Schwa

Language and Linguistics

Fall 2024 | Issue 31

Schwa: Language and Linguistics is published by the Department of Linguistics and English Language at Brigham Young University. Its contents represent the opinions of the authors and editors and not necessarily those of the faculty or administration of Brigham Young University. We are always looking for submissions and staff members. If you are interested, please visit us at schwa.byu.edu.

© 2024 by Brigham Young University. All rights reserved.

Cover design by Shaylyn Lawyer

Contents

iv Staff

v About *Schwa*

vii Editor's Note

1 Tag Questions and Jane Austen
Naomi Christensen

17 United in Names: A Comparison of Naming Trends in Alabama and Vermont
Elijah Clark

29 Representation of African American English in American Media
Noah Frazier

51 The Nuance of the Serial Comma
Abigail Christensen

61 Reactions to Sexual Taboo Language in Married and Unmarried Latter-day Saints
Alayna Beck, Lia King, Megan Beals, and Lauryn Wilde

73 Linguistic Realization of *Yeísmo Rehilado* in Uruguay
Jacob Badal

Staff

Editor in Chief
Rachel Hart

Managing Editor
Shaylyn Lawyer

Senior Editors
Alyssa Brownrigg
Halli Romero
Ciera Griggs

Staff Editors
Elise Green
Elijah Clark

Design Team
Rachel Hart
Shaylyn Lawyer
Ciera Griggs
Elise Green
Elijah Clark

Faculty Advisor
Dr. Dirk Elzinga

About *Schwa*

We are an academic journal produced by the students of Brigham Young University. Our mission is to increase the amount and accessibility of linguistic scholarship—especially for those without graduate school experience—while simultaneously training editors and designers in the ways of modern publishing. Some of our articles are strictly theoretical and academic. Others are less technical and more personal in nature. Experiments, surveys, corpus analyses, and essays are all acceptable. We have published on all the following subdisciplines of linguistics and more:

- Phonetics, the perception and production of speech sounds
- Phonology, the system of speech sounds used in a given context
- Semantics, the meaning constructs of words and sentences
- Syntax, the structure of permissible and meaningful sentences
- Pragmatics, real-world language use and other speech-related actions
- Sociolinguistics, language variation based on sociological factors
- Psycholinguistics, the cognitive tasks necessary for language
- Fieldwork notes from living in a foreign language-speaking community
- Forensics linguistics, the role of language in law

We are always accepting submissions. Articles on any language are welcome, including cross-linguistic studies, but they must be written in English.

Our staff includes both editors and graphic designers. We extend an open invitation for new staff members. Go to our website at schwa.byu.edu to submit an article or join our staff.

Editor's Note

Human beings are truly amazing. As I learn more about linguistics, I marvel at the complexity of the human brain and its ability to process and create language. But language isn't the only thing we do well. I am constantly in awe of the resilience and perseverance that I see around me—in family members, friends, classmates, and even in myself. This semester, I have been especially grateful for the perseverance of the *Schwa* staff.

It was not an easy semester. Technical issues with our website had us scrambling for submissions at first, putting us far behind our normal schedule. But through it all, our staff rose to the occasion, working hard to get edits done. And because of that hard work, we managed to get back on track. I learned many things this semester, but one of the most important was how much more we can accomplish working together. And for that, I would like to thank each and every one of our editors. I'd especially like to thank my managing editor, Shaylyn, for helping me keep track of everything that needed to be done and stay sane doing it.

We are also immensely grateful for our authors, who have worked hard to implement edits and make deadlines, shaping their articles into what you see today. Thank you for helping us keep *Schwa* alive and thriving.

We are grateful for the Linguistics department and for our professors. We are especially grateful for our faculty advisor, Dr. Elzinga. Thank you for supporting us and helping us learn and grow.

And thank you, reader, for taking the time to read and marvel with us at the capacity of the human brain. We hope you enjoy issue 31 of *Schwa: Language and Linguistics*.

Rachel Hart
Editor in Chief

Tag Questions and Jane Austen

Naomi Christensen

In this paper, Jane Austen's use of tag questions is explored, specifically "is not he/she/it?" and "is he/she/it not?" Using multiple corpora, the author compares Austen's use of these structures with authors who wrote before, during, and after her time. Austen's use of "is not he/she/it?" is found to be significantly higher than all other authors studied. Her increased use does not match other groups of authors and suggests that she may have adopted this phraseology as part of her own unique and innovative style.

"She is loveliness itself. Mr. Knightley, is not she?" (Austen, 1815).
She is, isn't she? Is she not? Innit?

There is some debate among Austen scholars about whether the famous writer was linguistically conservative or innovative. Tieken-Boon van Ostade (2014) says, "Her usage is often idiosyncratic as well as, at times, conservative, though by no means always so." Might Austen's use of tag questions (as shown in the example above) have something to contribute to the debate? To answer that question, we need to first understand how Austen used tag questions. Was her usage typical for her time, was it innovative, or did it reflect an earlier period in the history of English?

English tag questions have taken many different forms since they first showed up in the language in the sixteenth century (Tottie & Hoffmann, 2009). The kinds of tag questions listed in the introductory lines of this article (with the exception of *innit*) can be called canonical tag questions, and are subordinate clauses that are made up of an operator (or helper verb), an optional negative particle, and a personal pronoun. These can be contrasted with invariant tag questions like "right?" (Achiri-Taboh, 2015) or "innit?" (as mentioned above), which is an emerging grammaticalized form of "isn't it?" in Multicultural London English. (Palacios Martãnez, 2015). These invariant tag questions, as their name suggests, do not vary based on the tense or person being referred to. For the purpose of this study, we will only be looking at variable or canonical tag question constructions with the operator *is* that contain the negative particle and that refer to the third person singular. After factoring in variations in syntax, which is the primary focus of this study, we end up with nine possible variables: *is not she*, *is not he*, and *is not it*; *is she not*, *is he not*, and *is it not*; and *isn't it*, *isn't she*, and *isn't he*. It should be noted that because the main clause counterparts of these tag questions are structurally identical, I chose not to filter them out of the data (compare "*Is not she* a fine young woman?" and "How much she must be admired in such a place as this—*is not she?*" (Austen, 1817)).

Regardless of sentence position, the unmarked form of these questions has always been Aux+Pro+Neg, as in "is she not" (Risannen, 1999). The construction Aux+Neg+Pro ("is not she") entered the language in Late Middle English and surged in Early Modern English, making up thirty percent of negative questions.

By Late Modern English (Austen's time), the construction had already started to die out, being replaced by the contraction with the same syntactical order, "isn't she." Eventually, the contracted form became the most prevalent, and the other two forms started to decline (Rissanen, 1999). If we place Austen's writings (1787–1818) in the context of Rissanen's data, we see that Austen was writing in a time of great change for the negative question. The Aux+Neg+Pro construction had been somewhat innovative a century earlier but was almost nonexistent a century later.

Thus, the purpose of this study will be to determine, first, what were tag questions and their main clause counterparts doing in Late Modern English and specifically in the nineteenth century? Second, did Austen use Aux+Neg+Pro (*is not she*) more than her contemporaries? Third, if Austen did use Aux+Neg+Pro more than her contemporaries, does this usage provide evidence of Austen being a linguistically conservative writer?

Methods

I first did a few corpora searches to determine the form of English tag questions over time. I searched tag question variations on Google NGrams, COHA, and BYU-COEME. From there, I used Excel to create visualizations of the uses of different tag question forms from the sixteenth century to present day. I chose BYU-COEME because it has a good number of sources from the Early Modern English period, which will help us understand more about what these forms were doing before Austen. COHA is an ideal corpus for Late Modern English and will provide insight into Austen's time period and beyond. Google NGrams is also a useful tool because it can give us a visual of how a certain structure is used over time by pulling from a variety of print and online sources. Additionally, it has sources from all the relevant time periods, making it a more complete overview. However, because Google NGrams has limitations, including that it does not provide examples and details the way COHA and BYU-COEME do, these searches mostly served to confirm and replicate the corpus searches and the data given by Rissanen.

I then focused on authors from the late modern English period (LModE), which spans the eighteenth and nineteenth centuries, by creating four different literature corpora on AntConc to situate Jane Austen among her contemporaries. I first created an Austen Corpus and then a corpus of works published before Austen,

which I have designated "Pre-Austen Corpus." Next, I included works published the same years as Austen, which I have designated "Austen Contemporaries Corpus." Finally, I put together a corpus of works published shortly after Austen's time, titled "Post-Austen Corpus." I then compared the results of each of these corpora with the results from COHA and BYU-COEME about the trends of the time in general, using all normalized frequencies per million words. A complete list of works in the corpora and search terms used is given in the appendix (figures 2.1 and 2.2, respectively).

Results

Google Ngrams would not allow me to separate results for "is not she" and "isn't she," so the only comparison I could make was between the Aux+Pro+Neg constructions and the Aux+Neg+Pro constructions, without considering whether the negator in the latter was written out or written as a contraction. Despite this, the data given still proved useful in my analysis. Graphs showing usage over time for both syntactic constructions, divided out by pronoun, are shown below. The y-axis represents what percentage of all three word phrases (or trigrams) this particular phrase makes up within their data.

Figure 1.1

Google Ngrams: She

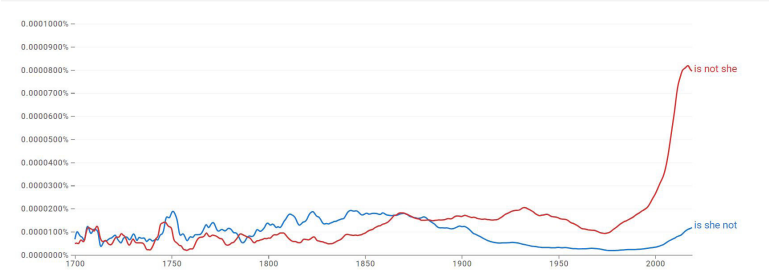


Figure 1.2

Google Ngrams: He

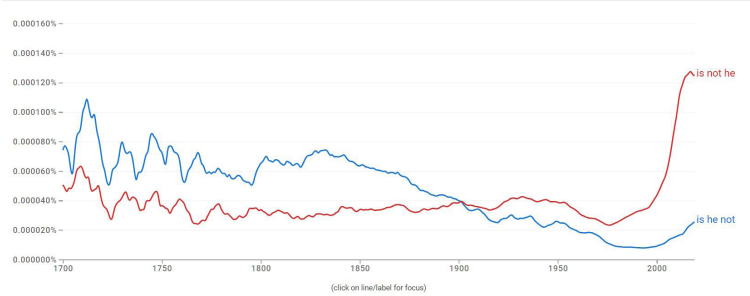
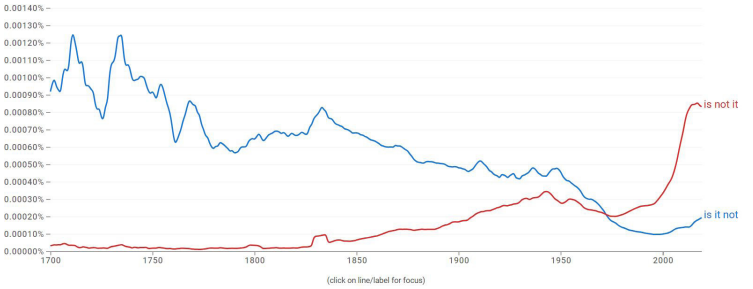


Figure 1.3

Google Ngrams: It



These graphs provide confirming evidence for Risannen's findings that Aux+Pro+Neg was the unmarked and most common form for most of the last few centuries. Aux+Neg+Pro became more common for *he* and *she* around the end of the nineteenth century and the beginning of the twentieth, with the switch for *it* happening a little later. These graphs also show a dramatic spike for the Aux+Neg+Pro constructions at the end of the twentieth century, which may be due to the rise in popularity of the contracted variety.

Risannen's findings were further confirmed by my searches on BYU-COEME, which also found Aux+Pro+Neg to be the most common construction in the Early Modern English period. The Aux+Neg+Pro constructions were present, but in very low quantities, with the contracted variety showing up at the end of the seventeenth century.

Figure 1.4

BYU-COEME Overview

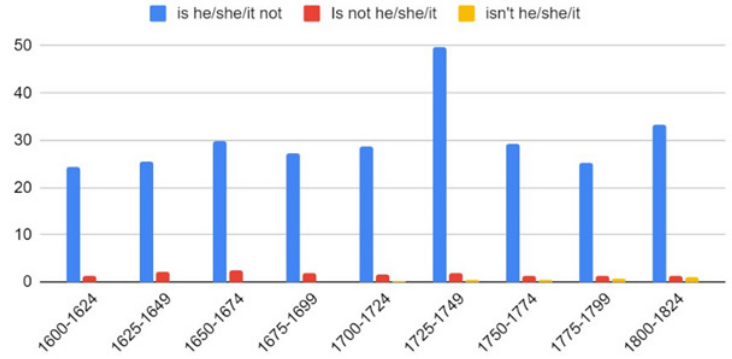


Figure 1.5

BYU-COEME Aux+Pro+Neg

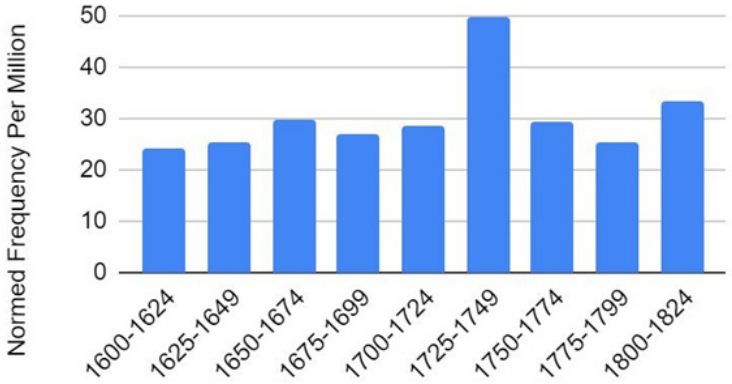


Figure 1.6

BYU-COEME Aux+Neg+Pro

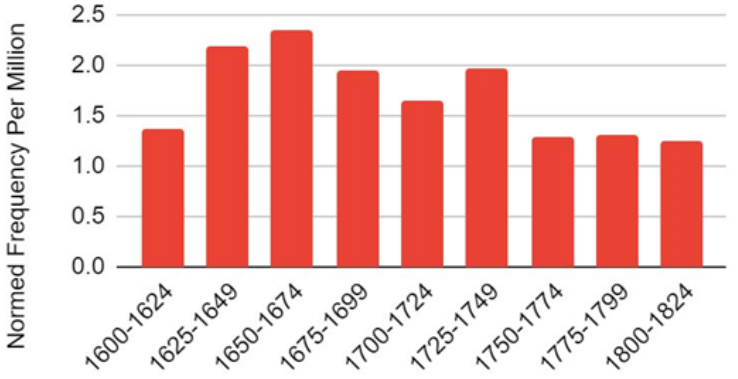
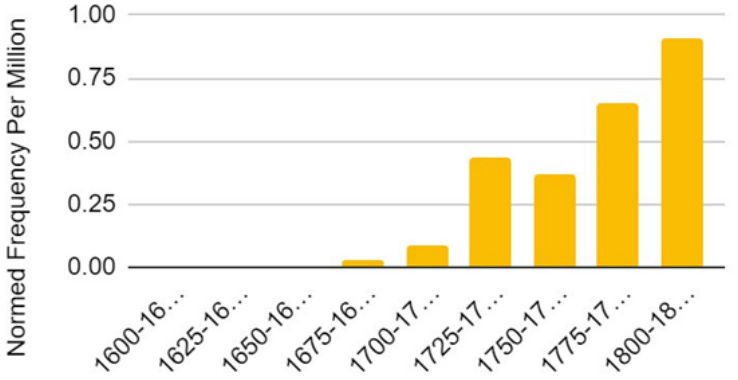


Figure 1.7

BYU-COEME Aux+Neg+Pro Contraction



As shown in the chart above, the use of the contraction *isn't he/she/it* increased dramatically between the seventeenth and nineteenth centuries. In fact, the contraction was the only tag question construction shown to have a statistically significant trend over time, with a p-value of < 0.001. The other two constructions did not show a significant trend for either increasing or decreasing over time.

My final general corpus search was in COHA, which provided data for the Late Modern English period. This search showed statistically significant data for all three constructions, with the

contracted construction increasing and the other two constructions decreasing over time. The graphs and their p-values are given below.

Figure 1.8

COHA Aux+Pro+Neg (p =< 0.001)

SECTION	ALL	1820	1830	1840	1850	1860	1870	1880	1890	1900
FREQ	5352	354	419	457	401	408	370	417	357	339
WORDS (M)	405	7.0	13.7	15.8	16.5	16.9	18.8	20.1	20.4	22.0
PER MIL	13.21	50.71	30.56	28.91	24.25	24.09	19.69	20.78	17.48	15.43
SEE ALL YEARS AT ONCE										

1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
280	299	203	177	168	154	177	116	77	82	97
23.1	25.7	27.7	27.4	28.7	29.1	28.8	29.9	33.1	34.8	35.5
12.12	11.63	7.33	6.46	5.86	5.29	6.14	3.89	2.32	2.35	2.74

Figure 1.9

COHA Aux+Neg+Pro (p =< 0.012)

SECTION	ALL	1820	1830	1840	1850	1860	1870	1880	1890	1900
FREQ	271	4	26	13	23	8	7	24	11	18
WORDS (M)	405	7.0	13.7	15.8	16.5	16.9	18.8	20.1	20.4	22.0
PER MIL	0.67	0.57	1.90	0.82	1.39	0.47	0.37	1.20	0.54	0.82
SEE ALL YEARS AT ONCE										

1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
10	7	12	14	11	13	5	11	18	15	21
23.1	25.7	27.7	27.4	28.7	29.1	28.8	29.9	33.1	34.8	35.5
0.43	0.27	0.43	0.51	0.38	0.45	0.17	0.37	0.54	0.43	0.59

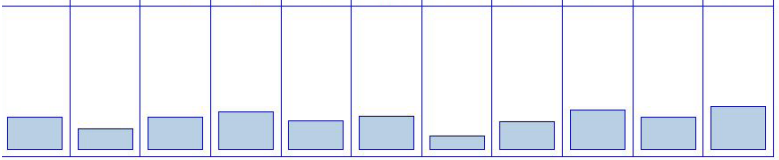
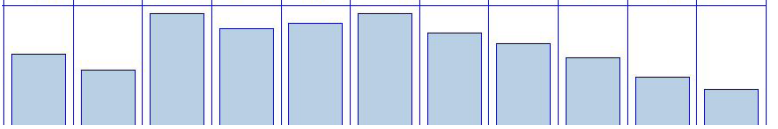


Figure 1.10

Aux+Neg+Pro Contraction (p =< 0.001)

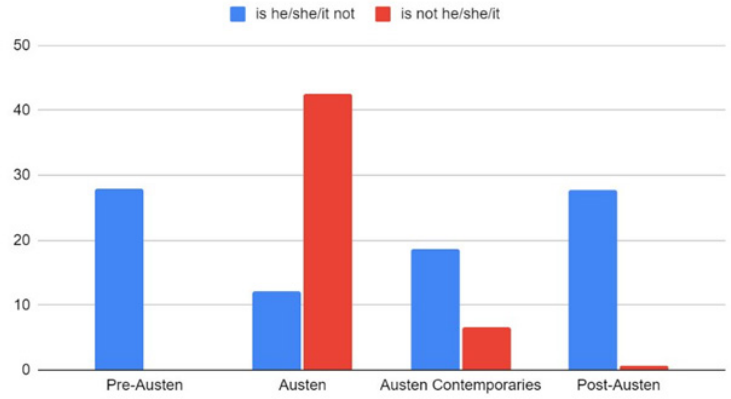
SECTION	ALL	1820	1830	1840	1850	1860	1870	1880	1890	1900
FREQ	21139	5	37	48	210	251	490	390	496	792
WORDS (M)	405	7.0	13.7	15.8	16.5	16.9	18.8	20.1	20.4	22.0
PER MIL	52.20	0.72	2.70	3.04	12.70	14.82	26.08	19.43	24.28	36.04
SEE ALL YEARS AT ONCE										

1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
1226	1116	2115	1854	2031	2229	1883	1767	1677	1378	1144
23.1	25.7	27.7	27.4	28.7	29.1	28.8	29.9	33.1	34.8	35.5
53.07	43.42	76.33	67.66	70.86	76.54	65.32	59.19	50.59	39.57	32.27



Finally, I examined my literature corpora for Austen and her contemporaries on AntConc. The data shown in the graph below is normalized and shows frequency per million words. The contracted construction did not occur in any of my four literature corpora.

Figure 1.11
AntConc



Discussion

Because the data from before Austen (found on BYU-COEME) did not show statistically significant trends over time, it seems that the balance between syntactic constructions of tag questions was fairly stable before Austen's time. The data from COHA, on the other hand, show that *is she not* and *is not she* constructions were both decreasing after Austen's time. These results together seem to suggest that Austen's time was important, but they make it difficult to say with certainty what tag questions were doing in her immediate context. This is because the decades in which she wrote lie at the end of BYU-COEME and at the beginning of COHA, meaning we have to jump between corpora to get a picture of the twenty years before and after her writing. This makes it hard to tell when *is she not* and *is not she* really started decreasing, as the differences between sources the corpora pull from can paint an unclear picture.

What we can say, however, is that Austen's use was atypical for her time period. The data from all corpora involved helps shed light on Jane Austen's stylistic choices. The data from AntConc in figure 1.11 above shows that not only did Austen use Aux+Neg+Pro significantly more than her contemporaries, but she also used it more than she used Aux+Pro+Neg, which does not occur for any of the other literature corpora. This is also surprising when compared to the data from Google Ngrams in figures 1.1–1.3, which

show that Aux+Neg+Pro was the less common construction. Additionally, looking at the 1800 to 1824 category in the data from BYU-COEME (figures 1.4 and 1.6) we can see that Aux-+Neg+Pro occurred less than two times per million tokens in this period, while Aux+Pro+Neg occurred more than thirty times per million token. Because the data from BYU-COEME and the Austen corpus were both normalized, we can directly compare Austen's use to the general trend of the time. The data shows that Austen used Aux+Pro+Neg only a little more than ten times per million tokens, and Aux+Neg+Pro more than forty, which shows a striking difference. These differences could perhaps be explained by the variation in genres of the works that the BYU-COEME corpus draws from; however, genre could not be used to explain the difference in frequencies between Austen and her contemporaries, because these corpora contain only literature.

Contrary to my hypothesis of Austen's use showing a tendency to be linguistically conservative, Austen's preference for Aux+Neg+Pro over Aux+Pro+Neg shows that in some ways, she was ahead of her time. The data from Google Ngrams shows that Aux+Neg+Pro constructions were not used more than Aux+Pro+Neg constructions until (at the earliest, in the data for *she*), the 1870s. This shows that Austen's preference to place the negator before the pronoun would not be adopted by the general populace until fifty years after her death. Even so, her language still would not have been representative of the time in a later period, because it seems that the Aux+Neg+Pro construction did not become the dominant form until it was a contraction. That is to say, *is not she* was never the most popular option, it was *is she not* and then *isn't she*, with *is not she* playing an intermediary role but always remaining in the minority. Interestingly, of the non-Austen literature corpora, the Austen Contemporaries Corpus showed the greatest use of the *is not she* construction, which might suggest that Austen's time, and maybe Austen herself, was pivotal to the syntactic change in this construction and to the reversal of the marked and unmarked forms. Discovering whether Austen actually influenced the syntactic change in these constructions or, if not, where she got her inspiration, is a topic for another study.

Conclusion

In summary, it seems that both the Aux+Neg+Pro and Aux-+Pro+Neg constructions were relatively stable in the centuries preceding Austen's life but started decreasing in favor of the Aux-+Neg+Pro contraction in the nineteenth century. Austen's preference for placing the negator before the pronoun is linguistically innovative (contrary to my hypothesis) and aligns better with language usage fifty or more years after her time. However, by the time this order was most common, it would have been a contraction. Thus, Austen's creating phrases like "But now, I am at such a distance from her—*is not it* hard, Miss Woodhouse?" (Austen, 1815) more than any other similar construction, is unique, both among her contemporaries and in relation to writers before and after her.

References

Achiri-Taboh, Blasius. (2015). A generalized question tag in English: Are English tag questions collapsing? *English Today, 31*(1), 48–54. https://doi.org/10.1017/s0266078414000546

Austen, Jane. (1817). *Emma*. *Project Gutenberg*. https://www.gutenberg.org/ebooks/158

BYU-Corpus of Early Modern English (BYU-COEME). Available online at https://lawcorpus.byu.edu/byucoeme.

Davies, Mark. (2008–) *The Corpus of Historical American English (COHA)*. Available online at https://www.english-corpora.org/coha/.

Google Ngrams. (2010–). Available online at https://books.google.com/ngrams/

Rissanen, Matti. (1999). Isn't it? or is it not? On the order of post-verbal subject and negative particle in the history of English. In Ingrid Tieken-Boon van Ostade, Gunnel Tottie, & Wim van der Wurff (Eds.), *Negation in the history of English*. De Gruyter. https://doi.org/10.1515/9783110806052.189

Tieken-Boon van Ostade, Ingrid. (2014). *In search of Jane Austen: The language of the letters*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199945115.001.0001

Tottie, Gunnel, & Sebastian Hoffmann. (2009). Tag questions in English: The first century. *Journal of English Linguistics, 37*(2), 130–161.

Appendix

Table 2.1

Works in Literature Corpora

1787–1818 Austen Corpus	1740–1787 Publications Pre-Austen Corpus	1787–1818 Publications Austen Contemporaries Corpus	1818–1850 Publications Post-Austen Corpus
Jack and Alice (1787)	*Pamela; or, Virtue Rewarded* by Samuel Richardson (1740)	*A Sicilian Romance* by Ann Radcliffe (1790)	*Ivanhoe* by Sir Walter Scott (1819)
Love and Friendship (1790)	*Clarissa; or, The History of a Young Lady* by Samuel Richardson (1748)	*The Mysteries of Udolpho* by Ann Radcliffe (1794)	*The Private Memoirs and Confessions of a Justified Sinner* by James Hogg (1824)
Lady Susan (1794)	*Tom Jones* by Henry Fielding (1749)	*Caleb Williams* by William Godwin (1794)	*The Last of the Mohicans* by James Fenimore Cooper (1826)
Northanger Abbey (1798–1799)	*The Life and Opinions of Tristram Shandy, Gentleman* by Laurence Sterne (1759–1767)	*The Monk* by Matthew Lewis (1796)	*The Hunchback of Notre-Dame* by Victor Hugo (1831)

1787–1818 Austen Corpus	1740–1787 Publications Pre-Austen Corpus	1787–1818 Publications Austen Contemporaries Corpus	1818–1850 Publications Post-Austen Corpus
Sense and Sensibility (1811)	*Rasselas* by Samuel Johnson (1759)	*Camilla* by Fanny Burney (1796)	*Oliver Twist* by Charles Dickens (1837)
Pride and Prejudice (1813)	*The Castle of Otranto* by Horace Walpole (1764)	*Belinda* by Maria Edgeworth (1801)	*Jane Eyre* by Charlotte Brontë (1847)
Mansfield Park (1814)	*The Vicar of Wakefield* by Oliver Goldsmith (1766)	*The Wild Irish Girl* by Lady Morgan (1806)	*Wuthering Heights* by Emily Brontë (1847)
Emma (1815)	*The History of Emily Montague* by Frances Brooke (1769)	*The Absentee* by Maria Edgeworth (1812)	*Vanity Fair* by William Makepeace Thackeray (1847–1848)
Persuasion (1818)	*Evelina; or, The History of a Young Lady's Entrance into the World* by Fanny Burney (1778)	*Frankenstein* by Mary Shelley (1818)	*The Scarlet Letter* by Nathaniel Hawthorne (1850)

Table 2.2

Search Terms by Corpus

AntConc	\bis not$ (^she|^he^|^it)\b \bis (^she|^he^|^it) not$ \bisn't$ (^she|^he^|^it)\b
COHA	is he/she/it not is not he/she/it is n't he/she/it
BYU-COEME	"is not he" "is not she" "is not it" "is he not" "is she not" "is it not" "isn't he" "isn't she" "isn't it"

United in Names

A Comparison of Naming Trends in Alabama and Vermont

Elijah Clark

This study examines the impact of religion on baby naming practices in the US by comparing the top twenty-five female names in Alabama and Vermont in 2000. Despite regional and religious differences, the top names in both states showed high similarity, mostly originating from Latin or Judeo-Christian traditions. This suggests a national homogenization of naming trends, with national preferences overriding regional and religious factors. Future research could expand to all states and include male names.

How does religion influence the names people choose for their children? This study seeks to explore the differences in naming practices of the top twenty-five names given to female Americans born in the year 2000 in Alabama and Vermont. I identify Alabama as the most religious state and Vermont as the least religious state within the US in that year. By examining US public records to determine the most frequently used names and analyzing the etymology of each name, I aim to identify trends in origins and ties to religion. Through data analysis, I seek to uncover any trends in naming practices affected by religious differences. This research has the potential to shine light on the social influences on what many consider the most precious sound in the world, their own name.

Literature Review

For this research project I wanted to look at the history of naming practices, particularly how religion influences given names in the US. In this literature review section, I paint a broad picture of this history with background knowledge from various sources.

Bozkurt and Mohammadi (2023) looked into the world of horror films, analyzing titles from both Turkey and the US. While seemingly far removed from the realm of everyday names, their work highlights the influence of cultural context. Just as horror films use titles to evoke fear, parents might choose names that reflect their cultural values or aspirations. This concept resonates with Pharr's (1993) exploration of African American naming trends. By examining how African American names diverged from white American trends, Pharr underscores the power of names to reflect cultural identity and social experiences.

Taking a historical perspective, Golub's (2024) study of names on an ancient Egyptian document sheds light on how naming practices evolve over time. This approach can be applied to the United States, where trends in popular baby names have shifted dramatically across the twentieth and twenty-first centuries. Slind's (2021) case study of a Norwegian immigrant community in Washington provides a fascinating example of this evolution. As the community assimilated, their naming choices transitioned from traditional Norwegian names to American ones, reflecting a broader cultural shift.

Valentová and Králik's (2022) work on Slovakian names offers a theoretical framework that strengthens the analysis of American naming trends. They distinguish between official names, nicknames, and popular baby names, each revealing different facets of culture and society. In the US context, popular baby names might reflect broad societal preferences, while nicknames within families offer a glimpse into more personal and intimate aspects of identity.

By drawing connections between these seemingly disparate studies, a richer understanding of American naming practices emerges. Horror film titles, ancient Egyptian documents, and immigrant communities all contribute to the story. These sources provide valuable tools for analyzing how American names have reflected cultural identity, historical shifts, and social dynamics throughout the twentieth and twenty-first centuries.

Very little research has been done on the differences of naming practices in the United States, and thus I found this a compelling topic to study.

Methodology

The US Social Security Administration (SSA) provides a publicly available dataset of the most frequently used names for newborns in the United States, categorized by year or decade. This highly reliable source offered a real-world dataset for my research. While it does not allow access to other demographic information about each child, it does permit me to analyze name frequencies by state and track the most popular female names in each. By using the state-level data from the SSA, I can compare it with other research on social, cultural, and demographic information gathered within each state.

To simplify the analysis and control for the variable of gender, I limited my research to female names within two chosen states. A more comprehensive study could explore the dynamics and potential differences between naming practices for boys and girls.

The data used is strictly from the year 2000. This limitation serves a specific purpose: to compare the two states at the same point in time. By reducing the number of variables, I can obtain a clearer picture of my target variable, which is the influence of religious practices on naming choices. The rationale for choosing data from the year 2000 is the abundance of religious practice data

available from that period. This includes surveys conducted by Gallup and Pew Research, as well as a religious census conducted by the Association of Religion Data Archives (ARDA) in 2000. This publicly available data is useful for confirming information about the religious differences between Alabama and Vermont.

For this project, I chose to focus on the top twenty-five female names in Vermont and Alabama. These two states were selected because I wanted to contrast their usage of Judeo-Christian biblical names.

Vermont consistently ranks as the least religious state in the United States by most metrics. In contrast, Alabama routinely ranks among the most religious states. While there are some exceptions (e.g., Utah's high church attendance rate due to a large Latter-day Saint population), research indicates that Alabama residents, as a collective group, tend to be more religious, identify as religious individuals more frequently, and value prayer more in their lives. While consistent weekly attendance is slightly lower in Alabama than Utah, Alabama boasts higher rates of biweekly or monthly attendance than any other state.

With this background information, Alabama and Vermont appear to be ideal states to test for the influence of religiosity on naming practices. It is important to acknowledge that not everyone in Alabama is religious, nor is Vermont devoid of religious people. This selection simply highlights the largest religious divide.

While there is some debate on many words in anamastics, the subfield of onomastics focused on proper names, there is fortunately very little debate about the origins of the most frequently used words. This was beneficial while looking at the twenty-five most frequently used names in Alabama and Vermont. As such, I was able to utilize several resources, including several websites such as ancestry.com, which polls information from babynames.com, and behindthename.com, as well as the Oxford dictionary of first names. By comparing data from each source, I was able to confirm consistent etymology of all thirty-three names. The notable exception was the name *Samantha*, which all sources acknowledged was under high scrutiny and debate.

The prevailing belief is that *Samantha* is the female equivalent of the name *Samuel*, with a highly productive suffix creating a feminine name. Although all sources acknowledged this widely held belief, they also stated possible exceptions. For the purpose

of this study, I labeled this name as having a Christian biblical origin in line with the view that it is a derivative form of *Samuel*.

Results

After compiling the data into a spreadsheet and tagging names by origin and meaning, I was able to identify some trends.

My first step involved examining the top twenty-five names from both Alabama and Vermont. I created an average rank for these names based on their individual rankings within each state. This average ranking does not consider the percentage or population of names, but rather focuses on the original rank in each state. Interestingly, these lists show remarkable similarity, even though the exact order is not identical. Many of the same names appear within the top ten of both states, as depicted in table 1 below.

Table 1

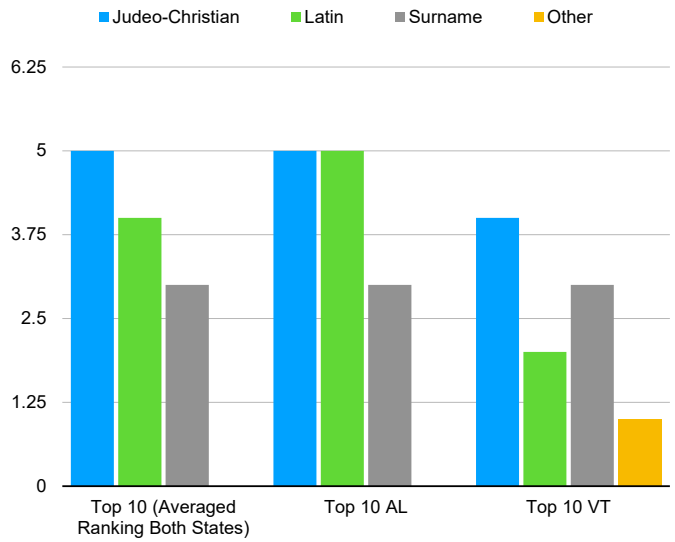
Female Name Rankings

Female Name	Alabama Rank	Vermont Rank	Average Rank
Emily	4	1	1
Hannah	1	4	2
Madison	3	6	3
Anna	2	14	4
Elizabeth	9	7	5
Alexis	6	12	6
Sarah	5	13	7
Taylor	7	11	8
Abigail	18	3	9
Olivia	20	2	10

The overlap extends beyond these ten names. In total, seventeen names appeared on both states' top twenty-five lists. Outside of these overlapping names, there were sixteen names that were unique to the top twenty-five of each individual list. This doesn't necessarily mean these names weren't used in the other state, but simply that they weren't among the top twenty-five most frequently used names in Alabama and Vermont for the year 2000. This resulted in a total of thirty-three names to analyze and draw conclusions from.

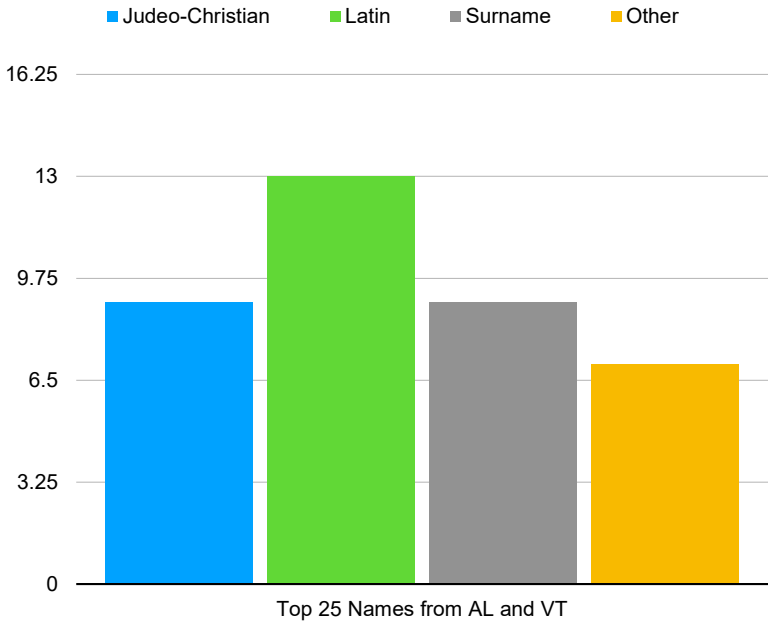
A significant trend I observed was the high prevalence of names of Latin origin (figure 1). This is partly because Latin can overlap with surnames and Judeo-Christian biblical names. However, there was no overlap between Judeo-Christian biblical names and surnames directly. Surnames could often be of Latin origin, and conversely, Judeo-Christian biblical names, although stemming from Hebrew or Aramaic, might be transformed and receive a Latin equivalent. Due to this overlap, there was a very high rate of Latin origins in the top twenty-five names from both states. In fact, nearly forty percent (thirteen out of thirty-three names) have some Latin origin. Nine of these names come from surnames, and nine come from Judeo-Christian biblical names. It is important to remember when reviewing the data that Latin can overlap with surnames or Judeo-Christian biblical names.

Figure 1



Beyond this general trend, we see a pattern within the top twenty-five names themselves. More frequent names, particularly those in the top ten by usage, tend to follow Judeo-Christian and Latin naming practices more closely. For example, five of the top ten names from Alabama and four of the top ten Vermont names all have Judeo-Christian biblical origins (figure 2).

Figure 2

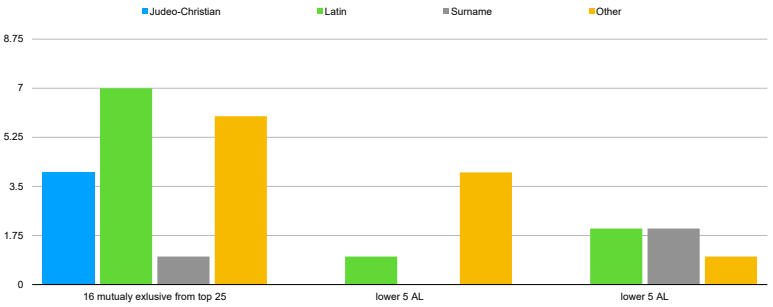


When we consider the potential overlap of Latin with Judeo-Christian names or surname origins, we recognize a strong influence of Judeo-Christian naming practices among the top ten most frequent names in these states. In fact, there is only one name in the top ten of both states that does not have Judeo-Christian, Latin, or surname origins. This outlier, *Emma*, has strong English roots and is associated with British royalty. There does not appear to be any Judeo-Christian or Latin origin to this name from an etymological standpoint.

As I examined the lower five names within the top twenty-five list, a pattern emerged of a decrease in biblical naming practices and an increase in other naming origins (figure 3). These include Irish, Ancient Greek, Native American, and Persian influences.

While these names are less common, they are still popular enough to be among the top twenty-five most frequently used names in each state.

Figure 3



The data reveals a trend of Judeo-Christian biblical names having a higher frequency in both Alabama and Vermont. These findings are consistent with the national average names. The top ten names in the year 2000 for the entire US are very similar to the average between Vermont and Alabama, with only three discrepancies out of the top ten. All three of these names which are not found in the US top ten are present within the top twenty-five of both states. This suggests that both Alabama and Vermont follow typical US naming patterns for the year 2000.

Conclusion

My analysis of naming practices in Alabama and Vermont during the year 2000 yielded some interesting results. The focus on the top twenty-five most popular names revealed a stark lack of variation between these two geographically distinct states. This suggests a potential homogenization of naming trends across the United States, with regional religious factors playing a minimal role in the selection of the most popular names.

However, it is crucial to acknowledge the limitations of this study. The scope, restricted to just two states, necessarily limits the claims of my findings. A more comprehensive understanding could be achieved by analyzing naming trends across all US states and territories. This broader analysis would allow for a more nuanced picture of regional variations and potential cultural influences.

Furthermore, the focus on the top twenty-five names may have caused an oversight in potentially revealing trends in less common names. It is possible that these less popular choices exhibit a stronger influence from regional or religious factors. Future research could improve upon this study by analyzing a broader spectrum of names within each state alongside more states.

Expanding the study to encompass all US states and territories would allow for a richer understanding of regional variations and potential cultural influences on naming practices. Additionally, investigating the potential correlation between religious affiliation and naming choices could shed light on the role of faith in this social practice. This could involve comparing the popularity of certain names with the religious demographics within each state.

Another intriguing avenue for future research would be a longitudinal study. By analyzing naming trends over time, researchers could identify how historical events, cultural shifts, and popular figures influence the popularity of certain names. Including both male and female names in such a study would provide a more comprehensive picture of societal naming preferences.

Finally, exploring the potential influence of socioeconomic background on naming choices could be a fruitful area of future investigation. This might require researchers to utilize data sources beyond what I was able to access for this research project. By incorporating information on socioeconomic factors, a more complete picture of the social and cultural forces shaping naming practices could emerge and reveal how religious practices interact with these broader social and cultural influences.

In conclusion, I believe this study does show that there is remarkably little variation between the most and least religious states in naming practices of females born in the year 2000. Although there are no noticeable trends identified leaning towards or against differences within these two states, it is a significant finding nonetheless that there was no noticeable variation between these two states.

References

Ancestry. (n.d.). *First name meaning: Mary*. (n.d.). Ancestry. Retrieved June 18, 2024, from https://www.ancestry.com/first-name-meaning/mary

Association of Religion Data Archives. (n.d.). *Browse data archive: Religion data*. The ARDA. Retrieved June 18, 2024, from https://www.thearda.com/data-archive/browse-categories?cid=B

BabyNames.com. (n.d.). Retrieved June 18, 2024, from https://babynames.com/

Behind the Name. *Samantha*. (n.d.). Behind the Name. Retrieved June 18, 2024, from https://www.behindthename.com/name/samantha

Bozkurt, Ferdi, & Mandana K. Mohammadi. (2023). A corpus-based socio-onomastic analysis on Turkish and American horror film naming. *CINEJ Cinema Journal*, *11*(2), 418–445. https://doi.org/10.5195/cinej.2023.589

Golub, Mitka R. (2024). A donation list from Elephantine: Judaean and Non-Judaean onomastic characteristics from the Persian period in Egypt, *Journal of Semitic Studies*, *69*(1), 135–160. https://doi.org/10.1093/jss/fgad014

Hanks, Patrick, Kate Hardcastle, & Flavia Hodges. (2006). *Oxford dictionary of first names* (2nd ed.). Oxford University Press. https://doi.org/10.1093/acref/9780198610601.001.0001

Lipka, Michael, & Benjamin Wormald. (2016, February 29). *How religious is your state?* Pew Research Center. Retrieved June 18, 2024, from https://www.pewresearch.org/short-reads/2016/02/29/how-religious-is-your-state/?state=alabama

Newport, Frank. (2015, February 15). Frequent church attendance highest in Utah, lowest in Vermont. *Gallup*. Retrieved June 18, 2024, from https://news.gallup.com/poll/181601/frequent-church-attendance-highest-utah-lowest-vermont.aspx

Pharr, Pauline C. (1993). Onomastic divergence: A study of given-name trends among African Americans. *American Speech*, 68(4), 400–409. https://doi.org/10.2307/455774

Slind, Marvin G. (2021). Child naming patterns as an indication of assimilation: A case study in Eastern Washington's Selbu community. *Journal of Family History*, *46*(2), 191–210. https://doi.org/10.1177/0363199020988582

Social Security Administration. (n.d.). Plain language at the Social Security Administration. Retrieved June 18, 2024, from https://www.ssa.gov/agency/plain-language/

Valentová, Iveta, & Ľubor Králik. (2022). Theory of proper names by Vincent Blanár and research into Slovak anthroponymy. *Onoma*, 57, 87–103. https://doi.org/10.34158/ONOMA.57/2022/6

Representation of African American Vernacular English in American Media

Noah Frazier

This research analyzes the representation of African American Vernacular English (AAVE) across genres of American film and television by investigating the relegated use of AAVE in the comedy, crime, and drama genres. Using quantitative, corpus-based methods in a cross-genre analysis, the author reaffirms the pattern of increased AAVE representation in the comedy genre identified by previous qualitative research methods and further identifies a marked increase of AAVE representation at the intersection of the crime and comedy genres. The author posits that American media's lack of AAVE representation beyond the crime and comedy genres bolsters harmful stereotypes against African Americans.

Linguistic research has found that discriminatory stereotypes against African Americans which originated historically in the colonial and antebellum eras of American history have been perpetuated and maintained in American media by the marginalization of African American Vernacular English (AAVE) to narrative driven genres. While prior research has primarily focused on the persistence of the Jim Crow stereotype and the lingering heritage of minstrel shows in comedies, sitcoms, and soap operas, there has been limited research on stereotypes against Africans Americans being perpetuated in other genres of television and film, and no research examining the subject with a cross-genre, corpus-based approach. This study addresses this gap in research through a cross-genre, corpus-based analysis of AAVE usage in American media and seeks to demonstrate how corpus linguistics can be applied to the discussion of genre-based linguistic discrimination as previous studies on the subject have typically involved qualitative research methods for the evaluation and comparing of films or television programs. By incorporating corpus-based methods in this study, a more quantitative understanding of stereotypes in American media is presented, which complements the existing literature on the subject, creating a more holistic understanding.

This study hypothesizes that there is a marked, increased usage of AAVE constructions and lexical items within the crime and drama genres of American television and film in comparison to other genres of television and film, comparable to the frequency of AAVE usage in the comedy genre. An increased usage or representation of AAVE in the crime or drama genres of television would indicate the perpetuation of the racist stereotype inherited from the colonial and antebellum eras of American history that African Americans have inherently violent demeanors. Understanding this misrepresentation of African Americans in American media provides a roadmap for rooting out persisting racist stereotypes.

Literature Review

Over the centuries, prevailing attitudes toward African influenced varieties of English have been documented in American media as "the mass media reflect[s] the state of knowledge, concern, and awareness of Black English" (Brasch, 1981, p. x.). From these attitudes, modern American society has inherited discriminatory

stereotypes against African Americans which are still pervasive in American media today (Bloomquist, 2015; Brasch, 1981; Bucholtz & Lopez, 2011; Khaghaninejad et al., 2019; Moody, 2020). Among the most persisting racist stereotypes in American media is the Jim Crow stereotype of a dim-witted fool and the Mandingo stereotype of the hypersexualized, violent Black man (Bloomquist, 2015; Khaghaninejad et al., 2019).

While previous historical eras have been overt in racial prejudices and depictions of Africans and African Americans in blackface productions or minstrel shows (Brasch, 1981), modern American media has been documented as linguistically marginalizing AAVE varieties, perpetuating the aforementioned lingering stereotypes (Bucholtz & Lopez, 2011; Khaghaninejad et al., 2019). One such example of marginalization of AAVE in media is the exclusion of nonstandard varieties of English from informational registers of English media, such as news broadcasts, while nonstandard varieties are maintained in narrative registers of English media, such as movies or sitcoms (Moody, 2020).

By relegating the use of AAVE to narrative driven media, harmful stereotypes are repeatedly perpetuated, "impact[ing]…the manner in which African Americans are perceived and treated by others" (Khaghaninejad et al., 2019). A prime example of harmful racial stereotypes being perpetuated by American media that has been extensively documented is the use of AAVE mimicked by White actors, termed *linguistic minstrelsy* (Bloomquist, 2015; Bucholtz & Lopez, 2011; Lopez, 2009, 2012). White actors' mimicking of AAVE for comedic effect and portrayal of hypersexuality further bolsters the stereotypes of African Americans being dim-witted, hypersexual, and violent (e.g., Jim Crow, Mandingo) and "may in fact be even more damaging…because racial stereotypes are hidden behind the parody of white [characters]" (Bucholtz & Lopez, 2011, p. 702).

While prior research has excellently documented stereotypes and AAVE usage in media, the thorough, qualitative methods utilized have typically narrowed findings to a single genre of media, often focused on the comedy genre (Bloomquist, 2015; Brasch, 1981; Bucholtz & Lopez, 2011; Moody, 2020). Previous research on this topic has primarily relied on qualitative methods, but some sociolinguists have advocated for incorporating corpus-based methods to augment sociolinguistic research (Kendall,

2011). Though corpus-based research methods have recently explored AAVE in American media, these studies have either honed into a specific genre (Khaghaninejad et al., 2019) or have not accounted for genre distinctions (Livskog, 2022). As such, adopting corpus-based methods in a cross-genre approach provides a more holistic view on AAVE usage in American media and the various racial stereotypes perpetuated by linguistic marginalization.

Data and Methods

The Corpora

To identify trends in AAVE usage across genres, this study relied on the Corpus of Contemporary American English's (COCA) database of film and television scripts and subtitles, creating a variety of sub-corpora to be representative of each genre analyzed. However, as COCA tags each script for multiple genres represented in the films and programs, this study created six subcorpora based on these genre tags, manually controlling for genre overlap. The subcorpora utilized were action, comedy, crime, comedy/crime, drama, and news.

Table 1

The Genre Corpora

Corpus	Number of Files	Number of Words
Action	3,671	17,218,496
Comedy	2,523	15,656,912
Crime	5,001	27,080,442
Comedy/Crime	267	2,487,076
Drama	4,056	24,188,483
News	5,000	10,038,478

Each of the subcorpora, with the exception of the news subcorpus, is composed of the movies and television shows that are primarily tagged for their genre and exclude any of the other genres being analyzed in this study. For example, the action corpus contains films marked as action/romance, action/sci-fi, action/thriller, but excludes action/comedy and action/crime. The comedy and crime corpora include scripts secondarily marked as comedy or crime (e.g., action/comedy, drama/crime), but any script that was

tagged for both comedy and crime were separated into the specific comedy/crime corpus. The news subcorpus is composed of transcripts of unscripted conversation from over a hundred television and radio programs. The news subcorpus serves as a representation of the non-narrative, informational register of American media in contrast to the narrative-driven genre subcorpora.

The AAVE Features

Each of the AAVE features analyzed in this study was chosen for its long-standing recognition as a marked feature of AAVE easily recognized by both insiders and outsiders of the African American community as a linguistic feature associated with African Americans. The lexical and syntactic features analyzed included *ain't*, *finna*, the non-rhotic n-word *n*gga*, the adjectival intensifier *ass*, *done* completion, multiple negation, habitual *be*, and the plural demonstrative *them*. For every query searched in the corpora, the results were manually checked for accuracy in representing AAVE constructions. When possible, each instance returned by a query was examined for accuracy. However, when queries returned more than two hundred instances, a random sample of two hundred instances was taken and examined for accuracy; the ratio of AAVE constructions to Standard American English (SAE) constructions was then utilized to extrapolate a usage frequency for that corpus query. See table 2 for the corpus search queries.

Ain't

While the lexeme *ain't* is not exclusively used by AAVE and can be found in other nonstandard varieties of English, the lexeme is strongly associated with African Americans and AAVE (Green, 2002). As such, searching for the lexeme *ain't* in American media gives indication to the distribution of nonstandard English varieties in American film and television. Considered alone, this single query would not provide sufficient evidence to the hypothesis, but will be considered in conjunction with the other AAVE features.

Finna

Etymologically derived from the SAE phrase *fixing to*, the preverbal marker *finna* indicates that an event is imminent. While the phrase *fixing to* is common through many varieties of English, *finna* is a standard and common AAVE spelling for the AAVE pronunciation of the phrase (Green, 2002). As such, any instance of

the lexeme *finna* within a script would be indicative of a character speaking AAVE.

Non-Rhotic N-Word

Considered to be one of the most inflammatory lexemes in the English Language, the n-word *n*gger*, and its non-rhotic variant *n*gga*, have a long, complex history deeply embedded in racism and discrimination. While much has been said on the topic by previous research, this paper recognizes the importance of the non-rhotic n-word as a strong indication of AAVE usage to signal Black cultural identity and solidarity, and recognizes that the term is most commonly employed by Black men (Rahman, 2012). Instances of the non-rhotic n-word are clear indications of Black representation in media.

Adjectival Intensifier *Ass*

While the use of the swear *ass* is used by most varieties of English, AAVE is unique in the syntactic construction of *ass* as an adjectival intensifier. These constructions are separate from common swears in SAE and other varieties of English (e.g., *smart-ass*, *dumbass*) in that the *ass* is an intensifier, constituent to a preceding adjective, and modifying to a following noun (e.g., *spicy-ass food*). Linguists have posited that this AAVE adjectival intensifier form of *ass* has undergone semantic bleaching, the process of reducing or losing lexical meaning, and functions more similarly to a grammatical morpheme (Mufwene, 1998).

***Done* Completion**

The preverbal marker *done* precedes past tense verbs to indicate the action is completed or that there has been a change of events or states. The AAVE *done* precedes *-ed* verb forms and is pronounced in an unstressed syllable (i.e., /dən/). This specific construction is syntactically and semantically different from the adjectival SAE form *done* (e.g., "Her homework is done") and the past participle form of *do* (e.g., "She has done her homework"/"She done her homework") (Green, 2002).

Multiple Negation

AAVE licenses the use of multiple negators within a single sentence, which directly contrasts with prescriptive grammatical rules in SAE (Green, 2002). While there is no limit to the number of negators that can be present in a single sentence, this study chose to identify three main syntactic constructions

which place two negators in quick succession (e.g., "He ain't never here").

Habitual *Be*

The aspectual marker be in AAVE indicates a habitual or iterative meaning to the modified verb (Green, 2002). While habitual *be* is one of the most studied, unique features of AAVE, developing a corpus query within COCA that maintained accuracy and precision proved challenging. Many attempted queries resulted in numerous false positives that were not indicative of AAVE usage. Limited in scope, the query utilized in this study to identify instances of habitual *be* resulted in too few instances to be considered representative of the syntactic construction's distribution. However, the limited instances elicited by these queries were included in the total count of AAVE features.

Them as a Plural Demonstrative

Many varieties of AAVE permit or prefer the lexeme *them* as a plural, distal demonstrative, in contrast to the SAE demonstrative *those* (Mufwene, 1998).

Table 2

Corpus Queries of AAVE Features

AAVE Features	Corpus Queries	Examples
Ain't	ai n't	"He ain't here"
Finna	finna	"I'm finna go"
Non-Rhotic N-Word	nigga	————————
Adjectival Intensifier Ass	ADJ ass NOUN	"We're done playing this stupid ass charade"
Done Completion	_p done _v?d	"You done moved on up"
Multiple Negation	ai n't n* nobody DO n't DO n't nobody	"He ain't never hurt nobody"
Habitual Be	PRON always be _v?g NAME always be _v?g	"He always be sliding out of his chair"
Plural Demonstrative Them	them _nn2	"Y'all need to leave them doughnuts alone"

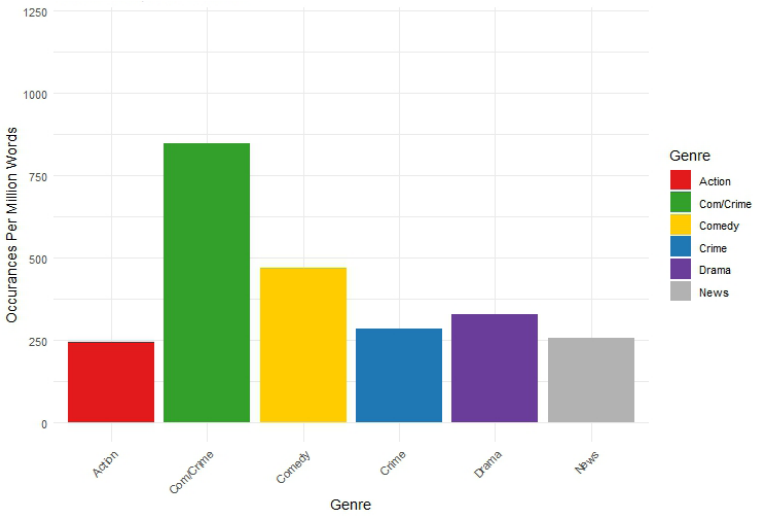
Results

High-Frequency AAVE Features

Occurring as high as 850 instances per million words, the most frequent AAVE feature observed in this study was the use of the negator *ain't*.

Figure 1

Ain't *Negation Usage*



Shown in figure 1 above, the comedy/crime genre displays a markedly higher usage of *ain't*. The comedy genre also utilizes *ain't* notably higher than other genres of media. The drama genre demonstrates a marginal increase of *ain't* usage, while the crime genre does not seem to have a notable difference in comparison to the action and news genres. These normalized frequency counts visualize the degree to which each genre represents non-standard English variety usage, with comedy/crime ranking the highest.

Mid-Frequency AAVE Features

Occurring up to 150 instances per million words, the following AAVE features of multiple negation and demonstrative *them* occurred notably less frequently.

Figure 2

Multiple Negation Usage

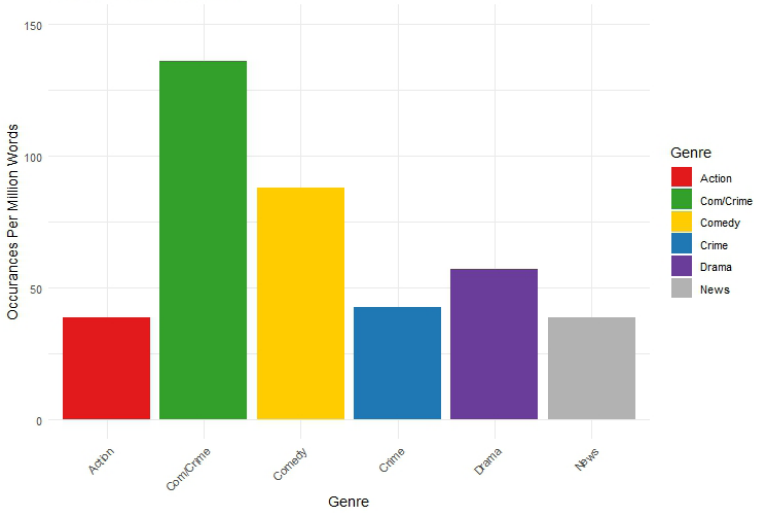
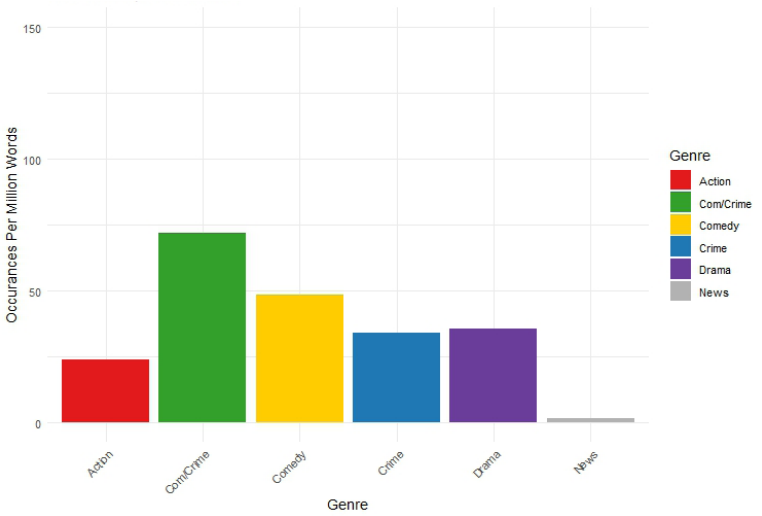


Figure 2 above visualizes the normalized frequencies of multiple negation constructions demonstrated in each media genre. Though multiple negation occurs less frequently than simple instances of the negator *ain't*, the data shows that multiple negation follows the same trends as previously seen. Action, crime, and news utilize the AAVE feature to a similar extent, drama portrays a marginal increase in usage, comedy has a relatively high usage, and comedy/crime has a markedly high usage of multiple negation. While the usage of negator *ain't* depicted in figure 1 simply demonstrates nonstandard varieties of English, figure 2 suggests that AAVE negation follows the same trend.

Figure 3

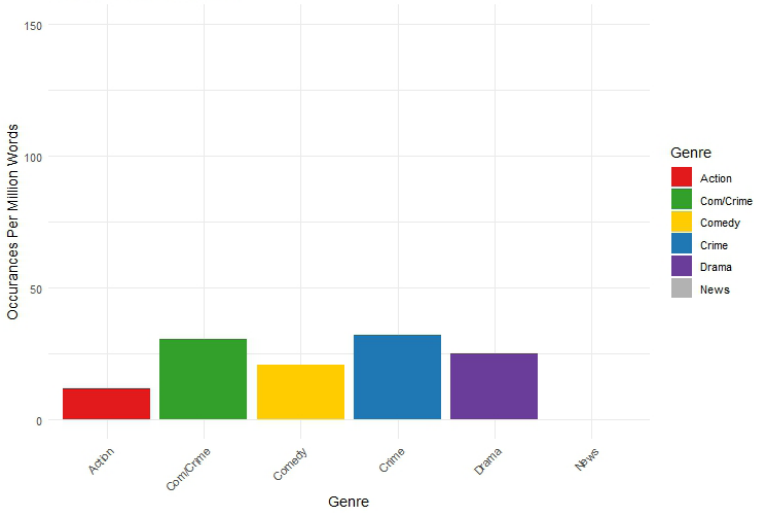
Demonstrative Them *Usage*



Repeating similar trends for the narrative genres demonstrated by previous AAVE features, demonstrative *them* drops suddenly to only a couple instances per million words in the news genre. One possible reason for the difference in usage of demonstrative *them* in the news genre is a lack of African American representation, or linguistic representation of AAVE, within the genre. Because the news genre is an authoritative and informational register, prior research would suspect that the genre would elicit fewer non-standard varieties of English (Moody, 2020). Interestingly, many of the false positives elicited by the query and discarded in analysis demonstrated the use of the third-person plural object pronoun *them* within the construction "many of them NOUN.PLURAL" when specifying the backgrounds and demographics of large groups (e.g., "Officials say sixty-four people died, many of them children"). Abundant use of this syntactic construction may overshadow the usage of the AAVE feature *them* as a demonstrative, but further research on the topic would be needed for empirical conclusions and explanations for this phenomenon.

Figure 4

The Non-Rhotic N-Word Usage

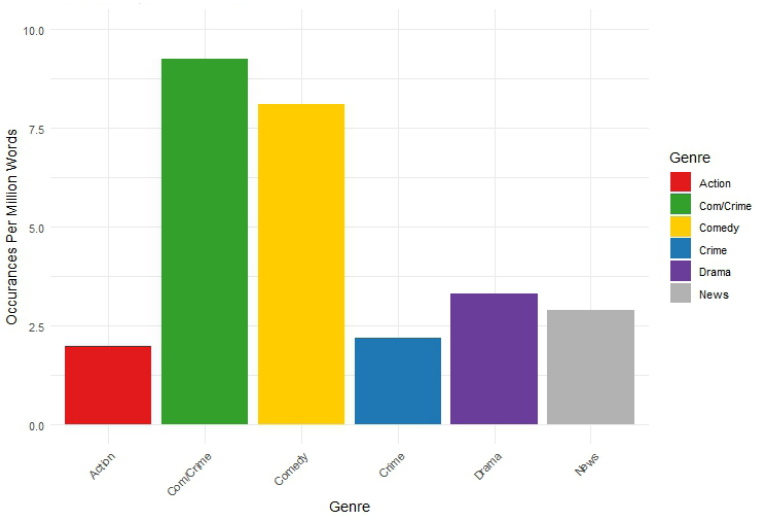


Breaking from the trends in data shown by previous AAVE features, the non-rhotic n-word is the only AAVE feature for which the crime genre elicited the most usage, though the usage of the non-rhotic n-word is similarly matched by the comedy/crime genre. This AAVE feature is likely censored from the news genre due to its inflammatory nature, resulting in the zero instances reported in the whole news subcorpus. Similarly, the inflammatory nature of the n-word, along with media censorship laws, are likely the reasons as to why none of the narrative genres show a notable spike in usage of this AAVE feature. Despite the relatively few occurrences of the non-rhotic n-word found in the narrative genre subcorpora, the marked, increased usage of the feature in the crime and comedy/crime genres are directly related to the representation of African Americans in these two genres as the term and its usage is distinctly reserved for African Americans and is typically taboo for other ethnicities and races (Rahman, 2012). These increased frequencies of the non-rhotic N-word in the crime and comedy/crime genres strongly suggest a higher representation of Black characters in these genres.

Low-Frequency AAVE Features

Occurring in less than ten instances per million words, the following AAVE features of adjectival intensifier *ass*, *done* completion, and *finna* are not commonly found in any genre subcorpus.

Figure 5

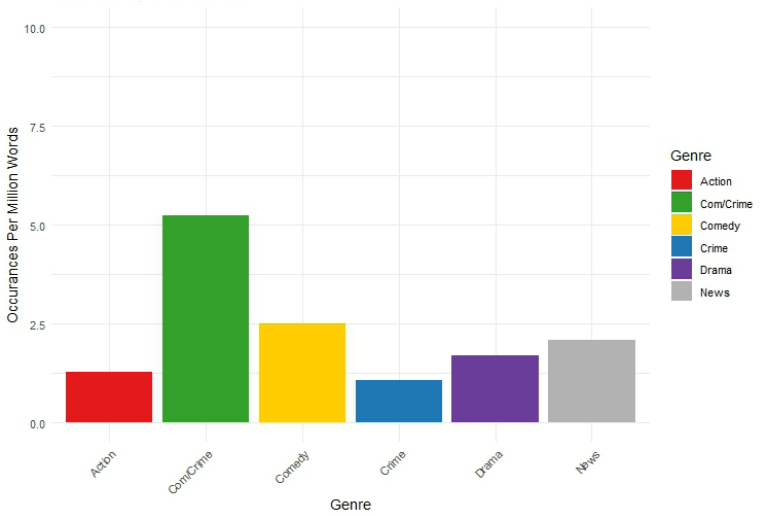


Adjectival Intensifier Ass *Usage*

Despite occurring substantially less often than other AAVE features analyzed in this study, the adjectival intensifier *ass* still follows the trend that the comedy/crime genre elicits the most usage of the feature, followed closely by the comedy genre. The remaining genres utilize the AAVE feature to similar degrees. However, while the general trends seen by previous AAVE features are demonstrated by adjectival intensifier *ass*, there is a notably larger proportional difference between the comedy/crime and comedy genres with the other genres, suggesting that this AAVE feature may be overrepresented in these two genres.

Figure 6

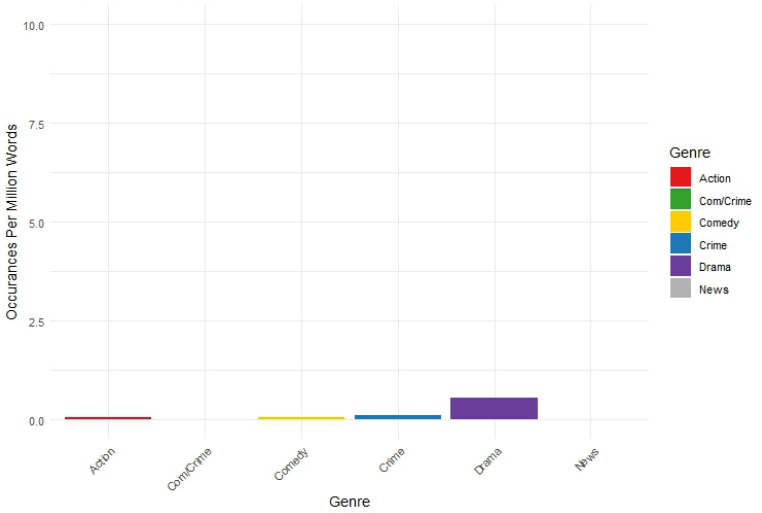
Done *Completion Usage*



Though *done* completion demonstrates the same general cross-genre trends identified by other AAVE features, the news genre shows a comparative increase in this specific AAVE feature. One possible explanation for this increase could be that the interviews included in the news subcorpus may be eliciting more contexts and scenarios appropriate for discussing completed events or a change in a state of events as interviews can influence which syntactic constructions interviewees rely on while answering interviewer questions (Cukor-Avila & Balcazar, 2019).

Figure 7

Finna *Usage*

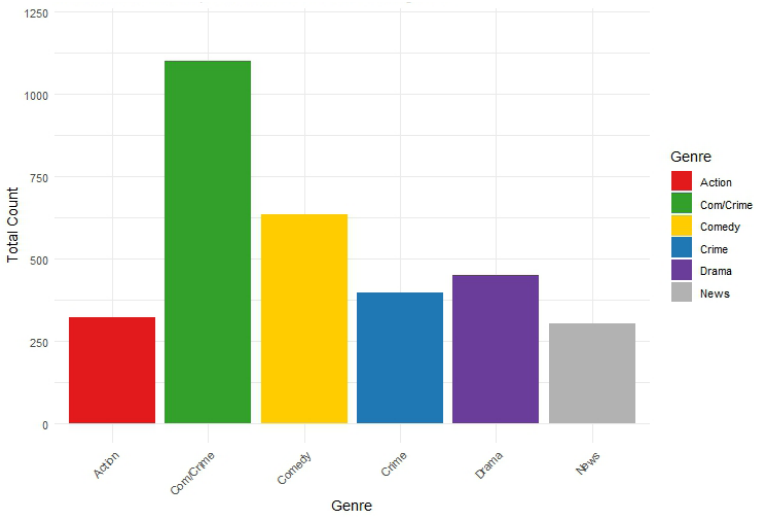


The lexeme *finna* occurs less than once per million words for every genre. As such, the frequency of *finna* does not follow the same trends identified in the analyses of other AAVE features. One possible explanation for the limited usage of *finna* across all the subcorpora is that the preverbal marker is genuinely a less common preverbal marker in AAVE and would not elicit many instances within a corpus of AAVE texts or speech, though more research would be needed to identify usage of AAVE features in strictly Black communications. As a less common preverbal marker from a nonstandard variety of English, *finna* would be rarely seen in American media. An alternative explanation for the lack of instances of *finna* in the subcorpora is a misrepresentation of African American speech patterns in the scripts and subtitles which comprise the corpus. As *finna* is etymologically derived from the phrase *fixing to*, the instances of AAVE *finna* may be recorded in the text files as *fixing*. With limited instances of *finna* identified in the subcorpora, the data is inconclusive if the cross-genre pattern would mimic the trends exhibited by other AAVE features. Though the results elicited few instances of *finna* usage across genres, the lack of *finna* usage indicates limited representation of AAVE in all genres of American media.

Total AAVE Features and General Trends

Figure 8

Total AAVE Features



Representing the overall usage of AAVE Features in American media, figure 8 above represents the combined total occurrences of all AAVE features elicited by all of the corpus queries, normalized per million words. Figure 8 indicates that the action genre represents AAVE to the same degree as news outlets do and the crime and drama genres represent AAVE slightly more often. While all the narrative-driven genres rank higher than the informational news register, the comedy register displays twice the frequency of AAVE features than is represented in news, and there is a marked, increased usage of AAVE features within the comedy/crime genre of American media. These general representations of AAVE features support the prior findings of researchers that American comedies utilize AAVE more regularly than other genres (Bucholtz & Lopez, 2012; Moody, 2020) and that American dramas also maintain an increased usage of AAVE (Khaghaninejad et al., 2019). While the crime genre does not display AAVE usage to the same degree as the comedy genre, as hypothesized, the combination of comedy and crime does elicit AAVE features considerably more often than any other genre analyzed.

Figure 9

Representation of AAVE features in America media genres

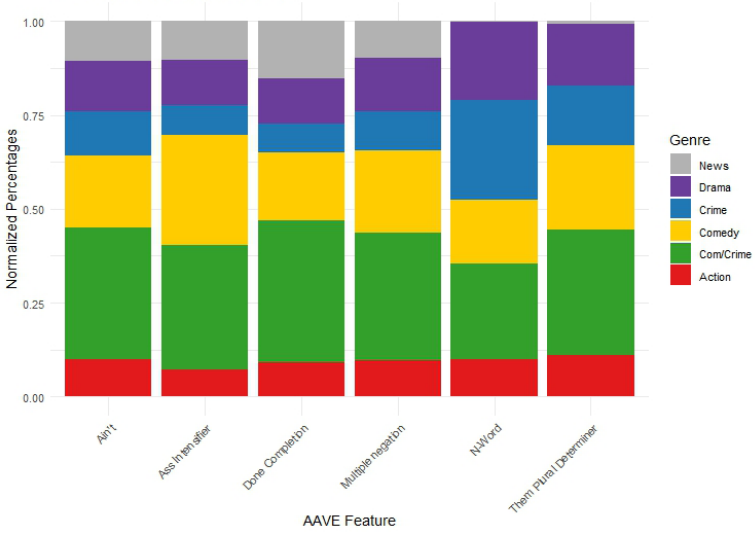


Figure 9 above represents the proportional usage of AAVE features by each genre of American media and illustrates the overall trends of AAVE usage frequency identified. Generally, each genre is relatively consistent in the degree to which they employ AAVE features. The action genre consistently demonstrates a lesser amount of AAVE features than other genres, often reflecting the usage of AAVE identified in the news. Though the crime genre regularly shows fewer instances of AAVE features, the genre displays a notable proportional increase in the use of the non-rhotic n-word, signaling a marked increase in African American representation within the crime genre.

Discussion

Previous researchers have posited that the comedy genre of American film and television has maintained the stereotypes developed in the antebellum era (such as the Jim Crow caricature) by exploiting African American speech patterns for humor (Bloomquist, 2015; Bucholtz & Lopez, 2011; Lopez, 2009, 2012; Moody, 2020). Results from this study's corpus queries are consistent with this stance. Though the comedy genre consistently ranks the second highest in AAVE usage, the genre shows a

considerable, proportional spike in the usage of adjectival intensifier *ass* while also showing a decrease in the usage of the non-rhotic n-word (see figures 4, 5, and 9). The spike in adjectival intensifier *ass* can be explained as the use of the AAVE feature to exaggerate and intensify scenarios for comedic effect. Ranking considerably lower in the usage of the non-rhotic n-word would indicate a smaller representation of Black actors, actresses, and characters signaling themselves as part of the Black community to other Black characters. Considered collectively, the frequency of AAVE features identified would support previous claims that the comedy genre of American media is perpetuating racial stereotypes against African Americans.

This study contests the idea that AAVE usage has been marginalized to narrative genres and excluded from authoritative registers of media (Khaghaninejad et al., 2019; Moody, 2020), proposing instead that AAVE is represented in narrative genres to the same degree as it is represented in informational registers. Although this study contests previous research that AAVE features are marginalized to narrative genres, this study does not dispute that there still exists possible underrepresentation of AAVE in all genres.

This article hypothesized that the crime and drama genres of American media would show marked increases in AAVE usage in comparison to other genres and comparable to AAVE usage in the comedy genre. Overall, neither genre proved to utilize AAVE features to the same degree as the comedy genre, despite utilizing AAVE more than the other genres of action and the news (see figure 8). However, the data illustrates consistent and substantial AAVE representation in the comedy/crime genre (see figures 1, 2, 3, 5, 6, and 8). Although the crime genre reported proportionally fewer AAVE features than other genres, the combination of comedy and crime consistently demonstrated the highest proportional usage of AAVE features comparatively (see figure 9). One possible explanation for this phenomenon is an increased representation of Black actors, actresses, and characters interacting with one another in the comedy/crime genre in contrast to the comedy genre. This explanation is supported by the evidence that the crime and comedy/crime genres exhibit the use of the non-rhotic n-word to a similar degree and are both utilizing the term more than the comedy genre (see figures 4 and 9). As the non-rhotic n-word is typically reserved for members of the Black community to signal solidarity and connection to the community and

is taboo for other ethnicities and races, the increase in frequency could indicate an increase in interactions between Black characters. As the non-rhotic n-word is also more commonly employed by Black men, these increased frequencies may also be indicative of more Black male characters being portrayed in these genres. As an amalgamation of the crime genre, a genre representing more Black characters, and the comedy genre, which elicits exaggerated rates of AAVE, the comedy/crime genre likely amplifies the usage of AAVE, thus reinforcing both negative stereotypes of AAVE speakers being dim-witted and of Black individuals having more violent dispositions.

Implications of AAVE Media Representation

The representation of AAVE, as highlighted in this study, suggests significant implications regarding the perpetuation and reinforcement of harmful racial stereotypes toward African Americans which have persisted in America for centuries. The marked use of AAVE in the comedy genre ties the linguistic patterns of African Americans to caricature, parody, and exaggeration, suggesting that AAVE is a source of humorous entertainment and undermines the Black experience in America. Furthermore, the proportionally high usage of AAVE in the crime/comedies further bolsters stereotypes against African Americans by conflating the comedic stereotype with the colonial and antebellum stereotype that African Americans are inherently dangerous and violent. By actively promoting racial diversity and representing AAVE positively in genres and characters void of the comedy and crime genre influences, American media can challenge these harmful stereotypes and draw attention to social prejudices.

Future Research

While this study considered several AAVE features, the research was restricted to lexical and morphosyntactic constructions which were represented in the texts of COCA. As the AAVE features considered were neither all-encompassing nor representative of phonological processes of AAVE, further research should be undertaken to analyze the phonological features of Black characters in a cross-genre analysis. Additionally, the scope of this study should be expanded to include more genres and genre overlaps than were considered in this study. Finally, the conclusions from this study that the crime and comedy/crime genres feature more

interactions between Black characters than other genres should be further researched and confirmed.

Conclusion

Results from this study indicate that while the comedy genre of American film and television elicits high usage of AAVE—a pattern consistent with previous research—the combination of the comedy and crime genres elicits a markedly higher usage of AAVE than other genres examined. Although racist prejudices may not be as overt as they were in historical representations of African Americans, these notable increases in AAVE in the comedy and comedy/crime genres perpetuate and reinforce racist stereotypes toward African Americans, insinuating that AAVE speech patterns indicate that characters are dim-witted and that African Americans have inclinations toward violence. Future media should be mindful of these pervasive stereotypes and seek to represent AAVE and African Americans positively and authentically in genres outside of these contexts to combat racial prejudice.

References

Bloomquist, Jennifer. (2015). The minstrel legacy: African American English and the historical construction of "Black" identities in entertainment. *Journal of African American Studies, 19*(4), 410–425. https://doi.org/10.1007/s12111-015-9313-1

Brasch, Walter M. (1981). *Black English and the mass media*. University of Massachusetts Press.

Bucholtz, Mary, & Qiuana Lopez. (2011). Performing blackness, forming whiteness: Linguistic minstrelsy in Hollywood film. *Journal of Sociolinguistics, 15*(5), 680–706. https://doi.org/10.1111/j.1467-9841.2011.00513.x

Cukor-Avila, Patricia, & Ashley Balcazar. (2019). Exploring grammatical variation in the corpus of regional African American Language. *American Speech, 94*(1), 36–53. https://doi.org/10.1215/00031283-7321989

Davies, Mark. (2008–). *The corpus of contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/.

Kendall, Tyler. (2011). Corpora from a sociolinguistic perspective. *Revista Brasileira de Linguística Aplicada, 11*(2), 361–389. https://doi.org/10.1590/S1984-63982011000200005

Khaghaninejad, Mohammad S., Mehrnoosh Dehbozorgi, & Mohammad A. Mokhtari. (2019). Cultural representations of Americans, Europeans, Africans and Arabs in American soap operas: A corpus-based analysis. *International Journal of English Language & Translation Studies*, 133–141.

Livskog, Felizia. (2022). *"What's up, bro?" A synchronic corpus study on the use of bro in TV shows* [Bachelor's thesis, Linnæus University]. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1793782&dswid=-5985

Lopez, Qiuana. (2009). Imitation or influence: White actors and Black language in film. *Proceedings of the Seventeenth Annual Symposium About Language and Society – Austin*, 110–120. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4e68d22d925d09d45946ea591437f92afb0efcd9

Lopez, Qiuana. (2012). *White bodies, Black voices: The linguistic construction of racialized authenticity in US film* [Doctoral dissertation, University of Texas at Austin]. University of Texas Libraries.

Moody, Andrew. (2020). Authority and authenticity in media Englishes and the Englishes of popular culture. In Ram A. Giri, Anamika Sharma, & James D'Angelo (Eds.), *Functional Variations in English, 37,* 99–110. Springer International Publishing. https://doi.org/10.1007/978-3-030-52225-4_7

Mufwene, Salikoko S., Guy Bailey, John Baugh, & John R. Rickford (Eds.). (1998). *African-American English: Structure, history, and use*. Routledge.

Rahman, Jacquelyn. (2012). The n word: Its history and use in the African American community. *Journal of English Linguistics, 40*(2), 137–171. https://doi.org/10.1177/0075424211414807

The Nuance of the Serial Comma

Abigail Christensen

Opinions on the usefulness of the serial comma vary throughout the writing community. This article examines the history of this divisive punctuation mark through corpora data, providing important context to an issue that dates back centuries. The author examines both sides of the issue and considers the advantages and disadvantages of serial comma use. The article ultimately argues for the necessity of the serial comma using style guide prescriptions, various newspaper publications, and a legal case as support.

The English language is a complex medium of communication, particularly when it is written. It has many nuances that can alter the meaning the writer is trying to convey or its clarity. One critical component of written English is punctuation. A single stroke can direct the entire meaning or connotation of a sentence, so these marks must be carefully chosen. One punctuation mark that has sparked controversy is the serial comma (also known as the Oxford or Harvard comma). This comma is one that is placed directly after the penultimate item in a list and often occurs before a conjunction (Garner, 2022, p. 981). Although some style guides argue against its use, the clarity the serial comma brings demonstrates its continued necessity in American English. This article will discuss the history of the serial comma, the recommendations varying style guides make about it, and individual cases that emphasize its continued importance in writing.

Historical Evolution

Punctuation has evolved greatly over time in the English language. For many years, written English was highly unstandardized and largely served as a way to record spoken words. Punctuation was occasionally used in a rhetorical sense (i.e., indicating pauses that would occur if the words were spoken aloud), but this use was very irregular and varied between manuscripts (Baron, 2001, pp. 22–24). However, the Renaissance (a period that spanned from about 1400 to 1650) brought about many new innovations, including the printing press. This machine allowed for mass production of written works, thus resulting in more standardized uses of punctuation (Calle-Martin & Criado-Pena, 2022). By the sixteenth century, the core punctuation marks of English (including the period, comma, colon, etc.) were established and discussions regarding the necessity of the serial comma began.

The serial comma has a history rich with change and varying opinions over time. According to a corpus study including both British and American English that tracked the inclusion of the serial comma from 1500 to 1999, inclusion was preferred over omission as early as the sixteenth century, with 61.09 percent of cases including the comma. However, its use was fairly irregular due to its newness and the general lack of standardization in written English at this time. These numbers of inclusion rose and hovered around 80 percent throughout the seventeenth and eighteenth centuries before rising to 91.8 percent by the

mid-nineteenth century (Calle-Martin & Criado-Pena, 2022). The increase in prescriptivism and attention to grammar over this period contributed to this rise in the inclusion of the serial comma. Writers began to view punctuation as a grammatical tool rather than a rhetorical one, and the inclusion of the serial comma helped to create a clearer, more grammatical structure to writing (Baron, 2001). There was also more investment in prescribed rules as style and usage guides came into greater prominence, so people began to follow the popular recommendation of including the serial comma.

These trends of increased prescriptivism, particularly of punctuation, continued into the twentieth century. In 1905, a man named Horace Hart (a major leader at the Oxford University Press) began requiring those who worked for him to implement the serial comma in their writing. Due to the prestige of the Oxford University Press, people began following its prescribed rule of including the serial comma in lists, and the comma became known colloquially as the Oxford comma (Bushnell, 2021). In the first half of the twentieth century, the inclusion of the serial comma remained high, presumably in part due to the influence of Hart and the Oxford University Press (Calle-Martin & Criado-Pena, 2022). This series of events demonstrates the influence that style guides have on the public and how they can influence actual usage.

However, the twentieth century also brought greater ideas of flexibility and individuality regarding grammar and, by extension, punctuation and the serial comma. In the latter half of the century, "it [was] fashionable to prefer a relatively open style of [punctuation]," which means that writers took greater liberties in applying prescriptive punctuation rules and instead chose to "open the doors to fashion" and "creativity of individuals" (Chafe, 1988, p. 416). As a result of this surge in individuality, the use of the serial comma in written works began to decline, sinking to 7.69 percent inclusion by 1999 (Calle-Martin & Criado-Pena, 2022).

Prescription by Style Guides

Despite its recent dip in actual usage, the serial comma is still prescribed by several style and usage guides, often on the basis of producing more clear writing. Prevention of ambiguity is one of the major justifications for any prescribed rule by a usage guide, as it is generally agreed that writing is best when the reader understands

what the author is trying to communicate. Omitting the serial comma can lead to several forms of ambiguity, including supposed appositives and misunderstood conjoinings (Garner, 2022, p. 984). Conjoinings are when two words are associated with each other through conjunctions such as *and* (or, in these cases, misunderstood to be associated with each other when they are not). When the serial comma is omitted, misunderstandings can come about when verbs are able to take both the final items in a list as their objects or when modifiers can modify both final items. For instance, in the sentence "He is an employee at the Ministry of Magic, a husband and father of three school-age children," a reader could misunderstand the sentence to mean that the man is married to the children as well as being their father. Although further thought would most likely resolve this misunderstanding, it still detracts from the flow of the piece, thus worsening the reader's experience (Garner, 2022, pp. 984–985).

Another common misunderstanding that results from omitting the serial comma deals with appositives. For instance, in the sentence "We went for a walk with our dogs, grandma and grandpa," *grandma* and *grandpa* could be misunderstood to be the names of the dogs, acting as an appositive phrase (Garner, 2022, p. 984). However, if the serial comma is added between *grandma* and *and*, the ambiguity is lessened, as it becomes clear that the writer is referring to a list of entities. This example illustrates that adding the serial comma can be a major contributor to limiting ambiguity. It clarifies the ways in which words are or are not related to each other, thus making the author's meaning more understandable.

In addition to Garner, other prominent usage guides like the *Chicago Manual of Style* also prescribe the serial comma. Chicago argues that using the serial comma prevents ambiguity, but also refers to previous style guides and opinions to defend its decision (Harper, 2017). Renowned usage guides such as *A Dictionary of Modern English Usage*, *The Oxford Guide to Style*, and *American Usage and Style: The Consensus*, along with others, have prescribed the serial comma for centuries. This longevity is used by *Chicago Manual of Style* and other prescribers of the serial comma to emphasize its continued importance. The comma clearly has value and should continue to be included to prevent the ambiguity that each of these usage guides warns about.

Legal Implications: A Case Study

The clarity provided through use of the serial comma is also critical in legal writing. Because of its purpose to outline laws and otherwise make arguments, "the chief hallmark of good legal writing is clarity," and "clarity requires proper (i.e., conventional) grammar and punctuation" (Osbeck, 2012, pp. 16–17). Since the serial comma has generally been agreed upon to limit ambiguity, it is therefore a critical component for lawyers to include in their writing, as has been shown by the results of legal cases.

The serial comma played a pivotal role in the decision of *O'Connor v. Oakhurst Dairy*. This court case involved a group of truck drivers who sued Oakhurst Dairy in 2014 for overtime wages. In Maine (the state where the lawsuit took place), workers are required "to be paid 1.5 times their normal rate for each hour worked after 40 hours" except for work related to "the canning, processing, preserving, freezing, drying, marketing, storing, packing for shipment or distribution of: (1) Agricultural produce; (2) Meat and fish products; and (3) Perishable foods" (Victor, 2017). The writers of this law chose not to include the serial comma between the words *shipment* and *or,* and the drivers' lawyers used this omission to convince the court that the law was ambiguous. They argued that the exceptions of overtime pay applied to the *packing* for distribution rather than the distribution of the goods themselves. As a result of this ambiguous law, the court ruled in favor of the truck drivers (who never packed the goods), and they were awarded around ten million dollars (Victor, 2017).

This example highlights the real implications of omitting the serial comma. Choosing not to include it led to a lack of clarity in law, and this ambiguity cost a company millions of dollars in damages. Without the confusion that the omission of the serial comma caused, the truck drivers would not have had a strong enough case to win in court, as is evidenced by the fact that they lost the case in the lower court (Victor, 2017). However, the appeals court decided that "the absence of a comma produced enough uncertainty to rule in their favor" (Victor, 2017). This legal example indicates the extensive difference that one punctuation mark can make in a piece of writing, demonstrating the importance of choosing these marks with precision.

Counterargument: Proscription of the Serial Comma

Despite its many benefits, the serial comma is not universally prescribed in writing. For instance, *The New York Times* "recommends against serial commas except when needed for clarity" (McCoubrey, 2017). They explain that omitting the serial comma except when absolutely necessary is a tradition in news writing, likely to help save time and effort during typesetting (Corbett, 2015). Garner also touches on this idea, stating that news sources "prefer to omit these commas to squeeze more words onto lines" (Garner, 2022, p. 985). Although this reasoning makes sense on a practical level, it is arguable that clear and accurate communication is more important than saving space when it comes to news writing. If people misunderstand what is said, the space or money saved is irrelevant.

The *Associated Press Stylebook* also recommends omitting the serial comma in most cases. However, like *The New York Times*, they include the caveat that writers should "include a final comma in a simple series if omitting it could make the meaning unclear" (Casagrande, 2023). Although following this more flexible rule may resolve some of the concerns with always omitting the serial comma, it still has flaws. First, this guidance allows for subjectivity on the part of the writer or editor. This can still lead to ambiguity because, as Garner explains, "people aren't necessarily good at seeing how others might misread their words" (Garner, 2022, p. 985). The meaning behind the words will always be clear to the author, so forcing them to make a judgment call regarding the serial comma's imperativeness could lead to unclear writing.

In addition, the lack of a clear-cut rule for the inclusion of the serial comma leads to inconsistency in writing. A major component of a good piece of writing is consistency and order throughout. If certain grammatical elements, including the serial comma, are not used consistently throughout a piece, "the reader may perceive the document to be disorderly and unprofessional or may be distracted by the inconsistencies" (Farkas & Farkas, 1981, p. 16). Choosing to include the serial comma only when the writer deems it necessary could evoke these thoughts in the reader, and the inconsistency could distract from the overall meaning of the document. Thus, it is a better practice to consistently use the serial comma than to use it sparingly.

Some opponents to the inclusion of the serial comma also argue that it is redundant due to the presence of the conjunction (Garner, 2022, p. 983). However, because of the alternative uses of both conjunctions and commas (such as the conjoinings and appositives discussed earlier), the serial comma is a necessary component for maintaining clarity in a sentence. It distinctly indicates the role that the various words play and their relation to one another. The serial comma's role in promoting clarity emphasizes that it is often not redundant and should be included. As many proponents of the serial comma argue, "including it *never* creates an ambiguity, whereas omitting it often does" (Garner, 2022, p. 982). This statement affirms that there is no harm done by including the serial comma, so it is always safer to include it to reduce the risk of creating an unclear sentence.

Conclusion

The serial comma debate has strong proponents on both sides, but its use is critical in preventing ambiguity. Including it takes up slightly more space to a written work, but the benefits of the clarity it provides are more than enough to offset this minor inconvenience. The serial comma provides a clearer sense of what each word in a sentence refers to, and it helps prevent small misunderstandings, distraction from the overall message, and ambiguity in law that can have major consequences.

These ideas regarding the serial comma also lend themselves to future analysis of actual usage. It would be interesting to observe how, when, and where regular people decide to include or omit the serial comma in their writing. In addition, the growing prevalence of digital communication and news reporting may alter the patterns of serial comma usage or prescriptivism in the years to come. The history of the serial comma is a rich one, and its future inclusion and exclusion is a very intriguing prospect to consider.

References

Baron, Naomi S. (2001). Commas and canaries: The role of punctuation in speech and writing. *Language Sciences, 23*(1), 15–67. https://doi.org/10.1016/S0388-0001(00)00027-9

Bushnell, John T. (2021). *What is an Oxford comma (and is it really optional)?—transcript*. Oregon State Guide to Grammar. https://liberalarts.oregonstate.edu/wlf/what-oxford-comma-oregon-state-guide-grammar#:~:text=The%20big%20shift%20in%20this

Calle-Martín, Javier, & Miriam Criado-Peña. (2022). 'A very pleasant, safe, and effectual medicine': The serial comma in the history of English. *Australian Journal of Linguistics, 42*(2), 165–183. https://doi.org/10.1080/07268602.2022.2123700

Casagrande, June. (2023, June 12). *A word, please: AP softens on the Oxford comma, but its rule against the punctuation mark can make sense*. Daily Pilot. https://www.latimes.com/socal/daily-pilot/opinion/story/2023-06-12/a-word-please-ap-softens-on-the-oxford-comma-but-its-rule-against-the-punctuation-mark-can-make-sense

Chafe, Wallace. (1988). Punctuation and the prosody of written language. *Written Communication, 5*(4), 395–426. https://doi.org/10.1177/0741088388005004001

Corbett, Philip B. (2015, June 23). *FAQs on style*. After Deadline Blog. https://archive.nytimes.com/afterdeadline.blogs.nytimes.com/2015/06/23/faqs-on-style-2/

Farkas, David K., & Nettie Farkas. (1981). Manuscript surprises: A problem in copy editing. *Technical Communication, 28*(2), 16–18. http://www.jstor.org/stable/43087342

Garner, Bryan A. (2022). *Garner's Modern English Usage* (5th ed.). Oxford University Press.

McCoubrey, Carmel. (2017, Mar 16). *For want of a comma*. The New York Times. https://www.nytimes.com/2017/03/16/opinion/for-want-of-a-comma.html

Osbeck, Mark K. (2011). What is 'good legal writing' and why does it matter? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1932902

Victor, Daniel. (2017, Mar 16). Lack of Oxford comma could cost Maine company millions in overtime dispute. *New York Times (Online)*. https://search.proquest.com/docview/1877858911

Reactions to Sexual Taboo Language in Married and Unmarried Latter-day Saints

Alayna Beck, Lia King, Megan Beals, and Lauryn Wilde

Taboo frequently surrounds sexual words in religious communities. This study examines the difference in reactions between married and unmarried Latter-day Saints to sexual language, hypothesizing that married individuals are more comfortable with it. Through recording the interviews, the researchers found no significant difference between reading times but many qualitative differences in the subjects' reactions. Married individuals reacted less to sexual words used in the study such as orgasm *and* penis. *More research is needed to understand this area of linguistics.*

Speaking about sex is often taboo in religious communities because abstinence is encouraged by many theological institutions (Navarro-Prado et al., 2023). The Church of Jesus Christ of Latter-day Saints (hereafter referred to as "the Church") is one such institution. We studied the difference between married and unmarried individuals—comfort level discussing sex may alter throughout various stages of life—who live in Utah county, a place where seventy-two percent of the population are members of the Church (Toone, 2023).

Our experiment was informed by Gao (2008): In their study, they analyzed perceptions in the acceptability of taboo language across gender in the sitcom *Sex and the City*. This medium portrays a fictional American experience, examining the root of taboo language in the United States. Their experiment dove into conversations used in the sitcom and specifically examined taboo language, including profanity, euphemism, and sexual language. One finding was that males tended to speak more directly about sex, while females preferred euphemisms. While other studies have been done on profane taboo language and euphemism, not much has been done with nonoffensive language about sex. This study helped us examine sexual language that is not necessarily explicit, as it looked at both explicit and nonexplicit sexual language.

We add to the research of taboo topics by asking the question "How do perceptions of language about sex vary between married and unmarried individuals in a religious community?" We hypothesize a higher comfort level among the married population due to an active sex life.

Literature Review

Previous research shows a connection between the Church and Saudi Arabia. Alqahtani et al. (2022) found that sex terms and swear words in Saudi Arabia are completely taboo, largely because of religious and cultural tradition, helping us understand how the religious community of Utah county affected our findings. Tamimi (2017) conducted a similar study in Iran, but focused specifically on taboo avoidance behaviors, such as code-switching and euphemisms, instead of vocabulary. This study was foundational in forming our hypothesis on which participants would be uncomfortable and how they might deal with the discomfort. More research can be conducted on how the acceptance of taboo

language is changing for younger generations compared to the cultural and social norms of the past. It has been noted that younger generations have begun to destigmatize taboo language and use it more commonly than before, redefining what talking about sex can look like. For instance, Downes (2022) focuses on the avoidance behaviors in children and schools, noting that schools have strayed away from explicit rules regarding taboo language and instead emphasize individual teacher preferences.

Many additional studies have been done on how religion affects taboo language from all over the world. Mocanu (2017) examined euphemisms in European Christianity and identified two categories: euphemisms avoiding topics viewed as sacred and euphemisms taken from politically correct language. Schlosser (2003) focused on religious privilege and the language that accompanies this advantage, especially as a member of a dominant group (e.g., Christians in America). Thus, regarding our research question, there are different opinions in the world of taboo language. We decided to utilize a younger pool of participants and study the implications of marital status through interviews.

Methodology

Participants were eighteen to thirty years old, residents of Utah County, and active members of the Church. We designed an experiment where they would read two passages of text out loud: a control passage and a passage with language about sex. Periodically throughout the experiment, participants would take a survey. For the pilot study, we selected fifteen participants that fell into this age range, though in a full study, a random sample should be chosen.

Before the experiment started, participants completed the first section of the survey (found in the appendix), which included demographic questions, a scale to rate their current mood, and a question about how comfortable they are with talking about sex.

After finishing the first section of the survey, the participant read out loud a passage about how to be successful in school. Each participant was timed and recorded. Once they finished, the participant returned to the survey to reevaluate their mood and comfort level with the researcher and was asked to provide any other comments. The participant then read out loud a passage about healthy sexuality. Once again, the participant was

recorded and timed. Again, they returned to complete the survey to indicate any changes in mood, comfort level, and comments. Both paragraphs were of similar reading levels and similar word counts, and the intention of both paragraphs was to inform. This kept the genre of the paragraphs similar to prevent any discrepancies. All data was compiled and analyzed individually, then compared to the group. An ANOVA test (an analysis of variance to determine significance) was run on the difference between each participant's first and second reading time.

Results

The pilot experiment involved fifteen participants: ten single individuals and five married individuals, all between the ages of eighteen and twenty-five. Across the single participants, there was generally a difference between the time it took to complete the first and second readings—participants 1, 5, 6, 7, and 10 read more quickly, while participants 8, 9, 11, and 12 read more slowly (table 1).

Table 1

Participant reading times and reactions

Participant ID	Gender	Marital Status	Reading 1	Reading 2	Reading 2 Reaction
1	F	S	1:06	1:00	laughter at the beginning
2	M	M	1:06	1:06	none
3	F	M	1:00	1:00	none
4	F	S	0:56	0:56	laughter throughout
5	M	S	1:03	0:59	pause before *vaginal*
6	M	S	1:13	1:03	sweating
7	F	S	0:47	0:45	laughter during the second half

Participant ID	Gender	Marital Status	Reading 1	Reading 2	Reading 2 Reaction
8	M	S	0:50	1:12	pause after *orgasms* and *masturbation*
9	M	S	1:33	4:17	laughter before starting; unwilling to say *arousal* and *orgasms*
10	F	S	1:11	1:09	laughter throughout, especially at *orgasm*
11	M	S	0:52	1:04	laughter at *penis*
12	M	S	0:54	1:01	laughter at *orgasm*
13	F	M	0:55	1:12	laughter at *orgasm*, *condoms*, *penis*, and *vagina*
14	M	M	1:00	0:58	none
15	M	M	1:10	1:08	none

Participant 4, who was single, did not show a change in reading time. Participant 9 showed the greatest difference, taking 1 minute 33 seconds during the first reading and 4 minutes 17 seconds during the second, nearly four times greater than any other participant. This participant also expressed unwillingness to say certain terms and had great difficulty finishing the passage. Among married individuals, participants 2 and 3 did not show any change in reading time, participants 14 and 15 had shorter reading times, and participant 13 had a longer reading

time. However, the difference in reading times across married and unmarried individuals was insignificant ($p = 0.528$).

In observing participants' reactions, we noticed that single individuals had more notable reactions to the second reading than married individuals did. All ten single participants experienced one or more of the following reactions: laughter, pauses, and sweating. Comparatively, four out of five married individuals did not have any notable reaction, but one individual (participant 13) laughed during the second reading.

While the passage was educational in content, it is interesting that all single individuals experienced a reaction to the text, while only one married individual did. Participants laughed or paused at various terms about sex (table 1), but universally we saw that language about sex caused visible discomfort in single participants. The reactions of laughter, pausing, and sweating support our hypothesis that married individuals are more comfortable talking about sex than single individuals are. This pilot was successful in achieving preliminary results and guiding the direction of future experiments.

Discussion

One potential problem with this pilot study is that all researchers involved were female, which could have affected the comfort level in both male and female participants. In the pilot, this could not be resolved because there were no male researchers that were familiar with the project and could perform the experiment, but in a full study this should be accounted for. The sampling from the pilot study was also not balanced between married and single individuals, so having a more even spectrum of participants would strengthen the study. Another potential issue is that the participants' relationship with the researcher may affect their comfort level. A researcher performing the experiment with a participant they have not met before may lead to different results than a researcher who knows the participant well. We are not sure the best way to address this variable, but it should be considered in the planning of the full experiment.

Some people may be naturally more comfortable talking about sex than others, so in future studies, it may be interesting to examine the effect of marriage on a participant's reaction to language about sex. Also, married individuals' comfort with sexual

language is likely affected by the length of time married. Participants 13 and 14 had been married for less than a month, while participants 2, 3, and 15 had all been married a year and a half to two years. In a full study, we recommend studying individuals who have been married for more than a year to give the individual enough time to become settled into married life (including possibly becoming more comfortable talking about sex).

Conclusion

Members of The Church of Jesus Christ of Latter-day Saints seem to treat language about sex as taboo language. Though the quantitative results of this study were insignificant, our initial question of how taboo sex terms are perceived differently between married and unmarried individuals was still answered through qualitative data. All single participants laughed, stuttered, and/or paused while reading the second paragraph out loud, displaying discomfort, while one out of five married individuals also expressed discomfort. These reactions offer some insight to responses to talking about sex, but there are many more sides to this story that should be explored in order to understand more wholly the effect marital status has on individuals' perceptions of language about sex.

References

Alqahtani, Salma M., Najat A. Busabaa, & Hala M. O. Salih. (2022). A change in Saudi attitudes from use of euphemism to taboo: A sociolinguistic study. *Journal of Teaching and Research, 13*(3), 561–571. https://doi.org/10.17507/jltr.1303.12

Downes, Lynn. (2022). "I say it but wouldn't be happy hearing it from a kid": Personal boundary parameters in responding to taboo language, swearing and -IST-ing use. *International Journal of Educational Research, 111*, 101906. https://doi.org/10.1016/j.ijer.2021.101906

Gao, Gao. (2008). *Taboo language in* Sex and the City*: An analysis of gender differences in using taboo language in conversation* [Master's thesis, Kristianstad University College]. http://urn.kb.se/resolve?urn=urn:nbn:se:hkr:diva-943

Mocanu, Mihaela. (2017). Taboo and euphemism in the religious language. *International Letters of Social and Humanistic Sciences, 75*(1), 1–9. https://doi.org/10.18052/www.scipress.com/ilshs.75.1

Schlosser, Lewis Z. (2011). Christian privilege: Breaking a sacred taboo. *Journal of Multicultural Counseling and Development, 31*(1) 44–51. https://doi.org/10.1002/j.2161-1912.2003.tb00530.x

Tamimi Sa'd, Seyyed H. (2017). A sociolinguistic analysis of taboos and euphemisms in an Arab community in Iran. *Dialectologia, 18*, 105–127. https://doi.org/10.1344/dialectologia2016.18.6

Appendix

Survey Questions

First and Last name

(This is for identifying your data and connecting it to the corresponding recordings. It will not be used in the study itself.)

What is your age?

What is your gender?

Male

Female

Other

What is your marital status?

Married with kids

Married

Single

Where would you rank your current mood?

Scale from 1 to 5

How comfortable are you with your interviewer?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

How comfortable are you with your talking about the topic of sex?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

STOP

You have completed the pre-experiment survey.

Please refer back to the interviewer and they'll give further instructions.

You'll finish this survey in a moment. Thank you!

Passage 1

A college education is beneficial in obtaining a well-paying career in a field that you enjoy. This can lead to a fulfilled and purposeful life. The first way to be successful in college is to simply show up to class. Being present in class can help you absorb and understand the material as well as knowing upcoming due dates. It is also important to show up to class prepared, either with a fully charged laptop or notebook and pencil to take notes. Many studies show that taking notes can improve one's ability to memorize the course information. Another way to prove successful

in higher education is to ask questions. It may be valuable to attend the professors office hours for clarification on a topic or assignment. That way, you can ensure you are understanding and applying the information correctly. Refer to the class syllabus for additional questions, clarifications, and ways to succeed. Lastly, practice time management. Try to not procrastinate until the night before a deadline. By avoiding procrastination, you can ensure that you are submitting your best work, not a rushed synopsis. It may also be beneficial to do your homework before it gets too late into the night, so your mind is working at 100% capacity.

Where would you rank your current mood after completing the first reading?

Scale from 1 to 5

How comfortable are you with your interviewer after reading the paragraph?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

How comfortable did you feel while reading the paragraph?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

How comfortable did you feel after finishing reading the paragraph?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

What else would you like to share about this experience?

STOP

Please refer back to the interviewer for your next instructions.

You'll do one more section of this survey after.

Passage 2

A healthy sex life can lead to many benefits, both emotional and physical. It can lead to an enhanced connection with a partner, higher self-esteem, and a longer lifespan. It is important to communicate with your sexual partners and understand how to foster arousal and, later, orgasms. Sex is not one size fits all. It is important to discuss with your partner what you like in the bedroom as well as ways to practice safe sex. Condoms or birth control are frequently used. Likewise, it is important to understand that everyone's sexual anatomy is a little bit different. Your sexual routine will likely differ from partner to partner. This is because each sex organ, namely the penis and vagina, experiences pleasure differently. There are also different forms of sex, namely oral, vaginal, and anal. That being said, many believe that your brain is the most important sex organ. It controls sexual response and arousal as well as housing sexual fantasies. Some people find pleasure in having sex with someone while others get to

know their bodies by masturbating. Learning how to orgasm on your own can make it easier to have one with someone else.

Where would you rank your current mood after completing the second reading?

Scale from 1 to 5

How comfortable are you with your interviewer after reading the second paragraph?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

How comfortable did you feel while reading the second paragraph?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

How comfortable did you feel after finishing reading the second paragraph?

Scale from 1 (extremely uncomfortable) to 7 (extremely comfortable)

What else would you like to share about this experience?

Linguistic Realization of *Yeísmo Rehilado* in Uruguay

Jacob Badal

This study investigates yeísmo rehilado*, the linguistic phenomenon that gives Uruguayan Rioplatense Spanish its unique sound. Specifically, this investigation examines the transition from the voiced [ʒ] to the voiceless [ʃ] postalveolar fricatives. Six audio samples were collected to investigate Uruguayan speakers. Using PRAAT software, instances of* yeísmo rehilado *were analyzed to assess pronunciation differences between urban and rural speakers. The findings reveal that the voiceless variant dominates in Montevideo. At the same time, rural speakers tend to retain the voiced form, with additional sporadic use of regular* yeísmo *[j] near the Brazilian border.*

In 2023, there were an estimated 559.1 million speakers of Spanish, making it the fourth most spoken language in the world (Eberhard et al., 2023). Due to the large geographic expanse of the Spanish language, there have arisen distinct varieties that differ in various regards, such as phonemically, morphologically, and lexically. These changes occurred due to the internal forces within the language or contact with various foreign languages across the history of Spanish.

One of the most notable phonemic shifts in the Spanish language after the *seseo*—the neutralization of the /s/ and /z/ sounds—is what is called *yeísmo* in Spanish linguistics. *Yeísmo* is a term that describes the neutralization and merging of the sounds traditionally represented by the letters /y/ and /ll/. Initially, in Old Spanish, the grapheme (the smallest functional unit of a writing system) /y/ represented either the voiced palatal approximant or the voiced palatal fricative, [j] and [ʝ], respectively. Meanwhile, the sound represented by the digraph (a symbol using two graphemes to represent one sound) /ll/ represented the voiced palatal lateral approximant [ʎ]. This process of neutralization caused the [ʎ] to merge with the phoneme [j], and now populations exhibiting *yeísmo* have completely dropped [ʎ] from their phonologies, but have maintained the grapheme /y/ and the digraph /ll/ in their writing system. Though it is speculated to have been a prominent dialectal feature by the eighteenth century, the precise diachronic process of this phenomenon remains uncertain; nonetheless, it is now a phonological feature currently used by over ninety percent of Spanish speakers (Kania, 2010; Coloma, 2011). The geographical distribution of this feature can also be seen in figure 1.

While *yeísmo* became a prominent phonological feature in the Spanish-speaking world, it underwent further shifts in the Southern Cone. In this region, which traditionally covers Argentina, Brazil, Chile, and Uruguay, there are the massive Río de la Plata and Paraná river basin, whose shores and hinterlands have developed a unique Spanish variety known as Rioplatense Spanish, predominantly spoken in Argentina and Uruguay. While it is difficult to ascertain an exact estimate, Rioplatense Spanish is spoken by millions of people, whose geographical distribution can be seen in figure 2. This variety of Spanish has gained many unique linguistic features due to influences from Brazilian Portuguese, Italian, and native languages such as Quechua and Tupi-Guarani, among others (Ramirez, 2005; Pešková et al., 2011). One of the most notable features of Rioplatense Spanish is the usage of the

voiceless and voiced postalveolar fricatives [ʃ] and [ʒ] in its *yeísmo* in place of [j], a phenomenon called *ʒeísmo, zheísmo, sheísmo,* or *yeísmo rehilado*; for the purpose of this paper we will use the latter denomination.

Figure 1

Geographical distribution of yeísmo

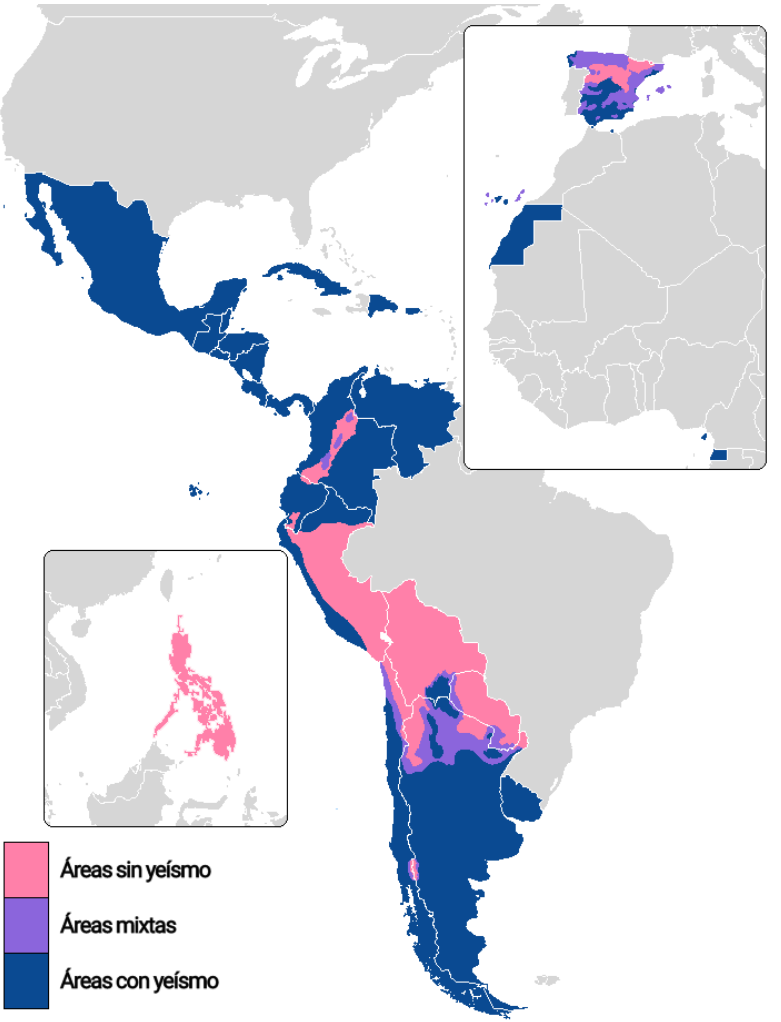
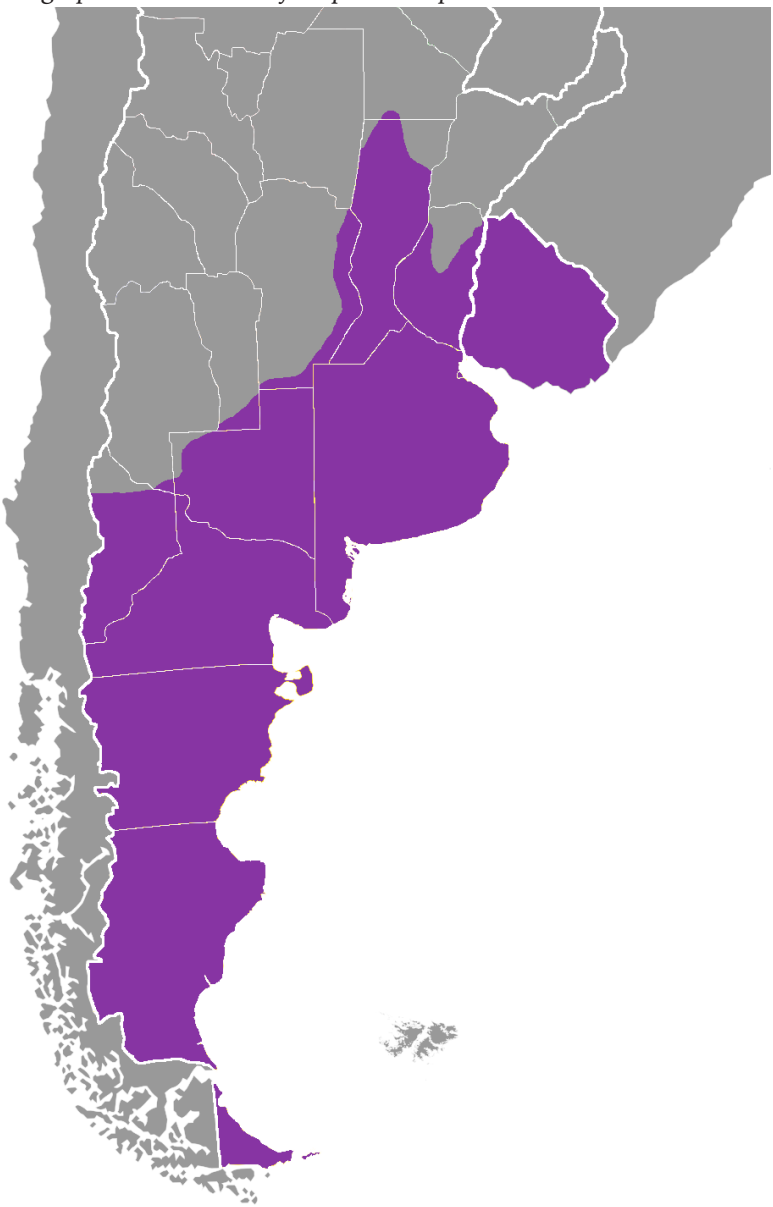


Figure 2

Geographical distribuion of Rioplatense Spanish



Background of *Yeísmo Rehilado*

Just as *yeísmo* refers to the use of [j] for the graphemes /y/ and /ll/, *yeísmo rehilado* refers to the pronunciation of these graphemes as [ʃ] and [ʒ]. Fortunately, the diachronic process of the sound change is much more documented than that of *yeísmo*, in part due to the area's relatively late period of colonization. The first documented appearances of *yeísmo rehilado* in Buenos Aires in the eighteenth century involved the phonemic shift of [j] to [ʒ] and was considered a "vulgar characteristic" of the speech of the lower class, but quickly became part of the prestige standard and spread through the entirety of the Río de la Plata region (Depré & Enghels, 2005). It appears that this transition of [j] to [ʒ] was relatively stable until about the 1930s to the 1940s, when there was another phonemic shift in Buenos Aires from [ʒ] to [ʃ]. This shift was again considered a lower class or pejorative form of speech as it was most likely caused by the large influx of rural immigrant workers to the capital, but the [ʃ] pronunciation was popularized by the presidency of Juan Domingo Perón, who was thought to be an advocate for the poor and working class (Staggs, 2019; Depré & Enghels, 2005).

In a similar manner as the [j] to [ʒ] shift, the move from [ʒ] to [ʃ], or the voiceless style of the *yeísmo rehilado*, spread to the other parts of the Río de la Plata. However, this expansion has been slower as the process of devoicing has been documented as nearly complete across all social classes in the younger population of Buenos Aires since as early as 1979, but only started to manifest in the younger generation and women in Montevideo as late as 2002 (Staggs, 2019).

Methodology and Hypothesis

The purpose of this article is to analyze the process of the voiced to voiceless transition of *yeísmo rehilado* in the Uruguayan variety of Rioplatense Spanish. This will be done by comparing recent videos of self-identified Uruguayan speakers from the capital of Montevideo to videos of interviews done with Uruguayan speakers in the more rural sectors. The analysis will take fifty instances of *yeísmo rehilado* from both sample sets and analyze them through PRAAT, a speech analysis program. It will then compare the rates of voiced and voiceless postalveolar fricatives.

The hypothesis for this analysis is that the speakers from the capital of Montevideo will exhibit complete or near-complete usage of the voiceless postalveolar fricative [ʃ]; in contrast, speakers from rural areas will use the voiced postalveolar fricative [ʒ] predominantly. Furthermore, it is hypothesized that this shift will spread radially from the capital, which will cause a higher percentage of voiceless *yeísmo rehilado* to appear closer to the capital. In contrast, there will be a lower percentage of the voiceless variety appearing in cities that are farther away.

Materials

Since constraints do not allow for fieldwork, this article makes use of videos posted on YouTube to analyze the phonemic variation in question. These videos were posted between six years and five months ago, all after the devoicing transition was documented in Montevideo in 2002. Videos in which the speakers discussed where they were from were used to ensure that the participants were from the correct locations to conduct the analysis properly.

The audio from these videos was analyzed in PRAAT to help in determining if the realization of *yeísmo* was [ʃ], [ʒ], or [j]. The first fifty instances of *yeísmo* were counted for each group, and an attempt was made to balance the counting between participants.

Participants

Three samples were analyzed for both speakers in Montevideo and speakers from the rural parts of the country; the rural samples were taken from varying parts of the country. The features of the participants are as follows.

Participants from Montevideo

Sample 1: Female. Approximate age: 20–30. From Montevideo.

Sample 2: Male. Approximate age: 20–30. From Montevideo.

Sample 3: Female. Approximate age: 15–20. From Montevideo.

Participants from Rural Areas

Sample 4: Male. Approximate age: 30–40. From Rivera.

Sample 5: Male. Approximate age: 60–70. From Melo.

Sample 6: Male. Approximate age: 40–50. From Durazno.

Results and Analysis

Looking at the results from the two groups of participants, there is a definite differentiation between them that followed the hypothesized general trend. The results can be seen in figure 3, showcasing urban versus rural usage and figure 4, centered on individual usage.

Figure 3

Urban vs. rural usage

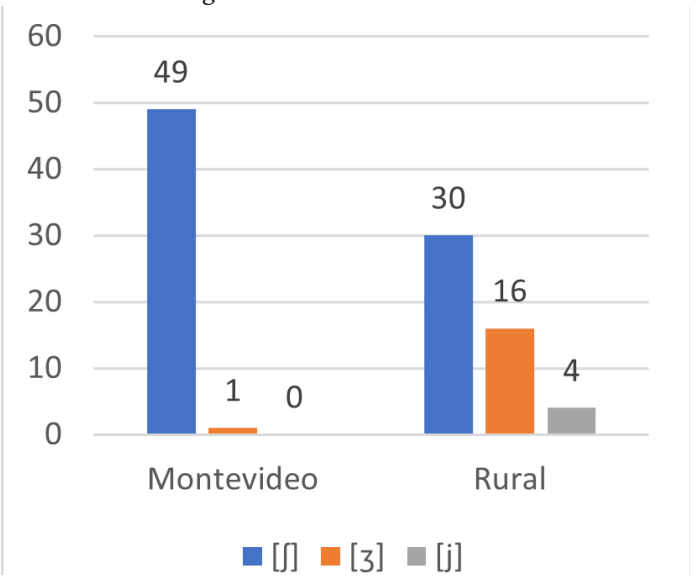
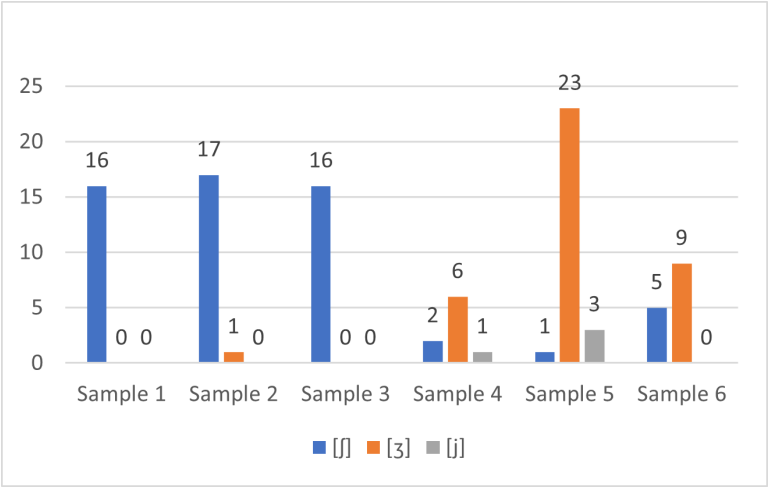


Figure 4

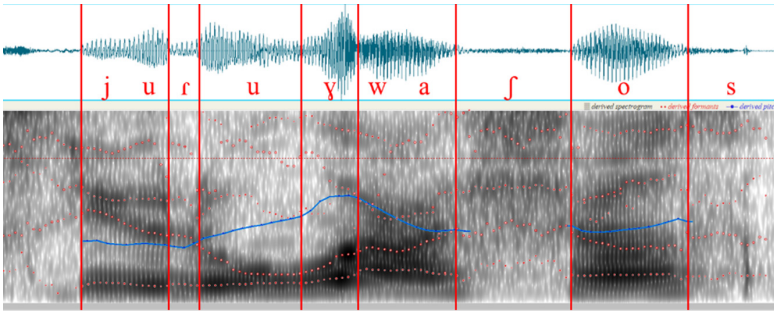
Individual usage



The speakers from Montevideo strongly manifested the voiceless *yeísmo rehilado*. The female participants used only the voiceless form, while the male participant used the voiced form only once—which correlates with the observations noted by Cecelia Staggs (2019). An example of the realization of the voiceless *yeísmo rehilado* can be seen in figure 5, which shows the phrase *y uruguayos* taken from sample 1.

Figure 5

Phonemic realization of y uruguayos



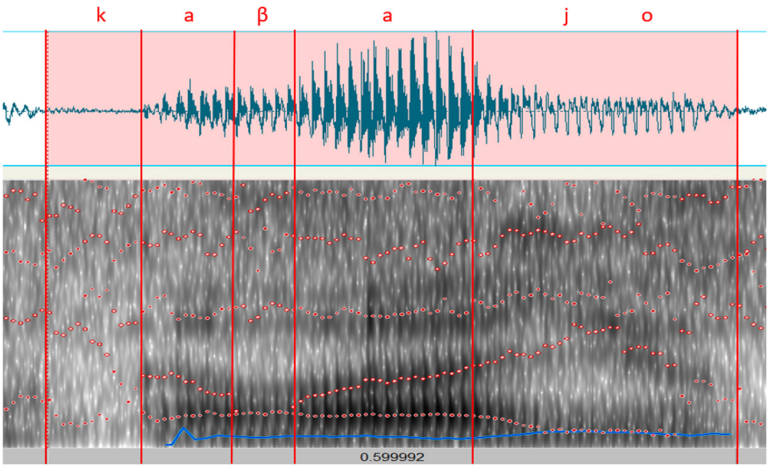
Since this phonemic realization has become so dominant in Montevideo, which is the largest city in the country and controls

all of the media production and dissemination, this variant of *yeísmo rehilado* has become the dominant form in media, which adds to the propagation of the feature in urban regions. This situation is further compounded by the constant influence of Argentine television being broadcast in Uruguay, and many Uruguayans attribute the introduction of the voiceless variant to this fact (Staggs, 2019).

Meanwhile, there was a greater variety of sounds in the rural *yeísmo*. As expected, the rates of voiceless *yeísmo rehilado* were lower than in the capital, totaling to only thirty instances, and there was a greater use of the voiced *yeísmo rehilado*. However, there was an additional usage of the regular *yeísmo* [j] that was not anticipated. For example, in sample 5, the participant consistently produced *caballo* (horse) as [ka'βajo] instead of the anticipated [ka'βaʃo] or [ka'βaʒo]. This can be seen in figure 6, which shows the word *caballo* taken from sample 5.

Figure 6

Phonemic realization of caballo



Noticeably, this usage of regular *yeísmo* occurred in samples 4 and 5, which were taken from speakers in Rivera and Melo, respectively, both of which are located near the Brazilian border. In this area, there is a continuum of Portuguese-Spanish varieties known as *fronteiriço* or as *dialectos portugueses de Uruguay*. One of the most notable phonological features of these varieties is for the Portuguese phoneme represented digraphically as /lh/ to change

from [ʎ] to [j]. For example, this would make *olho* (eye) shift from the standard form of ['oʎʊ] to ['ojo]—a feature which is considered the "most stigmatized variant of non-standard Portuguese" (Carvalho, 1997, p. 646). Therefore, it is reasonable to suppose that there are Portuguese lexical items and phonemic values that are entering the participant's speech.

Another interesting trend is that the participant in sample 6 regularly used voiced *yeísmo rehilado*, despite being from the relatively close Durazno, a city that is only 156 miles from the capital, compared to Rivera and Melo, which are 311 miles and 247 miles from the capital, respectively. If it is presupposed that the shift from the voiced to voiceless variant would spread radially from the capital, then it would be expected that speakers from Durazno would manifest the voiceless variety the most, yet the exact opposite occurred. This would suggest that more study on the exact nature of the spread of the voiced to voiceless *yeísmo rehilado* shift is needed.

Discussion and Conclusions

In reviewing the results of the video data analysis, they showed that the first hypothesis was correct and that there would be a significant difference in the usage of the voiced and voiceless *yeísmo rehilado* in that the younger generation from the Uruguayan capital of Montevideo uses almost exclusively the voiceless variant; meanwhile, the usage of the voiced variant was much more prevalent in rural populations outside of the capital. Nevertheless, the results of this study also showed that the spread of the voiceless form of *yeísmo rehilado* is not spreading in a uniform radial format and that the diffusion of this trait may be dependent on other factors such as education, age, and media usage. Further study in this area is needed to clarify the morphology of this trend.

Finally, while conducting the preliminary research into the topic of *yeísmo rehilado* in Rioplatense Spanish, all of the studies consulted discussed a binary difference in the *yeísmo rehilado* along the voiced/voiceless axis; however, it has become evident that in Uruguay, the phonological landscape is more complex than was previously thought. The results of the analysis have shown that there are, in fact, three realizations of *yeísmo* in the Uruguayan variety of Rioplatense: a voiceless *yeísmo rehilado* [ʃ], a voiced *yeísmo rehilado* [ʒ], and a regular *yeísmo* [j]. This raises further questions as to where all three variants are used geographically

and socioeconomically, if there are social preconceptions attached to the different realizations of *yeísmo* in Uruguay, and what the influences of Portuguese and other Indigenous languages are on these realizations.

References

Carvalho, Ana M. (1997). Variation and diffusion of Uruguayan Portuguese in a bilingual border town. *Actas do I Simposio Internacional sobre o Bilingüismo,* 642–651.

Coloma, Germán. (2011). Valoración socioeconómica de los rasgos fonéticos dialectales de la lengua española. *Lexis, 35*(1), 91–118. https://doi.org/10.18800/lexis.201101.003

Depré, Tristan-David, & Renata Enghels. (2005). *El yeísmo rehilado del cocoliche : Sobre las realizaciones de las aproximantes palatales en castellano porteño* (Publication No. 02549239) [Master's thesis, University of Ghent]. Lexis. https://revistas.pucp.edu.pe/index.php/lexis/article/view/1264

Eberhard, David M., Gary F. Simons, & Charles D. Fennig. (Eds.). (2023). *What are the top 200 most spoken languages?* Ethnologue. https://www.ethnologue.com/insights/ethnologue200/

Kania, Sonia. (2010). Documenting "yeísmo" in Medieval and Colonial Spanish texts. *Romance Philology, 64*(2), 223–234. http://dx.doi.org/10.1484/J.RPH.3.40

Pešková, Andrea, Ingo Feldhausen, & Christoph Gabriel. (2011). Fraseo prosódico en el español porteño. Evidencia de datos leídos y semispontáneos. *El español rioplatense: lengua, literatura, expresiones culturales* (Di Tullio, Angela, & Rolf Kailuweit, Eds.). Iberoamericana Editorial Vervuert.

Ramirez Luengo, José L. (2005). Contacto hispano-portugués en la *romania nova*: Aproximación a la influencia portuguesa en el español uruguayo del siglo XIX. *Res Diachronicae, 4*, 115–131.

Staggs, Cecilia. (2019). A perception study of Rioplatense Spanish. *McNair Scholars Research Journal, 14*(1), 47–56.

Figures

Figure 1: Tetrarca85. (2022). *Mapa con la distribución del yeísmo en el mundo hispanohablante* [Map]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Yeismo_en_Países_de_Habla_Hispana.png.

Figure 2: Fobos92. (2015). *Extención del dialecto rioplatense del idioma español* [Map]. Wikimedia Commons. https://en.m.wikipedia.org/wiki/File:Español_rioplatense.png

Video Material

Below are listed the videos that were used to collect data.

Alaska. (2021, September 21). *PALABRAS ARGENTINAS vs PALABRAS URUGUAYAS*. YouTube. https://www.youtube.com/watch?v=VyihWdE0Umg (8:02).

BBC News Mundo. (2017, November 10). *Cómo puedes diferenciar como hablan un argentino de un uruguayo*. YouTube. https://www.youtube.com/watch?v=fq8l6Fsg_IQ (1:43).

Canal 5 Uruguay. (2018, April 26). *Huellas - Peón de campo*. https://www.youtube.com/watch?v=w9C6bpVslcY (24:23).

Canal 5 Uruguay. (2018, June 21). *Huellas - Tropillero*. https://www.youtube.com/watch?v=Yj_TftptLFo (27:33).

Luancho. (2021, November 10). *Gaucho en Rivera*. https://www.youtube.com/watch?v=yBx5yRThVN0 (2:50).

Profe Nicole. (2023, June 28). *¿Argentina y Uruguay hablan igual?* https://www.youtube.com/watch?v=V2kBjHUETsw&t=68s (10:49).

Appendix

Sound Inventory of Rioplatense Spanish

Below is attached the sound inventory of Rioplatense Spanish. All IPA symbols that appear in parentheses only appear allophonetically and are not independent phonemes.

Consonants

	Bilabial		Labial		Dental		Alveolar	
Nasal		m						n
Stop	p	b			t	d		
Fricative		(β)	f			(ð)	s	
Lateral								l
Tap/Flap								ɾ
Trill								r

	Post-Alveolar		Palatal		Velar	
Nasal			ɲ			
Stop	tʃ				k	g
Fricative	ʃ	ʒ			x	(ɣ)
Lateral						
Tap/Flap						
Trill						

Vowels

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	