Sociophonetics of Hmong American English in Minnesota

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Eden A. Kaiser

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Benjamin R. Munson & Jeanette K. Gundel, Advisers

September, 2011

Acknowledgements

For their assistance great and small, my sincerest thanks goes to:

My dissertation committee: Benjamin Munson, for tirelessly reading and commenting on draft after draft, and for his guidance through the entire graduate school process; Katie Drager, for moral and statistical support; Jeanette Gundel, for being an ally and friend in the complicated world of graduate school; Bartłomiej Plichta, for sharing his world of knowledge about nasalization and field recording techniques, and for creating the best thing on the internet—Akustyk; and Bruce Downing, for his invaluable advice for working with the Hmong community in the Twin Cities, and his connections within the community.

My mentors, colleagues, and friends at the University of Minnesota, for their academic, personal, and feline support: Paula Chesley, Jean-Philippe Marcotte, Serguei Pakhomov, Robert DelMas, Nancy Stenson, Sara Schmelzer Loss, Ellen Lucast, Kateryna Kent, Kaitlin Johnson, Kari Urberg Carlson, Sarah Schellinger, Sarah Mack, Muhammad Abdurrahman, Rebecca Anthony, Cathryn Merla-Watson, Jim Gielissen, and Paul Ching.

Past and present administrative staff of Linguistics and Speech-Language-Hearing Sciences: Barbara London, Cathy Dziuk, and Brianna Schmidt.

Mentors and colleagues at other universities: Charles Boberg, Rika Ito, Christina Esposito, Mike Flynn, Dennis Preston, Mary Beckman, Jan Edwards, the English faculty at University of Michigan-Flint, and Louis Epstein.

All the participants of my studies, for their time and sincere efforts.

People outside of academia, for helping me keep perspective on life: Emily Marvin, Frances Veit, Maggie Kaiser, Dimi DeRose, Tom Kaiser, Doug Dalglish, my running buddies and the folks at the Running Room, and everyone at One Yoga.

Hmong community organizations: Hmong American Partnership, especially Eric Bestrom; and the Hmong Cultural Center.

Katie Levin, Kristen Nichols-Besel, Zach Pierson, Mitch Ogden, and the attendees of the 2011 Dissertation Writing Retreat at U of M Center for Writing; especially Kimberly Thomas-Pollei, for being my hard-working and enthusiastic writing buddy.

In addition, financial support for this dissertation was gratefully received from the Institute of Linguistics and the McKnight foundation, through a McKnight Presidential Fellowship awarded to Benjamin Munson.

Thank you, all!

Abstract

This dissertation is a sociophonetic analysis of the English spoken by Hmong Americans living in the "Twin Cities" of Minneapolis and St. Paul, Minnesota. The Twin Cities has the largest urban population of Hmong Americans in the United States. Through studies of production and perception of vowels involved in sound changes, I investigate whether Hmong Americans—a relatively new ethnic group in the United States—have established any elements of an ethnic dialect of English that communicates an identity that is uniquely Hmong American.

Sound changes are particularly fruitful objects of sociophonetic study as they provide a spectrum of potential indexical variables for speakers exposed to those sound changes. I examine Hmong Americans' participation in three sound changes: the Northern Cities Shift, the low back merger, and fronting of the high back vowel (/u/ or GOOSE). Their degrees of participation in those sound changes are compared to age-matched European Americans from the same area.

It was expected that the inferred tight-knit nature of Hmong Americans' social networks would cause a slower uptake of current regional and supra-regional sound changes versus the comparatively looser networks of many European Americans in the Twin Cities. Furthermore, the target population should presumably experience some influence in their English from the Hmong language. Crucially for this study, the Hmong language has phonemic nasal vowels whereas English does not. This L2 influence of phonemic nasal vowels was hypothesized to emerge in Hmong Americans' English as less nasalization overall, and to decrease the likelihood that they will engage in the Northern Cities Shift.

The results of the production study show that European American speakers seem to be participating in one supra-regional sound change, the fronting of the GOOSE vowel, to a greater extent than in the past, and to a greater extent than Hmong Americans. Two other sound changes, the Northern Cities Shift (a regional change) and the low back merger (a supra-regional change), show inconclusive evidence of adoption by either EA speakers or HA speakers.

The perception study, which was conducted with a new set of participants, aimed to uncover whether phonetic differences between Hmong Americans' and European Americans' vowel pronunciations are actually detectable by others. Words recorded during fieldwork were rated on a visual analog scale by listeners on several different dimensions of speakers' social characteristics, including ethnicity. It was found that although certain expected phonetic differences were not used to make judgments of speakers' ethnicities, other phonetic differences, some expected and some not, did indeed predict listeners' judgments of speaker ethnicity. Listeners seemed to use either formant values or vowel nasalization (or sometimes both) to judge speaker ethnicity, depending on vowel class, listener ethnicity, and listener birthplace.

Taken together, the results of the two studies provide evidence that Hmong Americans' vowel pronunciations are not simply Hmong-influenced imitations of vowels as spoken by European Americans, and that listeners, especially other Hmong American listeners, can use these complex yet systematic phonetic patterns to make accurate decisions about speakers' ethnicities.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview

> "The work of an intellectual is not to mould the political will of others; it is, through the analyses that he does in his own field, to re-examine evidence and assumptions, to shake up habitual ways of working and thinking, to dissipate conventional familiarities, to re-evaluate rules and institutions and ... to participate in the formation of a political will (where he has a role as a citizen to play)" (Foucault, 1988, p. 128).

This dissertation follows a long line of studies that explore the relationship between language and society. The intention for the current work is to acknowledge and re-evaluate the evidence that comes before it, by adding linguistic research on a new social perspective—that of Hmong Americans. It is hoped that by augmenting our understanding of how and why societies manipulate language, we may make more informed social (and political) decisions regarding ourselves, our neighbors, and those whose practices and customs are unfamiliar to us.

As a whole, Hmong Americans have immigrated relatively recently to the United States. Some are still learning the complex grammar and pronunciation of the English language. Many Hmong Americans, however, speak English fluently or even natively. This study investigates the speech of those who claim Hmong ancestry and have lived in the United States from age eight or before, and who are fluent in English and Hmong.

Although this study looks for English pronunciation differences between Hmong Americans and European Americans, it is not "Hmong-accented English" that is of primary concern here. Rather, I explore the different, but equally "American," accents of English spoken by young fluent speakers of the language, somewhat like a comparison between the accents of the East Coast and the West Coast.[1]

Learning what differences exist between two varieties of a language is interesting, but knowing *why* these differences exist can actually tell us something about the people who use these varieties. By means of original data collection, speech perception experiments, and detailed linguistic analyses, this dissertation explores the pronunciation differences between Hmong Americans in Minnesota and European Americans from the same region, and provides insight into why these differences may exist. Critically, the influence of the Hmong language is likely not the primary determining factor here. Rather, it is a complex set of factors relating to issues of sound change, speaker identity, and listener perceptions.

1.2 Language and social identity

1.2.1 Sociophonetics

Within a single language, words can be pronounced in subtly different ways. Some of this phonetic variation is seemingly random, and some is simply a result of the surrounding linguistic context. But some phonetic variation may signal to a listener that the speaker can be categorized into one or more social groupings. *Sociophonetics* investigates those speech sounds, called sociophonetic *variables*, that are believed to hold social meaning for a community of speakers.

Speech communities are groups of interacting individuals with a common interest, background, or goal, such as a high school clique or a group of computer programmers. Most people belong to multiple speech communities at the same time. To a large extent, one's speech communities together make up one's *social network*—a key agent in the

[1] The reader should keep in mind that within the field of linguistics, it is widely accepted that no single way of speaking is better or worse than any other. Each variety holds value for the group that uses it, and each is valuable to linguists as a source of information about how Language (the human capacity for linguistic communication) functions. Most importantly, the field of linguistics holds that "if a native speaker says it, it must be right:" other than speech errors, each minor variation between accents is no less "correct" than any other, despite what your English teacher may say.

innovation and spread of possible sociophonetic variables. Isolated or tight-knit social networks, where speech communities interact frequently and in many contexts with one another, but infrequently with outside groups, tend to adopt regional or supra-regional sound changes more slowly than outside groups. Tight-knit social networks also tend to promote linguistic divergence[2] from surrounding speech communities over time. Members of a tight social network may introduce an innovative linguistic form—a new word or pronunciation—which spreads through the group quickly, given their frequent interactions with one another. If this innovative form comes to be arbitrarily associated with a specific social category, it becomes a sociophonetic variable.

These sociophonetic variables may be used, whether consciously or not, to shape a speaker's social identity—that is, how others perceive the speaker's social characteristics. For this social message to be transmitted to others, there must be a shared understanding of the specific social meaning of the phonetic variation. One's social identity is formed, maintained, and changed through mutual understanding of meaningful signs, both linguistic and non-linguistic. Linguistic variation that is understood to be useful in shaping favored social identities is said to be a form of "capital" in the "linguistic marketplace" (Bourdieu, 1977). Using sociophonetic variants that are linked with more desired social attributes brings the speaker more social capital. In other words, sounding like "the cool kids" can make you a little bit "cool," too.

Coolness takes many phonetic forms, depending on the speaker, the listener(s), and the social context. For instance, the most valued social attributes in an interview setting will be different from those in a casual setting with close friends. Speakers are able to instantly adapt their valuation of a given sociophonetic variable for the appropriate linguistic marketplace. Furthermore, the actual use of sociophonetic variation is not straightforward; sociophonetic variables are gradient rather than categorical, and are only used by some people in some groups some of the time, to a greater or lesser extent (Foulkes & Docherty, 2006). The most suitable method of capturing these interesting but elusive patterns of sociophonetic variation requires "research agendas that

[2] Linguistic divergence occurs when there is infrequent between-group interaction, preventing the spread of innovative forms beyond the core network. With enough linguistic divergence, linguists label the group's speech a different ethnic, regional, or social dialect, depending on the defining social characteristic of the speech community.

are pattern-driven, rather than variable-driven, investigating the multidimensional patterns of factors woven together by speakers to construct social identities" (Hay & Drager, 2007, p. 90).

This social-constructionist view of defining variables through observation of patterns is a recent development in the short history of sociophonetics. The study of sociophonetics dates to the early-1960s, with Labov's (1962) analysis of phonetic variation stratified by social class and attitude toward the local economy in Martha's Vineyard, Massachussetts. Sociophonetics follows from the older linguistic variationist tradition, which seeks to quantify any linguistic variation and change vis-à-vis three dimensions: time, space, and social groupings. Generally only one of these dimensions is investigated at any one time, while the other two dimensions are kept constant. Many sociophonetic studies still explore linguistic variation with respect to broad social categories such as gender, age, social class, and ethnicity. Eckert (2005) calls these "First Wave" variation studies (e.g., Labov, 1966) . "Second wave" studies define social categories more locally or even individually, and view variation as affiliation with others (e.g., Eckert, 1989a; Labov, 1962). Studies that view variation as resources for constructing one's social identity are categorized as part of the "Third wave" (e.g., Benor, 2002; Podesva, 2004). The progression of these three waves (not necessarily happening in chronological order) illustrates researchers' increasingly complex ideas of social categories that are important in language variation.

Traditionally, sociolinguists quantified phonetic variables based on unaided observation alone. However, with the recent spread of readily available acoustic measurement tools, more objective scientific measurements are being made. And because of these tools, researchers have begun to investigate finer details in speech, revealing numerous insights about the kinds of variation that are important for indexing social meaning (e.g., Munson, Jefferson, & McDonald, 2006; Stuart-Smith, 1999; Zhang, 2005). Many of these newly discovered sociophonetic variants are so subtle that speakers are almost certainly unaware they are using them.

Some conscious awareness of sociophonetics does exist. In fact, public dialogue about sociophonetic variation (e.g., "a Southern drawl," "talking like a Valley Girl," "the gay lisp," and so on) is not uncommon. However, our actual social assessments of others based on their speech is a great deal more nuanced than these phrases suggest.

In an instant, we can often place speakers into several accurate social categories based on their sociophonetic variation. The phonetic properties that we use to base our social assessments on can be extremely subtle, and often below the level of consciousness. For example, we have trouble explaining how exactly we know that someone is Southern or Northern, or black or white, but studies suggest that we can very accurately estimate speaker region and race (Baugh, 1996; Munro, Derwing, & Flege, 1999; Preston, 1993; Purnell, Idsardi, & Baugh, 1999; Wolfram, 2000; Wolfram, Hazen, & Schilling-Estes, 1999)

However subtle or obvious, the transmission of a sociophonetic message requires the cooperation of at least two people: a speaker and a listener. A speech act must be both produced successfully (by the speaker) and perceived successfully (by the listener) to be considered fully successful. We can only claim the significance of a certain sociophonetic variable if the social message sent by the speaker via use of that innovative variable is accurately detected by the listener. Therefore, a complete investigation of sociophonetic phenomena should include both evidence of production and evidence of perception (see Hay & Drager, 2007; Thomas, 2002). Traditionally, experimental studies explore only one side of a linguistic exchange at a time: either production or perception, but not both. However, the types of questions that sociophoneticians ask, regarding social meaning of phonetic variation, can best be answered by a methodology that combines measures of production with measures of perception, taking into account the roles of both the speaker and the listener in a successful speech act.

In addition, sociophonetics requires another two-pronged methodology: an emphasis on both social context and controlled experimentation (Labov, 2006, p. 501). This joint focus, utilizing both qualitative and quantitative techniques, allows the researcher to gain new insights into both phonetics and sociolinguistics. Hay and Drager (2007, p. 90) explain the benefits of this combined methodology for the two individual disciplines, saying "[p]honeticians are increasingly turning to social variation as something that may hold the key to breakthroughs in the understanding of speech production and perception. Sociolinguists, in turn, stand to benefit from an increasingly sophisticated view of how individuals and groups are able to imbue their speech with social meaning." The combination of sociolinguistic and phonetic investigations, production and perception data, and qualitative and quantitative methodologies complement each other well,

in ways that provide a more complete picture of the patterns of language and social variation within a community of speakers.

1.2.2 Ethnolinguistic variation and Asian Americans

> "Our relationship to the Far East and Pacific is shaped to a great degree by the facts of nineteenth-century colonialism, in which the U.S., young in comparative terms, followed the European model in the way that smaller nations were overcome and dominated politically, economically, and socially. We have a history of dealing with the Asian world as a warehouse of persons and goods available to suit our own purpose and fill our own needs, a practice justified by the supposition that those people are inherently weaker" (Lippi-Green, 1997, p. 227).

One aspect of a person's social identity is ethnicity. Ethnic identity, like other aspects of one's social identity, is not fixed, but rather it is emergent and symbolic, the result of a dynamic process of construction (Hein, 1994). Race, in contrast, is determined by genetics and remains immutable. There may be linguistic patterns that vary by race, but language plays no part in *creating* one's race, nor any other part of one's biological identity.

The salience of different aspects of one's social identity will vary from person to person, but one's sociophonetic choices can communicate to others how one wishes to be categorized, ethnically and otherwise. In the United States, the dominant variety of English spoken by educated white European Americans is considered the "standard" dialect. Consequently, any other variety is labeled "non-standard," including the varieties often spoken by minority ethnicities.[3]

There is no opting out of speaking an ethnic dialect; every available option will associate you with one (or more than one) ethnicity. However, for people who need more ethnic identity choices, for whatever reason, code-switching is a common and effective way to construct a social identity that includes two or more ethnic associations. Code-switching involves the strategic and systematic use of different languages or dialects, mixed together in small segments (word-by-word) or larger segments (phrase-by-phrase

[3] For linguists, the terms "standard" and "non-standard" are more a reflection of the reality of "social capital" (in Bourdieu's (1977) terms) than of any prescriptive linguistic norms.

or greater). This dissertation does not address code-switching, as the focus here is on the use of the English language only.

In addition to code-switching, another linguistic strategy for constructing an individualized ethnic identity is to create a new ethnic dialect using innovative sociophonetic forms. This study investigates whether Hmong Americans, a relatively new ethnic group in the United States, have established any elements of a new ethnic dialect of English that communicates an identity that is uniquely Hmong American.

Asian American ethnolinguistic research

To my knowledge, there is only one existing analysis of Hmong American English (see Ito, 2010), which investigates a limited set of English vowels.[4] Hmong Americans, even classified simply as members of the larger pan-ethnic grouping of Asian Americans, have received little attention in the linguistic realm. According to the 2005–2009 American Community Survey (ACS), there are over 13 million people who identify as Asian (or Pacific Islander) Americans, which is about four percent of the total U.S. population (ACS, 2009). Lo and Reyes (2009, p. 5) state that despite the large and growing Asian American population in the U.S., scholars have neglected the linguistic characteristics of this group compared to Asian-descended popoulations of Canada, the United Kingdom, and Australia (cf. Li, 2002; Nishimura, 1997; Rampton, 1995; Tuc, 2003; Wei, 1994).

Lack of research on Asian Americans, and minority ethnicities in general, has engendered misinformation about whether ethnic minorities' speech contains aspects of dialects other than their respective ethnic dialects. There has been an unspoken assumption among early dialectologists that being an ethnic minority causes all other aspects of one's identity to fade away. Regional dialects, for example, were presumed to be spoken only by European Americans.

Speech of non-majority ethnic groups has not been examined for regional characteristics to the extent that speech of European Americans has. There are some exceptions to this (e.g., Frazer, 1996; Jones, 2003; Roeder, 2010; Yaeger-Dror & Thomas, 2010), but such studies are especially rare for Asian American ethnic groups. In fact, to my knowledge there are only five other studies exploring Asian Americans' speech for evidence of regional phonetic characteristics (Chun, 2009; Hall-Lew, 2009; Ito, 2010; H. Lee, 2001;

[4] Ito's (2010) study will be reviewed in Section 3.2.1 of Chapter 3.

Wong, 2007).[5]

Common assumptions about minority linguistic behaviors

It has been an explicit assumption in early (and not so early) sociophonetics that minority ethnic groups do not participate in majority sound changes to the same degree as the majority, or at all:

> "All speakers who are socially defined as white, mainstream, or Euro-American, are involved in the [sound] changes to one degree or another. But for those children who are integral members of a sub-community that American society defines as non-white Black, Hispanic, or native American the result is quite different. No matter how frequently they are exposed to the local vernacular, the new speech patterns of regional sound change do not surface in their speech" (Labov, 2001, p. 506).

If we take Asian Americans to be members of "non-white sub-communities," Labov is stating above that Asian American ethnic groups resist sound change, tending instead toward a static, non-regional, and non-standard ethnic dialect. In contrast, other research has shown that there are minority ethnic communities that are engaged in regional sound changes to the same degree as majority white European American communities (Thomas, 2007). And in some cases, a minority ethnic group may even be leading a regional sound change (Hall-Lew, 2009). These kinds of insights have become possible only as the field has begun to regard people of minority ethnicities as skillful users of language who utilize linguistic variation to construct a complex and nuanced social identity, worthy of careful investigation.

Another assumption about ethnic minorities that researchers have made concerns the speech of recent immigrants, including that of newly arrived Asian Americans. These

[5] Still, these studies (and my own) make an underlying assumption that a "legitimate" regional dialect is spoken by the European American majority and if other ethnic dialects of the same area contain regional features, they are simply imitations of the regional standard dialect. To counter this assumption, one would need to collect systematic long-term data from at least one non-white ethnic group from numerous dialect regions across the United States. As far as I know, this has not been done. For the present, I will continue to make this ethnocentric assumption, with the understanding that comparing a minority dialect to a local majority dialect is only one way of discovering whether the minority dialect is adopting regional features.

researchers have implicitly assumed for many years that as immigrant minority groups learn English, they are attempting to acquire a static, non-regional, mainstream variety, perhaps in an attempt to assimilate fully to a stereotypical middle-class European American lifestyle (Lo & Reyes, 2009) (e.g., Chi, 1999; Knoebel, 1986; P. Thao, 1999; Whitenack & Kikunaga, 1999). Often, the term "interlanguage" is used to describe anything similar to but not exactly the same as the non-regional mainstream variety, connoting the mainstream as the ultimate goal in second language acquisition (e.g., Huebner, 1983).

Lastly, there are two contradictory assumptions about Asian Americans in particular: the "forever foreigner" and "honorary white" stereotypes. Presumably, it would follow from the "forever foreigner" assumption that Asian Americans, no matter how native their English, never acquire a "legitimate" English dialect, whether static or changing, regional or non-regional, mainstream or vernacular. Possibly because of this assumption, much of the research regarding Asian American linguistic practices has centered on bilingualism, heritage language maintenance, code-switching, foreign-accented English, and accent perception (Lo & Reyes, 2009). The opposing but equally pervasive "honorary white" stereotype exhibits the assumption that all Asian Americans eventually "overcome" their ethnic minority status, or hope to do so, by assimilating to the American melting pot of middle-class European American linguistic and cultural norms (Lo & Reyes, 2009). Regarding the melting pot metaphor, Fadiman (1997, p. 183) says that "...the Hmong are what sociologists call 'involuntary migrants.' It is well known that involuntary migrants, no matter what pot they are thrown into, tend not to melt." Furthermore, involuntary migrants are often poor, which increases the likelihood that they will encounter racism in communities where they settle, and decreases the likelihood of quick and total assimilation.

The study of ethnolinguistic variation is built on many assumptions, and also on fighting against those assumptions, but there is not a great deal of research yet. Ethnic minorities, especially Asian Americans, are often assumed to be culturally and linguistically one-sided: either "fresh off the boat" ("FOB"), with a strong foreign accent, or "white-washed," with a standard European American dialect. The reality is, naturally, much more complex. Just as individuals' social identities differ from one person to the next, the ways of using language to convey those identities are equally numerous.

1.3 Hmong in the United States

1.3.1 Migration history

> "In the case of the Hmong refugees, escape from Laos or Vietnam was very traumatic and even quite dangerous. Long stays in refugee camps, predominantly in Thailand, were marked by severe and over-crowded conditions, deprivation of food and sanitary drinking water, as well as basic necessities for survival. The time spent in these refugee camps were a significant and meaningful part of the Hmong pre-migration experience" (Faruque, 2002, p. 2).

A growing ethnic group in the United States consists of the Hmong refugees from Laos and their descendants. The Hmong people have a very long history of being migrant minorities, with the earliest known records placing Hmong in China as far back as 2700 B.C. The Hmong in China were often pushed out of their villages by the expanding Chinese dynasty; some eventually migrated to Laos, Vietnam, and Thailand in the early nineteenth century (Hamilton-Merritt, 1993).

In the 1960s, the Hmong who settled in Laos found themselves involved in a "secret war" organized by the United States C.I.A. Many Hmong living in the remote mountainous areas near the Vietnam border were recruited to fight in Laos and block the passage of communist troops and supplies from North Vietnam into the south (Koltyk, 1998). Because of this war that most Americans had no knowledge of, the Hmong suffered extraordinary losses and were unable to return to their previous lifestyles.

After communist forces gained control of Laos in 1975, Hmong who had been allied with the United States, and even some who had not, were brutally hunted and tortured or shot by the new Lao People's Democratic Republic government (Hamilton-Merritt, 1993). According to Yang Dao, Hmong and former government official, the Lao government stated explicitly: "We must eradicate the [Hmong] minority completely" (D. Yang, 1982, p. 13). Fearing for their lives or livelihood, thousands of Hmong attempted to escape the country. Most who survived the escape attempt ended up in refugee camps in Thailand, across the Mekong River from Laos.

Before the Vietnam war, most Hmong lived in small villages of 25 to 30 houses and practiced swidden agriculture in the mountainous regions of Laos. They frequently

migrated to new locations for better farmland or to escape disease (C. Thao, 1982). In the refugee camps in Thailand, however, Hmong people had no choice but to live, work, and play in very close confines with other people, and not necessarily the relatives they were used to living with before. The refugee camps contained many more people than most had ever interacted with before, and their normal migrant agricultural lifestyle was now impossible (C. Thao, 1982). Families had been separated during their escape from Laos, and only sometimes were reunited in the refugee camps. The camps were often physically dangerous, disease-ridden, and mentally distressing (Hamilton-Merritt, 1993). The miserable living conditions induced many of the Hmong in the refugee camps to apply for safe passage to the United States, France, Canada, Australia, Argentina, and elsewhere. By 1983, over 50,000 Hmong were resettled in the United States; the Hmong resettled in the other countries totaled less than 10,000 at that point in time (Olney, 1986).

The first wave of Hmong refugees arrived in the United States between 1975 and 1983. Eventually, many Hmong families moved from their first, government-assigned locations to areas where there was a concentration of other Hmong people, mostly in California, Minnesota, and Wisconsin. When the Thai refugee camps officially closed in 1999, those living in the camps were ordered to return to Laos. Some did return to Laos, but others stayed in Thailand. Over 10,000 Hmong eventually sought refuge in a Buddhist temple in Central Thailand called Wat Tham Krabok, living in constant fear of deportation by the Thai government (Lor, 2009). In 2005 and 2006, the U.S. initiated the final resettlement of Hmong from Wat Tham Krabok (G. Y. Lee & Tapp, 2010), so some Hmong have arrived in the U.S. as few as five years ago.

In 2009, there were over 205,000 people who claimed Hmong ethnicity living in the United States, and 50,467 (about 25%) of those were in the metro area of St. Paul and Minneapolis, Minnesota ("The Twin Cities"), more than any other metropolitan area in the U.S. (ACS, 2009). For the 2010–2011 school year, children of Hmong ethnicity made up over 23% of students enrolled in St. Paul public schools (MN Dept. of Education).

1.3.2 Sociocultural background

> "The Hmong identity is determined today not only by the Hmong but also by other people, by factors outside their control, outside their culture and community" (G. Y. Lee, 2005, p. 12).

As with other recent immigrant groups, the Hmong have to negotiate the formation of an identity that draws from two different cultures: one of the old homeland left behind, and one of the new region. And in the case of the Hmong in the United States, the two cultures are extremely different.

The following cultural components are commonly said to characterize the Hmong in Laos, according to G. Y. Lee (1986, p. 55):

- shifting agriculture
- language with mutually intelligible dialects
- a strong belief in ancestor worship and animism[6]
- a division of labor according to family membership and sex
- a social structure based on kinship ties through patrilineage and clan systems
- a patrivirilocal pattern of residence
- a history of migration from southern China
- a long tradition of being stateless[7]

These characteristics have necessarily had to change somewhat as refugees adapted to their new lives in the U.S. and other countries, but clan interdependence, ancestor worship and animism, and key roles of male elders continue to define what "being Hmong" means to most Hmong Americans (G. Y. Lee, 1986; Trueba, Jacobs, & Kirton, 1990). Additionally, the ability to speak the Hmong language contributes to a sense of Hmong identity, not least because it allows them to communicate more effectively with

[6] Some Hmong in Laos were Christian, having been converted by missionaries.

[7] Although the Hmong in Laos were citizens of that country, it should be mentioned that many Hmong in Laos lived in the "highland" areas, having little to no contact with the ethnic Lao.

their parents and older family members (Ly, 1999). In fact, knowledge of the Hmong language may be the most common way to maintain an identity that is at least partly Hmong (Bosher, 1997). Speaking Hmong is an important part of the beauty pageants held at Hmong New Year celebrations in Minnesota (Vang, 2008).

Maintaining an identity that straddles two cultures can be rather complicated. For example, many Hmong in the United States have converted to Christianity for a variety of reasons, but this conversion is seen by some as a rejection of a true Hmong identity, which has been closely tied to ancestor worship and animism. The following quote illustrates some of the strategies and challenges of negotiating religion for a bi-cultural identity:

> "There's a lot of tension between Christian and non-Christian Hmong. There's a feeling that because you've converted to Christianity you're not as Hmong. I'm not sure how it is in American churches nowadays, but in Hmong churches we always talk about spirits. And spirits are very real to Hmong people—good spirits and bad spirits, people who've passed away, ancestors. At church, we say that a long time ago we used to believe in other spirits, but now God is our spirit. So there is a connection to the past, except you can't have both spirits because they're not alike. You have to choose one or the other, and this causes tension" (M. N. Moua, 1999, p. 43).

In this quote, Moua describes the compromise that Christian Hmong have come to regarding the pervasive notion of spirits in Hmong culture, and the tension that that compromise causes within the Hmong American community.

In addition to religion, many Hmong must choose between old and new customs regarding marriage and gender roles. According to folk custom, a person cannot marry someone of the same clan. As clans are identified by their surname, it is very rare to find someone married to another person with the same surname (G. Y. Lee & Tapp, 2010). On the whole, this is still being practiced in the U.S. today (D. Yang, 2009), although there are only about 20 different clan names among U.S. Hmong. The practice of polygamy, however, while fairly common in Hmong communities in Asia, is unusual in the West (G. Y. Lee & Tapp, 2010). There is also a custom of a "bride price," paid by the groom's parents to the bride's parents, and in the U.S. This custom, still

often practiced, also symbolizes a social commitment of the groom to the welfare of the bride and her family (Her, 2009). Another payment made at a marriage, the "seed money," is given to the bride and groom by the bride's parents. Part of this payment is understood to be used to buy a sacrificial animal for the bride's eventual funeral, and is still exchanged today among Hmong Americans in the U.S. (Her, 2009).

Traditionally, the Hmong social organization is patriarchal. While men are usually seen as the heads of their households, women are responsible for cooking, childcare, and passing down Hmong traditions to their children (Faruque, 2002). Elder males have usually been the clan leaders, and their opinions are sought out when important decisions need to be made. In the U.S., however, these traditional roles are frequently inverted, when women are the primary income-earners for the household, or when elders' poor English language skills do not allow them to take a leadership role in the community (Vang, 2008).

Hmong cultural values include an emphasis on group interests over individual interests, strong familial ties, and the clan system, which seems to have withstood the pressure to assimilate to a more American individual-centered focus. This high regard for family has contributed to strong, tight-knit Hmong communities. On the other hand, young Hmong people in the U.S. frequently go against Hmong cultural norms, putting marriages and child-bearing on hold while they pursue higher education and careers (Bosher, 1997).

Misunderstandings and prejudices have surfaced as a result of the vast cultural differences between the highly visible Hmong communities and their neighbors. It has been about 35 years since the first Hmong refugees settled in the United States, and the social and cultural beliefs and practices of the Hmong have been adapted to serve the needs of the community. Young Hmong people generally negotiate, often through conscious choice, a "middle ground between the two cultures that combine[s] elements from both" (Bosher, 1997, p. 599). But the perceived cultural differences have not disappeared, and neither have the prejudices against the Hmong (Hein, 1994; Koltyk, 1998). For example, a Hmong American man living in Wisconsin, Chai Soua Vang, was tried and convicted in 2005 of killing six white hunters in Wisconsin in 2004. After this, some people in Wisconsin put bumper stickers on their cars that read, "Save a deer, kill a Hmong" (K. Yang, 2009).

1.3.3 The Hmong language

> "I definitely want the Hmong culture to be a part of my life, and well as my children's when I have them. Learning the Hmong language is important" (S. Moua, 1999, p. 185).

The Hmong language is believed to be related to the languages of the Yao and She peoples in China, but to very few other existing languages. It is classified by some as a Sino-Tibetan language, by others as a Tai-Kadai language, and by still others as a language isolate (G. Y. Lee & Tapp, 2010). There are several mutually intelligible dialects of Hmong spoken in the U.S., the most prevalent of which are called White Hmong and Green (or Blue) Hmong.[8] The phoneme inventories vary slightly between these two dialects.

Words in Hmong are generally only one syllable, with obligatory onsets (except for several particles) and a prohibition against codas (except for a weak velar nasal [ŋ] after nasalized vowels and a weak glottal stop [ʔ] at the end of a low-falling creaky-toned syllable). There are seven "tones," including high, mid, low, mid-rising, high-falling, low-falling (creaky), and mid-low (breathy) (Ratliff, 1992; Whitelock, 1982).[9]

The Hmong language contains phonemic nasal vowels. Depending on the dialect, there are either two /ɛ̃, ɔ̃/ nasal vowels (for White Hmong) or three /ɛ̃, ã, ɔ̃/ nasal vowels (for Green Mong),[10] and there are six oral vowels /i, e, a, o, u, ɨ/.[11] There are also five or six diphthongs, depending on the dialect. (Ratliff, 1992)

In 2000, López reported that, according to the U.S. Bureau of the Census, 59.8% of all Hmong lived in linguistically isolated households, interacting with speakers of other languages very infrequently, if at all. It is not surprising, then, that 97.4% of all Hmong reported the ability to speak the Hmong language (López, 2000). Although

[8] The language is also spelled *Mong* for the Blue/Green dialect.

[9] These tones are indicated in the standard Romanized Popular Alphabet orthography with syllable-final letters: *b* (for high tone), *s* (for low tone), *v* (for mid-rising tone), *j* (for high-falling tone), *m* (for low-falling creaky tone), *g* (for mid-low breathy tone), and nothing for the mid tone. For example, the word for "spirit" is spelled *dab* and pronounced [da] with a high tone.

[10] Generally, "Hmong" refers to either the White dialect of Hmong alone or both White and Green dialects, and "Mong" refers to the Green dialect alone.

[11] Phonemic nasalization on vowels is indicated orthographically by doubling the vowel. For example, the word "Hmong" is spelled *Hmoob* in the White dialect, and is pronounced [m̰ɔ̃] or [m̰ɔ̃ŋ] with a high tone.

many bilingual young people go through a phase of feeling embarrassed about their heritage language, often refusing to speak it in public for fear of ridicule, most will eventually embrace their heritage language fluency, and even regret the lost years' of practice in it (Hinton, 2009; Shankar, 2011).

In recent years, young Hmong people have often attended Hmong language classes at churches, community centers, or universities. There is a general fear that young people are "no longer fluent in their mother tongue" (G. Y. Lee, 2005), even though their parents and grandparents may use Hmong exclusively at home. In fact, in a study of language patterns of Hmong Americans in Minnesota, Ito (2010) found that 54% of her second-generation participants were English-dominant, which reflects the common trend of second and third generation immigrants slowly becoming dominant in the majority language, and at the same time losing fluency in their heritage language (see Alba, Logan, Lutz, & Stults, 2002).

1.4 Linguistic assumptions and background

1.4.1 Wells' lexical set

Throughout this paper, I will be using Wells' (1982) lexical set to refer to categories of American English vowels that can be grouped under single phonemes. This alternative to the International Phonetic Alphabet (IPA) has been used frequently by phoneticians outside of the United States, especially in British-influenced traditions. This method is particularly appropriate for use in studies that deal with more than one dialectal realization of the same phonemic unit.

Tables 1.1 and 1.2 display the lexical set items that will be used in this paper to refer to American English phonemes. The starred items at the end of Table 1.1 are my own additions, used to distinguish subsets of the TRAP vowel: those that come before nasal consonants (STAND) and those that come before oral (non-nasal) consonants (CAB).

Table 1.1: Wells' lexical set: Front vowels

IPA	i	ɪ	e	ɛ	æ	æ(N)	æ(O)
Wells, 1982	FLEECE	KIT	FACE	DRESS	TRAP	STAND*	CAB*

Table 1.2: Wells' lexical set: Back vowels

IPA	u	ʊ	o	ʌ	ɔ	ɑ
Wells, 1982	GOOSE	FOOT	GOAT	STRUT	THOUGHT	LOT

1.4.2 How vowels are quantified

A great deal of sociophonetic variation is found in vowels,[12] but vowels are particularly difficult to characterize partly because, unlike consonants, no part of the tongue or mouth makes contact with anything else in the vocal tract, and the possible placements of the tongue within this space are almost limitless. Among all the advanced scientific tools that sociophoneticians use, the most important one for describing vowel variation is a two-dimensional representation of a mid-sagittal view of the mouth, called a "vowel quadrilateral" (shown in Figures 1.1 and 1.2). Placement of vowels on this quadrilateral is based on the typical location of the blade of the tongue when each vowel is produced. The first dimension of close (or high) and open (or low) represents the relative height of the tongue in the mouth; the second represents the relative position in the front/back dimension. For example, the vowel in the word FLEECE (IPA: /i/) is produced with the tongue very high and front in the mouth, relative to the other vowels, so it appears in the upper left corner of the quadrilateral.

Vowels' positions in the quadrilateral are quantified using acoustic measures of resonant frequencies of the vocal tract. The most important resonant frequencies for vowels are the first (called F1 or first formant) and the second (called F2 or second formant). The first formant varies inversely with vowel height, so high (close) vowels have low F1 values and low (open) vowels have high F1 values. The second formant varies inversely with vowel backness, so back vowels have low F2 values and front vowels have high F2 values. Ideally, each vowel can be uniquely identified by its F1 and F2 measures. The vowel in the word FLEECE, for example, will have a very low F1 and high F2, compared to the vowel in the word LOT (IPA: /ɑ/), which will have a very high F1 and low F2.

[12] In addition to tongue placement, there are even more sources of vowel variation: lip rounding, breathiness, nasalization, and more. This paper will address nasalization in Section 1.5.2 of Chapter 1, and in later chapters.

Figure 1.1: The vowel quadrilateral (Image: www.uni-bielefeld.de)

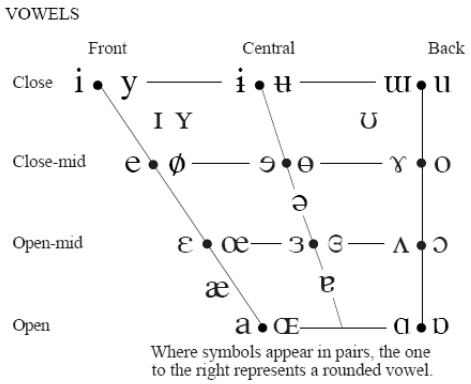
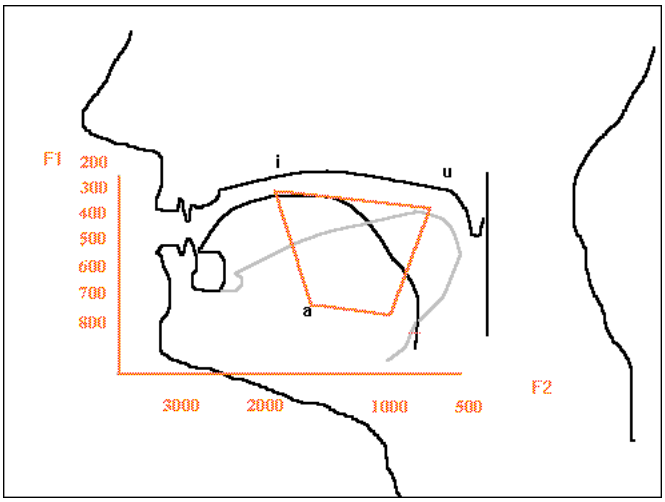


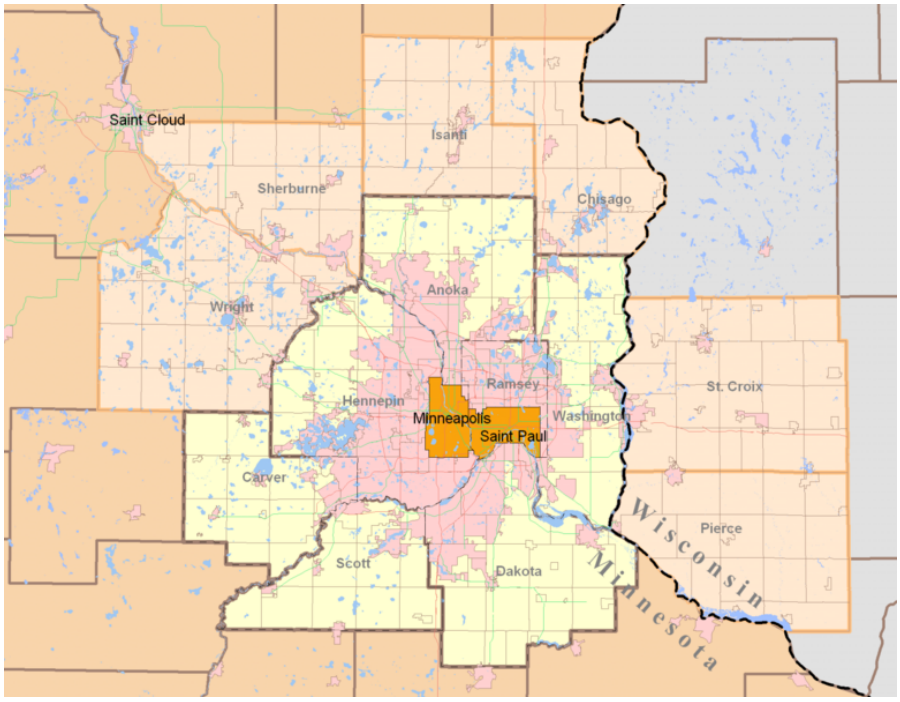
Figure 1.2: The basic vowel quadrilateral superimposed on a human pharynx (Image: www.uni-bielefeld.de)



1.5 Sociophonetic context in the Twin Cities, Minnesota

The Twin Cities is made up of St. Paul, which is the capitol of Minnesota; Minneapolis, which is the largest city in Minnesota; and the surrounding suburbs. The official metropolitan area of the Twin Cities consists of 13 counties, two of which are actually located in Wisconsin, as shown in Figure 1.3. The population of the metro Twin Cities area is over three million people, which makes it the 16th-largest metropolitan area in the United States.

Figure 1.3: The 13-county metro area of the Twin Cites (Image: wikipedia.org)



As the largest metropolitan area in Minnesota, the Twin Cities has significant linguistic influence over the rest of the state, which has been characterized in the past as having monophthongal GOOSE, GOAT, and FACE vowels; raising and offgliding of the TRAP vowel; the merging of the LOT and THOUGHT vowels; Canadian Raising; and some fronting of the GOOSE vowel (Allen, 1976; Ash, 1996; Ito, 2010; Labov, Ash, & Boberg,

2006; Thomas, 2001).[13] The monophthongal vowels are more typical of residents living in Northern Minnesota and rural areas of the southern half of the state, where there is still a stronger linguistic influence of the early European settlers from Germany and Scandinavia (Allen, 1976; Ash, 1996). However, the folk linguistic perception is that all Minnesotans speak like this to a greater or lesser extent, even in the urban areas of Minneapolis and St. Paul. For instance, take note of the following two excerpts from an archived internet forum:

> "Minnesota-ans [*sic*] have much stronger 'Fargo'ish accents than Wisconsinites. You can tell them apart, oftentimes, at UW-Madison, because they have such a thick accent" (Badger77, 2006).
>
> "...There is a SE Wisco accent that is different than the northern Wisco/Minny accent..." (Crankbaiter, 2006)

Although differences between Southeastern Wisconsin and the rest of Wisconsin are often noted among folk linguistic observations, Minnesota is more likely to be perceived as a homogeneous dialect area. In fact, for many outside the region, all midwestern speech is viewed as uniform (Frazer, 1993), as it is portrayed in the 1990 film "Fargo."

In 2009, 87.7% of Minnesotans were European American. In all but a handful of neighborhoods in the Twin Cities, European Americans were the majority ethnicity all across the state (ACS, 2009).[14] For the purposes of this project, we consider white European Americans to be the "mainstream" ethnicity. The use of this term does not imply a value judgment on any ethnicity; rather it is meant only to indicate the sociocultural majority in the region.

1.5.1 The low back merger

A unique opportunity for sociophoneticians comes along when the spread of a new pronunciation can be observed in real time. These changes in progress can inform the

[13] "Monophthongal" means that the vowel is long and unchanging in either the first or second formant. "Raising and offgliding" means that the vowel's first formant decreases (so the vowel raises) and it centralizes at the end of its pronunciation. "Canadian Raising" is the centralization of the first half of the diphthong /ai/ as in "nice," and sometimes /au/ as in "house" (Thomas, 2001).

[14] Twin Cities neighborhoods with European Americans as a minority ethnicity include areas of Brooklyn Center, Robbinsdale, Minneapolis, and St. Paul. In several of these neighborhoods, the Asian population constitutes the plurality ethnicity, although never the majority ethnicity.

ways linguists understand how sound changes spread: who spreads them, how quickly, in what contexts (both linguistic and social), and the motivation behind the spread of the changes.

The low back merger is one such opportunity, which sociophoneticians have studied extensively—possibly more than any other sound change in U.S. English. The low back merger is the loss of a phonemic distinction between the LOT and THOUGHT vowels so that they are generally indistinguishable in perception. This merger is a result of the LOT vowel moving to the THOUGHT vowel, vice versa, or both vowels moving towards one another.

This vowel shift is quite advanced throughout the U.S., occurring in both production and perception in many cities all over the country (Gordon, 2006; Hall-Lew, 2009; Labov et al., 2006),[15] with the Western states (and Canada) showing the greatest advancement in the sound change (Labov et al., 2006). According to the Atlas of North American English (ANAE; Labov et al., 2006), Minnesota is split on advancement of the change: Figure 1.4 illustrates that most of the ANAE speakers in Northern Minnesota produce merged low back vowels in all environments, whereas only two out of the four speakers from the Twin Cities do.[16]

In the Twin Cities, the merger has been shown to exist for many years. In fact, Allen, in his Linguistic Atlas of the Midwest (1973, 1976), says that he noticed the low back merger in University of Minnesota students' speech beginning in the 1940s and continuing through the 1970s. However, the ANAE noted that the low back merger seems to be receding in parts of the Midwest, including Southern Minnesota (Labov et al., 2006, p. 65).

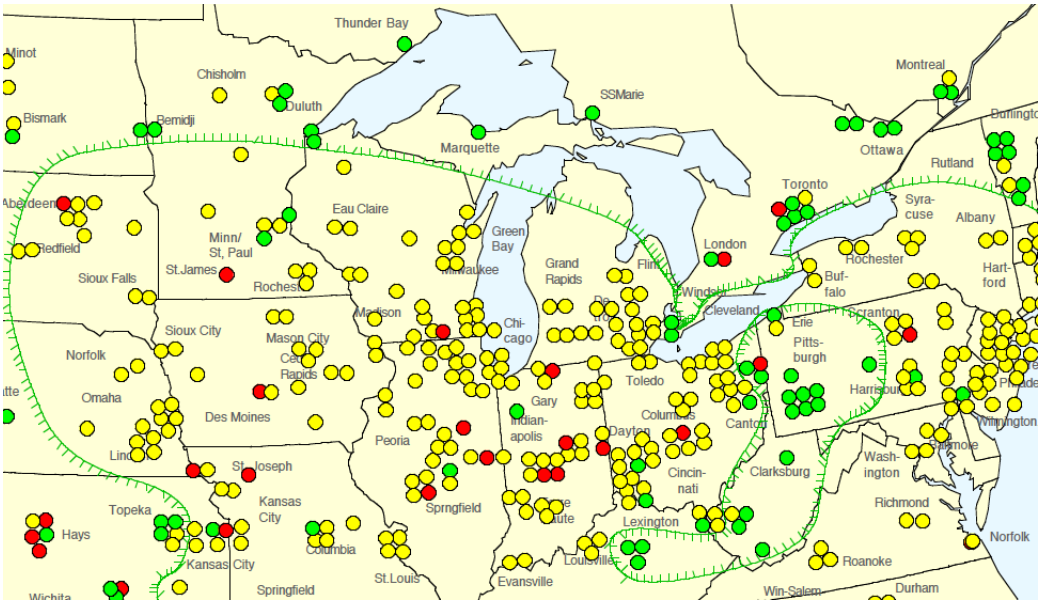
Another major sound change in the region, the Northern Cities Shift (NCS), also involves the LOT and THOUGHT vowels, among others, but the LOT vowel moves in somewhat opposing directions for the two shifts,[17] so it is unlikely that the LOT and THOUGHT vowels are involved in *both* the NCS and the low back merger at the same

[15] The low back merger usually occurs in perception before production (Labov et al., 2006).

[16] Figure 1.4 relies on color-coding to represent dialect differences visually. For readers viewing this in black and white, the dots shaded similarly to the dot under "St. Paul" are red, the dots shaded similarly to the two dots under "Eau Claire" are yellow, and the dots shaded similarly to the dot under "Thunder Bay" are green. See original PDF document for colors.

[17] In the low back merger, the LOT vowel usually moves back (and up), whereas in the NCS, the LOT vowel moves forward.

Figure 1.4: Low back vowels (LOT and THOUGHT) merged in production and perception (Image: Labov et al., 2006, p. 60): (Green = merged in all environments; red = merged before nasals only; yellow = not merged)

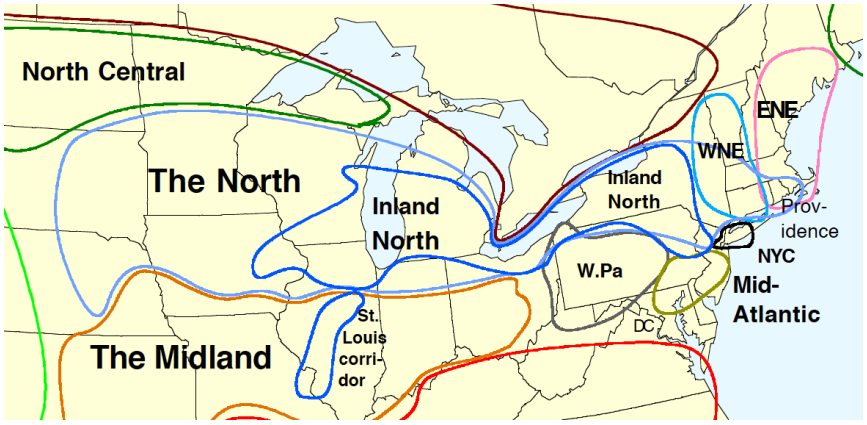


time. In fact, the ANAE calls the NCS a resistance strategy against the low back merger (Labov et al., 2006, p. 59).

1.5.2 The Northern Cities Shift

Minnesota lies just outside the dialect region known as the Inland North (see figure 1.5), which is defined by participation in the Northern Cities Shift. The NCS, first described by Labov, Yaeger, and Steiner (1972), is a sound change in progress centered around major U.S. cities near the Great Lakes: Syracuse, Rochester, Buffalo, Cleveland, Detroit, and Chicago, and extending east into New York state and northwest into Milwaukee and Madison. Labov et al. (2006) claim that the NCS is in its earliest stages in the Twin Cities.

Figure 1.5: The "Inland North" and surrounding dialect regions (Image: Labov et al., 2006, p. 148)



There are six individual vowel shifts recognized as part of the NCS, as follows:

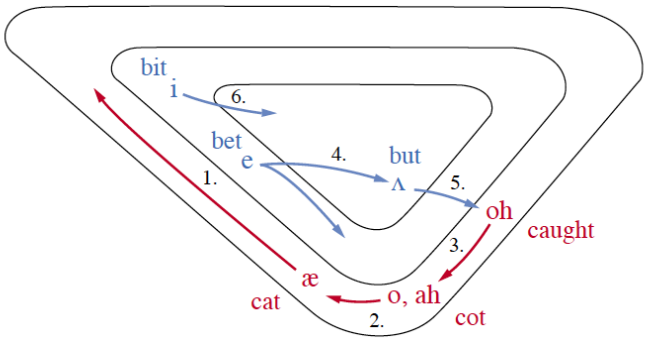
1. Raising of [æ] (the TRAP vowel)
2. Fronting of [ɑ] (the LOT vowel)
3. Centralization of [ɔ] (the THOUGHT vowel)
4. Lowering of [ɛ] (the DRESS vowel)
5. Backing of [ʌ] (the STRUT vowel)
6. Lowering of [ɪ] (the KIT vowel)

Usually, these steps proceed in order (as illustrated in figure 1.6), with raising of the TRAP vowel being observed first, followed by fronting of the LOT vowel, and so on.[18] There is no absolute consensus on the order of the last three steps.

The NCS has been shown to spread following the hierarchical pattern of Trudgill's (1974) gravity model of diffusion. This model predicts that sound changes spread from large city to nearby large city, and then to nearby smaller cities and more distant large

[18] In the St. Louis Corridor, however, the NCS has been shown to spread out of order. (Labov, 2007)

Figure 1.6: The Northern Cities Shift, according to Labov et al. (2006, p. 190) (Refer to vowel keywords used here instead of vowel symbols)



cities, although not all social networks in each city will adopt the sound change to the same extent. Most research shows that minority ethnicities are not adopting the NCS as fully as European Americans (or at all) (Frazer, 1996; Gordon, 2000b; Jones, 2003; Roeder, 2010; Samant, 2010; Thomas, 2001). In addition, Canadians, even in cities very close to the Great Lakes, are not adopting the NCS, most likely because of a "border effect," where sound changes do not spread readily over a country border regardless of proximity, which complicates the predictions of the diffusion model (Boberg, 2000).

Gender is another demographic factor that is believed to influence who adopts a sound change, with women being more likely to adopt a sound change compared to men (Labov, 1990, 2001). Women are especially likely to participate more fully when the sound change is "supralocal" (Milroy, Milroy, & Hartley, 1994). In other words, women tend to prefer the linguistic forms that bring them the most amount of linguistic capital in the widest context. For the NCS, as expected, women have usually been found to display the fullest adoption of the NCS before men do (Eckert, 1989b; Herndobler, 1977; Gordon, 1997).

Nasalization

The degree of vowel nasalization in non-nasal contexts has been shown to correlate with participation in the Northern Cities Shift (Plichta, 2005). Nasalization of speech sounds occurs as a result of the oral cavity (used for most speech sounds) coupling with the

nasal cavity. This coupling occurs as a result of the opening of the velopharyngeal port due to the lowering of the velum (at the back of the roof of the mouth). The velum must be lowered for air to flow out of the nasal cavity, creating the articulatory environment necessary for nasal consonant phonemes (which include /m/, /n/, and /ŋ/ in English). But these are not the only nasalized segments: it is a universal tendency to partially nasalize a vowel directly preceding a nasal consonant in anticipatory articulation, as in the vowel of a word such as *bent* but not *bet*. This ability to nasalize non-phonemically can then be extended to vowels before non-nasal consonants (as in *bet*), as is done by some speakers of the NCS (Plichta, 2005).

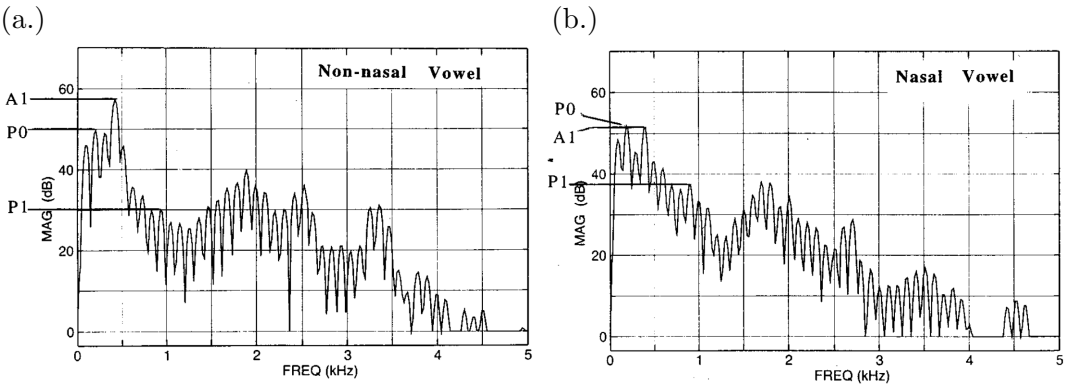
Some languages such as French also have nasal vowel phonemes (e.g., /ɑ̃/, /ɛ̃/, /õ/, etc), found in words like *monsieur*, *faim*, and *bon*. For languages with phonemic nasal vowels, it has been shown that phonemically oral (non-nasal) vowels are significantly less nasalized, in order to preserve the contrast between phonemic vowel nasalization and nasalization of a vowel due to articulation effects (Chen, 1997). Whereas, for languages with no phonemic nasal vowels, like English, vowel nasalization is theoretically more "available" for sociolinguistic variation, because it does not hold phonemic meaning for speakers.

There is evidence suggesting that the perceived vowel shifts in the NCS may be due in part to increased nasalization shifting listeners' percepts (Beddor & Hawkins, 1990; Chen, 1995, 1997; Kaiser & Drager, 2009; Kaiser & Plichta, 2009). According to Beddor and Hawkins (1990), when nasalization is increased digitally, the perception of high and mid vowels is lowered and that of low vowels is raised. This perceptual lowering of high and mid vowels corresponds to steps 3, 4, and 6 of the NCS, and the perceptual raising of low vowels (especially low front vowels) corresponds to step 1 (see Figure 1.6). This is not to say that the formants of these vowels are not shifting at all; in fact, it is possible that one speaker's increased nasalization is misperceived as vowel formant shifts by a listener, since nasalization can be such a subtle acoustic cue. That listener may then reproduce what they thought they heard as shifted formants (with or without accompanying nasalization). In this way the sound change could spread via misperception of vowel nasalization. One influential theory posits that these types of misperceptions are often the cause of sound changes (Ohala, 1981).

Perceived nasalization in speech is most likely related to the degree of velopharyngeal

opening which can only be measured directly using very invasive techniques. However, there are at least four documented non-invasive methods of measuring perceived nasalization: (1) nasal pressure (Shelton, Knox, Arndt, & Elbert, 1967; Weiss, 1954), (2) nasal vibration (Stevens, Kalikow, & Willemain, 1975; Stevens, Nickerson, Boothroyd, & Rollins, 1976), (3) nasal flow (Horii, 1980; Quigley, Shiere, Webster, & Cobb, 1964; Rothenberg, 1977), and (4) spectral characteristics (Beddor and Hawkins (1990), Chen (1997), among others). Of these four methods, the fourth is the easiest to implement, the most reliable across vowels and phonological contexts, and the most compatible with current acoustic phonetic theory (see Stevens, 2000). Chen (1997) describes two measurements based on spectral characteristics of nasal vowels called *A1–P0* and *A1–P1*. *A1* is the amplitude of the first (oral) formant, and *P0* and *P1* are the amplitudes of the first and second nasal peaks, respectively. See illustrations of these values in Figure 1.7 below.

Figure 1.7: *A1*, *P0*, and *P1* spectral measurements of (a.) a non-nasal vowel and (b.) a nasal vowel (Image: Chen (1997))



1.5.3 The GOOSE vowel: /u/-fronting

The GOOSE vowel has been undergoing fronting and diphthongization (moving closer to the FLEECE vowel for the first half of articulation) for many years in several dialects of English, including varieties of British English, Southern U.S. English, and in Philadelphia and California (Labov et al., 1972; Luthin, 1987; Veatch, 1991). Often, but not in

all cases, the GOAT vowel will front at the same rate as the GOOSE vowel, or slightly behind it (Hall-Lew, 2009). For speakers in the northern United States, the GOOSE vowel has long been pronounced as a monophthong in the back of the mouth (Ash, 1996).

According to Ash (1996, p. 4), there has recently been "considerable" fronting of the GOOSE vowel in the midwest, especially when the vowel is preceded by an apical consonant, as in *choose*, *shoot*, *soup*, and *doing*.[19] Allen (1973, 1976) provides further evidence that the GOOSE vowel has been fronting in the upper midwest since the 1970s. Speakers from Minnesota and Wisconsin front this vowel to a much lesser degree than speakers from other midwestern states, although "[s]ome speakers in the dominant cities of the area, Milwaukee and Minneapolis–St. Paul, do exhibit a high degree of fronting, [while] other speakers in those cities do not" (Ash, 1996, p. 21). Likewise, Labov (1991) asserts that, in Chicago, fronting of the GOOSE vowel can be found in the same individuals that participate in the NCS, although not for all speakers.

GOOSE vowel fronting is said to be a "global" sound change, affecting multiple social circles "rapidly and diffusely" (Fridland, 2008a), even across multiple ethnicities (Fought, 1999; Fridland & Bartlett, 2006). However, most folk linguistic perceptions of GOOSE fronting include relatively recent associations with California "Surfer Dude" or "Valley Girl" speech, as in words like *dude* or *spoon* (Fought, 1999; Fridland, 2008a; Hinton et al., 1987).

1.6 The aims and structure of this study

This dissertation explores the ways that Hmong Americans in Minnesota shape their ethnic identities through the strategic use of pronunciation variation in English. The majority speech community in Minnesota has been recently characterized by participation in several sound changes: the low back merger, the Northern Cities Shift, and GOOSE-fronting. I will investigate whether these sociophonetic variables are present in the English spoken by a group of Hmong Americans and a group of European Americans of the same age and from the same area of the state. Vowels will be quantified using primarily their first and second formant measurements. Nasalization, in relation

[19] Ash (1996) goes on to say that GOOSE vowels preceding /l/ tend to avoid fronting, as in the word *tool*.

to the Northern Cities Shift, provides a further avenue of investigation, beyond the two-dimensional first and second formant measurements, and will be explored using several measurement techniques.

There are two main questions motivating this research. The first question can be divided into two sub-questions.

1. Do Hmong Americans use language (specifically pronunciation differences) to establish, maintain, and change (peoples' perceptions of) their ethnic identities?

   (a) What are the phonetic differences between Hmong Americans and European Americans? Specifically, how do vowel pronunciations spoken in single words differ between the two groups?

   (b) Are these phonetic differences perceived by Hmong Americans and/or European Americans? In other words, can naive listeners accurately determine the ethnicity of Hmong American and European American speakers? What are the most important phonetic cues for these listeners in making accurate judgments?

2. What aspects of Hmong Americans' sociocultural and linguistic background are influencing their phonetic choices?

Research questions 1, 1a, and 1b will be addressed in Chapters 2, 3, 4, and 5. Chapter 2 explores the methodology used for studies of speech production, both of Hmong Americans and Europeans Americans, and Chapter 3 provides the results of these studies. Chapters 4 and 5 present the methodology and results of the speech perception studies.

Research question 2 will be addressed in Chapter 6, through the integration of results from speech production and speech perception, and taking into account the speakers' individual social and linguistic backgrounds. Chapter 6 will also provide a conclusion and suggestions for further research.

The answers to these questions may contribute new information to some of the bigger questions in sociolinguistics, including: How does language change? How does society use language to achieve social goals? For the subfield of diasporic linguistics,

the following questions will also be addressed: How do heritage languages adapt to new influences from other languages and cultures? How does the unique politics of a diasporic community affect the way language is used for sociolinguistic purposes?[20]

[20] See Hinrichs (2011) for more about the big questions in diasporic linguistics.

Chapter 2

Production Study Methods

2.1 Motivation and background

In this chapter and the chapter that follows, Hmong Americans' vowel productions are observed and compared to those of European Americans. I examine here Hmong American linguistic patterns by exploring the claim that minority ethnic groups do not participate in majority sound changes (Labov, 2001).

Specifically, during sociolinguistic interviews, I collected vowel tokens known to be involved in sound changes that are purportedly active in the Twin Cities, at least among European Americans. Vowels tokens were collected via wordlists from Twin Cities Hmong Americans and European Americans of the same age group. These vowel tokens were then quantified in several acoustic dimensions including formant values and nasalization.

2.1.1 Data collection

The first step in the process of analyzing sound patterns used by a speech community consists of collecting the raw materials: actual speech samples that include the sound patterns in question. We have a few options available for collecting appropriate speech samples. The first option is to simply ask speakers outright, "Would you please speak as you normally do?" This will probably not be very successful because we have called attention to their speech, which usually makes speakers switch to a more guarded, more prescriptively "correct" register.

In fact, many of the speech patterns that sociolinguists hope to observe and characterize disappear when speakers are concentrating on their speech. This is the so-called Observer's Paradox: the very speech patterns that we hope to observe vanish as soon as we begin to observe them (Labov, 1972, p. 10). The closest we may come to observing a natural, uninhibited register is when we are able to divert the attention of the speaker from their speech, through the use of a comfortable context (in a familiar environment, with familiar interlocutors) or by eliciting an exciting or calamitous topic of conversation.

Another option available to sociolinguists is the ethnographic method of participant observation. Conducted in contexts familiar to the speakers, this method allows the speakers to produce speech that is more relaxed, less controlled, and thus more valuable, at least from a sociolinguist's perspective. In addition, Eckert (2005) observes that "[e]mploying an ethnographic approach to determine locally relevant social categories enables researchers to break free from the mold of using predetermined categories that may or may not be relevant for a particular group of speakers." There are three major drawbacks to this method, however. The first is that it usually takes a great deal of time, and the second is that there is very little observer control over the speech samples collected. Many hours of participant observation may sometimes result in only a few analyzable speech patterns. The third drawback is that the observer has less control over the recording quality, which is particularly important when analyzing low-level phonetic differences.

A third option for observing speech is the traditional sociolinguistic interview. This is a more linguistically and acoustically controlled environment than participant observation, which leads to a greater concentration of useable speech samples in a shorter amount of time. Information-gathering strategies include direct questioning and reading narrative passages and wordlists. Through strategic structuring of the interview, it is possible to obtain samples of speech patterns in question without actually calling overt attention to those patterns. The greatest downfall of this method, though, is that speakers often pay greater attention to their mannerisms, including speech, resulting in more mainstream speech (i.e., less variation than is normally found for individual speakers).

Lastly, sociolinguistic data can be collected by means of carefully designed laboratory experiments. In such experiments, subjects provide judgments on previously recorded speech samples, rating their naturalness or social meaning, given certain social characteristics of the speakers. This method allows subjects to relax, knowing that their own speech is not being evaluated, which then allows them to concentrate on their judgments of others' speech. However, this method is not without its flaws: the experimenter cannot control for all of the world knowledge and acoustic cues that subjects use when making their judgments. Of the methods mentioned here, this one allows for the most control over the data (which is desirable), yet also takes place in the most unfamiliar environment (which is not desirable). The best laboratory experiments are ones designed to maximize useable data and minimize subject discomfort or task unnaturalness.[1] The laboratory experiment method is often most useful when combined with a non-laboratory method such as the sociolinguistic interview. In combining more than one method, the data become more valuable than the sum of their laboratory and non-laboratory parts.

2.1.2 Methodology for this dissertation

This dissertation is a combination of three methods: sociolinguistic interviews, group conversation recording, and a laboratory perception experiment. The participant observation and sociolinguistic interview questions allowed me to explore subjects' potential motivations for adopting certain phonetic patterns. The wordlists recorded during the interviews provided the desired phonetic data to be analyzed. And the perception experiment contributed information about how these phonetic patterns are perceived by listeners. Through the use of these three data collection methods, I hoped to gain insights not achievable by one method alone. The combination of these methods, along with powerful computational tools to aid in quantifying fine phonetic detail, allowed me to explore more fully why speakers sound the way they do. As Hay and Drager (2007, p. 89) claim, "it is the combination of all these approaches [sophisticated phonetic analyses, ethnographic approaches, and experimental techniques] that holds the key to an integrated understanding of how phonetic variation is produced, performed,

[1] See Thomas (2002) for more about how and why to design perception experiments for sociophonetic investigations.

and perceived in its social context."

It has been a central concern in sociophonetics to attempt to explain the motivations (either conscious or subconscious) behind certain phonetic choices that speakers make. In exploring motivations for linguistic choices, we cannot simply ask speakers, "Why do you speak in this way?" This is because speakers are often unaware of the phonetic pattern in question (as in, e.g., Fridland, 2008b; Kaiser & Plichta, 2009), speakers are not comfortable discussing the attitudes that motivated the use of that speech pattern, or the attitudes behind the phonetic choices are under the level of consciousness.

Through group conversation recording, I have gathered evidence of attitudes of identity and ethnicity based on interactions with other speakers. In sociolinguistic interviews, I have asked questions about attitudes toward local and regional pride, one's own ethnicity, racial relations in the Twin Cities, and language use; and I have recorded wordlists from subjects. The laboratory experiment explores whether the phonetic features in question can be perceived by others, i.e., the phonetic features can actually be used to index social meaning, rather than just being artifacts of articulation or acoustics. The methods of the first two procedures will be covered in this chapter, the results of the wordlist productions in Chapter 3, and the results of the ethnographic investigation in Chapter 6. The laboratory experiment is reported on in Chapters 4 and 5.

2.1.3 Choice of speakers

The target population of Hmong Americans was chosen for both linguistic and social reasons. Hmong Americans, as members of the pan-ethnic Asian American population, are a critical yet understudied part of the sociolinguistic variation found in the United States. Furthermore, the linguistic behaviors of Hmong Americans, in particular, may provide key insights into the cross-linguistic tendencies of phonemic nasalization influences on languages without phonemic nasalization (such as English), within the same individual and across a community of speakers. Socially, Hmong Americans tend to have more tight-knit social structures than other communities in the United States, which may have important sociolinguistic ramifications for rate of sound change adoption.

European Americans were chosen as the comparison group because they are the majority ethnicity in the Twin Cities. In addition, most ethnolinguistic investigations

in the United States have compared one or more minority ethnicities to white European Americans, and this study follows in the same vein.

To recruit participants, I emailed Hmong and international community centers in the Twin Cities, posted an ad on Craigslist, posted fliers on and off campus, and announced the study in two undergraduate linguistics classes.

2.1.4 Choice of location

About 50,000 Hmong Americans call the Twin Cities home—more than in any other urban area in the United States (ACS, 2009). The large number of Hmong Americans in the Twin Cities area means that they are more likely to be members of speech communities that include other Hmong Americans, and consequently they may be less likely to assimilate to European American linguistic and social practices.

Among European Americans, several important sound changes have been shown to be incipient or ongoing in this dialect region, including the low back merger, GOOSE-fronting, and the Northern Cities Shift. These sound changes provide multiple linguistic options along several acoustic and articulatory continua for speakers to use to identify with various social groups. With the availability of more linguistic options, there is a greater chance that communities of speakers have appropriated some of this available phonetic variation to index their identities. For these reasons, and for the sake of convenience, the Twin Cities was the ideal research site for this project.

2.2 Fieldwork

2.2.1 Procedure

The first phase of the fieldwork component consisted of sociolinguistic interviews. The first interview session with Hmong Americans lasted between 30 and 90 minutes, and was used to gather information about non-linguistic aspects of identity for each of the participants; the social networks of Hmong Americans in Minnesota; and personal experiences with racism, ethnic favortism, and discrimination. Language was not mentioned at all in the first session unless the participant brought it up on her or his own. (See interview questions in Appendix B). During this session, I also asked each participant to

read a list of pre-determined words near the end of the first session to ensure the target sounds occur a sufficient number of times in various phonological environments. (See wordlist in Appendix C.) Participants were asked to read the words aloud as each word was individually presented on a computer screen. Participants were instructed to read the words as if they were talking to a friend. The second session with Hmong Americans lasted about 30 minutes, and included questions about language habits, language attitudes, and relationships between language and their identities. The second interview sessions began after all of the first interviews had been conducted so that I could compile preliminary findings to share with the participants at the second interview session and gather their opinions on those findings.

The interviews with European Americans consisted of one session only, in which language-related topics were brought up only during the second half of the 30-minute interview, after the wordlist had been read. These interviews were conducted as similarly to the previous ones as possible, making changes where appropriate because of the condensed nature of the interviews. Questions given to the European American group are included in Section B.3 of Appendix B.

All interviews were conducted one-on-one with no one else present, usually in a small quiet room on the University of Minnesota campus. The equipment used included: a Marantz PMD660 Solid State Recorder (with built-in pre-amplifier), a head-mounted AKG MicroMic Series II C 420 cardioid condenser microphone, Sony MDR-NC6 noise-canceling headphones, and a Rolls PB23 Phantom Power Adapter.

All participants were given pseudonyms of their own choosing to protect their anonymity. All participants gave their verbal consent after reading the review-board-approved consent information sheet and asking any questions they had. Participants were aware that they were being recorded.

The second phase of the fieldwork component consisted of one hour of group conversation recording for three separate groups of participants. At least one of the participants in each group was Hmong American and had also participated in the first phase of the fieldwork. To facilitate conversation without actually speaking myself, I provided questions written on large notecards and placed on the table in front of the participants. These questions are listed in Section B.4 of Appendix B. Participants had the option of using these questions, or following the conversation wherever it led them.

The group conversations took place either at a local coffee shop (for Groups 1 and 2) or at the home of one of the participants (for Group 3). These sessions were purposely not held on the University campus in the attempt to maximize the comfort of the participants and the naturalness of their speech.

Group 1 included two original participants and four new participants. Group 2 included one original participant and two new participants. Group 3 included three original participants and one new participant. All participants in Group 3 were sisters or half-sisters.

For all of the group conversations, I used a Marantz PMD660 Solid State Recorder (with built-in pre-amplifier), a Shure 809 surface-mounted omnidirectional condenser microphone, Sony earbuds, and a Rolls PB23 Phantom Power Adapter.

2.2.2 Participant demographics

Fieldwork participants identified either as Hmong American and fluent in English (hereafter HA) or as European American (or mixed-race European American and another ethnicity) and native speakers of English (hereafter European American; EA). All participants had lived in Minnesota from age eight or before (except two HA participants; see Table 2.1) and were between the ages of 18 and 25 when the study took place. These restrictions were created in order to limit the scope of this project. The specific age range was chosen to maximize potential participation in age-graded sound changes. All participants' self-reported hometowns were within the 13-county Twin Cities metro area (except for two EA participants; see Table 2.2).

Hmong Americans

I collected data from 23 Hmong Americans who lived in the Twin Cities, whose demographic information is listed in Table 2.1. Subject 1 ("Carrie") was excluded from further analysis due to technical problems with the recording equipment during her interviews. Of the remaining 22 Hmong American subjects, 11 were born in Minnesota. Of the 11 subjects *not* born in Minnesota, some were born in other U.S. states such as Wisconsin (n = 5), Rhode Island (n = 1), and California (n = 1); one was born in France and three in Thai refugee camps. Several subjects (Sheng, Jade, Jeff, Julia, and Wendy) spent a significant amount of time in another city before moving to Minnesota.

These subjects are recorded as having two hometowns (one in Minnesota and one not in Minnesota). Table 2.1 summarizes the demographic profiles for all Hmong American participants. All cities listed in the table are located in Minnesota unless otherwise noted.

Table 2.1: Hmong American speakers from the Twin Cities, organized by birthplace

SPEAKER ID	PSEUDONYM	AGE	SEX	HOMETOWN	AGE OF ARRIVAL TO MN	WHERE BORN
4	Cynthia	18	F	Brooklyn Park	native	Minneapolis
6	Sunny	18	F	Columbia Heights	native	Minneapolis
8	Green	25	F	Eagan	native	Minneapolis
20	Morgan	20	M	Minneapolis	native	Minneapolis
10	Max	18	M	St. Paul	native	St. Paul
13	Gina	20	F	St. Paul	native	St. Paul
14	Mai	19	F	St. Paul	native	St. Paul
15	Gaozoua	20	F	St. Paul	native	St. Paul
17	Yoli	19	F	St. Paul	native	St. Paul
19	Spring	21	F	St. Paul	native	St. Paul
11	Blue	20	F	St. Paul	native	Rochester
5	Yue	22	F	Columbia Heights	2	Eau Claire, WI
18	Julia	23	F	Green Bay, WI & St. Paul	18	Green Bay, WI
3	Pang	20	F	St. Paul	2 mo.	Manitowoc, WI

(Table continued on next page)

Table 2.1 – continued from previous page

SPEAKER ID	PSEUDONYM	AGE	SEX	HOMETOWN	AGE OF ARRIVAL TO MN	WHERE BORN
7	Sheng	19	F	Sacramento, CA & St. Paul	8	Milwaukee, WI
12	Tasha	21	F	Brooklyn Center	2 mo.	Milwaukee, WI
21	Jiru	19	F	Minneapolis	4	Fresno, CA
23	Wendy	23	F	Fresno, CA & Brooklyn Center	13	Providence, RI
16	Jeff	21	M	Fresno, CA & Brooklyn Park	6	France
1	Carrie	29	F	Minneapolis	8	Thailand
2	Chee	23	F	St. Paul	1	Thailand
9	Jade	24	F	Sacramento & St. Paul	8	Thailand
22	Ruby	21	F	St. Paul	5	Thailand

European Americans

The comparison sample consisted of two males and seven females. Ethnicity of the EA group was not an explicit part of the recruitment process, but, given the demographics at the University of Minnesota where the study was conducted, I expected most, if not all, respondents would be European American when race/ethnicity was not an explicit recruitment factor. As it happened, only one of the nine EA respondents was Vietnamese American; the rest were European American or mixed-race Asian American–European American. I did not discard the data from the Vietnamese American participant because they did not differ significantly from the data of the others in the comparison sample.

All female participants' hometowns were located inside the officially-recognized 13-county metro Twin Cities area. One male was from just outside this area, and the other male was from Superior, Wisconsin, just across the border from Minnesota. Table 2.2 summarizes the demographic information for the European American speakers. Again, all cities listed in the table are located in Minnesota unless otherwise noted.

Table 2.2: Non-Hmong American speakers from the Twin Cities, organized by hometown

SPEAKER ID	PSEUDONYM	AGE	SEX	HOMETOWN	ETHNICITY	13-COUNTY METRO AREA?
31	Elizabeth	22	F	Blaine	Asian Am.	yes
24	Jessica	25	F	Edina	Mixed	yes
30	Rachel	19	F	Elk River	Euro. Am.	yes
26	Erica	20	F	Hopkins	Euro. Am.	yes
28	Ashley	22	F	Lindstrom	Euro. Am.	yes
25	Thursday	20	F	Minneapolis	Euro. Am.	yes
32	Emma	20	F	Minnetonka	Mixed	yes
29	Kid	21	M	Pine Island	Euro. Am.	no
27	Halden	22	M	Superior, WI	Euro. Am.	no

2.3 Predictions

The null hypothesis for this study is that HA speakers would show acoustic patterns that do not differ significantly from the EA comparison sample. However, I expected that I would reject this hypothesis. I anticipated HA speakers would tend to produce slightly less-shifted variants (but not fully standard American English, either) compared to EA speakers, for the relevant supra-regional and regional sound changes, for two reasons: (1) HA networks are likely more isolated than EAs, and this would predict changes being adopted slower than in looser networks (Milroy, 1980; Milroy & Gordon, 2003), and (2) the outside links that Twin Cities-based HA networks have are generally to other HA networks, many of which are located outside of the Inland North.

2.3.1 The low back merger

The low back merger, a sound change affecting the relative positions of the LOT and THOUGHT vowels (see Section 3.1.4 of Chapter 1), is prevalent in much of the country, including Northern Minnesota. Although there is anecdotal evidence of the merger in Minnesota from the 1940s through the 1970s (Allen, 1973, 1976), it is said to be currently receding in Southern Minnesota (Labov et al., 2006). If this is the case, we would expect to see the LOT and THOUGHT vowels moving away from one another in the current study, relative to the data in the Atlas. Typically, a supra-regional sound change such as this one is more likely to be adopted by speakers who are part of looser social networks, or who are not as proud to be natives of a certain region. However, since the low back merger may be receding, the sociolinguistic associations with this sound change have also likely shifted. It would follow that the merged vowels may be associated with older speakers, at least in Southern Minnesota, so younger speakers would be motivated to produce un-merged vowels.

Previous literature on the interaction between ethnicity and the low back merger has claimed that minority ethnicities produce LOT and THOUGHT vowels similarly to the majority, whether merged or not (Bernstein, 1993; Eberhardt, 2008; Hall-Lew, 2009; Hinton et al., 1987) or, more often, less merged than the majority (Bernstein, 1993; DeCamp, 1953; Hazen, 2005; Hinton et al., 1987; Ito, 2010; Labov et al., 2006; Thomas, 2001). Therefore, it was expected that Twin Cities HA speakers would produce less-shifted LOT and THOUGHT vowels compared to EA speakers. However, HA networks are probably more tightly knit, which tends to slow the adoption of sound changes, especially sound changes that are strongly associated with other ethnicities. If the low back merger is receding, we might actually see *more* shifted LOT and THOUGHT vowels among HA speakers.

Since the low back merger and the Northern Cities Shift both involve the LOT and THOUGHT vowels but their movement is nearly antithetical for the two sound changes, it is rare to find both the low back merger and the NCS advancing in the same speech communities. Therefore, we might also expect that, given the disappearance of the low back merger, the NCS, in turn, would have more freedom to spread in the Twin Cities.

2.3.2 The Northern Cities Shift

Theoretically, if the NCS were progressing in the Twin Cities as in most other NCS cities, we would expect to see TRAP-raising followed by LOT-fronting, then followed by THOUGHT-lowering, and so on (see Section 1.5.2 of Chapter 1). And in fact, the ANAE (Labov et al., 2006) has documented TRAP-raising for speakers living in the Twin Cities metro area, and has taken this to be a sign of the Northern Cities Shift spreading to this region. We would expect, then, that if the NCS is spreading to the Twin Cities, data collected more recently would show an even greater raising of the TRAP vowel, and some LOT-fronting and perhaps THOUGHT-lowering, relative to the data in the ANAE.

Gordon (2000a) reports that African American speakers do not participate in the NCS to the same degree as whites from the same area. However, Jones (2003) and Roeder (2006) have shown that African Americans and Mexican Americans, respectively, will accommodate white speech by displaying some NCS features when talking to whites. Given that I am white, one would expect some amount of accommodation by Hmong American speakers to the white majority dialect that they expect me to speak during the interviews, based on the dialect they hear me and their European American peers speaking.[2] This accommodation would be in addition to the NCS features that HA speakers may already use to convey social meaning through mutual understanding of certain sociophonetic variables.

2.3.3 Nasalization

The Northern Cities Shift usually begins with raising of the TRAP vowel, often accompanied by some fronting as well. As noted in Section 1.5.2 of Chapter 1, increased vowel nasalization tends to raise listeners' perceptions of low vowels. Specifically, when the TRAP vowel is digitally manipulated so that the percept of nasalization is increased, listeners are more likely to hear something closer to the DRESS vowel, whether or not they realize there is nasalization on the vowel (they most likely do not) (Beddor & Hawkins,

[2] As a native of Milwaukee, Wisconsin, I have been observed to pronounce vowels shifted toward NCS values, especially for the TRAP vowel. I have also noticed myself fronting the GOOSE vowel frequently in recent years. My LOT and THOUGHT vowels are completely distinct, though. If speakers were displaying accommodation to my own speech during interviews, they would show some adoption of the NCS and GOOSE-fronting, but no adoption of the low back merger. For more on accommodation to the interviewer's dialect, see Hay and Drager (2010)

1990; Kaiser & Plichta, 2009). Given Plichta's (2005) finding that vowel nasalization correlates with participation in the NCS, it is reasonable to assume that the NCS began with the misperception, reinterpretation, and innovative production of nasalization and a raised low front vowel, which is thought to be the cause of many other sound changes.[3] It was predicted, therefore, that speech communities displaying greater amounts of nasalization would also show a greater degree of NCS-shifted vowels, or at least for the TRAP vowel.

Hmong Americans, on the other hand, are not very likely to allow increased amounts of nasalization in their speech, if there is significant influence on their English from the Hmong language. Because the Hmong language contains phonemic nasal vowels it might be predicted that Hmong oral vowels are significantly less nasalized than, e.g., English vowels in similar contexts, as has been shown for French vowels compared to English vowels (Chen, 1997). Moreover, second language transfer theories (e.g., Flege, 1995) would predict that a Hmong learner of English might carry over the tendency to preserve nasalization contrast in vowels from Hmong to English by not nasalizing English vowels as much as a native speaker of English would.[4] Given the tight-knit nature of Hmong American communities, it would follow that these L2 transfer tendencies would influence patterns of speech for others in the community who speak English natively, and no longer need to retain such a contrast in vowel nasalization. Thus, we would expect less nasalization in Hmong American's speech, and consequently, less shifted NCS vowels, if indeed increased nasalization leads to the first step of the NCS. Variation in nasalization, a potential sociophonetic marker, may have originated in language transfer difficulties, but now might simply be a way for Hmong American speech communities to linguistically separate themselves from mainstream speech communities.

[3] The other steps of the NCS need not be motivated in the same way; they may simply follow from the first vowel's movement, in a typical chain shift pattern.

[4] It is an important empirical question (one that will not be evaluated here) as to whether native speakers of Hmong would *perceive* allophonically nasalized English vowels as equivalent to Hmong phonemically nasal vowels and then *produce* them with as much nasalization as their phonemically nasal vowels. For the purposes of the current work, all allophonically nasalized vowels in English will be assumed to be perceived as oral vowels by Hmong native speakers.

2.3.4 GOOSE vowel

It is undisputed that the GOOSE vowel is undergoing fronting all over the country, in some places for many years. Minnesota's GOOSE-fronting began more recently, and is still mostly limited to larger cities like Minneapolis and St. Paul (Ash, 1996). As my subjects are all from the metro Twin Cities area, I predicted that the GOOSE vowel would be more fronted than in previous studies.

Because GOOSE-fronting, like the low back merger, is a supra-regional sound change, I hypothesized that individuals whose social networks are looser and/or more outwardly-focused would tend to front the GOOSE vowel more than individuals with inwardly-focused, tighter social networks. Specifically, if my assumptions are accurate, European Americans should tend to front the GOOSE vowel more than Hmong Americans.

2.3.5 Consonants

As the main focus of this dissertation is on vowel variation, not a great deal of attention will be given to consonant variation. However, it is important to mention that if some vowel variation exists, some consonant variation is also to be expected. In particular, influence from Hmong L2 learners of English may be present in the speech of Hmong American native English speakers. This influence would likely include: word-final consonant cluster reduction, stopping of [θ] and [ð], difficulty with [l] and [r], and no plural marking or other word-final suffixations (Bliatout, Downing, Lewis, & Yang, 1988). If Hmong Americans tend to show consonant productions that resemble L2 learner pronunciations, this would be more evidence in favor of the other hypothesis that involves L2 learner influence—that of decreased nasalization.

In the next chapter, vowel productions will be reported and analyzed quantitatively and qualitatively, while consonant production trends will be described impressionistically at the end of the chapter.

Chapter 3

Production Study Results

The primary objective of the production study was to determine similarities and differences between the phonetic properties of English vowels produced by Hmong Americans and European Americans. In this chapter, I examine whether either group is participating more readily in several sound changes purported to be in progress in the Twin Cities: the low back merger, the Northern Cities Shift, and GOOSE-fronting; and whether the two groups differ with respect to the use of sociophonetic vowel nasalization. The motivation behind investigating these particular variables is provided in Chapters 1 and 2.

First, it is important to characterize the phonetic state of the speech of majority European Americans in the Twin Cities, who make up the comparison sample in this dissertation. Because we are dealing with sound changes, this means that we must also determine whether the sound changes have advanced or retreated, relative to a certain point in the past. The most feasible option for doing this involves collecting data from current EAs living in the Twin Cities, and compare that to data from Twin Cities EAs from some time ago.

The most comprehensive study of Twin Cities English vowels is found in the Atlas of North American English (ANAE; Labov et al., 2006), collected from Twin Cities residents over 16 years ago.[1] In order to show whether the sound changes are spreading

[1] To my knowledge, there are three other systematic investigations of Twin Cities English vowels besides the ANAE: Allen (1973, 1976) provides many impressionistic descriptions of vowel variation but does not include any quantitative measures of vowel formants; Thomas (2001) investigates vowel formants for only one speaker; and Ito (2010) provides a thorough quantitative description of a limited

in mainstream Twin Cities English, I compare data I collected in 2010 to data from the ANAE, which provides us with an approximation of the trajectory of the sound changes over a 16-year period. This is addressed by the first study I describe in this chapter. The second study in this chapter compares current (2010) EA vowels with HAs' English vowels.

3.1 Production study #1: Sound change among European Americans in the Twin Cities

3.1.1 Overview of study #1

This study was conducted in order to supplement our knowledge of the regional characteristics of majority Twin Cities English. In this study, data from the ANAE was compared to data that I collected from European American Twin Cities residents in 2010. To collect my data, I conducted face-to-face interviews, described fully in Section 2.2.1 of Chapter 2. Originally, I collected and analyzed data from nine speakers from Minnesota, whose demographic information is listed in Table 2.2 of Chapter 2. For this study, I excluded the two speakers whose hometowns were outside the metro Twin Cities area, and who also happened to be the only two male speakers in the EA sample.

3.1.2 Methods of the ANAE

As part of their massive North American survey of the English language, Labov et al. (2006) collected data from four speakers who were from the metro Twin Cities area. The data for these four speakers were recorded between 1993 and 1994. The ANAE Twin Cities speakers' demographics are outlined in Table 3.1.

set of vowels. Therefore, the ANAE remains the best option for a comparison of different speaker groups over time.

Table 3.1: ANAE Twin Cities speakers

SPEAKER ID	AGE	SEX	HOMETOWN	ETHNICITY
a	31	M	Minneapolis	Unknown
b	24	F	Minneapolis	Euro. Am.
c	57	F	St. Paul	Euro. Am.
d	26	F	Stillwater	Native Am.

Most of the data for the ANAE were recorded over the telephone. Telephone recordings have been found to have lower first formants overall, higher first formants for the DRESS vowel before nasal consonants, and lower first formants and higher second formants for the FLEECE vowel (Labov et al., 2006). All acoustically analyzed tokens for the ANAE were taken from primary phrase and syllable positions, using several types of elicitation methods: minimal pairs, semantic differential items, wordlists, and spontaneous speech. Vowel formants were measured at their "central tendency," which is defined as "the trajectory of the tongue during its articulation" (Labov et al., 2006, p. 38). Vowel means for each speaker were calculated for all vowels with three or more tokens per vowel. The data were log-mean normalized using the Nearey method, with slight modifications, including a set group log mean (called the G-value) based on data from 345 of their 439 acoustically analyzed subjects.

3.1.3 Data analysis

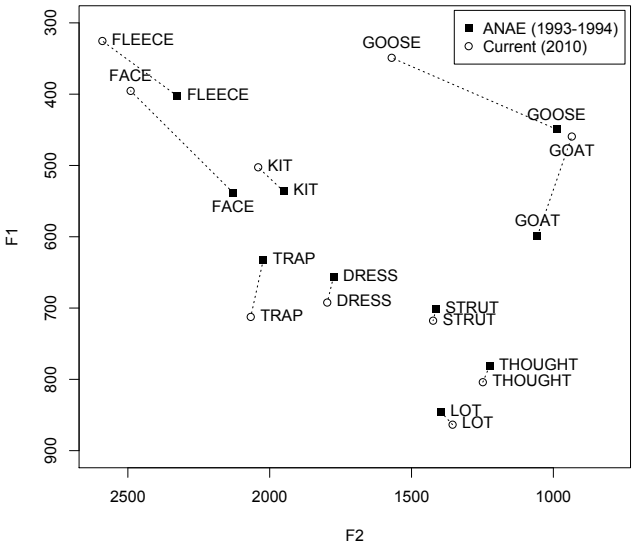
For the current study, analyzed tokens were taken from primary syllable position, from wordlists recorded during sociolinguistic interviews (described in Chapter 2). Formant measurements were made at the midpoint of each token, and means were calculated for vowels with three or more more tokens. The formant values were log-mean normalized using the same group log mean value as the ANAE (see above), to aid in comparison.

LPC analyses were semi-automated using Akustyk (Plichta, 2008) for Praat (Boersma & Weenink, 2008). Formant measurements and nasalization measurements were then automatically recorded by Akuskyk and checked by hand for accuracy. Akustyk also linked multiple speaker characteristics to each token automatically.

I plotted the data from both corpora in a two-dimensional F1/F2 vowel quadrilateral

(Figure 3.1), which allows us to observe differences involving vowels' first and second formant values, although other dimensions of speech such as nasalization cannot be observed using this visualization method. In Figure 3.1, the vowel means for the ANAE corpus are indicated by the squares, and my 2010 corpus means are indicated by the circles.

Figure 3.1: ANAE (1993-4) vs. current (2010) data: TELSUR G-value normalized vowel formant means



Cohen's d

The number of items in each group in this study are too small to allow for any meaningful calculation of significant difference between corpora. Instead, I calculated the Cohen's d values for both formants of each vowel, which gives a representation of the relative distance between the distribution of formant values for each corpora.

$$\text{Cohen's d} = \left| \frac{\bar{N} - \bar{O}}{stdev(N+O)} \right|$$

Larger Cohen's d values indicate greater separation between corpora formant means for a particular vowel. The schematic in Figure 3.2 represents two hypothetical distributions with greater separation, and therefore a larger Cohen's d value, compared to the schematic in Figure 3.3, which shows less separation between the two hypothetical distributions.

Figure 3.2: Hypothetical formant distributions with greater separation

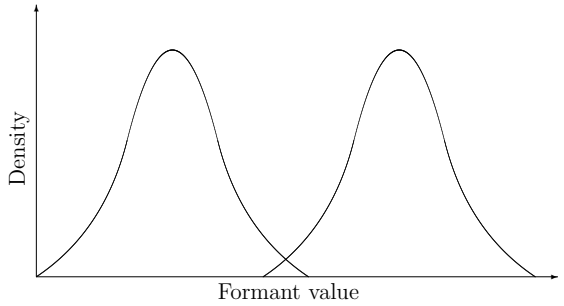
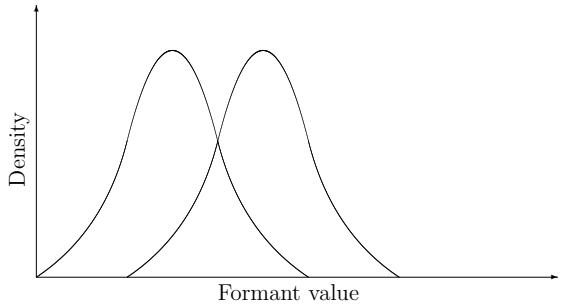


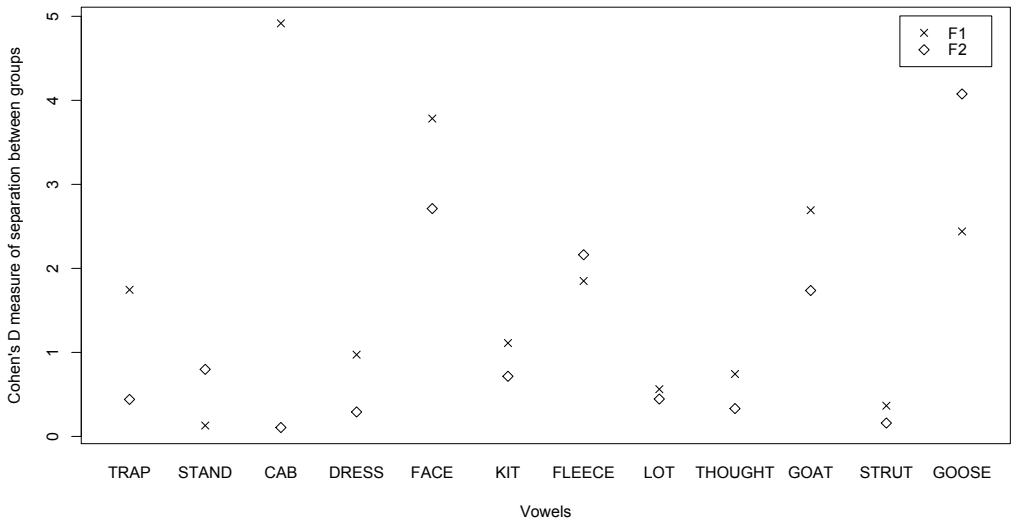
Figure 3.3: Hypothetical formant distributions with less separation



In order to compare the formant distributions for each vowel's first and second formants (F1 and F2) between the two corpora, I plotted the Cohen's d values on the vertical axis, and each vowel listed individually on the horizontal axis. This is shown in Figure 3.4. I split the TRAP vowel into those before nasal consonants (STAND) and

those before oral consonants (CAB).

Figure 3.4: Cohen's d calculation of separation of vowel means for F1 and F2 between ANAE and current corpus



3.1.4 Results

Based on the traditional F1/F2 plot in Figure 3.1, the most general observation we can make is that for the vowels that are different between the two corpora, the speakers in the current corpus pronounce most vowels more peripherally than the speakers in the ANAE. This likely due to hyperarticulation associated with wordlists, as vowel means were calculated from wordlist data only in the current corpus, as opposed to different types of elicitation methods in the ANAE. In other words, speakers tend to move vowels further away from one another when they pay closer attention to their speech, and the data collection method of the current corpus was more likely to make people pay attention to their speech than that of the ANAE.[2]

[2] Jacewicz, Fox, and Salmons (2006) found that vowels given prosodic prominence (such as those spoken in wordlists) are more likely to show evidence of sound changes. Therefore, while the data elicitation methods differed between the two corpora, those differences should actually highlight any

Many of the differences between corpora can be accounted for by conceding a certain amount of hyperarticulation in the current corpus compared to the ANAE corpus. However, the GOOSE vowel is the most obviously not headed toward the periphery, and in fact seems to be following the GOOSE-fronting shift that is being observed in many regions throughout the country (as noted in Section 1.5.3 of Chapter 1). The differences in the GOOSE vowel are explored more in this section below.

The largest Cohen's d value shown in Figure 3.4 is the F1 of TRAP before oral consonants (CAB). The second-largest Cohen's d value is indicated for F2 of the GOOSE vowel. The third- and fourth-largest Cohen's d values are indicated for the F1 and F2 of the FACE vowel, but since both of these apparent-time changes were in the direction of the periphery, based on Figure 3.1, it is unclear how much of it is attributable to hyperarticulation. We will investigate several vowels in more detail now, including TRAP, GOOSE, and the two vowels involved in the low back merger (LOT and THOUGHT).

The low back merger

It was predicted in Chapter 2 that the low back merger would be receding, according to claims made in the ANAE. However, based on Figure 3.1, there seems to be very little change in either the LOT vowel or the THOUGHT vowel. It appears that the low back merger has neither advanced nor receded in the sixteen-plus years since the ANAE data were collected.

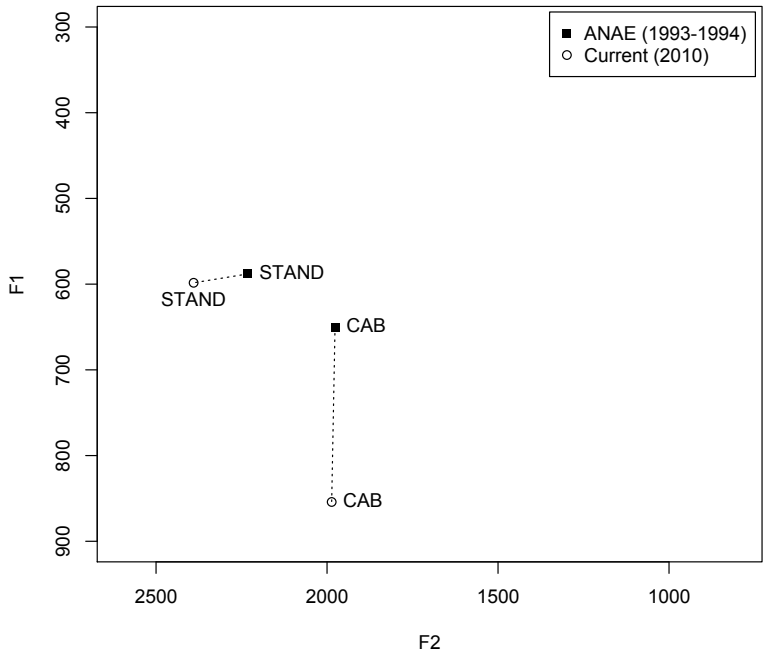
The Northern Cities Shift and nasalization

As the closest metro area west of the Inland North, the Twin Cities should start to see more and more adoption of the Northern Cities Shift, especially since the low back merger has probably not advanced, leaving the low vowels freer to move.

The NCS usually begins with a raised and slightly fronted TRAP vowel. The TRAP vowel in Figure 3.1 seems to be headed just slightly toward the periphery, which is to be expected. However, when the TRAP vowel is split into STAND (before nasal consonants only) and CAB (before oral consonants only), something more interesting emerges. This new plot is shown in Figure 3.5.

sound changes currently in progress.

Figure 3.5: The TRAP vowel split into nasal and oral contexts



In Figure 3.5, STAND and CAB have moved further apart in the current data, compared to the ANAE data. Although some of the movement to the periphery is expected, we should keep in mind that the F1 difference between corpora for CAB had the largest Cohen's d value of any vowel (as shown in Figure 3.4).

The GOOSE vowel

The GOOSE vowel is well-documented as undergoing fronting in areas all over the U.S. (Hall-Lew, 2009; Labov et al., 1972; Thomas, 2001), so it is not surprising to see it fronting in the Twin Cities as well. The traditional stereotypical GOOSE vowel in Minnesota is monophthongal and typically very far back in the mouth, so this GOOSE-fronting is in direct opposition to the traditional variant. The backed GOOSE pronunciation is more likely to be associated with residents living in Northern Minnesota and

those living in more rural areas, so the fronted GOOSE variant is anecdotally seen as a more urban, "cosmopolitan" pronunciation. Compared to the data from 16 years ago, we observe more GOOSE-fronting among today's young European American Twin City-ites.

3.1.5 Interim summary

Before drawing any conclusions, we should keep in mind some of this study's limitations, including the methodological differences between the two corpora being compared, and the limited pool of subjects. Some of the methodological differences that may have distorted the comparisons include: the attention to pronunciation and formality that wordlists often cause, since wordlists were the only elicitation method for the current corpus but just one of four elicitation methods for the ANAE. Another concern is the location of vowel formant measurements: either at the midpoint (for the current study) or at the central tendency (for the ANAE).

Despite these limitations, this study has provided evidence of expected and unexpected changes in the regional dialect of residents of the Twin Cities. The low back merger seems not to have changed since 1993-1994. As a potential next stop for the NCS, it is expected that the Twin Cities would avoid the low back merger. On the other hand, the advancement of the low back merger has been documented for many years in the Twin Cities, so that might lead us to expect more advancement in the current data compared to the ANAE. Because we observe no change in the vowels involved in the low back merger, we have no evidence for the merger advancing or receding. One possible explanation for this result is that the NCS and the low back merger have been "fighting" for control of these vowels and have consequently caused the vowels to come to a stand-still.

The GOOSE vowel seems to be more fronted in 2010, compared to the ANAE. As one of the vowels that carries a strong conscious stigma of unsophistication (one of the most emphasized when Minnesota speech is caricatured), it is not surprising that the GOOSE vowel has undergone a shift to a more "cultured-sounding" fronted variant. The GOAT vowel often fronts with the GOOSE vowel in other areas of the country, but in the Twin Cities, the GOAT vowel does not seem to be heading in the same direction at this point. Other studies have pointed out that GOOSE and GOAT do not always front together in

all cases, with GOOSE more often fronting than GOAT (e.g., Hall-Lew, 2009).

Another interesting and somewhat unexpected finding concerns the TRAP vowel and its phonetic realization in nasal/oral contexts. Before nasal consonants, the TRAP vowel (called STAND) is slightly more fronted in the current data compared to the ANAE, but before oral consonants, the TRAP vowel (called CAB) is much lower than in the ANAE. STAND and CAB may be moving away from one another, with CAB shifting downward, possibly as a reaction against another stigmatized pronunciation: the "Minnesota-A"; and STAND shifting upward, possibly due to a nasalization-assisted sound change.

Finally, and perhaps most interestingly, none of the vowels associated with the NCS have moved in the expected direction, especially with the CAB vowel moving in the *opposite* direction. This may be evidence that TRAP-raising could have been acting as a red herring for earlier investigations of Twin Cities speech, looking very much like the first step of NCS-adoption, while in reality being simply TRAP-raising and nothing else.

3.2 Production study #2: Hmong Americans in the Twin Cities

3.2.1 Overview of study #2

According to my interviewees, Hmong Americans in the Twin Cities generally seem to be vaguely aware that they speak differently than people of other ethnicities, although saying *how* their speech differs was extremely difficult for them. The aim of this study of Hmong American and European American young people's speech was to determine whether HAs are participating in relevant sound changes in the Twin Cities, to a greater or lesser extent than their EA peers. Unlike for study #1, I collected all of the data used in study #2. The EA subjects analyzed here are the same seven EA women from the 2010 data used in study #1 above. The HA subjects analyzed here are the women listed in Table 2.1 of Chapter 2, other than "Carrie" (Subject 1). I chose to limit study #2 to women only for a more accurate comparison between the two speaker groups, since the only EA speakers I interviewed from the Twin Cities metro area were women. Furthermore, women are a sensible choice for this study because they tend to lead most sound changes (see Labov, 2001).

Currently, the only other existing study of Hmong American English phonetic characteristics is Ito (2010). In that study, the author reports that for young white Americans, TRAP is minimally fronted (much less than expected), DRESS is lowered and backed (similar to St. Louis' adoption of only certain NCS features—not in the traditional order of adoption), and there is a reduction of distance between LOT and THOUGHT for men only (in apparent merger). For HA speakers, the author found that TRAP and DRESS formants were similar to the "local norm" for younger speakers with lower age of arrival to the U.S., and for speakers with greater social mobility; and LOT and THOUGHT remain separate for most speakers (especially for those with greater social mobility, which was unexpected). From these findings, Ito concludes that L1–L2 interference and social interaction patterns may contribute to HAs' patterns of accommodation to the local majority speech, and that the Hmong American community in the Twin Cities is diverse—demographically and linguistically.

The goals of the study reported in Ito (2010) were not the same as the goals of the current study, and as such, Ito did not investigate several variables that are of interest in the current study. For example, Ito was not concerned with the influence of nasalization on the TRAP vowel, so the TRAP vowel was not measured separately by CAB and STAND allophones. Ito found only minimal fronting of TRAP by whites in her study, but if she had broken down the TRAP vowel further, she might have observed more interesting patterns between HAs and EAs, and even within the HA group. In addition, Ito did not investigate the GOOSE vowel.

3.2.2 Data analysis

As in study #1 above, all vowel tokens analyzed for study #2 were in primary-stressed syllables only, from wordlist data. Formants were measured at each vowel's midpoint, and normalized based on Lobanov's (1971) speaker-intrinsic, vowel-extrinsic method. Acoustic analysis were conducted as above, using Praat and Akustyk, and checked by hand for accuracy.

In Figure 3.6, the vowel means for the two groups are plotted on a two-dimensional vowel quadrilateral, with HA vowel means indicated by triangles and the EA vowel means by circles, as above. Because a different normalization method was used in study

#2, the axes are labeled differently from those in study #1.[3] F1 is again represented on the vertical axis, with higher F1 values toward the bottom of the figure; and F2 on the horizontal axis, with higher F2 values toward the left side of the figure.

Figure 3.6: Vowel plot comparing Lobanov-normalized Hmong American and European American means (axes are not scaled)

3.2.3 Results

Compared with study #1, the differences between the vowel pronunciations of the two groups in study #2 (in Figure 3.6) are smaller in magnitude. Part of this has to do

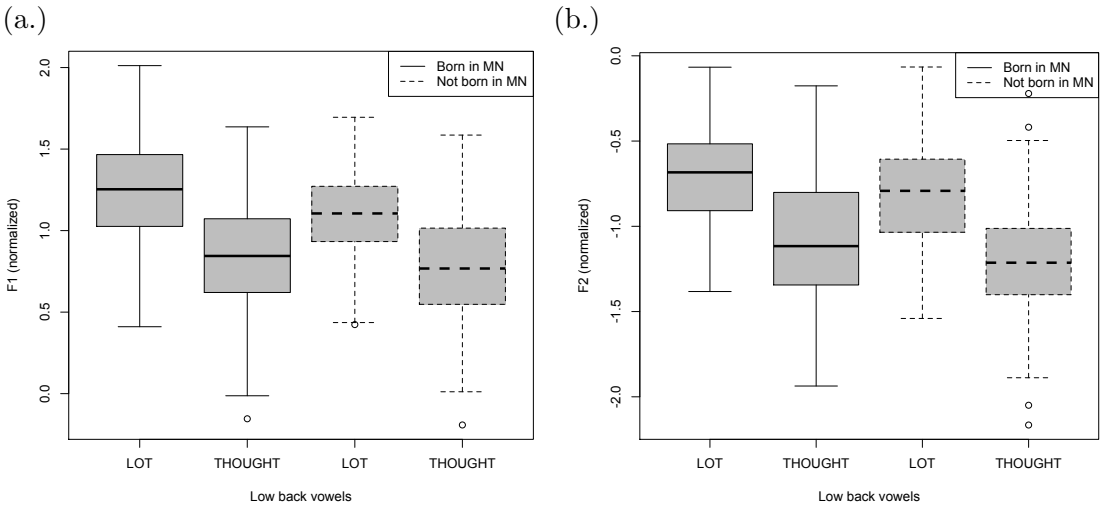
[3] The Lobanov normalization method was used here because it has been reported by Adank, Smits, and Van Hout (2004) to be superior to other normalization methods in several respects. For study #1, it was necessary to use the normalization method used in the ANAE for more accurate comparisons. As no direct comparisons are made between the data in studies #1 and #2, this normalization discrepancy should not be a problem.

with the methodology and subject demographics (other than ethnicity) being almost identical for both groups. This demographic and methodological similarity between groups allows us to more confidently attribute meaning to the differences that we do observe, if any.

The low back merger

As in Ito (2010), Figure 3.6 seems to show more separation between the low back vowels for HAs than for EAs, with EAs lowering and fronting THOUGHT more than their HA counterparts. However, quantitative analyses of the Cohen's d values for the separation of LOT and THOUGHT means for each speaker showed that there was no significant difference between the two ethnicity groups, for either F1 or F2. The linear models for Cohen's d values predicted by ethnicity group are provided in Appendix E: in Table E.1 for the Cohen's d for F1 and in Table E.2 for F2.

Figure 3.7: Boxplots comparing (a.) F1 and (b.) F2 means for the low back vowels (LOT and THOUGHT) by birthplace for HA speakers only



Within her HA sample, Ito (2010) found greater separation between the LOT and THOUGHT vowels for speakers with greater social mobility. I did not find that social mobility of HA speakers caused a significant difference between Cohen's d values for either F1 or F2. However, a linear model predicting F1 from HAs' birthplace achieved

a significant coefficient for the predictor variable. This model is given in Table E.3 of Appendix E, along with a model predicting F2 from birthplace, which did not approach significance in Table E.4. From Figure 3.7, we can see that those HA speakers born in Minnesota have more separation between the F1 means of LOT and THOUGHT than those not born in Minnesota.

The Northern Cities Shift

Our vowel of interest for the NCS is the TRAP vowel, which does not appear to be very different between the two groups in Figure 3.6. But when the TRAP vowel is split into its nasalized/non-nasalized allophones, as in Figure 3.8, we can observe something interesting. The CAB vowel is not different at all between the two groups, but the STAND vowel shows more of a difference between the two groups. Figure 3.9 shows the same data in boxplots. Recall that nasalization tends to raise the perception of vowels, and may cause a chain reaction that leads to actual raising of the vowel in speech production. So it is not surprising that the partially nasalized vowel variant STAND is pronounced higher (i.e., with a lower F1) than the non-nasalized variant CAB.

Also recall that speakers of languages like Hmong tend to maximize the contrast between nasal and oral vowels. Since both STAND and CAB in English are oral vowels (although STAND is *partially* nasalized), both should have low amounts of nasalization for Hmong speakers. For the HA speakers, because Hmong is their first language, and because they are often exposed to Hmong-accented English, we might expect to see an influence of the Hmong language in these speakers' English. Again, since the Hmong language has nasal vowels, speakers would be expected to use less nasalization on oral vowels, in other words, *all* English vowels. And, as expected, the HA speakers seem to be raising STAND less than the EA speakers.

Linear models of the Cohen's d measures for each speaker show a main effect of the ethnicity coefficient for both F1 and F2. In other words, the nasal/oral context causes a significant differential in EAs' F1 and F2 of TRAP, but less of a differential for HAs. The linear model coefficients are listed in Tables E.5 and E.6 of Appendix E.

Figure 3.8: Vowel plot comparing the TRAP vowel split into nasal (STAND) and oral (CAB) allophones by ethnicity)

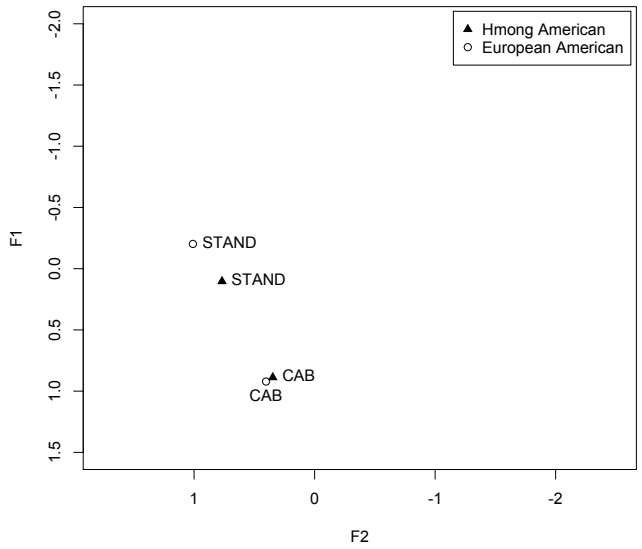
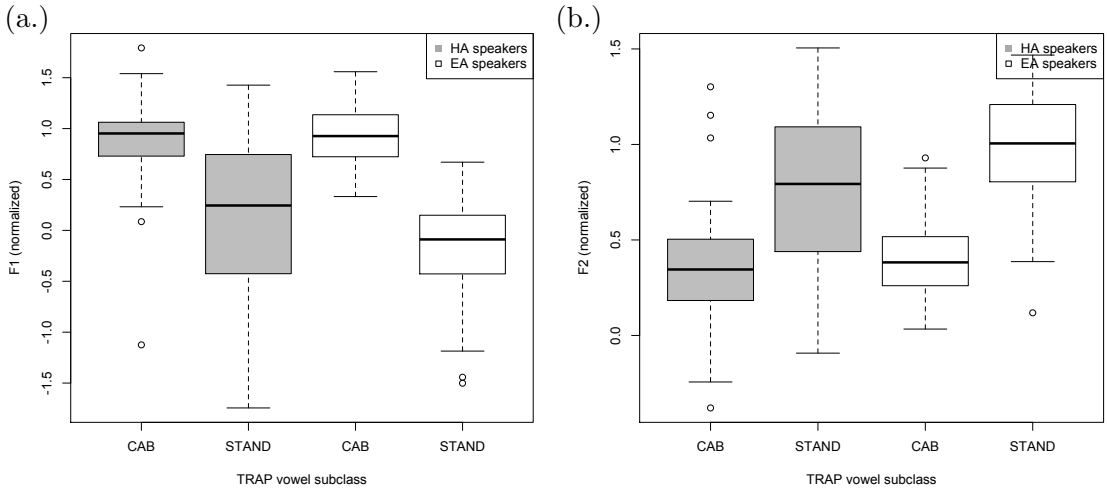


Figure 3.9: Boxplots comparing (a.) F1 and (b.) F2 means for the TRAP vowel split into nasal (STAND) and oral (CAB) allophones by ethnicity

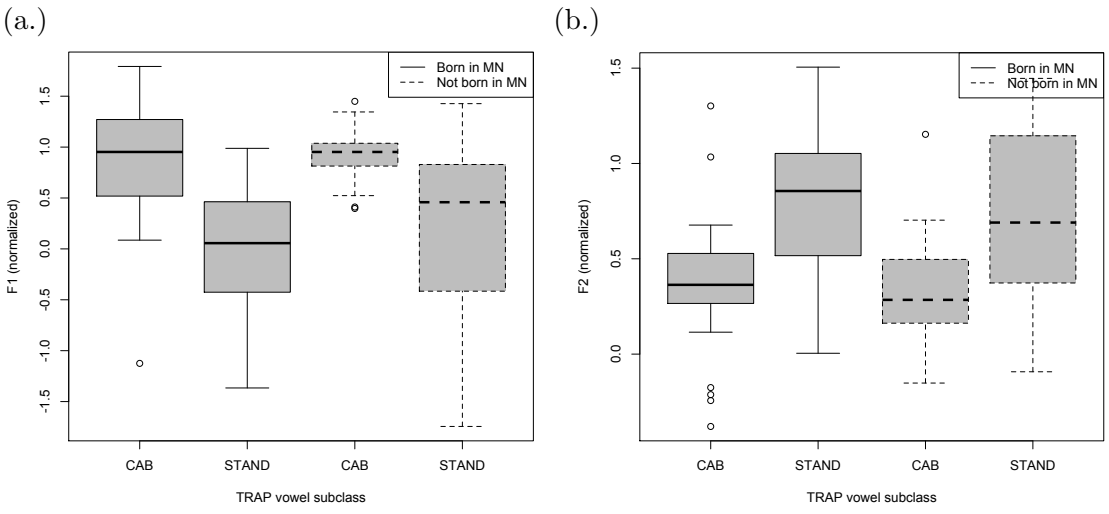


This might suggest that Hmong Americans are being influenced by the Hmong language, tending to nasalize less overall because of the presence of phonemic nasalization in Hmong. This lack of nasalization in HAs may result in listener misinterpretation of lower vowels (higher F1 values), and subsequent speaker production of lower vowels.

Yet, in a recent study done in Michigan with Mexican Americans, Roeder (2010) found the same pattern, where the majority whites were raising TRAP before nasal consonants more than the minority group. Crucially, Spanish, unlike Hmong, does not have nasal vowels. This suggests that perhaps an avoidance of TRAP-raising may not be due to speakers' exposure to a language with nasal vowels. Perhaps participation in TRAP-raising, for Mexican American and Hmong American communities, may simply be a phonological pattern useful for aligning or distancing oneself linguistically and socially from the majority speech community.

In addition, a very similar pattern can be observed, for F1 but not F2, when comparing two groups *within* the HA cohort: those born in Minnesota compared to those not born in Minnesota. As expected, those born in Minnesota more closely resemble EAs' TRAP-raising patterns, as illustrated in Figure 3.10 (a).

Figure 3.10: Boxplots comparing (a.) F1 and (b.) F2 means for the TRAP vowel split into nasalized (STAND) and non-nasalized (CAB) vowels by birthplace for HA speakers only



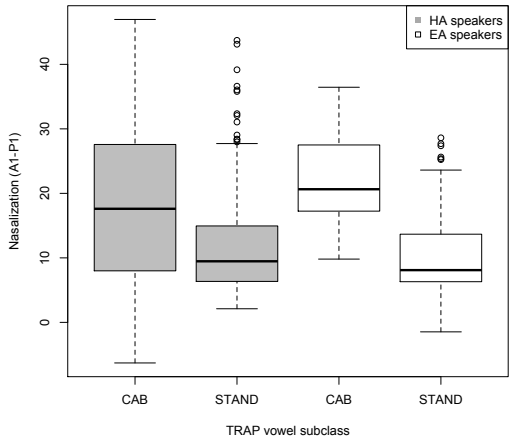
Linear models of the Cohen's d measures for HA speakers by birthplace show a main effect of the ethnicity coefficient for F1 but not for F2. In other words, the nasal/oral context causes a significant differential in HAs' F1 of TRAP, based on their birthplace, but the F2 means for the two groups did not differ significantly. Raising but not fronting seems to differentiate the two groups of HA speakers. The linear model coefficients are listed in Tables E.7 and E.8 of Appendix E.

Nasalization

Averaging over all vowels, nasalization levels did not differ between ethnicities. However, taking only the TRAP vowels into account, levels of *A1-P1* did differ between STAND and CAB, as expected, and also differed between ethnicities. Linear models of *A1-P0* and *A1-P1* by TRAP vowel sub-class and ethnicity as fixed effects were built, with speaker and word as random variables. These models are shown in Tables E.9 and E.10 of Appendix E. For *A1-P0*, although the coefficient for the TRAP sub-class (levels: STAND and CAB) was significant, the ethnicity coefficient did not reach significance.[4] For *A1-P1*, on the other hand, the coefficient for the interaction between TRAP sub-class and ethnicity was significant. From the boxplot in Figure 3.11, we can see that the direction of nasalization is unexpected: overall, there is more acoustic distinction between STAND and CAB for the EA speakers than for the HA speakers. Furthermore, most of the difference between the two ethnicity groups lies in the non-nasalized CAB vowel, rather than the nasalized STAND vowel.

[4] The coefficient for the interaction term was removed from the model because it did not add a significant amount of predictive power.

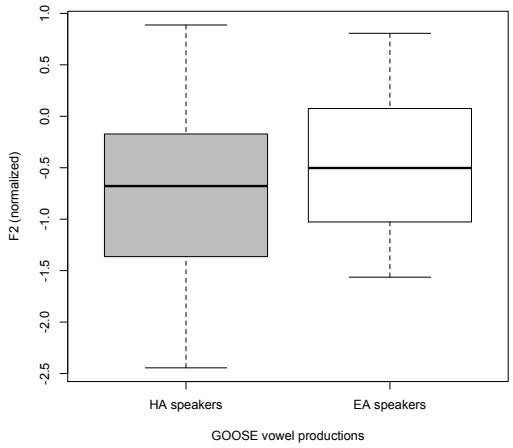
Figure 3.11: Nasalization (measured using *A1-P1*) for nasalized/non-nasalized TRAP vowel sub-classes (STAND and CAB) by speaker ethnicity (Lower *A1-P1* = more nasalization)



The GOOSE vowel

Qualitatively, we see in Figure 3.6 that European American women tended to front the GOOSE vowel more than Hmong American women, as expected. Fronting is indicated acoustically by a higher second formant. However, in a linear mixed-effects model of F2 with ethnicity as a fixed effect and speaker and word as random effects, the ethnicity coefficient did not reach significance. The data are represented in Figure 3.12 and model coefficients are listed in Table E.11 of Appendix E. The coefficient for speaker birthplace did not contribute significantly to the model, so it was removed.

Figure 3.12: GOOSE vowel F2 productions by speaker ethnicity



Consonants

Before I knew what to expect in terms of language-transfer issues for Hmong-influenced English, I observed and took note of several non-standard features of the HA speakers in the study: labiodentalization of /θ/ and /ð/ (the "th" sound pronounced as "f" or "v"), both word-medially and word-finally; /l/-vocalization (/l/ pronounced like [w]); word-final consonant cluster reduction (e.g., "des" instead of "desk"); and occasional dropping of the plural marker and third-person singular marker (both "-s").

All but the first observation can be attributable to Hmong language transfer to

English, according to Bliatout et al. (1988); we would have also expected /θ/ and /ð/ to be pronounced like [t] or [d] rather than [f] or [v].

3.3 Summary

In 2010, I recorded speech from seven European Americans from the Twin Cities. In Section 3.1 of this chapter, those speakers' vowel pronunciations were compared to speakers from the same area recorded 16-plus years ago by Labov et al. (2006). It was found that two out of three sound changes in question have not proceeded in the expected manner. The vowels involved in the low back merger, LOT and THOUGHT, have shifted very little since 1993-1994. The raising of the TRAP vowel, usually the first step in the Northern Cities Shift, seems to have moved in the opposite direction as expected. However, this lowering is driven exclusively by the CAB vowel—the sub-class of TRAP before oral consonants only. The last sound change in question, GOOSE-fronting, behaves as expected: current speakers produced more-fronted GOOSE vowels than speakers from 1993-1994.

In Section 3.2, the same EA speakers recorded in 2010 were compared to a group of demographically similar Hmong American speakers. Pronunciations of the low back vowels were expected to be closer for HA speakers than for EA speakers, based on previous literature. No significant difference was found for the LOT and THOUGHT vowels between the two ethnicity groups. However, HA speakers who were born in Minnesota tended to produce *more* separated vowels than HA speakers born outside Minnesota. Because it is still unclear whether the low back merger is receding or advancing in the Twin Cities, it is difficult to determine why this pattern was observed. The most likely conjecture is that the low back merger *is* receding among EAs, and HAs born in Minnesota align themselves more closely with EAs than HAs not born in Minnesota. This echos Ito's (2010) finding that HAs with greater social mobility—thus, those that interact more with EAs—show greater separation between the low back vowels.

For the raising of the TRAP vowel, HAs' and EAs' pronunciations differed only for the STAND vowel, which was not as raised for HAs as for EAs. At first glance, this may seem to be attributable to HAs' significant exposure to a language with phonemic nasal vowels, which may lessen nasalization overall, which in turn lowers the perception

and then production of STAND in particular, compared to EAs' STAND productions. However, similar findings in a Mexican American speech community (Roeder, 2010) suggest that heritage language is not the cause. Furthermore, birthplace among HA speakers was again a significant factor in the separation of the STAND and CAB vowels: more raised STAND vowels were produced by HAs who were born in Minnesota. This result further supports the hypothesis that Hmong language background is not the primary factor in whether a speaker raises STAND compared to CAB, because most of the Hmong American speakers in this study acquired English around the same time (head start or kindergarten), regardless of whether they were born in Minnesota.

In addition to a less raised STAND vowel, HAs actually produced *more* nasalization than EAs for the CAB vowel. Based on Plichta (2005), non-nasalized vowels should show more sociophonetic nasalization for NCS speakers than for non-NCS speakers. The STAND and CAB results, taken together, provide evidence that the NCS, if it has started advancing in the Twin Cities, is not following the typical pattern of adoption, and that neither HAs nor EAs can be said definitively to be adopting the NCS sooner. It would be worthwhile to investigate this phenomenon with more minority communities in the Northern Cities area, and then outside of the area, to see when speakers raise and/or nasalize TRAP and when they do not.

For GOOSE-fronting, EA speakers also displayed a trend to front the GOOSE vowel more than HA speakers, as expected. However, this trend did not reach significance in the linear mixed-effects test.

Several consonant differences were observed impressionistically, with HAs trending toward non-standard variants and EAs toward standard variants. Three out of the four non-standard consonant pronunciations can be attributed to heritage language transfer, but the fourth (labiodentalization of /t/ and /d/) cannot. This finding is further evidence suggesting that at least some differences reported here between HA and EA speakers may not be related to heritage language background. Instead, some (or possibly all) differences may be sociophonetic in nature, i.e., the differences are strategically maintained or created in order to index some aspect of a speaker's identity. The anecdotal evidence I have collected in the sociolinguistic interviews further suggests that some HA speakers may be patterning their speech, to a certain extent, after African American English, or "talking ghetto" (as several Hmong American participants said

of their more rebellious acquaintances). In fact, all of the differences noted here are characteristic of African American English, and could therefore be attributed solely to appropriation of African American English features by speakers of another minority ethnicity. Reyes (2005) has also found evidence of Asian American youth appropriating African American phrases, thereby efficiently transmitting an urban youth style. A similar phenomenon has been documented by Rampton (1995) for people of Asian heritage living in England and "crossing into" the local Afro-Caribbean English Creole.

In Chapter 6, the findings presented in this chapter will be integrated with the subsequent perception study results for a more complete picture of HA linguistic practices in Minnesota.

Chapter 4

Perception Study Methods

A perception experiment was conducted using excerpts of speech from the participants in the production study described in Chapters 2 and 3. The purpose of the perception experiment was to determine whether any of the inter- and intra-ethnic acoustic variation reported on in Chapter 3 is actually perceived by listeners and used to make judgments about speakers' social and demographic characteristics.

4.1 Main perception study procedure

4.1.1 Subjects and locations

For the speech perception study, a new set of participants was recruited. None of the speakers from the production study participated as listeners in the perception study, most importantly because the audio tokens being used in the perception study were recorded by speakers in the production study. All participants in the speech perception study were between the ages of 18 and 50; native speakers of North American English (or, if Hmong American, fluent in English and Hmong); and had no history of speech, language, or hearing problems; all according to self-report. These restrictions were created to minimize the interference of hearing loss or language-related differences between subject-listeners.

All European American participants (n = 27) completed the experiment at the University of Minnesota. Most EA participants were students or staff at the University

of Minnesota. These participants completed the experiment individually, seated in a double-walled sound-proof booth. Instructions and question prompts were presented on a computer screen in front of the participants. Normalized audio files were played to listeners on over-ear headphones. Participants input their responses by clicking with a computer mouse on a Visual Analog Scale displayed on the computer screen (see example in Figure 4.1).

Hmong American participants (n = 12) completed the experiment in a small, windowless room at the Hmong American Partnership (HAP), a local Hmong community organization in St. Paul, Minnesota. This location was chosen because it was more convenient for most Hmong participants than the University of Minnesota. The experiment was administered with Dell Latitude laptops, over-ear headphones, and a computer mouse to input responses. One to three people completed the experiment at a time due to time restrictions.

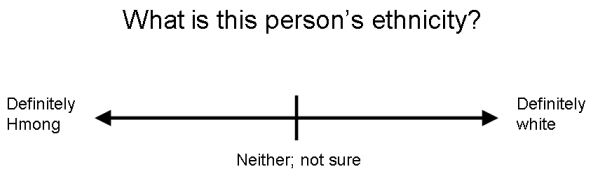
4.1.2 Recruitment

European American participants were recruited using fliers advertising the study, posted on and around the University of Minnesota campus. For recruitment of Hmong Americans, emails were sent to speakers from the production phase of the study, asking them to invite any eligible friends or family members to participate. Recruitment emails were also sent to friends with connections in the Hmong community, to Hmong language instructors at the University of Minnesota, and to Hmong and International community centers and churches in the Twin Cities area. Fliers were also posted on and around campus. None of these methods elicited any interest in the study among Hmong Americans; HA participants proved more difficult to recruit for this particular study, even though recruitment using the same methods for the production study had been successful. Therefore, an arrangement was made with the Hmong American Partnership to recruit participants through HAP's community connections, and allow participants to complete the experiment at HAP's St. Paul location.

4.1.3 Visual Analog Scale

This perception experiment utilizes the Visual Analog Scale (VAS) to elicit gradient judgments from subjects. This method is useful for measuring characteristics or attitudes that fall somewhere along a subjective continuum. The social construct of ethnicity is both subjective and gradient, so the VAS is an appropriate tool to measure people's judgments of others' ethnicity. Figure 4.1 illustrates one VAS elicitation prompt screen used in this experiment.

Figure 4.1: Prompt screen for ethnicity question (Actual size about three times larger)



4.1.4 Test stimuli

The speech data serving as test stimuli were extracted from the wordlists recorded during the one-on-one interviews with both HAs and EAs, described in Chapter 2. Four words were chosen from each of the five most promising vowel classes, according to the results described in Chapter 3: STAND, CAB, LOT, THOUGHT, and GOOSE. Nine filler words were also chosen at random. Stimuli were RMS-normalized to reduce the non-speech differences between them, in order to ensure that listeners were basing their input judgments on the acoustic information relevant to this study.

4.1.5 Study design

Participants in this experiment were asked to listen to single-word tokens played one at a time over headphones. For each word, listeners were asked to judge where the speaker of that word falls on a continuum for a given question prompt by indicating their judgment on a Visual Analog Scale. The question prompts that participants were asked to make judgments about included: speaker's age, speaker's ethnicity, speaker's pride in family heritage, whether the speaker was born in Minnesota or not, and the ethnicity of the speaker's friends. These questions were chosen to explore several likely sociophonetic influences.

The experiment was blocked by question, so subjects answered the same question for about 100 different stimuli before moving on to the next question. Subjects were offered a short break between each question block. With five question blocks total, subjects provided judgments on a total of about 450 unique tokens. Several tokens were repeated for accuracy and several tokens were resynthesized to sound more nasalized. A short practice block preceded the experiment, presenting the same questions and question prompt screens that subjects would see later in the experiment, but the words and speakers in the practice items were not used in the rest of the experiment. Demographic information for subjects and open-ended questions about what they used to make their judgments were collected at the end of the perception experiment. The experiment took about 45 minutes to complete, and subjects were offered cash compensation for their participation.

Demographic question blocks

The experiment consisted of six blocks: five demographic question blocks and one "word-use" block. The following demographic question prompts were presented in blocks of about 100 tokens each, in three different orders (see Table 4.2 below for ordering scheme):

1. What is this person's age?
    - Endpoints: "20 years old" & "40 years old"
    - Midpoint: "30 years old"
2. Where was this person born?

- Endpoints: “Definitely in Minnesota” & “Definitely not in Minnesota”
- Midpoint: “Not sure”

3. What is this person’s ethnicity?
    - Endpoints: “Definitely Hmong” & “Definitely white”
    - Midpoint: “Neither; not sure”
4. Are most of this person’s friends Hmong?
    - Endpoints: “Yes, most are definitely Hmong” & “No, most are definitely not Hmong”
    - Midpoint: “Not sure”
5. Is this person proud of their family heritage?
    - Endpoints: “Yes, definitely proud” & “No, definitely not proud”
    - Midpoint: “Not sure”

All answers to question prompts for the perception experiment were elicited using the VAS format. For example, the question about speaker ethnicity was presented to participants using the prompt screen shown in Figure 4.1 above. All prompt screens are provided in Appendix D.

Word-use block

After the five demographic question blocks, participants were asked to judge whether people of Hmong ethnicity were more or less likely to say certain words, compared to whites. The purpose of this portion of the experiment was to determine whether any words biased listeners to assume speaker ethnicity based on the perceived likelihood that a certain ethnicity uses the word more often. The ten words chosen are listed in Appendix D. Judgments were again elicited using the VAS format. The two endpoints of the VAS line were: “Used more by Hmong ethnicity” and “Used more by white ethnicity.” The midpoint was: “Used equally by both.” An example prompt screen is provided in Figure D.6 of Appendix D. No audio was played during this portion of the experiment.

A word on terminology used

In the perception experiment, the words used to refer to the two ethnicity groups in question were "Hmong" instead of "Hmong American" and "white" instead of "European American." These terms, although not as accurate, were preferred for their shorter reading (and thus processing) times, compared to the more correct terms (New, Brysbaert, Segui, Ferrand, & Rastle, 2004; Rayner, 2009). Furthermore, I gathered from the interviews that among Hmong Americans themselves, the most widely used ethnic term is "Hmong" rather than "Hmong American."

Speaker groups and listener groups

Speakers were split into two groups (Groups 1 and 2), counterbalanced for ethnicity, sex, and length of time in Minnesota. Table 4.1 shows these speaker groups. Listeners were divided randomly into six groups (1a, 1b, 1c, 2a, 2b, and 2c). Three listener groups (1a, 1b, and 1c) heard tokens from speakers in Group 1, and three listener groups (2a, 2b, and 2c) heard tokens from speakers in Group 2. Prompt questions were blocked in three different random orders (a, b, and c), so listeners in Groups 1a and 2a had the same question order, as did 1b and 2b, and so on, as shown in Table 4.2.

Table 4.1: Speakers divided into two counterbalanced groups

	Group 1	Group 2		Group 1	Group 2
Hmong Am.	Blue	Chee	**Euro. Am.**	Ashley	Elizabeth
	Cynthia	Gina		Emma	Jessica
	Gaozoua	Green		Erica	Kid
	Jiru	Jeff		Halden	Thursday
	Julia	Max		Rachel	
	Mai	Sheng			
	Morgan	Spring			
	Pang	Sunny			
	Ruby	Tasha			
	Yue	Wendy			
		Yoli			

Table 4.2: Three question block orders for perception experiment

Block number	Order **a**	Order **b**	Order **c**
1	Born	Age	Friends
2	Ethnicity	Heritage	Age
3	Heritage	Born	Born
4	Friends	Ethnicity	Heritage
5	Age	Friends	Ethnicity

4.2 Predictions

4.2.1 Inter- and intra-ethnic speech perception

I predicted that listeners would identify the ethnicity of the speakers at an above-chance rate, confirming that there are perceptible differences between the speech of the two ethnicities. I also predicted that the words in the STAND, CAB, and GOOSE vowel classes would enable more accurate ethnicity judgments from listeners compared to LOT and THOUGHT vowel classes, because the trends observed in Chapter 3 indicated that the vowel means for the first three vowel classes are acoustically different between the two ethnicity groups.

Additionally, I predicted that Hmong American listeners would be better at identifying Hmong American speakers than would listeners of other ethnicities, because they have more exposure to one another's speech and can thus better differentiate HA speech from non-HA speech. This prediction may seem indisputable, but there are surprisingly few empirical illustrations in sociophonetic literature of the role of previous exposure on speech perception (see Eckert, 2008).[1]

[1] One such empirical illustration (Graff, Labov, & Harris, 1986) found that previous exposure appeared to have caused African American listeners to become more inaccurate in identifying whether a speaker was African American. They speculated that this finding was due to task unnaturalness rather than previous exposure having an unexpected effect.

4.2.2 Perception of vowels involved in sound changes

Vowels believed to be involved in age-graded sound changes should differ by speaker's birthplace and age. But since it was found in Chapter 3 that some purported sound changes in the Twin Cities may not actually be taking place (anymore), it was likely that none of these questions would elicit clear differences for any words, no matter what vowel class.

4.2.3 The role of nasalization

In Chapter 3, I illustrated that there was a significant difference between HAs' pronunciations of the TRAP vowel in nasal and oral contexts, compared to EAs' pronunciations of the same vowel. Specifically, there was less separation of STAND and CAB (nasalized and non-nasalized sub-classes of TRAP) for HA speakers than for EA speakers.

Therefore, the tokens of STAND with a higher F1 and a higher F2 (lower and less front) should correlate with a tendency to judge speakers as Hmong, if indeed this characteristic is used by listeners to differentiate Hmong from non-Hmong in real life. Furthermore, *A1-P1* was found to be significantly different between ethnicities for the CAB vowel, so tokens of CAB with higher *A1-P1* values should elicit more judgments of "Hmongness."

4.2.4 The role of birthplace

Intra-ethnically, place of birth was found to be an important factor in speech production among Hmong Americans. Specifically, for the STAND vowel, HAs born in Minnesota tended to pattern more closely with EAs born in Minnesota, compared to their HA peers not born in Minnesota. If listeners are aware of this (at a subconscious level), it would follow that listeners' perception of speakers' birthplace will be affected by the formant values of each STAND token. Listeners should judge lower F1 and lower F2 tokens to be produced by EA speakers and HA speakers born in Minnesota, compared to very different tokens produced by HA speakers not born in Minnesota.

For the LOT and THOUGHT vowels, on the other hand, HA speakers born outside of Minnesota more closely resembled the EA speakers, who tended to have less separation between the two vowels, compared to HA speakers born in Minnesota.

4.2.5 The role of word-use judgments

If a word is judged to be more likely to be used by a certain ethnicity, then the overall judgments of Hmong talker ethnicity for that word should be higher than for words judged to be more likely to be used by whites.

4.3 Auxiliary perception study: Trained nasality judgments

An auxiliary experiment was conducted with five professional speech-language pathologists. All participants were between the ages of 18 and 50, native speakers of North American English, and had no history of speech, language, or hearing problems, according to self-report. This experiment was designed to provide a secondary measure of nasality in each of the tokens used in the main perception experiment.[2]

For most speech-language pathologists, typical levels of nasality observed on a daily basis are on a scale of normal to disordered. Here, the nasality differences between tokens were much more subtle, and participants were alerted to this unconventionality at the beginning of the experiment.

Participants completed this study individually, in a double-walled sound-proof booth at the University of Minnesota. As with the EA participants in the main perception experiment, instructions and question prompts were presented on a computer screen in front of the participants. Normalized audio files were played to listeners on over-ear headphones. Perceived nasality for every token (all words from the main perception experiment, from both speaker groups) was judged via the same basic protocol as the main perception experiment, using just one prompt screen, shown in Figure D.7 of Appendix D. Participants input their responses by clicking with a computer mouse on the Visual Analog Scale displayed on the computer screen, along with the prompt, "Please rate the nasality of this token." The endpoints of the VAS line were: "Not nasal at all" and "Extremely nasal." The midpoint was: "Moderately nasal." The experiment took about one hour to complete. At the end of the experiment, participants were asked

[2] Due to unforeseen equipment complications, the production data audio signal was not as clear as would be hoped for a completely reliable acoustic measure of nasality involving *A1*, *P0*, and *P1*, as in Chen (1997).

to fill out a demographic questionnaire very similar to the one mentioned in Section 4.1.

For recruitment, several emails were send to friends and friends-of-friends who were employed in Twin Cities' speech clinics.

Chapter 5

Perception Study Results

5.1 Overview of perception study

The central concern of the perception study is the presence or absence of detectable phonetic differences between Hmong Americans' English pronunciations and those of age-matched European Americans from the same region. As described in detail in Chapter 4, the perception study consists of two separate experiments completed by two different sets of participants: the main perception experiment which investigates listener perception of between- and within-ethnicity phonetic differences, and the auxiliary nasality experiment which provides a perceptual validation of the acoustic nasality measures for each token used in the main perception experiment.

Single-word audio tokens recorded by participants in the production study (who are listed in Tables 2.1 and 2.2 of Chapter 2) were played to a new set of HA and EA listeners for the main perception experiment. The listener-participants were asked to judge several demographic characteristics of the speakers, including speaker ethnicity, place of birth, age, friends' ethnicity, and pride in family heritage. Judgments were elicited using a VAS line on a computer screen. Only one demographic question was answered per block, for a total of five question blocks of about 100 tokens each. The data analysis for the main perception experiment is discussed in Section 5.2 and the results in Section 5.3.

At the end of the main perception experiment, several questions were asked to ascertain listeners' opinions on whether certain words were more or less likely to be

used by people of Hmong ethnicity and white ethnicity. Section 5.4 provides the results of this word-use sub-study.

For the nasality experiment, five trained speech-language pathologists listened to the same tokens used in the main perception experiment and were asked to judge the nasality of each token. They answered the same question for all tokens (about 1000 total). The results of the auxiliary nasality experiment are given in Section 5.5.

5.2 Main experiment data analysis

5.2.1 Responses and distributions

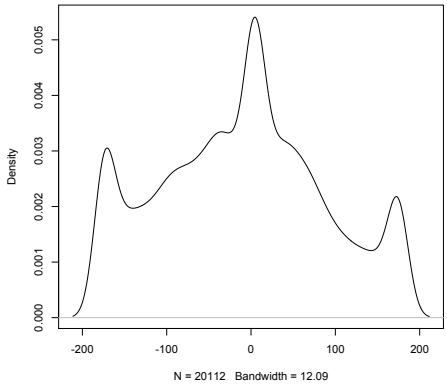
The response variable here is the horizontal position on the VAS line of every mouse click for all answers to the demographic questions in the main perception experiment. Of all mouse clicks from all participants in the main perception experiment (n = 20115), there were only three mouse clicks more than 70 pixels above or below the VAS line. Those three clicks were excluded from all subsequent analyses, due to the fact that they were likely made in error, as compared to the clicks that happened within 70 pixels of the line. For all questions other than the "age" question, the response variable was then converted to a zero-centered continuous number, from -175 (indicating the left endpoint of the line) to 175 (indicating the right endpoint of the line). For the age question, the response variable was scaled to accurately represent the approximate speaker age that listeners indicated on the VAS line for each token.

The overall response distribution was continuous but not fully normally distributed, as shown in Figure 5.1. Rather, the distribution is tri-modal, with the three modes corresponding to the three visible landmarks on the VAS line: two endpoints and a midpoint. Typical transformations (such as log and negative inverse) only made the distributions less normal, and trimming the endpoints would have eliminated many crucial datapoints, so no transformations were made on the response variable.

5.2.2 Predictor variables

Of primary importance in this study is *which speakers* are judged to sound Hmong American. If those speakers judged to sound HA are *actually* HA, we have evidence

Figure 5.1: Density plot for main perception experiment response variable



that there are real phonetic differences between HAs and EAs that clue listeners into a useable linguistic distinction along ethnic lines. Secondarily important is *which listeners* are better at identifying HA speakers and *which vowels* allow the most accurate speaker ethnicity judgments. These questions will be addressed in the main speaker ethnicity model in Section 5.3.

We would also like to know *why* speakers are identified as HA or EA. There are many possible explanations, including social and phonetic ones. Speaker ethnicity judgments could be due to speakers' presumed or actual age, presumed or actual place of birth, presumed or actual social group, presumed or actual pride in their ethnicity, use of sociophonetic nasalization, or any number of other possible speaker or listener characteristics. These speaker characteristics are tested for their influence in Section 5.6.

Linear mixed effects models were built with and without each potential random-effects factor. To determine whether each factor should be included in the final model, anovas were conducted comparing the model fit between the reduced models and the full model. If there was a significant difference between the reduced and the full model, the random effect in question was kept in the final model, but if there was no difference between the reduced and full model, the random effect was removed from the final model.

Listener, *speaker*, *item*, *trial index*, and *trial index by subject* were originally included in the main models as random-effects factors. Only the anovas testing *trial index by subject* did not reach significance, so that factor was removed from the final models. *Speaker ethnicity* and *listener ethnicity* were tested as a between-subjects fixed-effect factor to determine whether the speaker's actual ethnicity influenced listeners' responses about perceived speaker ethnicity, and whether listeners were better at correctly identifying speaker ethnicity when it was the same as their own. *Speaker birthplace* and *listener birthplace* were also tested for similar reasons. *Vowel class* was tested as a fixed-effects factor to determine whether certain vowels were more likely than others to differ sociophonetically between the two speaker ethnicities.[1]

Model coefficients for all fixed-effects factors are provided in Appendix E. Readers may wish to refer to the full models while reading the next sections.

5.3 Results

A linear mixed-effects model was created for each of the five questions asked in the perception study to model the response variable using the predictor variables discussed in Section 5.2.2 above.[2]

5.3.1 Ethnicity question

Listeners answered the question "What is this person's ethnicity?" by clicking on the VAS line whose endpoints were -175 (labeled "Definitely Hmong") and and 175 (labeled "Definitely white"). The midpoint of 0 was labeled "Neither; not sure."

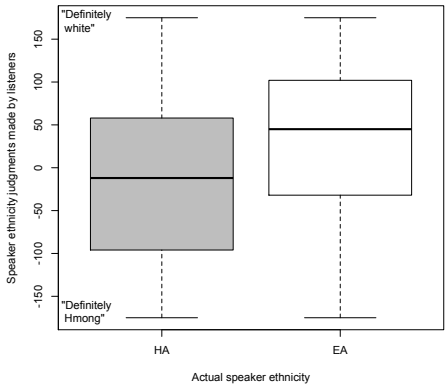
Most importantly, it was found that listeners judged HA speakers' ethnicity significantly differently from how they judged EA speakers' ethnicity, and both speaker groups' ethnicities were judged in the expected (i.e., correct) directions. The model coefficient for the main effect of speaker ethnicity was significant at $p<0.01$. Median VAS responses for the "ethnicity" question from all listeners were -12 for the HA speakers

[1] A series of preliminary analyses explored whether response time was meaningfully related to any of the random or fixed effects. In general, listeners tended to respond faster over time as the experiment progressed, as one might expect. However, there was no significant influence of any of the fixed effects on RTs, so it was removed from the models.

[2] See Drager (2011) for a thorough explanation of the mixed-effects linear regression method and a defense for its use in perception studies.

and 45 for the EA speakers (on the scale of -175 to 175 from "Definitely Hmong" to "Definitely white"). The main effect of speaker ethnicity is illustrated in Figure 5.2. The grey box in this plot represents HA speakers while the white box represents EA speakers.

Figure 5.2: Perceived speaker ethnicity responses, by actual speaker ethnicity

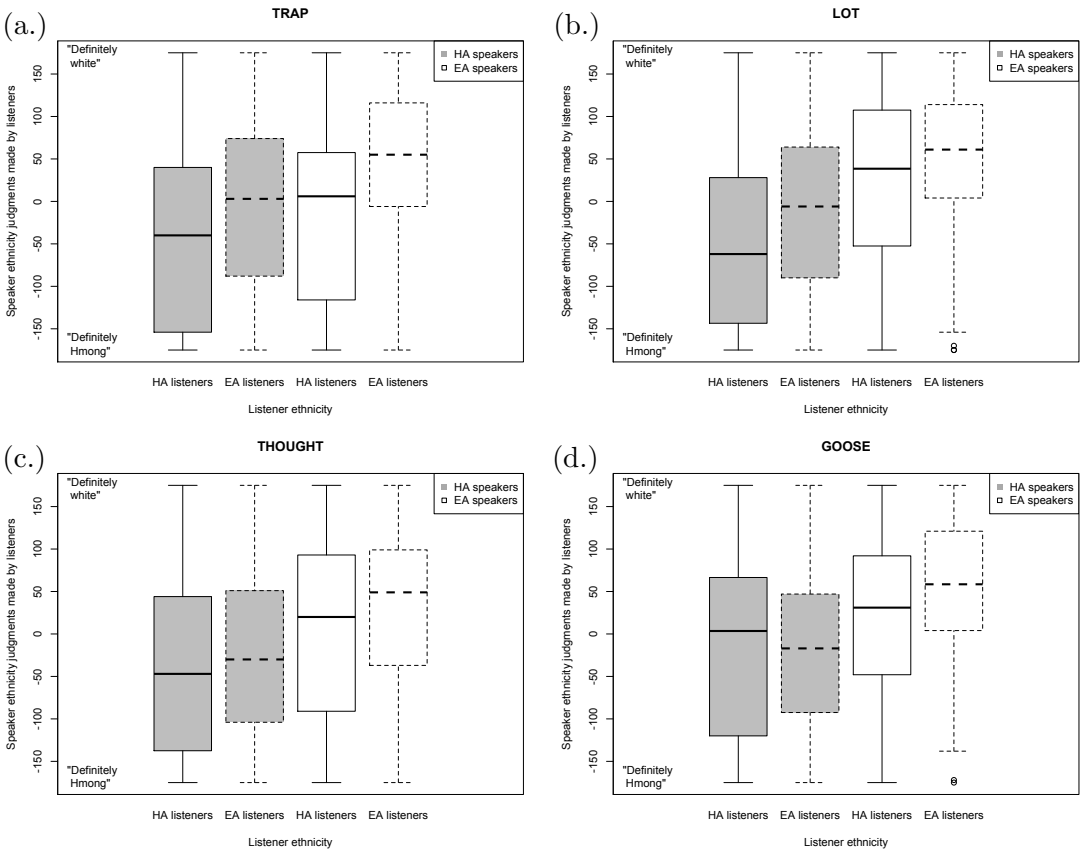


The model coefficients (found in Table E.12 of Appendix E) for the three-way interactions between speaker ethnicity, listener ethnicity, and vowel class were significant for the THOUGHT vowel and the GOOSE vowel, and marginally significant for the TRAP vowel, but not for the LOT vowel. The interactions are plotted for each of the four vowels in Figure 5.3. For these four plots, the solid-line boxes represent HA listeners and the dashed-line boxes represent EA listeners. As before, the grey boxes represent HA speakers and the white boxes represent EA speakers.

For the TRAP and THOUGHT vowels, the interaction plots in Figure 5.3 (a) and (c) indicate that EA listeners are more likely to label EA speakers as sounding more "white." This result appears to have arisen due to a combination of two forces: first, a bias in labeling any speaker a member of your own ethnicity regardless of the speaker's actual ethnicity, and second, a perceptual advantage in correctly identifying members of your own ethnicity.

For the GOOSE vowel, the interaction seems to be driven by an unexpected lack of accuracy in HA listeners' ability in identifying the ethnicity of HA speakers. In other

Figure 5.3: Perceived speaker ethnicity responses by listener ethnicity and actual speaker ethnicity, for the (a.) TRAP vowel, (b.) the LOT vowel, (c.) the THOUGHT vowel, and (d.) the GOOSE vowel



words, HA speakers don't sound as "Hmong" when pronouncing the GOOSE vowel, but only for HA listeners.

5.3.2 Age question

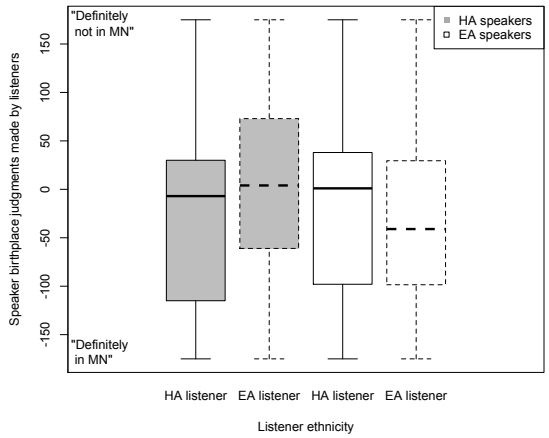
Listeners answered the question "What is this person's age?" by clicking on the VAS line whose endpoints were "20 years old" and "40 years old." The midpoint of 0 was labeled "30 years old." None of the coefficients in the model for this question were significant, as shown in Table E.13.

5.3.3 Born question

Listeners answered the question "Where was this person born?" by clicking on the VAS line whose endpoints were -175 (labeled "Definitely in MN") and 175 (labeled "Definitely not in MN"). The midpoint of 0 was labeled "Not sure." The model coefficients for the "born" question are found in Table E.14. The coefficient for the interaction between speaker and listener ethnicity was significant; see Figure 5.4 for the interaction plot.

Three of the four medians for the interaction between speaker and listener ethnicity are very similar, all very close to the midpoint of 0; but the fourth median, that of perceived birthplace judgments of EA listeners for EA speakers, seems to be driving this significant interaction. Here, EA listeners are rating EA speakers as being more likely to be born in Minnesota (regardless of whether the speaker was *actually* born in Minnesota). Another interpretation of this data is that HA listeners do not have a clear stereotype for what Minnesotans should sound like, whereas EA listeners do, and it includes something that the EA speakers were doing but the HA speakers were not.

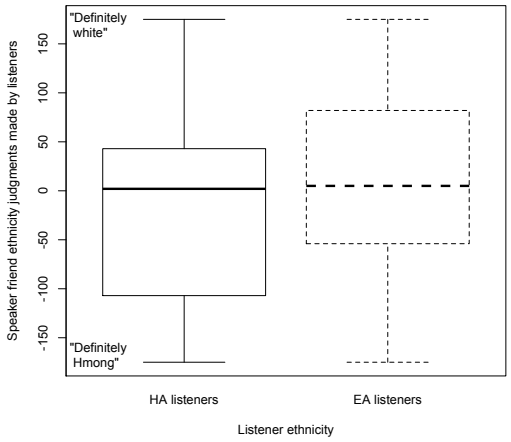
Figure 5.4: Perceived speaker birthplace responses, by speaker ethnicity and listener ethnicity



5.3.4 Friends question

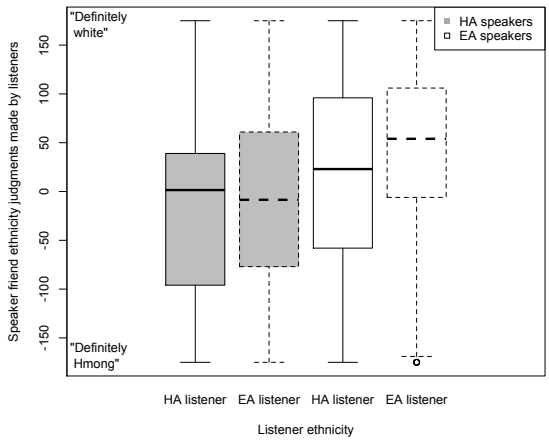
Listeners answered the question "Are most of this person's friends Hmong?" by clicking on the VAS line whose endpoints were -175 (labeled "Yes, most are definitely Hmong") and 175 (labeled "No, more are definitely not Hmong"). The midpoint of 0 was labeled "Not sure." The model coefficients for the "friends" question are found in Table E.15. The coefficient for the interaction between listener ethnicity and vowel class approached significance for the GOOSE vowel only. Average speaker friend ethnicity judgments by listener ethnicity are plotted for the GOOSE vowel in Figure 5.5. Both listener groups' median judgment hovered around zero, but, unsurprisingly, HA listeners' mean judgment indicated a tendency for listeners to categorize speakers' friends as sounding more Hmong, regardless of speaker ethnicity.

Figure 5.5: Perceived speaker friend ethnicity responses, by listener ethnicity, for the GOOSE vowel



The coefficient for the interaction between speaker ethnicity and listener ethnicity was also significant. In Figure 5.6, we see that, in general, listeners judge speakers' friends to be of the same ethnicity as speakers' actual ethnicity. However, EA listeners were more likely to engage in this pattern, compared to HA listeners who tended to click closer to the midpoint of the VAS line, indicating uncertainty or ambivalence.

Figure 5.6: Perceived speaker friend ethnicity responses, by speaker ethnicity and listener ethnicity



5.3.5 Heritage question

Listeners answered the question "Is this person proud of their family heritage?" by clicking on the VAS line whose endpoints were -175 (labeled "Yes, definitely proud") and 175 (labeled "No, definitely not proud"). The midpoint of 0 was labeled "Not sure." The model coefficients for the "heritage" question, none of which were significant, are found in Table E.16.

5.3.6 Response correlations between questions

A correlation matrix (shown in Table 5.1) was created for the mean responses to individual audio tokens, collapsed across listeners and ethnicity groups, for each of the five questions. According to this table, the responses to the "ethnicity" question were significantly correlated with all the other questions, especially the "born" question and the "friends" question, although the magnitude of the associations was weak. The correlation between the "ethnicity" question and the "born" question is negative, indicating that a "Definitely Hmong" response for one question correlated with a "Definitely not

in Minnesota" response for the other question. The correlation between the "ethnicity" question and the "friends" question goes in the expected direction: a "Definitely Hmong" response for one question correlated with a "Yes, most are definitely Hmong" response for the other question.

The other highly significant (negative) correlation was between the "born" question and the "friends" question, completing the triangle of correlation between three questions: "ethnicity," "born," and "friends."

Table 5.1: Correlation table for average VAS rating for each question type

	Ethnicity	Age	Born	Friends
Age	0.11**			
Born	-0.30***	0.04		
Friends	0.38***	0.06	-0.32***	
Heritage	-0.07*	0.06	0.09**	-0.07*

Overall, the "ethnicity" question elicited the strongest responses from listeners. The other four questions ("age", "born","friends", and "heritage") did not seem to be as relevant for listeners (or overlapped quite a bit with the ethnicity question) other than one interesting finding for the "born" question, discussed in Section 5.3.3, and one ineteresting finding for the "friends" question, discussed in Section 5.3.4.

From the main perception experiment, we can gather that there was a perceptible phonetic difference between HA speakers and EA speakers, and that the ethnicity of the listener affected how perceptible those phonetic differences were. Namely, if the ethnicity of the speaker matched the ethnicity of the listener, then the listener was more likely to respond in the expected (correct) direction.

5.4 Auxiliary study: word use

A subset of 10 of the total words tested were included in a short auxiliary experiment at the end of the main perception experiment. The words included: *ancestor*, *bamboo*, *coffee*, *daughter*, *father*, *mom*, *moon*, *new year*, *pet*, and *song*. These words were judged

by listeners for their likelihood to be used by Hmong Americans or whites, using a similar VAS line to elicit judgments.

A linear mixed-effects model was constructed, using word-use judgment as the response variable and listener ethnicity and word as fixed-effects predictor variables (see Table E.19 in Appendix E). According to the model, the two-way interactions between listener ethnicity and two different words (*mom* and *new year*) had significant coefficients. The main effects of several words also had significant coefficients, including *coffee*, *father*, *moon*, *pet*, and *song*.

Table 5.2 lists the mean likelihood of use judgments for each word, where 0 = "Used equally by both," negative = "Used more by HAs," and positive = "Used more by whites"). Figures 5.7 and 5.8 provide this information in graphical form.

Table 5.2: Mean word use judgments made by perception study listeners

Word	HAs	EAs	Grand mean
ancestor	-109.9	-105.9	-107.1
bamboo	-137.6	-88.6	-103.7
coffee	125.5	119.0	120.7
daughter	-30.3	-36.3	-34.8
father	-9.6	-66.4	-52.2
mom	-29.0	85.4	59.2
moon	-10.7	-35.3	-29.1
new year	-103.9	-2.7	-28.7
pet	112.2	78.8	88.1
song	18.3	17.1	17.4

The word-use judgment auxiliary experiment was created to determine if listeners' responses to questions in the main perception experiment were influenced by whether they thought a word was more likely to be used by Hmong Americans or European Americans. A linear mixed-effects model was created to test this (details of the model are shown in Table E.20 of Appendix E).

Figure 5.7: Word-use judgments by word and listener ethnicity, Part 1

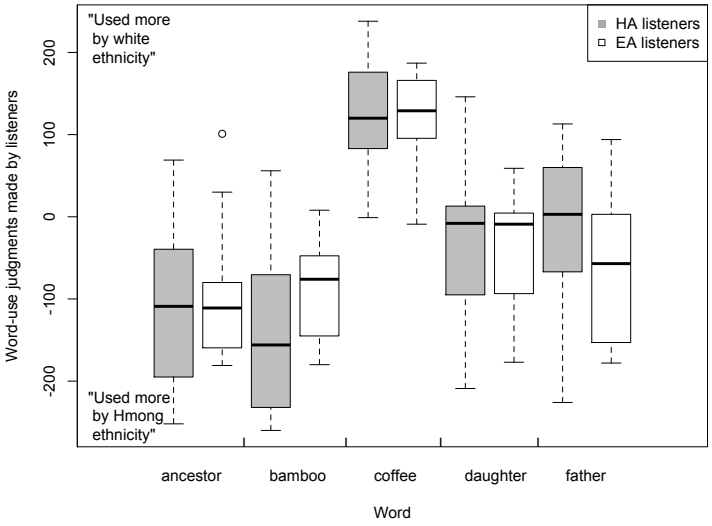
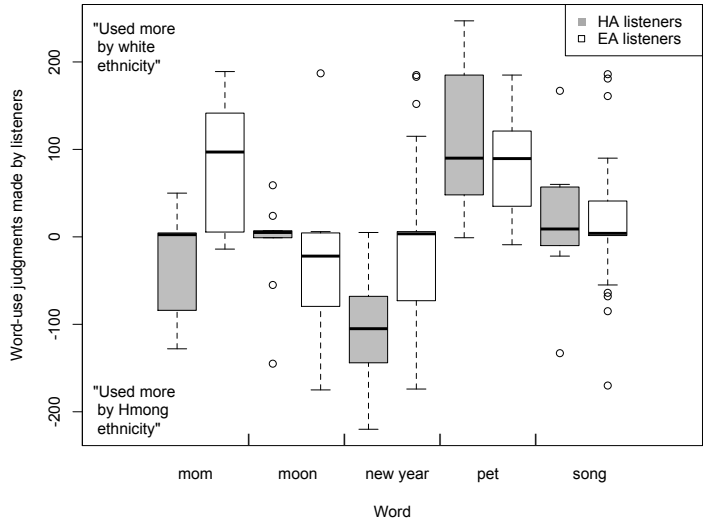


Figure 5.8: Word-use judgments by word and listener ethnicity, Part 2



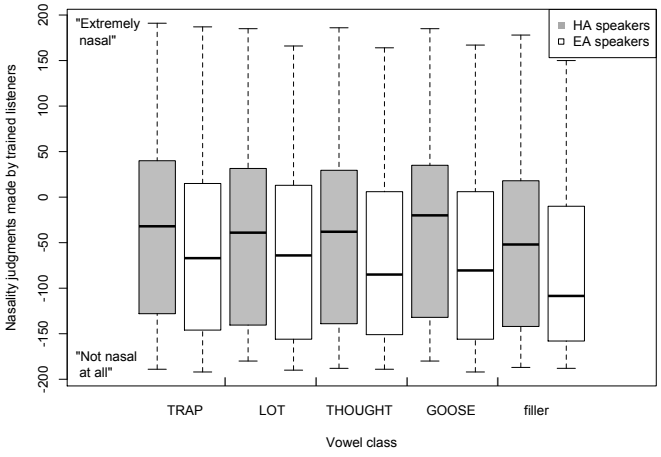
The model coefficients for the main effect of word-use judgments and all interactions involving word-use judgments are not significant in the model of listener judgments of speaker ethnicity. Therefore, we have evidence that listeners' biases of which ethnicities use which words more do not seem to affect listeners' judgments of speaker ethnicity for those words.

5.5 Auxiliary study: nasality

Two linear mixed-effects models were fit using speaker ethnicity and vowel class as fixed-effects factors; and subject (listener), token, speaker, and trial index as random-effects factors. The first model included all four vowel classes and filler words. The second model included only the words in the TRAP vowel class, split into their CAB and STAND sub-classes. These models are provided in Appendix E, Tables E.17 and E.18.

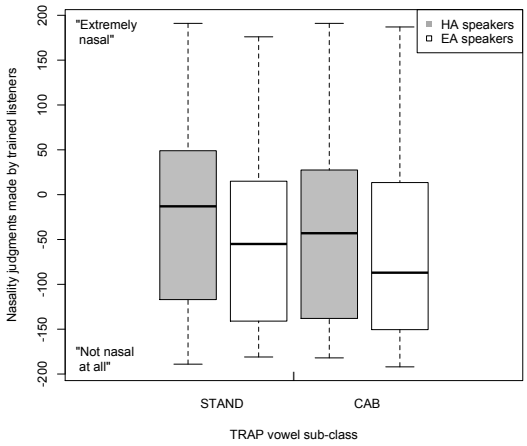
For the first model, the coefficient for the main effects of listener ethnicity and vowel class were found to be significant. The plot in Figure 5.9 illustrates that, for all vowels, HA speakers' productions are more likely to be rated as having more nasality compared to EA speakers' productions.

Figure 5.9: Nasality judgments by vowel class and speaker ethnicity: all vowels



For the second model, the coefficient for the main effect of vowel class (either STAND or CAB) was significant. The plot in Figure 5.10 shows that, unsurprisingly, CAB tokens are rated as having lower nasality than STAND tokens.

Figure 5.10: Nasality judgments by vowel class and speaker ethnicity: TRAP vowel only



The fact that HAs were rated as more nasal across all vowels suggests that either HAs *do* produce more nasalized vowels, despite the presence of phonemic nasal vowels in their heritage language,[3] or that the trained speech-language pathologists who made the nasality ratings were basing their judgments on other phonetic cues that differed between ethnicities, such as, for example, voice quality or fundamental frequency. Although the listeners are clinically trained to judge nasality, the scale on which they do so is usually much more extreme. Therefore, the task was at least somewhat unnatural for them, and this may have caused them to utilize other similar cues in addition to nasalization, in order to increase discriminability between tokens.

[3] It is not likely that HAs actually produced all vowels with more nasalization, given the results of the production study (given in Section 3.2.3 of Chapter 3), showing that nasalization measurements overall did not differ significantly between ethnicities.

5.6 Production–perception dynamics

In this section, I will integrate the perception results for the ethnicity question discussed above with some of the acoustic measures from the production study. The goal of this section is to elucidate some of the acoustic cues that listeners use when making judgments about a speaker's ethnicity.

Linear models predicting estimated speaker ethnicity judgments were constructed for each vowel class with random intercepts for listener, speaker, token, and trial index. Fixed effects included the first and second formant values and degree of nasality.[4]

5.6.1 The LOT and THOUGHT vowels

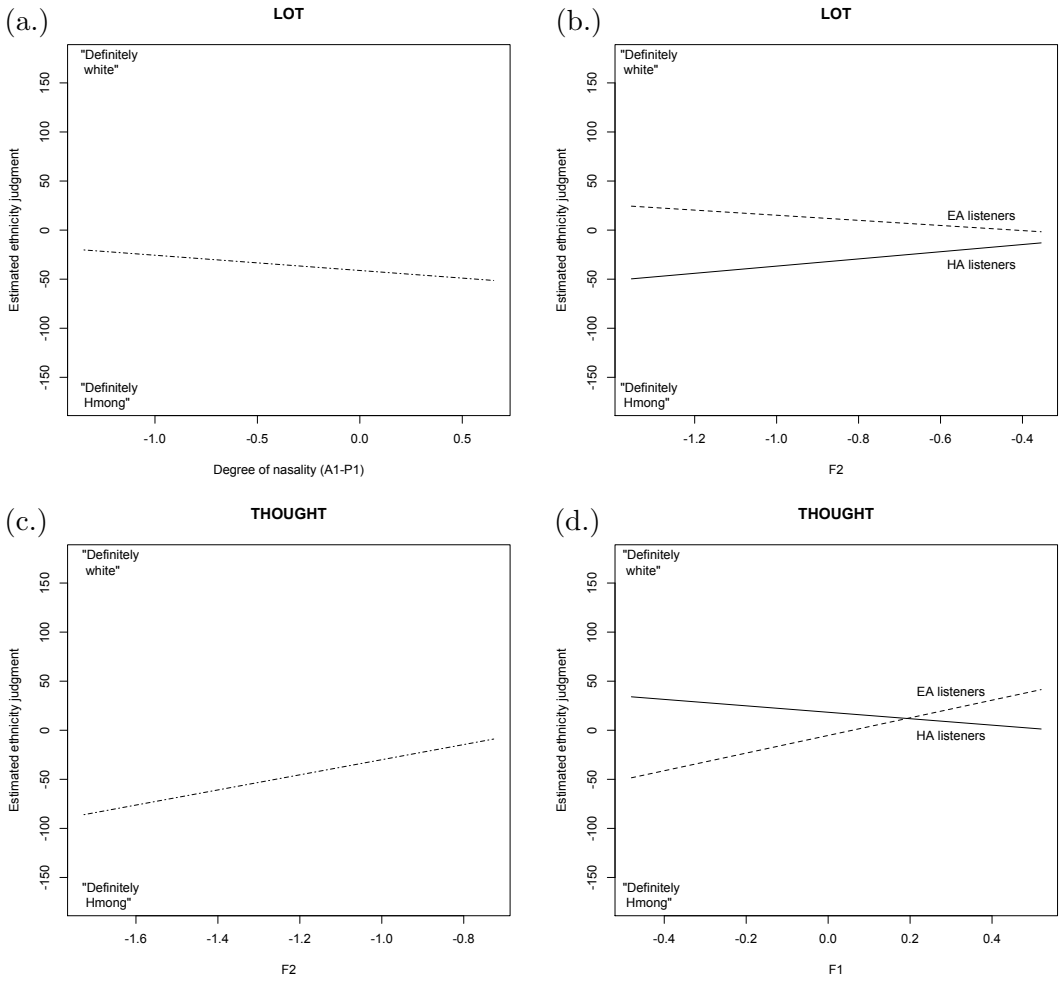
For the vowels involved in the low back merger, listener ethnicity, listener birthplace, nasality, and formants values contribute to the estimation of speaker ethnicity judgment responses by listeners. The models for the LOT and THOUGHT vowels are provided in Tables E.24 and E.25 of Appendix E, and the significant coefficients are illustrated in Figure 5.11.

Nasality did not contribute statistically significantly to the models of the LOT and THOUGHT vowels. For the LOT vowel, the coefficient for the main effect of *A1-P1* approached significance, shown in Figure 5.11 (a). For this vowel, greater nasality led to a trend towards more judgments of white speaker ethnicity. No other coefficients involving nasality approached significance for either vowel.

The coefficient for second formant values was significant for both the LOT and THOUGHT vowels. For LOT, there was a significant interaction between listener ethnicity and F2, with more back tokens eliciting judgments of one's own ethnicity, illustrated in Figure 5.11 (b). For the THOUGHT vowel, judgments were heavily skewed toward "Hmong" judgments, but the most front tokens of THOUGHT approached the midpoint of the VAS line, as shown in Figure 5.11 (c).

[4] For the acoustic measure of nasality, both *A1–P0* and *A1–P1* (Chen, 1997) may theoretically contribute some predictive power to the models. Therefore, each model will include one or both nasality measures if they are meaningful for that particular model.

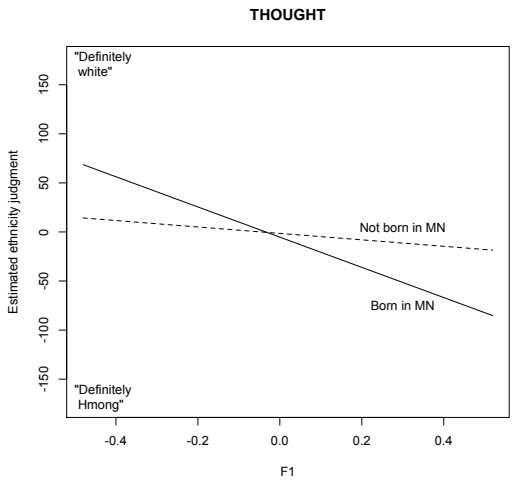
Figure 5.11: Estimated speaker ethnicity responses for (a.) the main effect of *A1-P1* for the LOT vowel, (b.) the interaction of listener ethnicity and F2 for the LOT vowel, (c.) the main effect of F2 for the THOUGHT vowel, and (d.) the interaction of listener ethnicity and F1 for the THOUGHT vowel



The interaction between listener ethnicity and F1 for THOUGHT (Figure 5.11 (d)) was unique because it was the only significant interaction wherein listeners were more likely to judge speaker ethnicity to be different from their own, especially if the token had a lower F1. Only the lowest THOUGHT vowels (thus with higher F1 values) elicited speaker ethnicity judgments similar to listeners' own ethnicity. This interaction is driven primarily by the EA listeners' responses; the relationship between F1 and listener ethnicity for HA listeners was not as strong.

The coefficient for listener birthplace, in an interaction with F1, approached significance for the THOUGHT vowel. This interaction, driven by listeners born in Minnesota, is illustrated in Figure 5.12. Those listeners born in Minnesota seemed to use F1 more often as a cue for speaker ethnicity than those listeners not born in Minnesota. For listeners born in Minnesota, a higher token of THOUGHT (lower F1 value) led to more judgments of "Hmongness."

Figure 5.12: Estimated speaker ethnicity responses for the interaction of listener birthplace and F1 for the THOUGHT vowel



5.6.2 The TRAP vowel

Formant values, nasality, and listener ethnicity all played a part in accurately predicting how listeners judged speaker ethnicity for the TRAP vowel, including its STAND and CAB sub-classes. The model coefficients for the TRAP model are provided in Table E.21 of Appendix E. For the model of the TRAP vowel itself (taking STAND and CAB together), only the fixed factor for listener ethnicity approached significance. We know from the previous section that listeners tend to judge speakers' ethnicity as being the same as their own, so this finding is not surprising. Instead, breaking the vowel up into its STAND and CAB sub-classes provides a better picture of what listeners use to make their judgments.

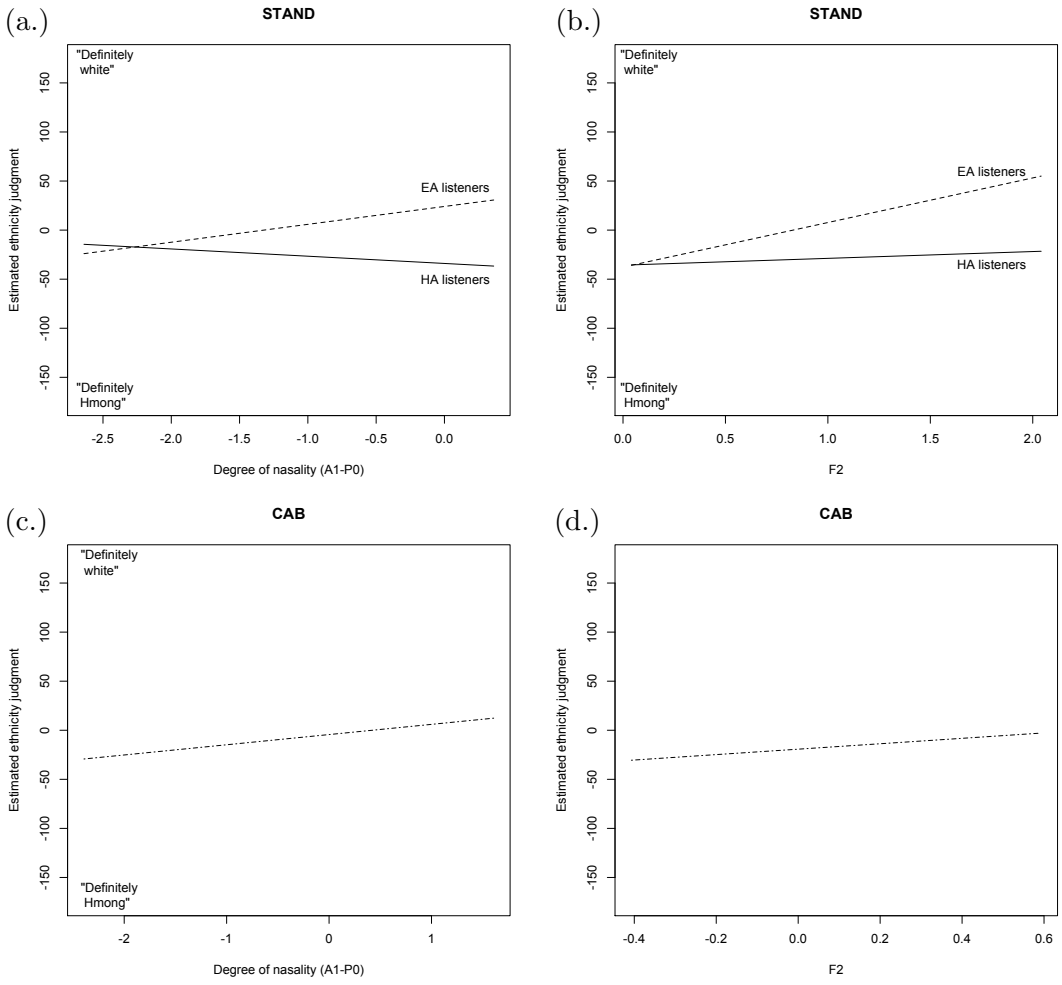
The STAND and CAB sub-classes of the TRAP vowel are expected to behave differently from one another. Therefore, two additional models were run, using only the STAND vowels in one model and only the CAB vowels in the other. The model coefficients for these are shown in Tables E.22 and E.23. For the STAND vowel model, both listener ethnicity groups judged the most front, least nasal tokens to be produced by speakers of the listeners' own ethnicity. The coefficient for the two-way interaction between listener ethnicity and degree of nasality (measured by *A1-P0* here) approached significance, and is illustrated in Figure 5.13 (a). Degree of nasality decreases as *A1-P0* and *A1-P1* increase, so we can see that as degree of nasality increases in this interaction, speaker ethnicity judgments by listeners of both ethnicities converge near the midpoint of the VAS line. When a STAND token has a low degree of nasality, EA listeners are more likely to judge a speaker as being the same ethnicity as themselves. The line for HA listeners is nearly flat in this figure, so there is a very weak relationship between listener ethnicity and *A1-P0* for HA listeners. The coefficient for interaction between F2 and listener ethnicity of STAND, illustrated in Figure 5.13 (b), also approached significance, and again this interaction was driven by EA listeners' responses. As F2 increases (resulting in a more front vowel), HA listeners' judgments do not change, whereas EA listeners tend to judge more front vowels as being produced by EA speakers. Because HA listeners' ethnicity responses overall were accurate, but nasality and F2 do not seem to affect their responses, we have evidence that HA listeners are using another phonetic cue not measured in this study (such as voice quality) to determine speaker ethnicity.[5]

[5] "Voice quality" is commonly thought to be a perceptible difference between ethnicities, even for

For the CAB vowel model, no interactions contributed a significant amount of predictive power. However, the coefficient for the main effect of nasality (*A1-P0*) was significant. Figure 5.13 (c) illustrates that tokens with more nasality tended to elicit slightly more judgments of "Hmongness," unlike for the LOT vowel. The coefficient for the main effect of F2 also approached significance. This is illustrated in Figure 5.13 (d): the more fronted CAB tokens were slightly more likely to elicit judgments of "whiteness."

untrained listeners. However, voice quality is difficult to quantify, and very few acoustic studies measure voice quality accurately and successfully.

Figure 5.13: Estimated speaker ethnicity responses for (a.) the interaction of listener ethnicity and *A1-P0* for the STAND vowel, (b.) the interaction of listener ethnicity and F2 for the STAND vowel, (c.) the main effect of *A1-P0* for the CAB vowel, and (d.) the main effect of F2 for the CAB vowel

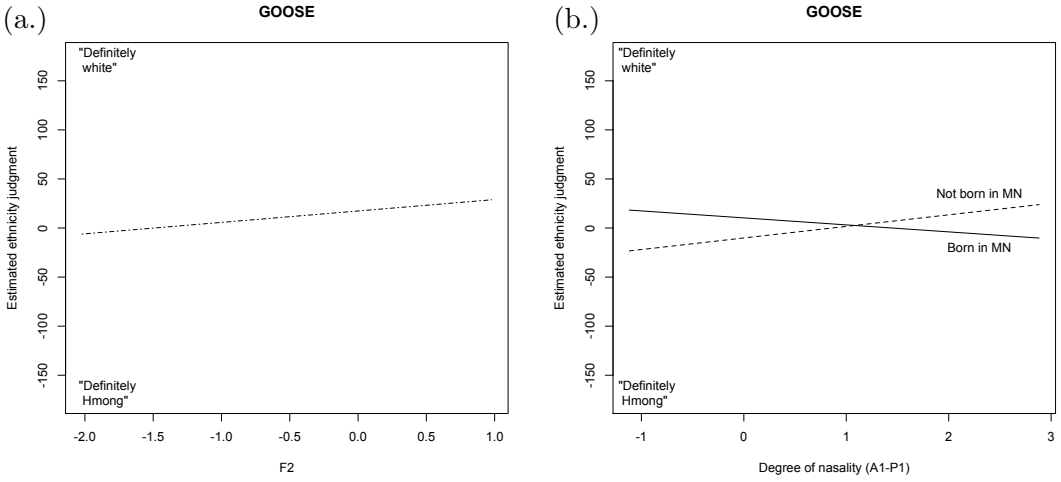


5.6.3 The GOOSE vowel

The model for the GOOSE vowel is provided in Table E.26. There was a significant coefficient for the main effect of F2, as expected. Judgments of "white" speaker ethnicity were more likely to be elicited by more front GOOSE tokens. This is illustrated in Figure 5.14 (a).

The coefficient for the interaction between listener birthplace and *A1-P1* was also significant. This is shown in Figure 5.14 (b). The tokens with extreme *A1-P1* values (either very nasal or not nasal at all) elicited responses closer to the endpoints of the VAS line. A higher degree of nasality caused listeners from Minnesota to tend to hear more "white" speakers, but listeners not from Minnesota tended to hear more "Hmong" speakers for the same tokens.

Figure 5.14: Estimated speaker ethnicity responses for the GOOSE vowel. Figure (a.) shows the main effect of F2 and Figure (b.) shows the interaction between listener birthplace and degree of nasality (*A1-P1*)



5.7 Summary

The perception study reported on in this chapter included three sub-studies completed by two different sets of listener-participants. The first sub-study is referred to as the "main perception experiment," and was completed by 12 Hmong Americans and 27

European Americans. Participants rated single-word audio tokens for five different questions about speakers' demographic profiles. The second sub-study, the "auxiliary word-use study," was completed by the same group of participants, who rated words on their likelihood to be used more by either Hmong Americans or European Americans. The third sub-study was completed by a new set of participants, five clinically-trained speech-language pathologists. In the "auxiliary nasality study," these participants rated the nasality of the same single-word audio tokens from the main perception experiment.

In Section 5.3, the results of the main perception experiment were reported by question. The question asking about speaker ethnicity elicited the strongest responses from listeners. The responses were overall quite accurate: HA speakers were usually judged to sound more like they were HA, and EA speakers like they were EA. However, there was a general bias for listeners to choose their own ethnicity, so when responses are broken down by listener ethnicity, we see that response means for HA listeners are closer to the "Hmong" end of the spectrum, regardless of speaker ethnicity, compared to response means for EA listeners, and vice versa.

Three questions were weakly correlated with one another: the "ethnicity" question, the "born" question, and the "friends" question. These three questions also produced significant results, unlike the other questions ("age" and "heritage"). For the "born" question ("Where was this speaker born?"), HA listeners seem not to have a clear stereotype of what Minnesotans sound like, whereas EA listeners do, and they judge EA speakers as sounding more Minnesotan than HA speakers, for all vowel types.

Several vowel-specific trends emerged from the data. For the TRAP and THOUGHT vowels, listeners followed the general same-ethnicity bias and the trend to identify ethnicity accurately. For the GOOSE vowel, HA listeners showed an unexpected lack of accuracy in identifying HA speakers, whereas EA listeners showed the expected accuracy and same-ethnicity bias. Additionally, the GOOSE vowel was the only vowel class to elicit a significant same-ethnicity bias from the "friends" question ("Are most of this person's friends Hmong?").

The word-use experiment was reported on in Section 5.4. It was found that seven out of ten words were significantly likely to be judged to be used more by one ethnicity or the other; two of those words (*mom* and *new year*) were judged differently by the two listener ethnicity groups, perhaps due to EAs' relative lack of knowledge about HA

culture. However, none of these word-use biases had a significant effect on how listeners judged speaker ethnicity in the main perception experiment.

In Section 5.5, results of the nasality experiment were reported. Overall, trained listeners tended to judge HA speakers as sounding more "nasal" than EA speakers. It is likely that the listeners in this study were listening for more than just subtle dialectal nasalization. Because their jobs require them to listen for nasality that is more extreme, they may have used other features of voice and speech to aid in perception of differences between tokens.

Finally, in Section 5.6, the perception results discussed above were integrated with the production measures from Chapter 3. Formants values, nasality, listener ethnicity, and listener birthplace were found to significantly contribute to the prediction of listeners' perception experiment responses. For the vowels involved in the low back merger (LOT and THOUGHT), it was found that the furthest back tokens of THOUGHT (most separated from LOT) were more closely associated with HA speakers, supporting Ito's (2010) finding that HAs' low back vowels are more separated than EAs'. The CAB subclass of the TRAP vowel was judged to be more "Hmong"-sounding as nasality increased and F2 decreased, contrary to my predictions, but consistent with the results of the production study in Chapter 3. The GOOSE vowel, as expected, was judged to be more "white"-sounding for the most front tokens.

Furthermore, listener ethnicity or birthplace interacted significantly with several acoustic measures for the low back vowels, the STAND vowel, and the GOOSE vowel, showing that listeners' background also influenced their perception of speakers' phonetic features, and in turn, their perception of speakers' ethnicity.

Chapter 6

Discussion and Conclusion

6.1 General summary

This dissertation is a sociophonetic analysis of the English spoken by Hmong Americans living in the Twin Cities. Minnesota is home to 66,181 Hmong Americans—more than any other state in the U.S., other than California. The Twin Cities has the largest urban population of Hmong Americans in the U.S. (ACS, 2009). Many Hmong Americans who live in the Twin Cities, especially the elderly and newly-arrived, interact primarily or exclusively with other Hmong Americans. I chose to investigate a group of college-age Hmong Americans who are fluent in English and Hmong, and who interact with other Hmong Americans as well as non-Hmong Americans, to allow for potential linguistic influence from both social groups in the target population. It was expected that the inferred tight-knit nature of Hmong Americans' social networks would cause a slower uptake of current regional and supra-regional sound changes versus the comparatively looser networks of many European Americans in the Twin Cities. Furthermore, the target population should presumably experience some influence in their English from the Hmong language. Crucially for this study, the Hmong language has phonemic nasal vowels whereas English does not. This L2 influence of phonemic nasal vowels was hypothesized to emerge in Hmong Americans' English as less nasalization overall, and this should decrease the likelihood that they will engage in the Northern Cities Shift.

In Chapter 1, the background necessary for understanding the rest of the dissertation was presented. This included information about the study of sociolinguistics

and sociophonetics in particular; Asian American sociolinguistics; the history of the Hmong American diaspora and the Hmong language; and the history and current state of American English in Minnesota, especially with regards to participation in regional and supra-regional sound changes in progress. In Chapter 2, I outlined the experimental procedure for the collection of phonetic data ("the production study") from Hmong Americans and age-matched European Americans from the same region. The fieldwork was conducted in a small, quiet room at the University of Minnesota. Participants (i.e., "speakers") were aware that they were being recorded, but all efforts were made to put speakers at ease so as to collect the most natural tokens of speech as possible. Chapter 3 reported the results of the production study, including a comparison of current data with ANAE data collected over 16 years ago. European American speakers seem to be participating in a supra-regional sound change, the fronting of the GOOSE vowel, to a greater extent than in the past, and to a greater extent than Hmong Americans. Two other sound changes, the Northern Cities Shift and the low back merger, show inconclusive evidence of adoption by either EA speakers or HA speakers. Chapters 4 and 5 presented the methods and results of the laboratory component of this dissertation (called "the perception study"). The perception study, which was conducted with a new set of participants (i.e., "listeners"), aimed to uncover whether phonetic differences between Hmong Americans' and European Americans' vowel pronunciations are actually detectable. Words recorded during the fieldwork became the experimental items for the perception study. These words were played to listeners over headphones and rated on a visual analog scale by listeners on several different dimensions of speakers' social characteristics, including ethnicity. It was found that although certain expected phonetic differences were not used to make judgments of speakers' ethnicities, other phonetic differences, some expected and some not, did indeed predict listeners' judgments of speaker ethnicity. Taken together, the results of the two studies provide evidence that Hmong Americans' vowel pronunciations are not simply Hmong-influenced imitations of vowels as spoken by European Americans, and that listeners, especially other Hmong American listeners, can use these complex yet systematic phonetic patterns to make accurate decisions about speakers' ethnicities.

6.2 Discussion

Numerous studies have shown in the past that ethnicity is an important part of one's social and linguistic identity. Based on those previous results, I expected to find differences between the vowel pronunciations of Hmong Americans (HAs) and European Americans (EAs) of the Twin Cities sampled in this dissertation, especially for vowels involved in sound changes. I predicted that HA speakers would tend to produce slightly less-shifted variants compared to EA speakers, for two reasons: (1) HA networks are likely more isolated than EAs, and this would predict changes being adopted slower than in looser networks (Milroy, 1980; Milroy & Gordon, 2003), and (2) the outside links that Twin Cities-based HA networks have are generally to other HA networks, many of which are located outside of the Inland North.

Given recent sociolinguistic studies of the so-called second and third waves (see Section 1.2.1 of Chapter 1), I suspected that the production and perception of the vowels in question would be more nuanced than simple categorization into "Hmong American" and "European American" categories. In fact, my results showed that production of vowels is acoustically and articulatorily complex, and that perception involves more than a simple two-way choice between one ethnicity or the other.

In what follows, I provide a summary of the qualitative impressions of the social networks of Hmong American and European American participants, based on their responses to questions asked during the interviews. Then, I provide summaries of the results of the current investigations divided into expected and unexpected categories, and, if the results were unexpected, conjecture as to why these results were obtained.

6.2.1 Qualitative impression of Hmong American and European American social networks

A tight-knit social network is one with many interconnections between members and few connections to other networks, as can be found among residents of small, rural towns. Networks of this kind are usually slow to spread sound changes coming from other networks, so members of tight-knit networks often lag behind in sound changes in progress. Members of tight-knit networks are also particularly prone to developing pronunciation patterns that differentiate them from members of other networks; language naturally

changes over time, but in different directions for disconnected groups.

Previous research on Hmong and Hmong Americans' cultural and social practices (discussed in Chapter 1) indicate that Hmong Americans should have social networks that are very tightly knit, especially among the oldest members of the community. In this dissertation, I have assumed that Hmong Americans in the Twin Cities have more tightly knit social networks than European Americans, on average. Still, the structure of Hmong Americans' social networks is an empirical matter, and one that I have attempted to answer through the sociolinguistic interview questions listed in Appendix B. The following paragraphs in this section summarize the ethnographic findings of the study that relate to participants' social networks. To limit the scope of this chapter, as in other chapters, I will focus on characteristics that typify ethnicity groups rather than characteristics of individual speakers.

Hmong American participants usually had very close relationships with their immediate family. HAs also had close relationships with some extended family members, as well, even if those family members lived in another state. Extended family members often visited the Twin Cities for the Hmong new year celebration in the winter, or the popular Hmong soccer tournament in the summer. It was also common for HAs' out-of-state relatives to visit for family gatherings such as weddings or funerals. For EAs, it was common to have close relationships with immediate family members, but rarely with one's extended family.

Hmong Americans often stated that their closest friends were siblings or cousins. One HA participant, Chee, said that the majority of her social circle was her family. This was extremely rare among EAs. When HAs had close friends outside of their relatives, most of these friends were other HA university students. EAs' friends were usually other students (either from their high school or the University of Minnesota), and mostly white—although one EA participant, Thursday, said that some of her closest friends were Ethiopian, and another, Jessica, said her closest friends included Hispanic, Indian, and African American ethnicities. Some HA participants had social connections to EAs through work and classes, but most of these connections were perfunctory and did not include meaningful interactions. Some HA participants who attended Christian churches had more connections to EAs through church if the church was not primarily Hmong congregants, which was only occasionally the case.

Many HA participants had exposure to varieties of language other than just Minnesotan English. HA participants usually had family members in Thailand, Laos, or other states in the U.S., and some had siblings or close cousins living in California or Wisconsin that they still talked with or visited regularly. For EA participants, it was less likely that they had strong connections with family or friends outside of Minnesota. Some EAs had moved around within the U.S., but they did not maintain many meaningful connections with other regions. Nevertheless, EAs did not express any strong desire to remain in Minnesota in the future, whereas HAs almost always did. As Chee said, "Minnesota is home sweet home; I couldn't imagine moving anywhere else." Another HA participant said that even if she did want to move to another state, her family would make her move back right away.

Several HA participants had personally experienced racism and discrimination, and most HA participants knew someone who had experienced racism and discrimination. The general feeling among HA participants, though, was that currently they did not personally experience negative effects of racism on a daily basis. Participants were aware of negative portrayals of the Hmong in the media, but this was not perceived to affect most of their daily interactions with other people in Minnesota. It was not uncommon for HA participants to say that they were generally treated *more* favorably because of their ethnicity. One male HA participant, Morgan, said that "everyone wants to be your friend because you're different." However, many HA participants mentioned feeling more uncomfortable in other states where fewer people of Hmong ethnicity live.

Although the general consensus among HA participants was quite favorable towards Minnesota, the only clear cause of this positive feeling that I could deduce was the concentration of other Hmong in the Twin Cities. One question I asked about favorite sports teams, which was meant to probe state-wide affiliations, only received a strong response from two HAs, both of whom were fans of several Minnesota professional sports teams.

6.2.2 Expected results

The GOOSE vowel

The country-wide phonemic shift involving the fronting of the GOOSE vowel (and sometimes GOAT) is the only vowel shift investigated in this dissertation that behaves as expected. Because the shift has been said to be spreading through Minnesota (see, e.g., Ash, 1996), I expected to measure this difference by comparing the ANAE data, from 1993-1994, and my own data, from 2010. The difference between F2 means for the two corpora was one of the largest differences measured for all vowels. Furthermore, the difference was in the expected direction; that is, current European Americans in the Twin Cities produced a more front version of the GOOSE vowel than the speakers from the ANAE.

It was also predicted that individuals whose social networks are looser and/or more outwardly-focused will tend to front the GOOSE vowel more than individuals with inwardly-focused, tighter social networks. In other words, European Americans should front the GOOSE vowel more than Hmong Americans. Again, the measured difference between the two speaker groups (now HA and EA, both from 2010) was one of the largest differences measured for all vowels, and again in the expected direction. Although the test of significance of an F2 difference between the two ethnicities was not conclusive, the quantitative trend was also in the expected direction. Unsurprisingly, listeners in the perception study were significantly more likely to identify the most front GOOSE tokens as being produced by "white" speakers.

Traditionally, a backed GOOSE vowel in Minnesota has been associated with rural residents and those with lower socioeconomic status. It is possible that avoidance of GOOSE fronting is used by speakers in this investigation to index a working class background, which is more common among the HA speakers than the EA speakers.

Interestingly, it was found in the perception study that EA listeners were able to more accurately identify speaker ethnicity than HA listeners when listening to the GOOSE vowel. Perhaps this is because EA listeners (and speakers) are more aware that there are several options for producing this vowel because their looser social networks allow exposure to a wider variety of linguistic options. HA listeners may be less exposed to more fronted variants of GOOSE, and therefore they either do not recognize the difference

when they hear it, or they recognize the difference but are not aware of who might use one variant over another.

Other predictions

As predicted in Chapter 4, listeners were able to accurately identify the ethnicity of speakers, confirming that there are perceptible differences between HA and EA speakers. There was one exception to this: for the GOOSE vowel, HA listeners were not as accurate at identifying HA speakers. Notably, this vowel was the only one found to be clearly proceeding in the expected direction of the vowel shift. Perhaps because HAs are not as active in the GOOSE-fronting shift, more- or less-shifted variants do not hold strong sociolinguistic meaning for them.

In addition, both HA and EA listeners were often better at identifying speakers of the same ethnicity as themselves. This holds true even after taking into account listeners' bias for choosing their own ethnicity regardless of actual speaker ethnicity.[1] This accuracy for identifying speakers of one's own ethnicity was expected because listeners are assumed to have more exposure to the speech of people of the same ethnicity, and thus recognition should be facilitated.

6.2.3 Unexpected findings

The low back merger

In Chapter 2, it was predicted that because the low back merger is said to be receding in Southern Minnesota (Labov et al., 2006), we would expect to see the LOT and THOUGHT vowels moving away from one another in the current study, relative to the data in the ANAE. Furthermore, because the merger may be receding, the production of merged vowels should be associated with older speakers, and thus younger speakers (like the ones in this study) would be motivated to produce un-merged vowels. However, the results of production study #1 (reported on in Chapter 3) suggest that the low back vowels have not moved a great deal since 1994. It is clear to me (as someone who produces separate low back vowels) that some speakers in the Twin Cities have mostly or completely

[1] This same-ethnicity bias is tempered by the results for the speaker friend ethnicity judgments, where, again, EA listeners were more likely to choose "Definitely white" for EA speakers' friends but HA listeners were also slightly more likely to choose "Definitely white" for EA speakers' friends.

merged low back vowels, whereas other speakers' LOT and THOUGHT vowels are quite separate. On average, however, it appears that the low back vowels in the Twin Cities are close but not merged, and have been this way for years. It is possible that the relative position of the low back vowels does not have strong sociolinguistic associations, and therefore there is little motivation for speakers to alter their pronunciations of these vowels.

It was also hypothesized that HAs in the Twin Cities would produce less-shifted low back vowels (or more shifted, if the merger were receding) than their EA age-matched counterparts, due to the tight-knit social structure of Hmong Americans that often slows adoption of sound changes that originate from outside the speech community. The results of production study #2 suggest that there is no significant difference between the two ethnicity groups sampled for production of the low back vowels. However, it was found that, within the Hmong American sample, speaker birthplace played a significant role in predicting the separation of LOT and THOUGHT, with those born in Minnesota producing vowels that were more separated than for those not born in Minnesota. If indeed the low back merger is receding, albeit too slowly to measure by comparing current data with the ANAE, this may mean that the low back vowels are more merged for speakers in other areas of the country, with whom non-Minnesota-born Hmong Americans identify more strongly.

The perception study revealed that listeners may in fact have some sociolinguistic associations with placement of the low back vowels. F2 was the strongest acoustic cue (of the ones investigated) that listeners used to differentiate HA speakers from EA speakers: the furthest back tokens of THOUGHT were strongly associated with HA speakers, whereas the furthest front tokens did not present listeners with strong associations to either ethnicity. It was also found that listeners (but only those born in Minnesota) tended to identify lower F1 values for THOUGHT (those further away from LOT) with speakers who are HA. On the other hand, it was also found that THOUGHT tokens with the lowest F1 values were most likely to be judged to be produced by a speaker of the other ethnicity, by both EA and HA listeners. For the LOT vowel, those with higher F2 values—further back tokens, closer to the THOUGHT vowel—were more likely to be judged to be produced by one's own ethnicity.

Overall, these results suggest that although there may not be great acoustic shifts

among the LOT and THOUGHT vowels of Minnesotans, these two vowels may still index some social meaning. For speakers born outside Minnesota, more-merged vowels index "like me" and less-merged vowels index "not like me." For speakers born in Minnesota, though, there was a stronger connection between ethnicity and low back vowel placement.

The Northern Cities Shift

In Chapter 4, it was hypothesized that if the low back merger is not advancing in the Twin Cities, the Northern Cities Shift would have more freedom to begin to spread. The usual first step of the NCS is the raising and some fronting of the TRAP vowel, followed by fronting of LOT and lowering of THOUGHT. Although the ANAE has documented TRAP-raising for speakers living in the Twin Cities, I have found no evidence that the NCS is advancing in this region. In fact, the TRAP vowel before oral consonants (CAB) seems to be *lowering*, relative to the ANAE Twin Cities data. Based on studies of the influence of nasalization on vowel production and perception (e.g., Beddor, Krakow, & Goldstein, 1986; Kaiser & Plichta, 2009; Krakow, Beddor, Goldstein, & Fowler, 1988), it is not surprising that the CAB sub-class is produced lower than the STAND sub-class. Yet, compared to the ANAE, the current vowel means for CAB are a great deal lower—more than would be expected given the STAND–CAB difference found in the ANAE. In Chapter 5, I hypothesized that this CAB-lowering result may suggest a reversal of the NCS. More likely, the TRAP-raising measured by the ANAE in 1994 was not actually the first step of the NCS, given the immobility of the LOT and THOUGHT vowels—the vowels moving in the second and third steps of the NCS—during the past 16 years.

While this result calls into doubt the presence of the NCS in the Twin Cities, the TRAP vowel does seem to be undergoing *some* kind of shift related to nasalization: the non-nasalized CAB vowel is moving further away from the nasalized STAND vowel. When we compare current EA speakers' productions with HA speakers' productions, we see that it is the F1 and F2 values of the STAND vowel that differentiates between the two ethnicity groups. EA speakers produce higher and more front STAND vowels than HA speakers. Therefore, the tokens of STAND with a greater degree of nasalization and/or a higher F1 and higher F2 (lower and less front) should correlate with a tendency to judge talkers as Hmong. Yet, when we bring the perception study results to bear on

this issue, the picture becomes more complicated. There are no clear acoustic cues that all listeners use to judge ethnicity for the STAND vowel. For the CAB vowel, HA speakers tend to nasalize this vowel more than EA speakers; and furthermore this nasalization difference is recognized and used to make ethnicity judgments by both clinically-trained listeners and the non-trained listeners.

There is conflicting evidence as to whether minority ethnicities participate in the same vowel shifts documented in European American speech communities (see, e.g., Gordon, 2000a; Jones, 2003; Roeder, 2006). I argue that accommodation can be quite nuanced and strategic: some vowel features can be used to signal accommodation while other features of the same vowel can be used to signal an independent social characteristic. Here we find that HAs' first formant values of CAB align themselves with EAs, while their nasalization levels of the same vowel sub-class distinguish them from the majority. For the STAND vowel, on the other hand, listeners did not agree on what features make a speaker sound more "Hmong" or more "white," other than a further back STAND vowel indicating more "Hmongness." These results do not indicate why a difference was observed between the F1 means of the STAND vowel produced by the two ethnicities.

Nasalization

The two most surprising, most unexpected results of this study involve nasalization of the TRAP vowel. First, it is surprising that HA speakers seem to nasalize TRAP more than EA speakers, and listeners use this to differentiate between ethnicities. Second, degree of nasalization differs between the two ethnicities for the CAB sub-class, but there is no sign of this in the traditional vowel quadrilateral plot in Figure 3.6, despite the documented perceptual-raising (and eventual production-raising) influence of nasalization on F1 of low vowels. The HA speakers in this investigation grew up speaking only the Hmong language usually until kindergarten. And even after kindergarten, most HA speakers still spoke Hmong in the home with their families, especially to the older members of their families. Many of them are still exposed to Hmong on a daily basis, with some still speaking it every day. Despite considerable exposure to the Hmong language, there seems to be no nasalization-dampening effect of Hmong phonemic vowel nasalization on HA speakers' English vowels. In fact, the evidence presented in this dissertation suggests

that HA speakers actually have higher nasalization levels for the CAB vowel compared to the EA speakers. The perception study and auxiliary study of the perception of nasality supported the findings of the production study, showing that this difference in the nasalization of CAB is perceptible and is used to make judgments about speaker ethnicity.

Recall that higher nasalization levels often lead to perception (then production) of raised low vowels, such as TRAP approaching DRESS. Since we found higher levels of nasalization for HA speakers, it was predicted that this group, then, would also be further along in the Northern Cities Shift. This is also not the case, according to the results reported in Chapter 3. Furthermore, the perception results in Chapter 5 validated this finding: listeners were not able to differentiate between speakers' ethnicities based on F1 cues for the TRAP vowel.

A possible explanation for why HA speakers did not have lower TRAP nasalization levels was introduced in Section 3.1.4 of Chapter 3. Mexican Americans in Detroit avoided TRAP-raising (despite otherwise almost complete adoption of the NCS) (Roeder, 2010), but not because of heritage language influence; Spanish does not have phonemic nasal vowels. The explanation given in Roeder (2010) was that "pronunciation of [TRAP], in particular, reflects unique patterns of history and identity" (p. 163). But this still does not explain why HA levels of nasalization were *higher* than EA levels. Nasalization has great potential for use as a sociophonetic variable—it is readily available and can be easily but subtly incorporated into any utterance. I fully expect nasalization is used sociophonetically by many speech communities. However, there is a dearth of research on sociophonetic nasalization because it is difficult to perceive and difficult to measure acoustically.

As for the finding that greater levels of nasalization did not lead to greater participation in the NCS, we may simply say that the NCS is not yet progressing in the Twin Cities. The TRAP vowel does indeed show phonetic differences between the two ethnicities, but it is not behaving as expected for the NCS, either longitudinally or between networks. Therefore, if the TRAP vowel is not in the process of moving, its use as a sociophonetic variable is more flexible and more likely to be used in non-NCS ways.

Other findings

For HA speakers, place of birth was found to be an important predictor for TRAP vowel articulation. HA speakers born in Minnesota patterned more like the EA speakers (who were also born in Minnesota). Apparently, being born in Minnesota aligns HA speakers at least partially with the majority ethnicity in Minnesota. However, the perception study did not show that both ethnicity and place of birth together were important predictors for any vowels. For STAND, ethnicity was an important predictor in an interaction with acoustic factors, but place of birth was not. Therefore, it appears that place of birth is used by Hmong Americans to categorize themselves in production but not perception.

For listeners, speaker birthplace is important too, but for different reasons. For HA listeners, there was no relationship between speaker ethnicity and perceived speaker birthplace. EA listeners, on the other hand, were more likely to rate EA speakers as being born in Minnesota. Essentially, these listeners may have been conflating "sounds like me" with "native (to Minnesota)" and "doesn't sound like me" with "not native (to Minnesota)," regardless of where the listeners themselves were born.

It was found in Section 5.4 of Chapter 5 that seven out of ten words tested were significantly more likely to be judged to be used by one ethnicity or the other, and that, generally, the two ethnicity groups agreed on these judgments. Two of those seven words, *mom* and *new year*, were judged differently by the two ethnicity groups. For the word *mom*, the HA median was close to zero, whereas the EA median was near the "white" end of the spectrum. Perhaps EAs assumed that HA familial relations were more formal, and *mom* was too informal for HAs' use. For *new year*, the EA median was close to zero, and the HA median was near the "Hmong" end of the spectrum. In HA culture, the new year is one of the most anticipated events of the year, and is cause for excitement and celebration. If EAs are not aware of this fact, they may be more likely to assume that neither ethnicity uses *new year* more than the other. The auxiliary word-use study results contributed to another unexpected finding: listeners have measurable stereotypes about HA and EA speech, and they also have measurable stereotypes about HA and EA word use, but these two stereotypes do not interact.

Although the main focus of this dissertation was vowel variation, several consonants were observed to vary between ethnicity groups, as well. In fact, consonant variation

was impressionistically more apparent than vowel variation. In particular, HA speakers' consonant productions, when different from the regional standard, often resembled certain features of African American English. There is a precedent already set for appropriation of AAE features by non-African American speech communities in the United States, "marking urban youth subcultural participation" (Reyes, 2005, p. 509). This is in line with some of the HA participants' comments about a tendency for HA youth to "talk ghetto." The vowel variation found in this work does not bear out an AAE-appropriation trend, but does not refute it, either. Rather, taking vowel and consonant variation together, we see a more complete picture of the complex identities represented by sociophonetic variation in the speech of Hmong American youth in Minnesota.

From this research, we can also gain insights into some of the bigger questions in sociolinguistics. Regarding the question of how language changes, we know that language change can be measured, but the characteristics of language that are changing, or are influencing the change, may be subtle or unexpected. Social identity is not one-dimensional, and if treated as such, valuable phonetic and social information may be lost. From the diasporic linguistics perspective, we have learned that heritage language influence on the matrix language is not straightforward—it is not simply foreign-accented English that Hmong Americans speak. The question of how diasporic politics affects language use has only been briefly touched on here, providing evidence that Hmong Americans in Minnesota may consider themselves somewhat politically aligned with European Americans and African Americans at the same time. More research on the social and political structures of the Hmong American community is needed to properly address this last question.

6.3 Further research

As with most research projects, the scope of this work was limited by time and resource availability constraints; this dissertation should most definitely not be taken as the final word. This topic would benefit from future studies with more subjects of greater variation in linguistic and demographic backgrounds, especially different education and socioeconomic levels. Most importantly, this study needs a follow-up investigation that takes a traditional ethnographic approach, in order to elucidate the social groupings

and practices that hold meaning for Hmong Americans in Minnesota.

In addition, another perception experiment should be designed that expands the speaker ethnicity choices to include African Americans as well as Hmong and European Americans. This protocol change would be more likely to model the actual choices that HA speakers make when they choose one sociophonetic form over another.

It would be insightful to eventually extend this research to Hmong American populations in other areas, especially areas under greater influence of the Northern Cities Shift. Then, I would like to see the research extend even further to investigate nasalization and regional sound changes for other similar minority groups.

Finally, I hope that this study will inspire more investigations of the cross-linguistic influences and sociophonetic uses of nasalization.

6.4 Limitations

This dissertation follows the typical ethnolinguistic strategy of comparing minority and majority speech communities, and taking the majority ethnicity to be the "matrix," "standard," or "comparison sample." This approach is admittedly ethnocentric, but it still has some value, especially for studies of minority populations who interact frequently with members of the majority ethnicity.

It may not be the case that any differences found in dialects between non-European American and European American populations signal the non-European American population must be consciously or unconsciously asserting their non-European American identity through their linguistic choices. In any population of speakers, from the most mainstream to the most marginalized, the linguistic choices made by the speakers must be evaluated with respect to the choices that are available to those speakers. In other words, for a speaker who has been exposed to only minor variations of a mainstream dialect, very small linguistic differences can reflect great differences in social attitudes. For speakers of vernacular dialects, who are usually exposed to a mainstream variety as well, they have a greater range of linguistic choices when it comes to deciding how their speech will sound. Members of non-majority ethnic groups also have an additional reason for utilizing ethnic identity markers that majority groups do not have: asserting pride in one's ethnic background, and thereby legitimizing or validating an important

part of one's identity.[2]

Other limitations of this work include sub-optimal recording environments and sound quality, especially for the acoustic analysis of nasalization; gender imbalances; and small sample sizes. In order to limit the scope of the project, individual differences were not explored; this should be done in future research.

6.5 Final words

Despite the limitations of field and laboratory research in general, and of my personal implementation of those methods in particular, I believe this dissertation contributes valuable knowledge to several areas of study. For dialectologists, this work provides an up-to-date view of American English as it is spoken in the Twin Cities, and how it has changed within the last sixteen-plus years. My exploration of acoustic and sociophonetic roles of nasalization in the Northern Cities Shift may influence how researchers view the NCS, and more generally, nasalization as a possible sociophonetic variable and the perceiver's role in sound changes. Laboratory phonologists may benefit from this study's integration of field research and laboratory methods. To the broader field of sociolinguistics, this work provides much needed data on the speech of a relatively understudied Asian American ethnic group, as well as an analysis of whether and why they adopt regional sound changes. This study is also important for linguistic anthropologists studying minority accommodation to an ethnic majority dialect or exploring whether Asian Americans conform to certain ubiquitous linguistic stereotypes. Finally, the investigation of Hmong American attitudes toward their own ethnicity and ethnic relations will make valuable contributions to the field of ethnic and cultural studies.

This work required countless hours of research and preparation, and countless more hours working with people who have their own lives and schedules. I have endured setbacks related to equipment failures, scheduling conflicts, cultural barriers, methodological adaptations, and more. As demanding as this project has been, it has been even more rewarding.

It is said that knowledge is the first step toward ending discrimination. If this dissertation serves only to augment public knowledge of Hmong Americans, an important and

[2] See Eckert (2008) for more discussion on this topic.

growing minority ethnic group in the United States, then it will have been a worthwhile endeavor.

References

ACS. (2009). *2005–2009 American Community Survey 5-year Estimates*. U.S. Census Bureau. Retrieved August 24, 2011, from http://factfinder.census.gov.

Adank, P., Smits, R., & Van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, *116*, 3099–3107.

Alba, R. D., Logan, J., Lutz, A., & Stults, B. (2002, August). Only English by the third generation? Loss and preservation of the mother tongue among the grandchildren of contemporary immigrants. *Demography*, *39*(3), 467–484.

Allen, H. (1973). *The linguistic atlas of the Upper Midwest, volume 1*. Minneapolis, MN: University of Minnesota Press.

Allen, H. (1976). *The linguistic atlas of the Upper Midwest, volume 2*. Minneapolis, MN: University of Minnesota Press.

Ash, S. (1996). Freedom of movement:/uw/-fronting in the Midwest. In J. Arnold, R. Blake, B. Davidson, S. Schwenter, & J. Solomon (Eds.), *Sociolinguistic variation: Data, theory, and analysis: Selected papers from NWAV 23 at Stanford* (pp. 3–25). Stanford, CA: CSLI Publications.

Badger77. (2006, May 17). Midwestern accents... [Msg 49]. Message posted to http://www.skyscrapercity.com/showthread.php?t=350336.

Baugh, J. (1996). Perceptions within a variable paradigm: Black and white racial detection and identification based on speech. In E. W. Schneider (Ed.), *Focus on the USA* (pp. 169–182). Amsterdam: Benjamins.

Beddor, P. S., & Hawkins, S. (1990). The influence of spectral prominence on perceived vowel quality. *The Journal of the Acoustical Society of America*, *87*, 2684.

Beddor, P. S., Krakow, R., & Goldstein, L. (1986). Perceptual constraints and phonological change: a study of nasal vowel height. *Phonology Yearbook*, *3*, 197–217.

Benor, S. B. (2002). Sounding learned: The gendered use of /t/ in Orthodox Jewish English. *Penn Working Papers in Linguistics: Selected Papers from NWAV 2000*, 1–16.

Bernstein, C. (1993). Measuring social causes of phonological variation in Texas. *American Speech*, *68*(3), 227–240.

Bliatout, B. T., Downing, B. T., Lewis, J., & Yang, D. (1988). *Handbook for teaching Hmong-speaking students.* Folsom, CA: Folsom Cordova Unified School District, Southeast Asia Community Resource Center.

Boberg, C. (2000). Geolinguistic diffusion and the U.S.–Canada border. *Language Variation and Change*, *12*, 1–24.

Boersma, P., & Weenink, D. (2008). *Praat: doing phonetics by computer.* [Computer program]. Version 5.0.36, retrieved 20 October 2008 from http://www.praat.org/.

Bosher, S. D. (1997). Language and cultural identity: A study of Hmong students at the postsecondary level. *TESOL Quarterly*, 593–603.

Bourdieu, P. (1977). The economics of linguistic exchanges. *Social Science Information*, *16*(6), 645–668.

Chen, M. (1995). Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *The Journal of the Acoustical Society of America*, *98*, 2443.

Chen, M. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, *102*, 2360.

Chi, M. (1999). Linguistic perspective on the education of Chinese-American students. In C. Park & M. Chi (Eds.), *Asian-American education: Prospects and challenges* (pp. 18–46). Westport, CT: Bergin & Garvey, Greenwood Publishing Group.

Chun, E. (2009). Speaking like Asian immigrants: Intersections of accommodation and mocking at a U.S. high school. *Pragmatics*, *19*, 17–38.

Crankbaiter. (2006, May 18). Midwestern accents... [Msg 54]. Message posted to http://www.skyscrapercity.com/showthread.php?t=350336.

DeCamp, D. (1953). *The pronunciation of English in San Francisco.* Unpublished doctoral dissertation, University of California, Berkeley.

Drager, K. (2011). Speaker age and vowel perception. *Language and Speech*, *54*(1), 99–121.

Eberhardt, M. (2008). The Low-Back Merger in the Steel City: African American English in Pittsburgh. *American Speech*, *83*(3), 284–311.

Eckert, P. (1989a). *Jocks and burnouts: Social categories and identity in the high school.* New York: Teachers College Press.

Eckert, P. (1989b). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1*, 245–268.

Eckert, P. (2005). *Variation, convention, and social meaning.* Paper presented at the annual meeting of the Linguistics Society of America, Oakland, CA.

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*(4), 453–476.

Fadiman, A. (1997). *The spirit catches you and you fall down: A Hmong child, her American doctors, and the collision of two cultures.* New York: Farrar, Straus and Giroux.

Faruque, C. J. (2002). *Migration of Hmong to the midwestern United States.* Lanham, MD: University Press of America, Inc.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Foucault, M. (1988). The concern for truth. In L. Kritzman (Ed.), *Politics, philosophy, culture: Interviews and other writings 1977–1984.* New York: Routledge.

Fought, C. (1999). A majority sound change in a minority community: /u/-fronting in Chicano English. *Journal of Sociolinguistics*, *3*, 5–23.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*(4), 409–438.

Frazer, T. C. (1993). Problems in Midwest English: Introduction and overview. In T. C. Frazer (Ed.), *"Heartland" English: Variations and transition in the American Midwest* (pp. 1–19). Tuscaloosa: University of Alabama Press.

Frazer, T. C. (1996). Chicano English and Spanish interference in the midwestern United States. *American Speech*, 72–85.

Fridland, V. (2008a). Patterns of /uw/, /ʊ/, and /ow/ fronting in Reno, Nevada.

American Speech, *83*(4), 432–454.

Fridland, V. (2008b). Regional differences in perceiving vowel tokens on Southerness, education, and pleasantness ratings. *Language Variation and Change*, *20*(01), 67–83.

Fridland, V., & Bartlett, K. (2006). The social and linguistic conditioning of back vowel fronting across ethnic groups in Memphis, Tennessee. *English Language and Linguistics*, *10*, 1–22.

Gordon, M. (1997). *Urban sound change beyond city limits: The spread of the Northern Cities Shift in Michigan.* Unpublished doctoral dissertation, University of Michigan.

Gordon, M. (2000a). Phonological correlates of ethnic identity: Evidence of divergence? *American Speech*, *75*(2), 115–136.

Gordon, M. (2000b). Small-town values and big-city vowels: A study of the Northern Cities Shift in Michigan. *American Speech*, *84*(1).

Gordon, M. (2006). Tracking the low back merger in Missouri. In T. E. Murray & S. B. Lee (Eds.), *Language variation and change in the American Midland: a new look at "Heartland" English* (pp. 57–68). Philadelphia: John Benjamins.

Graff, D., Labov, W., & Harris, W. A. (1986). Testing listeners reactions to phonological markers of ethnic identity: A new method for sociolinguistic research. In D. Sankoff (Ed.), *Diversity and Diachrony* (pp. 45–58). Philadelphia: John Benjamins.

Hall-Lew, L. (2009). *Ethnicity and variation in a San Francisco neighborhood.* Unpublished doctoral dissertation, Stanford University.

Hamilton-Merritt, J. (1993). *Tragic mountains: The Hmong, the Americans, and the secret wars for Laos, 1942-1992.* Bloomington, IN: Indiana University Press.

Hay, J., & Drager, K. (2007). Sociophonetics. *Annual Review of Anthropology*, *36*, 89–103.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865–892.

Hazen, K. (2005). Mergers in the mountains: West Virginia division and unification. *English World-Wide*, *26*(2), 199–221.

Hein, J. (1994). From migrant to minority: Hmong refugees and the social construction

of identity in the United States. *Sociological Inquiry*, *64*(3), 281–306.

Her, V. K. (2009). Animal sacrifice and social meanings in Hmong American funerals. In G. Y. Lee (Ed.), *The impact of globalization and trans-nationalism on the Hmong.* St. Paul, MN: Center for Hmong Studies.

Herndobler, R. (1977). *White working-class speech: The East Side of Chicago.* Unpublished doctoral dissertation, University of Chicago.

Hinrichs, L. (2011). The Sociolinguistics of Diaspora: Language in the Jamaican Canadian Community. In J. R. Sullivant (Ed.), *Texas Linguistics Forum: Proceedings of the Nineteenth Annual Symposium About Language and Society–Austin* (Vol. 54).

Hinton, L. (2009). Loss of heritage languages in the United States. In A. Reyes & A. Lo (Eds.), *Beyond Yellow English: Toward a linguistic anthropology of Asian Pacific Americans* (pp. 331–346). New York: Oxford University Press.

Hinton, L., Moonwomon, B., Bremner, S., Luthin, H. W., Van Clay, M., Lerner, J., et al. (1987). It's not just the Valley Girls: A study of California English. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the thirteenth annual meeting of the Berkeley Linguistics Society* (pp. 117–127). Berkeley, CA: Berkeley Linguistics Society.

Horii, Y. (1980). An accelerometric approach to nasality measurement: A preliminary report. *The Cleft Palate Journal*, *17*(3), 254–261.

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English.* Ann Arbor, MI: Karoma Publishers.

Ito, R. (2010). Accommodation to the local majority norm by Hmong Americans in the Twin Cities, MN. *American Speech*, *85*(2), 141–162.

Jacewicz, E., Fox, R., & Salmons, J. (2006). Prosodic prominence effects on vowels in chain shifts. *Language Variation and Change*, *18*(03), 285–316.

Jones, J. (2003). *African Americans in Lansing and the Northern Cities Vowel Shift: Language contact and accommodation.* Unpublished doctoral dissertation, Michigan State University.

Kaiser, E., & Drager, K. (2009). *A sociophonetic investigation of nasalization, vowel perception, and speaker characteristics.* Poster presented at the International Nasal Conference, Montpellier, France.

Kaiser, E., & Plichta, B. (2009). *Influence of nasalization on vowel perception.* Presented

at the American Dialect Society Annual Meeting, San Francisco.

Knoebel, S. (1986). *Hmong production of word-final obstruents in English as a function of language transfer, universal tendencies, and attention paid to speech.* Unpublished master's thesis, University of Minnesota.

Koltyk, J. (1998). *New pioneers in the heartland: Hmong life in Wisconsin.* Boston: Allyn & Bacon.

Krakow, R., Beddor, P., Goldstein, L., & Fowler, C. (1988). Coarticulatory influences on the perceived height of nasal vowels. *The Journal of the Acoustical Society of America*, *83*, 1146.

Labov, W. (1962). *The social history of a sound change on the island of Martha's Vineyard, Massachusetts.* Unpublished master's thesis, Columbia University.

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

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, *2*, 205–254.

Labov, W. (1991). The three dialects of English. In P. Eckert (Ed.), *New ways of analyzing sound change* (Vol. 5, pp. 1–44). New York: Academic Press.

Labov, W. (2001). Principles of linguistic change: Social factors. In *Language in society* (Vol. 2). Malden, MA: Blackwell Publishers.

Labov, W. (2006). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics*, *34*(4), 500–515.

Labov, W. (2007). Transmission and diffusion. *Language*, *83*(2), 344–387.

Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English: Phonetics, phonology, and sound change: a multimedia reference tool.* New York: Walter De Gruyter.

Labov, W., Yaeger, M., & Steiner, R. (1972). *A quantitative study of sound change in progress: Report on National Science Foundation Contract GD-3287.* Philadelphia: U.S. Regional Survey.

Lee, G. Y. (1986). Culture and adaptation: Hmong refugees in Australia. In G. L. Hendricks, B. T. Downing, & A. S. Deinard (Eds.), *The Hmong in transition.* New

York: Center for Migration Studies and The Southeast Asian Refugee Studies of the University of Minnesota.

Lee, G. Y. (2005). *Hmong post-war identity production: Heritage maintenance and cultural reinterpretation.* Paper presented at the ICCOM Forum on Armed Conflict and Conservation, Rome, Italy.

Lee, G. Y., & Tapp, N. (2010). *Culture and customs of the Hmong.* Santa Barbara, CA: Greenwood.

Lee, H. (2001). *Korean Americans as speakers of English: The acquisition of general and regional features.* Unpublished doctoral dissertation, University of Pennsylvania.

Li, G. (2002). *East is east, west is west?: Home literacy, culture, and schooling.* New York: Peter Lang Publishing Inc.

Lippi-Green, R. (1997). *English with an accent.* New York: Routledge.

Lo, A., & Reyes, A. (2009). Introduction: On Yellow English and other perilous terms. In A. Reyes & A. Lo (Eds.), *Beyond Yellow English: Toward a linguistic anthropology of Asian Pacific Americans* (p. 3-20). New York: Oxford University Press.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, *49*, 606–608.

López, M. (2000). The language situation of the Hmong, Khmer, and Laotian communities in the United States. In S. McKay & S. Wong (Eds.), *New immigrants in the United States* (pp. 232–262). Cambridge: Cambridge University Press.

Lor, L. (2009). The Thai-Hmong and Hmong refugees in Thailand: The politics of human rights. In G. Y. Lee (Ed.), *The impact of globalization and trans-nationalism on the Hmong.* St. Paul, MN: Center for Hmong Studies.

Luthin, H. W. (1987). The story of California (ow): The coming-of-age of English in California. In K. M. Denning, S. Inkelas, F. C. McNair-Knox, & J. R. Rickford (Eds.), *Variation in language: NWAV-XV at Stanford (Proceedings of the fifteenth annual conference on New Ways of Analyzing Variation in Language* (pp. 312–324). Stanford, CA: Department of Linguistics, Stanford University.

Ly, N. (1999). We are the freedom people. In D. Urbaniak Lesch & J. L. O'Donoghue (Eds.), (pp. 124–130). St. Cloud, MN: Sentinel Printing Company.

Milroy, L. (1980). *Language and social networks.* Oxford: Blackwell.

Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and interpretation* (Vol. 13). Walden, MA: Blackwell.

Milroy, L., Milroy, J., & Hartley, S. (1994). Local and supra-local change in British English: The case of glottalisation. *English World Wide*, *15*, 1–33.

MN Dept. of Education. (n.d.). *Enrollment data for 2010–2011 school year.* [Website] Retrieved on 31 May 2011 from http://www.education.state.mn.us.

Moua, M. N. (1999). We are the freedom people. In D. Urbaniak Lesch & J. L. O'Donoghue (Eds.), (pp. 40–51). St. Cloud, MN: Sentinel Printing Company.

Moua, S. (1999). We are the freedom people. In D. Urbaniak Lesch & J. L. O'Donoghue (Eds.), (pp. 180–187). St. Cloud, MN: Sentinel Printing Company.

Munro, M. J., Derwing, T. M., & Flege, J. E. (1999). Canadians in Alabama: A perceptual study of dialect acquisition in adults. *Journal of Phonetics*, *27*(4), 385–403.

Munson, B., Jefferson, S., & McDonald, E. (2006). The influence of perceived sexual orientation on fricative identification. *The Journal of the Acoustical Society of America*, *119*(2427).

New, B., Brysbaert, M., Segui, J., Ferrand, L., & Rastle, K. (2004). The processing of singular and plural nouns in French and English. *Journal of Memory and Language*, *51*(4), 568–585.

Nishimura, M. (1997). *Japanese/English code-switching: Syntax and pragmatics* (Vol. 24). New York: Peter Lang Publishing Inc.

Ohala, J. J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick, & M. F. Miller (Eds.), *Papers from the parasession on langugae and behavior—Chicago Linguistic Society* (pp. 178–203).

Olney, D. P. (1986). Population trends. In G. L. Hendricks, B. T. Downing, & A. S. Deinard (Eds.), *The Hmong in transition.* New York: Center for Migration Studies and The Southeast Asian Refugee Studies of the University of Minnesota.

Plichta, B. (2005). *Interdisciplinary perspectives on the Northern Cities Chain Shift.* Unpublished doctoral dissertation, Michigan State University.

Plichta, B. (2008). *Akustyk.* [Computer program]. Version 1.8, retrieved 16 August 2008 from http://bartus.org/akustyk/.

Podesva, R. (2004). *On constructing social meaning with stop release bursts.* Paper presented at the Sociolinguistics Symposium 15, Newcastle upon Tyne.

Preston, D. R. (1993). Folk dialectology. In D. R. Preston (Ed.), *American dialect research* (pp. 333–377). Amsterdam: Benjamins.

Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, *18*, 10–30.

Quigley, L., Shiere, F., Webster, R., & Cobb, C. (1964). Measuring palatopharyngeal competence with the nasal anemometer. *Cleft Palate Journal*, *1*, 304–313.

Rampton, B. (1995). *Crossing: Language and ethnicity among adolescents.* New York: Longman Publishing Group.

Ratliff, M. (1992). *Meaningful tone: A study of tonal morphology in compounds, form classes and expressive phrases in White Hmong.* Northern Illinois University, Center for Southeast Asian Studies.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Reyes, A. (2005). Appropriation of African American slang by Asian American youth. *Journal of Sociolinguistics*, *9*(4), 509–532.

Roeder, R. V. (2006). *Ethnicity and sound change: Mexican American accommodation to the Northern Cities Shift in Lansing, Michigan.* Unpublished doctoral dissertation, Michigan State University.

Roeder, R. V. (2010). Northern Cities Mexican American English: Vowel production and perception. *American Speech*, *85*(2), 163–184.

Rothenberg, M. (1977). Measurement of airflow in speech. *Journal of Speech and Hearing Research*, *20*, 155–176.

Samant, S. (2010). *Adolescents and the Northern Cities Vowel Shift.* Unpublished doctoral dissertation, University of Michigan.

Shankar, S. (2011). Asian American youth language use: Perspectives across schools and communities. *Review of Research in Education*, *35*, 1–28.

Shelton, R., Knox, A., Arndt, W., Jr., & Elbert, M. (1967). The relationship between nasality score values and oral and nasal sound pressure level. *Journal of Speech,*

Language and Hearing Research, *10*(3), 542–548.

Stevens, K. (2000). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Stevens, K., Kalikow, D., & Willemain, T. (1975). A miniature accelerometer for detecting glottal waveforms and nasalization. *Journal of Speech, Language and Hearing Research*, *18*(3), 594–599.

Stevens, K., Nickerson, R., Boothroyd, A., & Rollins, A. (1976). Assessment of nasalization in the speech of deaf children. *Journal of Speech, Language and Hearing Research*, *19*(2), 393–416.

Stuart-Smith, J. (1999). Voice quality in Glaswegian. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (Vol. 99, pp. 2553–2556).

Thao, C. (1982). Hmong migration and leadership in Laos and in the United States. In B. T. Downing & D. P. Olney (Eds.), *The Hmong in the West.* Center for Urban and Regional Affairs, University of Minnesota.

Thao, P. (1999). Mong linguistic awareness for classroom teachers. In C. Park & M. Chi (Eds.), *Asian-American education: Prospects and challenges* (pp. 237–262). Westport, CT: Bergin & Garvey, Greenwood Publishing Group.

Thomas, E. R. (2001). *An acoustic analysis of vowel variation in New World English* (Vol. 85). Durham, NC: Duke University Press.

Thomas, E. R. (2002). Sociophonetic applications of speech perception experiments. *American Speech*, *77*(2), 115–147.

Thomas, E. R. (2007). Phonological and phonetic pharacteristics of African American Vernacular English. *Language and Linguistics Compass*, *1*(5), 450–475.

Trudgill, P. (1974). Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society*, *3*, 215–246.

Trueba, H. T., Jacobs, L., & Kirton, E. (1990). *Cultural conflict and adaptation: The case of Hmong children in American society.* Bristol, PA: Falmer Press, Taylor & Francis Inc.

Tuc, H. (2003). *Vietnamese-English bilingualism: Patterns of code-switching.* New York: Routledge Curzon.

Vang, C. Y. (2008). *Hmong in Minnesota.* St. Paul, MN: Minnesota Historical Society Press.

Veatch, T. C. (1991). *English vowels: Their surface phonology and phonetic implementation in vernacular dialects.* Unpublished doctoral dissertation, University of Pennsylvania.

Wei, L. (1994). *Three generations, two languages, one family: Language choice and language shift in a Chinese community in Britain* (Vol. 104). Clevendon: Multilingual Matters Ltd.

Weiss, A. (1954). *Oral and nasal sound pressure levels as related to judged severity of nasality.* Unpublished doctoral dissertation, Purdue University.

Wells, J. (1982). *Accents of English 3: Beyond the British Isles.* Cambridge: Cambridge University Press.

Whitelock, D. (1982). *White Hmong language lessons.* Minneapolis, MN: Southeast Asian Refugee Studies Project, Center for Urban and Regional Affairs, University of Minnesota.

Whitenack, D., & Kikunaga, K. (1999). Teaching English to native Japanese students: From linguistics to pedagogy. In C. Park & M. Chi (Eds.), *Asian-American education: Prospects and challenges* (Vol. 111, pp. 110–130). Westport, CT: Bergin & Garvey, Greenwood Publishing Group.

Wolfram, W. (2000). On constructing vernacular dialect norms. In A. Okrent & J. Boyle (Eds.), *Chicago Linguistic Society 36: The panels* (pp. 335–358). Chicago: Chicago Linguistic Society.

Wolfram, W., Hazen, K., & Schilling-Estes, N. (1999). *Dialect change and maintenance on the Outer Banks* (Vol. 81). Tuscaloosa: University of Alabama Press.

Wong, A. (2007). Two vernacular features in the English of four American-born Chinese. *University of Pennsylvania Working Papers in Linguistics*, *13*(2), 217–30.

Yaeger-Dror, M., & Thomas, E. R. (Eds.). (2010). *African American English speakers and their participation in local sound changes* (Vol. 94). Durham, NC: Duke University Press.

Yang, D. (1982). Why did the Hmong leave Laos? In B. T. Downing & D. P. Olney (Eds.), *The Hmong in the West.* Center for Urban and Regional Affairs, University of Minnesota.

Yang, D. (2009). Hmong culture is Hmong soul. In G. Y. Lee (Ed.), *The impact of globalization and trans-nationalism on the Hmong.* St. Paul, MN: Center for

Hmong Studies.

Yang, K. (2009). The experience of Hmong Americans: Three decades in retrospective review. In G. Y. Lee (Ed.), *The impact of globalization and trans-nationalism on the Hmong.* St. Paul, MN: Center for Hmong Studies.

Zhang, Q. (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, *34*(03), 431–466.

Appendix A

Acronyms

Care has been taken in this thesis to minimize the use of acronyms, but several acronyms were used for the most frequent terms in order to conserve space. The following table defines acronyms used.

A.1 Acronyms

Table A.1: Acronyms

Acronym	Meaning
ACS	American Community Survey, published by the U.S. government to provide data for the years between the 10-year censuses
ANAE	Atlas of North American English, by William Labov, Sharon Ash, and Charles Boberg (2006)
EA	European American (broadly defined)
F1	First formant
F2	Second formant
HA	Hmong American
LPC	linear predictive coding
NCS	Northern Cities Shift

Appendix B

Interview Questions

B.1 HA first interview session

1. What city/neighborhood do you live in? How long have you lived here? Do you like living here?
2. How long has your family/ancestors lived in [St. Paul]? In the Twin Cities? In Minnesota? Do you think they like it?
3. Where were you born? Your parents? Your grandparents?
4. How much contact do you have with people from other parts of the country? Do you have relatives in other parts of the U.S.? What areas specifically?
5. Are you a student now?
6. Are you employed? What kind of work do you do? How long have you been doing that?
7. Are either or both of your parents employed? What kind of work do they do?
8. How many siblings do you have? What is their age range?
9. What neighborhood of [St. Paul] do you live in now? Have you always lived in this neighborhood?

10. Do you like living in this neighborhood? What do you like about it? What do you dislike about it?
11. Where would you choose to live if you could live anywhere in the world?
12. What do you like to do when you're not working/at school?
13. Where do you hang out when you're not working/at school?
14. Who do you hang out with when you're not working/at school?
15. When you are working/at school, who are you around most?
16. What kinds of movies/books/music do you like?
17. Are you a Vikings fan? A Twins fan? A Gophers fan? Other sports teams?
18. Do you think the Twin Cities is ethnically diverse?
19. Are minority ethnicities in the Twin Citites generally treated favorably or unfavorably?
20. What do you consider your ethnicity to be?
21. Is your ethnicity an important part of who you are?
22. Have you ever discussed your ethnicity with anyone?
23. Do you think your ethnicity makes other people treat you favorably or unfavorably? In what circumstances?
24. Do other peoples' opinions affect how you feel about your ethnicity? What are those opinions?
25. Who are the people you speak with every day?
26. Who are the most important people in your life?
27. Who do you ask to do favors for you frequently? Who do you do favors for frequently?

28. If you were to move, who would you ask for help?

29. Where do you see yourself in 10 years? (career, family, living location, etc.)

B.2 HA second interview session

1. Do you speak any other languages besides English? How comfortable are you in those other languages?

2. Did you grow up speaking Hmong in the home?

3. Do you speak Green Hmong or White Hmong?

4. Did you take ESL classes when you were younger? For how many years?

5. What percent of your day (on average) do you use English? What percent Hmong?

6. How comfortable are you speaking English? When did you learn English?

7. What language(s) do your parents speak?

8. What language(s) do you use with your friends?

9. What language(s) do you use at work/school?

10. When you were growing up, how did you feel about the way you spoke? How about now? If your feelings have changed, why do you think that is?

11. When you were growing up, how did you feel about the way your parents spoke? How about now? If your feelings have changed, why do you think that is?

12. How would you describe the way native Minnesotans speak English? How is it different from the rest of the U.S.?

13. Is there a difference in the way Hmong Americans speak compared to other people?

14. If you're talking on the phone to someone you haven't met, do you think you'd be able to tell if they are from Minnesota? If they are Hmong American?

15. How would others describe the way you speak English?

16. Is it important for you to be able to read and write in Hmong?
17. Is it important for you to teach Hmong to your children?

B.3 EA interview session

1. What city/neighborhood do you live in? How long have you lived here? Do you like living here?
2. How long has your family/ancestors lived in [St. Paul]? In the Twin Cities? In Minnesota? Do you think they like it?
3. Where were you born? Your parents? Your grandparents?
4. Are you a student now? Highest education level attained?
5. Are you employed? What kind of work do you do?
6. Do you think the Twin Cities is ethnically diverse?
7. What do you consider your ethnicity to be?
8. Are minority ethnicities in the Twin Citites generally treated favorably or unfavorably?
9. Do you know any Hmong Americans? Are you aware of any stereotypes surrounding this ethnic group?
10. Who are the people you spend the most time with at school, work, during free time? What are their ethnicities?
11. Who are the most important people in your life? What are their ethnicities?
12. Do you speak any other languages besides English? How comfortable are you in those other languages?
13. How often do you hear other languages being spoken in the Twin Cities?
14. How would you describe the way native Minnesotans speak English? How is it different from the rest of the U.S.?

15. Can you tell if there a difference in the way Hmong Americans speak compared to people of other ethnicities?

16. If you're talking on the phone to someone you haven't met, do you think you'd be able to tell if they are from Minnesota? If they are Hmong American?

B.4 Group conversation session

- Who are your biggest role models? Why?
- What are some of your future aspirations for education, career, family, etc.? How important is it that your family approve these aspirations?
- What advantages can you think of for knowing how to speak Hmong? How about English?
- Have you, or has someone you know, had experiences with racism at school, work, or elsewhere?
- How are your family's values different from those of other Hmong families? From those of non-Hmong families?
- Are gender role differences still very strong in Hmong culture? Are these changing at all in your own family?
- Do you value the ability to speak Hmong? Why? Do you know anyone who does not value it as much as you?
- Do you think that you and/or your Hmong friends sound different from people of other ethnicities? If so, what are some of these differences?
- What does it mean for a Hmong person to "sound American"? When is it a good thing? When is it a bad thing? Has this ever been said about you?
- When do you feel the most "Hmong"? When do you feel the most "American"? Is it hard to draw your identity from these two separate cultures?

Appendix C

Wordlist Items

Table C.1: Wordlist items elicited in production, part 1

LOT	TRAP	THOUGHT	FACE	DRESS	FLEECE
body	ancestor	ball	baby	chest	evil spirit
father*	back	call	nature	head	greetings
hop	bad	coffee	rain	hot pepper	meat
hot	cat	daughter	stay	leg	sea
hot pepper	dad	dog	United States	mend	tea
mom	dance	doll		met	teeth
pot	family	fog		neck	tree
shaman*	hand	frog		pet	
stop	happy	hog		tend	
top	land	moth		vegetable	
	mad	song		you're welcome	
	rat	talk		hello	
	sad	toss			
	stand	aunt			
	thank you				

* These words are not technically in the LOT lexical set, but are generally pronounced alike in most varieties of U.S. English (Hall-Lew, 2009)

Table C.2: Wordlist items elicited in production, part 2

KIT	GOAT	FOOT	GOOSE	STRUT	filler
big	cold	cooking	bamboo	brother	cow
chicken	ghost	foot	cute	country	house
fish	old	good	food	culture	Laos
kid	soul	good morning	moon	cup	mountain
little			new year[1]	fun	mouth
mix				mother	child
pig				onion	China
river				son	eye
sick				sun	goodbye
sing				touch	hi
sister				uncle	rice
sit				young	Thailand
					United States
					farm
					sorry
					spirit
					story

Table C.3: Wordlist items tested in perception

LOT	CAB	STAND	THOUGHT	GOOSE	filler	
father	bad	ancestor	coffee	bamboo	chest	mix
mom	back	family	daughter	food	cow	pet
pot	happy	land	song	moon	cup	story
stop	mad	stand	talk	new year	fish	teeth
					house	

[1] Technically, "new" is in a different lexical category, but as Ash (1996) has noted, in the Midwest, the class of words containing "new" are not pronounced consistently differently from the GOOSE class. None of my subjects pronounced "new" noticeably differently from the GOOSE class.

Appendix D

Perception Experiment Prompts

D.1 Demographics judgments

Five different question prompts were presented in question blocks of about 100 tokens each (several tokens were repeated for reliability). All question prompts for the perception experiment were presented in Visual Analog Scale format. The five prompt screens are provided here in Figures D.1, D.2, D.3, D.4, and D.5.

Figure D.1: Prompt screen for "age" question

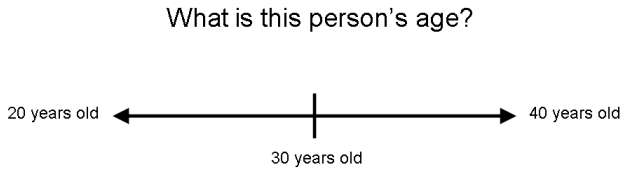


Figure D.2: Prompt screen for "birthplace" question

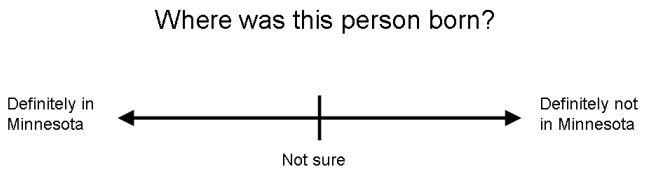


Figure D.3: Prompt screen for "ethnicity" question

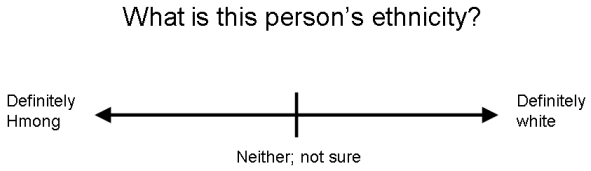


Figure D.4: Prompt screen for “friends” question

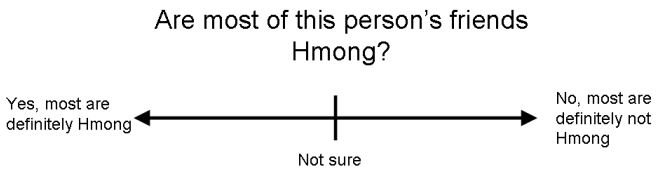
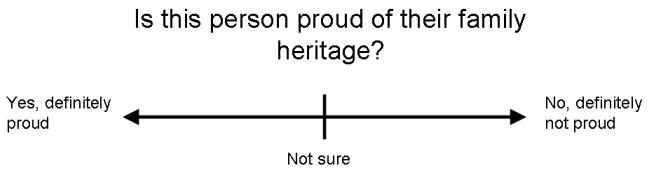


Figure D.5: Prompt screen for “heritage” question

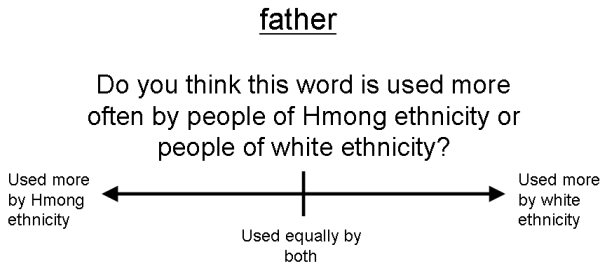


D.2 Word-use judgments

At the end of the main perception experiment, participants judged whether people of Hmong ethnicity were more or less likely to say certain words, compared to whites. Ten words were chosen for these judgments, including: *ancestor*, *bamboo*, *coffee*, *daughter*, *father*, *mom*, *moon*, *new year*, *pet*, and *song*.

For the word *father*, for example, participants saw the prompt screen in Figure D.6.

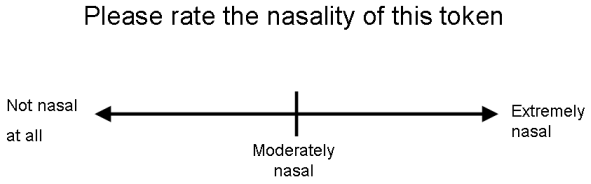
Figure D.6: Prompt screen for question about use of *father*



D.3 Nasality judgments

Perceived nasality for all tokens was judged by trained speech-language pathologists, using a VAS line on the prompt screen shown in Figure D.7.

Figure D.7: Prompt screen for nasality judgments



Appendix E

Linear mixed-effects model coefficients

E.1 Production study models

Table E.1: Cohen's d separation of F1 (normalized) of LOT and THOUGHT by ethnicity: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.33269	0.03351	9.929	5.65e-10
ethnicityEA	0.02180	0.06458	0.338	0.739

Table E.2: Cohen's d separation of F2 (normalized) of LOT and THOUGHT by ethnicity: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.37940	0.03117	12.171	9.31e-12
ethnicityEA	-0.03711	0.06007	-0.618	0.543

Table E.3: Cohen's d separation of F1 (normalized) of LOT and THOUGHT by birthplace for Hmong Americans only: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.41994	0.04447	9.444	3.56e-08
birthplaceNotMN	-0.16578	0.06129	-2.705	0.015

Table E.4: Cohen's d separation of F2 (normalized) of LOT and THOUGHT by birthplace for Hmong Americans only: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.43485	0.04513	9.635	2.66e-08
birthplaceNotMN	-0.10537	0.06221	-1.694	0.109

Table E.5: Cohen's d separation of F1 (normalized) of STAND and CAB by ethnicity: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.27189	0.09064	14.032	4.6e-13
ethnicityEA	0.39453	0.17469	2.259	0.0333

Table E.6: Cohen's d separation of F2 (normalized) of STAND and CAB by ethnicity: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.20326	0.06836	17.603	3.18e-15
ethnicityEA	0.29702	0.13174	2.255	0.0336

Table E.7: Cohen's d separation of F1 (normalized) of STAND and CAB by birthplace for Hmong Americans only: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.5138	0.1282	11.808	1.29e-09
birthplaceNotMN	-0.4596	0.1767	-2.601	0.0187

Table E.8: Cohen's d separation of F2 (normalized) of STAND and CAB by birthplace for Hmong Americans only: fixed effects

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.3002	0.1043	12.466	5.6e-10
birthplaceNotMN	-0.1842	0.1438	-1.281	0.217

Table E.9: Nasalization (*A1-P0*) of STAND and CAB by ethnicity: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-5.3616	0.8994	-5.962	0.0001
ethnicityEA	0.5479	1.4110	0.388	0.6352
TRAPcatSTAND	-2.7460	0.8411	-3.265	0.0026

Table E.10: Nasalization (*A1-P1*) of STAND and CAB by ethnicity: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	12.456	1.357	9.179	0.0001
ethnicityEA	-1.348	1.673	-0.806	0.4146
TRAPcatSTAND	5.776	1.824	3.166	0.0034
ethnicityEA:TRAPcatSTAND	4.461	1.968	2.266	0.0270

Table E.11: GOOSE vowel normalized F2 productions by ethnicity: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-0.13294	0.08542	-1.556	0.0001
ethnicityEA	0.01099	0.01120	0.981	0.5022

E.2 Perception study models

Table E.12: Speaker ethnicity judgments: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-34.7330	11.8389	-2.934	0.0030
TalkerEthnicity1	46.5324	15.1961	3.062	0.0030
ListenerEthnicity1	34.0885	12.3918	2.751	0.0050
WordCategoryGOOSE	24.2720	10.9927	2.208	0.0234
WordCategoryLOT	-17.0237	10.6103	-1.604	0.1248
WordCategoryTHOUGHT	1.3202	10.0812	0.131	0.8524
WordCategoryTRAP	-0.7314	8.4621	-0.086	0.9930
TalkerEthnicity1:ListenerEthnicity1	-10.7486	12.6117	-0.852	0.3976
TalkerEthnicity1:VowelClassGOOSE	-22.1840	20.2352	-1.096	0.2580
TalkerEthnicity1:VowelClassLOT	14.7305	18.4818	0.797	0.4512
TalkerEthnicity1:VowelClassTHOUGHT	-23.0466	18.6673	-1.235	0.1856
TalkerEthnicity1:VowelClassTRAP	-20.5910	16.6191	-1.239	0.1964
ListenerEthnicity1:VowelClassGOOSE	-41.5758	12.5856	-3.303	0.0008
ListenerEthnicity1:VowelClassLOT	5.8682	12.5507	0.468	0.6786
ListenerEthnicity1:VowelClassTHOUGHT	-26.8669	11.8611	-2.265	0.0216
ListenerEthnicity1:VowelClassTRAP	-0.6703	9.9758	-0.067	0.9750
TalkerEthnicity1:ListenerEthnicity1:VowelClassGOOSE	58.0250	23.3956	2.480	0.0136
TalkerEthnicity1:ListenerEthnicity1:VowelClassLOT	8.2555	21.7525	0.380	0.6724
TalkerEthnicity1:ListenerEthnicity1:VowelClassTHOUGHT	47.8378	21.6953	2.205	0.0262
TalkerEthnicity1:ListenerEthnicity1:VowelClassTRAP	33.1293	19.0815	1.736	0.0916

Table E.13: Speaker age judgments: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-28.541	11.133	-2.563	0.0114
TalkerEthnicity1	14.623	13.592	1.076	0.2520
ListenerEthnicity1	-8.665	10.204	-0.849	0.3582

Table E.14: Speaker birthplace judgments: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-26.147	9.263	-2.823	0.0050
TalkerEthnicity1	16.524	10.503	1.573	0.1256
ListenerEthnicity1	40.755	9.098	4.480	0.0001
VowelClassGOOSE	-6.793	5.282	-1.286	0.1628
VowelClassLOT	-5.752	5.204	-1.105	0.2436
VowelClassTHOUGHT	-6.769	5.195	-1.303	0.2064
VowelClassTRAP	-15.040	4.180	-3.598	0.0001
VowelClass1:ListenerEthnicity1	-54.157	6.922	-7.824	0.0001

Table E.15: Speaker friend ethnicity judgments: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-13.0412	9.0027	-1.449	0.1452
TalkerEthnicity1	37.6060	11.8768	3.166	0.0028
ListenerEthnicity1	5.7006	8.3417	0.683	0.4658
VowelClassGOOSE	-21.5644	9.2481	-2.332	0.0186
VowelClassLOT	1.6215	8.8355	0.184	0.8264
VowelClassTHOUGHT	-9.0053	9.2247	-0.976	0.2748
VowelClassTRAP	-5.9341	7.2919	-0.814	0.4288
TalkerEthnicity1:ListenerEthnicity1	14.3424	6.9959	2.050	0.0364
ListenerEthnicity1:VowelClassGOOSE	19.6253	10.5110	1.867	0.0656
ListenerEthnicity1:VowelClassLOT	-12.4310	10.1310	-1.227	0.2168
ListenerEthnicity1:VowelClassTHOUGHT	-0.4037	10.4135	-0.039	0.9964
ListenerEthnicity1:VowelClassTRAP	13.5741	8.3321	1.629	0.1166

Table E.16: Speaker pride in family heritage judgments: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-18.746	8.169	-2.295	0.0144
VowelClassGOOSE	4.855	4.436	1.095	0.2326
VowelClassLOT	-4.511	4.546	-0.992	0.3138
VowelClassTHOUGHT	-6.133	4.451	-1.378	0.1894
VowelClassTRAP	2.706	3.525	0.768	0.3580
TalkerEthnicity1	7.013	9.673	0.725	0.4888

Table E.17: Nasality judgments – all tokens: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-53.622	27.369	-1.959	0.0590
TalkerEthnicity1	-31.815	11.701	-2.719	0.0064
VowelClassGOOSE	17.678	5.616	3.148	0.0004
VowelClassLOT	11.749	5.646	2.081	0.0154
VowelClassTHOUGHT	11.544	5.598	2.062	0.0194
VowelClassTRAP	19.204	4.526	4.243	0.0001
TalkerEthnicity1:VowelClassGOOSE	3.126	10.361	0.302	0.7662
TalkerEthnicity1:VowelClassLOT	16.298	10.355	1.574	0.0776
TalkerEthnicity1:VowelClassTHOUGHT	3.935	10.265	0.383	0.6582
TalkerEthnicity1:VowelClassTRAP	10.388	8.331	1.247	0.1574

Table E.18: Nasality judgments – TRAP tokens only: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-44.23	28.17	-1.570	0.0884
TalkerEthnicity	-15.11	16.06	-0.941	0.3150
TRAPcatSTAND	19.89	6.25	3.182	0.0004
TalkerEthnicity:TRAPcatSTAND	-12.84	11.66	-1.102	0.1860

Table E.19: Word-use judgments: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-110.246	30.720	-3.589	0.0090
ListenerEthnicity1	4.357	27.398	0.159	0.8882
Wordbamboo	-27.337	42.491	-0.643	0.5012
Wordcoffee	235.223	43.584	5.397	0.0008
Worddaughter	79.406	44.321	1.792	0.0964
Wordfather	100.184	44.321	2.260	0.0464
Wordmom	81.681	45.222	1.806	0.0938
Wordmoon	99.030	44.321	2.234	0.0528
Wordnewyear	5.807	44.321	0.131	0.8944
Wordpet	221.923	43.584	5.092	0.0014
Wordsong	127.736	45.224	2.825	0.0234
ListenerEthnicity1:Wordbamboo	44.633	37.450	1.192	0.2264
ListenerEthnicity1:Wordcoffee	-10.371	38.686	-0.268	0.7942
ListenerEthnicity1:Worddaughter	-9.851	39.514	-0.249	0.8196
ListenerEthnicity1:Wordfather	-60.739	39.514	-1.537	0.1268
ListenerEthnicity1:Wordmom	109.578	40.522	2.704	0.0092
ListenerEthnicity1:Wordmoon	-28.400	39.515	-0.719	0.4684
ListenerEthnicity1:Wordnewyear	96.987	39.619	2.448	0.0154
ListenerEthnicity1:Wordpet	-37.027	38.793	-0.954	0.3490
ListenerEthnicity1:Wordsong	-4.699	40.525	-0.116	0.9166

Table E.20: Speaker ethnicity judgments – including word-use judgment effect: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-33.646185	11.525095	-2.919	0.0032
ListenerEthnicity1	17.478939	11.447785	1.527	0.1098
Worduse	-0.047242	0.060102	-0.786	0.4806
TalkerEthnicity	31.984381	15.707196	2.036	0.0372
ListenerEthnicity1:Worduse	0.090150	0.071665	1.258	0.2418
ListenerEthnicity1:TalkerEthnicity	26.500281	12.365440	2.143	0.0388
Worduse:TalkerEthnicity	-0.005267	0.105028	-0.050	0.9806
ListenerEthnicity1:Worduse:TalkerEthnicity	-0.019966	0.124021	-0.161	0.9130

Table E.21: Speaker ethnicity judgments for the TRAP vowel, with acoustic factors: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-27.674	14.765	-1.874	0.0448
ListenerEthnicity1	28.863	16.086	1.794	0.0552
z.a1p0	3.926	6.292	0.624	0.5418
z.f2norm	4.011	11.099	0.361	0.6726
ListenerEthnicity1:z.a1p0	9.286	6.964	1.333	0.1662
ListenerEthnicity1:z.f2norm	18.879	12.796	1.476	0.1428

Table E.22: Speaker ethnicity judgments for the STAND allophone of the TRAP vowel, with acoustic factors: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-41.310	25.078	-1.647	0.0814
ListenerEthnicity1	17.358	28.761	0.604	0.4936
z.a1p0	-7.389	12.368	-0.597	0.6166
z.f2norm	6.939	17.617	0.394	0.6444
ListenerEthnicity1:z.a1p0	25.616	14.236	1.799	0.0828
ListenerEthnicity1:z.f2norm	38.522	20.000	1.926	0.0644

Table E.23: Speaker ethnicity judgments for the CAB allophone of the TRAP vowel, with acoustic factors: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-35.65	15.24	-2.339	0.0100
ListenerEthnicity1	37.80	15.19	2.489	0.0054
z.a1p0	10.40	5.18	2.008	0.0372
z.f2norm	27.64	15.40	1.794	0.0680

Table E.24: Speaker ethnicity judgments for the LOT vowel, with acoustic factors: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-9.198	24.733	-0.372	0.7054
z.a1p0	10.073	6.271	1.606	0.0968
z.a1p1	-15.540	9.293	-1.672	0.0782
ListenerEthnicity1	-10.876	25.417	-0.428	0.6816
z.f2norm	36.653	26.978	1.359	0.1744
ListenerEthnicity1:z.f2norm	-62.721	29.267	-2.143	0.0462

Table E.25: Speaker ethnicity judgments for the THOUGHT vowel, with acoustic factors: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	71.26	32.82	2.171	0.0194
ListenerBornMN1	19.82	24.79	0.800	0.4006
z.f1norm	-32.79	23.36	-1.403	0.1172
ListenerEthnicity1	-23.63	25.77	-0.917	0.2878
z.f2norm	77.06	22.81	3.379	0.0004
ListenerBornMN1:z.f1norm	-48.58	28.87	-1.683	0.1070
z.f1norm:ListenerEthnicity1	76.79	30.03	2.557	0.0102

Table E.26: Speaker ethnicity judgments for the GOOSE vowel, with acoustic factors: fixed effects

	Estimate	Std. Error	t value	pMCMC
(Intercept)	-27.128	14.802	-1.833	0.0464
ListenerBornMN1	20.465	16.633	1.230	0.2178
z.a1p1	11.791	8.678	1.359	0.2070
z.f2norm	11.631	5.653	2.058	0.0252
z.f1norm	-14.010	9.700	-1.444	0.1058
ListenerBornMN1:z.a1p1	-18.920	9.702	-1.950	0.0498