PA001 identified as the city where they are from “Pittsburgh” because “it is so well-known, so I cannot really imagine having to use a label.” While these discursive comments begin to show the perceptions of the region and diverse regional identities of individual speakers, they do not show any connections across the speakers. Do similar regional identities have a pattern based on the county the individuals reside in?

Appalachia is filled with diverse individuals and states of mind, which is shown through these perceptual comments, but a deeper look into these regional variables will provide insight into patterns based on these regional perceptions and ambiguous identities. In Figure 11, this geo-spatial map shows the distribution between perceived location of Appalachia and counties across the informants. Further showing that informants, as a

majority, located PA and their hometowns as a part of the region of Appalachia, with only 3 informants (PA013, PA018, PA019) not including PA as a part of Appalachia. From Figure 11, it seems there is a common group that is consistently labeling their counties as a part of Appalachia (indicated in bright green) ranging from Lawrence County, Mercer County, Clarion County, and Erie County. This consistency is not seen in Erie County though, with some informants labeling the region as "Appalachia," "Rust Belt" (indicated by the red pin), and "Great Lakes" (indicated by the blue pin) while the majority have labeled it as "Appalachia and other" (indicated by the light green pin). In addition, the informant (PA019), farthest east in the state, located Centre County and PA as being a part of the Midwest (indicated by the yellow pin) and not Appalachia. From this map, it seems that informants near Lake Erie seem to indicate multiple regions that converge in this area, such as the Rust Belt, Great Lakes, and Appalachia. A reasoning for this could be the converging boundaries within this area and influence of other larger cities along with the cultural values attached to them influencing Erie County. This multitude of labels is not as prevalent as one moves further south in PA, where all informants labeled PA as a part of Appalachia. Through this map, there seems to be consistent patterns based on the inclusion of PA in Appalachia, but this consistency is not observed when considering regional identity (see Figure 12).

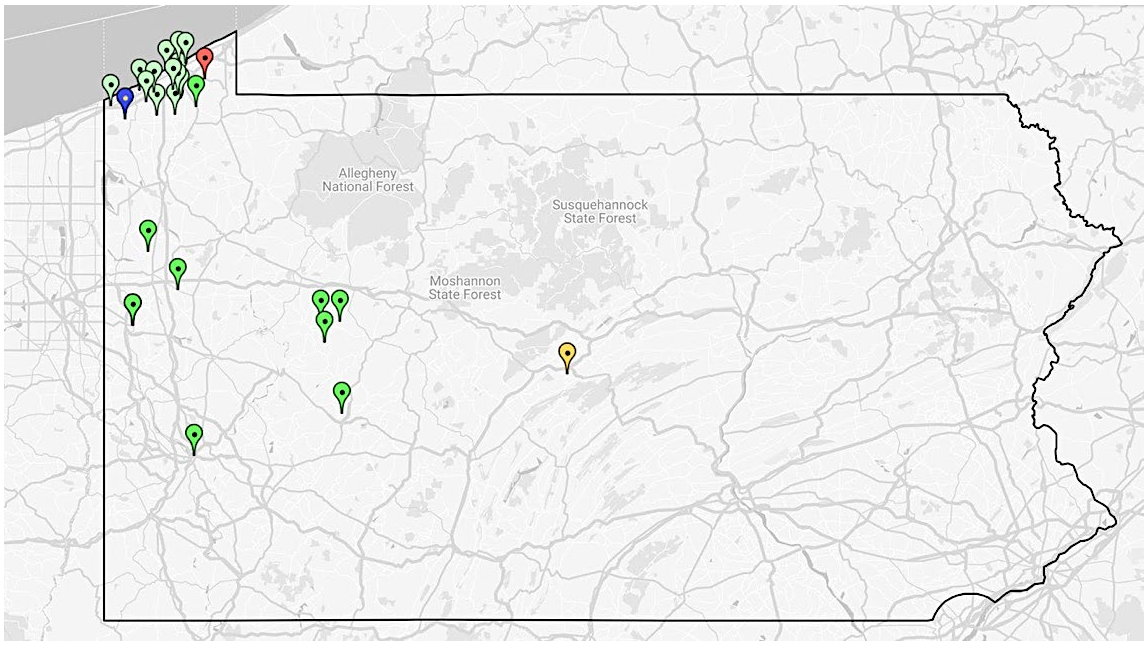

Figure 11. Map of Informants Inclusion of PA in Appalachia

Figure 12. Map of Informants Regional Identity

From Figure 12, one can see 4 informants, in four different counties, identifying as Appalachian (PA002, PA003, PA015, and PA022). These informants are from Lawrence County, Clarion County, Erie County, and Indiana County respectively. Along with this, 8

informants identified as Appalachian and another regional label, from Clarion County, Erie County, and Lawrence County. Both the groups that identified as "Appalachian" and "Appalachian and other" are exclusively from these 4 counties. Erie County had more informants identify themselves as Great Lakes, which can be expected since they are geographically closest to Lake Erie and this county also had the only informant that labeled themselves as part of the "Rust Belt." This can be due to Erie, PA being historically and culturally embedded within the Rust Belt and the societal values that are indicative of this industrial history along with the connection to closely located Rust Belt cities, such as Detroit, Buffalo, and Chicago. The informants that regionally identified as part of PA are in four different counties, Erie County (PA014), Clarion County (PA011), Mercer County (PA009), and Allegheny County (PA001). This distribution of regional identity is inconsistent and there seems to be no pattern, even when taking into consideration demographic information of each of the informants. This inconclusive pattern is also prevalent with only one informant regional identifying as Midwestern (PA019), but this is not able to be investigated within the scope of this study due to a lack of representativeness of all 52 ARC counties in PA. Through looking at geospatial distribution of the perceived location and regional identity of the informants, there seems to possibly be a connection between certain geographical and societal variables between locating Appalachia and defining one's regional identity, but Figure 13 begins to combine these two perceptual categories to investigate this connection closer.

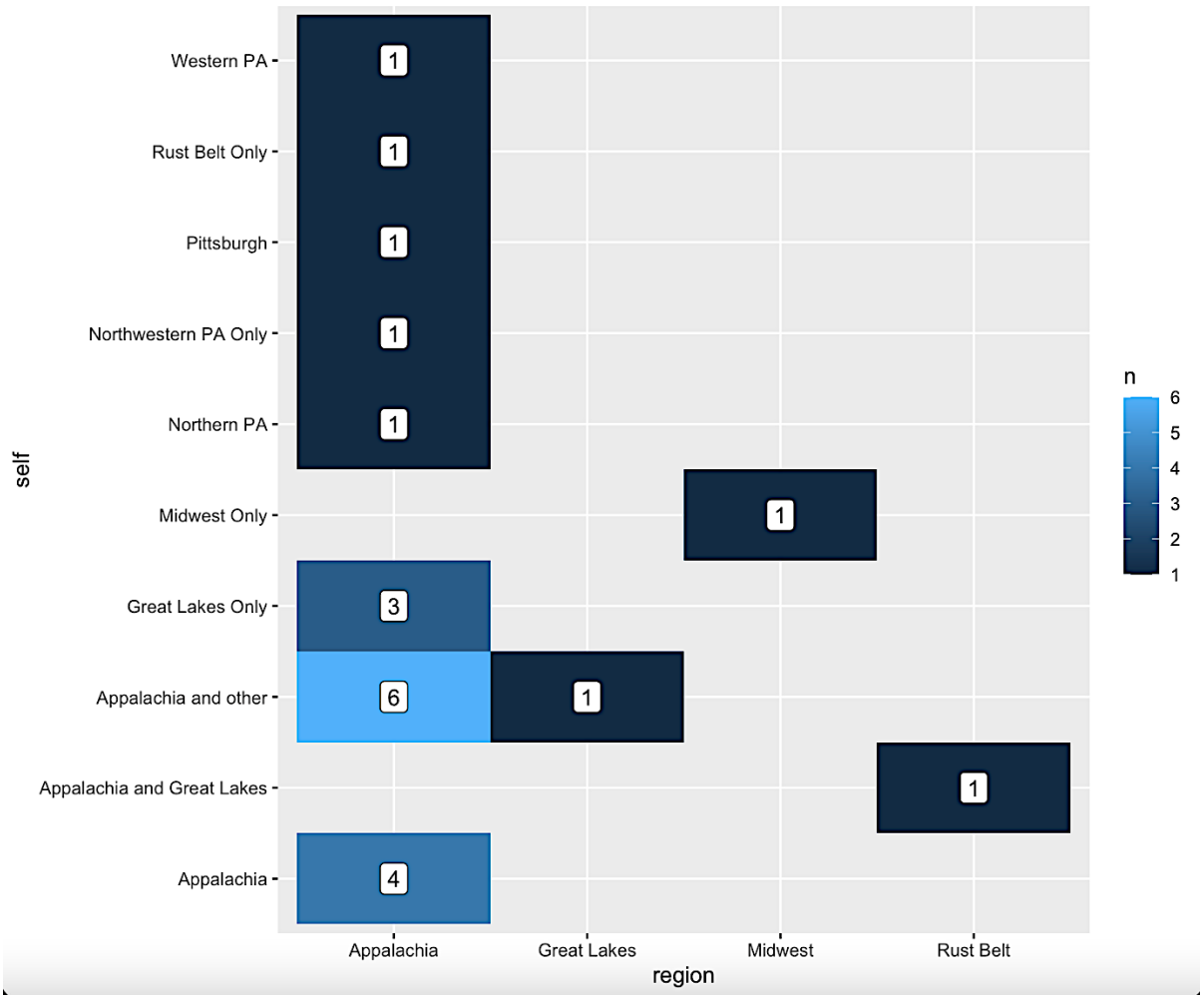

Figure 13. Cross Tabulation of Perceptual Comments

To further demonstrate these regional identity groupings, a cross-tabulation chart (Figure 13) was created highlighting the groupings of informants based on their perceptual responses to the location of the region and the speakers' self-regional identification as seen in Tables 13 and 14. One can see the 4 informants (as indicated by n) that indicated that PA is a part of Appalachia and regionally identified as such. Along with this, 6 informants located PA as a part of Appalachia, but regionally identified Appalachia and other. 8 informants located PA as a part of Appalachia, but 3 identified as a part of the Great Lakes, 3 identified as different parts of PA, and 1 identified as a part

of the Rust Belt. From this visualization, 18 informants located PA as a part of Appalachia, but all had diverse regional identities. 1 informant located PA as a part of the Great Lakes, but regional identified as Appalachian and other. 7 informants in total identified as Appalachian and other, but only 1 did not locate PA as a part of Appalachia. In addition to informants identifying as Appalachia and other, but labeling the region by a different term, was one informant who labeled the region as Rust Belt but identified as Appalachian and specifically Great Lakes. Lastly, one informant consistently identified the region of PA as Midwest and identified as such. Overall, by combining these perceptual comments, one can see the varying perceived locations of Appalachia and the varying identities across the informants. From this combination, there seems to be an ambiguous, but also fluid identity within these areas that is immensely diverse and varies on multiple contexts. These identities seem to correspond to the ambiguity of a singular regional term for these counties in PA in relation to the state and surrounding areas within the United States. It begs the question of whether this ambiguous identity is attached to the individuals within the region, the region itself, or both. Along with this, it shows a prevalence of Appalachia as a location and identity that varies in infinite contexts. Further analysis into the discursive comments by informants when asked to describe what comes to mind when discussing Appalachia may begin to shed light on this ambiguous Appalachianness and orientations that informants have toward this Appalachianness.

The discursive perceptual comments from the informants show a variation in perceptions of Appalachia, like the varying lens of Appalachia and ARC counties in PA (mentioned in the literature review) seen in society, pop culture, and literature. Appalachia

is seen as "a pejorative term that no one uses anymore or would identify as" (PA005) and as an exciting region because there is finally "a desire and necessity to care about Appalachia because everything impactful has come from these areas along with new up and coming research" (PA021). Speakers also comment, when asked about what they think of when they hear about the region, on the numerous perspectives and stereotypes seen within society about Appalachia such as an area of "poverty and forgotten industry" (PA003) with several "hardships"(PA022) where "moonshine drinkin" (PA014) and "drug addicted" (PA19, PA021) "hillbillies" (PA013), "almost rednecks and in-between southerners" (PA019) are "separating themselves culturally" (PA008) while playing "banjos on their porch" (PA002) and "fiddle music"(PA022) all in an area that they call "home and are closely tied to" (PA002), PA004, PA008, PA017, PA022). All acknowledging and discussing the stereotypes within Appalachia as a whole region, but also when discussing specifically Northern Appalachia.

While they also describe and acknowledge these stereotypes, informants who did identify as Appalachian, also express their Appalachian identity that "people are playin' banjos on the front steps of porches my hometown too" (PA002) and "we faced industrial devastation that will never be rebuilt" (PA010) along with wealthy outsiders "destroying the land, going to the bank, and then never being seen again" (PA011). The informants feel the "sense of home" (PA002 & PA004) and identify as part of the "hardworking" (PA PA002, PA003, PA021) "honest folk that value nature and family"(PA 002, PA003, PA004, PA010). Within their discursive comments, they also comment that Appalachia look different here, it looks "weird," (PA009) and doesn't seem "so special" (PA009 & PA019) compared to others. Maybe a "different subset of Appalachia" (PA022), but still

Appalachia. While these are only a few of the discursive comments, they begin to show the multiple orientations towards this Appalachianness within these PA counties and a wide variety of speaker perceptions of their own region.

By transforming these discursive comments in a makeshift corpus, Biterm Topic Modeling was conducted on all the discursive comments to see the connection between and across these regional labels described and discussed by informants. The discursive comments, a majority of which are shown above throughout the analysis, were all compiled into a spreadsheet in separate lines that were numbered in chronological order to compare the terminology used by all the informants to describe these regional labels. This analysis will not be focusing on specific discursive comments by individual informants, but rather compile all the comments into one output. After compiling these comments into a spreadsheet, they were then processed to parse and filter words. Through this process, tokenization occurred, and all function words were removed, so only content words were remaining. After this, the file was stemmed to create a standardized and consistent output of terms that was not repetitive. Through this process, term frequency analysis and topic modeling were able to be conducted. Within the topic modeling, topics were extracted based on co-occurrence patterns (bi-terms) across all the discursive comments. These extracted topics represent semantically related terms and are considered as a probability distribution of words within a single document (in this case the single document is all the discursive comments compiled into one spreadsheet). These probabilities from the co-occurrences indicate the topic of each grouping. The highest probabilities are the highest co-occurrences of the term in each topic and serve as a label for the topic. Ten topics were extracted from this process and are labeled based on the top 3 highest probabilities within

in extracted topic. These topics were “people,” “geographic features,” “features of housing,” “industry,” “Great Lakes,” “history,” “Rust Belt” “terms and words,” “home and family,” and “Midwest.” Out of the ten topics, the term Appalachia is noted within 5 of them at varying probabilities. Through these topics (see Figure 14), one can see these co-occurrences of Appalachia within these topics.

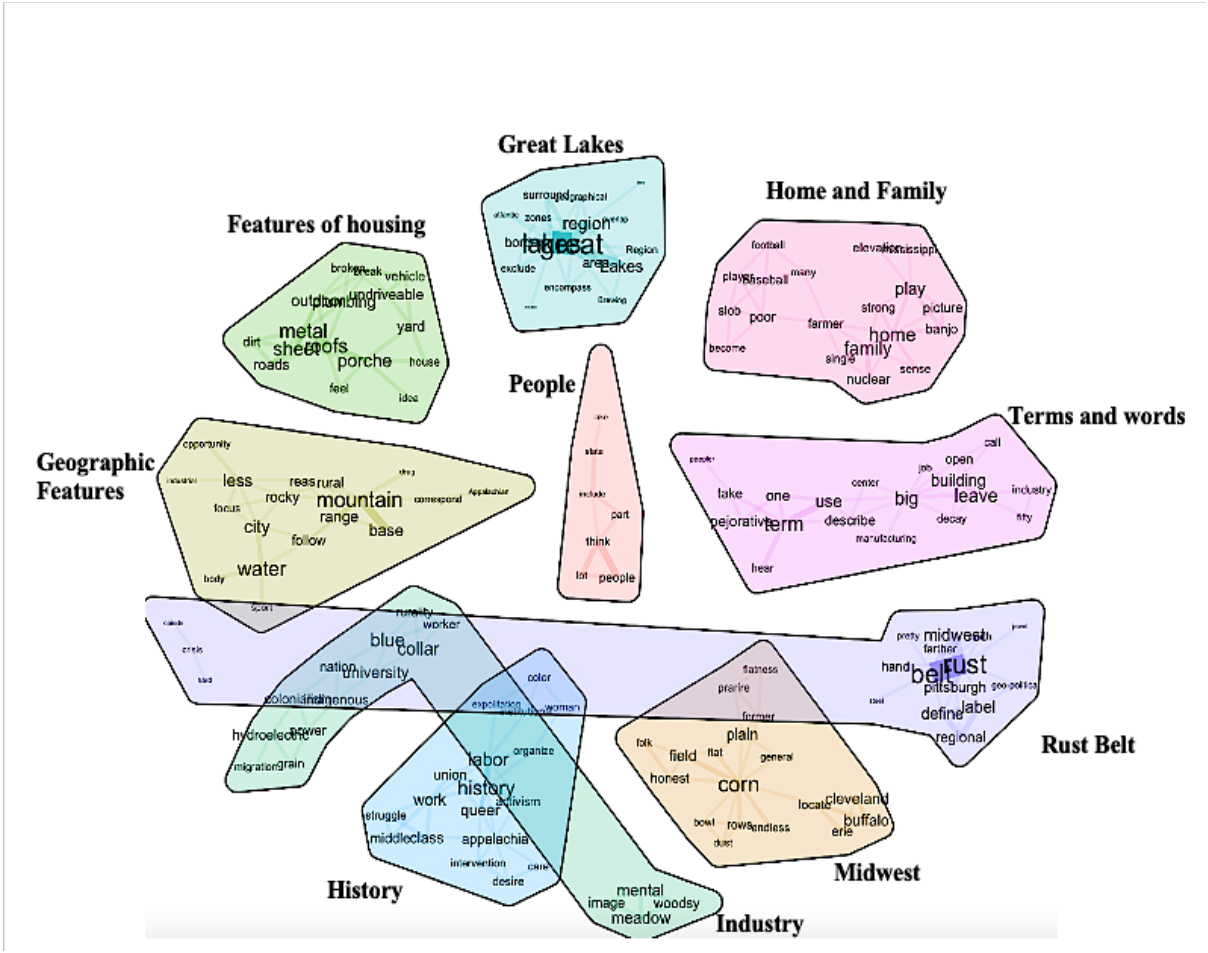

Figure 14. Biterm Topic Modeling of Discursive Comments

From this visual, there is a clear connection between terms used by speakers to describe Appalachia, the Midwest, and the Rust Belt. This clear connection is shown by the overlapping of terms (and visual shape distribution) co-occurring within the topics

extracted and the term “Appalachia” and “Appalachian” appearing multiple times each of these categories. While these three topics are overlapping, the Great Lakes region does not seem to have the same number of co-occurrences and is not overlapping, even though informants were labeling the region and themselves as Appalachia/Appalachian and Great Lakes. This could prove as evidence that while there is a combination of this perceived location and regional identity, the descriptions of both regions and regional identities are vastly different when speakers discuss it, furthering adding to the ambiguity of which term is fitting or if there is an alternative that is more representative of their own regional identity. The similar descriptions of Appalachia, the Midwest, and the Rust Belt could possibly be an explanation for the ambiguous and fluid identities since the speakers themselves describe these regions in a similar manner through these topics. These similar descriptions relate back to regional perceptions of the region, people in the region, and discursive comments made by the informants to explain the locations, perceptions of these regions, and their own regional identity. While the objective of this investigation is not to solve the mystery of these various identities, these similar descriptions of Appalachia, Midwest, and Rust Belt could be a factor creating this ambiguity and multiplicity of the speakers’ own regional identities. This also could be an explanation and lead to a discussion about if the speakers’ identities are ambiguous or if the identity of the region itself is ambiguous, or if this ambiguous Appalachianness is an identity within itself.

CHAPTER 5. CONCLUSION

5.1 Production, Perception, and Identity

From the above data, one can see that highly indexed Appalachian Englishes features are present in these ARC counties in PA historically and currently. There is variation across the frequency of usage of some features demonstrating that some features are not as prevalent currently as they were in the past. While some phonetic features such as offglide reduction and initial voiced theta deletion are not as prevalent now as they were in the past, still these features are being produced and heard among Northern Appalachian speakers. Along with this, morphosyntactic features such as a-prefixing, double modals, and non-standard past tense are not as prevalent, but other morphosyntactic features such as gerund variation are occurring frequently in addition to the varying lexical items. Demonstrating the production of Appalachian Englishes features in present speakers across various ages, education, and counties.

Along with present day interviews showing Appalachian Englishes production, one can see that speakers are acknowledging their inclusion in Appalachia when locating Appalachia with variation across speakers perceiving themselves as Appalachian. PA speakers discuss various similarities pertaining to geography, social variables, politics, and perceptions to Appalachia, but still there is a variety of fluid regional identities among speakers. This raises the question if this ambiguity across speakers, pertaining to regional identity, is ambiguous based on the regional identity of speakers, the region itself, or if this ambiguous Appalachianness is an identity within itself. In addition, this sheds light that this ambiguous Appalachianness may not even be ambiguous Appalachianness, but rather

ambiguity tied to the perceived location of this region and any identity labels that pertains to it.

While this is a question for individuals themselves and certainly not a question that a study of this size could clearly answer, it could be many factors that entertain each of these issues of ambiguity. These factors, just like the many definitions of Appalachia, could vary depending on just about any single context. The reasonings for this variation are a mystery able to be solved by a larger dataset with a larger scope and representativeness, but from this study, the perceptions of Appalachia and influences from other regional boundaries seem to possibly blur these identities and create a fluid ambiguous identity with only some informants identifying as only Appalachian, others including another term to combine with Appalachian, hinting at ambiguous Appalachianess of their own identity or the region itself. While others choose another term without any connection to Appalachia. There still seems to ambiguity present across speakers who labeled themselves as other regional terms, struggling to be certain if these labels even accurately apply and represent themselves and their hometown, but choosing one to identify as for the purposes of this interview. This ambiguity could be based in the overlapping of regional identities and boundaries or rather, an ambiguity based on the absence of identities tied to the region. This question of what an *ambiguous identity* means within the context of this study demonstrates multiple possibilities and theories that require further investigation into this region and the speakers that call it home. Through these varying identities one can see the diversity within this region, further affirming that Appalachia and its subregions are not a homogenous entity. With respect to how PA fits into the region that is Appalachia, this linguistic and perceptual investigation into PA speakers in these areas, historically and currently, shows

the immense diversity of voices, fluid identities, varying regional perceptions, and wide range of orientations that demonstrate the variety of ideas, stances, states of mind, and identities that make up this wonderful region of Appalachia.

CHAPTER 6. LIMITATIONS AND IMPLICATIONS

6.1 Limitations

Due to the nature of this study and time frame of data collection and analysis, there are many future considerations of this data that need to be taken into consideration with respect to the methodology. For a better investigation into historical perceptual information, looking into the LAP audio recordings and full interview transcripts, may reveal more perceptual and regional identity information about the LAP informants. This would indeed give a clearer explanation into the connection between language, place, and identity on a historical scale within these PA counties. Along with this, the LAMSAS audio files would obtain a better picture of the use of these linguistic features throughout the whole LAP interview, giving a tremendous insight about natural speech across these informants. Lastly, a more expansive examination of the LAMSAS informants of PA, could also expose a larger set of Appalachian Englishes features that are not covered within this study. Along with data collection limitations, there is limitation within the diversity of the demographic information of the informants. The LAMSAS PA informants are overwhelmingly male (with only 7 female informants) and all informants are white, per the LAP informant metadata. These factors affect the dataset greatly when using social variables, but future studies may look to fill in these demographic gaps in a variety of ways. Within LAMSAS, these methodological limitations from this study in no way are representative of the breadth and depth of the LAP data across all regions, which continues to be an important resource for regional and social variation.

In addition to the historical data, there are many limitations that deal with the time frame and structure of the current interviews of these informants. Due to the time

frame and the COVID-19 pandemic, finding volunteers to participate in the interviews was difficult and affected the representativeness of the counties within this region. Only 7 out of the 52 ARC counties in PA were represented in this study, which was a result of many factors, but future studies may want to ensure the counties are much more varied and implement a variety of marketing and advertising techniques to increase the number of participants. Due to these limitations, the PA counties being studied are centrally located within Western, PA and Central, PA, not indicative of the 45 other ARC counties. In addition to representativeness of counties being a limitation throughout this study, there is also a lack of diversity in respect to ethnicity. While this data, compared to the LAP, is diverse across age, education, and gender, there is an alarming lack of diversity in respect to ethnicity, similar to the LAP. While this is a pilot investigation into these areas linguistically and perceptually that has intentions of being continued on a much larger scale, the representativeness of PA counties and the speakers living in these counties will need to be made priority for this data to be effective and accurate in showcasing and investigating these voices and identities.

6.2 Implications

While this is truly only the beginning of showcasing these voices and identities within PA, this study has several crucial implications and exciting possibilities within various fields of social sciences, humanities, literature, information sciences, and many more. With the study being highly interdisciplinary, it can be built upon to create future progress looking at Appalachian PA counties in a variety of contexts and perspectives. These contexts and perspectives can range from historical studies, linguistic studies,

cultural studies, oral histories, and even investigations into rootedness and the fluidity of rootedness within these Northern Appalachian regions (Reed, 2016). This study can also serve as a beginning methodological framework on procedures of combining linguistic and perceptual data into a survey-style interviews on a much smaller scale that includes regional language variation, but also variation of perceptions and identities. Along with this, altering the regional lens could provide insight to the other regional identities and converging label boundaries that speakers identify as within these regions and other regions across the United States.

Experiencing and learning these speakers and their language varieties will begin to characterize and locate Appalachian PA and the ambiguous and fluid Appalachianness within it, representing their own Appalachia even when it is not labeled as such. Researchers need to rely on historical and current data from these speakers, distancing themselves from the outsider views, opinions, and perspectives that perpetuate generalizations and stereotyping. By embracing this fluid and ambiguous Appalachianness, researchers and society will begin how PA fits into the region that is Appalachia.

APPENDICES

APPENDIX 1. DEMOGRAPHIC INFORMATION QUESTIONNAIRE

Demographic Information

Age:

What is the highest level of education completed?

Occupation:

City:

State:

Hometown:

APPENDIX 2. BRIEF AUTOBIOGRAPHICAL ORAL HISTORY

Discussion topics:

Tell me about yourself: What are your hobbies? Did you do these hobbies when you were younger?

What is a childhood memory you have in PA?

Did you ever do anything fun in the winter as a kid? What about in the summer? Spring? Fall?

APPENDIX 3. PRODUCTION AND PERCEPTION INTERVIEW QUESTIONS

Production Questions:

Geography

What are words for different roads? What if the roads are main roads in towns? What about in cities? What if they are roads in rural areas?

Are there different terms for roads that you must pay money to use?

What are some main roads around here?

What are different bodies of water in PA?

What do you call small thin bodies of water? Are there different names? Does the name change if it is salt water? What about fresh water?

What if the water is moving, is there different names for that? Does the name change if it is salt water? What about fresh water?

What about larger thin bodies of water? Are there different names? Does the name change if it is salt water? What about fresh water?

What if the body of water is moving, is there different names for that? Does the name change if it is salt water? What about fresh water?

What do you call large bodies of water that normally seen on peoples' property? Are there different names if they are man-made or natural? What if it isn't not on someone's property?

What are some common small and large bodies of water in the area? Do they have different names?

What do you call land that is higher in elevation and rocky?

What if the land is high, but not rocky?

What if the land is high and flat?

What do you call flat low-lying land with no elevation?

What about flat low-lying land with grass?

What do you call areas that are between elevation?

What are some different geographic features around PA?

What are some mountain ranges around the Northeast? In PA?

Language Features

If you want someone to bring you a book that is yours, what would you say? Can you bring___?

If you were having a conversation with someone yesterday, you would say I____ to them yesterday

What about if you are having a conversation tomorrow? I _____to them tomorrow

What about if you are having a conversation right now? I _____to them right now

If you are about to get up and make some food, what would you say? I___ to make some food

What would you say if you are not done doing something and someone asks if you are done or not?

If you saw two people grab something and someone asks you who grabbed it, what would say?

If a dog is next to you on his and lays down right now, you would the dog is___.

If wind started to blow hard and speed up, you would say the wind is_____.

If carolers came to the house and started to sing, you would say the carolers are_____.

If it was storming bad, but then it started to clear up, you would say the weather is _____up.

If you are about to get up and go somewhere you would say, I am _____.

If someone asks you do to do something and you don't want to be mean, but don't want to do it, what would you say?

If it is the 25th of December, what greeting would you use to greet family and friends?

If you were supposed to do something, but it slipped your mind and you did not, what would you say?

If you are telling someone about a story of when you were doing something and then something else happened while you were doing something, how would you tell the story? Ex.) ___I lost the remote, the dog jumped on the couch

If you tell someone you are going to do something, but you aren't sure when you are going to do it, you would say? I will do it _____ I can.

As someone is leaving your house, when you leave if you hope to see them again, you would say? Does this change if more than one person is leaving? Have you heard other people use different words?

What do you call illegal alcohol? Have you heard other terms for this?

Past tense of climb? Have you heard other terms?

Past tense of dive? Have you heard other terms?

What is the word for the animal that is smaller than a squirrel?

What do you call milk that has gone bad?

If you have some pencils that you wanted to keep together, but you did not have a bag, what common household item would you use?

If you need to do laundry, you would say the clothes ___.

Can you think of other terms or phrases that you have heard from this area that you do not hear in other places?

Places and People

If someone is from Ireland, what would say they are? If they are a man? If they are a woman?

What do you call someone that is very outdoorsy and likes to be in the outdoors?

If people do not have a lot financially, what would you say to describe that they do not own a lot?

What are the states that border PA?

What are major cities in PA?

What are some main regions in the US?

Perceptual Questions:

Where is Appalachia in your opinion?

What states are in Appalachia?

Why did you pick these states?

Do you think someone else would have a different opinion? If so, why?

If you were to split up Appalachia into northern, central, and southern, how would you split it up? Why did you pick these boundaries?

When you think of this region, what are some things that come to mind?

What areas are a part of the Midwest Region? Do you think someone else would say otherwise?
When you think of this region, what are some things that come to mind?

What areas are a part the Rust Belt Region? Do you think someone else would say otherwise?
When you think of this region, what are some things that come to mind?

What areas are a part of the Great Lakes Region? Do you think someone else would say otherwise? When you think of this region, what are some things that come to mind?

What do you call the region you are from?

Do you think someone who is not from the region would call it a different name?

What are other names you have heard this region of PA called?

APPENDIX 4. BLANK MAP FOR INTERVIEWS

REFERENCES

Anderson, B. L., Cramer, J., Dumas, B. K., Flanigan, B. O., & Montgomery, M. (2014). Needed research on the Englishes of Appalachia. Southern Journal of Linguistics, 38(1), 1-30.

Appalachian Regional Commission. (2020). *About the Appalachian Regional Commission*. Appalachian Regional Commission. Retrieved March 28, 2022, from https://www.arc.gov/about-the-appalachian-regional-commission/ .

Appalachian Regional Commission. (2021, November 15). *Pennsylvania Strategy Plan*. Appalachian Regional Commission. Retrieved December 11, 2021, from https://www.arc.gov/resource/pennsylvania-strategy-statement/.

Appalachian Regional Commission. (2015). *Pennsylvania 2015-2018 Appalachian Development Plan*. Appalachian Regional Commission. Retrieved March 28, 2022, from https:// https://www.arc.gov/resource/pennsylvania-appalachian-development-plan/ .

Autumn C. James, & Ryan D. James. (2016). Economic Integration and the Appalachian Experience in the Rust Belt-Sun Belt Transition. *Journal of Appalachian Studies*, *22*(2), 160–186. https://doi.org/10.5406/jappastud.22.2.0160

Barnes, J. H., and Sevon, W. D., (2014), The geological story of Pennsylvania (4th ed.): Pennsylvania Geological Survey, 4th ser., Educational Series 4, 44 p. 2-43

Cooper, C. A., Knotts, H. G., & Elders, K. L. (2011). A Geography of Appalachian Identity. *Southeastern Geographer*, *51*(3), 457–472. http://www.jstor.org/stable/26228971

Cramer, J. (2020). Identity and Representation in Appalachia: Perceptions in and of Appalachia, Its People and Its Languages. In *Appalachian Englishes in the twenty-first century* (pp. 69–83).(K. Hazen, Ed.). West Virginia University Press.

Cramer, Jennifer. (2012). What and where is Appalachian English? The perceptions of outsiders. Paper presented at the annual meeting of the Appalachian Studies Association, Indiana, PA, March 23–25.

———. (2013). Styles, stereotypes, and the South: Constructing identities at the linguistic border. *American Speech* 88 (2): 144–67.

———. (2014). Is Shakespeare still in the holler? The death of a language myth. *Southern Journal of Linguistics* 38 (1): 195–207.

———. (2018). Perceptions of Appalachian English in Kentucky. *Journal of Appalachian Studies*, *24*(1), 45–71. https://doi.org/10.5406/jappastud.24.1.0045

Drake, R. B. (2001). The "Discovery" of Appalachia. In *A History of Appalachia* (pp. 119–130). University Press of Kentucky. http://www.jstor.org/stable/j.ctt2jcv7t.12

Erie Apparel. (2022). *It's Okay to Love Erie Shirt*. Erie Apparel . Retrieved March 30, 2022, from https://erieapparel.co/products/its-okay-to-love-erie-tee-1.

Evanini, Keelan (2008) "A shift of allegiance: The case of Erie and the North / Midland boundary," University of Pennsylvania Working Papers in Linguistics: Vol. 14 : Iss. 2 , Article 10. Available at: https://repository.upenn.edu/pwpl/vol14/iss2/10

Ferrence, M. J. (2019). *Appalachia north: A memoir*. Morgantown: West Virginia University Press.

Ferrence, M. (2012). You Are and You Ain't: Story and Literature as Redneck Resistance. *Journal of Appalachian Studies*, *18*(1/2), 113–130. http://www.jstor.org/stable/23337710

Geologic map of Pennsylvania (2007). Modified from Pennsylvania Geological Survey, 2007, Geologic map of Pennsylvania (3rd ed., 4th printing, slightly revised): Pennsylvania Geological Survey, 4th ser., Map 7, scale 1:2,000,000, and Sevon, W. D., compiler, 2000, Physiographic provinces of Pennsylvania (4th ed.): Pennsylvania Geological Survey, 4th ser., Map 13, scale 1:2,000,000.

Hanks, W. F. (1999). Indexicality. *Journal of Linguistic Anthropology*, *9*(1/2), 124–126. http://www.jstor.org/stable/43102444

Hasty, J. D., & Childs, B. (2021). Investigating Appalachian Englishes: Subregional Variation in the New Appalachia. *Journal of Appalachian Studies*, *27*(1), 69–88. https://www.jstor.org/stable/10.5406/jappastud.27.1.0069

Hasty, J. D. (2020). Just What and Where are Appalachian Englishes? In *Appalachian Englishes in the twenty-first century* (pp. 3–19). (K. Hazen, Ed.). West Virginia University Press.

Hayes, A. (2017). Place, Pedagogy, and Literacy in Appalachia. *English Education*, *50*(1), 72–89. http://www.jstor.org/stable/26492555

Hazen, Kirk, and Ellen Fluharty. 2004. Defining Appalachian English. In *Linguistic diversity in the South: Changing codes, practices, and ideology*, edited by Margaret Bender, 50–65. Athens: University of Georgia Press.

Hazen, K. (Ed.). (2020). *Appalachian englishes in the twenty-first century*. West Virginia University Press.

Hazen, K. (2020). Grammar Across Appalachia. In *Appalachian Englishes in the twenty-first century* (pp. 36–51). (K. Hazen, Ed.). West Virginia University Press.

House of Representatives, Congress. (2007, December 30). *40 U.S.C. IV – Appalachian Regional Development*. [Government]. U.S. Government Publishing Office. https://www.govinfo.gov/app/details/USCODE-2007-title40/USCODE-2007-title40-subtitleIV

Johnstone, & Pollak, C. (2016). Mobilities, Materialities, and the Changing Meanings of Pittsburgh Speech. Journal of English Linguistics, 44(3), 254–275. https://doi.org/10.1177/0075424216654519Johnstone, Andrus, J., & Danielson, A. E. (2006). Mobility, Indexicality, and the Enregisterment of "Pittsburghese." Journal of English Linguistics, 34(2), 77–104. https://doi.org/10.1177/0075424206290692

Kretzschmar, W. A., McDavid, V. G., Lerud, T. K., & Johnson, E. (Eds.). (1994). *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. The University of Chicago Press.

Kurath, H., Hansen, M. L., Bloch, B., Bloch, J., & In Kurath, H. (1939). *Handbook of the linguistic geography of New England*. Providence: Brown University.

Lewis, R. L., & Eller, R. D. (1999). Beyond Isolation and Homogeneity: Diversity and the History of Appalachia. In D. B. Billings, G. Norman, & K. Ledford (Eds.), *Back Talk from Appalachia: Confronting Stereotypes* (pp. 21–44). University Press of Kentucky. http://www.jstor.org/stable/j.ctt2jcp5m.6

Linguistic Atlas Project. Lexington, KY: University of Kentucky. https://www.linguisticatlasproject.org/.

Luhman, R. (1990). Appalachian English Stereotypes: Language Attitudes in Kentucky. *Language in Society*, *19*(3), 331–348. http://www.jstor.org/stable/4168150

Meade, M. R. (2019). *In the shadow of "king coal": Memory, media, identity, and culture in the post-industrial pennsylvania anthracite region* (Order No. 13881312). Available from ProQuest Dissertations & Theses Global. (2273838297). http://ezproxy.uky.edu/login?url=https://www.proquest.com/dissertations-theses/shadow-king-coal-memory-media-identity-culture/docview/2273838297/se-2?accountid=11836

McCann, E. J. (1998). Mapping Appalachia: Toward a Critical Understanding. *Journal of Appalachian Studies*, *4*(1), 87–113. http://www.jstor.org/stable/41446367

Montgomery, M. (2007). Appalachian English. In M. MONTGOMERY & E. JOHNSON (Eds.), *The New Encyclopedia of Southern Culture: Volume 5: Language* (pp. 42–45). University of North Carolina Press. http://www.jstor.org/stable/10.5149/9781469616629_montgomery.10

Montgomery, M. (2013). The Historical Background and Nature of the Englishes of Appalachia. In A. D. Clark & N. M. Hayward (Eds.), *Talking Appalachian: Voice, Identity, and Community* (pp. 25–53). University Press of Kentucky. http://www.jstor.org/stable/j.ctt2jckk2.5

O'Neill, B. (2009). The Paris of Appalachia : Pittsburgh in the twenty-first century . Carnegie Mellon University Press.

O'Neill, B. (2011, December 8). *Yes, we and yinz are part of Appalachia*. Pittsburgh Post-Gazette. Retrieved December 16, 2021, from https://www.post-gazette.com/opinion/brian-oneill/2011/12/08/Yes-we-and-yinz-are-part-of-Appalachia/stories/201112080418

Pennsylvania Department of Conservation and Natural Resources. (n.d.). PaGEODE. Retrieved December 15, 2021, from https://www.gis.dcnr.state.pa.us/pageode/

Raitz, K. B., & Ulack, R. (1981). Cognitive Maps of Appalachia. *Geographical Review*, *71*(2), 201–213. https://doi.org/10.2307/214188

Reed, P. (2020). Phonological Possibilities in Appalachian Englishes. In *Appalachian Englishes in the twenty-first century* (pp. 20–35). (K. Hazen, Ed.). West Virginia University Press.

Reed, P. E. (2016). Sounding Appalachian:/ai/monophthongization, rising pitch accents, and rootedness. Unpublished doctoral dissertation. University of South Carolina, Columbia.

Sandow, R. M., & McKinney, G. B. (2010). "Grudges and Loyalties Die So Slowly": Contested Memories of the Civil War in Pennsylvania's Appalachia. In A. L. SLAP (Ed.), *Reconstructing Appalachia: The Civil War's Aftermath* (pp. 269–292). University Press of Kentucky. http://www.jstor.org/stable/j.ctt2jctwj.14

Shuylkill County Historical Society. (2021). *I Speak Two Languages*. Shuylkill County Historical Society. Retrieved March 30, 2022, from http://schuylkillhistory.org/product/i-speak-two-languages/ .

Tamasi, S., Antieau, L. (2014). Language and Linguistic Diversity in the US: An Introduction. United States: Taylor & Francis.

Ulack, R., & Raitz, K. (1981). Appalachia: A Comparison of the Cognitive and Appalachian Regional Commission Regions. *Southeastern Geographer*, *21*(1), 40–53. http://www.jstor.org/stable/44370716

Ulack, R., & Raitz, K. (1982). Perceptions of Appalachia. Environment and Behavior, 14(6), 725–752. https://doi.org/10.1177/0013916582146005

Watt, D., & Llamas, C. (2017). Identifying Places: The Role of Borders. In C. Montgomery & E. Moore (Eds.), *Language and a Sense of Place: Studies in Language and Region* (pp. 191-214). Cambridge: Cambridge University Press. https://doi:10.1017/9781316162477.011

Wolfram, W., & Schilling-Estes, N. (2016). *American English: Dialects and variation*. Wiley Blackwell.

VITA

Crissandra George was born in Hawthorn, PA and then spent her childhood between Erie, PA and Hawthorn, PA. She graduated from the University of Mississippi in 2019 and received a Bachelor's in Linguistics, Spanish, and a minor in Swahili. She has worked for the Linguistic Atlas Project for a combined 6 years holding various positions from undergraduate student worker to graduate research assistant. Upon publication of this thesis, she will have completed a M.A. in Linguistic Theory and Typology from the University of Kentucky. She is currently finishing her M.S. in Library and Information Science at the University of Kentucky.