Spring 2019

The Distribution of [ʍ]: an Acoustic Analysis of Sociophonetic Factors Governing the Wine-Whine Merger in Southern American English

Keiko Bridwell

THE DISTRIBUTION OF [ʍ]:
AN ACOUSTIC ANALYSIS OF SOCIOPHONETIC FACTORS GOVERNING THE *WINE*-*WHINE* MERGER IN SOUTHERN AMERICAN ENGLISH

by

Keiko Bridwell

Baccalaureus Artium et Scientiae
University of South Carolina, 2017

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Arts in

Linguistics

College of Arts and Sciences

University of South Carolina

2019

Accepted by:

Amanda Dalola, Director of Thesis

Elaine Chun, Reader

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ACKNOWLEDGEMENTS

There are a number of people to which I owe deep gratitude for their support and contributions toward making this thesis a reality. First and foremost, I would like to thank my advisor and thesis director, Dr. Amanda Dalola. Her mentorship and guidance has been invaluable in shaping my identity as a linguist, in discovering my niche in research, in building my academic and professional skillset, and in inspiring a confidence in my own work that made this project possible. I also want to thank Dr. Elaine Chun, who challenged me to think from new perspectives in a way that deepened the substance of my research, and whose input broadened the scope of this project in a valuable way.

I would also like to express my gratitude toward my fellow students in the USC linguistics department for their support and interest in my research. Their encouragement and enthusiasm for hearing each new "hw" installment never failed to reignite my own passion for the topic.

Finally, I would like to sincerely thank my home community in Spartanburg, SC, for their support and willingness to contribute to my research. I deeply appreciate the generosity of the many people who agreed to serve as participants, people who gave of their time and their memories in order to make this thesis possible. In hearing their stories, I have grown to appreciate even more deeply the community I hope to represent through my own work. What I have learned from them goes far beyond anything this or any other paper will ever convey.

ABSTRACT

In US English, the merging of the voiceless labiovelar glide [ʍ] and its voiced counterpart [w] has been an ongoing process over the past century, originating in central port cities on the Atlantic seaboard and gradually spreading to include the bulk of the continental US. While described by many sources as still present in Southern American English, the so-called *wine-whine* merger shows evidence of nearing its completion as its usage becomes increasingly rare even within the Southeast, even as the segment [ʍ] is interpreted as a feature of Southern speech. Despite this fact, very little research has been conducted on the merger, with our knowledge of its extent largely attributable to broad dialect studies which do not focus on any particular linguistic feature. Therefore, the present study utilizes the recorded speech, collected via sociolinguistic interview, of speakers of Southern White American English. In the this paper, the resulting data is submitted to a sociophonetic analysis, in which the presence, duration, and COG of [ʍ] are compared across both demographic and linguistic variables to determine the factors governing its appearance and realization. Results reveal a strong, non-linear relationship between [ʍ] and age, as well as the existence of two sets of social patterns of usage, whereby [ʍ] is simultaneously associated with rurality and localization, as well as the overt prestige of education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

In the field of English dialectology, the *wine-whine* merger refers to the process by which the voiceless labiovelar glide [ʍ], the traditional pronunciation of orthographic <wh>, assimilates to voiced [w]. The distinction between these two sounds, once considered standard, has been lost in many varieties of English worldwide, resulting in the creation of homophonous pairs such as *wine/whine, weather/whether,* and *witch/which*. In the United States, in particular, this merger has been in progress over the past century, such that the geographical distribution of [ʍ] is now largely restricted to the Southeast (Kurath & McDavid, 1961; Labov et al., 2006).

Despite its association with Southern American speech, [ʍ] does not appear to function as a dialectal feature in the same sense as pre-nasal vowel raising (i.e. the *pin-pen* merger) or monophthongization. Speech recordings from 1997 to the present indicate that the merger has continued to progress throughout the United States, such that [ʍ] has been largely eliminated among younger speakers regardless of geographic region (Bridwell, 2018). As discussed in Chapter 3, an analysis of Internet discourse further suggests that [ʍ] is associated with not only region, but age, to the extent that it has become associated with ideological positions such as conservatism among speakers who have adopted the merger, while those who preserve [ʍ] tend to associate its use with

"correct" speech. Such evidence suggests the likelihood of variation governed by age within Southern American English, as well as the potential for more socially nuanced variation depending on the speaker's desired stance. Furthermore, the extended transitional stage in which [ʍ] exists in the South presents the possibility for transitional evidence within speakers, as well, making this feature a valuable target for investigating the nature of linguistic change and the contexts it first affects.

The current study approaches the topic of between- and within-speaker variation in <wh> pronunciation through a sociophonetic analysis of recorded interviews. Its primary focus centers around nature of the relationship between age and [ʍ]-usage, with other demographic and identity-based variables investigated as potential influences on variation between individuals, and linguistic contextual variables as a source of variation within speakers. Uniquely to this study, "variation" here refers not only to the binary opposition between [ʍ] and [w], but to the strength of aspiration in [ʍ], measured across a wide variety of acoustic variables.

This paper is organized as follows. Chapter 2 provides a description of the historical and current geographic distribution of [ʍ], with a particular focus on the methodology and results of previous dialectal surveys. In Chapter 3, the influence of social factors on pronunciation is presented, followed by a presentation of the results from an analysis of attitudes toward [ʍ] in online discourse. Chapter 4 includes a discussion of the principles of sociophonetic analysis, and the grounds for phonetic measurement techniques used in the current study.

This treatment of the relevant literature is followed by a description of the experimental design and variables under consideration in Chapter 5. In Chapters 6 through 8, results from the two major portions of the study are presented and compared. Chapter 9 closes with a discussion of the implications.

Over the course of this paper, it will be demonstrated that the *wine-whine* merger is indeed nearing completion in speakers of Southern White American English, and that its presence among older speakers is governed by a combination of demographic traits and linguistic attitudes. From the data presented, it will be argued that the indexical value of [ʍ] shifts according to age group, being associated with rural Southernness among the oldest speakers, educatedness among middle-aged speakers, and simply dropped by younger speakers. Emergent from the analysis is also evidence for multiple variants of [ʍ], themselves determined by social factors, and clues as to the underlying representation of [ʍ] in the phonemic inventories of its producers.

CHAPTER 2

PREVIOUS RESEARCH ON THE *WINE*-*WHINE* MERGER: HISTORICAL AND CURRENT DISTRIBUTION

2.1 HISTORICAL BACKGROUND

The loss of [ʍ] in American English can be traced back to Britain, where variation in its pronunciation has arguably existed since Middle English (Minkova, 2004). English <wh> is derived from Proto-Indo-European *k^w, which became [x^w] in Germanic through the application of Grimm's Law, by which voiceless stops became fricatives (Chambers, 2002). In Old English orthography, this sound was spelled as <hw>, and possibly analyzed as the biphonemic sequence /hw/ or /xw/ (Minkova, 2004)[1]. By the time of Middle English, the spelling had metathesized to <wh>, and the pronunciation had further weakened to [ʍ] (Chambers, 2002). Minkova (2004) suggests that a shift to voiced [w] may have taken place as early as the thirteenth century in the south of England, with an orthographically-motivated shift back to [ʍ] occurring in the sixteenth and seventeenth centuries. By the eighteenth century, however, the two sounds had merged or re-merged, and the [w] pronunciation has been standard in Britain since that time, with the exception of Scotland and Ireland (Kruse, 2016). While [ʍ] has been

[1] In many linguistic sources, [ʍ] is transcribed as the sequence [hw]. Although Roach (2009) states that there is no clear reason to avoid treating it as a separate phoneme, Minkova (2004) offers a persuasive analysis equating /hw/→/w/ with other instances of medieval "h-dropping", such as /hn/→/n/ and /hr/→/r/. Since this paper deals with phonetic detail rather than underlying phonemes, the symbol [ʍ] is preferred (however, see the discussion on pg. 51).

maintained by some in careful speech (Wells, 1982), verse-reading (Cruttenden, 1994), or highly prestigious speech (see Chapter 3), it is no longer a part of mainstream RP or British English (Upton, 2008; Roach, 2009; Cruttenden, 2014).

While the evolution of <wh> has been well-documented in British English, its transition in American English has been less extensively studied. [w] pronunciations have been attested since the 18th century, as evidenced by the fact that lexicographer Noah Webster (1758-1843) advocated for the use of [ʍ] as a method of distancing American speech from British (Forgue, 1986). However, to date, the only research on the geographical demarcation of the *wine-whine* merger has been confined to two dialect surveys: Kurath & McDavid's *Pronunciation of English in the Atlantic States* (1961), and Labov, Ash, & Boberg's *Atlas of North American English* (2006). These are discussed in more detail below.

2.2 KURATH & MCDAVID (1961)

In *The Pronunciation of English in the Atlantic States* (PEAS), survey data collected in the 1940s was used to create a dialect map of linguistic features in the states along the eastern coast. This map revealed three loci of [ʍ]-loss: 1) a large coastal area of the Middle Atlantic States, from the Hudson Valley to Chesapeake Bay, including Greater New York and Philadelphia and extending inland; 2) a narrow coastal strip of New England, from Boston to the Kennebec in Maine; and 3) a narrow coastal strip of South Carolina and Georgia, including Charleston and Savannah. The authors took this as evidence that [ʍ] was in use at the time the American colonies were established, and

that the later pronunciation was brought to American coastal cities through trade with Britain. Supporting this hypothesis were results for the word "wharf", a seaside term, which was widely pronounced with initial /w/ outside the three regions above. Kurath & McDavid also noted class differences in [ʍ] usage, whereby "cultured" speakers within these areas preferred [ʍ].

2.3 LABOV, ASH, & BOBERG (2006)

From 1992 to 1997, the Telsur Project gathered linguistic data via telephone surveys from 762 speakers in the continental United States and Canada. This study focused on the frontiers of linguistic change by collecting data solely from urban centers, with particular care taken to interview young females. The results, presented in *The Atlas of North American English* (ANAE), combined the observed distributions of a wide variety of linguistic features to identify the borders and characteristics of North American dialect regions.

Results specifically relating to the distinction between *whale* and *wail* revealed that as of 1997, [ʍ] was largely restricted to the Southeastern states: specifically southern West Virginia, western Virginia, North Carolina, South Carolina, and Georgia; parts of Alabama, Mississippi, and Arkansas; and Dallas, Texas (see Appendix A). Following this study, [ʍ] has frequently been treated as a feature of Southern American English. However, this feature did not appear to be particularly robust in any area, and was only preserved by 53.8% of informants within the isogloss (Labov et al., 2006).

2.4 CHAMBERS (2002)

Although no studies have taken a diachronic approach to the course of the *wine-whine* merger in US English, either through longitudinal research or by taking the age of the speaker into account, at least one such study has been conducted on Canadian English. Using data from the Dialect Topography of Canada (Chambers, 1994), Chambers (2002) graphs the trajectory of the *wine-whine* merger in four regions preserving [ʍ]. Within those regions, speakers in the two oldest groups (in their 70s and 80s) represent the starting point, where [ʍ]-usage was approximately 62% in both cohorts. For each subsequent 10-year cohort, [ʍ] decreased by approximately 10% until reaching speakers in their 20s (13.4%) and teens (9.4%), as the merger leveled off to near-completion.

These results have a number of implications for the same phenomenon in US English. First, as in the US, <wh> pronunciation was never unanimous; even at its highest and most stable point, only 62.3% of speakers used [ʍ]. The move toward [w], then, represents not a radical swap from one phoneme to another, but a shifting of relative frequencies. Second, the trajectory of the Canadian *wine-whine* merger exhibits a regular and long-term S-curve: there is no sharp cutoff between age groups, but rather a slow and relatively constant rate of change over time. Finally, Chambers compares three previous surveys that address Canadian English [ʍ]: Scargill & Warkentyne (1972), DeWolf (1992), and Chambers (1998). These surveys, taken circa 1970, 1980, and 1990 respectively, indicate that rates of [ʍ]-usage remain stable across speakers within the same age cohort over time. On the basis of this observation, examining current rates of

[ʍ]-usage should provide relatively reliable estimates of its past usage, as well, and should allow us to be able to estimate the point at which any shift occurred.

It is important to note that, although the aggregate data from Chambers's study display a clear and regular S-curve, it is by no means clear that the same pattern should be observed in the US. When broken down by geographic region, two out of the four areas (Montreal and Golden Horseshoe) showed similar patterns of regular change, while the other two, Ottawa Valley and Quebec City, were more variable. In Ottawa Valley, the rate of decrease progressed more slowly until the 20-year-old cohort, in which there was a rapid increase; in Quebec City, [ʍ]-usage remained at the 60% baseline until the 40-year age group, and the shift rapidly took place over two decades. Chambers explains this as being due to the highly urbanized nature and dense population of the first two regions, which allow for greater facility of linguistic diffusion, while the latter two include a variety of locations, from cities to towns to rural areas. In the US, [ʍ] has largely already been lost in the largest metropolitan areas, and is mainly restricted to the South, which tends to be more rural (Labov et al., 2006). It is then likely that the patterns of [w]-diffusion may differ from area to area, may depend on the isolation of the location, or may even be dependent on social indices of [ʍ] which are not present in Canadian English.

2.5 CORPUS FINDINGS

The first direct observation of the intersection of US dialect region and age in <wh> variation was drawn from corpus data, using recordings from the International

Dialects of English Archive (IDEA) (Bridwell, 2018). From a collection of 113 recordings of speakers across a variety of age groups and dialect regions, analyses revealed the clear presence of both regional and age effects on <wh>-pronunciation, such that [ʍ]-usage decreased with age in southern and western dialect regions, and was nearly absent elsewhere. Specifically, speakers in the three northern regions (New England, Inland North, and North Central) predominantly produced [w], regardless of age, while age effects were visible in the Midlands, West, and South, but most clearly in the South. However, the rate of [ʍ] production approached 0% as DOB increased for all regions, such that all younger speakers used [w], and the only difference between regions lay in the speech patterns of older and middle-aged speakers.

This study confirmed the results of the *wine-whine* merger observed by Labov et al. (2006), such that Southern speakers were the only ones to use [ʍ] a substantial portion of the time. The isogloss on the ANAE map was largely supported, with the exception that speakers from all areas of Texas appeared to exhibit [ʍ], and not just Dallas. However, other Southern areas excluded from the isogloss, such as Kentucky and Louisiana, did not include tokens of [ʍ]. This study, however, took the additional step of adding age-related variation to the geographic map, and reflected the usage of speakers across the urban-rural spectrum, rather than those from large cities alone.

Additionally, the study described investigated variation within speakers, such that all tokens of <wh> words among speakers who exhibited variation between [ʍ] and [w] were examined for patterns along four axes: semantic content, sentential position, following vowel, and word frequency. The only variable to show significant variation

was vowel, such that instances of <wh> preceding [aɪ] were less likely to be produced as [ʍ] than those preceding other vowels. However, the extremely low sample sizes for content words and words occurring sentence-initially left open the possibility of effects in those areas.

2.6 PRESENT STUDY MOTIVATION

While previous studies have examined the regional distribution of the *wine-whine* merger among other linguistic variables, to date there has been no comprehensive study of this variable in American English, nor an investigation of its social correlates. The current study seeks to address that gap through a small-scale investigation of [ʍ]-usage in the environment where its trajectory of variation may expected to be most robust: the rural South. In addition to its focus on age, this study seeks to identify other social sources of variability, including gender, Southernness, and formality, as well as variability due to linguistic context. Finally, while all previous studies have all treated <wh> pronunciation as a binary variable, the current study aims to provide a more nuanced look at its distribution across variables by measuring the acoustic properties of its production, in terms of fricative duration and center of gravity (COG).

CHAPTER 3
SOCIOLINGUISTIC ATTITUDES TOWARD [ʍ]

In light of the distributional findings presented in the previous chapter, it is evident that the *wine-whine* merger must be viewed in terms of the intersection of geographic and age-based variation, rather than either alone. As a result of this complex distribution, the pronunciation of <wh> has the potential to invoke a broad set of ideologies, both in favor of the merger and of the traditional pronunciation. In the following chapter, an analysis of Internet discourse surrounding this topic is presented with the aim of identifying and exploring a collection of these ideologies, as revealed by commenters' reactions to YouTube videos and language forum posts.

3.1 THEORETICAL FRAMEWORKS

On the basis of dialect survey results (Kurath & McDavid, 1961; Labov et al., 2006), corpus data (Bridwell, 2018), and reported observations, evidence indicates that the use of [ʍ] vs. [w] to pronounce <wh> is a linguistic feature that varies both by region and the age of the speaker, such that it may be described as a dialect feature of the older generation and the American South. As such a dialect feature, it carries the potential to develop ideological significance, coming to define membership within its representative speech communities, particularly as the larger culture is perceived as drifting from the values embodied within that community (Johnstone, 2018). As mainstream American

speech increasingly abandons [ʍ], then, the potential increases for it to develop such a significance, and to become representative of Southern or traditional values.

With the rise of mass media, as well as increasing mobility between geographic regions, exposure to multiple dialects is common, and speech forms associated with a particular dialect may become resources for expressing identification with that community (Johnstone, 2018). Conversely, rejection of such a community or the values it represents will frequently be accompanied by rejection of its salient linguistic features. On the other hand, sociolinguistic studies indicate that phonological features acquired early in life (i.e. accent) are often difficult to suppress, even when the speaker desires to do so (Johnstone, 2018); many tokens of [ʍ] by someone whose idiolect naturally contains this sound may be unconscious.

Regardless of intent, accent can be indicative of social information, serving to index a speaker as a member of a particular demographic. As defined by Bucholtz & Hall (2005), indexicality refers to "the creation of semiotic links between linguistic forms and social meanings," which may arise through a variety of processes, including "the use of linguistic structures and systems that are ideologically associated with specific personas and groups" (p. 593-594). For example, the usage of tag questions has been associated with female language, and "g-dropping" may index rurality or blackness. As shown by Inoue's (2004) discussion of Japanese "women's language," it is not even necessary for the group under consideration to actually produce the relevant features, as long as there is a social memory or perception that they habitually do so.

The indexing of identity (e.g. gender, ethnicity, regionality, etc.) may take place directly, with a direct mapping from linguistic form to demographic information, or

indirectly, such as by directly indexing a stance or act that in turn indexes identity. For example, tag questions are often used to index a hesitant stance, which further indexes femininity (Ochs, 1992). In the present case, [ʍ] directly indexes Southernness (among other characteristics); as discussed later, this allows it to indirectly index stances associated with the American South, such as conservative or right-wing political leanings. The notion of multiple, related indices is explored in detail by Eckert (2008), who defines this *indexical field* as "a constellation of ideologically related meanings, any one of which can be activated in the situated use of the variable" (p. 454). In Eckert's analysis, the use of a single linguistic feature may index any of a variety of meanings, depending on the identity of the speaker, context of the speech, and ideologies of the audience, among other factors. As a result, a linguistic feature may develop from an indicator of membership in a population, or *first-order indexical*, to a marker of character, or *second-order indexical*, which occurs when the feature is "internalized in speakers' own dialectal variability to index specific elements of character" (Eckert, 2008, p. 463). Because of this capability for a phonological feature to become linked with ideological meaning, speakers who wish to align or disalign with such an ideology may adopt or drop such a feature, leading to linguistic change (Silverstein, 2003).

The usage of [ʍ] here deviates from normal indexicality in one important context: comedy. In the "Cool hWhip" *Family Guy* scene, [ʍ] is presented as a comical and incorrect pronunciation. When a form is appropriated from its original context and given new meaning, this is referred to as *recontextualization* (Johnstone, 2018). In this case, [ʍ] is taken from the context of its natural production, in which it indexes Southernness or age, and given a new sense in which it represents ignorantly hyperarticulated speech.

Although potentially interpretable as a second-order indexical, developed from the first-order indices of "Southern" or "elderly", the following discussion will argue that this is in fact an independently constructed perception, created by speakers unaware of the original indexical values, yet existing within the same framework of discourse as perceptions based on a conscious awareness of the demographic groups associated with [ʍ].

Finally, the following analysis consists not only of an examination of the characteristics and stances which may be implicitly indexed by [ʍ], but the stances explicitly taken toward its usage. *Stance*, as a term, may be defined as "an act of evaluation owned by a social actor" (Du Bois, 2007, p. 173); it encompasses subjective orientations to discourse, as well as their performance in subsequent discourse. As modeled by Du Bois (2007), stance involves a "triangle" of alignment, by which two subjects evaluate a single object and position it ideologically with respect to themselves. If the second subject evaluates the object in the same way as the first, then they simultaneously align themselves with the first subject. This phenomenon is particularly noticeable in the YouTube comments analyzed below, which frequently reference the stance taken toward [ʍ] by the video creator and express their agreement or disagreement.

3.2 DATA AND METHODS

For the purposes of this analysis, discourse data surrounding the pronunciation of <wh> were collected from a variety of online spaces, including two YouTube pronunciation guide videos and their comments, the comments section of a YouTube clip from *Family Guy*, and five threads on language forums responding to questions about the

pronunciation of <wh>. This variety of sources allows for the collection of data from speakers who might be expected to encompass a broad range of ideologies towards language usage: from those belonging to communities with a shared intellectual interest in language, to educators teaching English pronunciation to non-native speakers, to viewers of popular entertainment commenting on its humorous aspects. By analyzing the stances taken in comments such as these, it is possible to arrive at a picture of the social variables indexed by a linguistic feature, both from the perspective of those who passively experience it and those who consciously observe its distribution.

The two pronunciation guides, published by Rachel's English and ConfidentSpeech respectively, are aimed at teaching the pronunciation of American English to non-native speakers. The Rachel's English channel contains over 500 videos, with a total of over 75 million views and 1.3 million subscribers; the video being examined has over 35 thousand views. ConfidentSpeech, on the other hand, reflects the output of a New York-based accent-reduction company, with a focus on the business world, and only has 45 videos with a total of 209 thousand views (14 thousand for the present video). Both videos under investigation attracted native and non-native speakers of English, but for the purposes of this study, comments reflecting personal experience with the pronunciation of <wh> were targeted.

The language-forum comments were derived from four sources: English Language and Stack Exchange (2 threads), Antimoon, Daily Writing Tips, and WordReference. Of these, ELSE and WordReference are question-and-answer forums dedicated to language-related topics, which are discussed by both native and non-native speakers of English (who typically identify whether they are native speakers of the

language under discussion). Antimoon is a site founded to teach English to non-native speakers; however, the thread under investigation was titled "Do you pronounce the 'H' in 'White'?" and only appeared to attract answers from native speakers. Finally, Daily Writing Tips provides advice for writers on the grammar and style of English. The current article, in answer to a question about the pronunciation of <wh>, is structured around the phonological history of <wh> and the author's pronunciation habits, and elicited responses from multiple readers about their own pronunciation. All seven sources are summarized below, in Table 3.1.

Table 3.1 Online sources for discourse analysis

Abbr.	**Source**	**Type**	**Title**
A	Antimoon	ESL forum	Do you pronounce the "H" in White?
CS	ConfidentSpeech	YouTube comments (language video)	What, Where, When, Why - How to pronounce "wh"
CW	Cool Whip	YouTube comments (humor video)	family guy cool whip
DWT	Daily Writing Tips	writing article comments	Pronouncing Words that Begin with WH
ESE1	English Language and Usage Stack Exchange	language forum	Pronunciation of 'Wales' and 'whales' in Scotland
ESE2			Is it affected to pronounce the 'h' in wh- words such as 'what'?
RE	Rachel's English	YouTube comments (language video)	How to Pronounce WH Words -- what, why, which -- American English
WR	WordReference	language forum	Pronunciation: When is the "h" in "wh" words pronounced?

3.3 SOCIAL AND REGIONAL DISTRIBUTION OF [ʍ]

At the most straightforward level, an analysis of YouTube comments, blog comments, and language forum posts on the pronunciation of <wh> can indicate the ways in which public discourses imagine the social distribution of [ʍ]. In this section, comments were identified in which people identified their personal pronunciation of <wh> (or other speakers' pronunciations that they had personally heard), and named the demographic group producing the sound. The results are shown below, in Table 3.2.

Table 3.2 Distribution of <wh> pronunciation by demographic data

Distinguishes [ʍ] and [w]		
Self	*Region*	*Nations:* Canada (2), Ireland (2), U.S. (2), England (1), Scotland (1) *Regions:* Southeast U.S. (2), mid-America (1) *States:* Tennessee (2), Arizona (1), Arkansas (1), Kentucky (1), Texas (1) *Cities:* Liverpool, (1), Newport (1)
	Age	born 1989 (1)
	Other	EFL learner (4), prestigious school (1), taught by nuns (1), "reasonable" self-taught pronunciation (1) no information (10)
Other	*Region*	*Nations:* Scotland (16), Ireland (8), U.S. (3), Britain (1), New Zealand (1) *Regions:* South[east] U.S. (3), Northern U.K. (1), Southern Midwest (1), outside Southeast U.S. (1) *States:* Arkansas (2), Oklahoma (2), Texas (2), New York (1) *Cities:* London (1)
	Age	*Relatives:* parents (3), grandfather (1), mother, (1), great-aunt (1) *Groups:* American media in 1960s or earlier (1), Canadians in 1950s-60s (1), "people of a certain age" (1)
	Other	posh/snobby/affected (5), conservative (1), "Southern American prestige dialect" (1), Received Pronunciation (1), WASP (1), Taiwanese speakers (1), English major (1), choir (1)

		Individuals: Groucho Marx (1), Mammas and the Papas (1), Terry Jacks (1), Trevor McDonald (1)
Does not distinguish [ʍ] and [w]		
Self	*Region*	*Nations:* England (5), Ireland (1), U.S. (1) *Regions:* mid-America (1) *States:* New York (2), California (1), not Arkansas (1)
	Age	14 years old (1)
	Other	black (1), WASP (1) no information (3)
Other	*Region*	*Nations:* England (7), U.S. (2), Australia (1), Canada (1), New Zealand (1), Scotland (1), Wales (1) *Regions:* South England (1), outside Southern U.S. (1) *States:* Rhode Island (1), California (1) *Dialect:* General American English (1)
	Age	*Relatives:* son/daughter/children (4), grandson (1), mother (1) *Groups:* my [teenage] generation (1), current American media (1), younger speakers (1)
	Other	acquaintances (2), middle class or lower (1), German non-native speaker (1)

The regional results above indicate that within the United States, [ʍ] is most commonly claimed to be produced by speakers from the South, specifically the Southeast, but also including states such as Arkansas, Kentucky, Oklahoma, Tennessee, and Texas. In fact, whenever another state or area was mentioned (see Table 3.3), there was almost always some caveat suggesting incomplete or anomalous usage: for example, the speaker from Arizona claimed to have difficulty pronouncing [ʍ] after the word "the" (WR comment 33), and the individual from New York who produced [ʍ] was both a member of the older generation (the commenter's mother) and an English major who strongly valued correct enunciation (DWT comment 15).

Table 3.3 Exceptions to Southern localization of [ʍ]

DWT #15	[…] Both my parents were born in **New York** (USA), and so was I. My mother was an English major and chastised me mercilessly for not blowing out candles when pronouncing "Why," "When," Whale," etc. The whole issue seriously disgusted me, and I would never in a million years pronounce the H ("HWY," "HWEN," "HWALE"). I'm sorry, but it sounds pretentious to me. […]
RE #91	I grew up in **midAmerica** and it is a wh. The difference, is that it is a soft w vs. a hard w. We have never said the H in front of the W, but always after the W in a soft w sound. Just like SH, we never thought that the H should come in front of the S, so why think it should come in front of the W? But it definitely changes the w to a soft sound, not Whuah but just a soft w, where the lips are not pursed hard as rocks before making the w sound. When the lips are pursed hard, that is called a hard w. […]
WR #33	[in bio: **Phoenix, AZ**] Just a thought ... I've noticed that /hw/ is very hard to pronounce following the word 'the' ... and that for myself I pronounce 'hweels and deals' but 'the weels go round' ... I wonder if many people actually mix up the two pronunciations.

As a whole, the above results indicate that [ʍ] may be produced over a slightly wider range than suggested by Labov et al. (2006), but that its general localization to the South is borne out by self-reported claims and reflected in the public perception of its distribution. The age-related data, on the other hand, provide new evidence that the generational pattern, previously only informally observed, accurately reflects the linguistic situation. The production of [ʍ] was most commonly associated with members of an older generation (parents, grandparents, or the general public in the 1960s), while [w] was identified as being produced by younger speakers. Five commenters mentioned that the inter-generational difference could be observed in their own families, with quotes such as "I have been teased by my grandson because I pronounce the wh" (CS comment

1), and “My parents both use the aspirated sound more than me, and my kids don’t use it at all” (DWT comment 17).

One difficulty does arise in interpreting the comments, in that (as shown by the prevalence of references to Scotland and Ireland), a large number of commenters were from the United Kingdom or other English-speaking countries, while others were foreign language learners of English. It was not always possible to identify the nationality or native language of each commenter, so that some of the age-related results may represent a generational divide that is not equivalent to that in the United States. However, it is worth noting that in both tutorial videos, which teach the production of American-accented English, [ʍ] was identified by the video creators as an acceptable but “old-fashioned” pronunciation of <wh>, suggesting that the above results hold true in the US.

3.4 PRO-[W] IDEOLOGIES

3.4.1 Constructing [ʍ] as outdated

As mentioned above, both English tutorial videos described [ʍ] as an acceptable, but older, pronunciation of <wh>. The narrator of the ConfidentSpeech video, evidently with the intent of presenting an objective, usage-based description of English, states that, “More and more, the trend in recent decades has been *not* to pronounce the W-H in any special way. Now most English speakers in both North America and the United Kingdom pronounce W-H with the same sound represented by the letter W.” This is followed by a statement licensing both pronunciations: “What is the best way to pronounce these words beginning with W-H? It makes absolutely no difference. Both are considered perfectly correct.”

The above quote reflects a stance in which the existence of multiple acceptable linguistic forms is acknowledged, a perspective identified by Chun (2017) as *pluralism*. Here, neither pronunciation of <wh> is given explicit superiority. Although [w] may seem to be implicitly favored by its identification as the most common pronunciation in current usage, it is interesting to note that in the above quote, "what" is pronounced as [ʍʌt]. This is not remarked upon, and it is difficult to tell whether this reflects a preference on the part of the ConfidentSpeech language-teachers for [ʍ], or simply an unconscious reflex of the narrator's idiolect.

In "Rachel's English," on the other hand, Rachel directly offers her opinion that the [ʍ] sound is old-fashioned:

Rachel:	I also think that the way my mom says it – white [waɪt] –
Mom:	White [ʍaɪt]
Rachel:	– is a little old-fashioned.
	…
Rachel:	Okay guys, so that's *my* opinion on how to say W-H words.

In both videos, therefore (although with different epistemic strengths), the dichotomy between [ʍ] and [w] is presented as representing an axis of time, in which [ʍ] is mapped onto "past" and [w] onto "present", as in Figure 3.1.

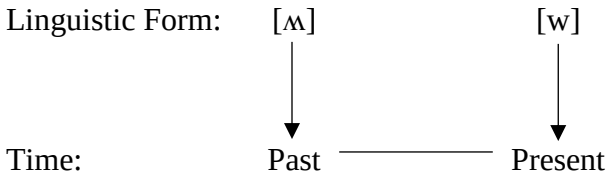

Figure 3.1 <wh> indices modeled for time

A few commenters take a similar position, through either overt support ("I totally agree, it's a dated way of saying it!") (RE comment 10) or through linguistic evidence ("[w] is much more common, exclusive among the younger generations") (RE comment 53). This position is rejected by some who maintain the [ʍ] sound, as will be discussed later; however, some who use [ʍ] accept and take up the prescribed designation: "in Texas we actually still do pronounce them **the old way**" (CS comment 5). This comment is particularly interesting, in that it creates a link between Texas and tradition, mapping Texas onto a past point in time, as modeled in the diagram below.

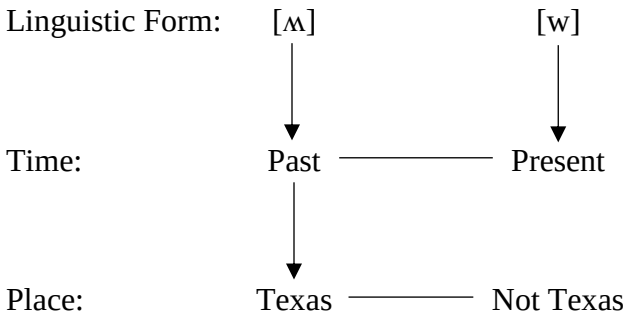

Figure 3.2 <wh> indices modeled for time and place

Although both videos state that [ʍ] is a correct pronunciation of <wh>, labeling it as old-fashioned or rare privileges the alternative pronunciation, further ensuring that [w] will be acquired by new speakers of English. Rachel even overtly recommends this, stating that there is no need for English language learners to bother with the optional [ʍ] sound, since it will be easier for them to acquire a smaller number of phonemes. By

referring to [ʍ] as "old-fashioned," "dated," or "archaic," then, speakers have the ability to recognize its position as having once been the standard pronunciation, while also providing support for [w] in modern-day speech.

3.4.2 Humor

While an analysis of the discourses surrounding a linguistic feature in language-centered settings can provide valuable information about its distribution and acceptability at an academic level, an examination of its use in popular culture, which appeals to a broad range of audiences with no explicit linguistic interest, can reveal even more about the commonly-held perceptions that surround it. Perhaps even more than its overt description as outdated, the clearest indicator that [ʍ] is no longer commonly used by younger speakers is the degree to which it may be used for comedic effect, and the popularity of such comedy. The pronunciation of <wh> as an over-emphasized [ʍ] is the subject of a running gag in the TV show *Family Guy*, beginning with an episode airing in December 2006, in which Stewie annoys Brian by asking for pie with "Cool hWhip." This joke proved extremely popular, being carried over into later episodes and a Wheat Thins commercial, and has led to "cool hwhip" entering the popular lexicon. For this portion of the analysis, comments on the most-watched YouTube video of the "Cool hWhip" scene (currently over 1 million views) were analyzed for their positioning towards the "hw" sound. As shown below, viewers appear to take up [ʍ] not as a regional variant, but as a comically incorrect one, suggesting that for the young demographic to which YouTube and *Family Guy* appeals, [ʍ] no longer exists in their experience as a plausible phoneme.

Table 3.4 "Cool hWhip" comments

Content	Freq.	Examples
Repetition of "cool whip"	35	165 Cool wHip 195 Kewl hwip 203 Coohwhip
Extension of "hw" to other words	6	61 This makes me laugh no matter hwat. 216 huwite people love cool huwip
Comedy of the joke/scene	25	36 One of my favourite Family Guy moments. 115 Ha ha ha ha ha ha xDDD cool "h-wip"
"hw" is incorrect	8	75 Stewie pronounces Cool Whip as Cool Hwhip and Brian is correcting him 183 i used to think this was a joke, but it turns out there's a bunch of people who actually talk this way with 'w's', what the hell? 212 well stewie um...go to speech class it's helping me and it might help you"
"hw" is correct	2	209 Lol the way Brian said Cool whip"
Personal experience with "hw"	8	174 I always pronounce cool whip like that and my friends are just confused. 202 This is me and my mom in the morning when I'm trying to ask for whipped cream on pancakes
Association of "hw" with a demographic group	6	16 My mother pronounces it the same way. I didn't realize this until Thanksgiving. When she said it, my husband, my boys, and I all burst out laughing. No one else got the joke. 95 Well. Stewie does have an old- school British accent, so it's not surprising he pronounces it that way.
Reference to another source	16	64 This is what I think of when I hear Kevin Spacey speak 92 Reminds me of 'The safe word is Whiskey' 206 When Jared Taylor says "hwhite".

As shown above, the most prominent response to the video was that of expressing amusement, either through direct statement or through the recontextualization of "cool whip" as the speaker's own (humorously pronounced) words. This humor did not appear to arise, for the most part, from any association of [ʍ] with a particular group or ideology, but rather from the perception that it represents an absurd, hypercorrected pronunciation of <wh>. On the contrary, when other figures were named, humor typically went unmentioned, except for the possible implication that those figures, in adopting a Stewie-like pronunciation, were revealing their own (potentially humorous) ignorance. This seems to indicate, then, that the majority of viewers who find the scene funny do not come into frequent contact with speakers who produce [ʍ], and may even be unaware that it is a possible pronunciation.

The above inference is borne out by the fact that only one commenter identified as a speaker who used [ʍ]. There was also little indication that most commenters were aware of its regional and geographic distribution: a few mentioned that their older family members used [ʍ], but never generalized this to an entire age group, and none mentioned the American South (although two referenced Stewie's British accent as a possible source of the [ʍ] sound, despite the fact that <wh> and <w> are no longer distinguished in British English). This is perhaps a reflection of the younger audience that is likely to watch *Family Guy* clips on YouTube, who may have spent their entire lives in a linguistic environment in which [ʍ] has been unused. As a result, such listeners, whose primary experience with [ʍ] is now through the *Family Guy* joke, are even more likely to perceive [ʍ] as a deviant and incorrect form, one which should be avoided in their own speech and that may index stupidity or pretentiousness in others.

3.4.3 Invoking negative personas

In addition to being perceived as outdated as incorrect, a subset of comments analyzed in this paper suggest that for some speakers, [ʍ] is associated with explicitly negative ideologies and stances, such that its production may serve to invoke the image of a specific negative persona. Of the outside references in the "Cool hWhip" comments, 7 referred to other humorous "mispronunciation" jokes (5 to the Julien Smith "Malk" video, and 2 to the similar "whiskey" scene in *Hot Rod*). The other 9 mentioned characters or public figures that use [ʍ], including Jared Taylor, Bob Ross, Johnny Cash, and Kevin Spacey (as Frank Underwood in *House of Cards*). Of these figures, Jared Taylor is the founder and editor of *American Renaissance*, a white supremacist online magazine, while Frank Underwood is a corrupt Southern politician. While [ʍ] does not appear to be the most salient feature of their respective dialects (in view of the fact that they were each mentioned only twice in the "Cool hWhip" comments), it may still become linked with negative ideologies, by association with them or with other figures.

Evidence of this perception was observed in several comments (collected in Table 3.2) which labeled [ʍ] as sounding "conservative", "WASP", or "snobby". One commenter even explicitly states that "Only the Southern American prestige dialect makes the distinction in the US" (SE1 comment 4). However, the most overt association of [ʍ] with negative ideologies occurred in the following exchange (A comments 14-18):

Guest [1]:	My relatives from Arkansas pronounce the hw fully. I thought it was so charming when they would say "hwaht" (white). I wish I had hw in my accent, but alas.
Guest [2]:	I also find it charming.

Not from Arkansas: **When I hear someone pronounce the H and they say "Hwaht" for "white" I expect them to put on a white sheet and lynch me.**
Guest [?]: Does that mean you're black?
Not from Arkansas: As the ace of spades.

Here, "Not from Arkansas" positions [ʍ], in combination with monophthongized [ai], as indexing a variety of cultural features that include the strong likelihood of racism. In contrast to the previous commenters, who claim to find [ʍ] appealing despite not possessing it in their own dialects, "Not from Arkansas" distances themselves from those who use it to the greatest extent possible, even using their screen name to identify as a non-member of that speech community. For any readers of this comment who distinguish [ʍ] and [w], then, this comment may suggest that they should exclude the phoneme from their idiolect, if they wish to be perceived as racially tolerant, or even to avoid evoking feelings of fear in others. A compiled description of the indices of [ʍ] may then be modeled as below:

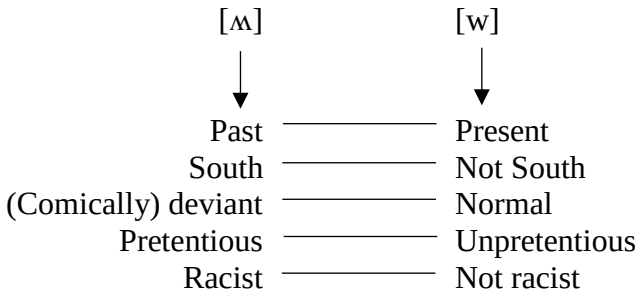

Figure 3.3 <wh> full index model

3.4.4 Indexical field

Although the perception of [ʍ] as humorously incorrect appeared to be largely held by speakers with little awareness of its social indices, this stance was not held fully separate from discourses surrounding these variables. In the comment "When Jared Taylor says 'hwhite'" (CW comment 206), the humor of Stewie's deviant pronunciation is associated with a real-life figure, such that the image of the *Family Guy* scene is evoked when he speaks. This, then, contributes to the indexical mapping of [ʍ] in two directions: by association with a white supremacist figure, the feature is further linked racism, traditionalism, and Southernness; by association with *Family Guy*, it is linked with ignorance and pretentiousness. Although comments like this one, explicitly bridging both perceptions of [ʍ], are rare, its overall mapping creates a setting similar to that seen in Figure 3.4, which visualizes the second-order indexicals (i.e. character traits) that may be attributed to speakers based on their pronunciation of <wh>.

When multiple indices within the boundary shown below are activated at once, and come to be associated, a picture emerges of the South (and by extension users of [ʍ]) as not simply traditional, but "backwards". This may contribute to a perception of [ʍ] as an embarrassing dialectal feature, as evidenced by a comment on the WordReference forum: "There are certain English dialects which pronounce wh differently from w, but the majority don't. Moreover, most speakers of those dialects are (painfully) aware of that fact" (WR comment 74). According to this view, speakers who produce [ʍ] are not only consciously aware of doing so, but are also aware that [w] is the predominant pronunciation in the U.S., and can be expected to accept that their own speech is less

correct. Though this claim will be challenged in the following section, it does reveal the existence of negative ideologies surrounding the production of [ʍ].

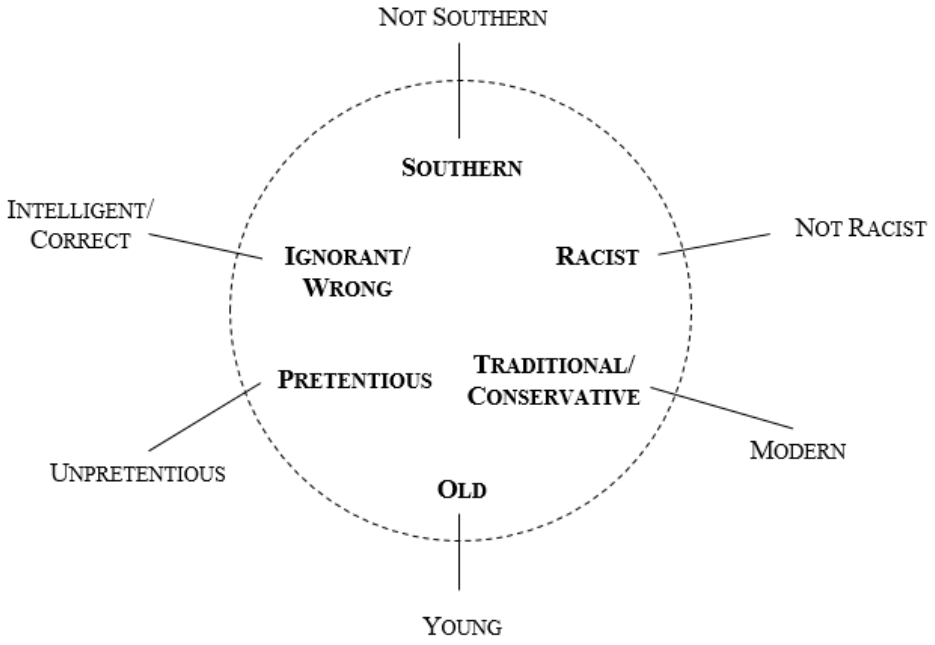

Figure 3.4 <wh> indexical field; ***bold = [ʍ]****, plain = [w], black lines = individual axes of indexical meaning, dotted line = field of [ʍ] indices*

3.5 PRO-[ʍ] TRADITIONALISM

While the presence of "pro-[w]" ideologies can be identified in a large proportion of the comments, the strongest and most aggressive stances taken toward the topic originate from speakers who preserve the distinction between [ʍ] and [w], and maintain that all other speakers should do so as well. These arguments were almost entirely confined to the comments for the YouTube language tutorials, since linguistic

prescriptivism appeared to be considered inappropriate on language forums (two posts favoring the preservation of minimal pairs appeared on English Stack Exchange, but were both downvoted) (ESE2 answers 5-6). YouTube viewers, in contrast, are not expected to abide by academic guidelines, and so were free to react to the creators' authoritatively-stated claims that [ʍ] is considered "old-fashioned" and that [w] is equally or more correct. Their arguments typically fell into one of four categories, as shown in Table 3.5:

Table 3.5 Traditional prescriptivist ideologies in YouTube comments

Content	**Freq.**	**Examples**	
the younger generation is wrong/uneducated	4	CS #7	I believe that misspeaking something fir a long time should not make it acceptable; that's why we have schools. Unfortunately, English teachers are often taught incorrectly. I am not one who jumps from a bridge just because everyone else is doing it.
		CS #12	I pronounce it the right way WH cause I was taught English correctly. […]
		CS #15	Today, most English speakers are illiterate.
		RE #94	The vast majority of us have never been taught how to enunciate English correctly because it has been dropped from schools in both the United States and Canada. Correct enunciation (not "pronunciation") for "wh" is "hw".
a distinction between <wh> and <w> is necessary to distinguish minimal pairs	4	CS #2	i think we should encourage the traditional pronunciation to avoid the wear/where (etc) confusion..
		RE #66	I use the 'wh' and it is the right way in England, Canada, Australia etc. - however like the 'ing' changed/ changing to the 'in' the 'wh' is going the same way -

			however it does show the difference between weather and whether in this case the ‘wh’ has value showing the difference between these two words. […]
		RE #75	I agree with the mother cause that's the proper way to say it. Like if you why and don't pronounce it like the mother then your saying the letter Y. So Y do you to this Wen you can say it properly and pronounce the H!! […]
		RE #92	Your mother is correct, of course.Watt is not the same as what. Nor is wail the same as whale.
pronunciation should mirror orthographic representation	3	RE #75	[…] My were you saying the letter Y for? It's Why?! If it was Y for asking a question it would be spelt that way. Congrats to the mother, I wanna meet her one day and just High five her
		RE #95	[…] With this logic of ignoring the combo letters, we may as well stop saying Shhh for SH and start saying SSS instead […]
<wh> should be pronounced with its historical pronunciation	1	RE #96	I have always pronounced ‘wh’ the way your mom does. I researched it, and it is the original pronunciation. We are right. We preserve the ‘wh’ sound and everyone else lost it.

However, another set of commenters assumed readers’ acceptance of [ʍ] as the genuine pronunciation, and instead addressed the reason for its falling out of use. Almost exclusively, they determined that [ʍ] was being lost because schools are failing to perform their job of educating children in the English language.

This category of comments reflects an understanding of language in which correct speech is not acquired naturally, but must be explicitly taught in order for children to

learn it properly. One commenter (CS comment 7) further argued that children were learning English wrongly because their teachers were being badly taught: “Unfortunately, English teachers are often taught incorrectly.” According to this ideology, correct pronunciation is not something that can be passed down by the average adult; it requires a professional trained in the nuances of language. In this sense, the English language is not a shifting construct, but a stable fixture to be carefully cultivated and maintained to preserve its purity.

Despite the pushback in the above comments toward the idea that [ʍ] is incorrect or perceived negatively (even with such a mild label as “old-fashioned”), commenters generally appeared to not only fail to convince their audience of the superiority of [ʍ], but to reinforce the perception of a persona marked by conservatism and resistance to change. Prescriptivist comments were rarely engaged with, despite their frequently confrontational tone (e.g. “Today, most English speakers are illiterate” in CS comment 15). However, one commenter (responding to CS comment 12) explicitly denied the correctness of prescriptivism: “you’re actually saying it the wrong since language is defined by the majority, pronounce it any way you want but arguing that the old pronounciation is correct is always going to be a losing battle... unless of course wh makes a comeback” (CS comment 13). Since commenters advocating traditionalism frequently identified themselves as members of the older generation by referencing the degeneration of education over time, and showed themselves to be resistant to linguistic change, their comments served to verify the accuracy of the “old”, “traditional/conservative”, and possibly “pretentious” axes in Figure 3.4. This opens up the likelihood for the attribution of other negative characteristics within the indexical

field, such as racism or ignorance, which would further position [ʍ] as a desirable feature to avoid.

3.6 LINGUISTIC TOLERANCE

The final section of this analysis deals with the intersection of pro-[ʍ] and pro-[w] ideologies as exhibited in the form of *linguistic tolerance*. The stance described by this term was characterized by a perspective similar to pluralism (cf. Chun, 2017), in which both [ʍ] and [w] were recognized as acceptable pronunciations of <wh>, but differed in that it frequently involved the prioritization of one acceptable pronunciation over another. This phenomenon was most clearly observed in the YouTube tutorials, which explicitly stated that "both [pronunciations] are considered perfectly correct" (RE), but was also visible also in the comments of viewers, many of whom claimed to prefer the [ʍ] variant, but recognized other pronunciations as legitimate.

The most striking example of linguistic tolerance occurs in the "Rachel's English" video, at a point where Rachel and her mother get into a friendly argument about whose pronunciation is correct:

> Mom: I don't think it's old fashioned. [chuckles]
> Rachel: [laughs] What do *you* think it is?
> Mom: [smiling] I think it's the way to say W-H!
> Rachel: Oh, it's *a* way to say W-H.
> *[cut away from clip with mother]*
> Rachel: Grammar note: when you say *the* way, it means there's only one way. When you say *a* way, it means there is more than one way. Mom thinks her way is the *only* way, the *right* way. But *I* know that *both* are acceptable pronunciations.

In this exchange, Rachel affirms the correctness of both pronunciations for the benefit of her English-learning viewers, as well as native English speakers from a range of dialect backgrounds. However, in doing so, she also asserts her linguistic superiority over anyone less open to accepting both pronunciations as accurate, a stance which several viewers take exception to. One particularly challenges her claim that both are correct, saying, "I pronounce it the way your mom does. I know many younger people are pronouncing it your way, so maybe it's changing. You insist your way is correct also, but you don't indicate what your authority is" (RE comment 87). With this quote, the commenter presents their openness to linguistic change, while simultaneously proposing that a "correct" form (or forms) does exist, independent of majority usage, and requests evidence that [w] fits the requirements for such a form.

Another commenter, taking a different approach to linguistic tolerance, described their choice not to use [ʍ] due to the fact that it was laughed at by some speakers, claiming, "It's more 'correct' but the goal's to sound like an english--speaker rather than be correct, and in my opinion it's not worth the hassle" (RE comment 71). For this speaker, a correct form does exist, but the desire to be perceived as part of the English-speaking community is prioritized above prescriptive rules.

Finally, several commenters exhibited an interesting approach to the [ʍ] sound, one characterized by appreciation for its aesthetic qualities, regardless of whether they themselves used it or not. Such commenters described [ʍ] as "charming," "cute," or "beautiful," as shown below:

Table 3.6 Appreciation for [ʍ]

A #14	My relatives from Arkansas pronounce the hw fully. I thought it was so charming when they would say “hwaht” (white). I wish I had hw in my accent, but alas.
A #15	I also find it charming.
RE #5	Your mom's pronunciation way is so cute :)
RE #63	Your mom pronounces the ‘wh’s’ beautifully. It's a pity that many English speakers don't know about this subtle sound. For me, it makes ‘wh’ words earthier and more interesting. Old fashioned?...no way.
CW #94	Whip, when it was first spelt that way, was actually pronounced as Stewie pronounced it, and was probably spelt hwip (and pronounced the same way) just like hwere changed to today's spelling where. We no longer have that affricate sound. I wonder if the writers knew this :P :(as a linguist I shed a tear for all lost phonemes :(

In the first three comments listed above, [ʍ] is treated as appealing to outsiders, but also as quaint and as confined to a small community, whether that be geographic (A comment 14) or age-related (RE comment 5). While these communities appear to be viewed positively, they are still positioned as deviating from the linguistic norm. In RE comment 63, [ʍ] is not associated with a particular group, and is even explicitly labeled as not being old-fashioned, despite the fact that the commenter acknowledges it is not widely used. Here, the pronunciation of <wh> as [ʍ] is privileged not because it is the “correct” usage, but because it is intrinsically appealing. Finally, CW comment 94, posted by a self-identified linguist, offers information about the historical background of <wh>, and its original pronunciation as [ʍ]. Unlike RE comment 96 (see Table 3.5), the historical pronunciation is not treated as grounds for preserving [ʍ], but rather as an interesting etymological fact. The commenter, in stating “We no longer have that

affricate sound," displays the perception that [ʍ] belongs to the distant past (perhaps Old or Middle English), exhibiting either a lack of awareness that it is still preserved by some speakers today, or excluding them from the general population of English speakers. As a whole, then, acts of linguistic appreciation demonstrate positive attitudes toward [ʍ], but further position it as outside the norm of English language use in the U.S.

3.7 CONCLUSIONS

In summary, this analysis offers evidence that the sound [ʍ], once common in US English, has merged with [w] across much of the population, being now limited to Southern geographical regions and the older generation. Because of this distribution, [ʍ] may be perceived as outdated or comical, and may lead to the perception of those who produce it as pretentious, snobby, or (by association with the South) racist. Although a smaller community argues for the correctness of [ʍ], it presumably consists of members of the previously identified groups, and this discourse is not taken up by the larger online community. Even among those who take a pluralist approach, licensing both pronunciations of <wh> as legitimate, [ʍ] tends to be viewed as the marked variant: an acceptable linguistic form, but uncommon and lying outside of Standard American English.

The above ideologies shape an overall perception of [ʍ] in which it is understood to be non-standard, either as a regional or age-related variant, or as simply wrong. Perhaps most significantly for its future as a linguistic feature in American English is its portrayal as humorously incorrect, a stance which is mainly taken by forms of media which appeal to younger speakers. This positioning simultaneously serves to distance

[ʍ] from any social indices, as well as to delegitimize its production within the contexts where it naturally occurs. This second effect, as modeled in the focus of this analysis, positions [ʍ] as an ignorant pronunciation of <wh>, and combines with other indexical axes to mark those speakers who use it as "backwards." Both these effects make [ʍ] less likely to be taken up as a means to signal in-group membership, and more likely to be perceived as simply incorrect. If understood as such, it is unlikely to be adopted by younger speakers, hastening the merger which is already in place.

Although speakers exist who support [ʍ] as the correct (and only correct) pronunciation of <wh>, they typically do so via authoritative claims that do not align with their audience's personal experiences. By advocating for traditional language usage, rather than accepting its current state, they further promote the view of [ʍ] as indexing conservatism or outdatedness. This may encourage those who are in the process of losing the phoneme, or those who are speakers of multiple dialects, to avoid it in the future, as associated with this type of persona. Overall, therefore, although [ʍ] still exists among a substantial number of speakers, these appear to belong to a limited demographic that is decreasing in number, with its distribution likely to further reduce in the coming years: a phenomenon which may be attributed both to the phonological pressure of a decreasing number of [ʍ] exemplars, and to the social pressure of its association with predominantly negative traits.

3.8 PREDICTIONS

In terms of [ʍ] usage and perception among the group with which it is most closely associated, older Southern Americans, the above findings have several, often

contradictory, implications. If speakers are aware of the nexus of social meaning associated with [ʍ], then it is reasonable to assume that this feature might be used as a tool to express Southern or traditional identity, similarly to commonly recognized features of Southern American English such as monophthongization, "g-dropping," or the lexical item "ain't." In this case, we would expect to see greater usage of [ʍ] in informal contexts, especially those which excited strong emotion or elicited nostalgic associations with place. However, those speakers who reported using [ʍ] almost exclusively claimed to view it as the correct pronunciation, not a dialectal variant. Such speakers would be unlikely to treat [ʍ] as an intentional index of Southern identity, and would only use it as a natural part of their phonemic inventory, with little variation across contexts.

On the other hand, many speakers who do not natively possess [ʍ] appear to have some level of conscious awareness of its social indices. It is therefore possible that, even if the presence of [ʍ] were not used to express Southernness, its absence (i.e. the use of [w]) might be used to distance a speaker from a Southern persona. This could further contribute to the phoneme's diachronic loss, as it would mean that those committed to a local identity would make little effort to preserve the sound, while those who disliked that identity would actively avoid it.

Finally, those speakers who possessed [ʍ] repeatedly stated their belief in its "correctness", frequently calling on the concept of bad education to explain its current decline. This raises the possibility that, rather than treating [ʍ] as a marker of local identity, those speakers who possess it will use it to express education level, authority, or formality. If that is the case, then we would expect to see variation in <wh>-pronunciation among older speakers in the opposite direction to that described

previously, with [ʍ] produced more frequently or more strongly in formal or high-attention contexts.

The current study seeks to understand the underlying indices and ideologies associated with [ʍ] by determining which of the above scenarios most closely aligns with reality. To achieve this, a two-task design is used, such that participants begin with a more formal reading task, and progess to an informal interview. By comparing patterns of [ʍ]-production across tasks, this study aims to tease apart the relative importance of the regional and age components of [ʍ]-indexicality among those speakers who actively utilize it to express their identity.

CHAPTER 4
THE SOCIOPHONETIC APPROACH

4.1 VARIATIONIST FRAMEWORKS

Beginning with Labov's (1963) study of diphthong-raising in Martha's Vineyard, the goal of sociophonetics, and more broadly variationist linguistics, has been to uncover socially-driven patterns in the variable use of linguistic forms via empirical observation. Studies conducted within this framework rely on the understanding that, rather than exhibiting a categorical preference for one linguistic form over another, much of social variation stems from stratified levels of relative frequency between social groups (Labov, 1966). As a result, quantitative analysis is necessary in order to determine patterns of variation. Following Labov, Yaeger, & Steiner's (1972) seminal work on the Northern and Southern Cities Vowel Shift, acoustic data has grown increasingly central to sociophonetics, inasmuch as it can precisely indicate the distance between a feature participating in a shift and its final form, as well as provide evidence of the degree to which an individual participates in this shift.

Although the majority of sociophonetic studies are synchronic; that is, they display a snapshot of language at a single point in time, Labov (1975) also popularized the notion of "the use of the present to understand the past." The *apparent-time hypothesis*, in particular, is based on the idea that, when other factors are held constant,

linguistic differences between generations will mirror the actual timeline of diachronic change (Bailey, 2002). Age-stratified variation may therefore be taken as an indicator of a change in progress. As described in Chapter 2.3, Chambers (2002) has used this approach to look at the *wine-whine* merger in Canadian English, with results from a 1994 dialect survey indicating a regular decrease in [ʍ]-usage over time, such that the ratio of speakers using [ʍ] dropped by approximately 10% with each successive 10-year cohort.

The apparent-time hypothesis relies on the assumption that language reaches a stable point in adolescence and does not subsequently change, an assumption which has been disproven in some contexts. Age-grading, or the adoption of specific speech patterns in order to bring one's own speech in line with adult norms, was documented by Macaulay (1977), who demonstrated that by the age of 15, middle class children ceased to produce glottal-stop variants of /t/ (an indicator of lower socioeconomic status), and persisted in this new usage until adulthood. However, Chambers argues that such a change will present as a dramatic shift around adolescence, and that the Canadian <wh> data follow a linear pattern that is indicative of a change in progress (2002).

On the other hand, longitudinal studies investigating the speech of prominent personalities have revealed gradual changes in speech that occur long after adolescence. In an analysis of Queen Elizabeth II's Christmas broadcasts over a 50-year period, Harrington (2006) found that the queen exhibited tensing of the final vowel in "happy" (i.e. [ɪ:] to [i:]), a shift in line with that of the general British population. This suggests that speakers may adjust their speech patterns well into their lifespan.

With respect to [ʍ], Chambers (2002) tested for the presence of within-speaker diachronic change through a review of data collected from three unrelated studies of Canadian English. He found that the percentage of [ʍ]-usage within a single age-cohort remained stable from 1970 to 1990, even as additional younger groups increasingly switched to [w]. For the current study, although it is acknowledged that within-speaker change may still apply to [ʍ]-usage in varieties of American English, any changes in adult speech toward community norms should result in a smaller differential between generations. Therefore, any error should lie on the conservative side.

4.2 PHONETIC MEASURES

Although the acoustic coordinates of vowels are well understood to differ across individuals, context, and register, evidence shows that consonants also exhibit similar variation in the same situations. In unstressed syllables and informal speech, consonants will show shortened duration, lower COG, and reduced articulatory precision, indicating that they undergo reduction in the same way as vowels (Van Son & Pols, 1996). Fricatives, in particular, show differences in duration and spectral information on the basis of factors including word-position, stress, focus, and task type (Crystal & House, 1988; Silbert & de Jong, 2008; Maniwa & Jongman, 2009). Furthermore, multiple studies have demonstrated that the duration and spectral shape of fricatives are significantly affected by social factors, including gender (Gordon et al., 2002; Heffernan, 2004), sexuality (Zimman, 2017), socioeconomic status (Stuart-Smith, 2007), or L1 (Dalola, 2017). For these reasons, this study uses acoustic data not only to categorize

segments as tokens of [ʍ] or [w], but to investigate how individuals fall on a continuum of relative aspiration strength.

The methods of fricative analysis for this study, which utilize the spectral moments defined by Forrest et al. (1988), involve the collection of spectral data at each quartile point (25%, 50%, 75%), of the segment (cf. Erker, 2010; Lee & Jongman, 2016; Dalola & Bridwell, in press; Dalola & Bridwell, in progress). This method was chosen due to the fact that the acoustics of [ʍ], a glide, may be affected by the identity of the following vowel. Since it was initially unclear whether a standard measurement at the midpoint of the segment (cf. Fuchs & Toda, 2010; Gordon et al., 2002; Koenig et al., 2013) would be early enough to accurately show the characteristics of pure [ʍ], measurements at each quartile were analyzed. Due to the fact that oral cavity size is irrelevant for frontal fricatives (Schwartz, 1968), biologically-motivated gender differences are not expected to occur; therefore no normalization is conducted for these measures.

CHAPTER 5
METHODOLOGY

5.1 EXPERIMENTAL DESIGN

5.1.1 Participants

To test the effect of age on [hw]-occurrence and quality in speakers of Southern White American English, the current study utilizes data collected from sociolinguistic interviews. Twenty-six (26) participants, native speakers of English who self-identified as having been born and raised in the South, served as informants (see Appendix B). These were selected so as to balance the resulting corpus by age and gender, with the goal of interviewing four speakers (two male, two female) to represent each decade. This ideal corpus was closely approximated, as shown in Table 5.1.

Notably, researchers who have studied [ʍ], or American dialects in general, have primarily gathered data from white speakers. The same is true of the current study, which sought to draw from a demographic with robust [ʍ], a sound which is not common in African-American English (AAE). However, this paper recognizes the existence of regional and sociolinguistic variation in AAE, as well as ethnic-based variation within geographic regions, and seeks to make no claims about the "South" as a whole that are not generalizable from the data at hand. For the remainder of the paper, the variety of English under consideration will be referred to as Southern White American English

(SWAE), and will be held to represent no broader group than white speakers in the Southeastern United States.

Table 5.1 Informants by decade, gender, and date of birth (DOB)

DECADE	**FEMALE**		**MALE**	
	DOB	*ID*	*DOB*	*ID*
1930s	1936 1939	P8 P4	1933 1936	P27 P7
1940s	1942 1947	P6 P9	1941 1948	P21 P18
1950s	1953	P20	1950 1951 1951	P3 P13* P16
1960s	1960 1960 1963	P22 P25 P1*	1962 1964	P5 P2
1970s	1977	P19	1975 1979	P15 P26
1980s	1984 1989	P17 P23	1984 1986	P24 P10
1990s	1997 1997	P11 P14		
TOTAL		**13**		**13**

**Missing/low data in one portion of sociolinguistic interview*

5.1.2 Stimuli and procedures

Each individual participated in a sociolinguistic interview consisting of three parts: (1) a reading task, (2) an informal interview, and (3) a set of demographic

questionnaires. Recordings of the reading task and interview were obtained via a Shure WH20 head-mounted microphone and an Olympus Linear PCM LS-10 Recorder.

The reading portion of the procedure consisted of six narrative passages displayed on an Asus R515M laptop computer via PowerPoint presentation software. Passages contained a total of 43 <wh> tokens drawn from 37 unique words, with the words *what, when, where, why, which,* and *whether* occurring in both sentence-initial and sentence-medial position. (Passages are presented in Appendix C, and properties of the target words in Appendix D.) Informants were directed to read aloud as naturally as possible, and to move from passage to passage at their own pace.

The reading task was followed by an interview conducted by the researcher, a 23-year-old native speaker of SWAE from Spartanburg, South Carolina, who suppressed her own pronunciation of [ʍ]. The questions for the interview were drawn from a series of biographical questions (see Appendix E) intended to elicit narratives about past memories, a setup designed to focus the attention of participants on the content of their speech rather than its linguistic features, and to evoke nostalgic associations of place and time that would encourage localized and informal speech. Informants produced an average of 34.2 words containing orthographic <wh> (SD=19.3). Following the interview, informants filled out a demographic questionnaire (see Appendix F).

5.2 MEASUREMENTS

Recordings were collected with an Olympus Linear PCM LS-10 Recorder via a Shure WH20 head-mounted microphone, digitized at 44 kHz, and downloaded into Praat

(Boersma & Weenik, 2017). Interview files were transcribed by the researcher, who identified all instances of orthographic <wh> (excluding words with initial [hV], such as *who* and *whole*) and coded each token for word-level variables. Each instance of <wh> from the reading and interview recordings was subsequently identified and delimited in Praat (Boersma & Weenik, 2017) and inspected for the *presence of frication*, a binary variable in which any labial glide with measurable frication was counted as a token of [ʍ], and all tokens with no frication were counted as tokens of [w].

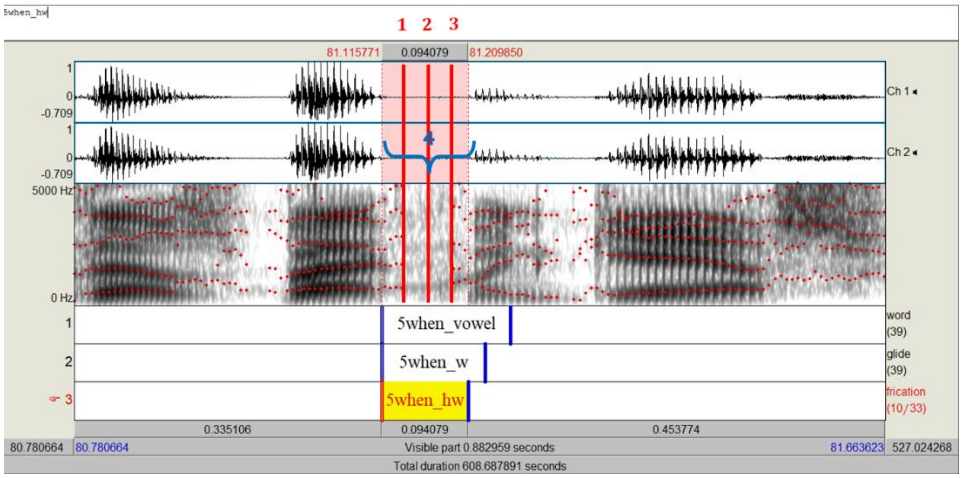

Figure 5.1. Location of COG measurements for token "when"

For all tokens containing voiceless frication, measurements of *duration* (in seconds) were taken from the onset of aspiration to the onset of voicing. Four measurements were also taken for the *center of gravity (COG)* of each [ʍ] token: Praat scripts were used to divide each period of frication into quartiles, and the COG extracted

at 25%, 50%, and 75%. The fourth measurement was obtained by creating a spectral slice over the full duration of the frication and extracting the COG (see Figure 5.1).

5.3 INDEPENDENT VARIABLES

5.3.1 Speaker-level variables

Between speakers, six primary variables were examined for their effects on the five dependent variables named above: (1) *date of birth*, (2) *gender*, (3) *city recorded*, (4) *rurality*, and (5) *education*. These were operationalized as follows:

Date of birth. The informant's year of birth (continuous variable).

Gender. *Levels = female, male.* Informants were asked to record their gender on the demographic survey. 12 respondents answered "female", 12 answered "male", 1 answered "woman", and 1 answered "♂". "Male" and "♂" were grouped together as *male*, and "female" and "woman" as *female.*

City recorded. *Levels = Spartanburg, Columbia.* All interviews were conducted in Spartanburg, SC and Columbia, SC. Of the 19 participants recorded in Spartanburg, 15 were born in that area; the remaining 4 were born in Georgia (2), North Carolina, and the lower part of South Carolina respectively. 17 of these participants described Spartanburg or a local area in Upstate South Carolina as "home", the other named South Carolina in general as his home. The 7 participants recorded in Columbia came from a more diverse set of backgrounds: 2 were born in Columbia, 1 in Charleston, SC, 1 in Greenwood, SC, 1 in Georgia, 1 in North Carolina, and 1 in Virginia. 4 of these

participants described Columbia as “home”, another described both Texas and Columbia as home, and the others described Georgia and Virginia as home respectively.

Rurality. *Levels = rural, small town, suburban, large town.* Informants’ upbringing was classified as having taken place in one of the above area-types on the basis of their identified hometown and responses to the biographical interview

Education. *Levels = high school, college, graduate school.* Based on information from the demographic questionnaire, informants were divided into three education levels. The level *high school* included informants with a high school degree or less; *college* includes those with an undergraduate degree or degree in progress, and *graduate school* includes those with a graduate degree or degree in progress.

5.3.2 Linguistic variables

At the word level, seven variables were examined: (1) *semantic content*, (2) *phrasal position*, (3) *word frequency*, (4) *following vowel quality*, (5) *preceding segment*, (6) *word position*, and (7) *minimal pair status*. These were operationalized as follows:

Semantic content. *Levels = content word, function word.* Content words are taken to be nouns, verbs, adjectives, and descriptive adverbs (e.g. *wheezily*), while function words are taken to be pronouns, conjunctions, and adverbs with grammatical function (e.g. *when, where*).

Phrasal position. *Levels = phrase-initial, phrase-medial, phrase-final.* The onset of a phrase was prosodically determined, such that words beginning a new utterance were considered phrase-initial.

Word frequency. The SUBTLEXus corpus (Brysbaert & New, 2009) was used to determine word frequency, using the most frequently occurring lemma of the word (i.e. the frequency of *wheels* would be determined by the frequency of *wheel*). This was translated to the Zipf value, also provided in the corpus. Zipf's law states that in a large corpus of words, the frequency of any word will be inversely proportional to its rank in the frequency table. A Zipf normalization allows for a more linear interpretation of word frequency effects, by setting similarly ranked words close to each other on a frequency scale, rather than allowing proximity to the upper end of the frequency spectrum to exaggerate differences in usage. For example, *what* (the most common word used in the current study) appears 9842.45 times per million words in the SUBTLEXus corpus, while the second most common word, *why*, appears 2248.76/million times. The Zipf normalization reduces the apparent gap between the usage of these two words, giving them values of 6.99 and 6.35, respectively. On the opposite end of the spectrum, rare words with apparently small differences in usage will show larger differences in their Zipf values: in this study, *whelk* and *whisk* have SUBTLEXus frequencies of 0.04 and 0.57 per million respectively, and Zipf values of 1.77 and 2.77.

Following vowel quality. *Levels = [i], [ɪ], [eɪ], [ɛ], [æ], [a], [ʌ], [ɔ], [ai], [ɝ].* Vowels were coded both phonemically and phonetically, such that tokens which might be expected to exhibit the Southern American English features of pre-nasal raising or monophthongization were examined and categorized according to their pronunciation.

Preceding segment. <wh> words were coded for the identity of the preceding segment with respect to manner and voicing. Levels for manner are *no segment, stop,*

fricative, affricate, nasal, liquid, and *vowel*, while levels for voicing are *no segment, voiceless,* and *voiced*.

Word position. *Levels = initial, non-initial*. When <wh> occurred at the beginning of a word boundary, it was coded as initial; when it occurred word-medially, it was coded as non-initial.

Minimal pair status. *Levels = none, rare pair, common pair*. This variable describes whether the pronunciation of a word with [ʍ] forms a minimal pair with a word containing [w]; i.e. whether its pronunciation with [w] would be an instance of the *wine-whine* merger. If a pair existed, but relied on an obscure word (e.g. *wheel, weal*), then the level *rare pair* was used.

5.4 SEGMENT VS. SEQUENCE

As discussed in Chapter 2.1 (see footnote on pg. 4), the underlying representation of the voiceless labiovelar fricative may be either /ʍ/ or /hw/, with little clear evidence from Modern English to support either alternative. In these initial chapters, [ʍ] has been used in order to avoid the implication that the sound under discussion consists of voiceless glottal frication followed by a glide, or that [ʍ]-loss is necessarily associated with "h-dropping." However, while tokens do exist in which the spectral qualities of [ʍ] are similar to a voiceless [u], it was found that this varied between participants, and that many tokens showed spectral qualities more similar to [h]. Indeed, the nature of the current analysis, which focuses on isolating the voiceless frication within a segment, lends itself toward a sequential representation of [hw], in that the majority of tokens

consisted of aspiration of some sort followed by a glide. Furthermore, results from Chapter 3 indicate that the majority of listeners perceive [ʍ] as a [hw] sequence, with only one commenter describing a difference between a "hard w" [w] and "soft w" [ʍ]. In light of this, and also in light of the nature of certain allophonic variants appearing in the data, [hw] will be used in the following analysis. The issue of the segment/sequence debate will be returned to in Chapter 9.

CHAPTER 6
RESULTS: READING TASK

6.1 STATISTICAL TESTS

For the reading task, each of the 26 participants produced 43 tokens of <wh>-words. Including repetitions and excluding all audio-corrupted tokens, this resulted in a total of 1064 tokens, with a range of 38 to 45 tokens per participant.[2] One word token, *wharf*, was observed to have an extremely low rate of [hw]-realizations: only 2 out of 26 speakers (P13 and P20) pronounced it with aspiration. Since this word was observed to behave anomalously as early as the 1940s, it was removed from the present analyses so as not to skew the data, and statistical tests conducted on the remaining 1040 tokens.

Data for each dependent variable was examined in terms of five individual variables: *date of birth, gender, semantic content, phrasal position,* and *word frequency*, which were determined to be of primary interest on the basis of previous research. All twelve independent variables (see Chapter 5.3) were subsequently submitted to a mixed regression model in order to determine the relative and combined effects of the speaker- and word-level variables. Based on the hypothesis that <wh>-pronunciation varies not only across social categories and linguistic contexts, but is subject to variation across

[2] One exception to this existed: a recorder malfunction obscured a substantial part of P13's reading audio, so that only the 10 final tokens remained. These 10 tokens were included in the overall total.

individuals, *participant* was treated as a random effect. All other variables were treated as fixed effects and submitted to a stepwise regression to determine their place, if applicable, in the final model. All regressions were computed in the statistical tool R (R Core Team, 2019) using the function *lmer()* from the package *lmerTest* (Kuznetsova et al., 2017). The resulting visuals were generated using the *effects* package (Fox, 2003).

6.2 PRESENCE OF [hw]

6.2.1 Individual variables

Analyses of the continuous variables *date of birth (DOB)* and *word frequency* were performed by fitting a robust linear model (RLM) for *presence*, using the function *rlm()* from the package *MASS* (Venables & Ripley, 2002). Categorical variables were tested using two-way Kruskal-Wallis analyses of variance, using the function *kruskal.test()*.

An RLM revealed a significant effect of date of birth on [hw]-production [$t(1038)=-18.619$, $p<.0001$], such that the likelihood of a word being realized with [hw] decreased by 1.3% for every one year increase in the participant's date of birth. However, as shown in Figure 6.1, this relationship was not necessarily linear. Rather, the majority of participants showed high rates of [hw]-usage until reaching a birthdate of roughly 1970, after which only one participant (P10) pronounced more than 20% of tokens as [hw]. The years between 1960 and 1970 appear to represent a transition period, in which one participant showed high usage of [hw], one used [hw] about half of the time, and three showed little to no usage of [hw].

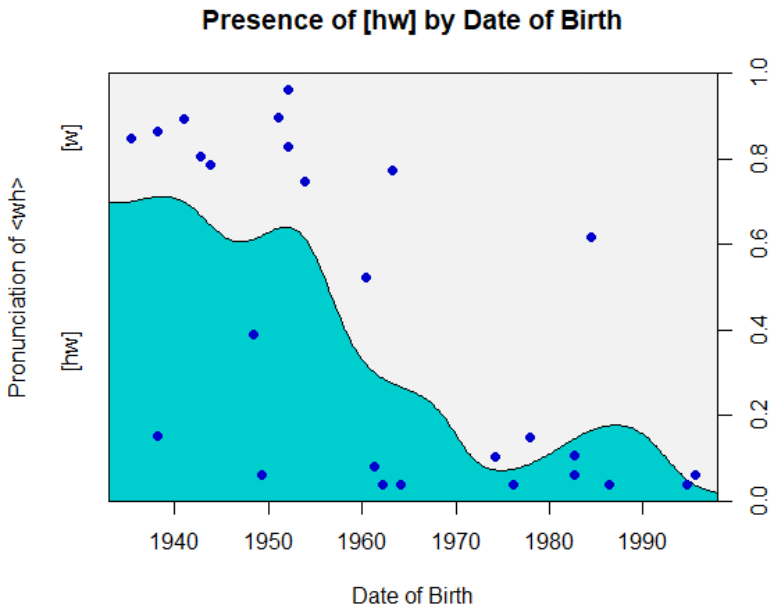

Figure 6.1 Presence of [hw] by date of birth. Dots represent the percentage of [hw] tokens used by individual participants

A chi-squared test revealed a significant relationship between *presence* and *gender* [$\chi^2(1)=22.569$, $p<.0001$], such that men were more likely (47.3%) to use [hw] than women (32.6%). For *semantic content* and *phrasal position*, chi-squared tests revealed no significant relationship with [hw]-presence [$\chi^2(1)=0.609$, $p=.435$; $\chi^2(2)=0.703$, $p=.704$]. The distribution of these three categorical variables is illustrated below, in Figure 6.2.

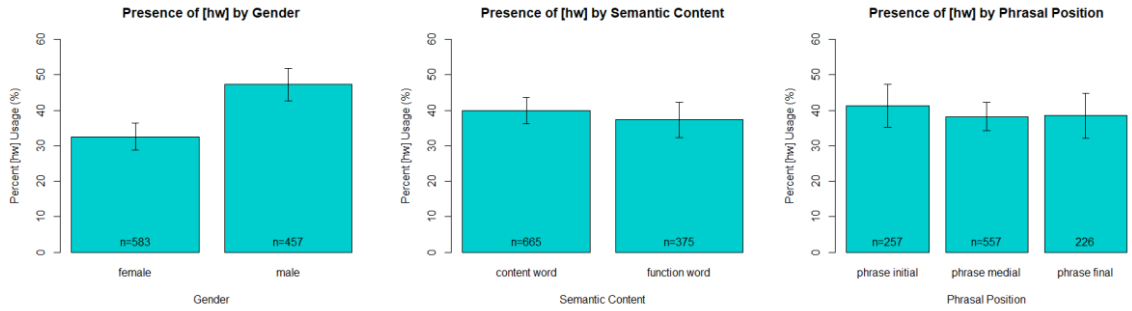

Figure 6.2 Presence of [hw] by (a) gender, (b) semantic content, and (c) phrasal position. Error bars represent 95% confidence intervals

With respect to *word frequency*, an RLM revealed no significant relationship between *presence* and *word frequency* [t(1038)=-0.649, *p*=.258].

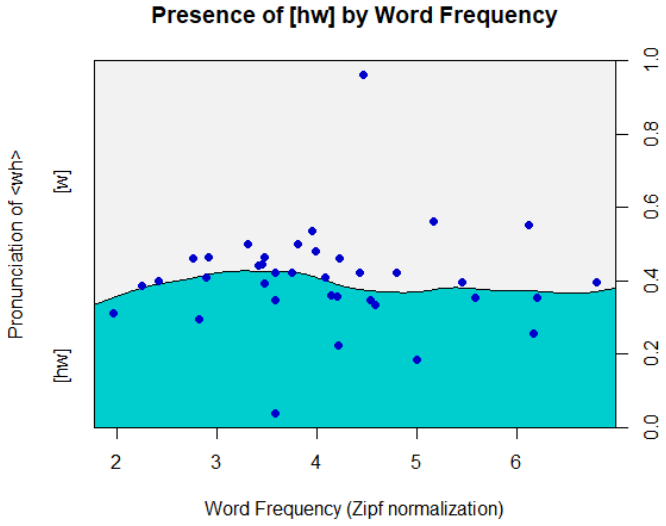

Figure 6.3 Presence of [hw] by word frequency. Dots represent the percentage of [hw] tokens produced for unique words

6.2.2 Multiple logistic regression

Results of a stepwise logistic regression predicting the presence of [hw]-production, shown below in Table 6.1, selected a best-fitting model that included the following predictors in the following order: *word position, date of birth, city recorded*, and *semantic content*.

Table 6.1 Mixed logistic regression model for presence

	Odds ratios (lower, upper CI)	p-values
(Intercept)	**2.518e13 (3.230e8, 1.972e18)**	**.0000******
Word initial – Initial	**1.139 (1.078, 1.203)**	**.0000******
Date of birth	**0.984 (0.979, 0.990)**	**.0000******
City recorded – Columbia, SC	**1.426 (1.107, 1.839)**	**.0134***
Semantic content – Content word	**1.062 (1.020, 1.106)**	**.0033****

Within the above model, in which *participant* was treated as a random effect, there was a main effect of *word position*, such that <wh> was 1.14 times more likely to be pronounced as [hw] when it occurred initially in a word, versus in medial position. There was also a main effect of *date of birth*, such that for every year that date of birth increases, the likelihood of [hw]-production decreases by 1.6%. There was also an effect of *city recorded*, such that those participants recorded in Columbia, SC were 1.43 times more likely to produce [hw] than those in Spartanburg, SC. Finally, there was an effect of *semantic content*, such that content words were pronounced with [hw] 1.06 times more often than function words. These effects are summarized in Figure 6.4.

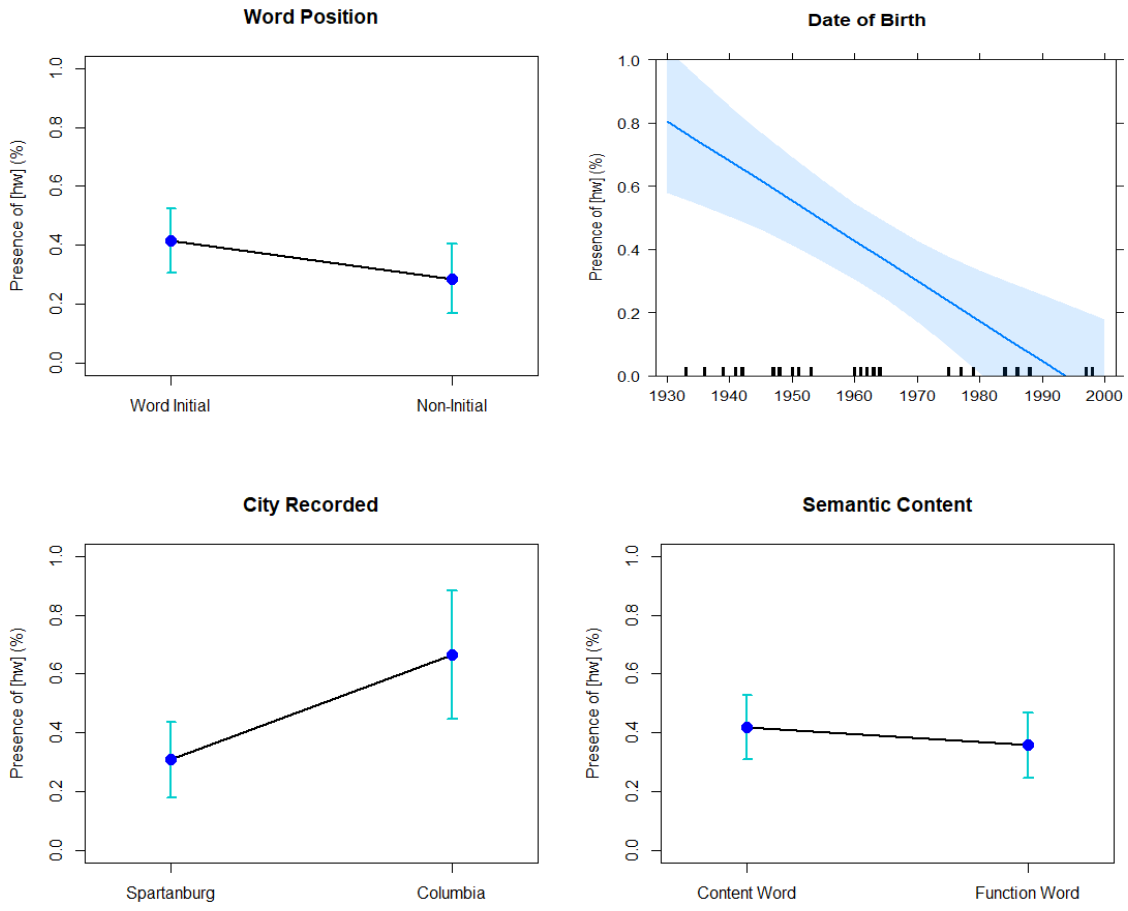

Figure 6.4 Effects of a logistic regression for presence of [hw]

6.3 DURATION OF [hw]

In order to determine whether variation within [hw] tokens could be predicted by demographic or linguistic variables, analyses of *duration* and *COG* were performed on those tokens (n=378) which exhibited some degree of voiceless frication. In addition to removing all tokens coded as [w], another subset of tokens which contained frication but

also a voicing band were also excluded. The issue of these "voiced [hw]s" will be returned to in later sections.

6.3.1 Individual variables

Due to the non-normality of the data, the analysis for *gender* was performed by conducting a Kruskal-Wallis one-way analysis of variance, using the function *kruskal.test()*. The within-subjects categorical variables, *semantic content* and *phrasal position*, were tested via a one-way repeated measures analysis of variance (ANOVA), with *participant* treated as a random effect, using the function *aov()*. Comparisons within each variable were made using the Mann-Whitney U test, the non-parametric equivalent of a t-test, using *wilcox.test()*. For the continuous variables *date of birth (DOB)* and *word frequency*, a robust linear model (RLM) was fit for the data, using the function *rlm()* from the package *MASS* (Venables & Ripley, 2002). Results for all five individual variables are visualized in Figure 6.5.

As shown in Figure 6.5a, an RLM revealed a significant relationship between *duration* and *date of birth* [t(376)=-2.1952, p=.014], such that a one-year increase in date of birth was accompanied by a .0002s (0.2ms) decrease in fricative duration. For *gender*, a Kruskal-Wallis one-way analysis of variance revealed a significant difference between male and female speakers with respect to duration [$\chi^2(1)=7.06$, p=0.008], such that men produced longer frication for [hw] than women.

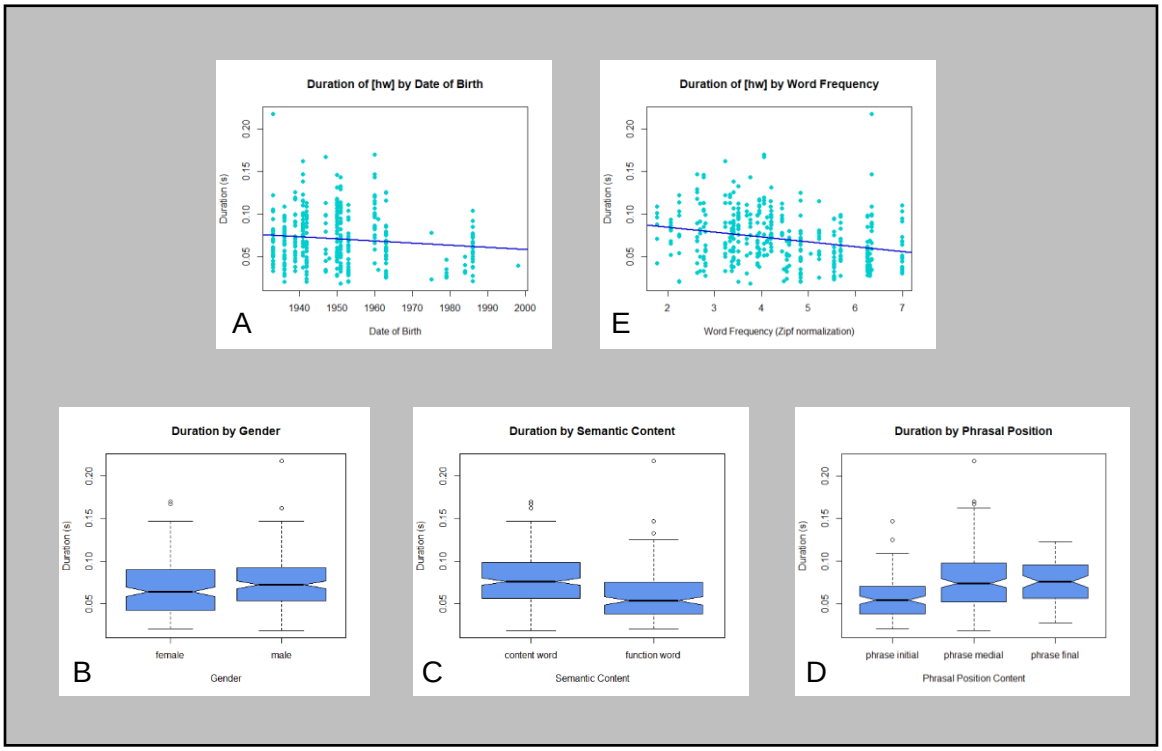

Figure 6.5 Duration of [hw] by (a) date of birth, (b) gender, (c) semantic content, (d) phrasal position, and (e) word frequency. Boxplot notches represent a 95% confidence interval for the median.

Among the word-level variables, a one-way repeated-measures ANOVA revealed a significant relationship between *duration* and *semantic content* [$F(1,15)=23.10$, $p=.0002$], such that content words were produced with longer frication than function words. A second repeated-measures ANOVA revealed a significant relationship between *duration* and *phrasal position* [$F(2,29)=9.23$, $p=.0008$], such that words occurring phrase-initially were produced with shorter frication than those in medial position [$W=6596.5$, $p<.0001$] and final position [$W=2442.5$, $p<.0001$]. There were no significant durational

differences between words occurring medially and finally [W=7578.5, *p*=.712]. Lastly, an RLM revealed a significant relationship between *duration* and *word frequency* [t(376)=-5.77, *p*<.0001, such that an increase of one unit on the Zipf scale (i.e. an increase in word frequency) was accompanied by a 0.006s (6ms) decrease in duration. (See Chapter 5.3 for a discussion of how to interpret the Zipf scale.)

6.3.2 Linear regressions

Despite previously statistically significant effects of date of birth and gender, the inclusion of *participant* as a random effect eliminated all subject-level variables from the linear regression models used to predict duration. The stepwise linear regression procedure selected a best-fitting model with only *previous voicing* as a predictor, such that tokens in which <wh> followed a voiced segment were produced with 0.020s longer frication relative to those which had no immediately preceding segment, and tokens following a voiceless segment showed a tendency to be produced with longer duration than those with no preceding segment.

Table 6.2. Mixed linear regression model for duration (previous voicing)

	Estimate (lower, upper CI)	p-values
(Intercept)	**0.054 (0.045, 0.062)**	**.0000******
Previous voicing – Voiceless	**0.007 (-0.001, 0.015)**	**.080(.)**
Previous voicing – Voiced	**0.020 (0.016, 0.028)**	**.0000******

Notably, three other variables emerged as alternative models: *word frequency, semantic content*, and *phrasal position*. These four variables, which show strong multicollinearity, were taken to be indexing the same phenomenon: the position of a word in an utterance. Words which occurred phrase-initially were coded as "no preceding segment" for previous voicing, and tended to be function words of high frequency. Out of the 103 tokens occurring phrase-initially, 95 had no preceding segment, and 86 were function words. Therefore, the regression model including phrasal position is also provided below, in Table 6.3.

Table 6.3 Mixed linear regression model for duration (phrasal position)

	Estimate (lower, upper CI)	p-values
(Intercept)	**0.055 (0.046, 0.063)**	**.0000******
Phrasal position – Medial	**0.016 (0.010, 0.023)**	**.0000******
Phrasal position – Final	**0.015 (0.007, 0.023)**	**.0002*****

According to the above model, tokens occurring phrase-medially are expected to be 0.016s longer than those occurring phrase-initially, and tokens occurring phrase-finally are expected to be 0.015s longer.

6.4 COG OF [hw]

Among the 13 speakers with more than 5 analyzable tokens of [hw] (i.e. excluding P8, P14, P15, P17, P18, P24, P25, and P26), COG values were largely similar

across timepoints, with a tendency to drop slightly as the fricative progressed, as shown in Figure 6.6. As such, the average COG over the course of the fricative was taken to be a good representative of COG at all timepoints, with the additional benefit of being less subject to instantaneous fluctuations in spectral composition, and is the metric used in the following analyses.

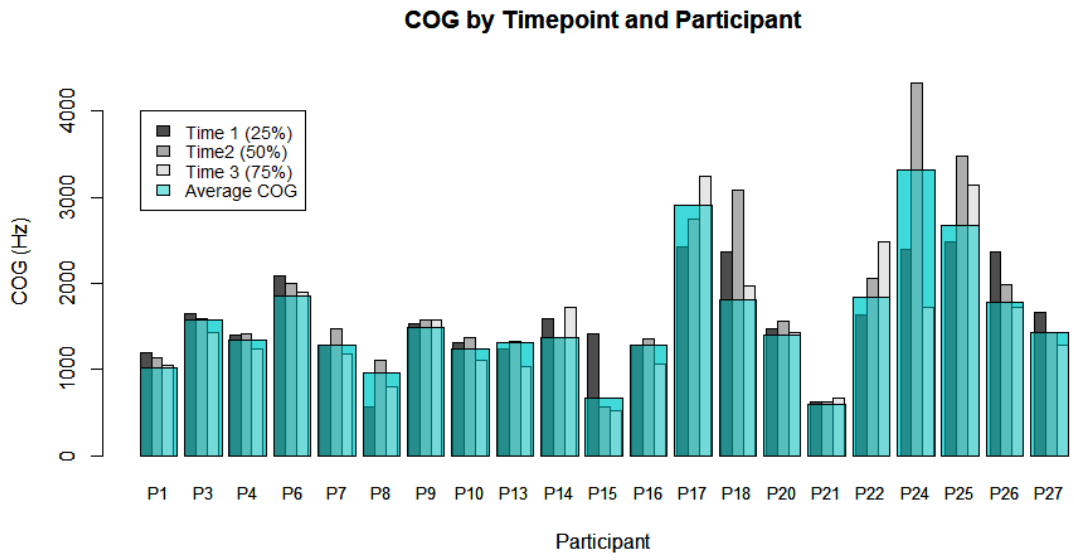

Figure 6.6 Mean COG values by participant at individual timepoints

6.4.1 Individual variables

All statistical tests of COG were conducted in the same manner as those for duration, using only tokens exhibiting voiceless frication and utilizing Kruskal-Wallis tests and robust linear models. Visualizations are shown in Figure 6.7.

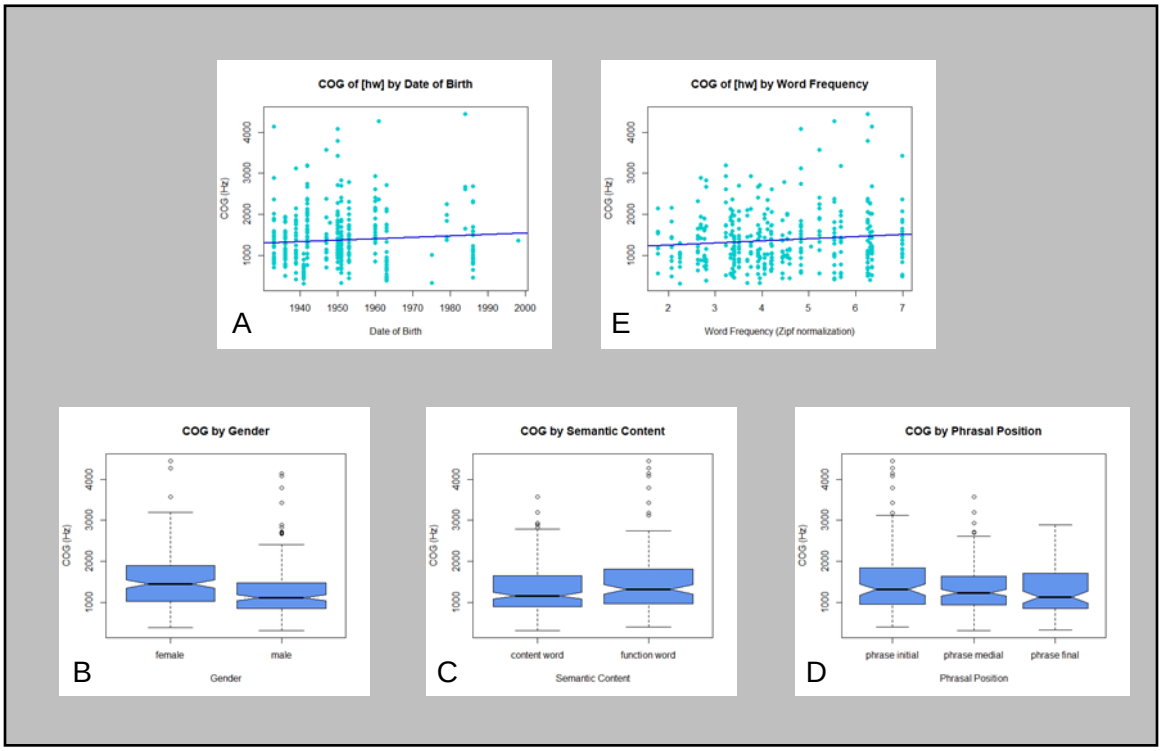

Figure 6.7 Mean COG of [hw] by (a) date of birth, (b) gender, (c) semantic content, (d) phrasal position, and (e) word frequency

Among the speaker-level variables, an RLM revealed no significant relationship between *COG* and *date of birth* [t(376)=-1.12, p=.133]. A Kruskal-Wallis one-way analysis of variance revealed a significant difference between male and female speakers with respect to COG [$\chi^2(1)=18.53$, $p<.0001$], such that women produced [hw] with higher COG values than men.

For word-level variables, a one-way repeated-measures ANOVA revealed trending COG differences between content and function words [F(1,15)=4.48, p=.051], such that function words had higher COG values than content words. A trending

relationship was also present between *COG* and *phrasal position* [F(2,29)=2.57, *p*=.094], such that phrase-initial words tend to be produced with higher COG than those in phrase-final position. Finally, an RLM revealed a trending but non-significant relationship between *COG* and *word frequency* [t(376)=-1.40, *p*=.081], such that COG increased as word frequency increased.

6.4.2 Linear regression

Since COG is a spectral cue that is known to be affected by coarticulation, the variables *vowel, previous manner,* and *previous voicing* were removed from the stepwise regression, so as to allow for the identification of patterns specific to [hw]. The resulting multiple linear regression, shown in Table 6.4, selected a best-fitting model that included the following predictors in the following order: *rurality*, *phrasal position*, and *education*. Within this model, there was a main effect of *phrasal position*, such that tokens occurring phrase-medially were produced 353 Hz lower than words occurring phrase-initially. Trending main effects also existed for *rurality* and *education*: speakers from large towns produced lower COG values than those from rural areas, and speakers with undergraduate or graduate-level education used produced lower COG values than those with a high school education.

Table 6.4 Mixed linear regression model for COG

	Estimate (lower, upper CI)	p-values
(Intercept)	**2044.591 (1577.125, 2492.540)**	**.0000******
Rurality – Small town	500.604 (-224.345, 1171.277)	.233
Rurality – Suburban	439.353 (-474.996, 1343.567)	.427
Rurality – Large town	**-931.661 (-1755.408, -90.613)**	**.085(.)**
Phrasal position – Medial	**-352.695 (-622.204, -94.807)**	**.010****
Phrasal position – Final	-239.803 (-585.677, 71.807)	.157
Education – College	**-813.977 (-1559.263, -42.022)**	**.083(.)**
Education – Graduate school	**-900.152 (-1698.165, -52.047)**	**.080(.)**
PP – Medial * E – College	**948.506 (403.496, 1503.075)**	**.001*****
PP – Final * E – College	348.804 (-277.909, 1019.177)	.301
PP – Medial * E – Graduate	**518.774 (-15.665, 1086.191)**	**.069(.)**
PP – Final * E – Graduate	237.236 (-394.354, 945.094)	.492
R – Small town * PP – Medial	-369.487 (-889.634, 117.932)	.157
R – Suburban * PP – Medial	**-637.715 (-1220.678, -71.773)**	**.032***
R – Large town * PP – Medial	242.964 (-176.691, 669.022)	.267
R – Small town * PP – Final	-193.171 (-830.855, 385.580)	.539
R – Suburban * PP – Final	-464.958 (-1169.840, 206.926)	.192
R – Large town * PP – Final	-45.241 (-546.323, 489.589)	.865

In addition to the main effects, a significant interaction was present between *education* and *phrasal position*, as shown in Figure 6.8, such that college-educated speakers' COG was highest in medial position, contrasting with the overall trajectory of initial > medial = final. There was also a non-significant trend for graduate-educated speakers to produce tokens in medial position with a higher COG than otherwise predicted.

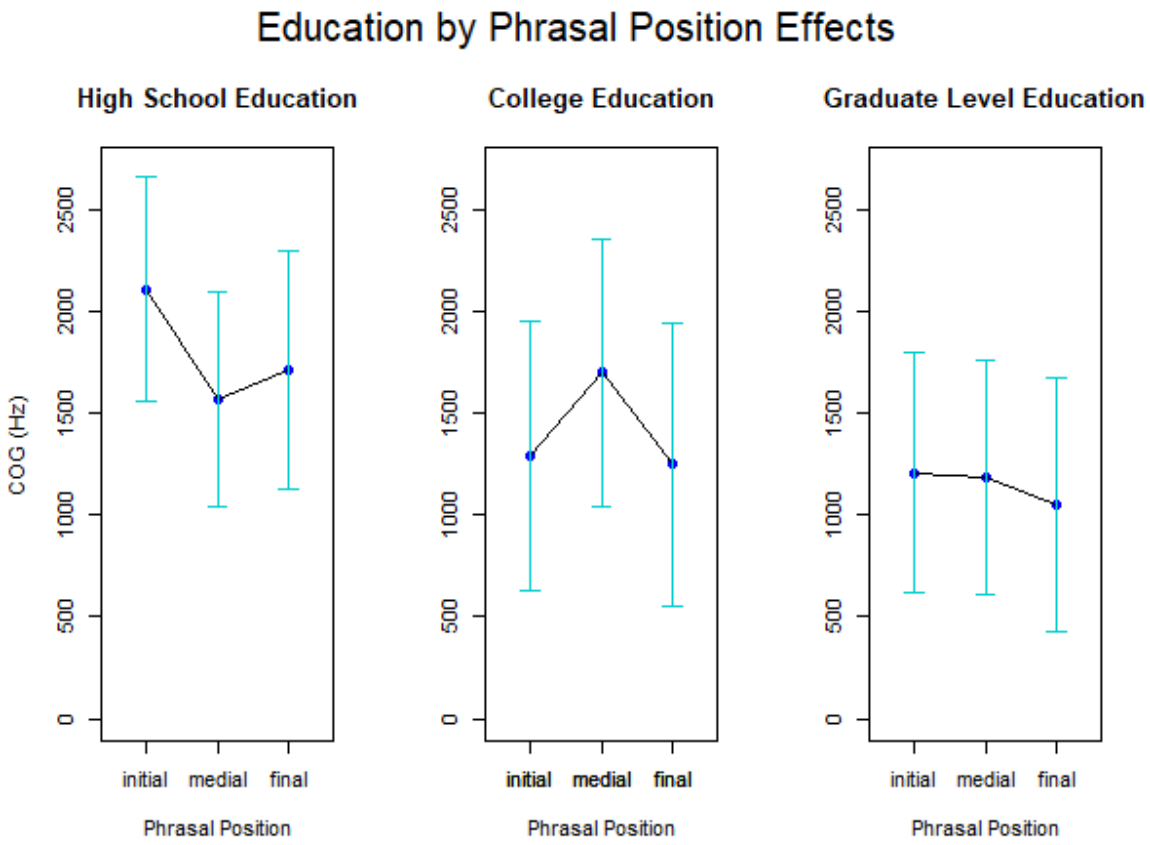

Figure 6.8 Interaction of education and phrasal position with respect to COG

A second significant interaction was present for *rurality* and *phrasal position*, as shown in Figure 6.9, such that suburban speakers had lower COG frequencies in word-medial position relative to words occurring elsewhere, compared to speakers from other backgrounds.

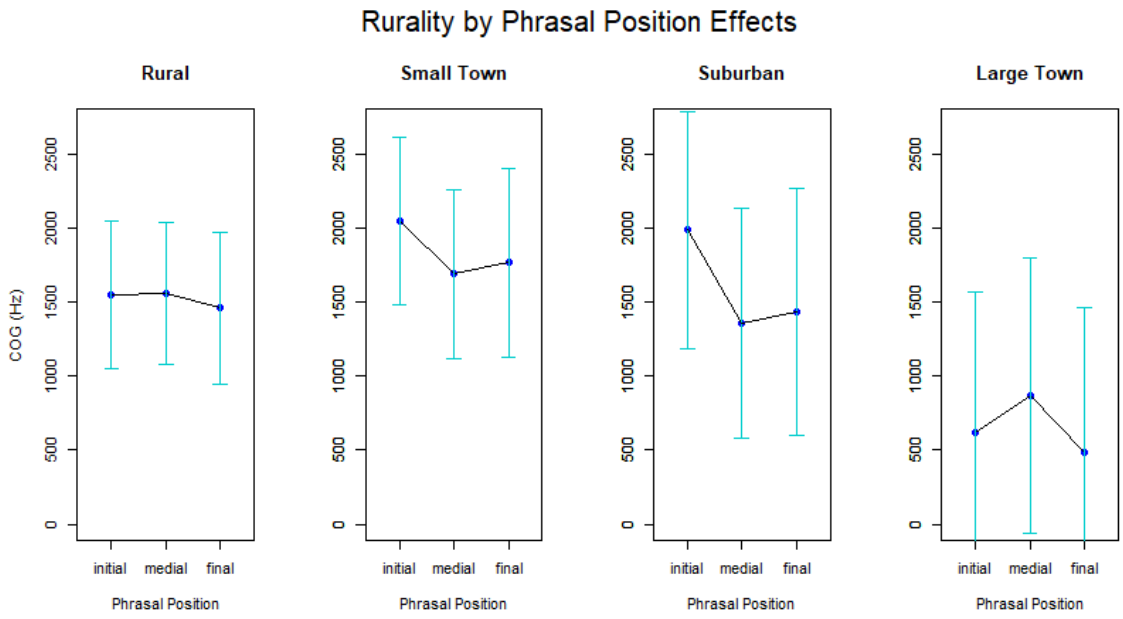

Figure 6.9 Interaction of rurality and phrasal position with respect to COG

6.5 VOICED [hw]

In the course of the previous analyses, a subset of data emerged in which participants produced tokens with audible or spectrally visible frication preceding the glide [w], but this frication also exhibited a voicing band throughout. Such tokens occurred 28 times, comprising 6.8% of the 406 tokens exhibiting some form of aspiration. In order to determine the allophonic distribution of these "voiced [hw]" variants, rates of voiced tokens were compared across each of the word-level variables which described the surrounding environment: *phrasal position, following vowel, previous manner, previous voicing*, and *word position*. Since not all [hw]-producing

participants exhibited [hw] voicing, this was followed by a logistic regression including both social and linguistic variables.

6.5.1 Individual variables

All analyses of contextual variables for *[hw] voicing* were performed by conducting a Pearson's chi-squared test of goodness of fit, using the function *chisq.test*. The results of these are shown below, in Table 6.5.

Table 6.5 Chi-squared tests of [hw] voicing

Variable	
Phrasal position	$\chi^2(2)=3.710$, $p=.157$
Following vowel	$\chi^2(8)=18.228$, $p=.020$*
Previous manner	$\chi^2(5)=57.703$, $p<.0001$****
Previous voicing	$\chi^2(2)=24.68$, $p=.0000$****
Word position	$\chi^2(1)=26.479$, $p<.0001$****

As shown in Figure 6.10, strong effects exist for *previous manner, previous voicing*, and *word position*, as well as a moderate effect for *vowel*. In terms of *previous manner*, tokens of [hw] following a nasal were most likely to be realized with voicing, followed by those preceded by liquids and then vowels, while [hw] tokens following an obstruent or beginning an utterance were never voiced. Similarly, [hw] tokens following voiced segments were occasionally realized with voicing, while those following no segment or a voiceless segment never were. For *vowel*, [hw] was most frequently voiced

when preceding [ɝ], followed by [i] and [a]. Finally, [hw] was realized with voicing a much greater percentage of the time when it appeared word-medially than word-initially.

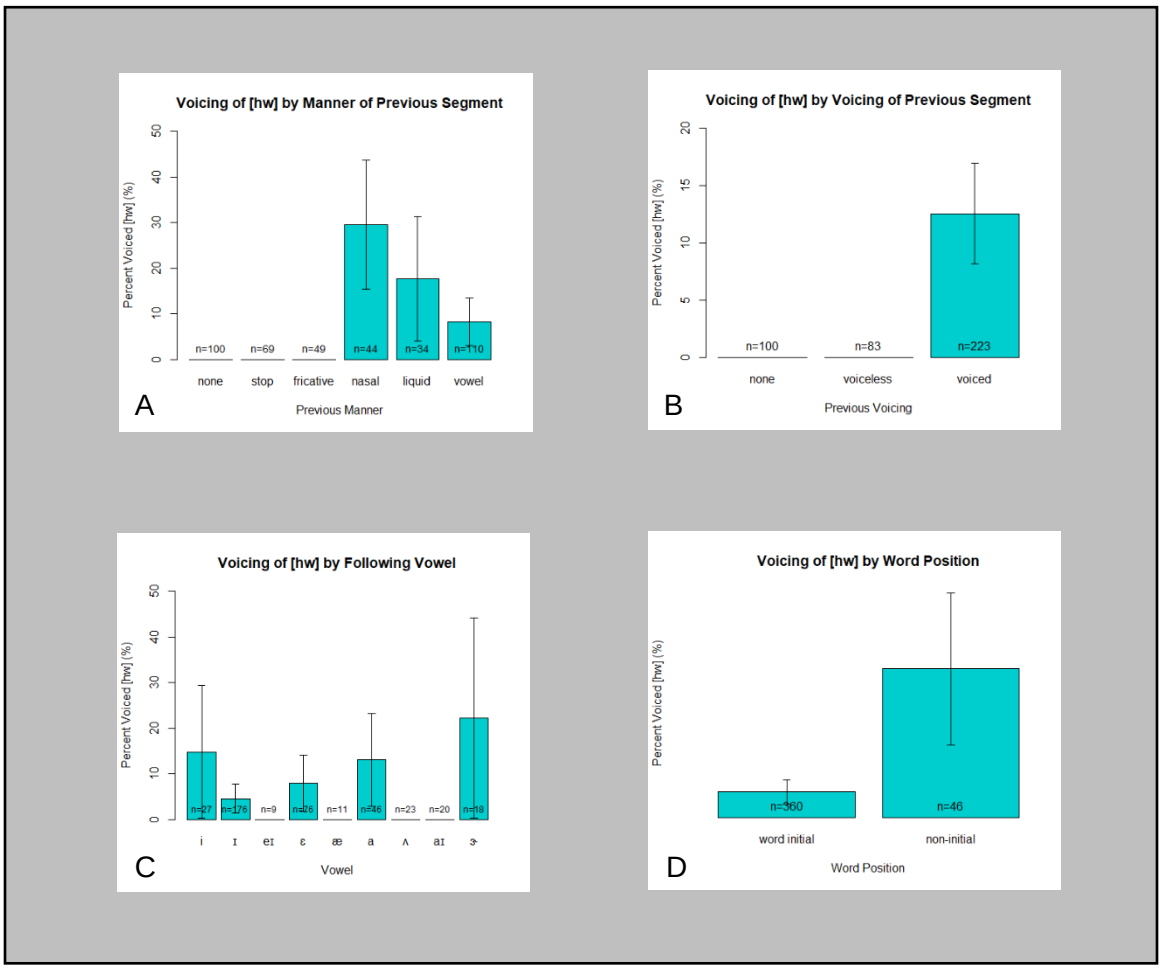

Figure 6.10 [hw] voicing by (a) previous manner, (b) previous voicing, (c) following vowel, and (d) word position

6.5.2 Logistic regression

Results of a stepwise logistic regression predicting the presence of [hw] voicing, shown below in Table 6.6, selected a best-fitting model that included the following predictors in the following order: *previous manner*, *word position*, *city recorded*, and *education level*. For this variable, it was found that the inclusion of *participant* as a random effect did not serve to increase the efficiency and predictive power of the model, so only fixed effects were included.

Table 6.6 Logistic regression model for voicing

	Odds ratios (lower, upper CI)	p-values
(Intercept)	**1.251 (1.107, 1.413)**	**.0003*****
Previous manner – Stop	**0.824 (0.666, 1.021)**	**.076(.)**
Previous manner – Fricative	1.000 (0.897, 1.115)	1.000
Previous manner – Nasal	**1.960 (1.579, 2.432)**	**.0000******
Previous manner – Liquid	1.138 (0.920, 1.408)	.230
Previous manner – Vowel	1.070 (0.982, 1.167)	.123
Word position – Initial	**0.799 (0.719, 0.889)**	**.0000******
City recorded – Columbia, SC	**0.823 (0.668, 1.014)**	**.067(.)**
Education – College	1.000 (0.874, 1.145)	1.000
Education – Graduate school	1.000 (0.886, 1.128)	1.000
PM – Stop * WP – Initial	**1.209 (0.984, 1.484)**	**0.071(.)**
PM – Fricative * WP – Initial	---	---
PM – Nasal * WP – Initial	**0.706 (0.578, 0.863)**	**.0007*****
PM – Liquid * WP – Initial	**1.220 (0.995, 1.496)**	**.056(.)**
PM – Vowel * WP – Initial	---	---
PM – Stop * CR – Columbia	**1.012 (0.810, 1.265)**	**.0000******
PM – Fricative * CR – Columbia	1.000 (0.785, 1.273)	1.000
PM – Nasal * CR – Columbia	**0.569 (0.439, 0.738)**	**.0000******

PM – Liquid * CR – Columbia	1.137 (0.860, 1.502)	.366
PM – Vowel * CR – Columbia	1.015 (0.837, 1.232)	.876
WP – Initial * CR – Columbia	**1.216 (1.033, 1.431)**	**.019***
PM – Stop * E – College	1.001 (0.818, 1.223)	.996
PM – Fricative * E – College	1.000 (0.803, 1.246)	1.000
PM – Nasal * E – College	**0.722 (0.568, 0.918)**	**.008****
PM – Liquid * E – College	**0.707 (0.529, 0.944)**	**.019***
PM – Vowel * E – College	0.950 (0.798, 1.131)	.561
PM – Stop * E – Graduate sch.	0.999 (0.806, 1.237)	.991
PM – Fricative * E – Graduate sch.	1.000 (0.796, 1.256)	1.000
PM – Nasal * E – Graduate sch.	**1.279 (0.997, 1.641)**	**.053(.)**
PM – Liquid * E – Graduate sch.	**0.716 (0.549, 0.932)**	**.013***
PM – Vowel * E – Graduate sch.	0.955 (0.804, 1.134)	.598

Within the above model, there was a main effect of *previous manner*, such that, relative to when it followed no segment, [hw] was 1.96 times more likely to be pronounced with voicing when it followed a nasal. There was also a trending effect by which a voiced [hw] was less likely to follow a stop than no segment (however, this appeared to be an artifact of other interactions, as there were no instances of voicing for either level in the data).

Another main effect was present for *word position*, such that tokens of [hw] occurring word-medially were 1.25 times more likely to be pronounced with voicing than those occurring word-initially. Finally, there was a main effect of *city recorded*, such that those participants recorded in Spartanburg, SC were 1.22 times more likely to produce voiced [hw] than those from Columbia, SC.

In addition to the above main effects, significant interactions were present between *previous manner* and *word position*, *previous manner* and *city recorded*, *word*

position and *city recorded*, and *previous manner* and *education*. These effects are visualized below, in Figures 6.11 to 6.13.

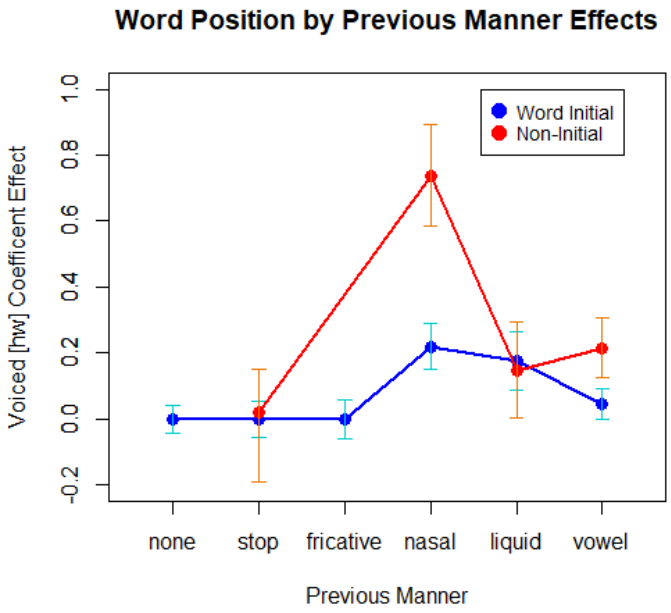

Figure 6.11 Interaction of manner and word position with respect to [hw] voicing

As shown by Figure 6.11, when all other effects are held constant, [hw] behaves similarly across word position when preceded by a stop or liquid (there were no isolated or fricative tokens in medial position). Nasal-preceded [hw] tokens, however, were much more frequently voiced in word-medial position than word-initial, and vowel-preceded tokens showed tendencies in the same direction.

Speakers recorded in both Spartanburg and Columbia, SC showed similar patterns of [hw] voicing after stops, liquids, and vowels. With respect to tokens following nasals, however, speakers from Spartanburg showed proportionally larger increases in voicing rates than speakers from Columbia.

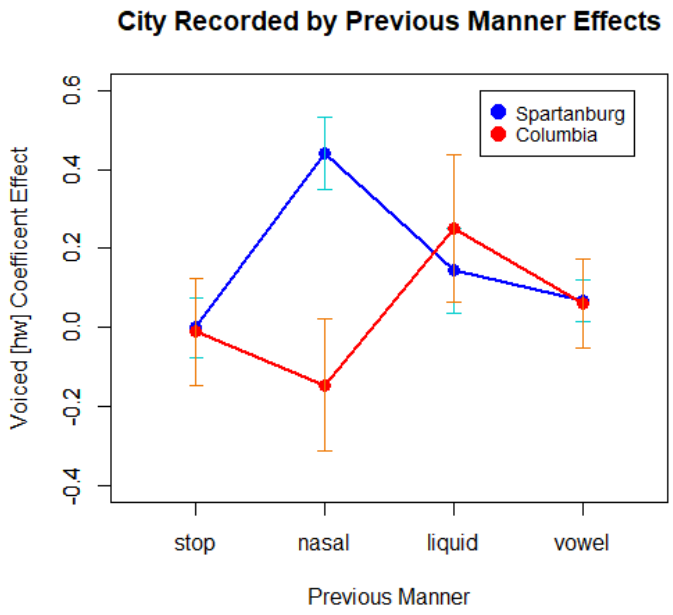

Figure 6.12 Interaction of manner and city recorded with respect to [hw] voicing

Across education levels, all speakers showed similar patterns for tokens following stops and vowels. High-school educated speakers showed greater voicing of tokens following liquids than college- or graduate-educated speakers, while all three groups

differed for tokens following nasals, according to the pattern *graduate* > *college* > *high school*.

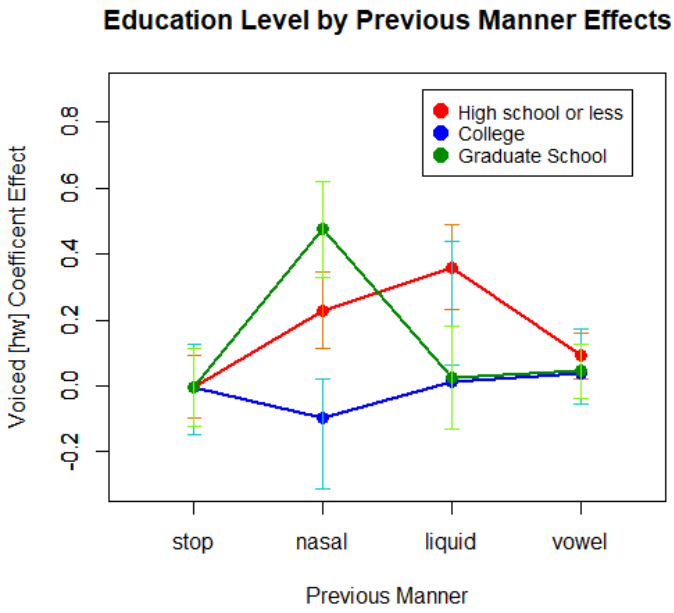

Figure 6.13 Interaction of manner and education with respect to [hw] voicing

Finally, an interaction existed for *word position* and *city recorded*. While no effects estimates were calculable for this interaction due to gaps in the data, analyses of raw percentages revealed that speakers from Spartanburg produced higher rates of voicing in word-medial position than speakers from Columbia.

CHAPTER 7

RESULTS: INTERVIEW

7.1 CORPUS PROPERTIES

From a total of 478 minutes of recorded discourse, participants produced a corpus of 855 measureable <wh> tokens (mean=34.2, SD=19.2), as shown in Table 7.1.

Table 7.1 Interview tokens per participant

Participant	Token Count	Participant	Token Count
P1	*no data*	P15	59
P2	42	P16	37
P3	19	P17	11
P4	37	P18	55
P5	46	P19	42
P6	19	P20	39
P7	97	P21	22
P8	16	P22	33
P9	38	P23	50
P10	43	P24	23
P11	21	P25	7
P13	31	P26	35
P14	8	P27	25

Out of the 855 above tokens, 31 unique words were identified. As shown in Table 7.2, the overwhelming bulk of the corpus consisted of those words which could be used to introduce a relative or dependent clause: *what, when, where, which,* and *while,* as

well as *whatever*, which frequently served as a filler word (i.e. in the end-of-sentence tag "or whatever").

Table 7.2. Interview corpus lexicon

Lemma	Wordforms	Frequency
anywhere		7
everywhere		9
meanwhile		1
nowhere		1
overwhelm	*overwhelmed*	1
somewhere		9
Whaley		2
what		174
whatever		50
whatsoever		1
wheat		2
wheel	*wheel*	8
	wheeled	1
	wheelers	1
	wheels	1
	wheely	1
when		320
whenever		3
where		95
whereas		1
wherever		1
whether		2
which		84
whichever		1
while	*while (noun)*	23
	while (conj)	16
whip	*whipped*	1
white	*white*	14
	Whitestone	1
	whitetail	1
	Whitey	1
why		22

As a result of the uneven distribution across both participants and words, no inferential statistics could be reliably conducted on the interview data. Therefore, all results will be presented descriptively, with no claims made as to their statistical significance.

7.2 PRESENCE OF [hw]

As shown in Figure 7.1, [hw]-usage started at approximately 50% among the oldest speakers and decreased with relative regularity until reaching those speakers who were born around 1960, where rates leveled off and neared 0%. There was a slight increase in the 1980s, largely due to one speaker (P10) with unusually high [hw]-usage.

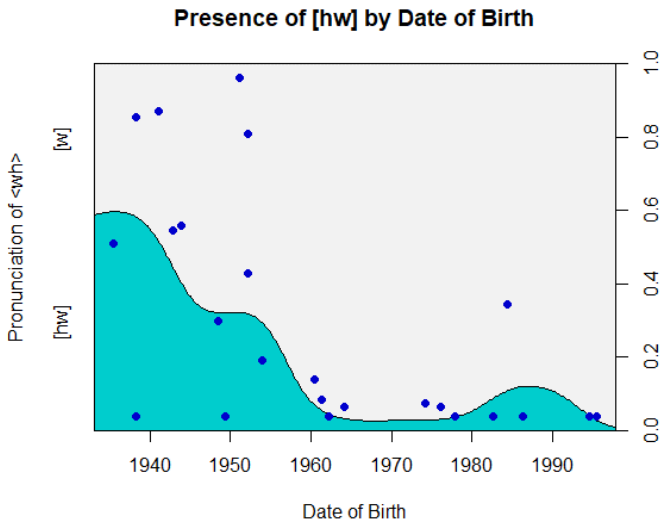

Figure 7.1 Presence of [hw] by date of birth. Dots represent the percentage of [hw] tokens used by individual participants

Results for the remaining social variables, displayed in Figure 7.2, indicated strong differences in [hw]-production by *gender*, *rurality*, and *education*. In terms of *gender*, men produced [hw] twice as frequently as women (33.9% vs. 16.5%). By location type, small-town speakers had the lowest rates of [hw] (10.8%), compared to rural speakers (38.6%), suburban speakers (29.8%), and large-town speakers (28.1%).

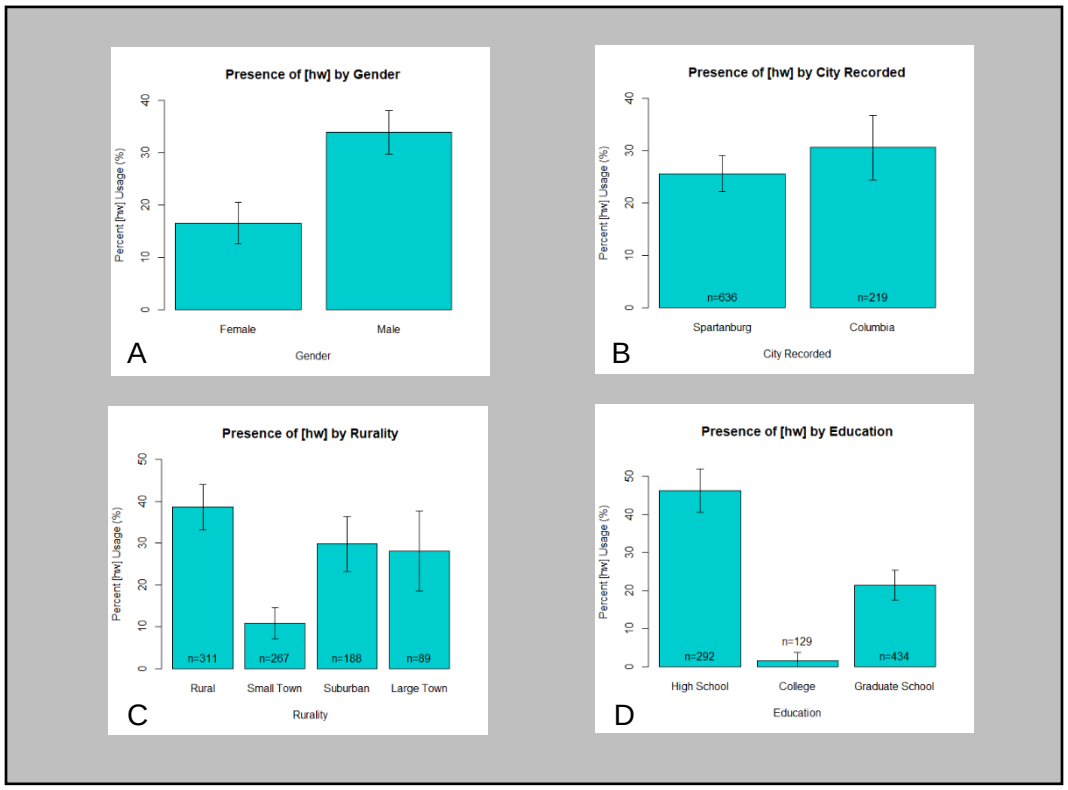

Figure 7.2 Presence of [hw] by social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

Finally, with respect to *education*, high-school-educated speakers produced [hw] the most (46.2%), graduate-educated speakers second-most (21.4%), and college-educated speakers least (1.6%). The extremely low levels among college-educated speakers reflected the fact that this tended to include the youngest speakers among the three groups, while the high-school-educated speakers tended to be the oldest.

Results for the word-level variables (see Figure 7.3) indicated potential differences due to *semantic content, phrasal position, vowel,* and *previous manner*. Content words tended to be pronounced with [hw] more frequently than function words (34.0% vs. 25.5%), while preceding nasals conditioned aspiration less frequently (14.3%) than no segment (31.0%), stops (30.1%), fricatives (33.1%), and possibly vowels (25.5%). Although the confidence intervals for phrasal position overlapped, largely due to the low number of phrase-final tokens, there was a considerable difference between rates of [hw]-usage for phrase-final words (37.8%) than words appearing in initial or medial position (28.8%, 23.8%).

For *vowel*, the levels of [æ] and [eɪ] were removed from the analysis, since they had only 1 and 4 tokens respectively, generating confidence intervals which ranged from 0 to 100%. Here, [ɪ] exhibited lower rates of [hw]-usage (21.7%) than [aɪ] (57.1%), as well as potentially lower rates than [ʌ] (30.7%) and [i], which had a [hw]-rate of 66.7% but only 9 tokens.

Across frequencies, there were no clear patterns where Zipf values exceeded 3.5; however, there appeared to be unanimous [hw]-production in rare words (see Figure

7.3c). However, this was revealed to be misleading: only 5 tokens where Z<3.5 existed, 3 of which were pronounced with [w], and 2 with [hw].

Figure 7.3 Presence of [hw] by linguistic variables: (a) semantic content, (b) phrasal position, (c) word frequency, (d) following vowel, (e) previous manner, (f) previous voicing, (g) word position, (h) minimal pair status

7.3 DURATION OF [hw]

As in the reading data, analyses of duration and were performed on those tokens (n=211) which exhibited some degree of voiceless frication, removing all tokens coded as [w] and those [hw] tokens which contained a voicing band.

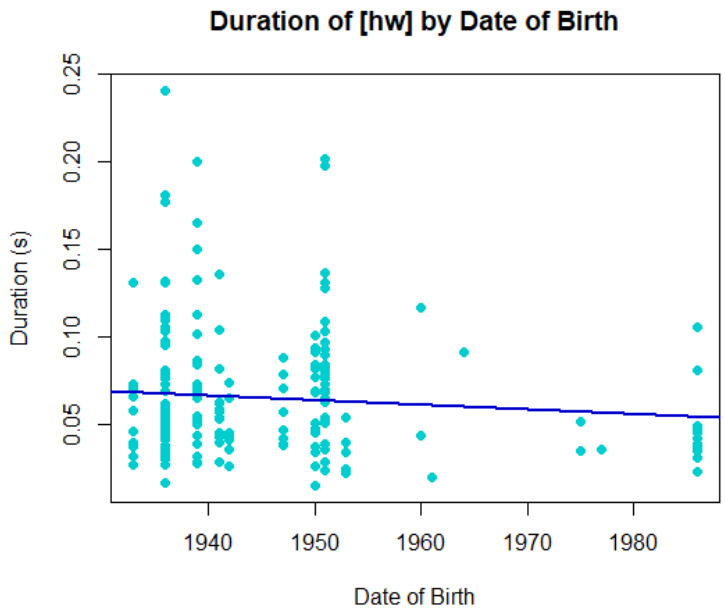

Figure 7.4 Duration of [hw] by date of birth

There was no correlation between [hw] duration and date of birth (r=-0.091), as visualized in Figure 7.4. Among the other social variables (Figure 7.5), there were potential effects of *city recorded*, *rurality*, and *education*. For *city recorded*, speakers from Columbia produced [hw] with longer frication (Mdn=0.068s) than those from

Spartanburg (Mdn=0.054s). By *rurality*, suburban speakers used longer frication (Mdn=0.073s) than those from other locations (rural: 0.055s, small town: 0.042s, large town: 0.054s). For *education*, graduate-educated speakers used longer frication (Mdn=0.066s) than high-school-educated speakers (Mdn=0.054s).

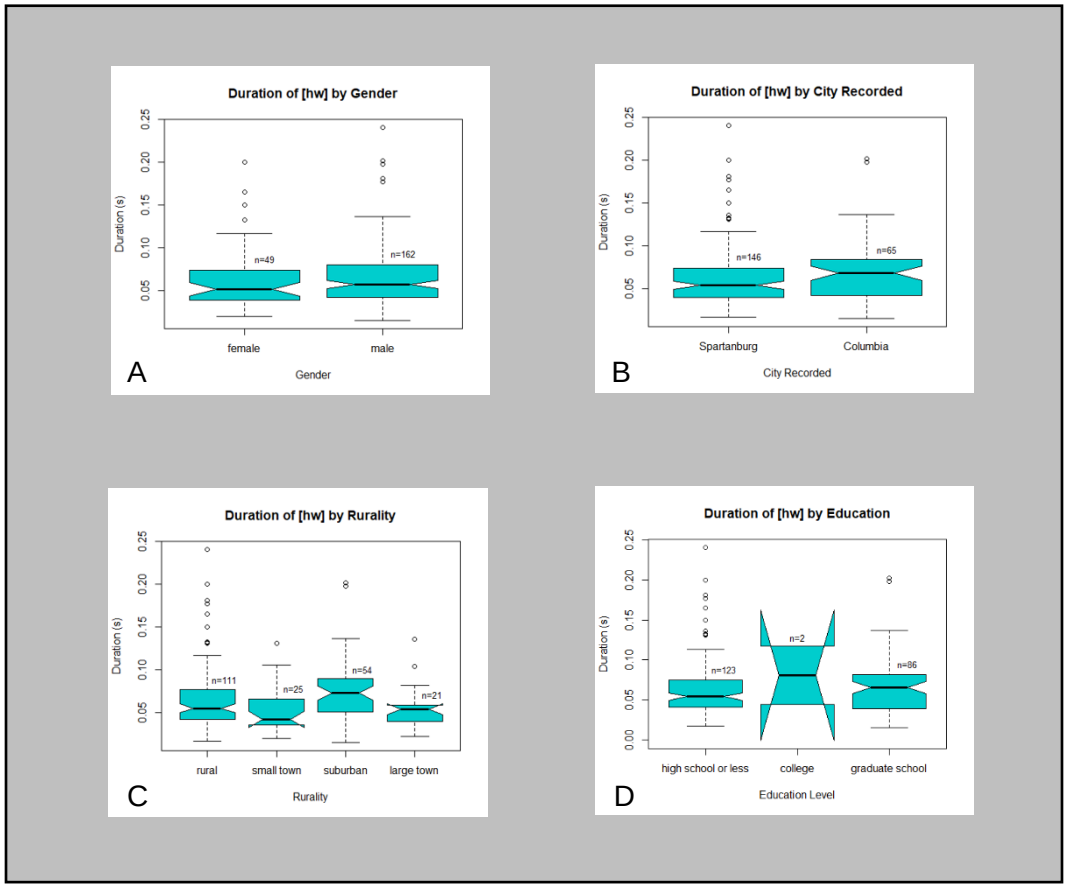

Figure 7.5 Duration of [hw] by social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

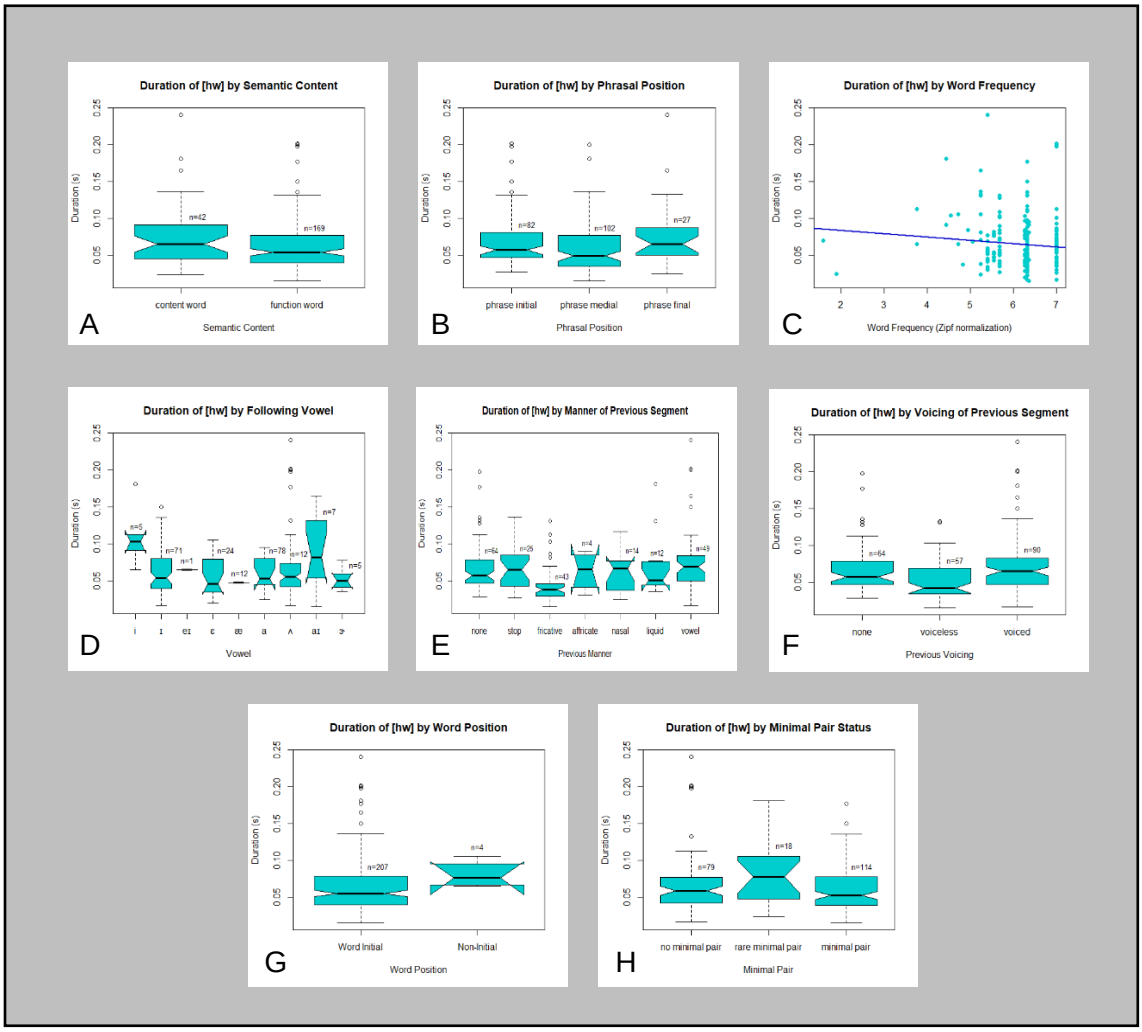

Figure 7.6 Duration of [hw] by linguistic variables: (a) semantic content, (b) phrasal position, (c) word frequency, (d) following vowel, (e) previous manner, (f) previous voicing, (g) word position, (h) minimal pair status

Among word-level variables, *semantic content*, *phrasal position*, *previous manner*, and *previous voicing* exhibited differences in [hw] duration across levels.

Content words (Mdn=0.065s) were produced with longer duration than function words

(Mdn=0.054s), phrase-final tokens (Mdn=0.065s) were longer than phrase-medial tokens (Mdn=0.049s), tokens following fricatives (Mdn=0.038s) were shorter than those following all other segments, and tokens following voiceless segments (Mdn=0.042s) were shorter than those following no segment (Mdn=0.057s) or a voiced segment (Mdn=0.065s). For *vowel*, [i] appeared to be preceded by longer frication (Mdn=0.103s) than all other vowels, but was represented by very few tokens.

While the graph in Figure 7.6c appears to show a substantial decrease in duration as frequency increases, the correlation between variables was very weak (r=-0.094), with a wide range of variation among all tokens with a Zipf value above 5.

7.4 COG OF [hw]

There was a weak correlation (r=-0.104) between COG and date of birth, such that COG decreased by 7.65 Hz with each one-year increase (see Figure 7.7).

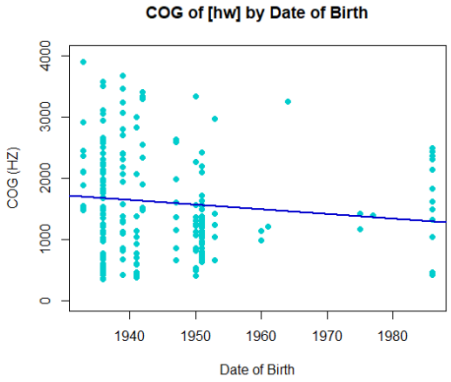

Figure 7.7 COG of [hw] by date of birth

Among the social variables, speakers from Columbia produced [hw] with lower COG than those from Spartanburg (Mdn=1232.147 Hz, Mdn=1526.922 Hz), and also with less variation (IQR=606.593 Hz, IQR=1374.018). Similarly, speakers from rural areas and small towns produced [hw] with high COG (Mdn=1611.186 Hz, 1855.658), while suburban and large town speakers used low COG (Mdn=1185.760 Hz, 1038.643 Hz). None of the word-level variables displayed COG values differing across levels (see Figure 7.9).

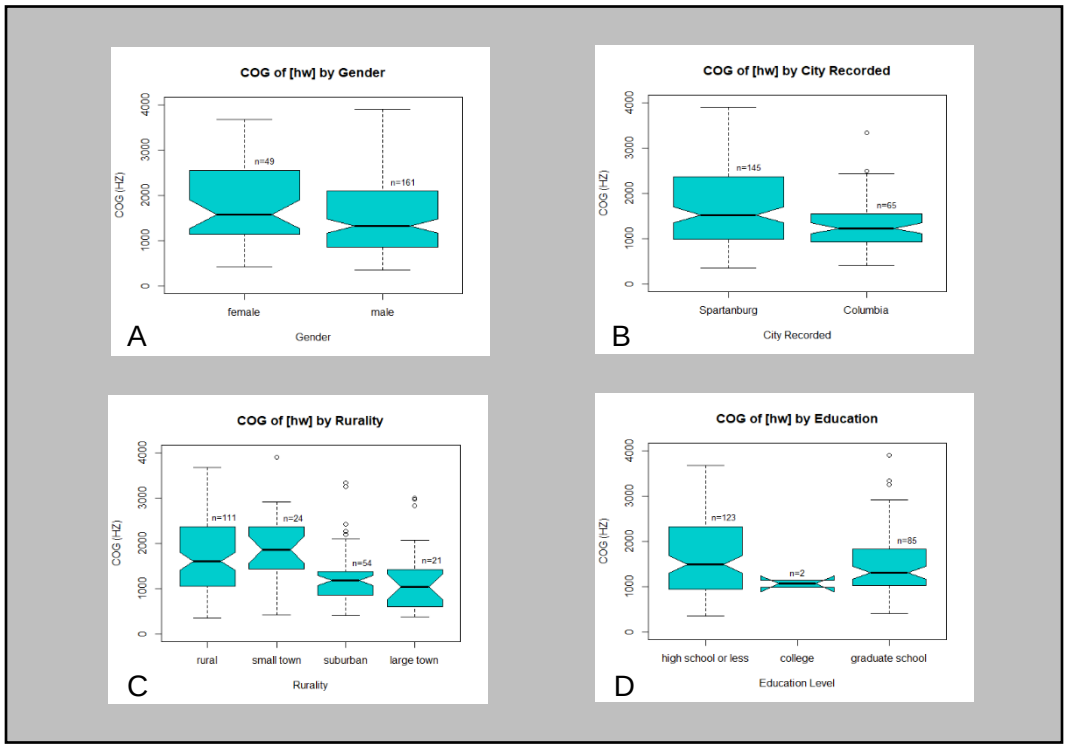

Figure 7.8 COG of [hw] by social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

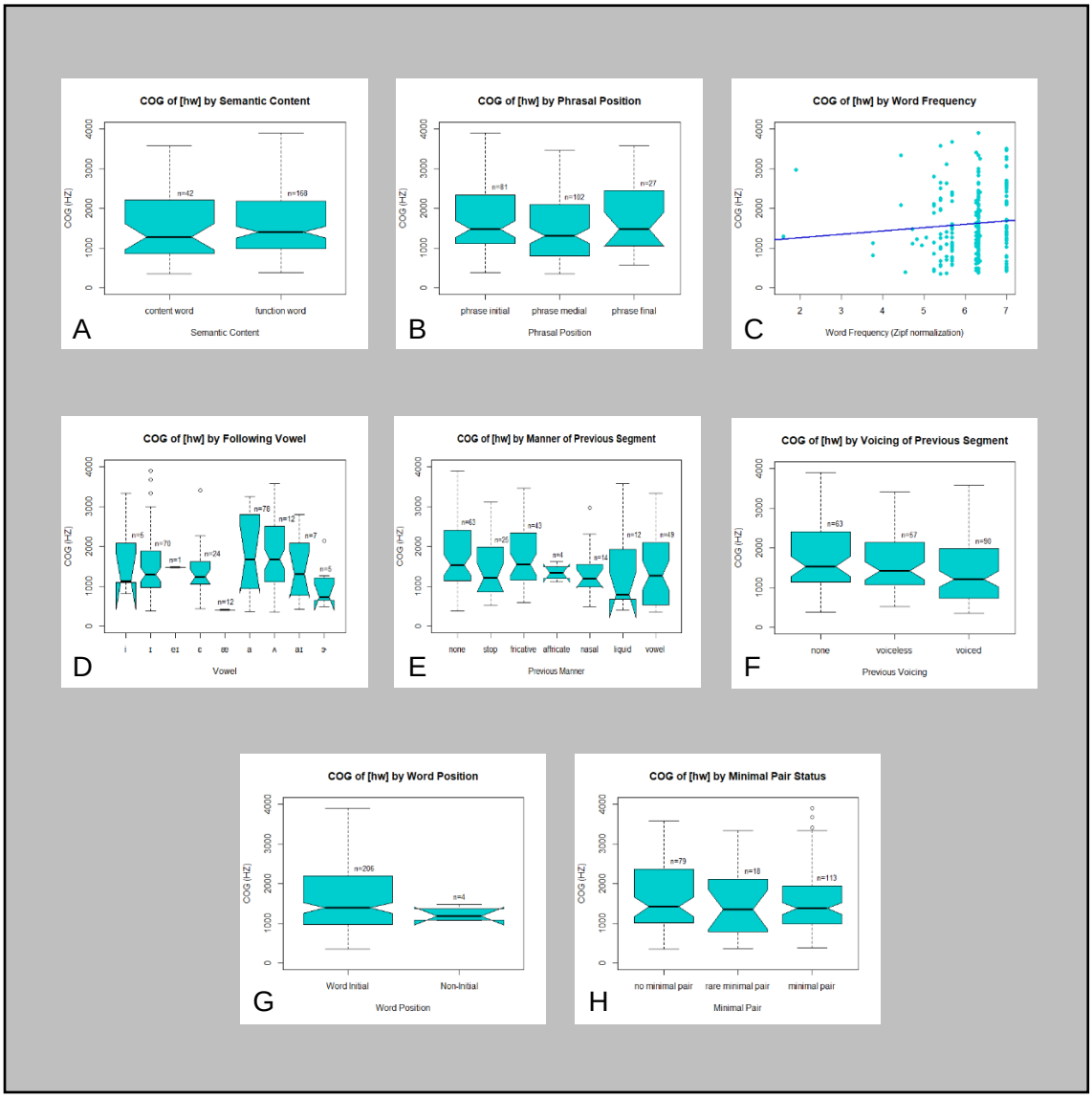

Figure 7.9 COG of [hw] by linguistic variables: (a) semantic content, (b) phrasal position, (c) word frequency, (d) following vowel, (e) previous manner, (f) previous voicing, (g) word position, (h) minimal pair status

7.5 VOICING OF [hw]

As in the reading data, a subset of those tokens coded as [hw] were characterized by a voicing band. Therefore, an analysis of *[hw] voicing* was performed on all tokens exhibiting frication (n=230).

As shown in Figure 7.10, voiced instances account for approximately 10% of [hw] tokens for speakers born from 1933 to 1960. After a date of birth of 1953, voiced tokens are not present in the corpus.

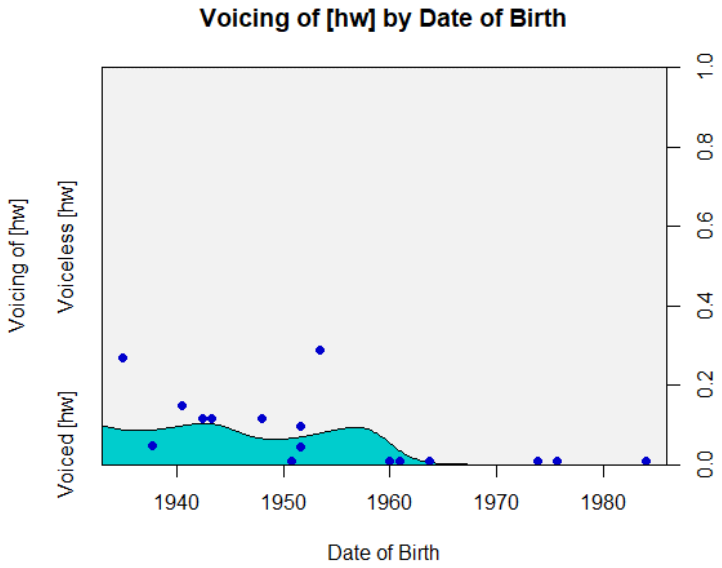

Figure 7.10 [hw] voicing by date of birth

Because of the scarcity of [hw] tokens, none of the social variables showed differences between levels at a 95% confidence level. However, tendencies were present

for *gender* and *city recorded*. In terms of gender, women voiced [hw] more than twice as often than men (14.0% vs. 6.4%). By city, speakers from Spartanburg used voiced [hw] more frequently than speakers from Columbia (10.4% vs. 3.0%).

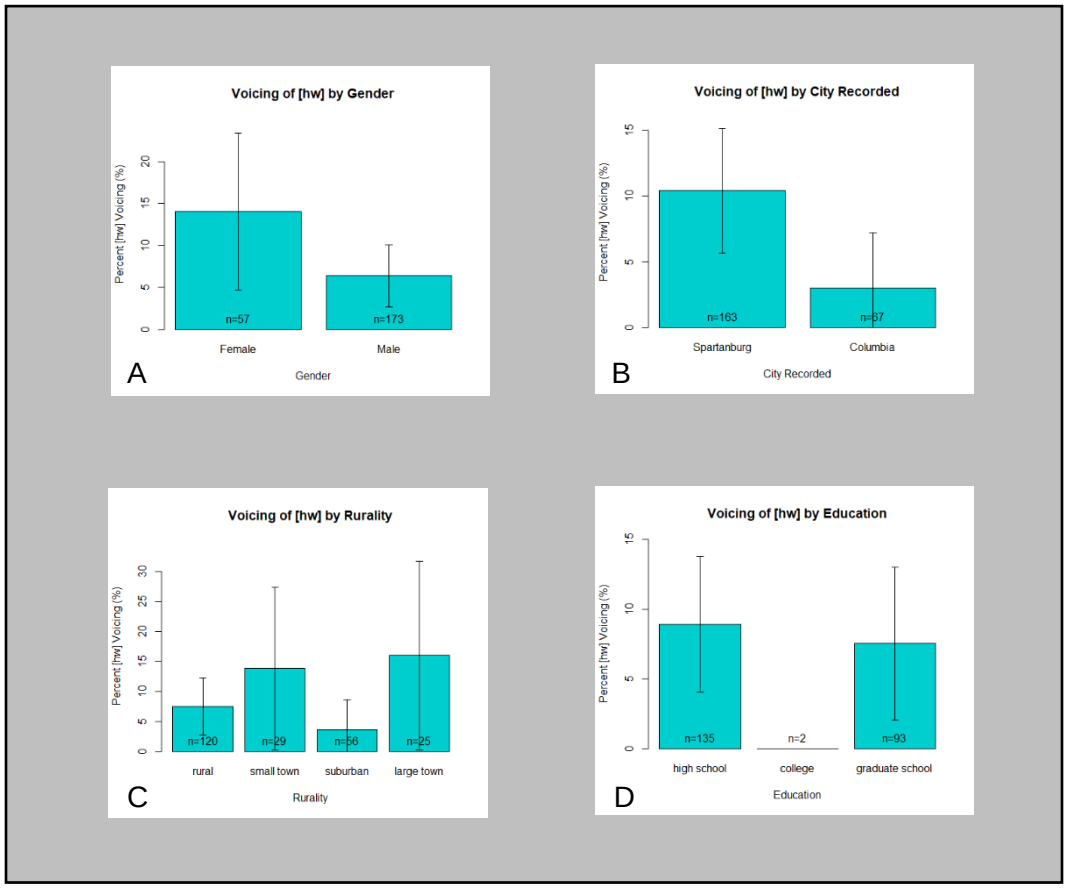

Figure 7.11 [hw] voicing by social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

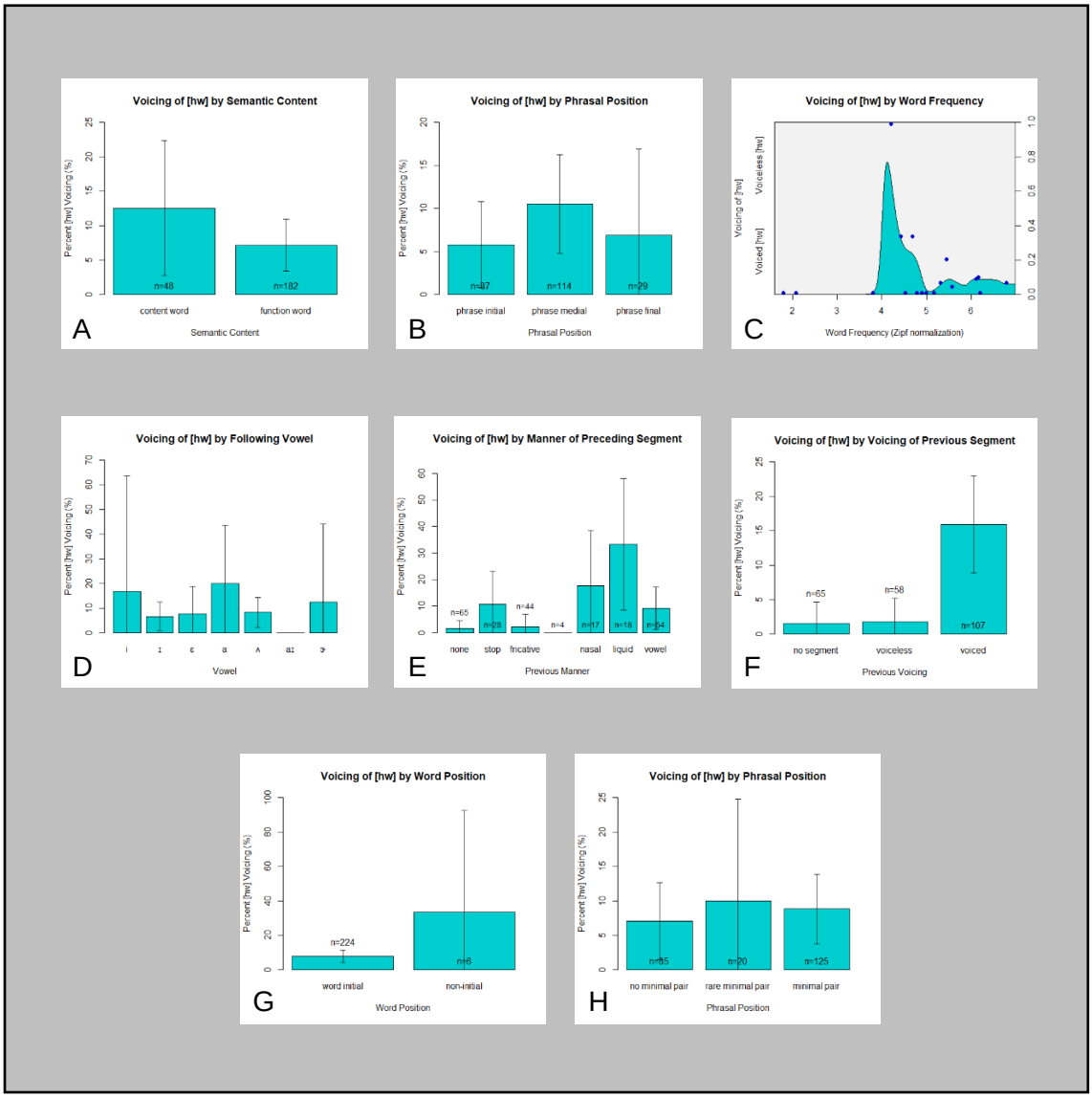

Figure 7.12 [hw] voicing by linguistic variables: (a) semantic content, (b) phrasal position, (c) word frequency, (d) following vowel, (e) previous manner, (f) previous voicing, (g) word position, (h) minimal pair status

As with the social variables, the low number of tokens available made it difficult to draw reliable conclusions for the majority of the linguistic variables. However, one

variable did produce clear results: [hw] voicing was most strongly conditioned by the voicing of the previous segment. 15.9% of tokens following a voiced segment were voiced, compared to less than 2% of tokens following a voiceless segment or no segment. This was further broken down by the manner of the previous segment: tokens following liquids were most frequently voiced (33.3%), followed by nasals (17.6%). By vowel, there was a potential tendency for [hw] preceding [aɪ] to be voiced less frequently than others. A potential tendency for word-medial tokens to be voiced more frequently than word-initial ones was also present, although difficult to determine since only 6 non-initial tokens existed.

CHAPTER 8
RESULTS: TASK COMPARISON

8.1 AREA OF OVERLAP

As shown below, rates of [hw]-usage tended to be much higher in the reading task than in the interview. This was true of almost all speakers, but was especially strong among those born from the years of 1950 to 1970.

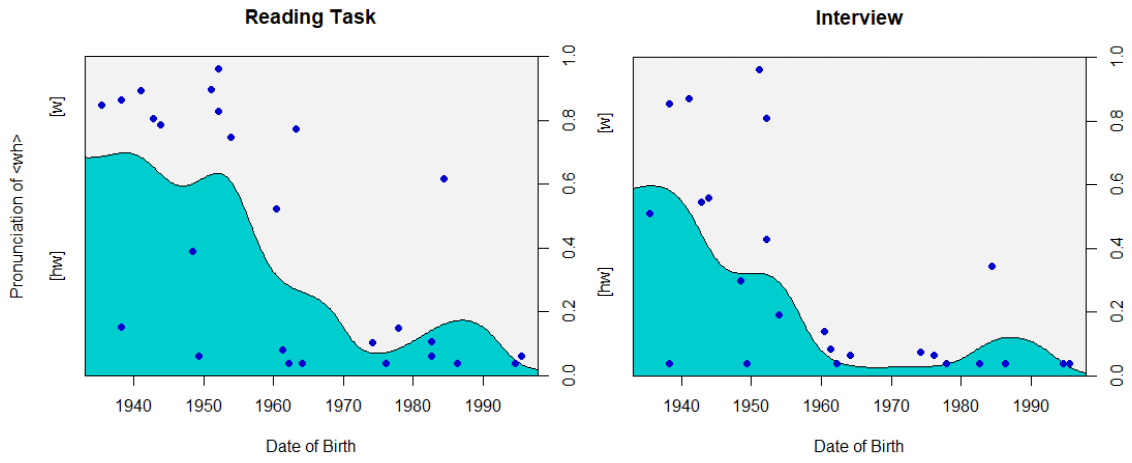

Figure 8.1 Presence of [hw] by date of birth and task type, from full dataset

Initially, this might appear to reflect a true task effect, indexing either formality or attention. However, as discussed in Chapter 7.1, the range of words occurring in the

interview corpus was much narrower than those from the reading task, with the bulk of tokens being question words such as *what* and *when*. Since an effect of semantic content was observed within both datasets, such that content words were more often aspirated than function words, it is possible that this differing distribution was simply an artifact of the words represented in each task, with 375 out of 1064 reading tokens being content words (35.2%), versus 141 out of 855 interview tokens (16.5%).

To account for the variation between datasets, therefore, tokens within the current comparative analysis were limited to those words which appeared more than 15 times in the interview corpus: *what, whatever, when, where, which, while,* and all forms of *white*. Together, these tokens accounted for 801 (93.7%) of all interview tokens, and for 318 (29.9%) of the reading tokens. This largely removed the issue of differing representation of semantic content: 23 of 318 reading tokens were content words (7.2%), and 90 of 801 interview tokens (11.2%).

8.2 PRESENCE OF [hw]

An overall comparison of the reading and interview tasks revealed that substantial differences were still present between tasks, such that [hw] tokens occurred more frequently in the reading data than the interview (38.4% vs. 27.1%), as shown in Figure 8.2.

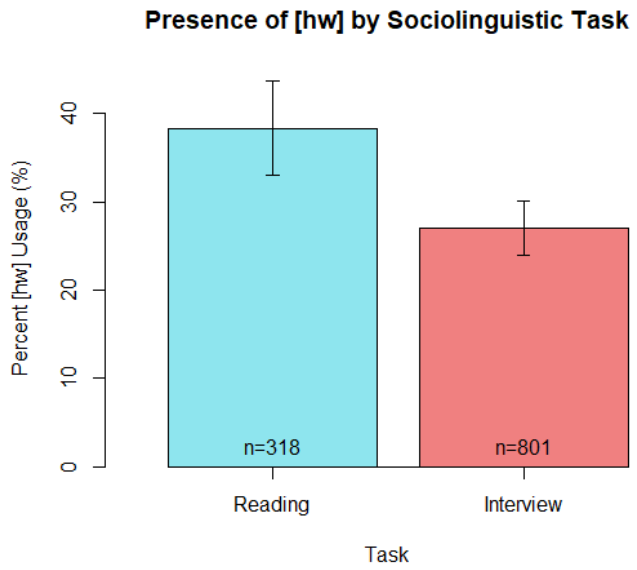

Figure 8.2 Presence of [hw] by sociolinguistic task

When task type was compared across date of birth, a pattern emerged whereby rates of [hw]-production were higher for the reading task than the interview for all ages (see Figure 8.3a). The largest differences across task type occurred between speakers born in the late 1940s and 1955. As shown in Figure 8.3b, the pattern of *reading* > *interview* was true across individual speakers, as well, with only three participants exhibiting more frequent usage of [hw] in the interview, and all three of these having low rates overall.

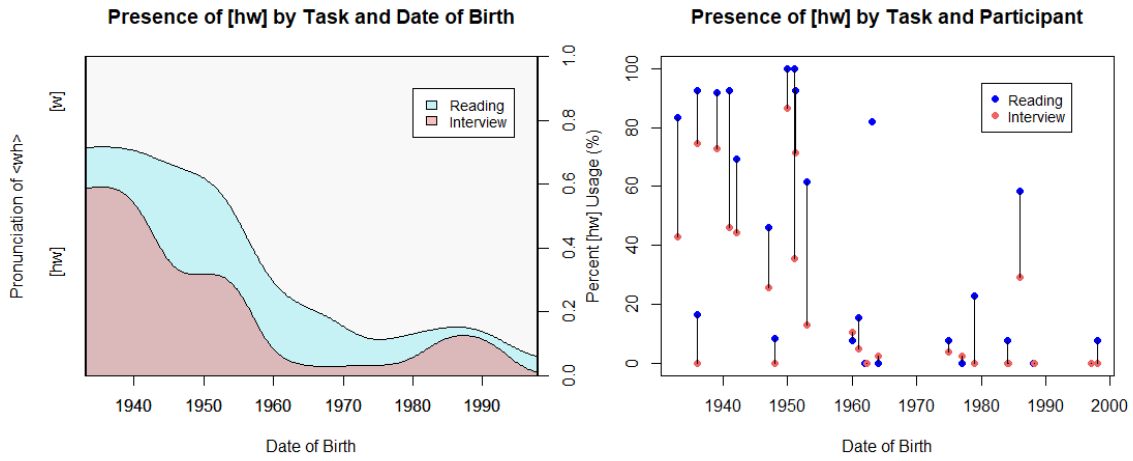

Figure 8.3 Presence of [hw] by task type and date of birth, according to (a) aggregate data and (b) individual participant

Following date of birth, the other social variables were examined for interactions with *task*. For *gender*, differences existed within and across task, but were consistent within each variable; the frequency of [hw]-usage decreased in the interview task at approximately the same rate for men and women (Figure 8.4a). A similar pattern was observed for *city recorded*, such that the rate of decrease across task was similar for the two locations (Figure 8.4b).

An interaction between *rurality* and *task* was present, such that rural speakers used [hw] equally across tasks, while speakers from large towns showed much higher rates of [hw] in the reading task than in the interview, and small town and suburban speaker exhibited modest task differences (Figure 8.4c).

Similarly, there was an interaction between *education* and *task* (Figure 8.4d), such that high-school educated speakers produced similar rates of [hw] across task, while college- and graduate-educated speakers used [hw] less often in the interview.

No significant interactions with task appeared to be present for the linguistic variables.

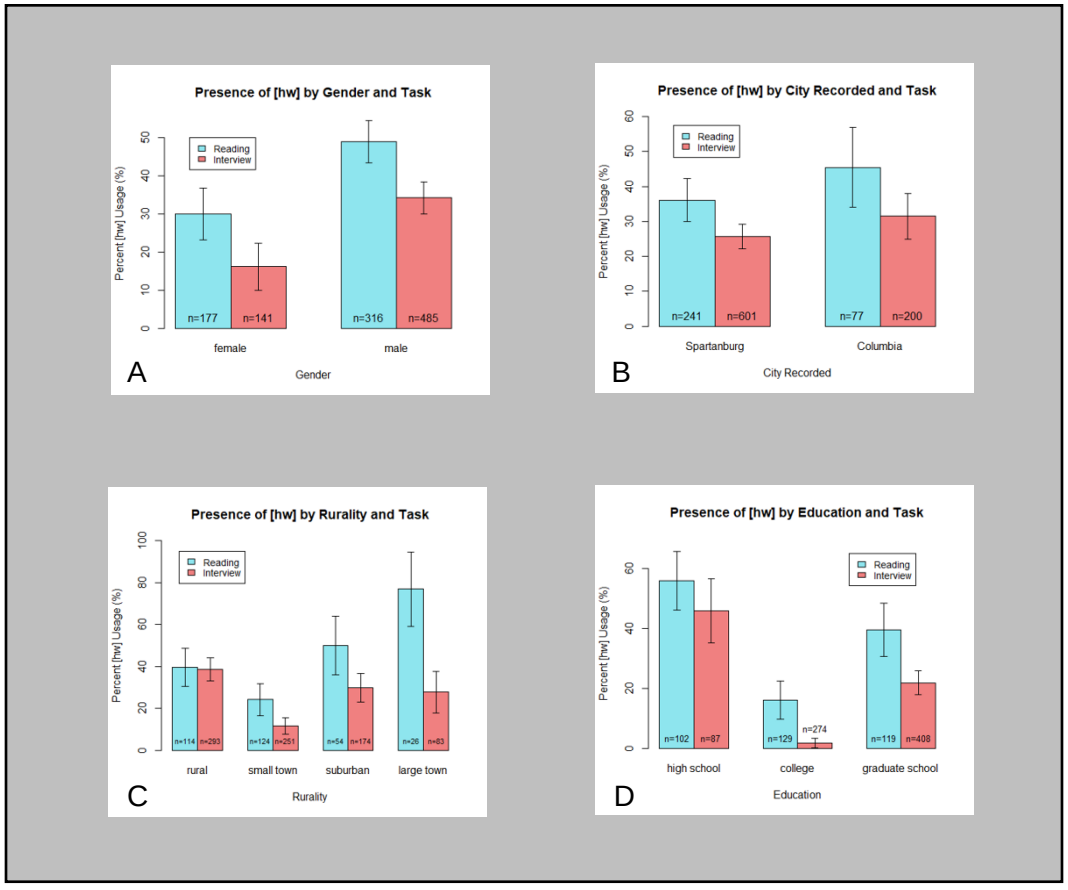

Figure 8.4 Presence of [hw] by task type and social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

8.3 DURATION OF [hw]

For the comparisons of duration and COG across task, all [w] tokens and instances of voiced [hw] were removed from the analysis, leaving 320 voiceless [hw] tokens. Among this dataset, there was no difference in [hw] length between the reading task and the interview (Mdn=0.053s vs. Mdn=0.054s).

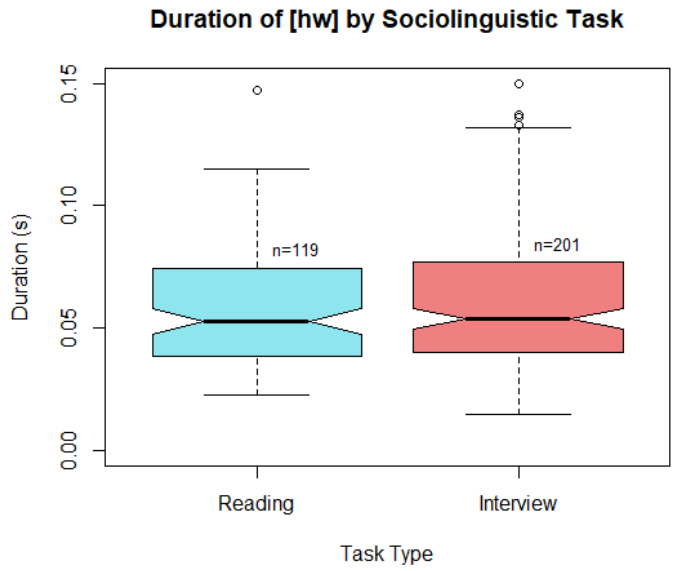

Figure 8.5 Duration of [hw] by sociolinguistic task

In the reading task, there was a significant linear decrease in duration as date of birth increased [F(1,117)=8.128, *p*=.005], such that an increase of one year was accompanied by a 0.0005s (0.5ms) decrease in frication length. No such relationship was present for the interview task; as shown in Figure 8.6a, [hw] duration was similar among

the oldest speakers across both tasks, and then diverged as date of birth increased, with duration decreasing for the reading task and remaining stable for the interview. However, when broken down by individual speakers, there was no pattern whether [hw] tokens were longer in the reading task or interview across date of birth.

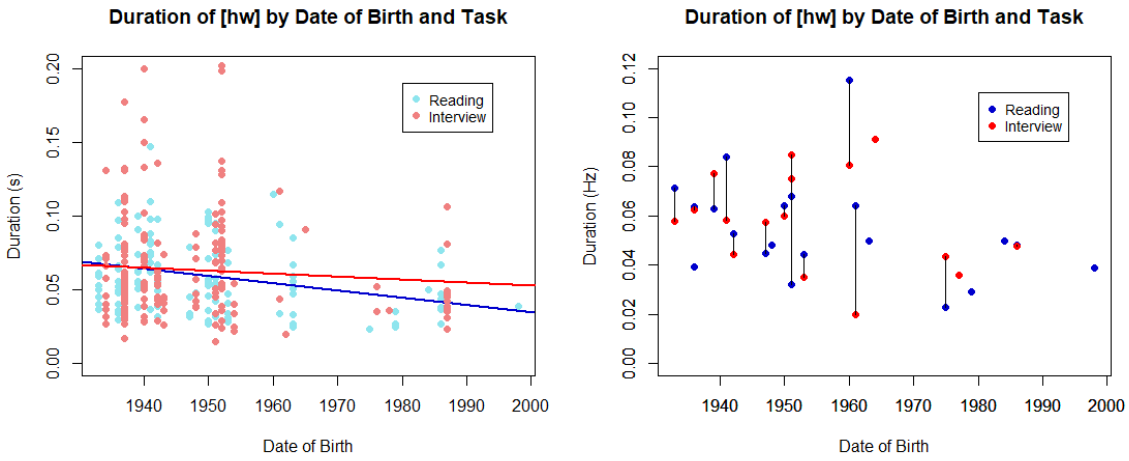

Figure 8.6 Duration of [hw] by task type and date of birth, according to (a) individual tokens and (b) averages by participant

There was no difference in duration between tasks when divided by gender, although men tended to produce longer tokens than women in both the reading task and interview (see Figure 8.7a). An interaction was present between *city recorded* and *task*, whereby speakers recorded in Spartanburg produced tokens of similar length across tasks (Mdn=0.052s, 0.054s), while speakers recorded in Columbia produced longer tokens in the interview (Mdn=0.068s) than the reading task (Mdn=0.054s). Similar results were

observed with respect to *rurality* and *education*, with the only significant differences across task taking place among suburban speakers (reading: 0.056s, interview: 0.073s) and graduate-educated speakers (reading: 0.053s, interview: 0.065s).

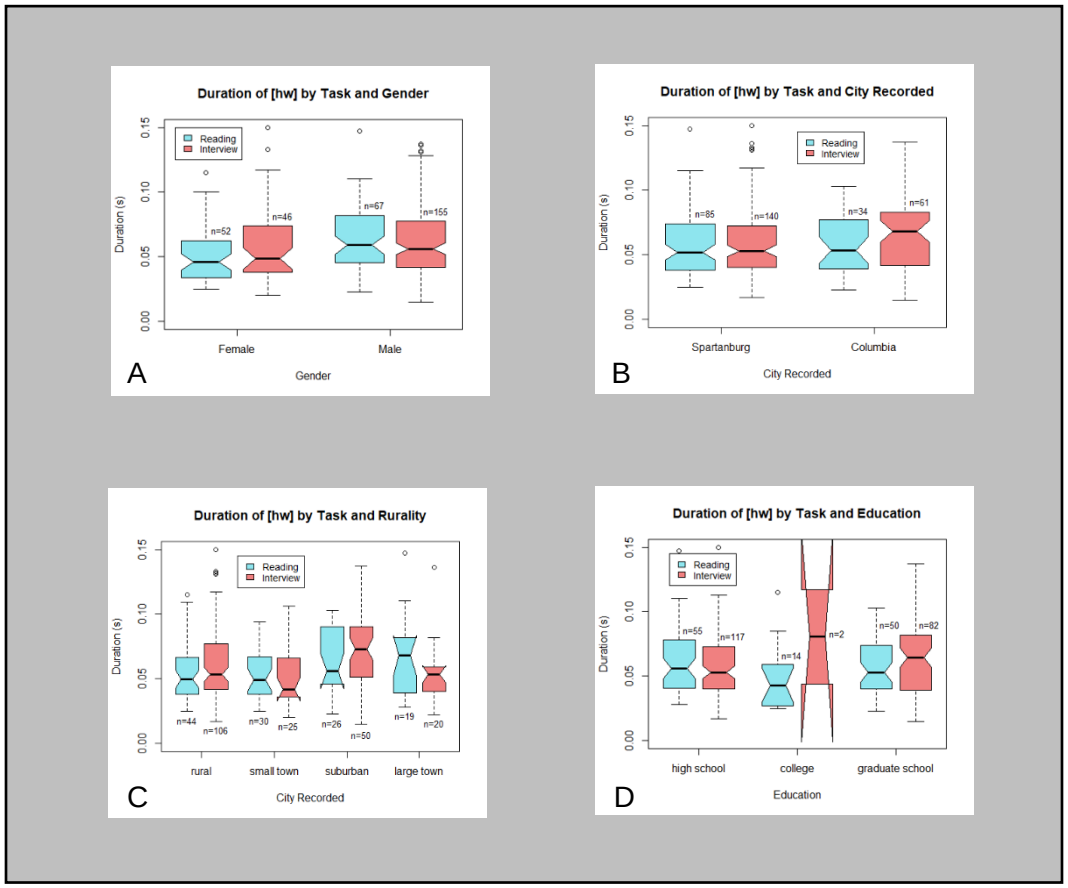

Figure 8.7 Duration of [hw] by task type and social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

Duration was compared across the linguistic variables of *previous voicing*, *previous manner*, *word frequency*, *semantic content*, and *phrasal position*, which had

previously been found to have a significant effect on frication length. Of these, only *semantic content* exhibited an interaction with task, such that there was a task difference among content words but not function words, with content words being produced with longer duration in the reading task (Mdn=0.074s) than the interview (Mdn=0.057s).

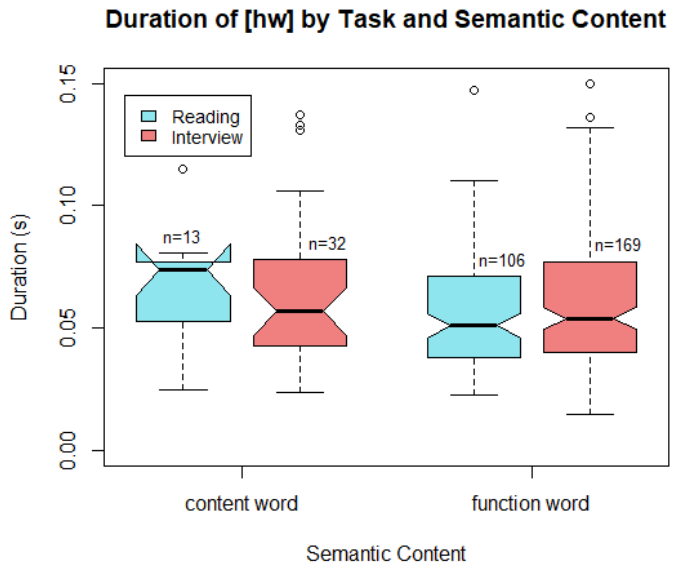

Figure 8.8 Duration of [hw] by task type and semantic content

8.4 COG OF [hw]

Overall, there was no difference in the COG of [hw] tokens from the reading task and the interview (Mdn=1333.029 Hz, Mdn=1403.709 Hz), as shown in Figure 8.9. However, the distribution across participants did differ between the two tasks, as shown by their interaction with *date of birth* (Figure 8.10).

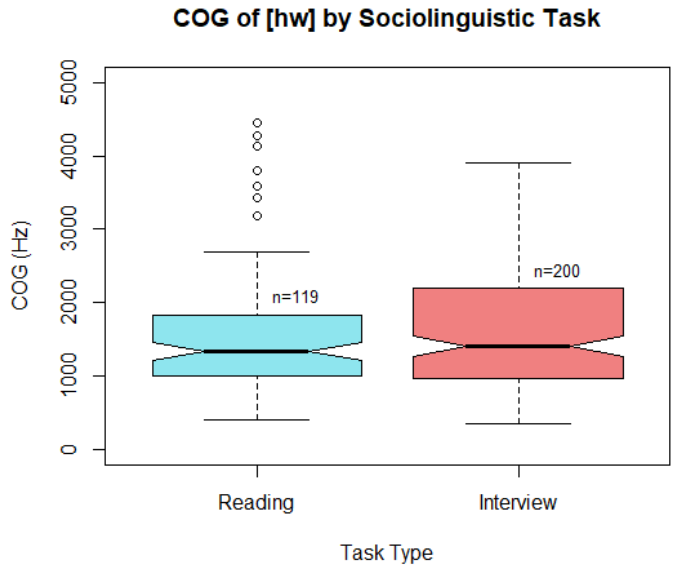

Figure 8.9 COG of [hw] by sociolinguistic task

There was a trend for COG values to increase as date of birth increased in the reading task, and to decrease as date of birth increased in the interview, as shown in Figure 8.10a. Regression models predicted an 8.86 Hz increase and 8.20 Hz decrease per year, respectively [$F(1,117)=3.174$, $p=.077$; $F(1,198)=2.529$, $p=.113$]. When broken down by participant, as in Figure 8.10b, speakers born before 1945 tended to have higher COG values in the interview, while speakers born from 1945 to 1965 tended to have higher COG values in the reading task.

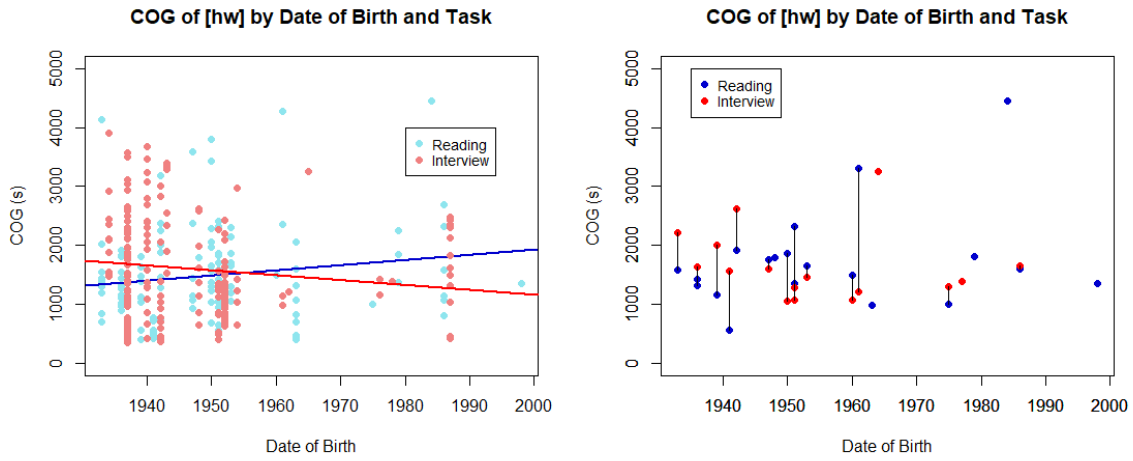

Figure 8.10 COG of [hw] by task type and date of birth, according to (a) individual tokens and (b) averages by participant

As shown in Figure 8.11, interactions were present for *city recorded*, *rurality*, and *education*, such that speakers from Spartanburg, rural speakers, small-town speakers, and high-school-educated speakers produced higher COG values for interview tokens than reading tokens. Speakers recorded in Spartanburg produced [hw] tokens with a median COG of 1310.144 Hz in the reading task and 1547.749 Hz in the interview, rural speakers had a median COG of 1414.609 Hz in the reading task on 1680.447 Hz in the interview, small-town speakers had a median COG of 1348.141 Hz in the reading task and 1855.658 Hz in the interview, and high-school-educated speakers had a median COG of 1282.450 Hz in the reading task and 1526.922 Hz in the interview. On the other hand, speakers recorded in Columbia and suburban areas produced higher COG values in the reading task. Columbia speakers had a median COG of 1503.256 Hz for reading tokens and

1231.147 Hz for interview tokens, and suburban speakers had a median COG of 1503.256 Hz for reading tokens and 1158.164 Hz for interview tokens. No significant interactions were present for word-level variables.

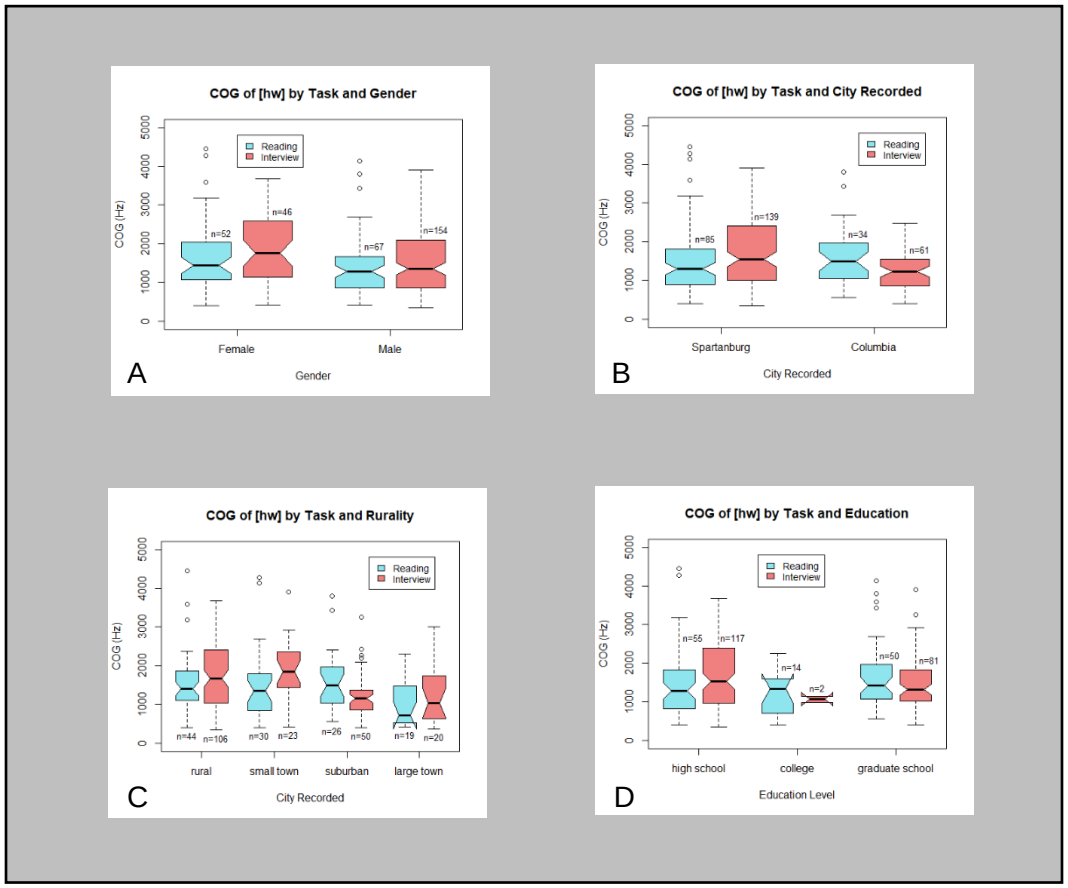

Figure 8.11 COG of [hw] by task type and social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

8.5 VOICING OF [hw]

A comparison of [hw] voicing across task type revealed that voiced tokens of [hw] were more common in the interview than in the reading task (7.4% vs. 2.5%), as shown in Figure 8.12.

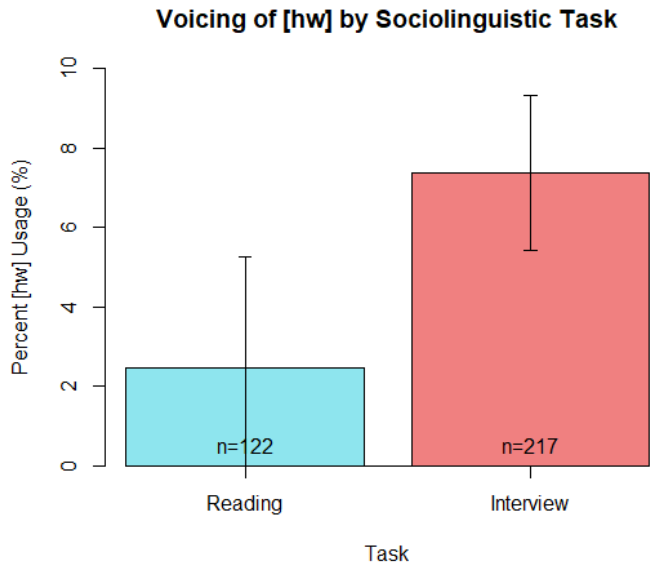

Figure 8.12 [hw] voicing by sociolinguistic task

No voiced tokens appeared among speakers born after 1953 (see Figure 8.13). Only three participants produced voiced tokens in the reading task, compared to nine participants who produced voiced tokens in the interview.

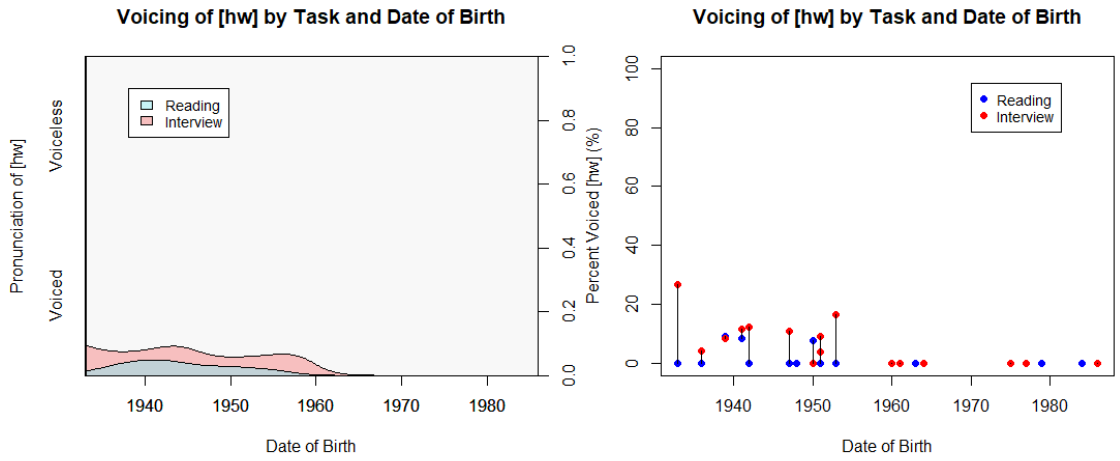

Figure 8.13 [hw] voicing by task type and date of birth, according to (a) individual tokens and (b) averages by participant

Although voiced tokens were invariably more common in the interview, the limited data meant that task differences were rarely significant within levels of independent variables. However, one nearly significant interaction was present for *city recorded*, such that rates of [hw] voicing were similar for Spartanburg speakers in the reading task (2.3%), Columbia speakers in the reading task (2.9%), and Columbia speakers in the interview (3.2%), while Spartanburg speakers voiced [hw] tokens with much higher frequency in the interview (9.1%). No significant interactions appeared to exist between *voicing* and linguistic variables.

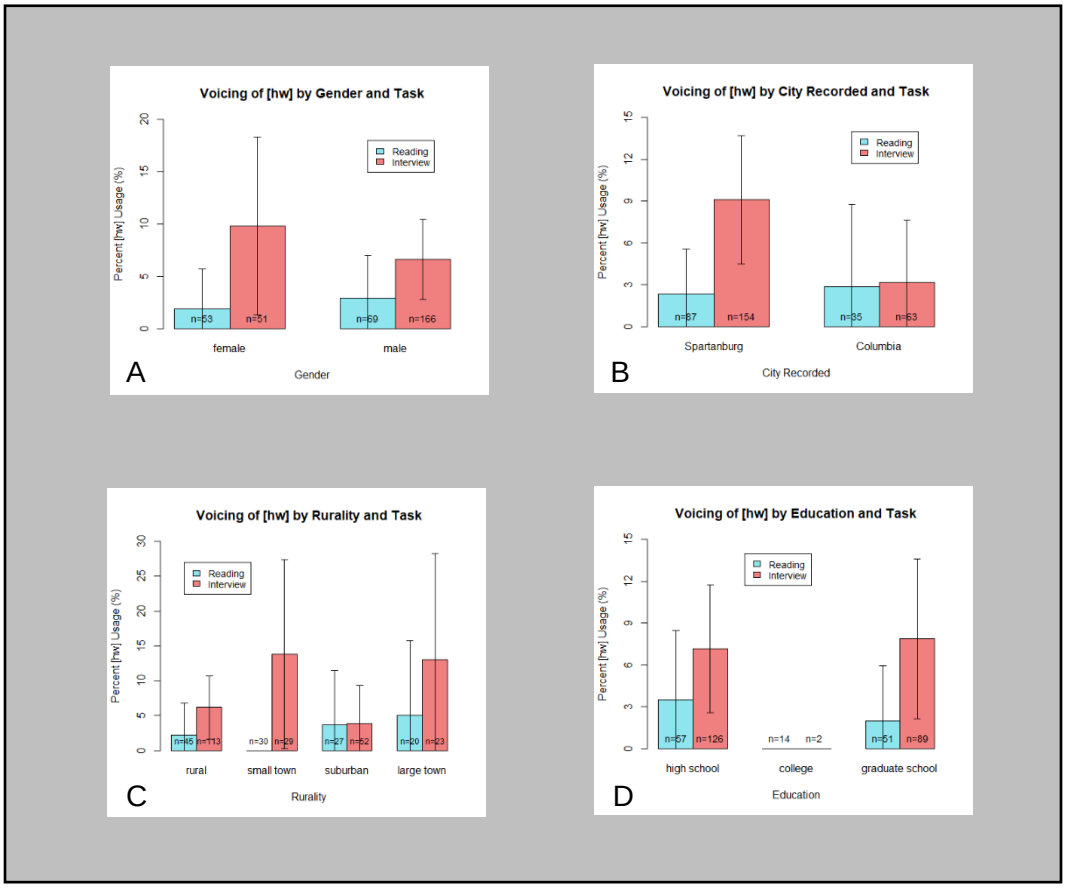

Figure 8.14 [hw] voicing by task type and social variables: (a) gender, (b) city recorded, (c) rurality, and (d) education

CHAPTER 9
DISCUSSION

9.1 RESULTS SYNOPSIS

9.1.1 Presence results

In the preceding analyses, four variables were targeted as metrics of variation in <wh> pronunciation: the presence, duration, COG, and voicing quality of aspiration. The first of these variables, *presence,* identified the sound used to express orthographic <wh>: [hw] or [w]. This variable was found to be strongly governed by date of birth, with rates dramatically decreasing among speakers born from 1960 to 1970.

To determine the best-fitting combination of variables contributing to the sampled distribution of [hw]-presence, a logistic regression was fit for the reading data. The resulting model included the variables of word position, date of birth, city recorded, and semantic content, such that older speakers and speakers from Columbia, SC were more likely to produce [hw], and that [hw] was more likely to occur when it occurred word-initially or in a word with semantic content.

Interview data showed lower overall rates of [hw]-production, consistent with the lower sample of content words, but a similar general pattern to that observed across speakers in the reading data. Differences across variable-level were observed for gender, rurality, and education, with higher rates of [hw]-usage among men, rural speakers, and

high-school-educated speakers. These findings are consistent with the general sociolinguistic expectation that nonmobile older rural males, or NORM speakers, are expected to employ the most traditional speech (Chambers & Trudgill, 1980). At the word level, differences were also present across semantic content, phrasal position, and previous manner, with content words and words in phrase-final position eliciting higher rates of [hw], and preceding nasals reducing the likelihood of its production. These first two findings may reflect the same phenomenon, as all words in final position were content words; the logistic regression from the reading data suggests that the stronger of the two predictors was semantic content, with words carrying greater semantic weight receiving a more careful pronunciation.

When subsets of overlapping tokens were compared across the reading and interview tasks, a pattern emerged in which reading tokens were more frequently pronounced with aspiration than those from the interview. Interactions were also present between rurality and task type, and education and task type, such that rural and high-school-educated speakers did not exhibit differences in [hw]-production across task type, while other groups did.

9.1.2 Duration results

For the reading data, a mixed linear regression selected a model with only one variable, which could be either previous voicing, word frequency, semantic content, or phrasal position. Specifically, words preceded by no segment, high frequency words, function words, and phrase-initial words were produced with shorter duration than their

corresponding alternate levels. These categories, which have a high level of overlap, reflect an overall tendency for words occurring at the beginning of an utterance, which tend to be highly common question words such as *what*, *when*, and *which*, to be pronounced less distinctly and with greater phonological erosion.

Highly similar results across word-level variables emerged from the interview data, with content words and phrase-final words produced with longer aspiration, and tokens following fricatives and voiceless segments produced with shorter aspiration. Although all subject-level effects in the reading data were eliminated when submitted to a mixed model, the raw numbers used in the interview analyses suggested the possibility that certain social variables might affect duration. Here, speakers recorded in Columbia, speakers from suburban areas, and graduate-educated speakers exhibited longer frication in [hw] tokens than those from other variable levels. Combined, these characteristics (highly educated and non-rural) suggest that the speech of these participants may have been less colloquial, and possibly more carefully enunciated, resulting in more pronounced, distinctive [hw] tokens.

No overall differences in duration were present between the reading and interview tasks. However, there were interactions between task type and each of the variables of city recorded, rurality, and education, whereby speakers recorded in Columbia, suburban speakers, and graduate-educated speakers produced significantly longer tokens in the interview than in the reading task.

9.1.3 COG results

For the reading data, a mixed linear regression selected a complex model predicting the COG of [hw] tokens, even after the contextual variables relating to surrounding segments were removed. This model included the variables of rurality, phrasal position, education, and the interactions of phrasal position and rurality and phrasal position and education. Specifically, phrase-medial tokens were produced with lower COG than phrase-initial tokens, while speakers from large towns tended to exhibit lower COG values, and high-school-educated speakers tended to exhibit higher COG values. Interactions revealed that college-educated speakers possessed significantly higher COG values in medial position, and suburban speakers possessed significantly lower COG values in medial position.

The interview data revealed similar results, with lower COG values produced by speakers from Columbia, suburban speakers, and large-town speakers. When the two tasks were compared, it was found that speakers from Spartanburg, rural speakers, small-town speakers, and high-school-educated speakers produced higher COG values for interview tokens than reading tokens, while other speakers tended to remain the same across tasks.

9.1.4 Voicing results

Out of the 406 reading tokens realized as [hw], 28 of these tokens were pronounced with voiced frication. A logistic regression model revealed that variation within the data could be explained by the variables of previous manner, word position,

city recorded, and education level, as well as the interaction of previous manner with each of the other variables. Main effects were present for previous manner, such that [hw] tokens following nasals were most likely to be voiced; word position, such that tokens occurring word-medially were more likely to be voiced than word-initial tokens; and city recorded, such that participants recorded in Spartanburg were more likely to voice tokens. Interactions revealed that tokens following nasals were disproportionately more likely to be voiced when in word-medial position, when produced by a speaker from Spartanburg, or when produced by a graduate-educated or higher-school-educated speaker, while tokens following liquids were more likely to be voiced when produced by a high-school-educated speaker.

The interview data revealed a slightly different, but closely-related conditioning variable, previous voicing, with almost all voiced tokens occurring after a preceding voiced segment. In this dataset, liquids were more likely than nasals to condition voicing, although, as in the reading data, both had higher rates than other preceding manners. Finally, there was a tendency for speakers from Spartanburg to voice [hw] tokens more than speakers from Columbia.

A comparison between the two task types showed that [hw] voicing was more common in the interview across all social variables. A particularly strong difference between the two tasks was present for speakers recorded in Spartanburg.

9.2 AGE-GRADED LOSS OF [hw]

Table 9.1 [hw] producers by decade

Decade	**ID**	**Date of Birth**	**[hw] > 10%**	**% [hw] producers**	**[hw] > 10%**	**% [hw] producers**
			Reading		***Interview***	
1930s	P27	1933	Yes	100%	Yes	75%
	P8	1936	Yes			
	P7	1936	Yes		Yes	
	P4	1939	Yes		Yes	
1940s	P21	1941	Yes	75%	Yes	75%
	P6	1942	Yes		Yes	
	P9	1947	Yes		Yes	
	P18	1948				
1950s	P3	1950	Yes	100%	Yes	100%
	P13	1951	Yes		Yes	
	P16	1951	Yes		Yes	
	P20	1953	Yes		Yes	
1960s	P22	1960	Yes	40%		0%
	P24	1961				
	P5	1962				
	P1	1963	Yes			
	P2	1964				
1970s	P15	1975		33%		0%
	P19	1977				
	P26	1979	Yes			
1980s	P17	1984		25%		25%
	P25	1984				
	P10	1986	Yes		Yes	
	P23	1988				
1990s	P11	1997		0%		0%
	P14	1998				

As described above, there was a strong relationship between [hw]-usage and age, such that occurrences of [hw] were rare among participants born after 1970, aside from

one participant born in 1986 whose usage patterned with older speakers. Unlike the pattern observed in Canadian English by Chambers (2002), this relationship was non-linear: in the reading task, rates of [hw]-production were concentrated within the 50-90% range until a birthdate of 1960, after which occurrences sharply dropped. On the other hand, Chambers's rates of [hw]-usage represented the percentage of a population possessing [hw]; if every participant in the current study were binarily categorized in this way, then the reading data would indeed appear more linear, as shown in Table 9.1. Interestingly, the interview data, which arguably reflect a closer approximation of real-world usage, maintain the abrupt cutoff at the 1960s. In both tasks, however, speakers born in the 1960s appear to be the first to show substantial [hw]-loss.

As evidenced by the differences between the reading and interview tasks, not all speakers who were aware of a distinction between [w] and [hw], or even possessed [hw] in their own phonemic inventory, produced this distinction with any kind of regularity. Furthermore, while some speakers showed a certain amount of predictable variation (e.g. P22 produced [hw] almost exclusively in content words), the majority showed less identifiable patterns, potentially governed by a combination of phonetic context, prosodic stress, and pragmatic intent, but not obvious from a listener's standpoint.

The fact that some level of this variation was present among speakers of all ages suggests that a certain amount of [hw]~[w] variation is a natural product of casual speech, and will pose few problems in comprehension for listeners. As such, as [w] has becoming increasingly mainstream in American English, speakers may increasingly acquire habits in production that approach that mainstream, even if they preserve [hw] as

an underlying representation in the mental lexicon. This is particularly likely to happen in commonly-used function words, which are easy to phonetically erode in any context without semantic loss: as an extreme example, the phrase *What are you doing?* can be reduced to *'cha doin'?* without confusing the listener; it is therefore unsurprising that [hw] could unconsciously become [w] in such a context. Indeed, several speakers, when the purpose of the study was explained following the interview, expressed surprise that they did not always use [hw], or that it was disappearing in the speech of those around them.

In light of this, it is noteworthy to consider those comments from online discourse (see Table 3.5) arguing for the necessity of preserving [hw] as a means to distinguish between minimal pairs such as *wear/where* and *whether/weather*. If [hw] were being retained in SWAE so as to preserve these distinctions, then we might expect to see more frequent (or stronger) [hw]-productions among words possessing a [w] ~ [hw] minimal pair. In reality, this phenomenon was not observed for any of the dependent variables, in any task. As many [w] ~ [hw] pairs (including the two mentioned above) consist of a [w] noun or verb and a [hw] question word or conjunction, it is indeed likely that syntactic and contextual factors render the need for such a distinction unnecessary, making [hw] a low-stakes speech sound for younger speakers to acquire or preserve.

9.3 SOCIAL CORRELATES OF [hw]

Across tasks, four social variables were found to contribute to the likelihood of <wh> being realized as [hw]: from the interview data, gender, rurality, and education

each displayed differences across levels; from the reading task, city recorded was included in the logistic regression.

Results from the interview showed that men, rural speakers, and high-school educated speakers used [hw] most frequently. Each of these groups is one which, according to traditional sociolinguistic expectation, is more likely than its counterparts to preserve traditional pronunciation. However, when broken down by participant, as in Table 9.2, two separate patterns emerged. According to one (Pattern A, shown in blue), speakers from more isolated areas and with less education pronounced <wh> as [hw] more frequently than others of the same gender and similar age. This pattern almost exclusively applied to speakers born before 1950 (although it arguably governed the results for P2, as well). On the other hand, according to Pattern B (shown in green), speakers with higher education and a suburban upbringing tended to show higher frequencies.

I argue that these two patterns represent two different phenomena, indexing two different sets of values attached to [hw]. Among speakers for whom Pattern A is active, [hw] is a natural feature of their native dialect: speakers who have had low contact with other regions or populations therefore preserve it to a higher degree. For speakers using Pattern B, [hw] may be a marker of "correct" pronunciation, used to index proper speech acquired through higher education.

Table 9.2 Interview [hw] rates by participant. Blue rows represent Pattern A, with darker rows indicating combinations of features contributing to greater [hw]-usage and lighter rows contributing to reduced [hw]-usage; green rows represent Pattern B. Features which are understood to be the governing factors for the individual participant are represented in bold.

ID	Date of Birth	Gender	Rurality	Education	% [hw]
P27	1933	Male	Small town	**Graduate**	42.9
P8	1936	Female	Rural	High school	0.0
P7	1936	Male	**Rural**	**High school**	74.2
P4	1939	Female	**Rural**	**High school**	75.7
P21	1941	Male	**Large town**	High school	46.2
P6	1942	Female	**Rural**	**High school**	47.4
P9	1947	Female	Rural	**Graduate**	23.7
P18	1948	Male	Small town	**Graduate**	0.0
P3	1950	Male	**Suburban**	**Graduate**	84.2
P13	1951	Male	Suburban	Graduate	35.5
P16	1951	Male	**Suburban**	**Graduate**	70.3
P20	1953	Female	Large town	High school	14.0
P22	1960	Female	Rural	College	9.1
P24	1961	Female	Small town	High school	4.3
P5	1962	Male	**Rural**	College	0.0
P2	1964	Male	**Suburban**	**Graduate**	0.0
P15	1975	Male	**Suburban**	**Graduate**	3.4
P19	1977	Female	Small town	**Graduate**	2.4
P26	1979	Female	Rural	College	0.0
P17	1984	Female	Rural	High school	0.0
P25	1984	Male	Small town	College	0.0
P10	1986	Male	Small town	**Graduate**	27.9
P23	1988	Female	Small town	Graduate	0.0
P11	1997	Female	Small town	College	0.0
P14	1998	Female	Small town	College	0.0

Further supporting the above analysis, duration results for the interview revealed that speakers who were more likely to produce [hw] according to Pattern B (suburban

speakers, graduate-educated speakers, and speakers from Columbia) also produced [hw] tokens with longer duration, that is, tokens which were stronger and more salient. If these speakers were using [hw] to project an "educated" identity, it makes sense that they would produce such tokens more deliberately and/or emphatically than those speakers for whom [hw] was simply a natural dialectal feature.

In the reading task, on the other hand, none of the above variables were significant; only the variable of city recorded contributed to the logistic regression model (although when considered as an isolated variable, male speakers did display higher [hw]-rates than female speakers). The results of the model showed that speakers from Columbia, when considered in light of the effects of word position, date of birth, and semantic content, produced [hw] more frequently than speakers from Spartanburg. However, the speakers recorded in Columbia (P3, P10, P11, P13, P14, P15, P16) did tend to be more educated and less geographically isolated, indicating that these results perhaps also reflected Pattern B.

9.4 ACOUSTIC VARIABILITY

While social variation was expected to primarily be reflected through the binary production of [hw] vs. [w], fine-grained acoustic detail was considered as well in an attempt to determine whether the gradual loss of [hw] takes place via the sharp deletion of frication, or through a weakening processes in which emphatic production of [hw] are replaced by less salient ones, and then further to a simple [w]. While there was little

evidence of such an effect taking place diachronically, variation was observed along other variables, as described in the following sections.

9.4.1. Duration as a measure of token strength

Results from both the reading and interview tasks indicated that [hw]-length covaried with a number of word-level variables, such that more prominent words (i.e. content words, words occurring phrase-finally) were produced with longer aspiration than words which might be expected to receive less emphasis. In the same way that a reduction from [hw] to [w] may be expected to take place first in function words, a reduction in the articulatory effort expended to produce a [hw] token appears to take place in similar contexts.

9.4.2 Voicing as evidence for a phonetic sequence

As discussed in Chapter 5.4, it has been unclear from previous studies whether the sound under consideration is stored in speakers' phonemic inventory as /ʍ/ or /hw/. From a historical perspective, Minkova (2004) has argued that /hw/→/w/ was part of a larger shift in Middle English whereby initial h-clusters (i.e. /hn/, /hr/, /hl/) were reduced to the second phoneme, and that /hw/ itself, never fully lost, was reanalyzed as standard more than once on the basis of the <wh> orthography. However, it is difficult to determine how closely this reflects the cognitive representations of modern (not to mention Southern American) speakers of English. While the spelling <wh> might seem to suggest a sequence of segments, English possesses several <h>-containing graphemes

that represent a single phoneme: <ch>, <sh>, and <th> bear no predictable relation to <c>, <s>, and <t>. Indeed, it is possible that the phonemic representation of <wh> words differs from speaker to speaker: the majority of Chapter 3 commenters represented the <wh> sound as <hw>, or spoke of "pronouncing the h", but one made a distinction between a "hard w" and "soft w": an evident layman's description of [w] and [ʍ].

While the present study does not claim to provide a conclusive answer to this phonological problem, a clue toward the underlying representation of [hw] was provided by the existence of a substantial number of voiced [hw] tokens in the corpus. If the distinction between <wh> and <w> is phonemic, then we would expect them to be represented as similar bundles of features, differing only with respect to [voice]. In this case, we would expect a "voiced <wh>" to be a voiced labiovelar fricative, i.e. [w]. In the corpus, however, several speakers produced tokens possessing both clear frication and a clear voicing band (see Figure 9.1), a sound which might be most accurately transcribed as [ɦw]. This variant showed strong signs of allophonic distribution: although it was not required to appear in any particular environment, its appearance was strongly governed by the surrounding context, such that it was almost always preceded by a voiced segment, and was particularly common in word-medial position.

Among these speakers, then, <wh> appears to be mentally represented as [hw], with the [h] undergoing voicing in appropriate environments. On the other hand, at least one speaker (P21) showed phonetic evidence of [ʍ] pronunciations in other contexts, and yet still exhibited a small number of voiced [hw] tokens. While no solution to this can be provided at present with confidence, one possibility is that these speakers possessed an

underlying representation of /hw/, yet realized it as [ʍ] in some contexts, possibly in itself as a form of lenition.

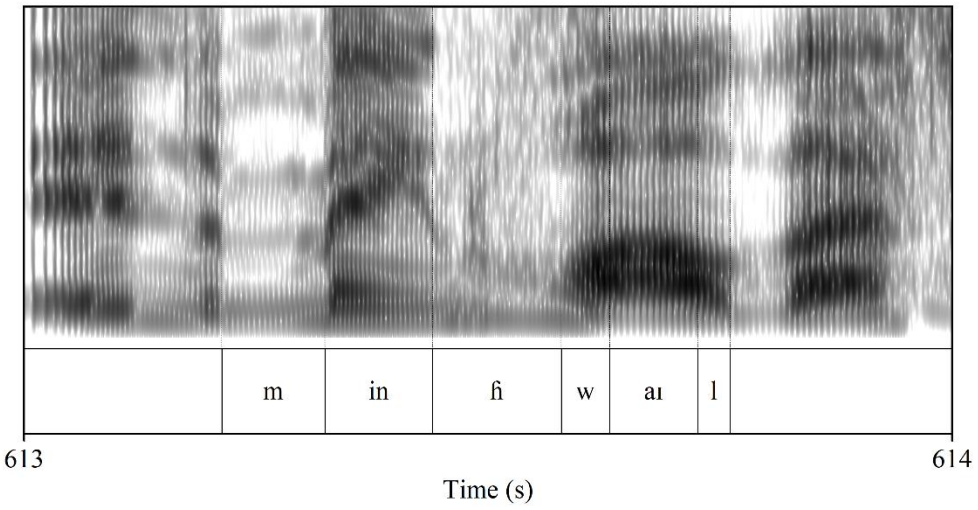

Figure 9.1 Voiced [hw] token of "meanwhile"

9.4.3 COG as a measure of allophonic quality

Prior to analysis, it was expected that higher COG values would reflect a stronger fricative intensity, with higher frequencies resulting from a more forceful pressure of air through the lips. However, as individual tokens were viewed, it became apparent that the frication produced in [hw] tokens varied in its quality and place of articulation, as shown in Figure 9.2: some tokens functioned as true voiceless labiovelar glides, with the aperiodic noise clustered around the expected loci of [u] formants, while others exhibited

a highly diffuse frication scattered across all frequencies, auditorily similar to [h]; others fell somewhere in between.

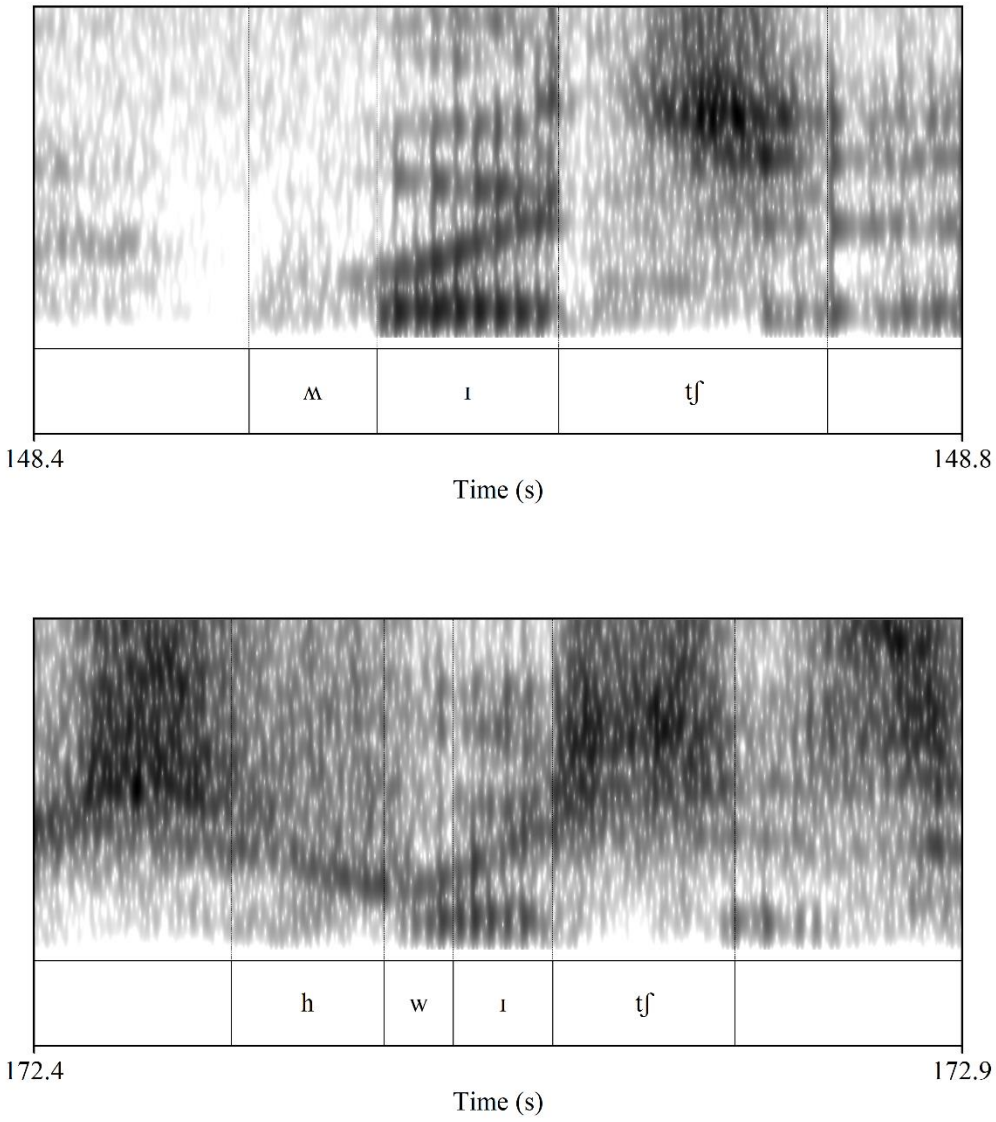

Figure 9.2 Allophonic variations of [hw] in "which": (top) labial fricative, (bottom) glottal fricative

As a result of these observations, a low COG value was taken to reflect a labial fricative, while a high COG value reflected the more widely distributed frication of a true [hw]. As described above, a number of social factors were found to play into COG, including city recorded, rurality, and education level. In terms of main effects from the linear regression fit for the reading data, speakers from large towns had a much stronger tendency toward labial frication than any other group, while high-school-educated speakers showed a greater tendency toward glottal frication than college- or graduate-educated speakers. According to the interview data, lower COG values (indicating more labialized frication) were produced by speakers from Columbia, suburban speakers, and large-town speakers.

Although teasing apart the distinction between these variants is beyond the scope of the present study, requiring a more accurate metric or combination of metrics than a simple COG value, the above results do seem to suggest that less localized speakers exhibited more labialized frication. Whether this was due to a weakening of [hw] due to contact with [w], a different underlying representation than in local speech, or an attempt to approach a more "cultured" pronunciation to reflect one's level of education is presently unclear.

9.5 FORMALITY EFFECTS

9.5.1 Presence of [hw] by task

Overall, rates of [hw]-production were significantly higher in the reading task than in the interview, both when the entire corpora were used and when tokens were

limited to those occurring at least 15 times in the interview (*what, whatever, when, where, which, while, white*). This was true not only of the aggregate data, but also of all but 3 individuals (each of whom had less than 12% [hw] tokens in both tasks, and whose percentage of [hw] tokens in the reading data was no more than 3% lower than their interview rate). Since these differences could not be attributed to differences in the words included in the corpora, they were taken to be a reflection of the nature of the task.

The reading task, in presenting sentences on a screen to be read aloud, was a performance-based task: that is, participants' focus was at least partially on ensuring that the correct forms of the words were produced. Due to this, and to the fact that it was the first task in the overall experiment, speakers were also likely to be less at ease during this portion, or to modify their speech in order to provide the best possible data, according to their perception of what the researcher would want. Furthermore, while participants were aware that a requirement for the was to have been born and raised in the Southern United States, and the researcher made no attempts to suppress her own SWAE accent, the strong stigmatization of Southern speech in broader culture may have led participants to modify aspects of their speech which they had overt awareness of, particularly during a task in which the obvious focus was on the qualities of their speech. The interview, on the other hand, was structured so as to seem as though participants were providing a record of their experiences with Southern culture, leading them to focus on the content of what they were saying, rather than their pronunciation. Participants also typically became significantly more comfortable over the course of the interview, often becoming noticeably animated.

As a result, the reading task may be understood as providing samples of formal, careful, and/or and high-attention speech, and the interview to provide samples on the familiar, casual, and/or low-attention end of the spectrum. Since [hw] tokens were more common in the reading task, the occasional merging of [hw] with [w] appears to have been a natural process in the speech of all participants, which they suppressed in more careful speech. This suggests that the [hw]-producers in this study considered [hw] to be the prestige variant, a marker of correct diction. Confirming this fact, rural and high-school-educated speakers as a group did not differ across task type, which other groups did: those participants with least exposure to outside speech used [hw] similarly across registers, while those with higher education and less geographic isolation used [hw] when speaking carefully, but [w] casually.

The implication of the above results is interesting, in that it counters the results from Chapter 3: [hw]-producers do not seem to perceive [hw] as a marker of Southernness or of the past. If they had done so, then [hw] should have occurred more frequently in the interview, where topics relating to the South and to childhood memories were approached in the form of a semi-casual conversation; instead, the opportunity to speak more casually resulted in more merged tokens. It therefore appears that the outside perception of [hw] as a marker of Southernness, old age, and traditionality, while potentially accurate, is not shared by those who produce the sound. Rather, they attach no social meaning to [hw], or one of cultured speech.

9.5.2 Duration and COG of [hw] by task

With respect to acoustic variation, the results were more complex. Exhibiting an opposite pattern to that observed above, speakers recorded in Columbia, suburban speakers, and graduate-educated speakers produced significantly longer tokens in the interview than in the reading task. These speakers, then, used [hw] less frequently in a casual context, but when they did produce it, did so more distinctly. It is possible that this reflects a less complete internalization of the sound, such that it does not appear consistently, but is produced with a full, unreduced articulation when it does appear.

COG results showed that speakers from Spartanburg, rural speakers, small-town speakers, and high-school-educated speakers produced higher COG values for interview tokens than reading tokens, while other speakers tended to remain the same across tasks. Interestingly, this group was the one which did not exhibit a task effect for the presence of [hw]: instead, the effect of task type appears to have conditioned allophonic variation. If higher COG is taken to reflect more [h]-like production, then these speakers produced more labialized, [ʍ]-like tokens when paying close attention to their speech, but leaned toward [hw] in a more casual context.

9.5.3 Anecdotal evidence

The subject P9 provided an interesting example of variation across tasks, in that she unintentionally provided a third task level through metalinguistic commentary. This speaker, a rurally-raised, graduate-educated female, pronounced fewer [hw] tokens than other speakers of similar age in the reading task (16 out of 43). Afterwards, she

mentioned that she had noticed a lot of <wh> words, giving “whippet” as an example of one she had seen. Notably, when reporting this, she pronounced “whippet” with [hw]; however, during the actual task, she had pronounced it as [w]. Furthermore, despite the potential for this comment to prime her toward using [hw] in the interview, her percentage was even lower during the second task (9 out of 39). At the end of the interview, when the purpose of the study was explained, she expressed surprise that she had ever pronounced <wh> as anything other than [hw].

This anecdote provides insight into the processes underlying [hw] variation in several ways. First, P9 was unaware that she was exhibiting variation at all. She possessed a conscious mental representation of [hw] for <wh> words, but more commonly produced it as a simple [w]. When paying the maximum amount of attention to a word (i.e. reporting on the presence of “whippet” and other words), she would be certain to use [hw], indicating that she understood this to be the correct pronunciation; however, when reading “whippet” in context, she used [w], despite its salience as an extremely uncommon word. In an even less attention-dependent setting, such as the interview, the likelihood of [hw]-usage reduced further: in the reading task, she pronounced “wheel” with [hw], but in the interview used [w] twice for the same word.

A partial explanation for this phenomenon may be found in P9’s occupation, that of a retired high school English teacher. As part of a generation overtly taught “enunciation” (see comment RE #94 on p. 30), followed by a career in education, it is likely that she strongly associated [hw] with <wh> on an intellectual level. However, after daily immersion in a population of speakers (i.e. high schoolers) that adopted the

merger in increasing numbers each year, it is also possible that she unconsciously absorbed elements of their speech and the speech of younger teachers.

While this explanation only addresses the [hw]-usage of a single participant, the same logic may apply to other speakers, as well. Many participants claimed to be unaware that [hw] was being lost, or to never have thought about it, while younger speakers who had already adopted [w] occasionally stated that they had never heard [hw], despite living in a community in which it frequently appeared. The low salience of this sound, then, may be responsible for its disappearance; since [hw] is rarely necessary for comprehension, and acoustically diffuse, its presence or absence may pass unnoticed by speakers without linguistic training, who simply produce the variant in relation to the exemplars stored from their own experience.

9.6 CONCLUSIONS

The results from this study present a set of patterns whereby [hw] is statistically associated with age, rurality, and low education, while simultaneously being perceived as cultured and reflective of correct education. As a result, although [hw] is naturally preserved in the speech of Southern speakers born prior to 1950 (particularly men, people raised in rural areas, and those with little education), it does not fully disappear after this point. Instead, it continues to be preserved by a shrinking number of speakers, particularly those who are highly educated. These speakers are more likely to use [hw] in careful and/or formal speech, and to more strongly enunciate it, indicating that they perceive it as the correct pronunciation. Ironically, this may lead to the production of

[hw], among these speakers, as an attempt to sound less Southern, while speakers of English from other regions perceive it as an index of Southernness. However, this phenomenon would only apply to a limited age group: speakers born after approximately 1970 appear to use [hw] very rarely regardless of social variables; it appears that among the younger demographic, the *wine-whine* merger is largely complete.

Acoustic results reveal the potential for further differentiation in [hw] categorization, with some tokens exhibiting labial frication concentrated at the [u] formants, some showing diffuse [h]-like frication, and some showing frication across all frequencies with bands of greater intensity at [u]. At present, it is unclear whether these variants reflect a true dichotomy in how phonemes or stored, or whether they simply represent a continuum of fricative strength; however, they do appear to covary with social categories, such that speakers with higher education and less localization used lower frication overall, and rural, less-educated speakers used lower frication in careful speech vs. casual speech. As such, [ʍ]-like tokens seem likely to be the prestige variant; however, additional analyses should investigate the nature of these different realizations and the most reliable metrics by which to quantify them.

9.7 FUTURE DIRECTIONS

While the present results indicated that [hw] was perceived by its users as having overt prestige, indexing education, the current study does not rule out the possibility that it might reflect Southern identity, as well, particularly among the older, more rural speakers that preserve it most fully, or among the younger speakers who reject its usage.

It would therefore be valuable to quantify informants' degree of identification with Southern culture through the use of a metric like that used by Reed (2016) to measure speakers' level of rootedness to their Appalachian locale. This could then be used to determine the relationship between one's Southern identity and usage of [hw] in order to determine its association, if any, with Southernness in the minds of Southerners.

Additionally, future studies could benefit from directly addressing the *wine-whine* merger within the sociolinguistic interview, so as to draw insights from the intuitions of native speakers about its distribution of usage, social indices, and the level of awareness surrounding its presence and ongoing loss. In doing so, it may be possible to gain an understanding, not only of the current state of the merger, but of the social and linguistic shifts triggering its final and rapid progression into the Southeastern US, as well as the remaining trajectory of the sound change's final stages.

REFERENCES

Bailey, G. (2002). Real and Apparent Time. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 349-372). Blackwell Publishing Ltd.

Boersma, Paul & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.36, retrieved from http://www.praat.org/

Bridwell, K. (2018). *The Distribution of [ʍ]: Geographic and Age-Related Factors Governing the* Wine-Whine *Merger.* Paper presented at LCUGA 5, University of Georgia.

Brysbaert, M. & New, B. (2009). Moving beyond Kucera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. Behavior Research Methods, *41*(4), 977-990.

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse studies*, *7*(4-5), 585-614.

Chambers, J. K. (1994). An introduction to dialect topography. *English World-Wide*, *15*(1), 35-53.

Chambers, J. K. (1998). *Whither withers: last gasp of the oldest sound change.* Presented at NWAV 27, University of Georgia.

Chambers, J. K. (2002). 14 Patterns of Variation including Change. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 349-372). Blackwell Publishing Ltd.

Chambers, J. K., & Trudgill, P. (1980). *Dialectology*. Cambridge University Press.

Chun, E. W. (2017). How to drop a name: Hybridity, purity, and the K-pop fan. *Language in Society*, *46*(1), 57-76.

Cruttenden, A. (1994). *Gimson's pronunciation of English.* London: Edward Arnold.

Cruttenden, A. (2014). *Gimson's pronunciation of English*. Routledge.

Crystal, T. H., & House, A. S. (1988). A note on the durations of fricatives in American English. *The Journal of the Acoustical Society of America*, *84*(5), 1932-1935.

Dalola, A. and Bridwell, K. (in press). The shape of [u]: Towards a Typology of Final Vowel Devoicing in Continental French. In *ICPhS 2019*.

Dalola, A. & Bridwell, K. (in progress): COG measures as a meaningful parameter in performing sociophonetic competence in L2 French.

Dalola, A., & Bullock, B. E. (2017). On Sociophonetic Competence: Phrase-final Vowel Devoicing in Native and Advanced L2 Speakers of French. *Studies in Second Language Acquisition*, *39*(4), 769-799.

De Wolf, G. D. (1992). *Social and regional factors in Canadian English: A study of phonological variables and grammatical items in Ottawa and Vancouver* (Vol. 2). Toronto: Canadian Scholars' Press.

Du Bois, J. W. (2007). The stance triangle. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, *164*, 139-182.

Eckert, Penelope (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*, 453-476.

Erker, D. G. (2010). A subsegmental approach to coda /s/ weakening in Dominican Spanish. *International Journal of the Sociology of Language*, *2010*(203), 9-26.

Forgue, G. J. (1986). American English at the Time of the Revolution. In *Dialect and Language Variation* (pp. 511-523).

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

Fox, J. (2003). *Effect Displays in R for Generalised Linear Models*. Journal of Statistical Software, *8(15)*, 1-27.

Fuchs, S. & Toda, M. (2010). Do differences in male versus female /s/ reflect biological or sociophonetic factors?. In S. Fuchs, M. Toda, & M. Żygis (Eds.) *Turbulent sounds: An interdisciplinary guide* (pp. 281-302). Walter de Gruyter.

Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, *32*(2), 141-174.

Harrington, J. (2006). An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics*, *34*(4), 439-457.

Heffernan, K. (2004). Evidence from HNR that /s/ is a social marker of gender. *Toronto Working Papers in Linguistics*, *23*.

Inoue, M. (2004). What does language remember?: Indexical inversion and the naturalized history of Japanese women. *Journal of Linguistic Anthropology*, *14*(1), 39-56.

Johnstone, B. (2018). *Discourse analysis*. John Wiley & Sons.

Koenig, L. L., Shadle, C. H., Preston, J. L., & Mooshammer, C. R. (2013). Toward improved spectral measures of /s/: Results from adolescents. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1175-1189.

Kruse, J. (2016). Accent variation reflected in the standard writing system of English. In Cook, V., & Ryan, D. (Eds.), *The Routledge Handbook of the English Writing System*, (pp. 175-187). Routledge.

Kurath, H., & McDavid, R. I. (1961). *The Pronunciation of English in the Atlantic States: Based upon the Collections of the Linguistic Atlas of the Eastern United States*. University of Michigan Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, *82*(13), 1-26.

Labov, W. (1963). The social motivation of a sound change. *Word*, *19*(3), 273-309.

Labov, W. (1966). *The Social Stratification of English in New York City*. Center for Applied Linguistics, Washington, DC.

Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English: Phonetics, phonology and sound change*. Walter de Gruyter.

Labov, W., Yaeger, M., & Steiner, R. (1972). *A quantitative study of sound change in progress* (Vol. 1). US Regional Survey.

Lee, H., & Jongman, A. (2016). A diachronic investigation of the vowels and fricatives in Korean: An acoustic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, *46*(2), 157-184.

Macaulay, R. K. (1977). *Language, Social Class and Education: A Glasgow Study*. Edinburgh: Edinburgh University Press.

Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America*, *125*(6), 3962-3973.

Minkova, D. (2004). Philology, linguistics, and the history of [hw]~[w]. In A. Curzan & K. Emmons (Eds.), *Studies in the History of the English Language II: Unfolding Conversations* (p. 7-46).

Ochs, E. (1992). Indexing gender. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon*, *11*, 335-358.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reed, P. E. (2016). "Sounding Appalachian: /aɪ/ Monophthongization, Rising Pitch Accents, and Rootedness". PhD thesis. University of South Carolina.

Roach, P. (2009). *English Phonetics and Phonology Fourth Edition: A Practical Course, 4th Ed*. Cambridge: Cambridge University Press.

Scargill, M. H., & Warkentyne, H. J. (1972). The Survey of Canadian English: A Report. *English Quarterly*.

Schwartz, M. F. (1968). Identification of speaker sex from isolated, voiceless fricatives. *The Journal of the Acoustical Society of America*, *43*(5), 1178-1179.

Silbert, N., & de Jong, K. (2008). Focus, prosodic context, and phonological feature specification: Patterns of variation in fricative production. *The Journal of the Acoustical Society of America*, *123*(5), 2769-2779.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & communication*, *23*(3-4), 193-229.

Stuart-Smith, J. (2007) Empirical evidence for gendered speech production: /s/ in Glaswegian. In J. Cole and J.I. Hualde (Eds.) *Laboratory Phonology 9* (pp. 65-86). Mouton de Gruyter: New York, USA,

Upton, C. (2008). Received pronunciation. *Varieties of English*, *1*, 237-252.

Van Son, R. J., & Pols, L. C. (1996, October). An acoustic profile of consonant reduction. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on* (Vol. 3, pp. 1529-1532). IEEE.

Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S, 4th Ed*. Springer, New York.

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

Zimman, L. (2017). Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and/s. *Language in Society*, *46*(3), 339-370.

APPENDIX A

MAP OF [ʍ] DISTRIBUTION

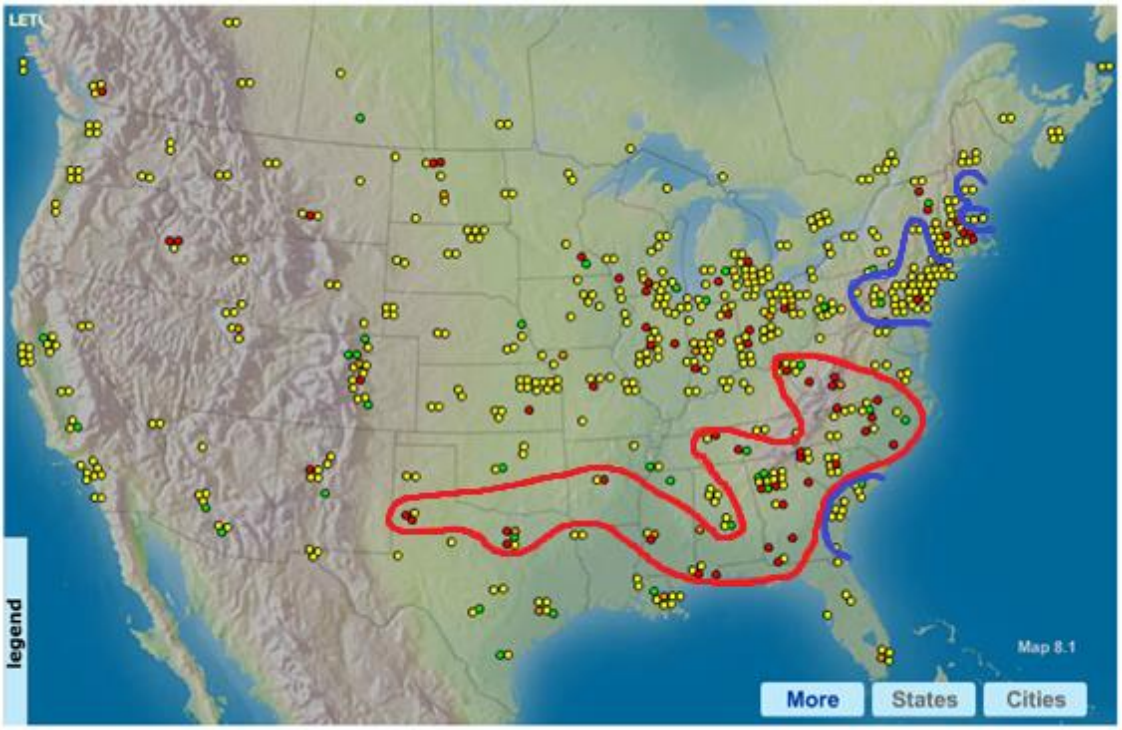

Figure A.1 Isogloss of the wine-whine *merger as mapped in the Atlast of North American English (Labov et al., 2006)*

APPENDIX B

INFORMANT DEMOGRAPHICS

Table B.1 Informant demographics

ID	Birth year	Gender	Birthplace	Hometown
P1*	1963	Female	Landrum, SC	Moore, SC
P2	1964	Male	Greer, SC	Spartanburg, SC
P3	1950	Male	Charleston, SC	Columbia, SC
P4	1939	Female	Moore, SC	Moore, SC
P5	1962	Male	Moore, SC	Moore, SC
P6	1942	Female	Greer, SC	Reidville, SC
P7	1936	Male	Reidville, SC	Reidville, SC
P8	1936	Female	Reidville, SC	Reidville, SC
P9	1947	Female	Moore, SC	Spartanburg, SC
P10	1986	Male	Hogansville, GA	Hogansville, GA
P11	1997	Female	Greenwood, SC	Columbia, SC
P13	1951	Male	Columbia, SC	Columbia, SC
P14	1998	Female	Charlotte, NC	Seabrook, TX
P15	1975	Male	Columbia, SC	Columbia, SC
P16	1951	Male	Newport News, VA	southeast VA
P17	1984	Female	Moore, SC	Moore, SC
P18	1948	Male	Fountain Inn, SC	Spartanburg, SC
P19	1977	Female	Waycross, GA	Spartanburg, SC
P20	1953	Female	Greenville, SC	Moore, SC
P21	1941	Male	Spartanburg, SC	Spartanburg, SC
P22	1960	Female	Spartanburg, SC	Spartanburg, SC
P23	1988	Female	Hartwell, GA	Moore, SC
P24	1960	Female	Wilkesboro, NC	Moore, SC
P25	1984	Male	Lyman, SC	Moore, SC
P26	1979	Female	Greenville, SC	Duncan, SC
P27	1933	Male	Lancaster, SC	SC

*Reading task only

APPENDIX C

STIMULI: READING PASSAGES

EXAMPLE PASSAGE:

The rainbow is the result of sunlight striking droplets of water in the air. These droplets bend the light so that it appears as many beautiful colors. According to legend, a pot of gold can be found at the end of the rainbow.

(Press the Enter key)

PASSAGE 1:

One October afternoon, Tina drove to the pet store. She had just moved into a new apartment that allowed animals, and was excited to pick out a pet. Inside the store, the saleslady was very friendly. "Which animal are you interested in?" she asked, gesturing toward a wide selection of cats, dogs, rabbits, and smaller animals. Tina examined a chubby hamster running on a wheel, a puppy with droopy ears, and a turtle with a swirly pattern on its shell, but none seemed quite right. Then she saw it: a beautiful calico kitten. It blinked at her with big blue eyes and twitched its whiskers. Tina's heart melted. She reached for the cat, but then saw the one sitting beside it. This one was a gray tabby cat with white paws. "I can't decide which one I like better!" she exclaimed. "Why not get both?" the saleslady suggested. On a whim, Tina agreed. A few minutes later, she was happily driving home with her two new kittens.

PASSAGE 2:

The schoolchildren ran happily around the playground. Some laughed as they swung from the monkey bars; others cheered as they watched the baseball game in the corner. Over by the swings, where Lisa and Katie were pushing each other higher and higher, Anna and Whitney had formed a gymnastics club. They were practicing handstands and cartwheels, and frequently falling down. Miss Taylor, the teacher on recess duty, hurried over to stop them before they injured themselves. The little girls whined as she told them that no flips were allowed, but the teacher remained firm. While her back was turned, though, she heard a loud thud. She gasped as she saw that a small boy named Joey had whacked his head on the slide. But the little boy only grinned at her and went back to playing.

Whispering a prayer of relief, Miss Taylor shook her head in frustration. She didn't know why she even bothered!

PASSAGE 3:

For a delicious homemade breakfast, follow this easy recipe. Preheat the oven to 200 degrees. Meanwhile, mix together whole wheat flour, sugar, baking powder, and salt. In another bowl, mix together milk, butter, and eggs. Add the dry ingredients to the milk mixture and whisk until moist. Spoon the batter onto a skillet over medium-high heat. Cook each pancake until bubbles form, then flip and cook until brown. Keep the pancakes warm in the oven, and serve with strawberries and whipped cream.

PASSAGE 4:

When I was young, I always spent the summer at my grandparents' farm in Kentucky. During the day, I climbed trees and helped around the house, and in the evenings, my grandfather would whittle toys for me out of pine and cedar wood. We liked to sit on the front porch together while he carved things like tops or whistles. One day, I asked him to teach me, so he showed me how to hold his knife carefully and carve away from my body. For my twelfth birthday, he gave me my own knife and a special stone to whet the blade, so that I could work whenever I felt like it. Still, my favorite memories are of sitting on the porch with Grandpa, carving together as he drank whiskey and I drank lemonade. I always enjoyed the time we spent together.

PASSAGE 5:

Jenna was getting ready to leave the house when she heard a huge commotion outside. Her dog, a whippet named Buster, was barking frantically. She ran to the door, but any intruders were nowhere to be seen. "What is it?" Jenna asked her pet. Buster only whimpered. Just then, Jenna caught a whiff of the dog. Immediately, she realized what had happened. Buster had been sprayed by a skunk. Jenna quickly put her plans on hold. Covering her nose, she wrapped Buster in a towel and placed him in the car, hoping she could have him cleaned at whichever vet or groomer was closest. Unfortunately, the highway was extremely crowded, and the trip stretched longer and longer. The smell in the car became overwhelming. Finally, they arrived at the vet's office. The receptionist warned her that the cleaning fees would be expensive, but Jenna was determined to get rid of the skunk spray whether or not she had to pay a high price.

PASSAGE 6:

The ocean waves crashed loudly on the rocks beneath the wharf. Salty droplets of water sprayed in the air. On the dock, cargo ships unloaded their crates from ports all along the coast. Some boats held timber or coal; others contained shipments of local crops. Where the smaller fishing boats sat, there was even more activity. One crew was unloading their daily catch of lobsters; another produced a load of oysters and whelks. Some of the fishermen were passing the time by trading tales about their adventures at sea. An old man with a long gray beard began to tell a group of young boys about the time he had been trapped in a whirlpool. Another sailor joined in with a story about a vicious storm somewhere in the Caribbean. A third claimed that a sea monster had once attacked his crew, making the boys laugh. They said that it must have been a whale or squid, but the fishermen insisted that real sailors saw monsters all the time. Whether or not the stories were true, both the boys and fishermen seemed to enjoy them.

APPENDIX D

STIMULI: READING TARGET WORDS

Table D.1 Properties of target words in reading passages

Passage	Wordform	Word Frequency		Part of Speech	Following Vowel	Preceding Segment	Phrasal Position
		Raw	***Zipf***				
1	Which	477.24	5.68	pronoun	/ɪ/	-----	initial
	wheel	27.06	4.43	noun	/i/	/ə/	final
	whiskers	2.33	3.37	noun	/ɪ/	/s/	final
	white	171.45	5.23	adjective	/aɪ/	/θ/	medial
	which	477.24	5.68	adjective	/ɪ/	/d/	medial
	Why	2248.76	6.35	adverb	/aɪ/	-----	initial
	whim	2.14	3.33	noun	/ɪ/	/ə/	final
2	where	1830.22	6.26	adverb	/ɛ/	-----	initial
	Whitney	3.25	3.51	noun	/ɪ/	/n/ or /d/	medial
	cartwheels	0.47	2.69	noun	/i/	/t/	final
	whined	1.63	3.22	verb	/aɪ/	/z/	medial
	While	349.43	5.54	conjunction	/aɪ/	-----	initial
	whacked	8.92	3.95	verb	/æ/	/d/	medial
	Whispering	8.10	3.91	verb	/ɪ/	-----	initial
	why	2248.76	6.35	adverb	/aɪ/	/oʊ/	medial
3	Meanwhile	15.92	4.20	adverb	/aɪ/	-----	initial
	wheat	5.75	3.76	noun	/i/	/l/	medial
	whisk	0.57	2.77	verb	/ɪ/	/n/ or /d/	medial
	whipped	13.16	4.12	adjective	/ɪ/	/n/ or /d/	medial
4	When	2034.10	6.31	conjunction	/ɛ/	-----	initial
	whittle	0.41	2.63	verb	/ɪ/	/d/	medial
	while	349.43	5.54	conjunction	/aɪ/	/ɹ/	medial
	whistles	15.45	4.19	noun	/ɪ/	/ɹ/	final
	whet	0.16	2.25	verb	/ɛ/	/ə/ or /u/	medial
	whenever	35.10	4.54	adverb	/ɛ/	/k/	medial

	whiskey	16.12	4.21	noun	/ɪ/	/k/	medial
5	when	2034.10	6.31	conjunction	/ɛ/	/s/	medial
	whippet	0.10	2.07	noun	/ɪ/	/ə/	medial
	nowhere	39.12	4.59	adverb	/ɛ/	/oʊ/	medial
5	What	9842.45	6.99	pronoun	/ʌ/	-----	initial
	whimpered	2.49	3.40	verb	/ɪ/	/i/	final
	whiff	2.49	3.40	noun	/ɪ/	/ə/	medial
	what	9842.45	6.99	pronoun	/ʌ/	/d/	medial
	whichever	3.20	3.51	pronoun	/ɪ/	/t/	medial
	overwhelming	4.92	3.69	adjective	/ɛ/	/ɪ/	final
	whether	67.14	4.83	conjunction	/ɛ/	/eɪ/	medial
6	wharf	1.27	3.11	noun	/ɔ/	/ə/	final
	Where	1830.22	6.26	adverb	/ɛ/	-----	initial
	whelks	0.04	1.77	noun	/ɛ/	/n/ or /d/	final
	whirlpool	0.61	2.80	noun	/i/	/ə/	final
	somewhere	111.53	5.05	adverb	/ɛ/	/m/	medial
	whale	11.25	4.05	noun	/eɪ/	/ə/	medial
	Whether	67.14	4.83	conjunction	/ɛ/	-----	initial

APPENDIX E

STIMULI: INTERVIEW QUESTIONS

The following list of questions was used as an outline for the interview portion of the study. Variation could possibly occur based on the answers given, and informants were asked to expand upon their answers where appropriate. Questions 1 and 6 always occurred first and last in the interview process respectively.

1) Where did you grow up?
    a) What was it like living there? What kind of experience did you have growing up?
2) What type of work did/do you do?
    a) How did you get involved with that occupation?
    b) What was your first job?
3) Describe your first car.
4) What was your first pet?
5) Describe the scene of the biggest snowstorm you can remember.
6) Tell about an interesting memory from when you were younger (ex. a fun vacation, a funny family story, a narrow escape)

APPENDIX F

DEMOGRAPHIC QUESTIONNAIRE

See following page

Participant # _____

DEMOGRAPHIC SURVEY

Gender _____

Year of birth _____

Place of birth (city, state) _____

Years lived in birth city _____ **Years in birth state** _____

Place of current residence (city, state) _____

Years lived in current city _____ **Years in current state** _____

If you have lived in any other locations for an extended period of time (2 or more years), list these below:

Place *Length of time*

_____

_____

_____

Which of the areas that you have lived in would you most strongly describe as "home"?

_____

What is the highest level of education you have completed? (check box)

- ☐ Less than high school diploma
- ☐ High school diploma or equivalent (eg. GED)
- ☐ Associate degree (e.g. AA, AS)
- ☐ Bachelor's degree (e.g. BA, BS)
- ☐ Master's degree (e.g. MA, MS, MEd)
- ☐ Professional degree (e.g. MD, DDS, DVM)
- ☐ Doctorate (e.g. PhD, EdD)

Places of education (city, state)

_____

_____

_____

What is your occupation? (If retired, list your previous occupation)

_____

Are you retired?	Yes	No
Are you currently a student?	Yes	No