SPONTANEOUS NASALIZATION: AN ARTICULATORY INVESTIGATION OF GLOTTAL CONSONANTS IN THAI

BY

SARAH E. JOHNSON

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Linguistics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Associate Professor Ryan K. Shosted, Chair and Director of Research
Associate Professor Chilin Shih
Assistant Professor Keiko Ishikawa
Professor Brad P. Sutton

Abstract

Vowel nasalization usually occurs through a two-step process whereby a vowel is nasalized via coarticulation with a nearby nasal segment; when the language later drops the nasal segment, a nasal vowel remains. Spontaneous vowel nasalization is a rare, peculiar form of nasalization that emerges in contexts that lack an historical etymological nasal (Blevins and Garrett, 1992; Matisoff, 1975; Ohala, 1975, 1974). Spontaneous vowel nasalization has been found to emerge in contexts with glottal consonants, low vowels, and transitionally breathy segments such as fricatives and aspirates. Blevins and Garrett (1992) classified these contexts as *rhinoglottophilia* ‘nose-larynx-affinity’, *rhinochthamalophilia* ‘nose-low/flat-affinity’, and *rhinosyrigmatophilia* ‘nose-whistle-affinity’. Many articulatory and acoustic explanations have been proposed for spontaneous nasalization. First, vowels may nasalize near glottal consonants because they are underspecified for velopharyngeal opening and thus may be produced with a lowered velum. Second, low vowels may nasalize because they can better tolerate the acoustic consequences of greater nasal coupling because the intraoral acoustic impedance of low vowels is naturally lower. Finally, speakers may nasalize vowels in breathy environments through misperception or enhancement because nasality and breathiness are acoustically similar (Matisoff, 1975; Ohala and Amador, 1981). Due to the acoustic similarities between nasalization and breathiness, e.g. high spectral tilt and weakened F1, disentangling the relative roles of physiological nasalization and breathiness during spontaneous nasalization is not possible using acoustic data alone. There have been few articulatory investigations that assess the origins of spontaneous nasalization (Clumeck, 1975; Ohala, 1971).

Thai arguably manifests an interaction among rhinoglottophilia, rhinochthamalophilia, and rhinosyrigmatophilia-based spontaneous nasalization. Thai is known to nasalize low and mid-low vowels after the voiceless glottal fricative /h/ and glottal stop /ʔ/, such that /hɛ˧/ [hɛ̃˧] ‘parade’ and /ʔaw˧/ [ʔãw˧] ‘to take’ are reported to sound nasal (Cooke, 1989; Matisoff, 1975; Noss, 1964). Vowels following /h/ are reportedly more susceptible to nasalization than those following /ʔ/ (Matisoff, 1975; Cooke, 1989). Furthermore, of the nasalized vowels, Cooke (1989) observed that /a/ is reported to sound more nasal than /ɛ/ and /ɔ/. In Thai it is possible that the velopharyngeal underspecification of glottal consonants, low tongue height, and breathy-nasal acoustic similarity (in proximity to /h/) may all play a role in nasalization. In order to deepen our understanding of these potentially interacting sources of spontaneous nasalization in Thai, this dissertation undertakes an extensive articulatory investigation of physiological nasalization and breathiness during spontaneously nasalized vowels in Central Thai.

Spontaneous nasalization has been an historically difficult topic of study because its analysis requires a combination of articulatory data that measures both nasalization and phonation quality. This dissertation integrates measures from state-of-the-art ultra-fast magnetic resonance imaging (MRI), aerodynamics, and electroglottography (EGG) to estimate physiological nasalization and breathiness. The speech of ten speakers of Central Thai was assessed using these tools: Four speakers were recorded using ultra-fast MRI and six speakers were recorded using aerodynamics and EGG. Speakers produced CV syllables that varied by onset consonant and vowel height within the same carrier phrase. Syllables with an onset glottal consonant /h, ʔ/ were varied at four different vowel heights; syllables with onset consonants /n, t^{h}, d/ were included as nasal and oral controls.

The major finding of this dissertation is that /h/-onset syllables are produced with greater physiological nasalization than /ʔ/-onset syllables. This finding suggests that, counter to previous claims that glottal consonants are generally underspecified for velopharyngeal opening in Thai, /ʔ/ is in fact specified, at least in a prosodically prominent context. Furthermore, using EGG we observed a slight increase in breathiness during vowels just after /h/, but not after /ʔ/, suggesting a potential relationship between nasalization and breathiness in Thai. The voiceless turbulence of /h/ induces coarticulatory breathiness during the following vowel; this breathiness may facilitate the perception of nasalization during the following vowel. This might explain why vowels after /h/ reportedly sound more nasal than vowels after /ʔ/ (Cooke, 1989; Matisoff, 1975). Further perceptual testing is needed to assess this possibility.

Furthermore, we observed inconsistent and often minimal variation in physiological nasalization during vowels of varying height in both nasalized and non-nasal context. This suggests that spontaneous vowel nasalization in Central Thai may now be similar to Northeastern Thai, a dialect where all vowels were reported to nasalize after glottal consonants approximately 40 years ago (Matisoff, 1975). Finally, during /hV/ syllables, all /h/ consonants are produced with greater nasal airflow than the following vowel. The onset /h/ appears to be the locus of nasalization that spreads to the vowel through coarticulation.

The results of this dissertation deepen our understanding of the production of spontaneous nasalization and contexts that facilitate its realization in Thai and perhaps cross-linguistically as well. Our results demonstrate that spontaneous nasalization in Thai is primarily attributed to rhinoglottophilia: velopharyngeal underspecification of onset /h/. Rhinosyrigmatophilia may potentially also play a role: the presence of breathiness after /h/ may enhance the percept of nasalization. Nasal coupling and breathiness may be integrated into a single acoustic object that the listener perceives as nasal. This possibility presents an intriguing example of the many-to-one problem, whereby many possible articulatory configurations may result in similar acoustic output (Maeda, 1990).

In memory of my grandfather, Bloice W. Yount who attended the University of Illinois at Urbana-Champaign in the early 1930's. After the passing of his father, Bloice returned home to manage the family farm and was unable to finish his degree. With this dissertation I help realize his dream.

Acknowledgments

I am grateful to my committee members for their support and guidance during this project. I thank Dr. Brad Sutton and the Magnetic Resonance Functional Imaging Lab for their collaboration and use of equipment in the Bioimaging Center at the Beckman Institute. I also thank Dr. Keiko Ishikawa for her insights about the larynx and for allowing me to borrow her electroglottography equipment.

I would like to especially thank my adviser, Dr. Ryan Shosted. His guidance, discussions, and invaluable feedback have not only benefited this dissertation but my professional development as a scholar over the years.

Thank you to the National Science Foundation Graduate Research Fellowship Program whose support I relied upon under Grant No. DGE–1144245.

I thank the Department of Linguistics for their support reimbursing participants of the aerodynamic experiment.

A special thanks to Pan Fry for his assistance in translating the Thai speech material in this dissertation, for his native speaker intuitions, and for his assistance recruiting participants. I thank Nancy Dodge and other technicians at the Beckman Bioimaging center for their invaluable assistance in creating and executing MRI protocols in the collecting of the MRI data for this dissertation.

I thank the many mentors I have had along the way. A special thanks to Dr. Nicole Wong for inspiring me to pursue phonetics research and training me to conduct my first experiment in the lab. I thank Dr. Marissa Barlaz for her invaluable help and support, for training me how to collect MRI data, and how to implement advanced statistical models. Without Dr. Barlaz's guidance and mentoring, many of the statistical models in this dissertation would not have been possible.

I thank two anonymous reviewers, Dr. Marc Garellek, and Dr. Taehong Cho, the Editor-in-Chief of the *Journal of Phonetics*, for their insightful comments and suggestions that helped strengthen the presentation of the MRI data. I also thank audience members at the 174th Meeting of Acoustic Society America for their comments.

A special thanks to my mom and my fiancé for their love, support, and encouragement over the years.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

In many languages of the world, nasalization and glottal configuration are related (Blevins and Garrett, 1992; Carignan, 2017; Garellek *et al.*, 2016; Matisoff, 1975; Ohala, 1975; Ohala and Amador, 1981). Because the physiological mechanisms underlying the production of nasal and glottal sounds are so different and clearly independent of one another, this pattern has been the subject of considerable interest among linguists. Researchers have asked whether speakers may confuse nasal sounds with certain glottal sounds or produce certain glottal sounds with an inherent degree of nasality (Blevins and Garrett, 1992, 1993; Matisoff, 1975; Ohala, 1971, 1974). Until recently, it has been difficult to simultaneously examine all the physiological mechanisms of the vocal tract at work, and so answers to these questions have remained elusive or partial.

Nasal vowels typically arise when an oral vowel is in proximity to a nasal consonant through coarticulation. For example, Latin UNUS 'one' became French [œ̃] *un* when the nasal consonant was lost (Ohala, 1980). The historical development of nasal vowels in the phonological context of other nasal speech sounds is typologically common and well-studied. Laboratory work on the acoustics, perception, and articulation of nasality has broadened and enriched our understanding of the diachronic and synchronic characteristics of oral-nasal vowel systems (Fujimura and Lindqvist, 1971; Maeda, 1993; Beddor, 1993; Shosted *et al.*, 2012a; Carignan *et al.*, 2015; Barlaz *et al.*, 2018). However, contextual coarticulatory nasalization is not the only source of diachronic nasal vowels: nasalization may also take place in contexts lacking any etymological nasal. This phenomenon, called spontaneous nasalization, is defined as, "[T]he emergence of distinctive nasalization on a vowel when there was no historical antecedent to it" (Ohala and Ohala, 1992, p. 1034). Examples include British English [hã:vəd] 'Harvard', Hindi [sãp] 'snake' from Sanskrit *sarpa*, and Thai [hɛ̃ː] 'parade' (Ohala, 1980). Spontaneously nasalized vowels have been found to develop in contexts with a glottal consonant (British English, Thai, Lao) and with a fricative or aspiration (British English, Hindi, Bzhedukh, Thai); most cases occur on low or mid-low vowels (Blevins and Garrett, 1992, 1993; Matisoff, 1975). See Blevins and Garrett (1992; 1993) for a review of cases of spontaneous nasalization.

While the articulatory and perceptual processes involved in typical synchronic and diachronic vowel nasalization are relatively well documented (Maeda, 1993; Beddor, 1993; Stevens, 2000; Shosted *et al.*, 2012b; Carignan *et al.*, 2015; Barlaz *et al.*, 2017, among many others), spontaneous nasalization remains somewhat mysterious. Several articulatory, aerodynamic, and perceptual explanations have been offered for this phenomenon. Spontaneous nasalization is thought to be induced by velopharyngeal underspecification of glottal consonants, misperception of nasalization facilitated by breathiness, and / or variations in velum position when the tongue is low (Blevins and Garrett, 1992; Matisoff, 1975; Ohala, 1993b, 1975, 1974, 1971). The problem of spontaneous nasalization is a complicated one because a comprehensive analysis of the phenomenon requires articulatory measures of both the velum and the larynx along with the associated acoustic output. Because nasalization and breathiness can produce similar acoustic effects, e.g. amplification of low frequency harmonics and F1 weakening, it is not possible to rely on acoustics alone to understand the articulation of spontaneous nasalization. Non-invasive measures of the posterior vocal tract can be difficult to obtain.

There is a long-standing problem in phonetics called the "many-to-one" problem whereby many possible articulatory configurations may result in similar acoustic output (Maeda, 1990). Speech perception is viewed as the ultimate target of speech production, such that articulation culminates in a single acoustic product that is perceived and interpreted by the listener (Diehl *et al.*, 1990; Kingston, 1991; Ohala, 1996). As long as the acoustic output of a particular sound category is similar across the population, variation in its articulation can be tolerated. Research has shown that nasalization generally is an interesting example of the many-to-one problem, as speakers often modulate position of the velum, tongue, lips, and pharynx to achieve the acoustic product of nasalization (Engwall *et al.*, 2006; Carignan *et al.*, 2015; Barlaz *et al.*, 2018; Shosted *et al.*, 2012a). Spontaneous vowel nasalization is another excellent example of this problem because, as explained in Chapter 2, configurations of the velum, larynx, and tongue can influence the perception of spontaneously nasalized vowels.

Ohala (1993b) argues that many cases of diachronic sound change arise from a synchronic change when listeners misperceive the speech signal, reinterpret the articulatory configuration, and produce the sound in a different way. Variations in production can arise due to hypocorrection, when a listener undercorrects a coarticulated production and reinterprets it as phonologically intended. Misperception and change may also occur when two speech sounds are acoustically similar and a listener misinterprets one segment for another (Garrett and Johnson, 2013; Ohala, 1993a). Many sound changes are errors in production and perception that can provide clues about the vital acoustic features necessary for successful categorical speech perception. These kinds of sound change can be thought of as "nature's speech perception experiment" (Ohala, 1993b, p. 1). By understanding how variations in articulation can induce such sound changes, we may gain a better understanding of the relationship between the production and perception of speech. With this dissertation I hope to contribute to our understanding of the relationship between the larynx and the velum in spontaneous nasalization by studying Thai, a language known to nasalize low vowels after glottal consonants Cooke (1989); Matisoff (1975); Noss (1964).

Chapter 2

LITERATURE REVIEW

2.1 Spontaneous nasalization: Classification and cross-linguistic observations

The study of spontaneous nasalization dates back to at least the early 1920's (Grierson, 1922; Turner, 1921). In these earlier writings, researchers use the term spontaneous nasalization to refer to nasalization that emerges where no nasal segment existed before. Regarding Gujarati phonology, Turner (1921) writes, "There seems from the earliest times to have been a tendency to pronounce vowels with the velum incompletely raised, which results in the vowel becoming nasalized." While other scholars later showed that Gujarati nasal vowels likely originated from a previous nasal segment and were thus not spontaneous (Blevins and Garrett, 1993; Grierson, 1922; Schwyzer, 1934), Turner's initial use of the term spontaneous nasalization set the stage for later work on nasalization of seemingly unexplained origin.

Matisoff (1975) coined the term *rhinoglottophilia*, 'nose-larynx-affinity' to describe a specific kind of spontaneous nasalization that is associated with glottal consonants. Matisoff discussed several examples from British English (Indo-European), Thai (Kra-Dai), Lao (Kra-Dai), Lahu (Sino-Tibetan), Lisu (Sino-Tibetan), and East Gurage (Afro-Asiatic) where vowel nasalization developed in proximity to /h/ and /ʔ/. In British English (Matisoff, 1970), syllables beginning with /h/ or a vocalic onset (no consonant) and containing the low vowel /a/ are nasalized. Examples include 'half' [hãːf], 'art' [ãːt], 'heart' [hãːt], and 'hour' [ãːə]. Languages with a similar nasalization process include Lahu and Lisu (Lolo-Burmese) where the vowels /ɔ/ and /a/ nasalize after /h/ or a vocalic onset. For example 'to bend' /ɔ/

becomes [ɔ̃] and ‘likeness, spirit’ /ɔha/ becomes [ɔ̃hã]. Two other examples include Thai and Lao where low vowels nasalize after /h/ and /ʔ/ so that [hɛ̃] ‘parade’ and [ʔãw] ‘to take’ are nasalized (Matisoff, 1975).

Matisoff also observed that most of these examples occurred during low or mid-low vowels, a phenomenon he termed *nasovocogenesis* ‘nasal-vowel-origin’. Ohala provides two potential explanations for rhinoglottophilia. The first is velopharyngeal underspecification (VPU) of glottal consonants. According to this explanation, glottal consonants are underspecified for velopharyngeal opening (VPO) because the place of primary constriction is posterior to the velum. Therefore, glottal consonants may be produced with large VPO, causing the surrounding vowels to nasalize through coarticulation (Ohala, 1974, 1975, 1971). Ohala’s second explanation for rhinoglottophilia involves only the fricative /h/ and attributes nasalization in proximity to /h/ as arising due to acoustic similarity between breathiness and nasalization, e.g. high spectral tilt (increased amplitude at low frequencies) and a weakened F1. The segment /h/ is often produced with transient physiological breathiness because laryngeal configuration during the transition from glottal frication to modal voicing is briefly conducive to breathy phonation (Blankenship, 2002; Miller, 2007; Stevens, 2000). Ohala reasons that because the glottis is open and some coarticulatory breathiness may be present, vowels surrounding /h/ may sound nasal. Finally, Matisoff and Ohala explain that low vowels may generally be produced with increased VPO because low vowels can better tolerate acoustic coupling with the nasal cavity; this is because low vowels have less intraoral acoustic impedance compared to high vowels and thus the acoustic effects of nasal coupling are less pronounced for low vowels (House and Stevens, 1956). We explore these explanations for rhinoglottophilia and nasovocogenesis in more detail in Sections 2.2.1 through 2.2.3.

Spontaneous nasalization has been observed in many contexts cross-linguistically (Blevins and Garrett, 1992). For example, Dingemanse *et al.* (2013) observed spontaneous nasalization in a survey of interrogative interjections, e.g. ‘huh?’, in ten typologically diverse languages: Cha’palaa /ʔaː/, Dutch /hɜ/, Icelandic /ha/, Italian /ɛː/, Lao /haː/, Mandarin

Chinese /aː/, Murriny Patha /aː/, Russian /aː/, Siwu /aː/, and Spanish /e/. They found that nearly all interjections surveyed were monosyllables composed of a glottal stop or /h/ followed by a nasalized low or mid-low vowel or a monosyllable with a nasalized low or mid-low vowel onset. The exception was Spanish /e/, although the specific dialect was not provided. They reported hearing impressions of nasalization during the interjections of most languages, particularly Lao, Mandarin, Murriny Patha, and Siwu (Dingemanse *et al.*, 2013).

Blevins and Garrett (1992) further stratified spontaneous nasalization into three categories: *rhinoglottophilia*, *rhinochthamalophilia*, and *rhinosyrigmatophilia*. Similar to Matisoff's (1975) earlier work, they define rhinoglottophilia, 'nose-larynx-affinity', as nasalization associated with glottal consonants. Rhinochthamalophilia 'nose-low/flat-affinity' is defined as nasalization associated with low vowels or pharyngeal consonants. This category subsumed Matisoff's original nasovocogenesis (low vowels) category but is broadened to include pharyngeals. The inclusion of pharyngeals was inspired by Eastern Gurage, where vowels near pharyngeal consonants often nasalize. Hetzron (1969) explains that nasalization near pharyngeals likely occurs because pharyngealization is produced with a lowered uvula and thus increased VPO.[1] Finally, Blevins and Garrett define rhinosyrigmatophilia 'nose-whistle/pipe-affinity' as nasalization associated with sibilants. This category was created because of observations that nasalization sometime originates near sibilants, as in Hindi, Bzhedukh, and Shapsegh. For example, in Bzhedukh and Shapsegh vowels are allophonically nasalized following aspirated sibilants but remain oral following unaspirated sibilants. For example, Bzhedukh /ʃ^{h}aʃ^{h}/ 'horse's milk' becomes [ʃ^{h}ɪ̃ʃ^{hn}].[2] (Colarusso, 1988) A similar process has occurred in Shapsegh where an older distinction between aspirated and unaspi-

---

[1] Pharyngeals can potentially be explained by both rhinoglottophilia or rhinochthamalophilia. The primary place of articulation is posterior to the velum, therefore the velum is possibly underspecified for pharyngeal consonants. However, it has been reported that a lowered uvula is an inherent articulatory feature of pharyngeal, which may also work to increase velopharyngeal area (Blevins and Garrett, 1992; Hetzron, 1969). The precise classification of spontaneous nasalization associated with pharyngeals is therefore still open to debate and requires further research.

[2] Colarusso (1988) reports nasalization during the aspirated coda. It is unclear whether the superscript /n/ is mean to indicate alveolar closure or the presence of nasalization during aspiration. Furthermore, it is unclear how salient nasalization during the aspirated coda is if the aspiration is voiceless.

rated sibilants has been lost. Instead, the aspirated sibilants have been reinterpreted as nasalization and a new nasal vowel allophone has emerged (Colarusso, 1988; Blevins and Garrett, 1992). Blevins and Garrett admit that this type of spontaneous nasalization likely has a similar explanation as the breathy-nasal misperception-based rhinoglottophilia associated with /h/. They reason that voiceless sibilants may have a similar glottal configuration as /h/ and may be produced with coarticulatory breathiness, thus sounding nasal.

While the classification system created by Blevins and Garrett (1992) covers most types of spontaneous nasalization, aspiration-based nasalization, as in Bzhedukh, Shapsegh, and Ponapean, do not precisely fit within the three *rhino-* categories. For example, Ponapean and Mokilese exhibit a process where voiceless geminate stops undergo nasal substitution: CC → NC. Blevins and Garrett (1993) explain that Ponapean and Mokilese geminate stops were preaspirated and that this aspiration was likely later reanalyzed as nasalization. They denote this change as such: CC → hC → NC. They reason that aspiration-based nasalization may involve a combination of rhinoglottophilia and rhinosyrigmatophilia because aspiration is /h/-like and thus may be acoustically similar to nasalization. We expound upon the similarities between /h/ and aspiration below.

In order to provide a more specific classification for aspiration-based nasalization, I propose the following minor simplification to Blevins and Garrett (1992)'s spontaneous nasalization classification scheme. The category rhinosyrigmatophilia should broadly include any voiceless fricatives and aspiration that are associated with nasalization. From an articulatory standpoint, this category includes segments that are produced with voiceless glottal frication, i.e voiceless turbulence channeled through the vocal tract that can induce coarticulatory breathiness during the transition into a surrounding modally-voiced vowel (Blankenship, 2002; Miller, 2007; Stevens, 2000). This redefinition would subsume all voiceless sibilants, fricatives (including /h/), and aspiration. The voiceless frication shared among these segments is important irrespective of the position of the supraglottal articulators. For example, /h/ and aspiration are produced with voiceless glottal frication, while the tongue may be in position for a surrounding segment. For example, English 'he' is produced with a high tongue during the /h/ in anticipation of the /i/, resulting in [çi] with a palatal fricative. The word 'hawk' is produced with the tongue low during the /h/, resulting in a true glottal fricative [h] with no substantial supralaryngeal narrowing. Lingual and labial position is similarly determined during consonant aspiration. For fricatives such as /s/ or /ʃ/, the tongue is at the primary place of articulation for the fricative such as on or behind the alveolar ridge. The key commonality among /h/, aspirates, and other voiceless fricatives is that all are produced with an open glottis and glottal frication that can exhibit coarticulatory breathiness during the transition with an adjacent, modally voiced vowel. As described in Section 2.2.3 in more detail, such environments are usually produced with coarticulatory breathiness that may facilitate the perception of nasalization (Ohala and Amador, 1981; Blankenship, 2002; Miller, 2007; Stevens, 2000).

It is important to note that this modified classification,/h/ can be described both within rhinoglottophilia and rhinosyrigmatophilia. This overlap is intended. /h/ is both a glottal consonant and a fricative that likely induces coarticulatory breathiness. Therefore, either VPU and/or breathy-nasal acoustic similarity may explain nasalization in proximity to /h/. As this dissertation will show, rhinoglottophilia and rhinosyrigmatophilia may play an important interacting role in Thai spontaneous nasalization (see Chapter 5.3). Table 2.1 provides a summary of types of spontaneous nasalization, the proposed acoustic/articulatory explanation, and any modifications made by the author. In Table (2.2) we provide a cross-linguistic overview of attested cases of spontaneous nasalization described above. In the following section (2.2) we provide a more detailed overview of articulatory and acoustic explanations for spontaneous nasalization.

Classification	Definition	Explanations
rhinoglottophilia	‘nose-larynx-affinity’ Nasalization associated with glottal consonants /h/ /ʔ/	VPU (Ohala, 1971, 1974)
rhinochthamalophilia	Nasalization associated with low vowels and pharyngeals	Low intraoral acoustic impedance of low vowels (Ohala, 1974; House and Stevens, 1956); lowered uvula during pharyngeals (Hetzron, 1969)
rhinosyrigmatophilia	Nasalization associated with voiceless glottal frication during fricatives and aspirates*	Breathy-nasal acoustic similarity (Ohala, 1974)

Table 2.1: Blevins and Garrett (1992) and Matisoff (1975)’s classification of spontaneous nasalization with the author’s modifications (marked by a star*), proposed explanations for spontaneous nasalization.

Type	Attested contexts
Rhinoglottophilia	/h/: British English, Thai, Laos, Lahu, Lisu /ʔ/: Thai, Lao
Rhinochthamalophilia	Low/mid-low vowels: British E., Lahu, Lisu, Mandarin*, Murriny Patha*, Siwu*, Thai, Lao Pharyngeals: East Gurage
Rhinosyrigmatophilia	Aspiration: Bzhedukh, Shapsegh, Ponapean, Mokilese Fricatives: British English, Thai, Laos, Lahu, Lisu, Hindi

Table 2.2: Cross-linguistic observations of spontaneous nasalization. A star* indicates that the only reported spontaneous nasalization in the language were interrogative interjections observed in a survey by Dingemanse *et al.* (2013).

2.2 Articulatory and acoustic explanations

2.2.1 Rhinoglottophilia

Ohala (Ohala, 1971, 1974) proposed an articulatory explanation for nasalization near glottal consonants: velopharyngeal underspecification (VPU). A segment is considered underspecified for a feature if it has no value for that feature; it may be produced with more than one value because it has none of its own (Keating, 1988). An example is lip rounding during English /t/. English /t/ is typically produced with the labial configuration of a surrounding vowel, so that /tu/ 'two' is produced as [t^{w}u] and /ti/ 'tea' is produced as [ti]. Because /t/ is phonetically underspecified for lip rounding and one may produce a perfectly acceptable /t/ regardless of lip configuration in English, coarticulation determines labial configuration(Keating, 1988). Likewise, glottal consonants such as /h/ and /ʔ/ do not necessarily require the same degree of velopharyngeal closure as buccal consonants. Ohala (Ohala, 1974) reasons that this is likely because glottal consonants require constriction or pressure build up behind the velum and therefore have no specification for velum position. Indeed, glottal consonants are often transparent to the phonological process known as nasal harmony (Walker, 2000; Blevins and Garrett, 1992). However, buccal consonants require pressure build up anterior to the velum. Leakage of air through the velum is antagonistic to the release burst of plosive consonants.

As introduced in Section 2.1, Dingemanse *et al.* (2013) performed a survey of interrogative interjections, or brief interjections meant to express confusion or request clarification, e.g. English 'huh?', in ten typologically diverse languages. They find that nearly all interjections are monosyllables composed of a glottal stop, /h/, or no consonants followed by a nasalized low or mid-low vowel. This observation is not surprising when we consider that the natural resting position of the velum is lowered and that glottal consonants do not require velar elevation (Hixon *et al.*, 2014). Many languages appear to spontaneously nasalize interrogative interjections that begin with a glottal consonant followed by a low vowel. However, rather than developing nasalization at a later time, it may be that these vowels were already nasalized by default because of the lowered natural resting position of the velum.[3]

Ohala (1971) used a nasograph to assess velar configuration during glottal consonants. A nasograph uses a light and a light sensor to detect degree of velar opening, similar to glottography. Ohala observed that glottal consonants /h/ and /ʔ/ were produced with a similar degree of velar opening as /n/ during nasal contexts in one English speaker. He reasons that glottal consonants in English "...allow the velar elevation to be determined by neighboring consonants and vowels.' (Ohala, 1971). Ohala further speculated that it may be possible to produce an acoustically acceptable version of glottal consonants in English, regardless of the state of the velum. He reasons that the build-up of air pressure behind a pharyngeal or glottal constriction would not be impeded by an opening in the velopharyngeal port Ohala (1975).

2.2.2 Rhinochthamalophilia

Blevins and Garrett (1992) coined the unique term *rhinochthamalophilia*, 'nose-low/flat-affinity', or nasalization induced by low vowels. What cases of spontaneous nasalization often have in common is that all either strictly occur with low vowels (Thai and Lao) or are more likely with low vowels (Lahu and Lisu). It has been observed that nasalization is more likely to occur with low vowels in many languages, for both acoustic and articulatory reasons (Bell-Berti *et al.*, 1979; Bell-Berti, 1993; Hajek and Maeda, 2000; Ruhlen, 1973). Researchers have observed that the degree of VP opening is often indirectly correlated with vowel height (Clumeck, 1973; Bell-Berti *et al.*, 1979, nasograph, endoscope), while there can be great individual variation across languages and speakers (Al-Bamerni, 1983; Clumeck, 1976).

Low oral vowels are even reported to accompany greater VP opening as compared to high oral vowels (Fant, 1960; Hiroto *et al.*, 1964, X-ray, cineradiography). Using MRI, Whalen

---

[3] A special thanks to Marc Garellek for his insights on this topic.

(1990) reports that the low oral vowel /a/ is produced with the velum approximately 3 mm lower than high vowels. One reason for this tendency may be physiological. The tongue and soft palate are connected via the palatoglossus muscle. When the tongue is lower, researchers have reasoned that pull on this muscle induces downward pressure on the velum, resulting in greater velum lowering (Hajek and Maeda, 2000; Ruhlen, 1973). While this suggests that velum lowering during low vowels is a passive process, electromyographic (EMG) study reveals that in American English vowels the levator veli palatini is actively engaged in lifting the velum during the production of /i/ and /u/ as compared to /a/ (Clumeck, 1975). EMG results show that these same speakers also contracted the palatoglossus muscle during /a/ to a greater extent than both /u/ and /i/. This suggests that during the production of high vowels in American English, the velum is actively raised for high vowels (via the levator veli palatini) and actively lowered and pulled forward for low vowels (via the palatoglossus).

An explanation for this difference in velum position is both acoustic and perceptual. Researchers have found that low vowels are better able to "tolerate" VP opening than high vowels in that the acoustic consequences of velum lowering appear to be greater for high vowels as compared to low vowels (Ohala, 1974). Given the same degree of nasal coupling in a low and high vowel, the nasal resonances will be weaker during the low vowel and will have less of a shifting effect on oral formants; low vowels are less likely to be confused with other vowels due to formant shift. House and Stevens (1956) explains that the degree of impedance within the vocal tract directly influences the acoustic effects of nasal coupling. As mentioned earlier, because the nasal cavity is smaller and radiates less acoustic energy than the oral cavity, its main role when coupled is to alter the primary oral acoustic output. When the impedance in the mouth increases in high vowels (i.e, when the tongue is close to the palate), the interacting acoustic output of nasal-oral coupling will be affected more by nasal coupling. This heightened interaction "...will occur in the frequency ranges where the difference in the impedance is greatest, particularly in the vicinity of the first vowel formant" (House and Stevens, 1956, p. 223). Ohala (1974) even argues that EMG evidence of active velum lowering during low vowels shows that low vowels actually require some velopharyngeal coupling as part of their acoustic contrast with high vowels. This may explain why low vowels are more likely to nasalize during sound change; they are already somewhat nasalized.

2.2.3 Rhinosyrigmatophilia

According to the source-filter theory of speech production (Fant, 1960), an acoustic source generated at the larynx is shaped by the resonant properties of the vocal tract. Modulation of anatomical parts within the vocal tract, including the pharynx, tongue, jaw, and lips, results in varying vowel qualities and consonants. The acoustic output of this resonance system can become complicated by the existence of branching cavities. Two of these branches include the nasal tract accessible via the nasopharynx, and the subglottal tract (sublaryngeal cavity and trachea) located below the laryngeal vibratory source. The nasal and subglottal tracts can become coupled by lowering the velum or opening the glottis for voiceless or breathy phonation, respectively. When these cavities are coupled with the oral tract and the acoustic vibrations generated in the larynx resonate within them, they produce a series of poles and zeros (formants and anti-formants). These poles and zeros interact with the oral formants produced via resonance within the primary oral tract (Stevens, 2000; Styler, 2017).

The acoustic consequences of nasal and subglottal coupling are distinct, yet similar in many ways. In many languages detailed previously in Section 2.1, nasalized vowels develop diachronically in contexts with increased coarticulatory breathiness such as voiceless fricatives and aspiration. Nasalization associated with voiceless frication may be due to misperception between the acoustic properties of nasalization and breathiness (Matisoff, 1975; Ohala, 1974). The physiological and acoustic underpinnings of such diachronic changes are currently not well understood.

The most widely accepted explanation for nasalization near a voiceless, wide-glottis segment involves misperception of the acoustic signal associated with nasalization and breathiness. The transition between articulations with a wide glottis like /h/ or aspiration and a following modally-voiced vowel typically involves a brief time where the vocal folds are in a configuration conducive for breathy phonation (Blankenship, 2002; Miller, 2007; Stevens, 2000). Blankenship (2002) observed increased breathiness during vowels adjacent to /h/ and aspiration. Stevens (2000)[91] explains:

> During speech production in some languages it often happens that that vocal folds are manipulated from an abducted configuration, for which the vocal folds do not vibrate, to a more adducted configuration, for which normal vocal fold vibration occurs. During this time interval, the vocal folds pass through a configuration for which breathy phonation occurs... The first few cycles of vibration tend to be smooth pulses that have weak high-energy but strong low-frequency energy, and there is a large average airflow during this time interval when breathy phonation is initiated.

This configuration, which may be described as a transition from voicelessness to voicing, is characterized by vocal folds that are vibrating but not yet touching, a direct current of air resulting from lack of vocal fold contact and / or the presence of a posterior glottal gap and a more coupled trachea. The perceptual affinity between breathy phonation and nasalization is thought to be high because both nasalization and breathiness result in weakening of the first oral formant (F1) with a wider bandwidth and a relative increase in amplitudes of lower frequency harmonics in comparison to high frequency harmonics (Garellek, 2014; Garellek and Keating, 2011; Gordon and Ladefoged, 2001; Stevens, 2000). The acoustic characteristics of both breathiness and nasalization are described below.

Nasalization

The acoustic features of nasalization are attributed to lowering of the velopharyngeal port and coupling of the nasal cavity and the oral cavity via the nasopharynx. Because the nasal cavity is smaller and radiates less acoustic energy than the oral cavity, its main role when coupled is to alter the primary oral acoustic output (House and Stevens, 1956). This coupling introduces additional nasal resonances that interact with the resonances of the vocal tract. The frequency of these poles and zeros (formants and anti-formants or resonances and anti-resonances) depends on several factors including the anatomy of the speaker (i.e., the size of the nasal cavity and sinuses), the degree of velum lowering, and tongue position. The general frequency ranges of the primary nasal resonances are approximately 200–400, 800–1100, 1200–1300, and 2000–2500 Hz (Johnson, 1997; Stevens, 2000; Chen, 1995; Styler, 2017).

The presence of nasal resonances cause the oral F1 to shift in frequency. While a shift in F1 of nasal vowels may in part be attributed to oropharyngeal adjustment (Carignan *et al.*, 2015; Shosted *et al.*, 2012b), the addition of nasal poles and zeros alters the acoustic power near oral formants and can change the formant's central frequency and width. A formant reduced in amplitude will have wider "shoulders" and hence a wider bandwidth (Styler, 2017). Formant frequency shifts depend on relative height of the vowel. In low vowels where the frequency range of F1 is typically above the first nasal pole (app. 400 Hz) the oral F1 will often shift lower in frequency. In high vowels where F1 is often near or below the nasal pole it may shift higher in frequency. This alteration is said to cause the vowel to "centralize". Furthermore, the coupling of the nasal and oral cavities increases the surface area of vocal tract, dampening all frequencies. (Johnson, 1997).

Each nasal pole is associated with a nasal zeros that are created when sound resonating within the nasal cavities reflects back into the oral cavity. When a zero is in phase with other frequencies in the signal they cancel each other out, resulting in local spectral valleys. The lowest frequency nasal zeros cited in acoustic studies of nasal vowels are approximately 600–900 Hz (associated with the lowest nasal pole of around 400 Hz), 1200–1400 Hz (associated with the pole of 800–1100 Hz) and 1500–2500 Hz (associated with the pole of 2000–2500 Hz) (Johnson, 1997; Stevens, 2000). Although a zero associated with the second nasal pole around 1200–1300 Hz undoubtedly exists, it is not discussed in the literature (Styler, 2017). The presence of the first nasal zero causes the oral F1 to weaken due to its proximity. See Table 2.3 for a list of nasal poles and zeros.

Lowering of the soft palate also induces relative weakening of high frequency harmonics

Nasal poles (formants)	Nasal zeros (anti-formants)
200-400 Hz	600–900 Hz
800-1100	1200-1400
1200-1300	not described
2000-2500	1500–2500

Table 2.3: Nasal poles and zeros in Hz

because the presence of nasal zeros introduced throughout the spectrum causes the amplitude of oral formants higher in frequency to decrease by about 1.6 dB per octave (Johnson, 1997; Fant, 1960; Pruthi and Espy-Wilson, 2005). The reduction of amplitude causes the spectral tilt of a nasal vowel to increase in absolute negative slope, where the high frequency components of the spectrum exhibit relatively lower amplitude as compared to oral vowels (Styler, 2017). Spectral tilt is a measure of the relative energy from low to high harmonics. If a spectrum exhibits relatively higher amplitude in lower frequency components and relatively lower amplitude in high frequency components, the spectral tilt is high. This is usually the case in nasalized and breathy (described below) vowels (Hillenbrand and Houde, 1996). Nasalized vowels exhibit higher spectral tilt than modal vowels. Creaky phonation exhibits relatively lower amplitude in low frequency components and higher amplitude in high frequency components as compared to modal vowels. The spectral tilt of creaky vowels is therefore low compared to normal and breathy vowels.

Breathiness

For breathy phonation, the vocal folds are less adducted than in modal (typical) phonation and an interarytenoid gap is often present. An interarytenoid gap results in a direct current of airflow during vocal fold vibration that allows for smoother vibratory transitions; there is often no cessation of airflow during the vibratory cycle as compared to modal voicing because the vocal folds do not touch (Hanson, 1995). Breathy vibration is described as being more sinusoidal in comparison to modal voice due to a lack of vocal fold contact (Holmberg *et al.*, 1988; Stevens, 2000; Simpson, 2012). When a glottal pulse is more symmetrical in the open-close phase of vocal fold contact, the spectral tilt of the resulting spectrum is steeper (Chasaide and Gobl, 1997; Wayland and Jongman, 2003). This causes the spectrum at lower frequencies to be relatively higher in amplitude resulting in increased spectral tilt (Bickley, 1982a; Garellek, 2014; Garellek and Keating, 2011).

However, this measure of spectral tilt can become complicated by the presence of aspiration noise. Hillenbrand and Houde (1996) explain that the turbulence of airflow present in breathiness can cause the periodic component of a breathy signal to relatively weaken at high-frequencies "[T]he presence of aspiration noise, which is stronger in the mid and high frequencies than in the lows, can result in a voice signal that is richer in high-frequency energy than nonbreathy signals" (p. 312). This may result in a spectral balance affected by two separate amplitude increases in both the high and low frequency ranges (Wayland and Jongman, 2003). Findings based on spectral tilt measures for breathiness have been mixed. Using measures of the amplitude of F1 relative to F2 and of F2 relative to F3, F4, and F5, Klatt and Klatt (1990) found no relation between rating of breathiness and spectral tilt. Stevens and Hanson (1994) and Hanson (1995) measured spectral tilt by the relative amplitude of the first harmonic (H1) and F3. Both studies found high spectral tilt (increased amplitude in H1) in breathy conditions as compared to modal. However Simpson (2012) finds that sex-specific differences in nasalization during low vowels makes the use of H1 and H2 amplitude as a measure of phonation unreliable. Some nasal coupling is always present on low vowels (see Section 2.2.2) and the lowest nasal pole affects the first few harmonics. Since males have a typical f0 of 120 Hz the nasal pole will enhance the second and third harmonic. A lower f0 means that harmonics will be lower in frequency and spaced out less than harmonics of relatively higher f0. In females, since f0 is usually higher (around 200 Hz), the nasal pole will typically enhance the first harmonic. Therefore, the use of H1 amplitude in measures of phonation are unreliable across vowels and genders. A broader spectral analysis is therefore advisable (Simpson, 2012).

An interarytenoid gap and a longer proportion of vocal fold opening during vibration (large open quotient) can also cause greater coupling of the trachea and vocal tract. This results in a series of poles and zeros affiliated with the trachea that interact with the existing oral resonances. The natural resonant frequencies of the trachea are about 600, 1550, and 2200 Hz in an adult male and 700, 1650, and 2350 Hz in an adult female, although there is variability in these values depending on individual speakers (Stevens, 2000; Ishizaka and Matsudaira, 1972; Cranen and Boves, 1987; Fant *et al.*, 1972).

The first subglottal (Sg) formant is typically less prominent than the higher subglottal formants because a partially adducted glottis creates acoustic losses in the lower frequency spectral region (Lulich *et al.*, 2011; Stevens, 2000). Because the first Sg resonance is in the frequency range of F1 (about 600 Hz) and the shoulder of the associated zero is rather wide (200–400 Hz) the prominence of oral F1 is reduced and its bandwidth increases. If the Sg resonance is higher in frequency than the vowel's oral F1, as would be the case for high vowels such as /i/ or /u/, we would expect the distortion from the Sg resonance to shift the frequency of F1 higher. If the Sg resonance is lower than the oral F1, as would be the case in low vowels like /a/ and /æ/, we would expect the Sg resonance to shift the frequency of F1 lower. The second Sg resonance, in the frequency range of 1400 to 1800 Hz, is more prominent and is known to create an acoustic boundary between the oral F2 of front and back vowels. When the F2 of a vowel is in proximity to the pole-zero pair of the second Sg resonance, the Sg zero will cause a distortion in F2, weakening it (Chi and Sonderegger, 2007; Stevens, 2000).

Breathy-nasal interaction, perception

The articulatory configurations for nasal and breathy sounds described above result in similar acoustic outputs where the first formant is weakened and has a larger bandwidth and there is increased spectral tilt as compared to oral, modally voiced vowels. See Table 2.4 for a list of articulatory and acoustic effects.

Attributes	Nasal cavity	Subglottal cavity
Coupling	velum lowering	breathy voicing
App. Resonances	400, 950, 1200, 2000 Hz	600, 1500, 2200 Hz
F1 strength	weakens F1	weakens F1
F1 position	low vowel: F1 ↓; high vowel: F1 ↑	low vowel: F1 ↓; high vowel: F1 ↑*
F1 bandwidth	wide	wide
Spectral tilt	high	high**

* Has not been empirically verified; modeling result only

** Spectral tilt may be complicated by aspiration noise. High spectral tilt during breathiness may result from subglottal resonances, sinusoidal vocal fold vibration, or both.

Table 2.4: Affects of Nasal / Subglottal coupling as compared to oral, modally-voiced vowels

Perceptual studies have shown that listeners can misperceive breathy phonation as nasal (Ohala and Amador, 1981; Lintz and Sherman, 1961). Ohala and Amador (1981) presented English listeners with iterations of the vowels just before before /s/, excised from their VC environments. Listeners rated these items as sounding more nasal than control stimuli. Ohala and Amador (1981) conclude that the transition from a voiced modal vowel into the voiceless fricative /s/ exhibits a brief breathy transition. Because breathiness and nasalization are acoustically similar, listeners misperceived the context as nasal. A perception study by Lintz (1961) tested listeners' judgments of nasality of excised vowels near voiced and voiceless plosives and fricatives; the fricative environments were judged as more nasal than plosive environments. In both perception experiments above, the coarticulation with a voiceless consonant with a wide glottis, in both cases voiceless fricatives, prompted the misperception of nasalization during an adjacent vowel. These findings can be explained by the acoustic consequences of subglottal coupling and breathiness outlined above. During the production of glottal fricatives, the glottis is spread and the subglottal cavity is coupled to the vocal tract. During the transition to or from a surrounding vowel we assume that the vocal folds are in a transient breathy configuration as they begin to vibrate (Stevens, 2000, pg. 91). This brief breathy transition can result in increased coupling with the trachea, introducing subglottal poles and zeros into the spectrum. This would result in weakening of F1, widening

F1's bandwidth, and possibly an increase spectral tilt. The frequency of F1 may even be affected by the presence of the lowest subglottal formant and may shift depending on the height of the vowel in question. Because these acoustic changes are similar to those resulting from nasalization, they may cause listeners to misperceive the sound as nasal.

Increased breathy phonation has been found to co-occur with, and possibly phonetically enhance, nasal sounds in languages like Yi and Southern French (SF) (Carignan, 2017; Garellek *et al.*, 2016). Phonetic enhancement is the use of secondary articulatory gestures in conjunction with primary gestures that are meant to enhance the perceptual saliency, i.e. perceptual distinctiveness from other phonemes, of the defining attributes of a phoneme (Stevens and Keyser, 2010). This may be accomplished by superimposing a gesture onto another to enhance the primary acoustic feature of a defining attribute. Garellek *et al.* (2016) studied three Yi languages and observed increased breathiness in nasal consonants and nasalized vowels as compared to non-nasal conditions. Similarly, using EGG Carignan (2017) found increased breathiness during nasal SF vowels as compared to oral vowels in most speakers. Breathiness and nasalization share important acoustic features, described in Section 2.2.3. In Yi and SF, increased breathiness during nasal phones may either be used as a strategy to phonetically enhance perceptions of nasality or may arise from speaker misperception. If the addition of breathiness were a phonetic enhancement strategy, it would likely work to enhance the percept of nasalization by introducing higher energy at low spectral components and further increase spectral tilt, two acoustic consequences of both nasalization and breathiness. See Sections 2.2.3 and 2.2.3 for further physiological and acoustic descriptions of nasalization and breathiness. Further study is needed to determine the acoustic features of simultaneously breathy and nasal speech and its effect on the percept of nasalization. While the cases of SF and Yi demonstrate the acoustic affinity between breathiness and nasalization, they are not examples of spontaneous nasalization. We may alternatively think of Yi and SF as cases of spontaneous breathiness in nasal contexts. Spontaneous nasalization, the focus of this dissertation, is defined as nasalization that emerges in phonetic contexts with no etymological nasal (Ohala and Ohala, 1992).

2.3 Central Thai

Central Thai, a tonal language from the Tai-Kadai family, is the primary language of Thailand spoken by about 50 million people (Gordon, 2005). Thai has five lexical tones, including three levels tones (high, mid, and low) and two dynamic tones (falling and rising). Thai phonology has 21 contrastive consonant phonemes, including two glottal consonants /h/ and /ʔ/ (Diller, 2004). Some scholars argue that /ʔ/ should not be included as a phoneme in Thai because it is predictable based on stressed position (Abramson, 1962; Harris, 2001). Thai syllables that begin with a vocalic onset (no onset consonant) are produced with a glottal stop in stressed position. Furthermore, Thai syllables with a low, short vowel are produced with a final glottal stop in stressed position. Glottal stop is often omitted during non-stressed syllables (Noss, 1964). For this reason, the current study only analyzes productions in a stressed and prominent prosodic position (see Section 3.0.2). Thai has nine vowel phonemes that each contrast with a long variant, thus providing a contrast of 18 Thai monophthong vowels Diller (2004).

Thai presents researchers with an excellent opportunity to study spontaneous nasalization because it potentially exhibits an interaction among rhinoglottophilia, rhinochthamalophilia, and rhinosyrigmatophilia. As introduced above, Thai is known to nasalize vowels after the voiceless glottal fricative /h/ and glottal stop /ʔ/, such that /hɛ/ [hɛ̃] 'parade' and /ʔaw/ [ʔãw] 'to take' are reported to sound nasal by Cooke (1989); Matisoff (1975); Noss (1964). Depending on the phonological analysis of Thai, it can be said that Thai nasalizes vowels after glottal consonants /h, ʔ/, or that Thai nasalizes vowels after /h/ and a vocalic onset. Regardless, if vowels reported to nasalize were all produced in stressed position, the phonetic environment would be produced with a glottal stop. It is unclear from previous research what stress position the reported nasalized vowels belonged to (Matisoff, 1975; Cooke, 1989; Noss, 1964). Thai nasalization is reported to affect only the low/mid-low vowels /ɛ/, /ɔ/ and /a/ and vowels following /h/ are more susceptible to nasalization than those following /ʔ/ (Matisoff, 1975; Cooke, 1989). Matisoff notes that the Thai /ɛ/ is a rather low vowel that is closer to /æ/, and thus can be phonetically transcribed as [ɛ̞] (Matisoff, 1975). Furthermore, of the low vowels Cooke (1989) observed that /a/ sounds more nasalized than /ɛ/ and /ɔ/. Haas reported that Thai dialects may differ in the contexts where spontaneous vowel nasalization may occur.[4] He reported that speakers of Northeastern Thai nasalize all vowels after /h/, not just low or mid-low vowels. However, in Central Thai, the language under investigation in this dissertation, vowel nasalization is only reported to affect low vowels.

To frame these observations within our spontaneous nasalization classification, Thai nasalizes low/mid-low vowels(rhinochthamalophilia), after glottal consonants (rhinoglottophilia), of which one of the glottal consonants is a voiceless fricative (rhinosyrigmatophilia). Therefore, it is possible that VPU, tongue height, and breathy-nasal acoustic similarity (in proximity to /h/) may all play a role in Thai nasalization. An articulatory investigation that assess velopharyngeal opening and phonation quality is necessary to tease apart these potentially interacting sources of nasalization in Thai.

2.4 Summary, Questions, and Predictions

While the proposed inducers of spontaneous nasalization have been explained in isolation, nearly all cases involve an interaction among two or more of the explanations described above. It is unclear how proposed inducers interact in a single language to facilitate nasalization. For example, Thai phonology arguably manifests an interaction among velopharyngeal underspecification, breathy-nasal acoustic similarity, and vowel-height-based spontaneous nasalization. Nasalization in proximity to Thai /h/ may be explained by VPU and breathy-nasal acoustic similarity because /h/ is both a glottal consonant and is produced with a voiceless transglottal airflow that can induce coarticulatory breathiness. However, nasaliza-

---

[4] Matisoff's personal communication, 1974

tion near /ʔ/ cannot be explained by breathy-nasal acoustic similarity because glottal stop is not similar to nasalization acoustically. The positive spectral slope of creakiness is opposed to the negative spectral slope of nasalization and breathiness. We would not predict that the laryngeal configurations of glottal stop would facilitate misperception with nasalization as breathy sounds would. Thus, nasalization in proximity to /ʔ/ can only explained by velopharyngeal underspecification and the presence of low vowels.

Recall that low vowels after /h/ are reported to sound more nasal that those after /ʔ/ (Cooke, 1989; Matisoff, 1975). It is unclear if vowels following /h/ are produced with greater VPO in Thai, or rather, if this context sounds more nasal due to breathy-nasal acoustic similarities. Are impressions of greater nasality after /h/ attributed to physiological breathiness, physiological nasalization, or some combination of the two? Furthermore, among the non-high vowels, Cooke (1989) observed that /a/ sounds more nasal than /ɛ/ and /ɔ/. Given the propensity for lower vowels to better tolerate the acoustic effects of nasal coupling than high vowels, a more nasal-sounding /a/ would almost certainly be produced with significantly greater physiological nasalization compared to the higher vowels /ɛ/ and /ɔ/.

2.4.1 Nasalization after glottal consonants

Matisoff (1975) and Cooke (1989) reported greater impressions of vowel nasalization after /h/ than after /ʔ/ in Thai. While there is a perceptual affinity between breathy and nasal sounds in terms of amplitude of the lower frequency harmonics and F1, glottal stop is not known to share acoustic similarities with nasalization. Are vowels after /h/ really produced with greater VPO than those after glottal stop or is it possible that the perceived greater nasalization after /h/ actually attributed to breathiness? If nasalization is actually greater, we would expect to observe larger VPO after /h/. However, we can conclude that impressions of increased nasalization after /h/ may be attributed to the presence of breathiness if VPO is comparable after /h/ and /ʔ/. Refer to the following section for exploratory considerations for breathiness after /h/. See Table 2.5 for a summary of VPO predictions.

	Predictions for increased VPO
Vowel	Low > mid-low > mid-high / high
Glottal	Glottal > non-glottal
/h/	/h/ > /ʔ/

Table 2.5: Predictions for increased VPO based on vowel height, glottal vs non-glottal consonants, and type of glottal consonant. Low/mid-low vowels after /h/ are reported to sound more nasal than those after /ʔ/. If VPO is responsible for this impression, we expect nasalized syllables with /h/ to be produced with greater VPO.

Exploratory considerations

We expect to observe at least some coarticulatory breathiness during the vowel after /h/. Breathiness that serves as a contrastive property of vowels has been found to be increased in magnitude and longer in duration compared to incidental coarticulatory breathiness (Blankenship, 2002). Breathiness has been observed during the nasalized vowels of Yi and Southern French as a possible enhancement feature for acoustic nasality (Carignan, 2017; Garellek *et al.*, 2016). If Thai also uses breathiness to enhance the percept of nasalization after /h/, we might expect there to be a greater degree of breathiness during nasalized Thai vowels compared to non-nasal vowel. For example, we may observe increased breathiness during low vowels after /h/, environments where the vowels are reported to sound nasal, compared to high vowels after /h/ that are not reported to sound nasal (Matisoff, 1975; Cooke, 1989). Increased breathiness could explain why low vowels after /h/ sound more nasal than those after /ʔ/. Even if low vowels after /h/ are produced with greater VPO than after /ʔ/, we still may observe increased breathiness during low vowels after /h/. That is, speakers may increase the degree or duration of breathiness during nasalized vowels after /h/ to facilitate perceptions of nasalization due to the acoustic similarities between nasalization and breathiness.

In sum, regardless of the state of physiological nasalization after /h/ and /ʔ/, we may observe differences in breathiness for the nasal-sounding low vowels after /h/ compared to high vowels after /h/ or vowels after /ʔ/. The degree and duration of observed breathiness after /h/ may provide clues about the status of this feature. If breathiness is similar during all vowels after /h/, both nasalized and oral, we can conclude that breathiness after /h/ is due to coarticulation with a the voiceless frication of /h/. However, if breathiness is greater or longer in duration during nasalized vowels after /h/ compared to non-nasal vowels after /h/, we can conclude that breathiness is likely used as an enhancement feature for the percept of nasalization during /hV/ syllables. Because our understanding of the role of breathiness in spontaneous nasalization is still developing, no concrete predictions are made at this point. All data regarding phonation quality will be analyzed in an exploratory nature.

A second consideration that is exploratory in nature concerns the relative timing of physiological nasalization in Thai. Based on previous literature concerning VPU of glottal consonants, we suspect that velum lowering may begin during the glottal consonants /h/ and /ʔ/ and continue throughout the duration of the vowel. That is, rather than nasalization that begins during the vowel, we may expect that VPO will increase during the glottal consonant of CV syllables and that the nasalization will continue into the following vowel. This finding would lend support to the velopharyngeal underspecification explanation for spontaneous nasalization. Because there is no phonetic specification for velopharyngeal configuration during glottal consonants, nasalization can begin before the vowel. However, it is likely that increased physiological nasalization is only heard during the vowel of a CV syllable because the glottal consonants /h/ and /ʔ/ are voiceless.

2.4.2 Glottal state and vowel height

Spontaneous nasalization has been found to not only occur in contexts with glottal consonants and wide-glottis segments, but during low vowels as well. Many researchers have observed a gradient behavior of VPO based on vowel height where VPO may be inversely correlated with tongue height. Does VPO also vary in a similar way based on the glottal state of immediately preceding consonants in syllables that otherwise are not reported to nasalize? For example, with all other factors held constant, would syllables beginning with /t^h/ be produced with greater VPO compared to syllables beginning with /t/? We do not predict such gradient changes in VP lowering based on glottal state. This is because misperception-based spontaneous nasalization likely requires an intermediate stage where breathiness is misperceived as nasalization and/or speakers take advantage of breathy-nasal acoustic similarity to enhance the acoustic distinction of certain syllables. In particular, Thai seems to manifest spontaneous nasalization where the presence of a wide-glottis consonant is not sufficient to induce nasalization; the consonant must also have a glottal place of articulation that is underspecified for VPO and thus allows for velum lowering. Most languages reported to have spontaneous nasalization manifest gradient vowel-height based variations of VPO near glottal consonants; only a few languages have been reported to develop spontaneous nasalization associated non-glottal voiceless wide-glottis consonants, e.g. Bzhedukh aspirated sibilants, Hindi sibilants, and Ponapean aspirated geminate obstruents. See Table 2.5 for a summary of predictions concerning VPO.

Chapter 3

STUDY 1: ULTRA-FAST MRI

Advances in ultra-fast MRI have allowed for imaging further back in the vocal tract in areas such as the pharynx (Hsieh *et al.*, 2013; Fu *et al.*, 2015a), the tongue (Smith *et al.*, 2013; Carignan *et al.*, 2015; Barlaz *et al.*, 2015), and the velum (Carignan *et al.*, 2015). These studies have contributed to the growing field of time-dynamic MRI analysis of time-varying speech physiology.

During an MRI scanning session subjects lay supine within the scanner. Supine posture is known to affect speech production, although conclusions about the extent and nature of this effect are mixed. In regards to the breathing apparatus, in supine position gravity facilitates greater expiration such that vital capacity is reduced by about 20%. Therefore, speech in a supine posture is shown to use a lower level in speakers' vital capacities (Hixon and Hoit, 2005; Hixon *et al.*, 2014). Stone *et al.* (2007) find that supine posture caused tongue dorsum retraction of at most 3 mm during speech in 7 out of 13 speakers, with two speakers exhibiting a more advanced tongue. They did not find significant differences in tongue position based on individual phonemes. Dietsch *et al.* (2013) report no difference in tongue position depending on upright or supine posture.

In regards to velar height, Whalen (1990) reports an effect of gravity on VP position in 4 adult subjects, while Kollara and Perry (2014) find no effect of gravity in children. Moon and Canady (1995) observe lower activation of the levator veli palatini muscle for subjects in supine position, suggesting that less activation is required for velopharyngeal closure when gravity pulls in the same direction (Hoit and Weismer, 2018). Finally, Vorperian *et al.* (2014) report an effect of supine posture on vowel formants, although the vowel space is still maintained, while Bae *et al.* (2014) find no significant effect of supine posture on vowel formants. While it would be ideal for speakers to be in an upright posture, this is not possible within the machine used in the current study. Ultra-fast MRI provides a high spatiotemporal resolution, noninvasive method for studying speech physiology. Possible differences in organ position based on posture will be kept in mind when interpreting the data.

3.0.1 Speakers

Four native speakers of Thai, two females and two males, were recruited to participate in a ultra-fast MRI study at the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. All four speakers grew up in Central Thailand. At the time of the experiment, each speaker's age and gender were: TF1 (female), 26 years old; TM1 (male), 23 years old; TM2 (male), 20 years old; TF2 (female), 23 years old. All subjects reported normal speech and hearing.

3.0.2 Materials

The materials were designed to examine the relationship of nasalization to syllable onset and vowel height. For the first set of tokens we constructed a list of monosyllabic real words and morphemes all containing a long low vowel /aː/; these differed only in onset consonant. Besides the two glottal consonants of interest, /h/ and /ʔ/, we included syllables beginning with consonants after which vowels are predictably nasalized or oral. Vowels after /n/ are reported to nasalize contextually in Thai (Beddor and Krakow, 1999), so we chose /naː˧/ as a syllable with predictable nasalization on the vowel. For predictably non-nasal conditions, three syllables chosen were /tʰaː˧/, /daː˧/, and /saː˧/. Furthermore, we compare productions of the vowels of syllables /tʰaː˧/ and /saː˧/ and /daː˧/ in order to determine whether the presence of coarticulatory breathiness from aspiration/voiceless frication (in /tʰaː˧/ and /saː˧/) can induce spontaneous vowel nasalization after a non-glottal consonant.

Spontaneous nasalization associated with aspiration (Ponapean languages) and /s/ (Hindi) has been reported in other languages (Blevins and Garrett, 1992; Ohala, 1974). The syllable /daː˧/ was included as a non-aspirated, non-glottal contrast. Nasalization is not reported after /tʰ, s, d/ in Thai, so far as we are aware.

Word	Gloss
/haː˧/	ฮา "to guffaw"
/hɛː˧/	แห่ "parade"
/heː˧/	เฮ "to swiftly flock"
/ʔaː˧/	อา "aunt"
/ʔɛː˧/	แอ "small / young"
/ʔeː˧/	เอ "pause (interjection)"
/naː˧/	นา "rice farm"
/tʰaː˧/	ทา "to apply / paint"
/daː˧/	ดา "to advance in a group"
/diː˧/	ดี "positive modifier"
/saː˧/	ซา "abate"

Table 3.1: Speech Materials

The second set of contrasts included syllables produced with glottal consonants, /h/ and /ʔ/, with vowels at three different heights: /a/, /ɛ/, and /e/. We chose this contrast following reports of varying degrees of nasalization based on vowel height in Thai (Cooke, 1989; Matisoff, 1975). Vowels /a/ and /ɛ/ are predicted to nasalize after /h/ and /ʔ/ while /e/ is not (Matisoff, 1975). We also selected the syllables /diː˧/ and /daː˧/ to compare VPO at differing vowel heights in predictably oral environments. We chose the syllable /diː/ because no native Thai word or morpheme /deː/ exists. The words were produced in the carrier phrase /pʰuːt˥˩ kʰam˧ waː˥˩ "__" ʔiːk˨ kʰraŋ˥ / "Say the word '__' again". See Table 3.1 for a complete list of syllables and their English glosses. All words and morphemes in Table 3.1 are produced with the mid tone except /hɛː˧/ *parade* for which no mid tone option existed.

Each item in Table 3.1 was presented in the carrier phrase to participants in the MR scanner and during a separate, noise-attenuated recording session in a sound booth. During both the ultra-fast MRI and separate audio recording sessions, three sentences were presented at a time and participants were instructed to produce all three sentences repeatedly for five minutes. Naturally, the number of repetitions per participant varied based on individual speaking rate. See Appendix A.1 for a list of the number of token repetitions per speakers.

3.1 Data acquisition and analysis

3.1.1 MRI acquisition

Participants were scanned while lying supine within a 3T Siemens Trio MRI scanner at the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. Participants' heads were gently restrained with a pillow and a 12-channel head coil. Participants spoke into a noise-attenuating MR-compatible optical microphone (Dual Channel-FOMRI; Optoacoustics, Or Yehuda, Israel). This acoustic signal was later used for segmentation of target words and syllables in Praat (Boersma and Weenink, 2018). However, due to noise introduced by the MRI system gradients, the acoustics recorded within the MRI were not analyzed. While the quality of the acoustics were sufficient for segmentation, they were too poor for spectral analysis. In accordance with the local IRB, informed consent was obtained from each speaker prior to scanning.

Initial scans included a localizer and a static T2–weighted Turbo Spin Echo (TSE) scan. The localizer is a short scan used to determine the orientation for the imaging sequences. The T2–weighted TSE scan produced a high-definition 3D image of the vocal tract to assist with manual slice placement for MRI scans. This scan was 0.8 × 0.8 × 0.8 with 320 × 240 × 192 voxels covering a field of view that is 256 × 192 × 153.6 mm per voxel. Dynamic MR images were reconstructed using the Partial Separability method (Fu *et al.*, 2015b, 2017; Liang, 2007). This method provides a nominal frame rate of 25 frames per second (fps) within scans composed of four simultaneously collected anatomical slices. The spatial resolution for each slice pixel is 2.2 × 2.2 *mm*, each image measuring 128 × 128 voxels. One slice was placed obliquely through the velopharynx at the place of attachment of the levator veli palatini muscle (Figure 3.1). The other three slices are not relevant to the current study and are therefore not reported. See Figure 3.3–3.6 for images of scan placement for each speaker. Three 5-minutes scans were performed to collect all speech samples; scans took approximately 15 minutes per speaker.

We cropped the area around the VPO imaged in the oblique slice (Figure 3.2) and converted the pixels to black and white (indicating absence and presence of tissue) using an automatically generated threshold from gray-level histograms (Otsu, 1979). The number of pixels were then added and multiplied by a factor 4.84 (2.2 × 2.2; the area of a pixel in mm^2) to obtain an approximation of velopharyngeal area in millimeters squared (Carignan *et al.*, 2015). This procedure was performed for every time sample of each token. Samples per token ranged from 4 to 12 samples, with a mean of 6. We observed no effect of token duration on VPO. See Appendix A.2 for an analysis of VPO based on token duration.

Although our measure of VPO is closely indicative of velum function, it is important to note that this measure is unlikely to capture complete velopharyngeal closure unless the slice placement exactly aligns with the point of contact between the velum and the wall of the nasopharynx during velum raising. This point of contact is difficult to obtain because slice placement is made using scans recorded when the participant is resting for five minutes with the velum lowered. Our most reliable and visible landmark for slice placement is the levator veli palatini muscle. Therefore, it is possible that when the velum is completely raised, as for the production of a plosive, our measure of VPO will still indicate some non-zero (positive) value (in mm^2). This is likely because the slice placement is just anterior to the place of contact for the velum and nasopharynx. We do not necessarily interpret these non-zero values to indicate coupling between the nasal and oro-pharyngeal cavities, but we do interpret the changing values in relative degree of velopharyngeal closure, with minimal values suggesting complete closure.

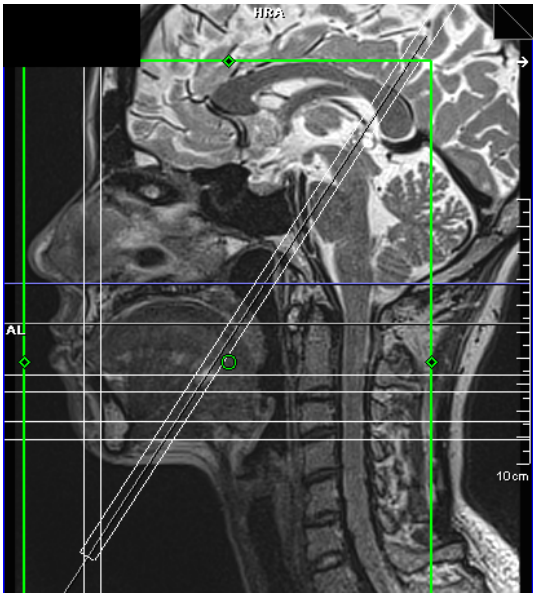

[redacted]

Figure 3.1: Slice orientation: oblique through VP port at levator veli palatini attachment (approximately 43 degrees)

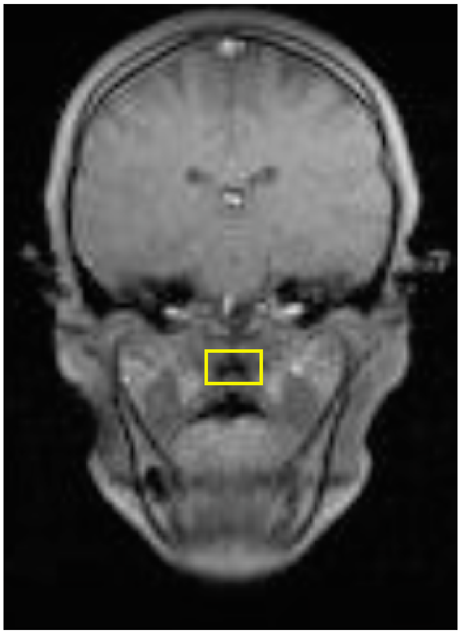


Figure 3.2: Ultra-fast scan example frame: VPO. The VPO region of interest is indicated with a yellow box

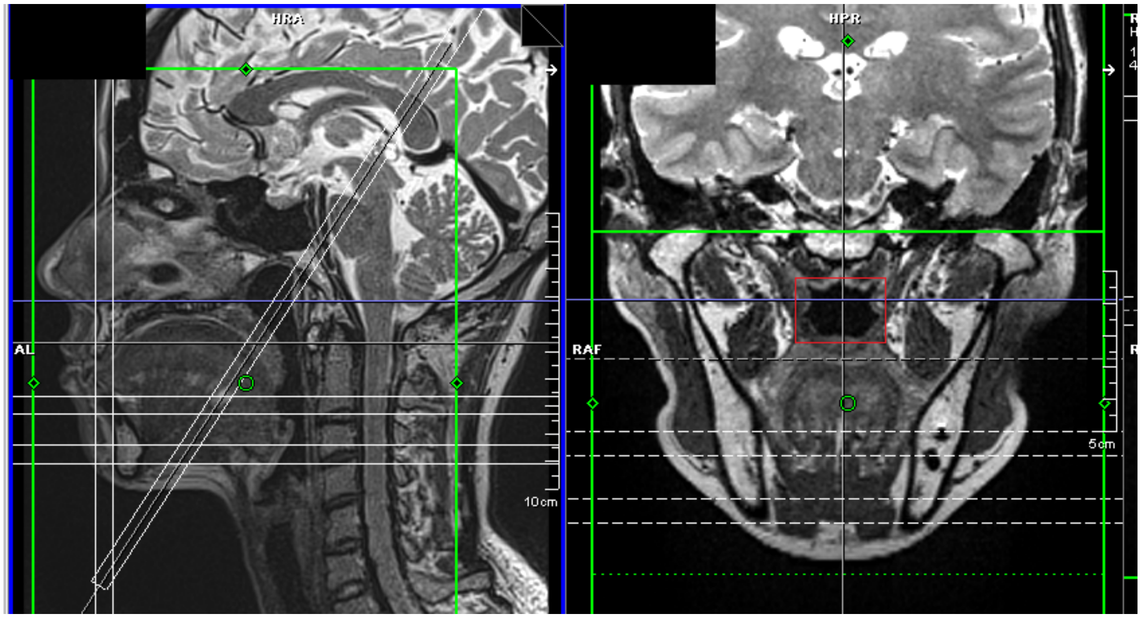

[redacted]
[redacted]

Figure 3.3: Scan orientation placement for speaker TF1. The image on the left shows the oblique scan placement through the levator veli palatini muscle. The image on the right shows the target slice, with velopharyngeal opening indicated within a red box.

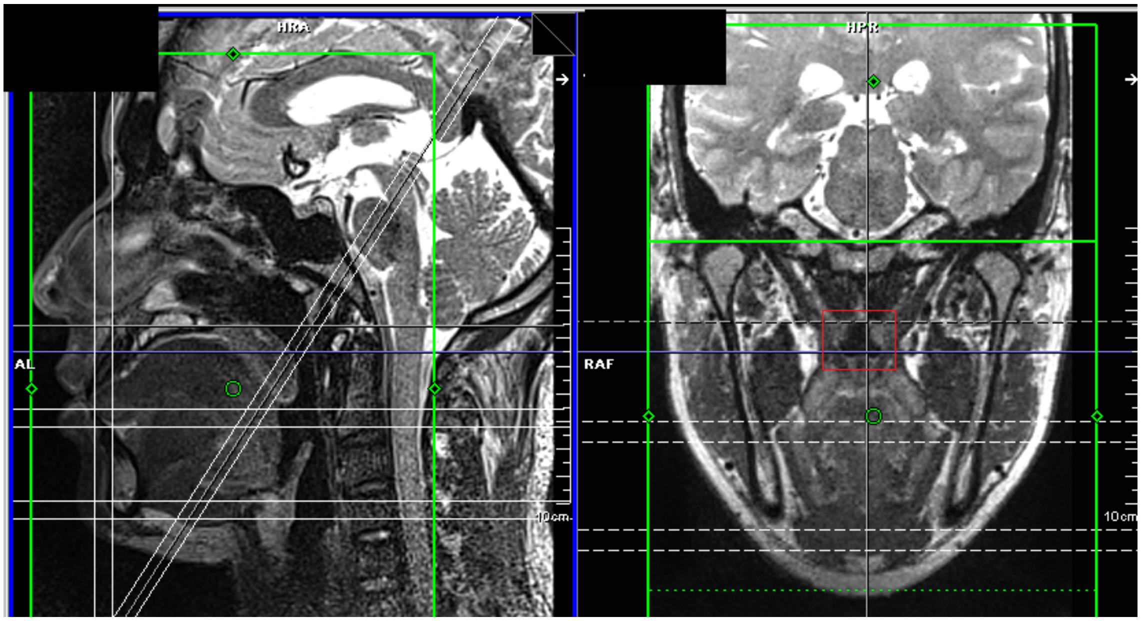

[redacted]
[redacted]

Figure 3.4: Scan orientation placement for speaker TM1. The image on the left shows the oblique scan placement through the levator veli palatini muscle. The image on the right shows the target slice, with velopharyngeal opening indicated within a red box.

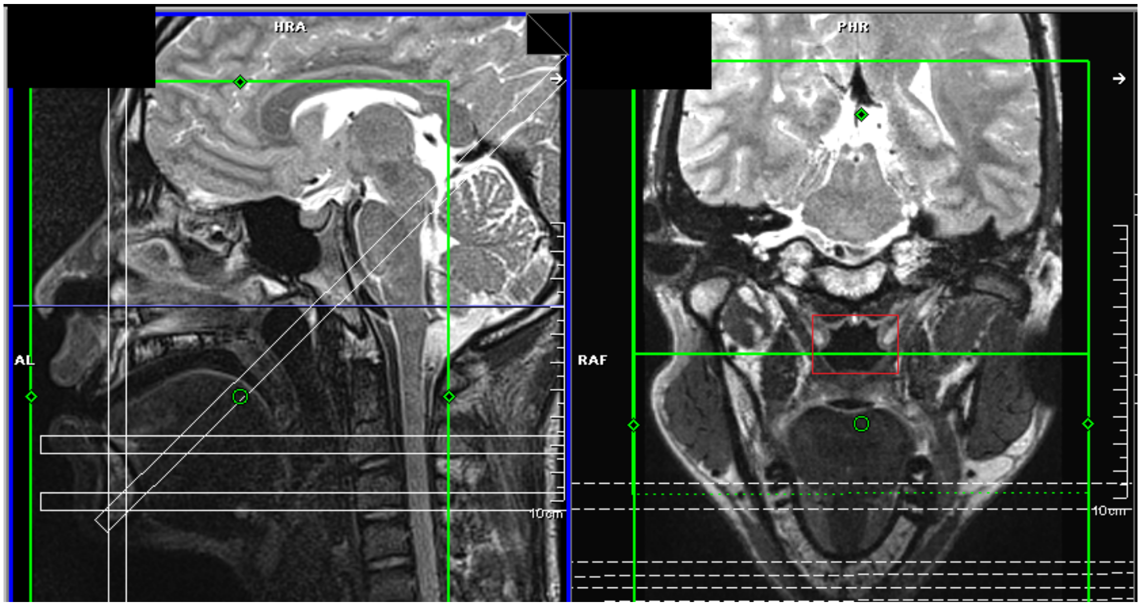

[redacted]
[redacted]

Figure 3.5: Scan orientation placement for speaker TM2. The image on the left shows the oblique scan placement through the levator veli palatini muscle. The image on the right shows the target slice, with velopharyngeal opening indicated within a red box.

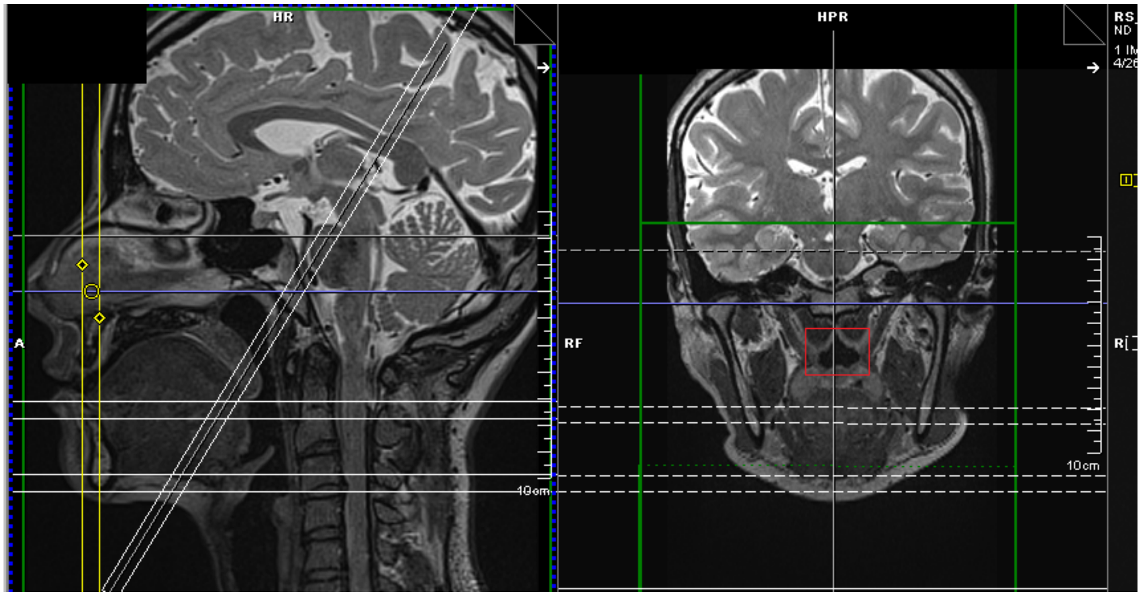

[redacted]

Figure 3.6: Scan orientation placement for speaker TF2. The image on the left shows the oblique scan placement through the levator veli palatini muscle. The image on the right shows the target slice, with velopharyngeal opening indicated within a red box.

3.1.2 Audio recordings

All speakers participated in a second, noise-attenuated audio recording within a week of the MRI session. Participants were recorded in a sound booth in the Illinois Phonetics and Phonology Laboratory. In order to approximate the experimental setting in the MR scanner, participants lay supine on a cot in the booth during the audio recording session and were instructed to remain still while speaking. Speakers were recorded at 44.1 kHz using a C520 headset microphone (AKG Harman, Stamford, CT). The acoustics were recorded using a PMD570 Solid State Recorder (Marantz Professional, Cumberland, RI) with a Grace m101 preamplifier (Grace Designs, Lyons, CO). Participants were instructed to repeat the same phrases as in the MRI for 3 minutes. See Appendix A.1 for a list of the number of tokens produced by each speaker.

3.1.3 Measures

Acoustic analyses

We assess phonation quality by measuring the difference in amplitude between the first and second harmonics, H1-H2. The greater the (positive) difference between H1 and H2, the greater the so-called spectral tilt (Bickley, 1982b; Kreiman *et al.*, 2007). H1-H2 is known to be correlated with open quotient, the temporal proportion of glottal opening during vocal fold vibration (Holmberg *et al.*, 1995). Breathy phonation is correlated with higher spectral tilt in part because vocal fold vibration is more sinusoidal than modal or creaky phonation. Furthermore, a lower spectral tilt is correlated with increased glottal constriction (Garellek, 2019). High and low H1-H2 has been shown to correlate with breathy and creaky phonation, respectively (Abramson *et al.*, 2015; Bickley, 1982b; Garellek and Keating, 2011; Gordon and Ladefoged, 2001). As breathiness is produced with a relatively more open glottis and/or more sinusoidal vibration, we expect breathy sounds to have relatively higher spectral tilt compared to modal or creaky phonation. We apply this method to detect breathiness during vowels with the caveat that variation in vowel height and gender can affect this measure. Simpson (2012) finds that some nasal coupling is often present during low vowels and the lowest nasal pole affects the first few harmonics. Since males typically have an f0 of 120 Hz, the nasal pole will enhance the second and third harmonic. In females, since f0 is higher (around 200 Hz), the nasal pole may enhance the first harmonic. H1-H2, henceforth H1*-H2*, was corrected based on formant frequency and bandwidth using algorithms written by Iseli and Alwan (2004) and Hawks and Miller (1995) in order to allow H1*-H2* comparisons across vowel quality and gender. We measured H1*-H2* at 20 time points per token. We applied the implementation of these algorithms available in VoiceSauce (Shue *et al.*, 2011).

3.1.4 Statistical methods

We analyzed time-series VPO and acoustic data using generalized additive mixed models (GAMMs) (Wood, 2006). This model allows for the inclusion auto-correlation parameters and has been previously applied to analyze time-series phonetic and linguistic data (Barlaz *et al.*, 2018; Desmeules-Trudel and Brunelle, 2018; Ko *et al.*, 2014; Nixon and Best, 2018; Sóskuthy, 2017; Wieling *et al.*, 2016; Wieling, 2018). It allows for the inclusion autocorrelation parameters for time-series phonetic and linguistic data. The models were created using the *mgcv* package in R (Wood and Wood, 2009). Due to the small number of speakers in the current study, a separate model was run for each speaker. The models included VPO or H1*-H2* values as the dependent variable. Each model included syllable as an independent variable and a smoothed function through normalized time (range 0–1). See Tables 3.2–3.5 for the full specifications of each model. An AR-1 correlation parameter was included in the model to account for any autocorrelation of the time-dynamic data, i.e. to prevent violation of the assumption of independence (Baayen *et al.*, 2018; Desmeules-Trudel and Brunelle, 2018; Wieling, 2018). The models were also constructed using restricted maximum likelihood (REML). All p-values were corrected using the Bonferroni method to reduce the likelihood of a Type I error from constructing separate models for each speaker. Finally, the model output, including time-series data plots, were visualized using the *itsadug* package in R (van Rij *et al.*, 2016).

3.2 Results

3.2.1 VPO

GAMMs of VPO through smoothed, normalized time revealed significant differences in VPO among syllables for every speaker (Tables 3.2–3.5). Time was a significant factor on VPO among syllables for speakers TM1 and TF2. The model outputs predict that VPO of /haː/ will be significantly different than all other syllables with the exception of /naː/ for speaker TF2 and /daː/ for speaker TF1. Figures 3.7–3.9 show mean model predictions with 95% confidence intervals for individual speakers. When the confidence intervals of two curves do not overlap they are significantly different from one another.

Figure 3.7 shows predicted model curves for every syllable with the vowel /aː/. The syllable /naː/ exhibits greater VPO than all other syllables in half the speakers (TM1 and TM2). In the remaining speakers /naː/ and /haː/ have similar VPO. For TF1, VPO of /haː/ is greater than /naː/ until about halfway through the syllable where it merges with /naː/. While /haː/ is produced with large VPO for TM2 during the initial half of the syllable, it overlaps with /daː/ during the latter half. Other than this exception, /t^{h}aː/, /saː/, and /daː/ manifest consistently lower VPO than /naː/ and /haː/, at the bottom of each speakers' VPO range.

The degree of VPO during /ʔaː/ is more variable. For TF1 and TF2, the VPO of /ʔaː/ is greater than predictably non-nasal syllables, /t^{h}aː/, /saː/ and /daː/, although /ʔaː/ and /daː/ briefly overlap at approximately 30% the vowel duration, before /ʔaː/ rises for TF1. TM1 and TM2 produce /ʔaː/ with low VPO, either overlapping with or lower than non-nasal syllables /t^{h}aː/ and /daː/. In sum, predictably nasalized syllables /naː/ and /haː/ are produced with high VPO in all speakers while predictably non-nasal syllables /t^{h}aː/, /saː/,

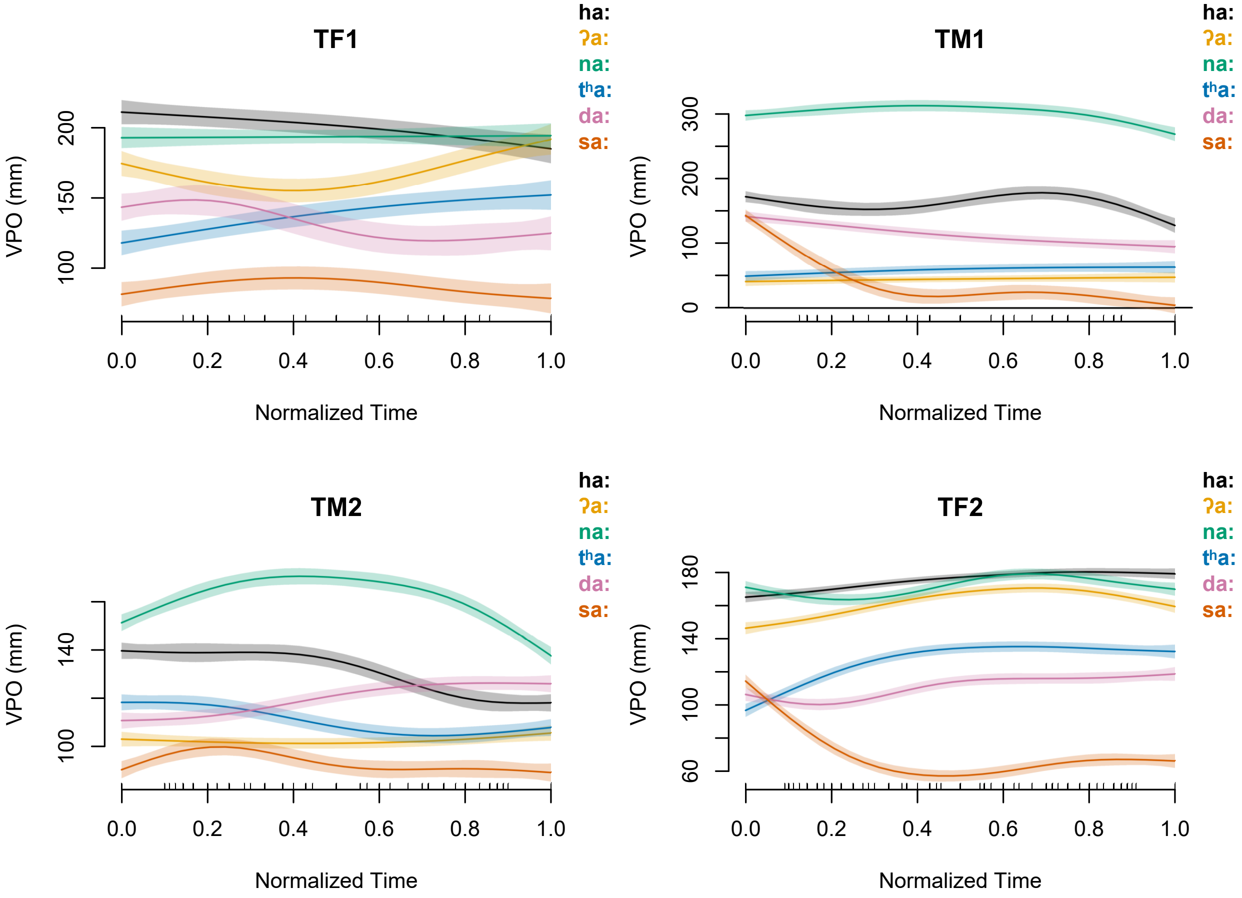


Figure 3.7: Predicted velopharyngeal opening in millimeters during the vowel of all syllables with the vowel /a/. Solid lines indicate the mean and shaded ribbons show the 95% confidence intervals. Note that y-axis ranges differ based on individual speaker ranges to improve legibility. Refer to y-axis values when comparing across-speaker differences.

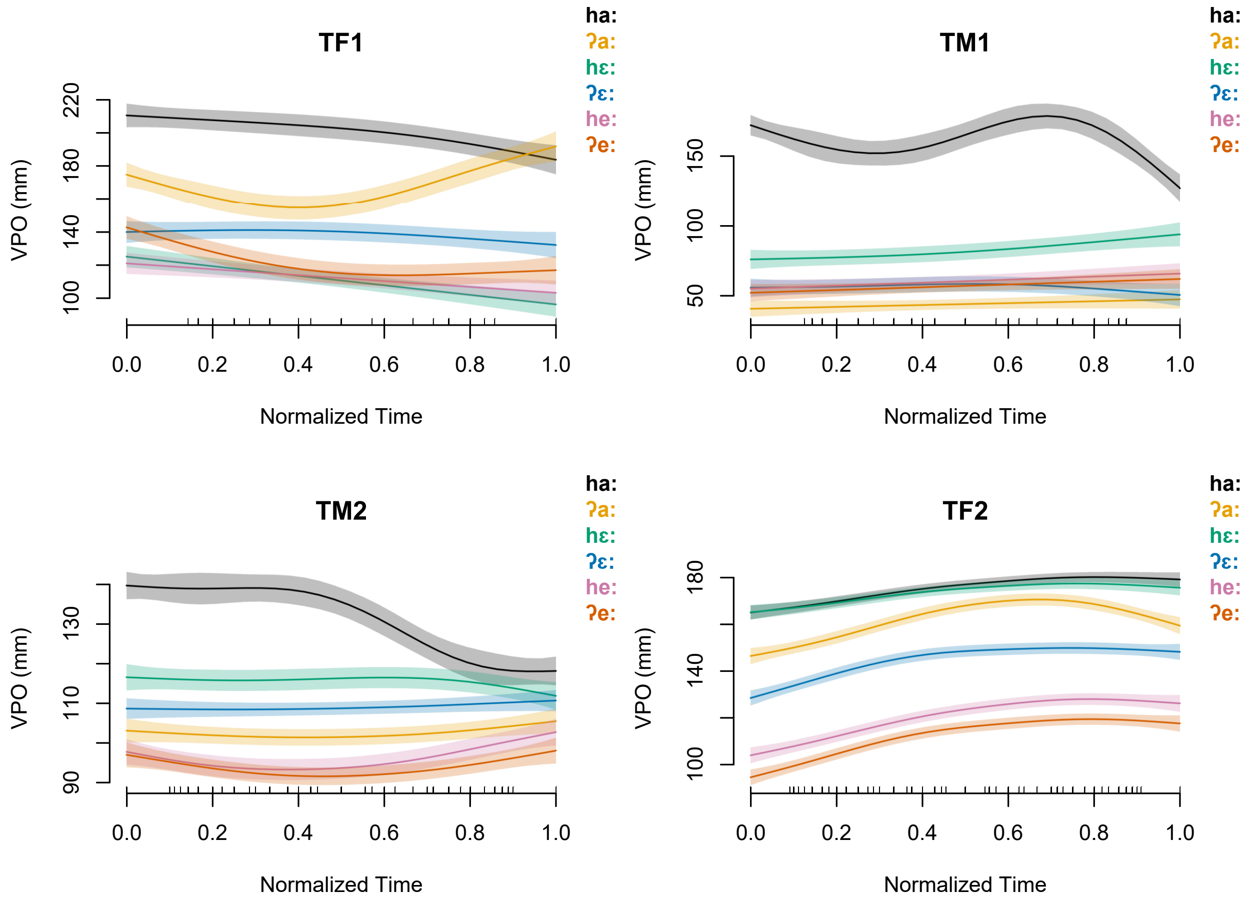


Figure 3.8: Predicted velopharyngeal opening in millimeters during the vowel of syllables beginning with /h/ and /ʔ/. Solid lines indicate the mean and shaded ribbons show the 95% confidence intervals. Note that y-axis ranges differ based on individual speaker ranges to improve legibility. Refer to y-axis values when comparing across-speaker differences.

and /daː/ are produced with low VPO. The syllable /ʔaː/, which is predicted to nasalize, is only produced with relatively high VPO, i.e. higher than /t^haː/, /saː/, and /daː/, for half the speakers.

Figure 3.8 shows predicted model curves for all syllables beginning with /h/ and /ʔ/. Predictably non-nasal syllables /heː/ and /ʔeː/ show the lowest VPO in all speakers, manifesting similar VPO values as non-nasal /t^haː/ and /daː/ within each speakers' respective ranges. /haː/ exhibits the highest VPO for all speakers. For TF2, /haː/ and /hɛː/ have similar VPO, and both overlap with /ʔaː/ at the temporal midpoint of the vowel. /hɛː/ manifests high VPO for TF2 only. In other speakers, the VPO of /hɛː/ is either moderate,

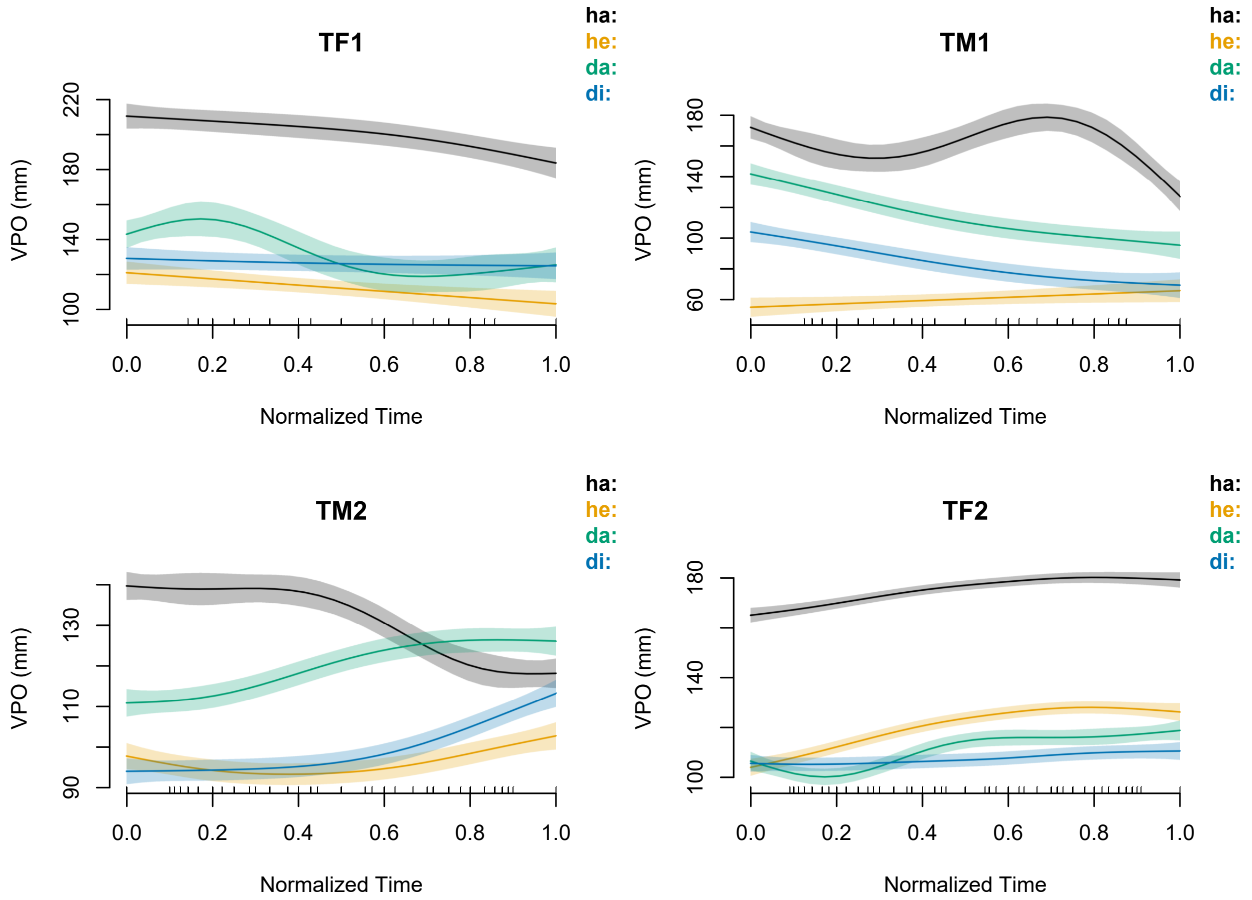


Figure 3.9: Predicted velopharyngeal opening in millimeters during the vowel of syllables with low and mid-high/high vowels after /h/ and /d/. Solid lines indicate the mean and shaded ribbons show the 95% confidence intervals. Note that y-axis ranges differ based on individual speaker ranges to improve legibility. Refer to y-axis values when comparing across-speaker differences.

i.e. lower than /haː/ but just greater than predictably non-nasal /heː/ and /ʔeː/ (TM1 and TM2), or low (TF1). The syllable /ʔɛː/ is moderate in three out of four speakers and low in one (TM1).

Finally, Figure 3.9 compares predicted VPO of vowels /iː/ and /aː/ after /d/. The syllables /heː/ and /heː/ are shown on these plots, as well, for comparison. Of these four syllables, only /haː/ is predicted to nasalize. We observe that /haː/ is produced with the greatest VPO, except for TM2 where the curve for VPO merges with /daː/ about halfway through the syllable. For every speaker, /haː/ is produced with much higher VPO than

/heː/. Most speakers also produce /daː/ with greater VPO than /diː/ for some duration of the vowel, except for TF2, who produces /diː/ with greater VPO for most of the syllable. TF1 also produces /diː/ with similar VPO as /daː/ about halfway through the syllable. In sum, while nasalized /haː/ is consistently produced with greater VPO than /heː/ (a syllable with a mid-high vowel), we observe greater variability in a high/low vowel comparison after /d/ than after /h/. Some speakers produce /daː/ with greater VPO than /diː/ while others do not.

	VPO (AR-1 = 0.25, r^2 = 0.61)				H1*-H2* (AR-1 = 0.36, r^2 = 0.27)			
Parametric coeff.	Est.	Std.Error	*t* value	*p* value	Est.	Std. Error	*t* value	*p* value
(Intercept)	199.911	2.203	90.741	< 0.001***	14.630	0.388	37.719	< 0.001***
Syllable /hɛː/	-61.911	2.939	-21.065	< 0.001***	-0.856	0.548	-1.560	0.475
Syllable /heː/	-29.108	2.915	-9.985	< 0.001***	2.352	0.548	4.289	< 0.001***
Syllable /ʔaː/	-77.240	2.909	-26.548	< 0.001***	-1.781	0.548	-3.251	0.004**
Syllable /ʔɛː/	-67.391	3.146	-21.422	< 0.001***	-4.519	0.632	-7.149	< 0.001***
Syllable /ʔeː/	-73.307	3.048	-24.050	< 0.001***	-1.373	0.548	-2.504	0.049*
Syllable /naː/	-89.234	3.043	-29.325	< 0.001***	1.838	0.541	3.398	0.002**
Syllable /tʰaː/	-87.769	3.048	-28.796	< 0.001***	-5.334	0.544	-9.807	< 0.001***
Syllable /daː/	-6.152	3.002	-2.050	0.162	-3.143	0.557	-5.645	< 0.001***
Syllable /diː/	-62.010	3.119	-19.879	< 0.001***	-15.130	0.636	-23.779	< 0.001***
Smooth terms	EDF	Ref.df	*F* value	*p* value	EDF	Ref.df	*F* value	*p* value
s(Time)	1.000	1.000	2.237	0.539	1.001	1.001	6.163	0.052
s(Time):Syll /haː/	1.865	2.270	2.702	0.1976	4.783	4.976	18.867	< 0.001***
s(Time):Syll /hɛː/	1.686	2.073	2.025	0.503	4.370	4.815	8.913	< 0.001***
s(Time):Syll /heː/	3.124	3.735	9.430	< 0.001***	4.369	4.814	44.525	< 0.001***
s(Time):Syll /ʔaː/	2.541	3.084	5.364	0.004**	2.974	3.595	4.464	0.006**
s(Time):Syll /ʔɛː/	2.465	2.980	2.956	0.141	4.333	4.794	19.257	< 0.001***
s(Time):Syll /ʔeː/	1.180	1.335	2.333	0.617	3.621	4.251	18.944	< 0.001***
s(Time):Syll /naː/	1.000	1.000	5.249	0.088	4.799	4.979	28.238	< 0.001***
s(Time):Syll /tʰaː/	1.000	1.000	3.913	0.192	3.032	3.658	7.247	< 0.001***
s(Time):Syll /daː/	1.000	1.000	2.066	0.603	3.802	4.411	3.655	0.041*
s(Time):Syll /diː/	1.888	2.302	1.295	1.000	2.032	2.659	2.325	0.194

Table 3.2: TF1 GAMM models. The VPO model was constructed using the formula: VPO ∼ s(Time) + Syllable + s(Time, by = Syllable). The H1*-H2* model was constructed using the formula: H1*-H2* ∼ s(Time) + Syllable + s(Time, by = Syllable)

3.2.2 H1*-H2*

GAMMs of H1*-H2* through smoothed, normalized time predict significant differences among syllables for all speakers (Tables 3.2–3.5). Overall, time is predicted to be a significant factor for H1*-H2* for each speaker. Figures 3.10–3.11 show mean model predictions with 95% confidence intervals for individual speakers.

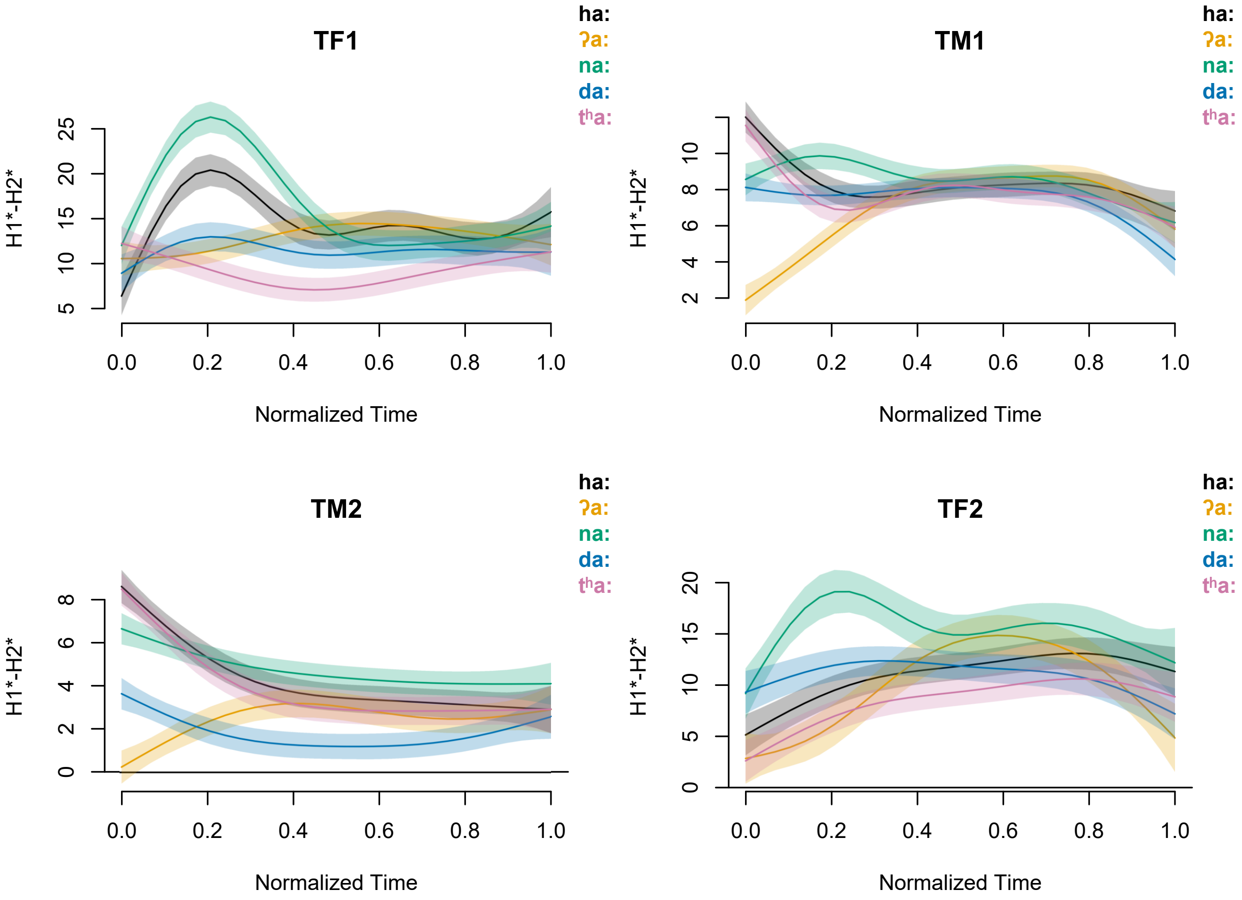


Figure 3.10: Predicted H1*-H2* during the vowel of syllables with the vowel /aː/. Solid lines indicate the mean and shaded ribbons show the 95% confidence intervals. Note that y-axis ranges differ based on individual speaker ranges to improve legibility. Refer to y-axis values when comparing across-speaker differences.

	VPO (AR-1 = 0.30, r^2 = 0.82)				H1*-H2* (AR-1 = 0.37, r^2 = 0.32)			
Parametric coeff.	Est.	Std. Error	t value	p value	Est.	Std. Error	t value	p value
(Intercept)	159.387	2.076	76.775	< 0.001***	8.357	0.164	51.125	< 0.001***
Syllable /hɛː/	-103.637	2.752	-37.654	< 0.001***	-3.011	0.232	-12.963	< 0.001***
Syllable /heː/	-115.256	2.770	-41.609	< 0.001***	0.954	0.232	4.112	< 0.001***
Syllable /ʔaː/	-102.284	2.841	-36.005	< 0.001***	-1.484	0.230	-6.447	< 0.001**
Syllable /ʔɛː/	-44.832	2.938	-15.257	< 0.001***	-2.536	0.229	-11.064	< 0.001***
Syllable /ʔeː/	-75.502	2.857	-26.430	< 0.001***	0.468	0.230	2.031	0.170
Syllable /naː/	-76.308	2.970	-25.694	< 0.001***	0.102	0.231	0.439	1.000
Syllable /tʰaː/	-98.908	2.935	-33.702	< 0.001***	-0.529	0.232	-2.279	0.090
Syllable /daː/	139.534	2.838	49.175	< 0.001***	-0.927	0.232	-4.001	< 0.001***
Syllable /diː/	-101.376	2.935	-34.539	< 0.001***	-4.470	0.231	-19.326	< 0.001***
Smooth terms	EDF	Ref.df	F value	p value	EDF	Ref.df	F value	p value
s(Time)	1.000	1.001	8.834	0.011*	4.384	4.689	19.855	< 0.001
s(Time):Syll /haː/	4.126	4.629	11.341	< 0.001***	3.851	4.420	12.620	< 0.001***
s(Time):Syll /hɛː/	1.991	2.445	3.807	0.089	3.659	3.923	28.708	< 0.001***
s(Time):Syll /heː/	1.000	1.000	10.447	0.005**	2.855	3.442	8.287	< 0.001***
s(Time):Syll /ʔaː/	1.000	1.000	11.588	0.003**	3.527	4.131	16.697	< 0.001***
s(Time):Syll /ʔɛː/	1.084	1.543	1.404	1.000	3.237	3.846	14.8034	< 0.001***
s(Time):Syll /ʔeː/	1.943	2.382	1.205	1.000	3.972	4.516	23.720	< 0.001***
s(Time):Syll /naː/	1.574	1.929	7.372	0.002**	4.270	4.727	8.560	< 0.001***
s(Time):Syll /tʰaː/	1.000	1.000	11.911	0.002**	4.332	4.765	11.387	< 0.001***
s(Time):Syll /daː/	3.486	4.094	11.034	< 0.001***	1.001	1.001	20.578	< 0.001***
s(Time):Syll /diː/	1.773	2.177	7.164	0.002**	4.786	4.963	36.298	< 0.001***

Table 3.3: TM1 GAMM models. The VPO model was constructed using the formula: VPO ∼ s(Time) + Syllable + s(Time, by = Syllable). The H1*-H2* model was constructed using the formula: H1*-H2* ∼ s(Time) + Syllable + s(Time, by = Syllable)

Figure 3.10 shows predicted curves of H1*-H2* for all syllables with the vowel /aː/. We observe that during the initial half of the vowel, the H1*-H2* of /haː/ is relatively high for all speakers. The syllables /naː/, /haː/, and /tʰaː/ exhibit higher H1*-H2* than /ʔaː/ and /daː/ during the initial half of the vowel in TM1 and TM2. This increased H1*-H2* may be a result of breathiness or an influence from nasalization on low frequency harmonics. Spectral tilt becomes similar in all syllables by the temporal midpoint of the vowel. The order of greatest H1*-H2* for these three syllables varies across speakers. The syllable /ʔaː/ is produced with the lowest H1*-H2* for two speakers at the start of the vowel. Low H1*-H2* may be associated with the presence of creakiness.

Figure 3.11 shows predicted H1*-H2* during vowels after /h/ and /ʔ/. We compare pairs of glottal consonants for three vowel heights, /ɛː/, reported to nasalize, and /eː/, not reported to nasalize. For each pair, H1*-H2* is significantly higher during most of the initial half of the vowel after /h/ compared to after /ʔ/ in all speakers except TF2. This suggests that the start of the vowel may be breathier after /h/. This is especially likely for /heː/ because

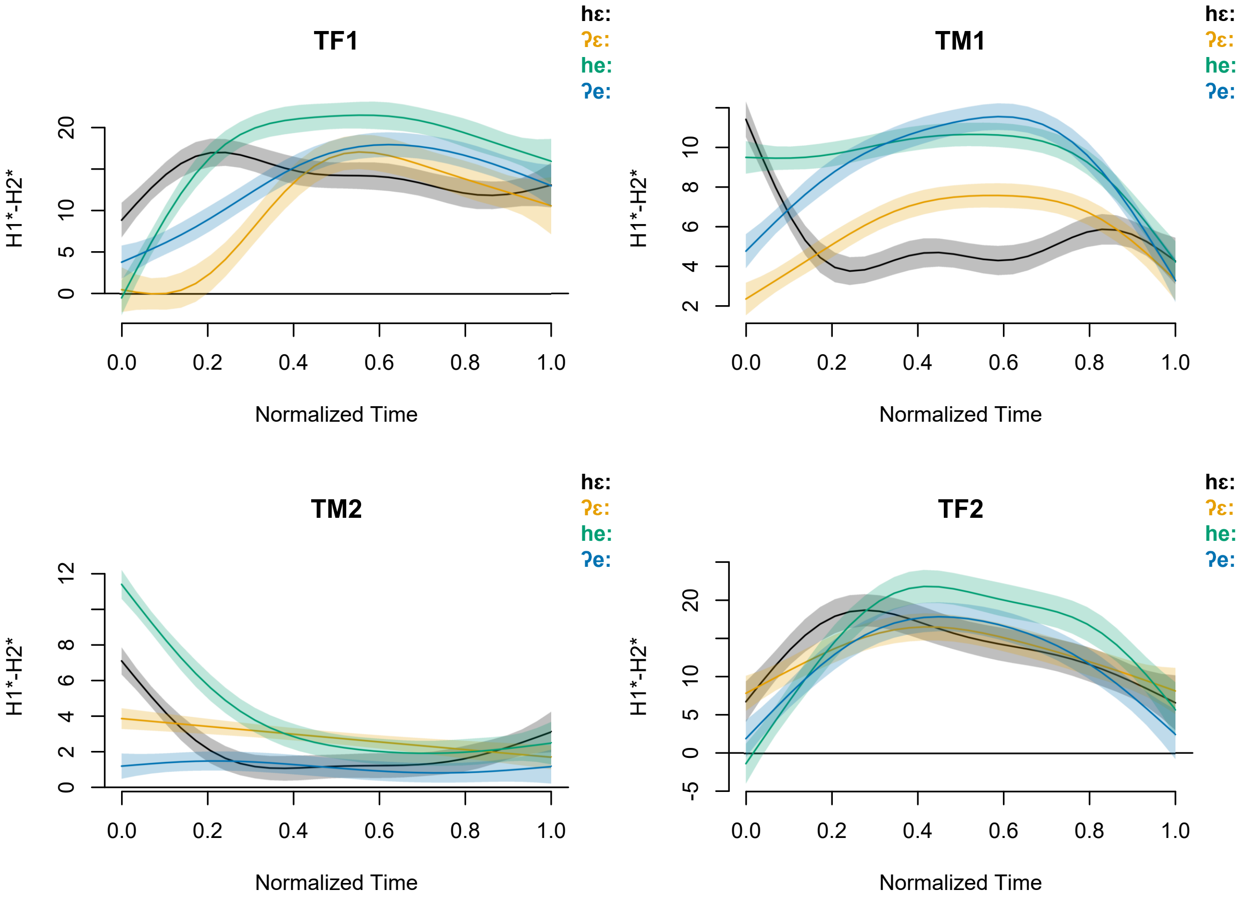


Figure 3.11: Predicted H1*-H2* during the vowel of syllables beginning with /h/ and /ʔ/ with the vowels /eː/ and /ɛː/. Solid lines indicate the mean and shaded ribbons show the 95% confidence intervals. Note that y-axis ranges differ based on individual speaker ranges to improve legibility. Refer to y-axis values when comparing across-speaker differences.

	VPO (AR-1 = 0.10, r^2 = 0.56)				H1*-H2* (AR-1 = 0.52, r^2 = 0.30)			
Parametric coeff.	Est.	Std. Error	*t* value	*p* value	Est.	Std. Error	*t* value	*p* value
(Intercept)	130.973	0.866	151.248	< 0.001***	4.277	0.175	24.446	< 0.001***
Syllable /hɛː/	-21.744	1.112	-19.559	< 0.001***	-1.988	0.242	-8.215	< 0.001***
Syllable /heː/	-28.103	1.135	-24.755	< 0.001***	-0.301	0.251	-1.199	0.923
Syllable /ʔaː/	-36.420	1.108	-32.880	< 0.001***	-1.826	0.247	-7.380	< 0.001***
Syllable /ʔɛː/	-11.323	1.181	-9.591	< 0.001***	-1.482	0.247	-6.009	< 0.001***
Syllable /ʔeː/	-30.921	1.154	-26.800	< 0.001***	-3.109	0.245	-12.692	< 0.001***
Syllable /naː/	-15.595	1.186	-13.147	< 0.001***	0.465	0.248	1.875	0.243
Syllable /tʰaː/	-34.057	1.173	-29.030	< 0.001***	-0.356	0.246	-1.451	0.587
Syllable /daː/	27.548	1.174	23.467	< 0.001***	-2.437	0.243	-10.013	< 0.001***
Syllable /diː/	-19.947	1.200	-16.617	< 0.001***	0.530	0.248	2.140	0.130
Smooth terms	EDF	Ref.df	*F* value	*p* value	EDF	Ref.df	*F* value	*p* value
s(Time)	1.752	2.101	1.619	0.727	1.044	1.078	15.760	< 0.001
s(Time):Syll /haː/	3.471	3.951	4.000	0.009**	3.375	4.010	13.114	< 0.001***
s(Time):Syll /hɛː/	1.000	1.000	3.025	0.328	4.164	4.687	23.460	< 0.001***
s(Time):Syll /heː/	1.754	2.147	2.186	0.468	3.865	4.461	42.783	< 0.001***
s(Time):Syll /ʔaː/	2.406	2.940	3.354	0.074	3.299	3.934	13.073	< 0.001***
s(Time):Syll /ʔɛː/	2.952	3.534	5.839	0.002**	0.000	0.001	0.185	1.000
s(Time):Syll /ʔeː/	2.607	3.150	7.902	< 0.001***	2.437	2.981	2.540	0.220
s(Time):Syll /naː/	1.409	1.920	2.145	0.175	2.417	2.958	1.924	0.420
s(Time):Syll /tʰaː/	2.495	3.010	4.290	0.005**	3.506	4.139	15.915	< 0.001***
s(Time):Syll /daː/	4.034	4.568	36.206	< 0.001***	3.015	3.635	5.858	0.001**
s(Time):Syll /diː/	3.125	3.658	2.884	0.176	1.000	1.000	0.094	1.000

Table 3.4: TM2 GAMM models. The VPO model was constructed using the formula: VPO ∼ s(Time) + Syllable + s(Time, by = Syllable). The H1*-H2* model was constructed using the formula: H1*-H2* ∼ s(Time) + Syllable + s(Time, by = Syllable)

speakers do not produce this syllable with a high degree of VPO. Therefore, increased H1*-H2* is likely attributed to breathiness in this case. H1*-H2* is similar across all syllables. In every speaker, the vowels after /h/ and /ʔ/ usually converge further into the duration of the vowel, i.e. the H1*-H2* of /hɛː/ converges with /ʔɛː/, and /heː/ converges with /ʔeː/. The exception is TM1, where /heː/ and /ʔeː/ merge and then /ʔeː/ exhibits higher H1*-H2*.

3.3 Discussion

We have presented comprehensive, time-varying imaging data of the velopharyngeal port during the production of Thai vowels. Our goal was to observe the relationships between vowel height, onset consonant, and VPO. A variety of hypotheses regarding these relationships have been forwarded over the years. We systematically tested these hypotheses directly using non-invasive, non-ionizing imaging technology. While there is clearly further work to be done—particularly in the detection of breathiness—to fully understand these relationships,

	VPO (AR-1 = 0.10, r^2 = 0.71)				H1*-H2* (AR-1 = 0.45, r^2 = 0.25)			
Parametric coeff.	Est.	Std. Error	t-value	p-value	Est.	Std. Error	*t* value	*p* value
(Intercept)	174.812	0.739	236.709	< 0.001	10.850	0.514	21.129	< 0.001
Syllable /hɛː/	-30.829	1.008	-30.572	< 0.001	2.541	0.746	3.406	0.003**
Syllable /heː/	-13.756	1.008	-13.653	< 0.001	3.818	0.737	5.179	< 0.001
Syllable /ʔaː/	-63.179	1.020	-61.924	< 0.001	-1.137	0.718	-1.583	0.454
Syllable /ʔɛː/	-63.541	1.044	-60.888	< 0.001	2.027	0.708	2.862	0.017
Syllable /ʔeː/:	-67.228	1.083	-62.076	< 0.001	1.080	0.718	1.504	0.530
Syllable /naː/	-1.716	1.110	-1.545	0.489	4.468	0.712	6.280	< 0.001
Syllable /t^{h}aː/	-55.001	1.069	-51.449	< 0.001	-2.492	0.733	-3.401	0.002**
Syllable /daː/	-3.513	1.049	-3.350	0.003**	0.040	0.745	0.054	1.000
Syllable /diː/	-49.173	1.052	-46.767	< 0.001	-12.633	0.715	-17.659	< 0.001***
Smooth terms	EDF	Ref.df	F-value	p-value	EDF	Ref.df	*F* value	*p* value
s(Time)	3.757	4.594	6.069	< 0.001***	3.755	4.282	5.930	< 0.001***
s(Time):Syll /haː/	1.000	1.001	0.619	1	1.000	1.000	12.774	0.001**
s(Time):Syll /hɛː/	2.807	3.396	2.614	0.163	3.797	4.390	5.531	0.001**
s(Time):Syll /heː/	3.284	3.907	6.511	< 0.001***	3.965	4.527	25.560	< 0.001***
s(Time):Syll /ʔaː/	2.522	3.076	2.109	0.373	3.724	4.325	8.905	< 0.001***
s(Time):Syll /ʔɛː/	4.178	4.682	4.429	0.007**	3.007	3.617	3.093	0.058
s(Time):Syll /ʔeː/	2.434	2.967	4.180	0.023*	3.440	4.061	15.015	< 0.001***
s(Time):Syll /naː/	1.000	1.000	1.660	0.791	4.194	4.690	3.794	0.050*
s(Time):Syll /t^{h}aː/	1.093	1.574	1.672	1	1.000	1.000	12.693	0.001**
s(Time):Syll /daː/	4.146	4.660	6.665	< 0.001***	0.000	0.000	0.001	1.000
s(Time):Syll /diː/	3.731	4.333	12.720	< 0.001***	3.236	3.858	5.885	0.001**

Table 3.5: TF2 GAMM models. The VPO model was constructed using the formula: VPO ∼ s(Time) + Syllable + s(Time, by = Syllable). The H1*-H2* model was constructed using the formula: H1*-H2* ∼ s(Time) + Syllable + s(Time, by = Syllable)

we offer the following summary of our work so far.

We observed that all speakers consistently produce the vowel of /haː/ with high VPO, while only one speaker produces /hɛː/ with high VPO. Therefore, while Matisoff (1975) reports that the word /hɛː/ ‘parade’ is nasalized in Thai, we conclude that currently only some speakers nasalize it. We observe similar variability in VPO after /ʔ/. Only two speakers, TF1 and TF2, produce low and mid-low vowels after /ʔ/ with greater VPO than predictably non-nasal vowels. For these two speakers, /ʔaː/ is also produced with greater VPO than /ʔɛː/. In sum, we observe that the lowest vowel /a/ is indeed produced with greater VPO than mid-low vowels predicted to nasalize after glottal consonants. However, this trend is not as gradient as Matisoff (1975) and Cooke (1989) originally suggested. While the low vowel after /h/ is consistently nasalized, we observe greater individual variation for the mid-low vowel; some speakers nasalize it and others do not.

A higher H1*-H2*, i.e. higher spectral tilt, value for the vowel of /haː/ than the vowel of /ʔaː/ suggests that vowels after /h/ may be produce with breathier phonation just after consonant release. The fact that this value falls slightly throughout the duration of the syllable /haː/ (while VPO does not) in some speakers, suggests that the vowel after /h/ may be breathy just after consonant release but that this breathiness generally decreases over time. Because nasalization can also increase spectral tilt, we cannot infer from acoustics alone whether elevated H1*-H2* during the remainder of the vowel is attributed to breathiness or high VPO.

This complicates our understanding of spontaneous nasalization in Thai in that there may be an interaction between breathiness and VPO in the percept of nasalization. Low vowels after /h/ are produced with greater VPO than low vowels after /ʔ/, although we observe greater variability in VPO during mid-low vowels. This difference in VPO may be responsible for reports of greater nasalization after /h/ (Matisoff, 1975; Cooke, 1989). However, the presence of breathiness during vowels after /h/ may affect the perception of nasalization. Further articulatory testing, such as EGG or laryngoscopy, is required to verify the phonation quality of vowels after /h/ and /ʔ/. Furthermore, perceptual testing is necessary to assess to what degree Thai speakers perceive vowels after /h/ and /ʔ/ as nasalized and whether or not the presence of breathiness affects this perception.

We do not always observe a significant effect of time on VPO for nasalized vowels. We initially considered the possibility that VPO may be largest during the onset glottal consonant, and that it may gradually fall during a following, coarticulatory nasalized vowel. However, we do not observe a gradual fall in VPO during the vowel for most speakers. Instead, VPO remains steady for most syllables. Further research is necessary to determine whether VPO is also large, or larger, during the preceding consonant.

We observed great variation in degree of VPO when comparing varying vowel heights after the same consonant. Some speakers produced low vowels with greater VPO and high/mid-high vowels with low VPO, while others produce high, mid-low, and low vowels with similar degrees of VPO. These findings suggest that vowel-height-based VPO modulation may not be attributed to passive lingual-velum muscle connection—for example, via the palatoglossus muscle—alone. If this were the case, we would expect to observe a consistent, gradient relationship between VPO and vowel height when comparing syllables after the same consonant. Instead, the more complicated pattern we observe may suggest an active process in response to differences in intraoral acoustic impedance associated with variable tongue height. When the tongue is high, the relative acoustic consequences of nasal coupling are stronger. Conversely, when the tongue is low, intraoral impedance is also low and the relative strength of the same nasal coupling has less of an effect on vocal tract acoustic output (House and Stevens, 1956). This means that greater VPO is required to cause low vowels to sound nasal, i.e., increased velopharyngeal coupling is required for nasal resonances to become salient in the acoustic signal. Speakers may decrease VPO during the production of higher vowels to compensate for increased intraoral impedance and/or increased VPO in low vowels. Indeed, Ohala (1974) suggested that low vowels may require some increased VPO to enhance their acoustic contrast with high vowels.

Our data show great individual variation in degree of VPO. In the case of /daː/ and /diː/, three speakers produce /daː/ with greater VPO than /diː/ throughout at least part of the vowel, while one speaker produced both vowels with the same degree of VPO. We may extrapolate from this behavior that the three speakers are either (1) actively compensating for differences in intraoral impedance or (2) enhancing the acoustic contrast between high and low vowels with variation in VPO.

In the first case, the speakers may increase VPO during low vowels to offset the relatively greater effect nasal coupling has during high vowels. Nasal resonances are relatively stronger during high vowels because intraoral acoustic impedance is high. By increasing VPO during low vowels, nasal resonances are strengthened in the acoustic signal. Without some degree of nasalization on low vowels, high vowels might actually sound more nasal in comparison (Ohala, 1974). In the second case, speakers may take advantage of the difference in intraoral impedance between high and low vowels by increasing VPO on low vowels. This would increase the acoustic differences between high and low vowels.

Finally, we observe no consistent difference in VPO following /tʰ/, /s/ /d/ across speakers. While the presence of aspiration/voiceless frication has been known to induce spontaneous nasalization in some languages (Colarusso, 1988; Blevins and Garrett, 1992; Schadeberg, 1989), coarticulatory breathiness alone does not appear to induce nasalization in Thai. While we observed elevated H1*-H2* after /tʰ/ in some speakers, that is likely associated with increased breathiness since we do not observe increased VPO (in the same speakers but on different occasions). Therefore, while both glottal consonants and voiceless fricative/aspirated consonants have been associated with spontaneous nasalization in other languages, Thai only spontaneously nasalizes vowels after glottal consonants. This suggests that in order for a vowel to spontaneously nasalize in Thai, the preceding consonant must be glottal, not just incidentally breathy. Velopharyngeal underspecification is likely a crucial factor in spontaneous nasalization in Thai.

We observe a complex set of differences in VPO based on onset consonant and vowel height in Thai. We reviewed three explanations for spontaneous nasalization: (1) vowel height; (2) velopharyngeal underspecification; and (3) breathy-nasal misperception. Our findings show that none of these explanations on its own is sufficient to fully explain spontaneous nasalization in Thai. Besides the coarticulatory nasalized /naː/, Thai speakers only manifest relatively large VPO with glottal consonants (explanation 2) and a low vowel nucleus (explanation 1). The fact that vowels after /h/ are produced with greater H1*-H2* just after consonant release suggests that there may also be influence from breathy-nasal misperception or enhancement (explanation 3). Further perceptual and laryngeal articulatory work is needed to explore this possibility. A major challenge in developing a full understanding of spontaneous nasalization is designing methods to test each proposed explanation in isolation. In a language like Thai where multiple articulatory variables may influence VPO and perceived nasalization, a study that considers articulatory measures of both VPO and the larynx, acoustic output, and native speaker perception, will be necessary. Our first step in understanding the puzzle of spontaneous nasalization in Thai reveals important pieces about the role of onset consonant and vowel height in measures of VPO and spectral tilt; many more pieces about laryngeal configuration and effects on perception have yet to be incorporated and explained.

Due to the small sample size of the present study, the weight and generalizablility of the results should be carefully considered. Furthermore, the speech samples analyzed in this data set are limited in scope to a single prosodic context that is phrase-medial and prominent. Research has shown that phonetic realization can vary based on prosodic context, i.e. boundary location and prominence/focus (**????**). For example, Cho *et al.* (2017) found that vowels in phrase-final, non-focused words were produced with more coarticulatory nasalization than in phrase-initial, focused words in English. It is possible that prosodic context may also affect spontaneous nasalization in Thai, e.g. we might observe increased nasalization during non-prominent contexts. Further study is needed to assess the effect of prosodic context on spontaneous nasalization in Thai.

Chapter 4

STUDY 2: AERODYNAMICS AND ELECTROGLOTTOGRAPHY

4.1 Methods

4.1.1 Speakers

While participants from the MRI study were invited to return for a second aerodynamic and EGG study, no participants from the MRI study were available. Six different speakers of Central Thai were recruited to participate in the aerodynamic and EGG study. All participants grew up in central Thailand near Bangkok, spoke Central Thai as a native language, were between 18-26 years old, and reported no speech or hearing problems. Four participants were female and two were male. All participants provided informed consent in accordance with local IRB guidelines.

4.1.2 Materials

While words from the MRI study were also used in the airflow/EGG study, many new items were added to increase the number of vowel height and consonant contrasts represented in Thai. The speech materials for this study are designed to vary in vowel height and onset consonant. The words and morphemes chosen are designed to vary minimally by tone and vowel length. All three low vowels reported to nasalize in Thai, /a/, /ɛ/, and /ɔ/ are included after both /h/ and /ʔ/. We also include /h/ and /ʔ/ paired with a high vowel, /e/. The syllable /naː/ is included as a predictably nasalized environment for comparison. /t^{h}/ at three vowel heights are included in order to asses nasal airflow during oral vowels at varying height.

Breathy and nasal levels are defined based on predicted allophonic nasalization of the vowel and laryngeal configuration of the preceding consonant. Levels include syllables with the following onsets: nasal breathy /h/, nasal non-breathy /n/ and /ʔ/, non-nasal breathy /tʰ/, and non-nasal non-breathy /d/. Classification also depends on vowel height. Syllables /haː,hɛː,hɔ/ are breathy nasal while /heː/ is breathy non-nasal; Syllables /ʔaː,ʔɛː,ʔɔ/ are non-breathy nasal while /ʔeː/ is non-breathy non-nasal. These levels are based on observations of Thai nasalization by Matisoff (1975) and Cooke (1989). These levels provide contrasts between the presence of breathiness, nasalization, and a non-nasal non-breathy control. See Table 4.1 for a complete list of words and morphemes. The words were produced in the same carrier phrase as in the MRI experiment: /pʰuːt˥˩kʰ am˧waː˥˩ "____" ʔiːk˨kʰ raŋ˦/ "Say the word '____' again".

Word	Gloss	Breathy/nasal level
/haː˧/	'to guffaw'	breathy/nasal
/hɔː˧/	'44th letter of Thai alphabet'	breathy/nasal
/hɛː˧˩/	'parade'	breathy/nasal
/heː˧/	'to swiftly flock'	breathy/non-nasal
/ʔaː˧/	'aunt'	non-breathy/nasal
/ʔɔː˧/	'crowd together'	non-breathy/nasal
/ʔɛː˧/	'small / young'	non-breathy/nasal
/ʔeː˧/	'pause (interjection)'	non-breathy/non-nasal
/naː˧/	'rice farm'	non-breathy/nasal
/tʰaː˧/	'to apply / paint'	breathy/non-nasal
/tʰɛː˧˥/	'genuine, true'	breathy/non-nasal
/tʰeː˧/	'to collect; to slant'	breathy/non-nasal
/daː˧/	'to advance in a group'	non-breathy/non-nasal

Table 4.1: Speech Materials. In the breathy/nasal column, 'nasal/non-nasal' refers to predictions about whether or not the vowel nucleus of a syllable is nasalized; 'breathy/non-breathy' refers to whether the vowel nucleus is predicted to undergo any coarticulatory breathiness from the preceding consonant. In this data set, vowels after /h/ and /$t^{/}$ are predicted to undergo some coarticulatory breathiness because these consonants are produced with voiceless transglottal airflow (Blankenship, 2002). Note that this does not refer to phonemically contrastive breathiness, but rather phonetic coarticulation.

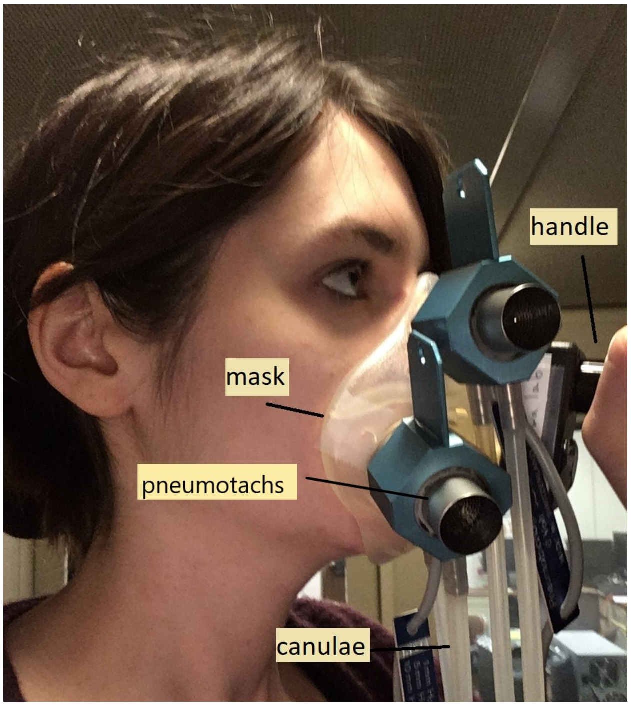


Figure 4.1: Airflow equipment diagram, demonstrated by the author.

4.2 Data acquisition and analysis

4.2.1 Aerodynamics

Data was collected with Acq*Knowledge* data acquisition and analysis software (BIOPAC, version 3.9.1). All data was sampled at 2000 Hz, the maximum allowable sampling rate with the MP100. Participants held a Glottal Enterprise OroNasal mask against the mouth and nose (Figure 4.1). This double-compartment mask is separated to capture oral and nasal airflow independently. Two sizes of mask were provided, an adult size and child size. In some cases the participants' face was too small to fit into the adult size mask and the child size mask was used instead. Two heated pneumotachs were inserted into vents on the mask. Rubber cannulae connected the pressure ports to a Biopac TSD160a pressure transducer that recorded +-12 cm H20. A BIOPAC AFT6 600 ml calibration syringe was used to calibrate the signal for both the oral and nasal mask compartments. The airflow signal was also rectified during each recording session by adjusting the signal to zero during a voiceless

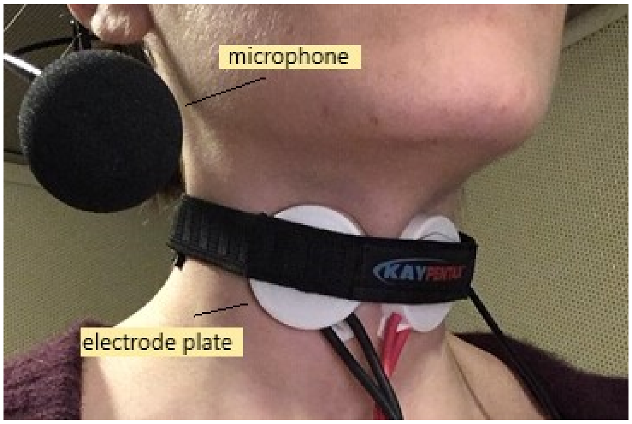


Figure 4.2: EGG electrode equipment, demonstrated by the author.

stop /k/.

While holding the mask against the mouth and nose, participants recited each sentence shown consecutively on a tablet screen. Participants produced each of the 13 items within the carrier phrase 12 times, resulting in 156 tokens per speaker.

4.2.2 Electroglottography and audio recordings

EGG and high quality acoustics were recorded in a separate recording immediately after participants performed the nasal airflow experiment. This sequence of tasks was chosen because the nasal airflow mask interferes with acoustic recording by reducing acoustic amplitude and adding mask resonances. Therefore, it was necessary to record high quality acoustics separately. EGG and acoustics were simultaneously recorded at 44.1 kHz. Acoustic data were collected with a C520 head-set microphone recorded at 44.1 kHz (AKG Harman, Stamford, CT). Acoustic data were collected on a PMD570 Solid State Recorder (Marantz Professional, Cumberland, RI) with a Grace m101 preamplifier (Grace Designs, Lyons, CO). The EGG signal was recorded with an electroglottograph, EG2-PCX2 (Glottal Enterprises, Inc., Syracuse, NY). Two electrode plates were placed on either side of each thyroid lamina with a conductive gel and were held in place with a soft band (Figure 4.2). The electrode plates measure electrical resistance across the glottis by emitting a low-voltage current across the vocal folds. When the vocal folds are touching this electrical current is better able to pass through (Baken and Orlikoff, 2000). This method allows us to approximate the time that the vocal folds are apart (open quotient) and together (closed quotient). Breathy phonation is characterized by a high open quotient. During breathy phonation the vocal folds are typically vibrating but only lightly touching during the brief closure phase (Esling and Harris, 2005).

4.2.3 Data analysis

Aerodynamics

Aerodynamic data included in the analysis are filtered nasal and oral airflow (l/s) over time. Airflow data was filtered using a 2nd order Butterworth low-pass filter in Matlab (2015b) using the *butter* and *filtfilt* functions. 20 equidistant airflow samples were taken during the syllable for analysis, 10 during the consonant and 10 during the following vowel. Nasal airflow proportion was also analyzed by dividing the nasal airflow during the vowel, normalized to 20 equidistant samples, by the sum of the nasal and oral airflow during the vowel (nasal / oral + nasal). Proportional nasal airflow during the vowel is a useful measure because fluctuation in subglottal air pressure will affect both nasal and oral airflow. Subglottal air pressure cannot be easily controlled throughout the experiment. Therefore, a measure of nasal airflow relative to oral airflow is helpful in ascertaining the relative strength of nasal airflow. However, it is also necessary to consider overall raw nasal airflow. Even if proportional nasal airflow is low, if raw nasal airflow is also quite high, i.e. both nasal and oral airflow are high to similar degrees, this suggests large VPO.

Before this division took place, all airflow data was adjusted within each speaker to eliminate slightly negative values. Due to minor electrical noise in the recording system, small negative values are sometime recorded when airflow is near zero, e.g. $-1x10^{-10}$ μv. The presence of a small negative value would make an analysis of nasal airflow proportion

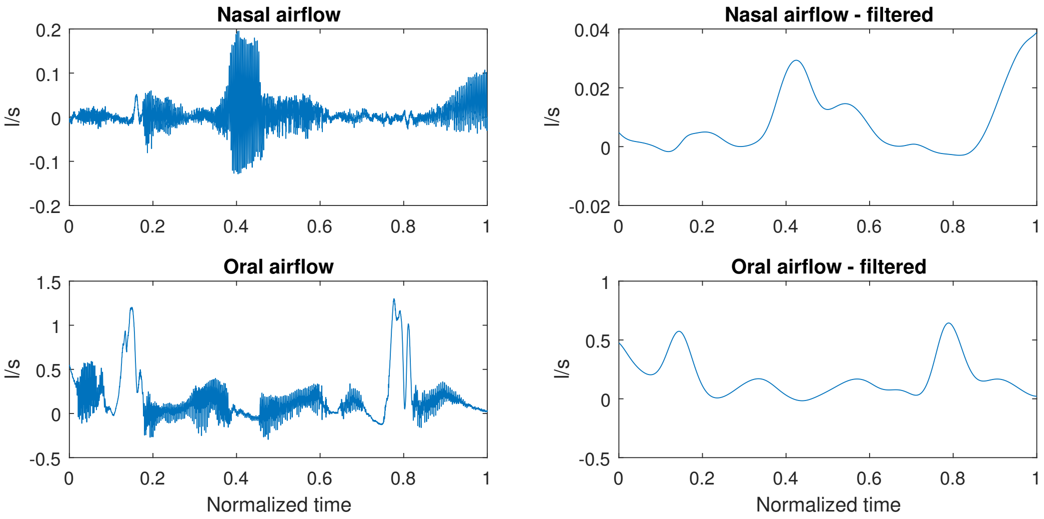


Figure 4.3: Raw and filtered nasal and oral airflow data. Production of /pʰuːt˥˩kʰ am˧waː˥˩ ‘naː˧’ iːk˧kʰ raŋ˥/, “Say the word ‘rice farm’ again.”

inaccurate. For example if nasal airflow were a very small negative number (essentially zero with minor error) and oral airflow were positive, a measure of nasal proportion relative to oral airflow would yield a negative number. To address this problem, nasal and oral airflow were corrected to be positive by subtracting the smallest airflow value of both oral and nasal airflow data (an extremely small negative number in all cases) from the airflow data.

Finally, the integral was taken of all samples within the consonant and vowel separately. This yielded a single measure of cumulative nasal airflow during both the consonant and vowel. For each token, the location of the maximum of integrated nasal airflow was determined and logged as either occurring during the onset consonant or the vowel.

EGG and acoustics

EGG data was analyzed by measuring the open quotient of the vocal fold cycle. Open quotient is the proportion of the time that the vocal folds are open rather than closed during vocal fold vibration. This measure was obtained using the DEGG method based on research by Henrich *et al.* (2004) on the use of the derivative of the EGG signal to calculate open quotient. A set of Matlab scripts by Michaud *et al.* (2017) were adapted to calculate open quotient using the EGG signal derivative. Open quotient was calculated at 20 equidistant time points through every vowel. See Figure 4.4 for an example of the EGG and DEGG signals of modal phonation. The timing of glottal closure corresponds to the highest positive peak in the DEGG signal, while the glottal opening corresponds to the lowest negative peak. Childers *et al.* have studied the physiological correlates of the DEGG opening and closing peaks by synchronizing the EGG and DEGG signals with ultra-high speed cinematography (inverse-filtered derived glottal flow and glottal area) (Childers *et al.*, 1983; Childers, 1983; Childers *et al.*, 1986, 1990). From these studies, researchers conclude that the peaks within the DEGG signal are reliable indicators of the opening and closing phases of the vocal folds in most contexts.

For some contexts such as extremely creaky phonation, a measure of open quotient is unobtainable because glottal opening is very gradual and occurs continuously along the vertical length of the vocal folds. An example of creaky vocal fold vibration without an open quotient can be found in Figure 4.5. During creaky phonation, closing at the inferior end of the vocal folds often occurs before the vocal folds fully open at the superior end. This pattern of movement has been referred to as air "bubbling" between the strongly adducted vocal folds (Zemlin, 1997). Michaud *et al.* (2017) argue that in cases where there is no distinguishing time between opening and closure, it is unintuitive to discuss the vocal folds as having an open quotient. Rather, other physiological aspects such as vibratory rate, glottal configuration, or glottal flow are more informative measures in this case. In the present data set, we find four phonetic contexts where phonation is so creaky that open quotient cannot be measured. These contexts include brief intervals (5-30 ms) just after release of glottal stop. The speech materials containing these contexts are /ʔaː/, /ʔɛː/, /ʔoː/, and /ʔeː/ in some, but not all, speakers. In cases where open quotient is unanalyzable due to extreme creakiness, values are left blank and the presence of creakiness is remarked upon in the figure legend. Cases of nonexistent open quotient were determined through visual inspection of all EGG and DEGG signal during analysis, such that the process of

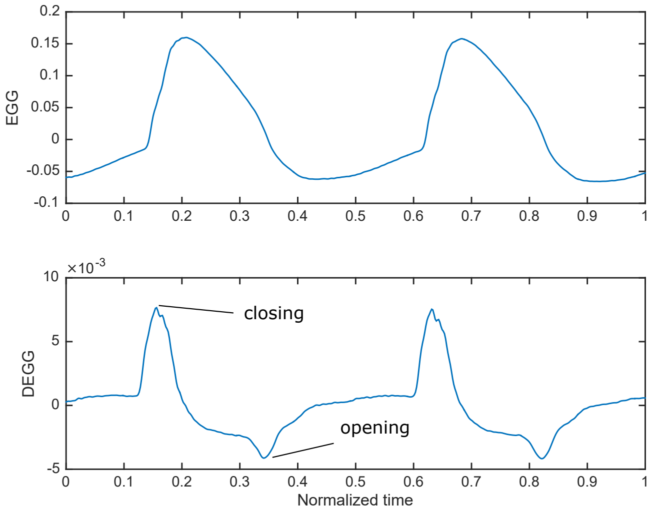


Figure 4.4: EGG and DEGG signals of modal phonation.

open quotient calculation was semi-automated. For most cases where open quotient was analyzable, the script automatically calculated open quotient using peak detection of the DEGG signal. Because the present study is primarily concerned with the identification of breathy phonation, we feel that this qualitative assessment of extremely creaky phonation for lack of a more direct physiological measure is sufficient.

Spectral tilt was measured using time-varying values of the energy ratios of low/high frequencies. This measure is a spectral metric that has been used to estimate nasality because it can detect the presence of extra resonances introduced at low frequencies (Pruthi and Espy-Wilson, 2005). The energy ratio of bands between 0–320 Hz and 320–5360 Hz is expected to be higher when both nasal resonances and subglottal resonances (associated with a wider glottis as with breathiness) are present (Carignan, 2017; Stevens, 2000). This measure was automatically extracted using a custom Praat script that calculated the energy ratio of 0–320 and 320–5360 bands (Boersma and Weenink, 2018). For the current study this full-spectrum measure of spectral tilt is preferred to a measure of H1-H2 that could otherwise be used to detect breathiness. Simpson (2012) finds that variations in vowel height and gender make the use of H1 and H2 amplitude an unreliable measure of phonation. Some

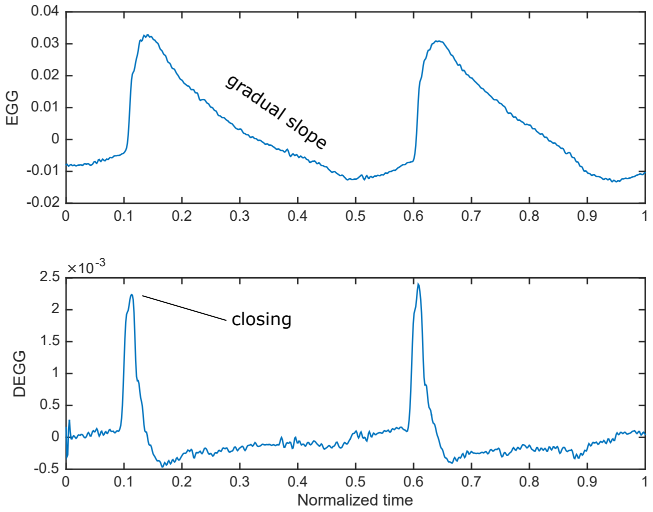


Figure 4.5: EGG and DEGG signals of creaky phonation.

nasal coupling is typically present during low vowels and the lowest nasal pole affects the first few harmonics. Since males have a typical f0 of 120 Hz the nasal pole will enhance the second and third harmonic. In females, since f0 is higher (around 200 Hz), the nasal pole will usually enhance the first harmonic. Therefore, the use of H1 amplitude in measures of phonation is unreliable across vowels and genders. A broader spectral analysis, e.g. a measure of energy ratio conducted in the current study, is therefore advisable (Simpson, 2012).

4.2.4 Statistical methods

We analyzed time-series airflow, EGG, and acoustic data using generalized additive mixed models (GAMMs) (Wood, 2006). This method has previously been used to analyze time-series nasal airflow data and other phonetic data (Barlaz *et al.*, 2018; Desmeules-Trudel and Brunelle, 2018; Ko *et al.*, 2014; Nixon and Best, 2018; Sóskuthy, 2017; Wieling *et al.*, 2016; Wieling, 2018); it allows for the inclusion of random effects and auto-correlation parameters for time-series data. The models were created using the *mgcv* package in R (Wood and Wood, 2009). Two models were constructed for the nasal airflow data. The first model contained raw nasal airflow values (l/s) as the dependent variable. Airflow data from the entire syllable (consonant and vowel) was included. The independent variables chosen were Syllable, Segment (consonant or vowel), and a smoothed function through Time. Smoothed functions through Time by Syllable and Segment were also included. Finally an interaction between smoothed Time and Syllable by Speaker was included in order to account for speaker-specific variation. The best model fit was determined using the *compareML* function in R. The independent variables were selected in a top-down manner, where one variable was excluded at a time and compared to a model that did not exclude the variable. It was determined that the optimal model, i.e. the model with the lowest AIC score, included all of the independent variable listed above. An AR-1 correlation parameter was included in the model to account for any autocorrelation of the time-dynamic data, i.e. to prevent violation of the assumption of independence (Baayen *et al.*, 2018; Desmeules-Trudel and Brunelle, 2018; Wieling, 2018). The final model was constructed using restricted maximum likelihood (REML).

The second GAMM model included proportional nasal airflow as a dependent variable. Independent variables included Syllable, a function through smoothed time, and interactions between Syllable and Time as well as Syllable and Time by Speaker. The model containing all of these independent variables was determined to be optimal compared to each test model where one independent variable was left out using the *compareML* function in R. Similar to the previous nasal airflow model, an AR-1 correlation parameter was included and the final model was constructed using REML.

The first model included nasal airflow values (l/s) during the vowel as the dependent variable, Syllable as an independent variable, a smoothed function through normalized time (0 to 1), and Speaker as a random effect. Interactions between smoothed time and Syllable and smoothed time and Speaker by Syllable were also included. Similar models were constructed for nasal airflow during the syllable consonant (l/s) and proportional nasal airflow (o to 1). An AR-1 correlation parameter was included in the model to account for any autocorrelation of the time-dynamic data, i.e. to prevent violation of the assumption of independence (Baayen *et al.*, 2018; Desmeules-Trudel and Brunelle, 2018; Wieling, 2018). The final model was constructed using restricted maximum likelihood (REML). Finally, the model output, including time-series data plots, were visualized using the *itsadug* package in R (van Rij *et al.*, 2016). These visualizations are useful for interpreting differences in the relevant measures, e.g. nasal airflow or EGG, across syllables.

Two more GAMMs were fitted, one with open quotient as a dependent variable and another with spectral energy ratio as a dependent variable. In both models, it was determined that the optimal independent model structure should include Syllable, a smoothed function through time, an interaction between Syllable and Time, and an interaction between Syllable Time by Speaker. In the case of the open quotient model, 1.5% of the data constituted NaN values (see Section 4.2.3). In the EGG model, NaNs were handled by using the na.gam.replace option that allows mean imputation of the missing values. The final EGG and acoustic GAMM model included an AR-1 correlation parameter and were constructing using REML.

Finally, the model output, including time-series data plots, were visualized using the *itsadug* package in R (van Rij *et al.*, 2016). These visualizations are useful for interpreting differences in the relevant measures, e.g. nasal airflow or EGG, across syllables and speakers.

Cumulative integrated nasal airflow data was analyzed with a linear mixed effects model using the *lmerTest* package in R. Integrated nasal airflow was the dependent variable with fixed effects including syllable, location of maximum integrated flow (MaxLoC), and segment type (consonant or vowel). Speaker was included as a random effect.

4.3 Results

4.4 Airflow

4.4.1 Time-series airflow

Figure 4.6 shows mean nasal airflow (l/s) of all syllables with an /aː/ nucleus. These syllables include /haː/, /ʔaː/, /naː/, /daː/, and /tʰaː/. Of these syllables, /haː/ and /ʔaː/ are predicted to nasalize in Thai (Matisoff, 1975). The onset consonant of /naː/ is nasal, while the vowel of /naː/ is expected to undergo coarticulatory nasalization (Beddor and Krakow, 1999). In Figure 4.6, confidence interval bars are based on the standard error (R Development Core Team, 2011). Each token is composed of 10 normalized, equidistant points during the onset consonant and 10 points during the following vowel. The boundary between the onset and vowel nucleus is indicated by a vertical dotted line. /h/ is produced with high nasal airflow, over twice as much as the next highest airflow – that of /n/. /ʔ/, /tʰ/, and /d/ are all produced with low nasal airflow, with /ʔ/ exhibiting the lowest amount of nasal airflow that is close to zero. During the following vowel portion, we observe high nasal airflow during the vowel of /haː/ that is again over twice as high as /naː/, but that gradually lessens throughout the vowels until nasal airflow is similar in the vowels of /haː/ and /naː/ about halfway through the vowel. Nasal airflow during the vowel of /tʰaː/ and /daː/ is close to zero, while airflow during the vowel of /ʔaː/ gradually rises after about 10% into the vowel. Although nasal airflow rises slightly during this vowel by about 0.01 l/s, it does not reach the level of /haː/ and /naː/.

Plots of individual speaker data in Figure 4.7 shows little variation among speakers. In general all speakers produce /haː/ with the greatest nasal airflow during the consonant and initial part of the vowel, followed by /naː/. Both syllables' nasal airflow gradually falls during the vowel to become similar. We observe some variability in /ʔaː/ across speakers. While usually similar to /tʰaː/ and /daː/ in all speakers, it is only slightly higher during

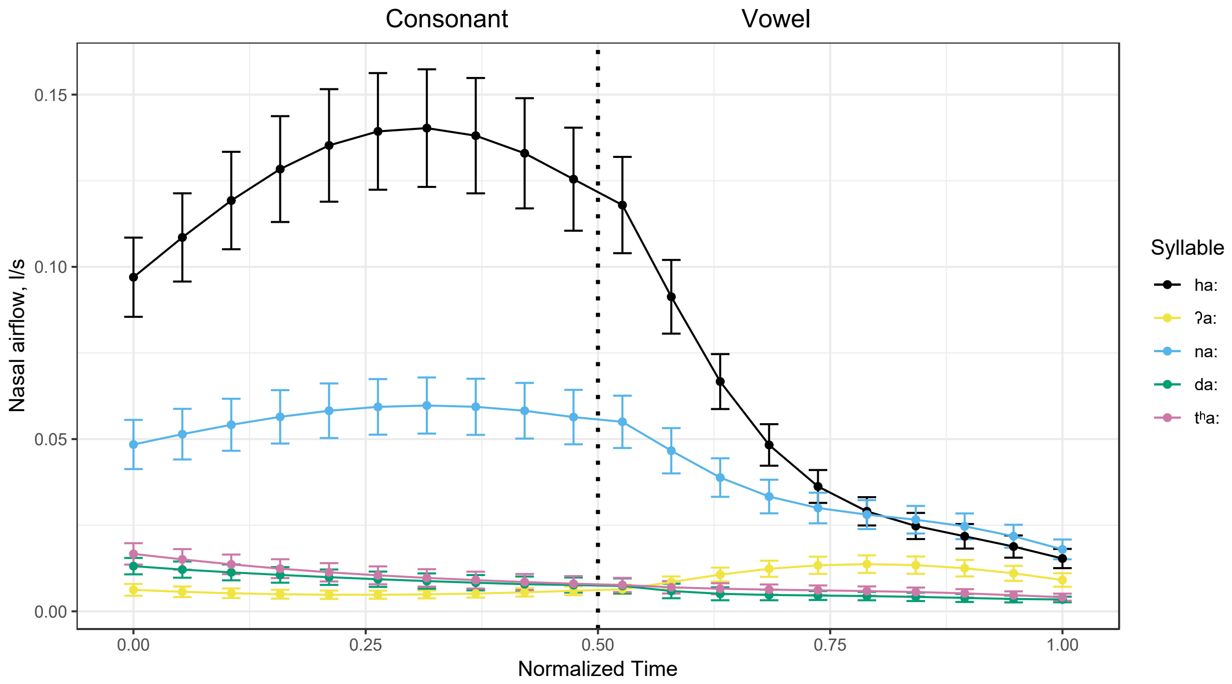


Figure 4.6: Mean nasal airflow of all syllables containing the vowel /a/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined.

the vowel for TF2 and TF4. The greatest interspeaker difference we observe is that the two male speakers, TM1 and TM2, exhibit overall reduced nasal airflow compared to the female speakers by about 0.1 l/s in the largest gap among productions of the consonant /h/. This may be because the two male participants, who happened to have larger faces, wore a larger mask than the female participants. The nasal compartment of the large mask is larger than the mask worn by the female participants. It is possible that the larger amount of space within the nasal compartment of the large mask allowed some nasal air to dissipate within the space of the compartment, such that there was weaker pressure exerted on the transducer. Without testing all participants with an identical mask, which is unlikely given differences in facial size, it is unclear whether differences in overall magnitude of nasal airflow between genders are attributed to mask size or actual nasal airflow produced. Regardless of the source of this difference, we do observe similarities in overall nasal airflow behavior in both males and females.

As discussed in Section 4.2.4, two GAMM models were constructed for nasal airflow data:

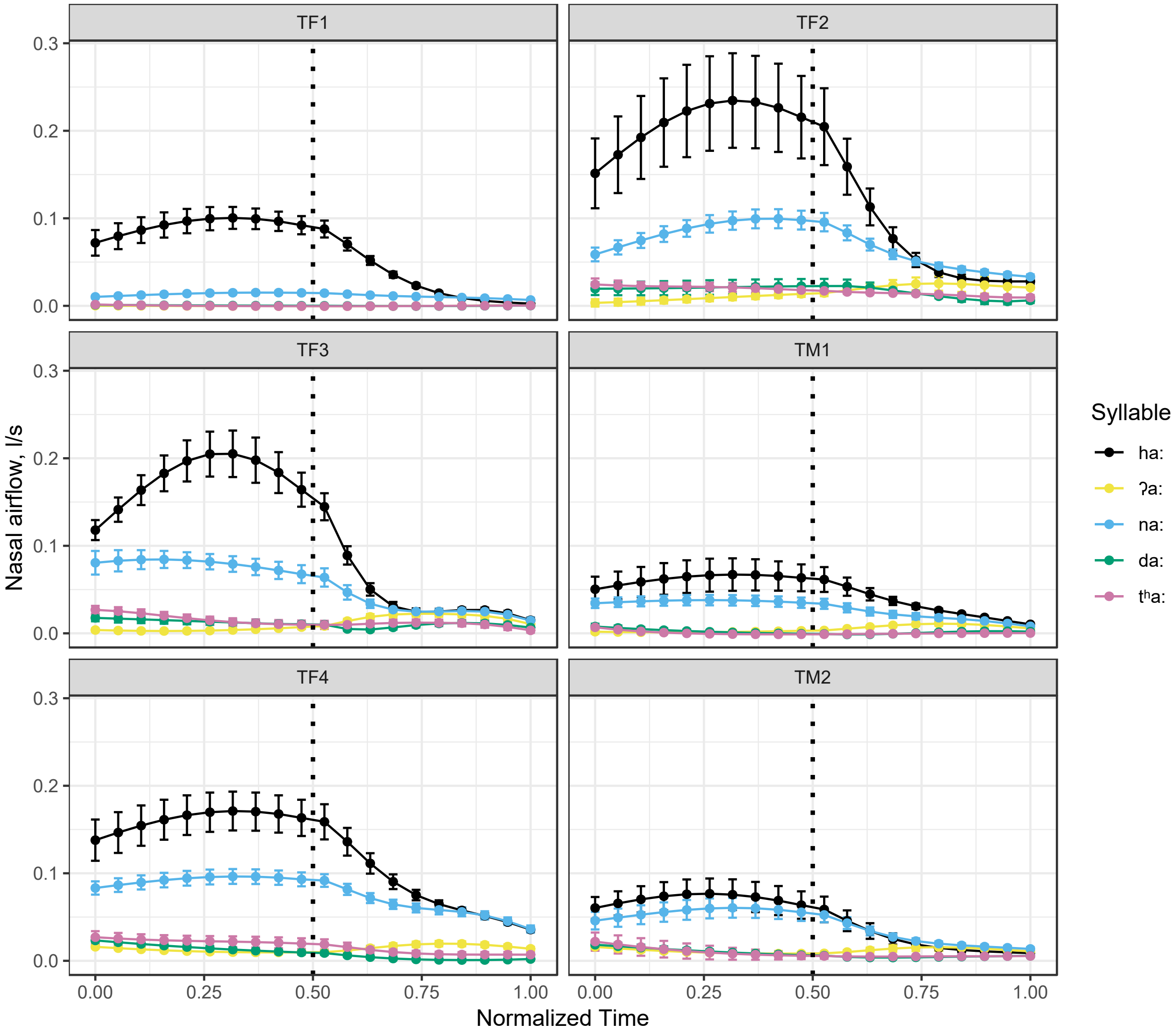


Figure 4.7: Mean nasal airflow (liters/second) during of all syllables containing the vowel /a/. Confidence interval bars are based on the standard error.

	Nasal airflow
(Intercept)	0.09(0.01)***
Syllable /hɛː/	−0.04(0.02)*
Syllable /heː/	−0.05(0.02)**
Syllable /hɔː/	−0.03(0.02)
Syllable /ʔaː/	−0.08(0.01)***
Syllable /ʔɛː/	−0.08(0.01)***
Syllable /ʔeː/	−0.08(0.01)***
Syllable /ʔaː/	−0.08(0.01)***
Syllable /naː/	−0.04(0.02)*
Syllable /daː/	−0.08(0.01)***
Syllable /t^{h}aː/	−0.08(0.01)***
Syllable /t^{h}ɛː/	−0.08(0.01)***
Syllable /t^{h}eː/	−0.08(0.02)***
SegmentVowel	0.00(0.00)
EDF: s(Time)	1.00(1.00)
EDF: s(Time):Syllable /haː/	5.33(6.38)***
EDF: s(Time):Syllable /hɛː/	3.25(4.08)**
EDF: s(Time):Syllable /heː/	1.69(2.36)
EDF: s(Time):Syllable /hɔː/	3.77(4.44)**
EDF: s(Time):Syllable /ʔaː/	1.00(1.00)
EDF: s(Time):Syllable /ʔɛː/	1.00(1.00)
EDF: s(Time):Syllable /ʔeː/	1.00(1.00)
EDF: s(Time):Syllable /ʔɔː/	1.00(1.00)
EDF: s(Time):Syllable /naː/	2.68(3.40)*
EDF: s(Time):Syllable /daː/	1.00(1.00)
EDF: s(Time):Syllable /t^{h}aː/	1.00(1.00)
EDF: s(Time):Syllable /t^{h}ɛː/	1.00(1.00)
EDF: s(Time):Syllable /t^{h}eː/:	1.00(1.00)
EDF: s(Time):SegmentConsonant	0.00(0.00)
EDF: s(Time):SegmentVowel	1.00(1.00)
EDF: s(Speaker,Time):Syllable /haː/	21.23(53.00)***
EDF: s(Speaker,Time):Syllable /hɛː/	12.11(53.00)***
EDF: s(Speaker,Time):Syllable /heː/	11.06(53.00)***
EDF: s(Speaker,Time):Syllable /hɔː/	27.92(53.00)***
EDF: s(Speaker,Time):Syllable /ʔaː/	0.00(53.00)
EDF: s(Speaker,Time):Syllable /ʔɛː/	0.00(53.00)
EDF: s(Speaker,Time):Syllable /ʔeː/	0.00(53.00)
EDF: s(Speaker,Time):Syllable /ʔɔː/	0.00(53.00)
EDF: s(Speaker,Time):Syllable /naː/	5.80(53.00)***
EDF: s(Speaker,Time):Syllable /daː/	1.53(53.00)
EDF: s(Speaker,Time):Syllable /t^{h}aː/	2.33(53.00)
EDF: s(Speaker,Time):Syllable /t^{h}ɛː/	2.05(53.00)
EDF: s(Speaker,Time):Syllable /t^{h}eː/	3.26(53.00)**
AIC	-97604.56
BIC	-96495.67
Log Likelihood	48942.47
Deviance	7.08
Deviance explained	0.80
Dispersion	0.00
R^2	0.80
GCV score	-48653.72
Num. obs.	20128
Num. smooth terms	29

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4.2: GAMM model output for nasal airflow was constructed using the algorithm: Nasal Air ˜s(Time) + Syllable + s(Time, by = Syllable) + Segment + s(Time, by = Segment) + s(Speaker, Time,by=Syllable, bs = ”fs”, m=1).

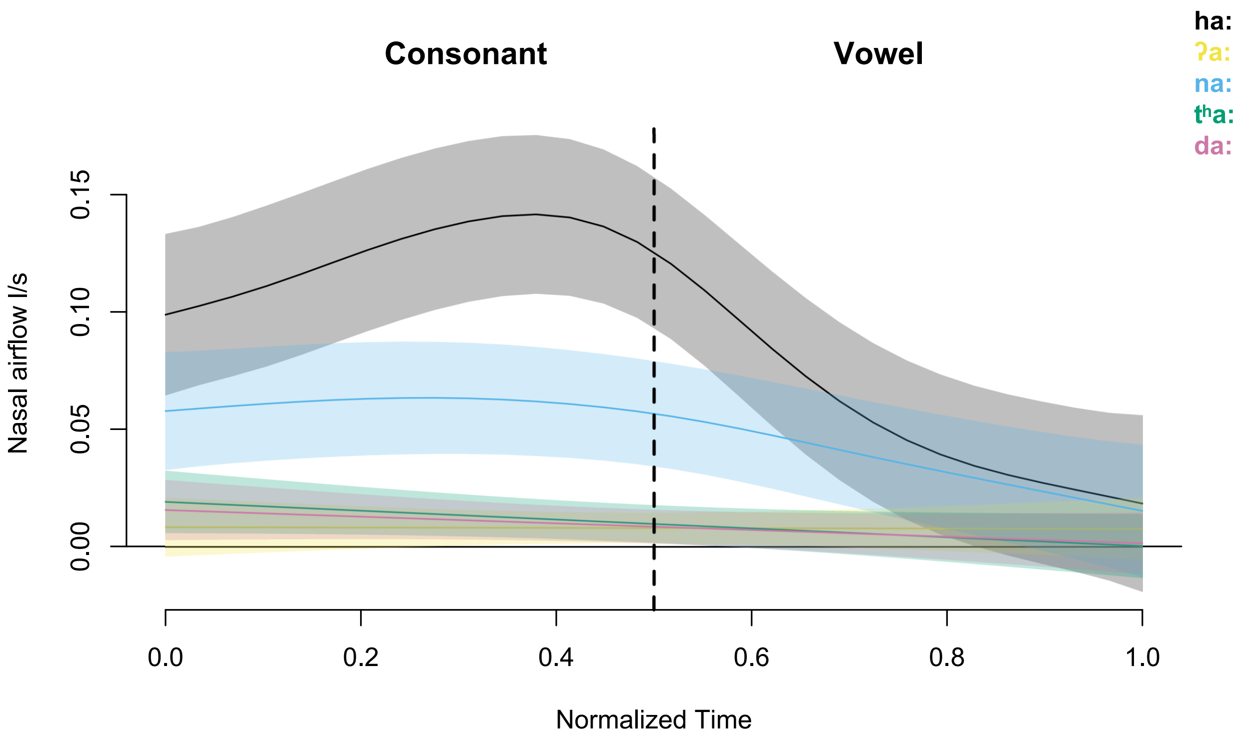


Figure 4.8: GAMM-predicted nasal airflow of all syllables containing the vowel /a/, liters/second. All speakers' data are combined.

one for nasal airflow during the entire syllable (onset and vowel) (AR-1 = 0.88, r^2 = 0.8), and one for proportional nasal airflow (nasal/nasal+oral) during the vowel (AR-1=0.94, r^2 = 0.64). The results of these GAMMs can be found in Tables 4.2 and 4.3.

According to the nasal airflow model (Table 4.2), nasal airflow during /haː/ is significantly different, than all other syllables with a vowel /aː/ nucleus, including /ʔaː/, /naː/, /tʰaː/, and /daː/ (Table 4.2). Nasal airflow is not predicted to vary based on segment type, i.e. onset consonant or vowel.

Recall from Figure 4.6 that while nasal airflow is higher for /haː/ than /naː/ during the initial portion of the vowel, they become similar in all speakers by halfway through the vowel. Indeed, Time is predicted to be an important factor for nasal airflow only for syllables /haː/ and /naː, but no other syllables (Table 4.2). Nasal airflow remains low or close to zero during the syllables /ʔaː/, /tʰaː/, and /daː/.

Figure 4.8 displays predicted nasal airflow values based on the GAMM in Table 4.2. In this figure, the predicted mean is surrounded by a 95% confidence interval. If the confidence interval does not overlap between two syllables, it can be said that they are significantly different from one another. We observe that predicted nasal airflow is similar to data presented in Figure 4.6 representing mean airflow with standard error bars. During the syllable /haː/ nasal airflow is greater than all other syllables during the consonant and at the initial portion of the vowel, until it gradually declines until about halfway through the syllable. Plots of individual speaker predicted data show little interspeaker variation (Figure 4.9). TM2 is predicted to produce /haː/ and /naː/ with similar high nasal airflow while TF1 is predicted to produce /naː/ with diminished nasal airflow.

Meanwhile, we observe a different trend in oral airflow for syllables with the nucleus /aː/ (Figure 4.10). While /tʰaː/ exhibited little nasal airflow during either the consonant or following vowel in Figure 4.6, /tʰaː/ is produced with the greatest oral airflow during both the consonant and vowel. Given that /tʰ/ is an aspirated consonant, produced with a large puff of air after stop release, this finding is not surprising. The syllable /haː/ is also produced with high oral airflow during both the consonant and vowel, although the magnitude of oral airflow during /tʰaː/ is greater than /haː/ during the latter half of the consonant and initial half of the vowel. This is likely due to the aspiration of /tʰaː/. Oral airflow of both /tʰaː/ and /haː/ fall gradually during the initial half of the vowel, such that oral airflow becomes similar with other syllables about halfway through the vowel. Other syllables, /ʔaː/, /naː/, and /daː/ are produced with lower oral airflow compared to /haː/ and /t ʰaː/. A review of Figure A.4 shows that individual speaker productions of oral airflow are similar to the grouped data in Figure 4.10. Only speaker TF2 shows a diminished contrast among syllables, although the mean oral airflow of /tʰaː/ and /haː/ are somewhat greater during the consonant and beginning of the following vowel. Given that both oral and nasal airflow are relatively higher during the syllable /haː/, we turn to a proportional analysis of nasal airflow during the vowel to determine the relative strength of nasal airflow.

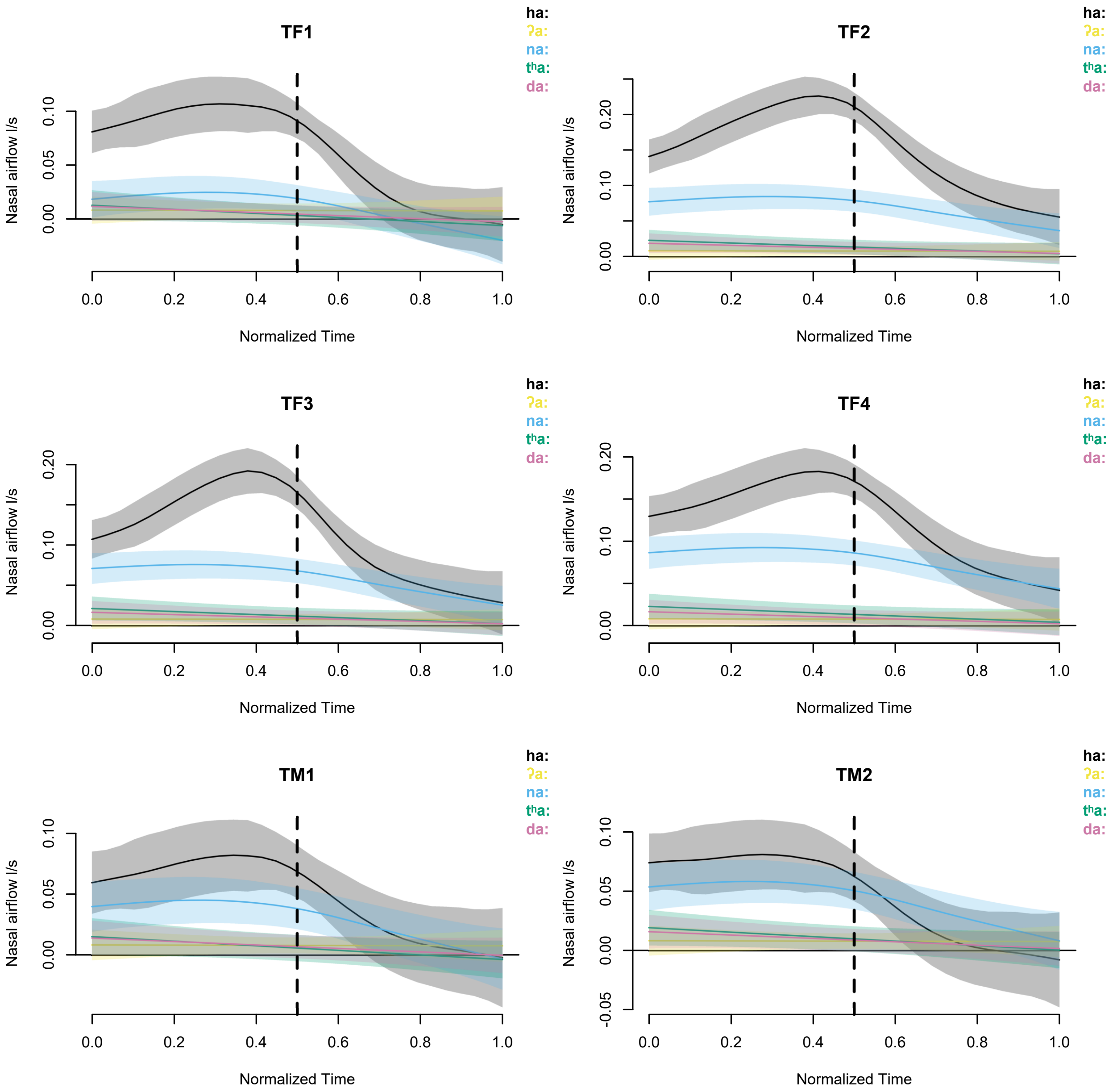


Figure 4.9: GAMM-predicted nasal airflow of all syllables containing the vowel /a/, liters/second.

	Proportion nasal (vowel)
(Intercept)	0.37(0.06)***
Syllable /hɛː/	−0.09(0.08)
Syllable /heː/	−0.13(0.08)
Syllable /hɔː/	−0.11(0.09)
Syllable /ʔaː/	−0.07(0.08)
Syllable /ʔɛː/	−0.11(0.08)
Syllable /ʔeː/	−0.10(0.08)
Syllable /ʔɔː/	−0.14(0.08)
Syllable /naː/	0.02(0.08)
Syllable /daː/	−0.17(0.07)*
Syllable /tʰaː/	−0.20(0.08)**
Syllable /tʰɛː/	−0.17(0.07)*
Syllable /tʰeː/	−0.17(0.07)*
EDF: s(Time)	1.00(1.00)***
EDF: s(Time):Syllable /haː/	1.00(1.00)*
EDF: s(Time):Syllable /hɛː/	1.00(1.00)***
EDF: s(Time):Syllable /heː/	1.00(1.00)**
EDF: s(Time):Syllable /hɔː/	1.00(1.00)
EDF: s(Time):Syllable /ʔaː/	1.00(1.00)*
EDF: s(Time):Syllable /ʔɛː/	1.00(1.00)*
EDF: s(Time):Syllable /ʔeː/	1.47(1.83)*
EDF: s(Time):Syllable /ʔɔː/	0.00(0.00)
EDF: s(Time):Syllable /naː/	2.74(3.54)**
EDF: s(Time):Syllable /daː/	1.00(1.00)*
EDF: s(Time):Syllable /tʰaː/	1.00(1.00)***
EDF: s(Time):Syllable /tʰɛː/	1.00(1.00)***
EDF: s(Time):Syllable /tʰeː/	2.05(2.65)***
EDF: s(Speaker,Time):Syllable /haː/	8.74(53.00)***
EDF: s(Speaker,Time):Syllable /hɛː/	5.84(53.00)***
EDF: s(Speaker,Time):Syllable /heː/	11.75(53.00)***
EDF: s(Speaker,Time):Syllable /hɔː/	18.53(53.00)***
EDF: s(Speaker,Time):Syllable /ʔaː/	11.46(53.00)***
EDF: s(Speaker,Time):Syllable /ʔɛː/	3.57(53.00)***
EDF: s(Speaker,Time):Syllable /ʔeː/	3.50(53.00)**
EDF: s(Speaker,Time):Syllable /ʔɔː/	3.49(53.00)***
EDF: s(Speaker,Time):Syllable /naː/	8.67(53.00)*
EDF: s(Speaker,Time):Syllable /daː/	4.01(53.00)*
EDF: s(Speaker,Time):Syllable /tʰaː/	3.77(53.00)***
EDF: s(Speaker,Time):Syllable /tʰɛː/	3.01(53.00)*
EDF: s(Speaker,Time):Syllable /tʰeː/	4.33(53.00)*
AIC	-46721.56
BIC	-45634.23
Log Likelihood	23498.22
Deviance	150.07
Deviance explained	0.64
Dispersion	0.05
R^2	0.64
GCV score	-23297.19
Num. obs.	20160
Num. smooth terms	27

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4.3: Proportion nasal (during the vowel): Proportional nasal airflow ˜s(Time) + Syllable + s(Time, by = Syllable) + s(Speaker, Time, by=Syllable, bs = ”fs”, m=1).

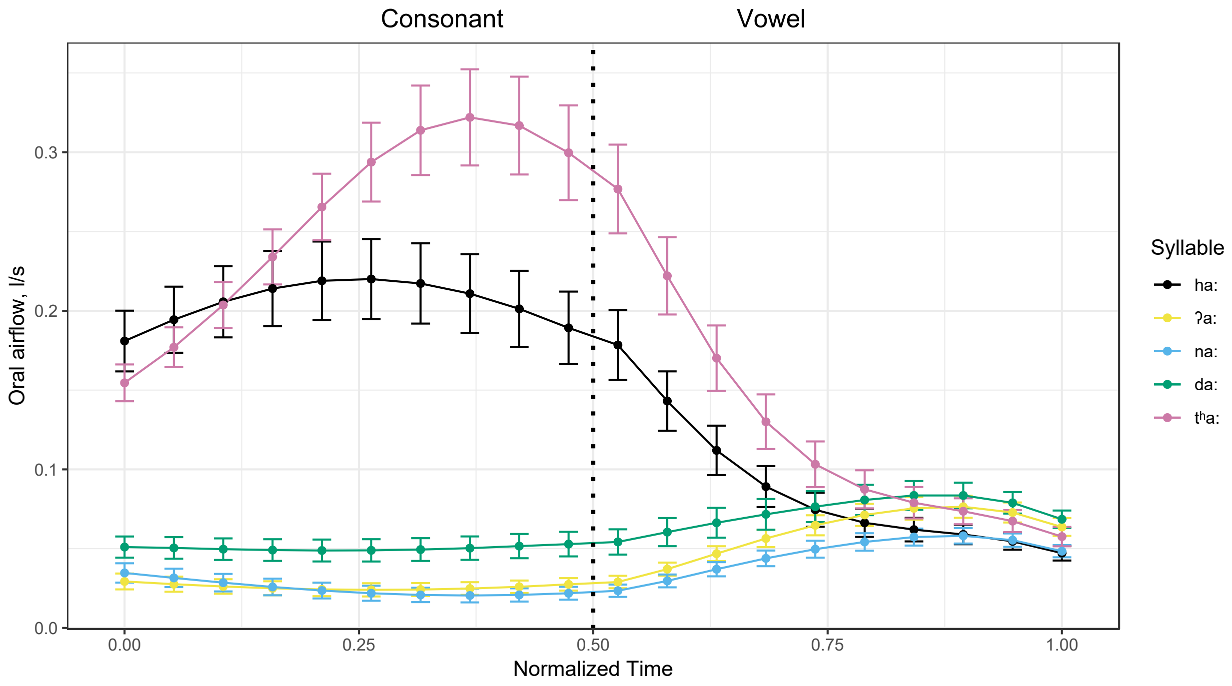


Figure 4.10: Mean oral airflow of all syllables containing the vowel /a/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined.

Figure 4.11 shows proportional nasal airflow (nasal/oral+nasal) during the syllable vowel of all syllables with the vowel nucleus /a/. According to the GAMM model (Table 4.3), /haː/ is not predicted to be overall significantly different from /naː/ and /ʔaː/ in proportional nasal airflow, while it is in comparison to /tʰaː/ and /daː/. Furthermore, TIME is a significant factor in the proportional nasal airflow of each syllable with the nucleus /aː/. At the beginning of the vowel, all syllables exhibit different proportional nasal airflow in the order: /naː/ > /haː/ > /ʔaː/ > /daː/ > /tʰaː/. While /haː/ exhibits the greatest nasal airflow during the beginning of the vowel, /haː/ is produced with less proportional nasal airflow than /naː/. Towards the center of the vowel, the proportional nasal airflow of /haː/ and /naː/ converge. /ʔaː/, /daː/, and /tʰaː/ converge at the center of the vowel as well.

An analysis of individual speaker variation of proportional nasal airflow in Figure 4.12 reveals that /naː/ is similarly large in proportional nasal airflow during the initial half of the vowel across all speakers. /haː/ is produced with the second highest value of proportional nasal airflow during the initial half of the vowel in TF1, TF3, and TF4, and manifests

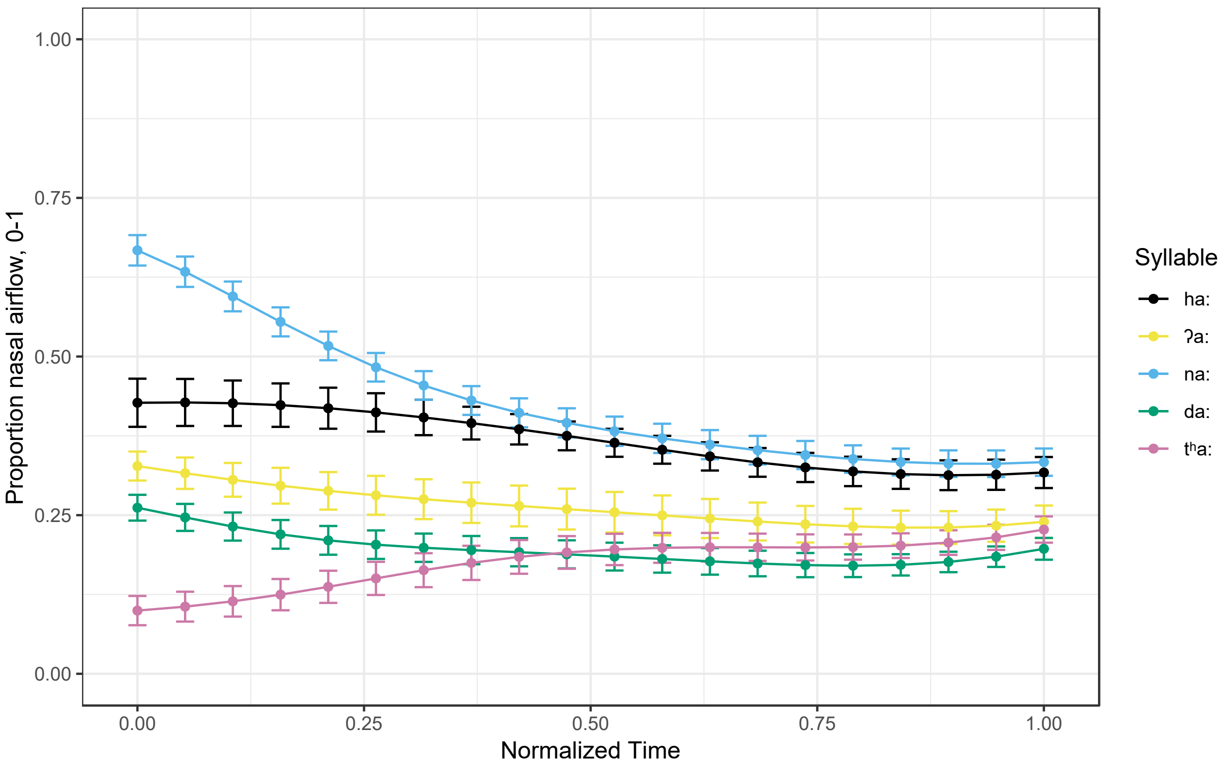


Figure 4.11: Mean proportional nasal airflow during vowel /a/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined

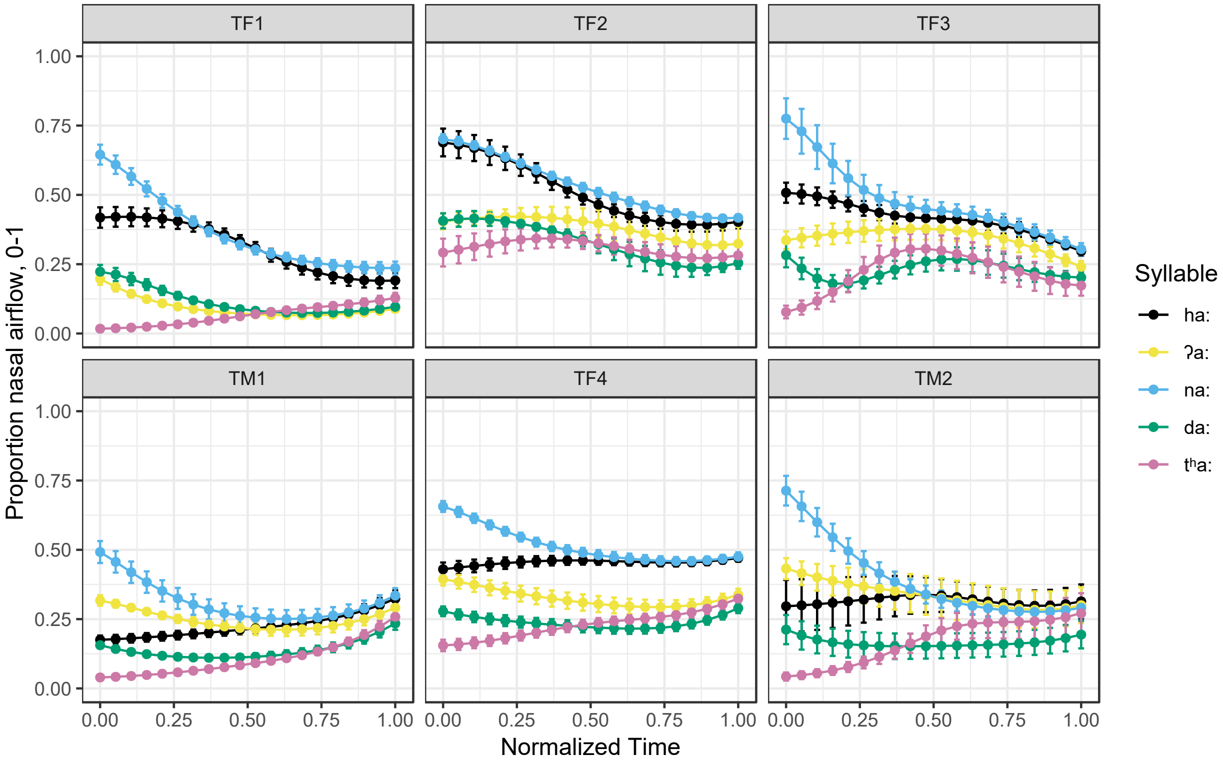


Figure 4.12: Mean proportional nasal airflow during the vowel /a/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral)Confidence interval bars are based on the standard error.

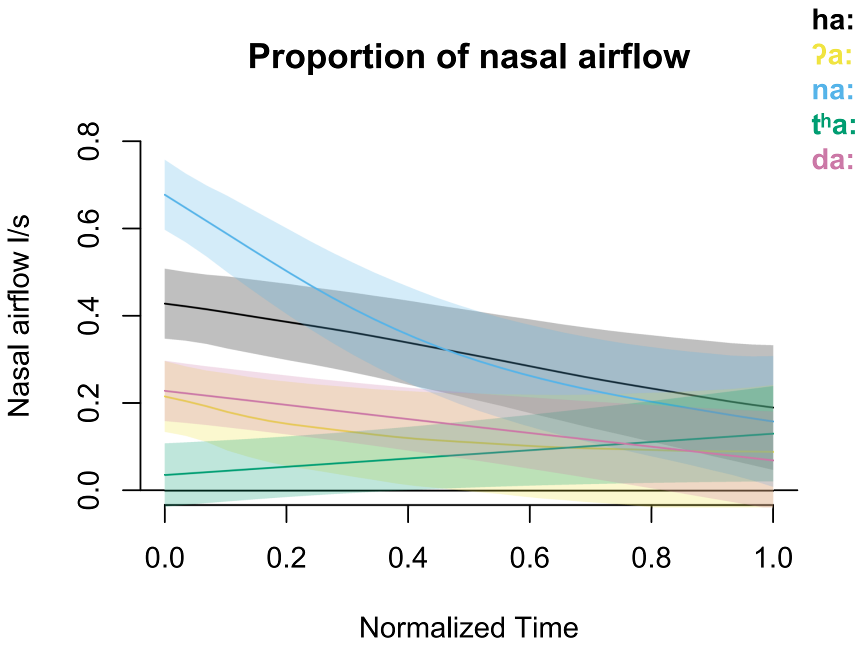


Figure 4.13: GAMM-predicted proportional nasal airflow during vowel /a/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined.

similar proportional airflow as /naː/ in TF2. In TM1 and TM2, /ʔaː/ is produced with greater proportional airflow than /haː/.

The predicted values of proportional nasal airflow can be found in Figure 4.13. We observe similar predicted proportional nasal airflow as in Figure 4.11 of actual proportional nasal airflow data. /naː/ is predicted to be produced with the highest proportional nasal airflow at the beginning of the vowel, followed by /haː/. /ʔaː/ and /daː/ are predicted to exhibit a similar degree of proportional nasal airflow, followed by /tʰaː/. Towards the center of the vowel, all proportional nasal airflow is similar. We observe some individual speaker variability in predicted proportional nasal airflow in Figure 4.14. All speakers except TF2 are predicted to produce /naː/ with the greatest proportional nasal airflow at the beginning of the vowel, while most speakers differ in the order thereafter. Only TF1 is predicted to produce /haː/ with the second greatest proportional nasal airflow at the start of the vowel,

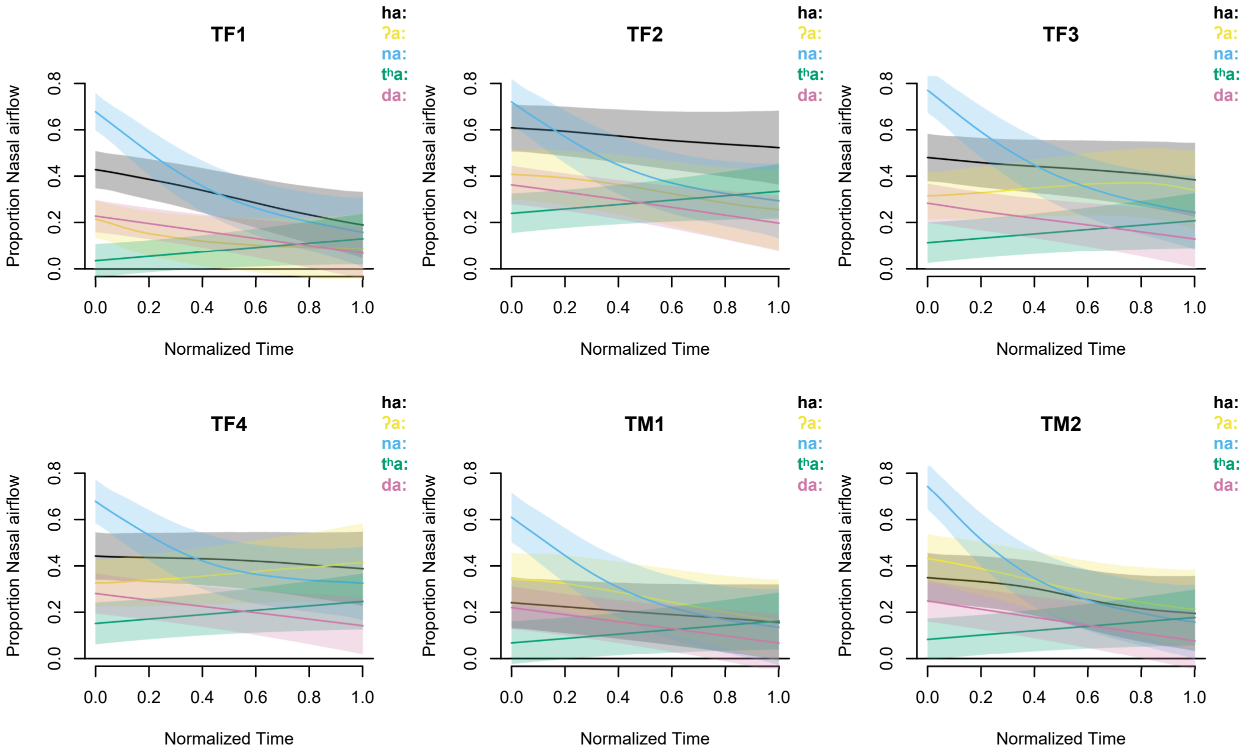


Figure 4.14: GAMM-predicted proportional nasal airflow during vowel /a/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined.

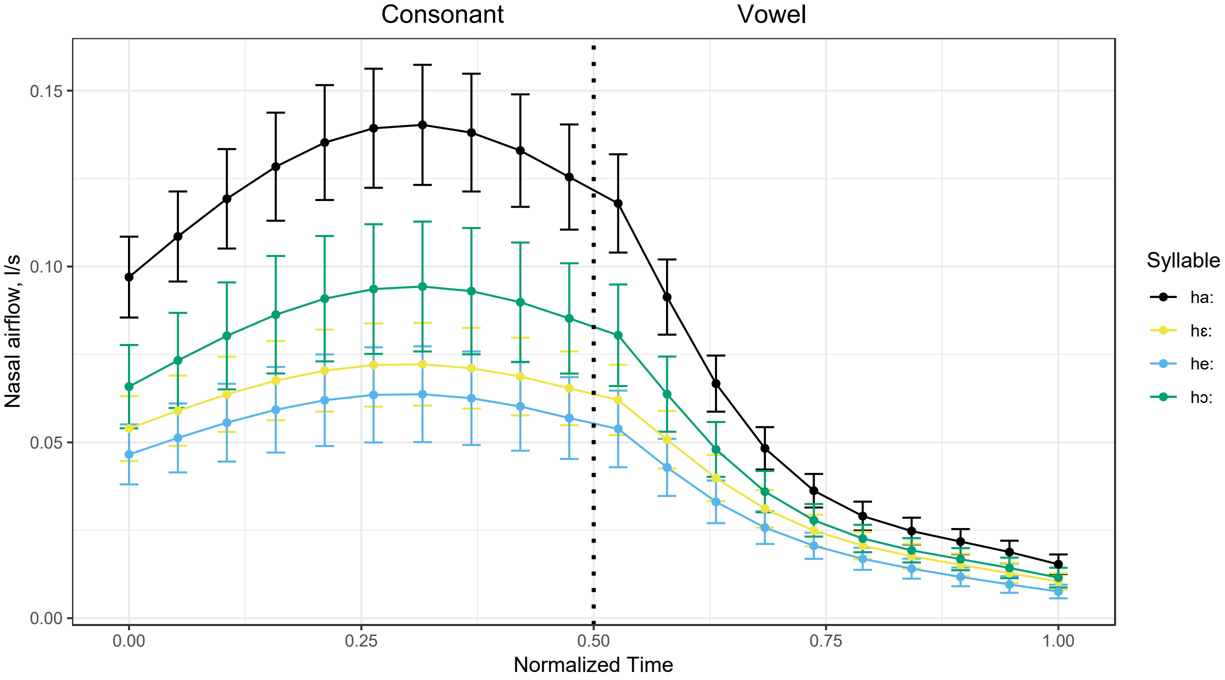


Figure 4.15: Mean nasal airflow of all syllables with the onset consonant /h/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined.

the rest of the speakers are predicted to produce /haː/ with a similar degree of proportional nasal airflow as other speakers.

We now turn to all syllables with the onset consonant /h/ in Figure 4.15. These syllables include /haː/, /hɛː/, /hɔː/, and /haː/, the initial three of which are predicted to nasalize in Thai (Matisoff, 1975). The onset consonant of /haː/ is produced with greater nasal airflow than all other syllables beginning with /h/. While the consonant of /hɔː/ is produced with greater mean nasal airflow than /hɛː/ and /hɔː/, the confidence intervals of all three overlap. Note that although /heː/ is not predicted to nasalize based on reports by Matisoff (1975), the consonant of /heː/ is actually produced with a similar degree of nasal airflow as /n/ in Figure 4.6.

All vowels after /h/ begin with elevated nasal airflow of 0.05 l/s or greater, values that are similar to the vowel of /naː/ or greater. The vowel of /haː/ is produced with greater nasal airflow than all other syllables until about 25% into the duration of the vowel where the confidence intervals of all vowels converge. By the end of the vowel all nasal airflow

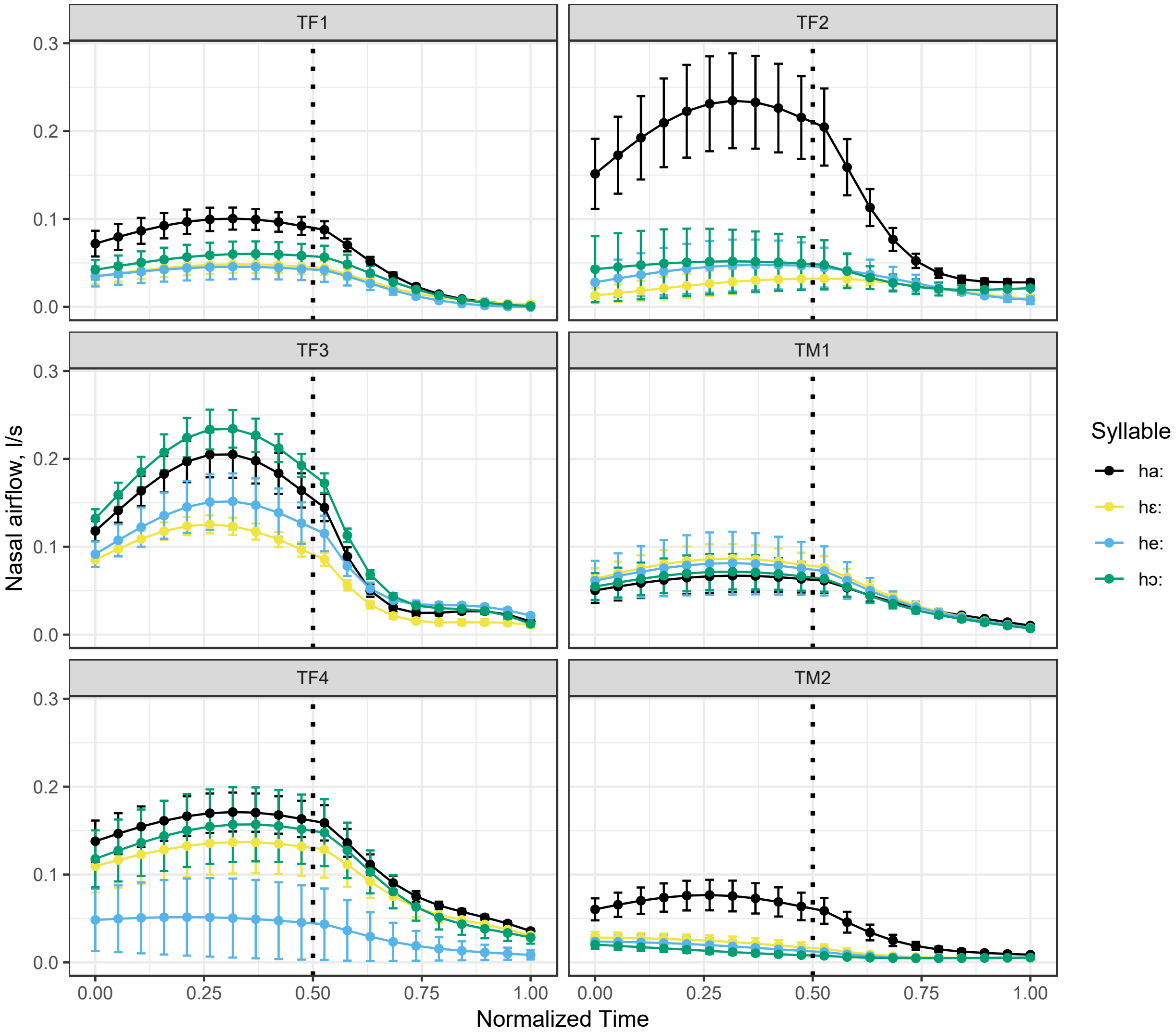


Figure 4.16: Mean nasal airflow of all syllables with the onset consonant /h/. Confidence interval bars are based on the standard error.

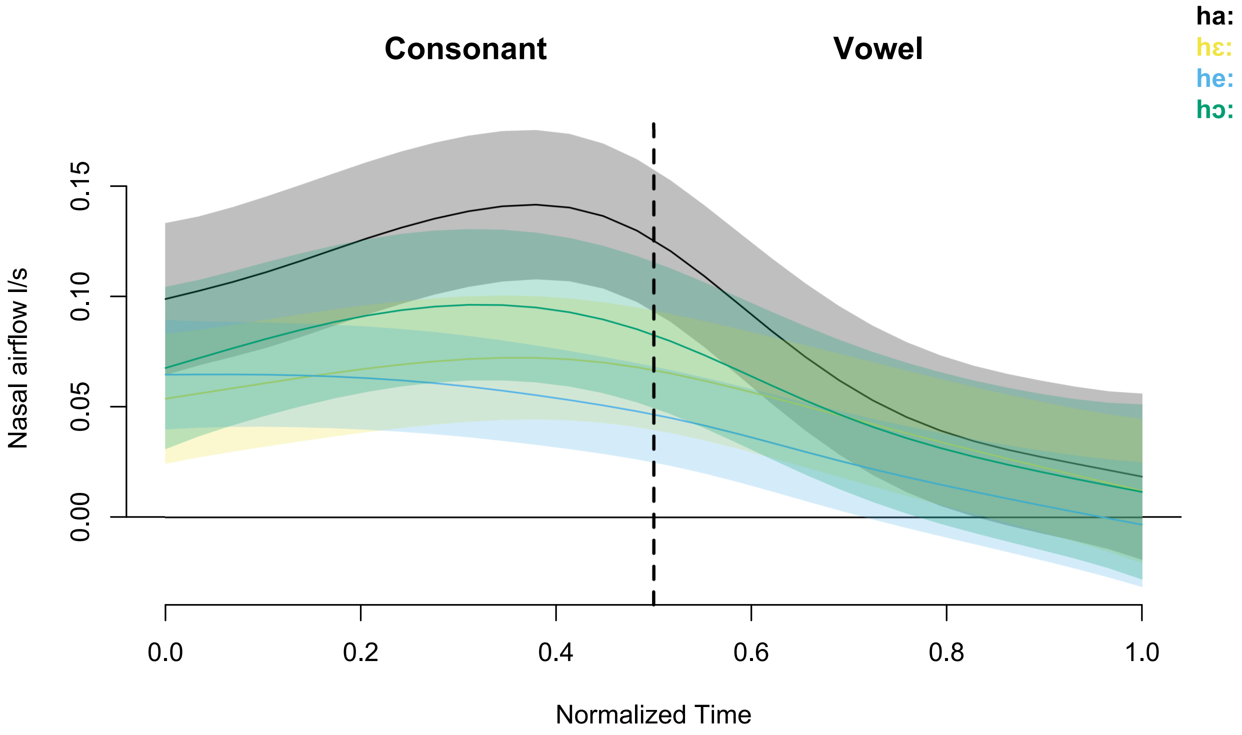


Figure 4.17: GAMM-predicted nasal airflow of all syllables with the onset consonant /h/, liters/second. All speakers' data are combined.

values overlap with non-nasal syllables such as /tʰaː/ and /daː/ (Figure 4.6). Figure 4.16 shows some individual variability in nasal airflow. Speakers TF1, TF2, and TM2 produce /haː/ with the greatest nasal airflow during the consonant and at the start of the following vowel. For the other three speakers, the nasal airflow of /haː/ overlaps with other syllables beginning with /h/. For TF4, the nasal airflow of /heː/ is less than all other syllables during the consonant and first half of the vowel.

The GAMM predictions of nasal airflow for syllables beginning with /h/ (Figure 4.17) show a more conservative prediction compared to the observed means and confidence intervals in Figure 4.16. We observe that the mean nasal airflow of the consonant of /haː/ is predicted to be greater than other syllables beginning with /h/ except for syllable /hɔː/. The predicted confidence intervals of nearly all syllables overlap. During the following vowel, predicted nasal airflow overlaps across all syllables.

We observe some individual speaker variation in the GAMM model predictions in Figure 4.18. While TF2 is predicted to produce the consonant and initial portion of the vowel of /haː/ with greater nasal airflow than other syllables, TF3 is predicted to produce /haː/ and /hɔː/ with greater nasal airflow during the consonant, and TF4 is predicted to produce /heː/ with the lowest nasal airflow during the consonant and vowel.

Syllables /hɛː/ and /heː/ are produced with greater mean oral airflow during the consonant and initial 20% of the vowel (Figure 4.19). An analysis of proportional nasal airflow during the vowel reveals that syllables /haː/ and /hɔː/ are produced with greater proportional nasal airflow than /hɛː/ and /heː/ (Figure 4.20). Plots of individual speaker variation in nasal proportion (Figure 4.21) show that all speakers except TM1 and TF3 produce /haː/ with greater nasal airflow proportion at the start of the vowel, while TF1, TF3, and TF4 vary in whether they also produce /hɛː/, /hɔː/, and /heː/ with high nasal airflow as well.

Finally, the GAMM-predicted proportional nasal airflow of syllables beginning with /h/ are shown in Figure 4.22. Predicted values are similar to observed mean proportional nasal airflow in that the airflow is predicted to be greater during the initial portion of the vowel for /haː/ and /hɔː/. Individual plots of predicted proportional nasal airflow show that while TF1 and TF2 are predicted to produce /haː/ with greater proportional nasal airflow for the first half of the syllables, all other speakers show complete overlap in confidence intervals (Figure 4.23). TF1 is also predicted to produce /hɔː/ with increased proportional nasal airflow during the first 10% of the vowel.

Figure 4.24 show mean nasal airflow during all syllables beginning with /ʔ/. These syllables include /ʔaː/, /ʔɛː/, /ʔɔː/, and /ʔɛː/. According to Matisoff's (1975) description of Thai, the first three syllables are expected to nasalize because these possess low/mid-low vowels, although perhaps to a lesser degree than syllables beginning with /h/. /ʔeː/ is not expected to nasalize. The syllable /haː/ is also shown in these plots for reference. We observe that nasal airflow is close to zero during all /ʔ/s. This finding is expected because the glottis is closed during /ʔ/, preventing egressive airflow. During the following vowel,

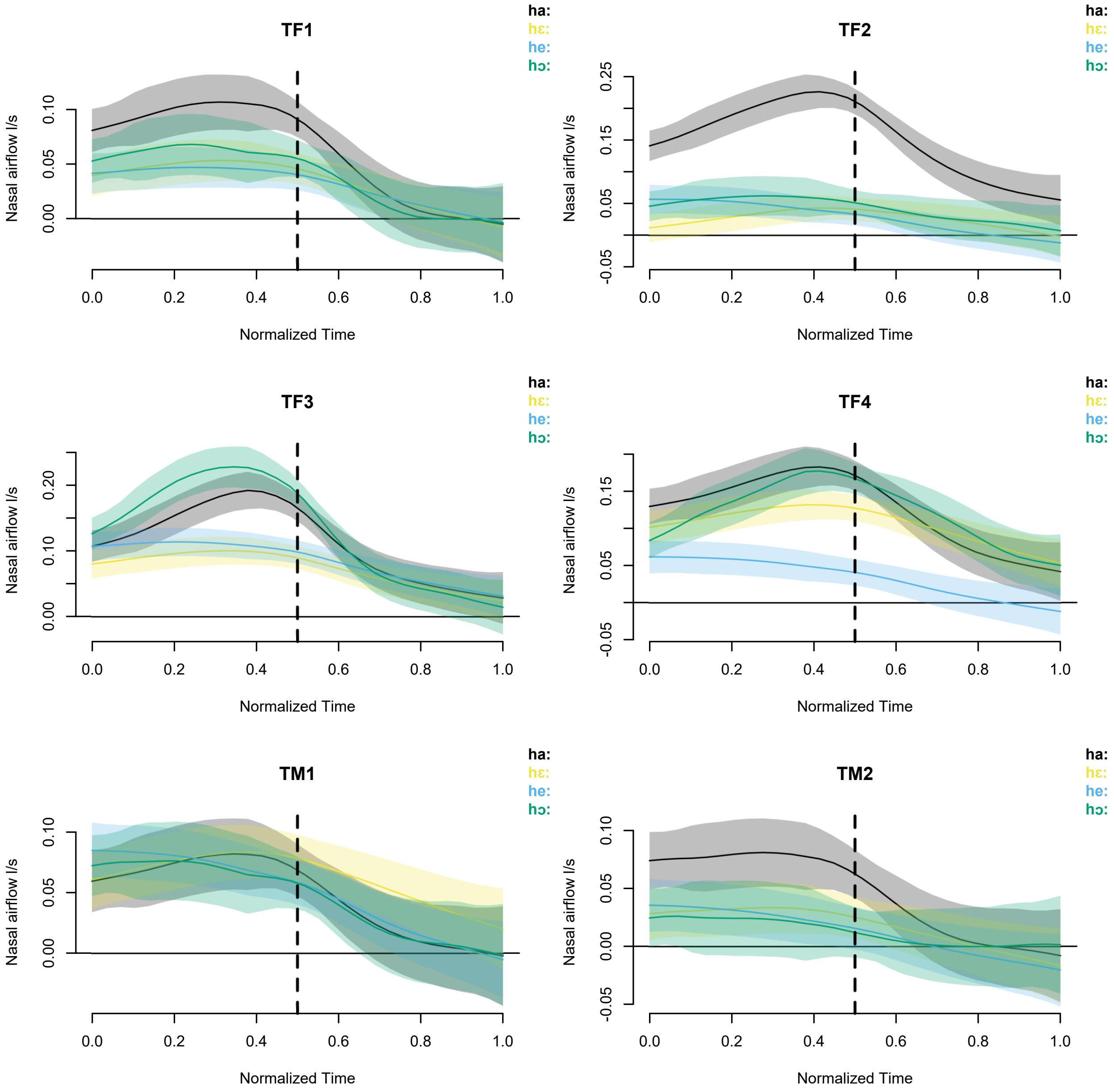


Figure 4.18: GAMM-predicted nasal airflow of all syllables with the onset consonant /h/, liters/second.

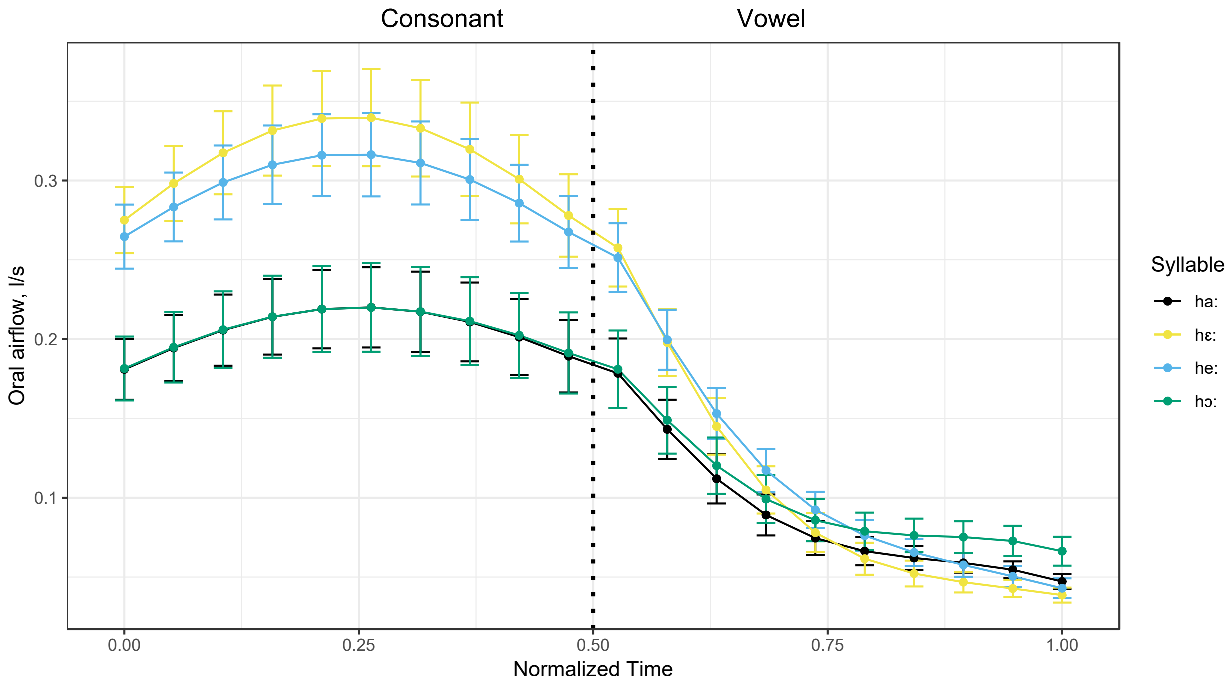


Figure 4.19: Mean oral airflow of all syllables with the onset consonant /h/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined.

nasal airflow remains low for all syllables beginning with /ʔ/. While the mean nasal airflow of /ʔaː/ rises somewhat during the vowel, its confidence interval usually overlaps with that of other syllables beginning with /ʔ/.

Figure 4.25 reveals that individual speakers produce nasal airflow similar to the combined graph (Figure 4.24). Nasal airflow is close to zero during /ʔ/ closure and also low during the following vowel. Only TF4 produces the vowel of /ʔaː/ with slightly elevated nasal airflow. Both combined (Figure 4.26) and individual speaker plots (Figure 4.27) of GAMM-predicted nasal airflow during syllables beginning with /ʔ/ are similar to observations of mean nasal airflow in Figures 4.25 and 4.24. The GAMM does not predict elevated nasal airflow for the vowel of /ʔaː/ for any speaker.

Meanwhile, oral airflow during the consonant of /ʔ/ is similarly low (Figure 4.29). During the following vowel oral airflow starts low and gradually rises to be either the same as, or in the case of /ʔɔː/, slightly higher than during the vowel of /haː/. A proportional analysis of nasal airflow shows that vowels after /ʔ/ exhibit similar proportional nasal airflow that is less

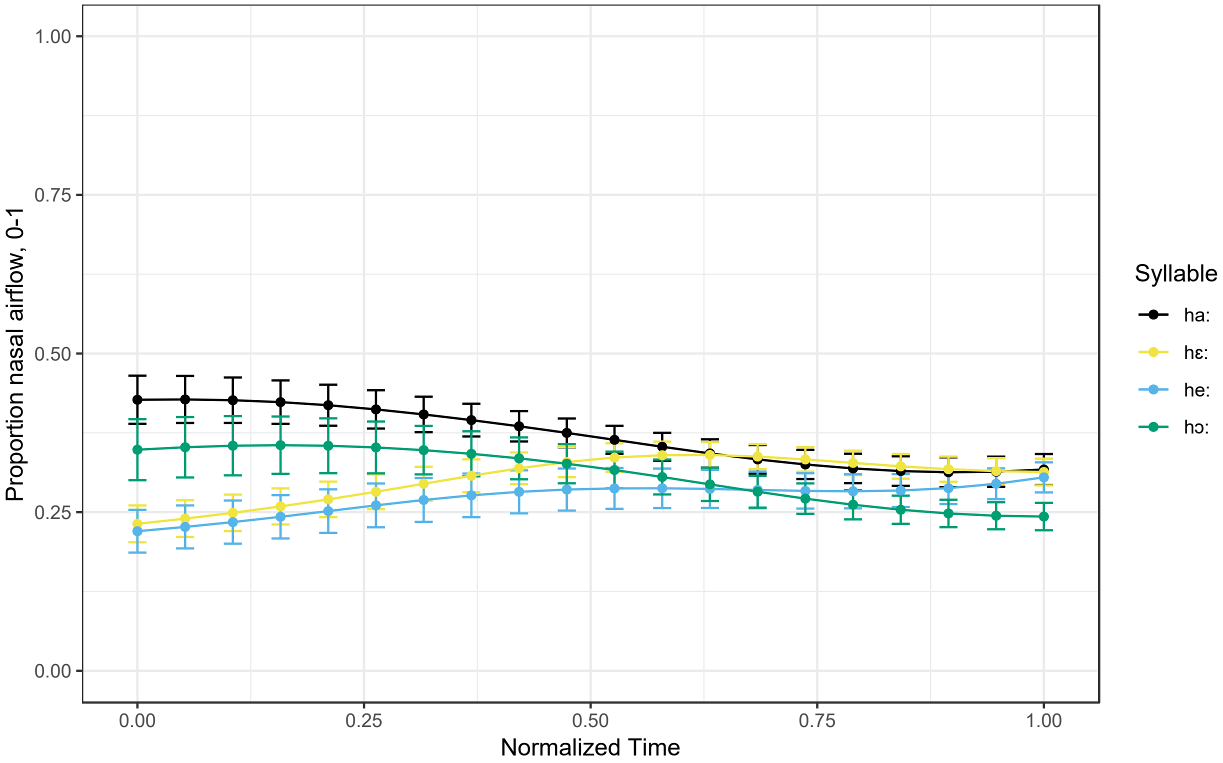


Figure 4.20: Mean proportional nasal airflow during vowels after /h/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral)Confidence interval bars are based on the standard error. All speakers' data are combined

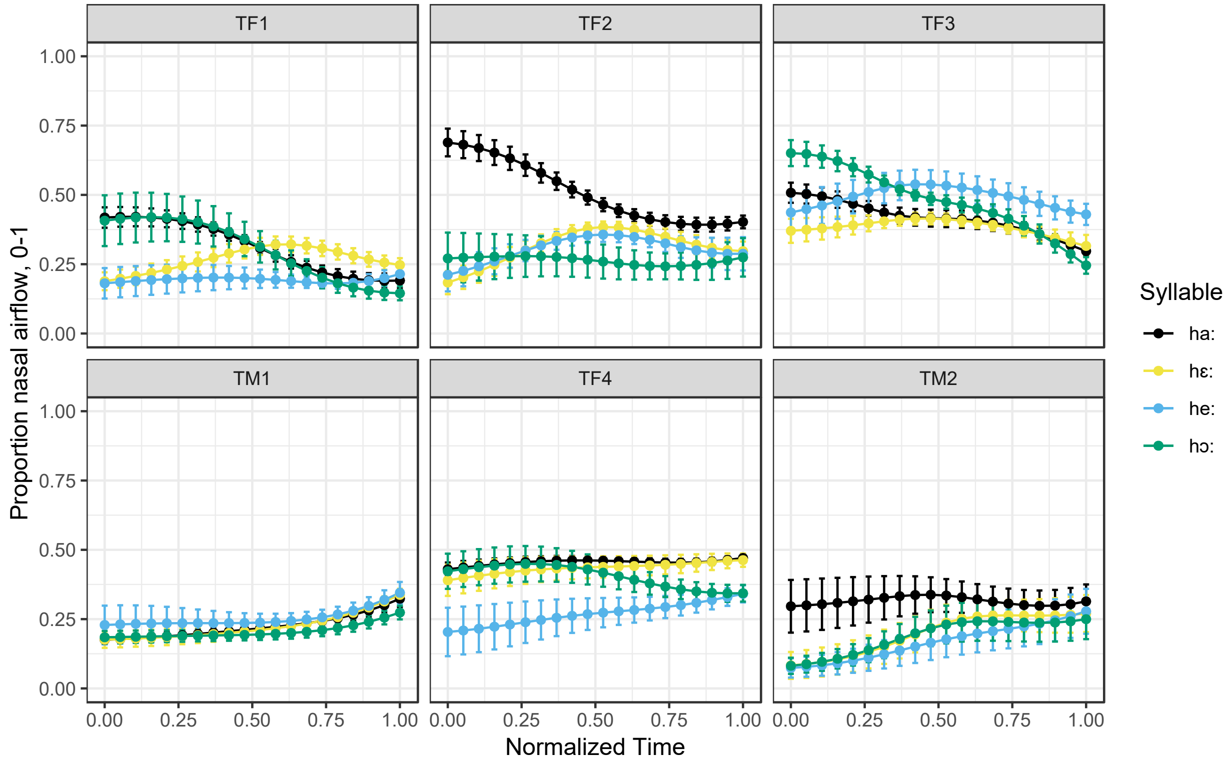


Figure 4.21: Mean proportional nasal airflow during vowels after /h/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral)Confidence interval bars are based on the standard error.

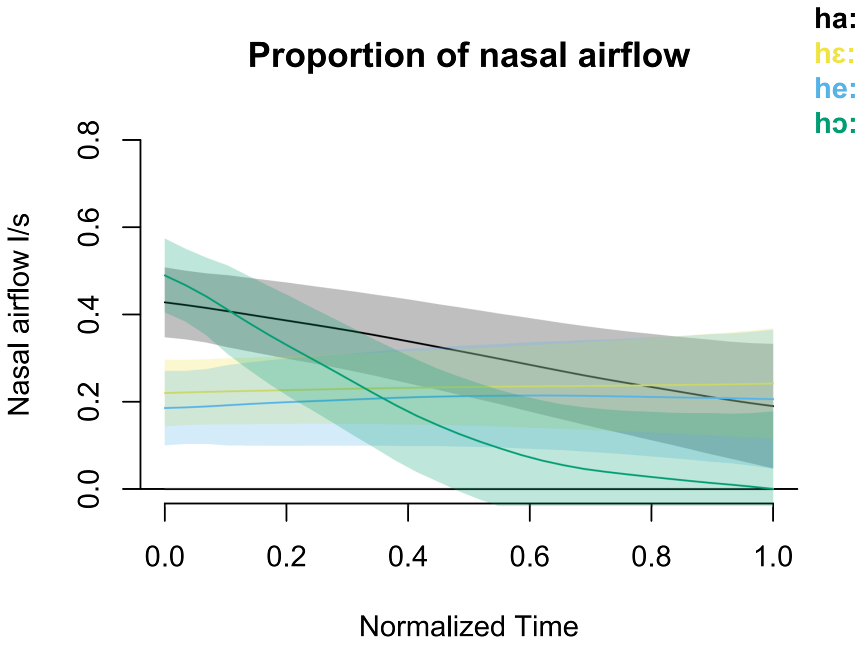


Figure 4.22: GAMM-predicted proportional nasal airflow during vowels after /h/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined.

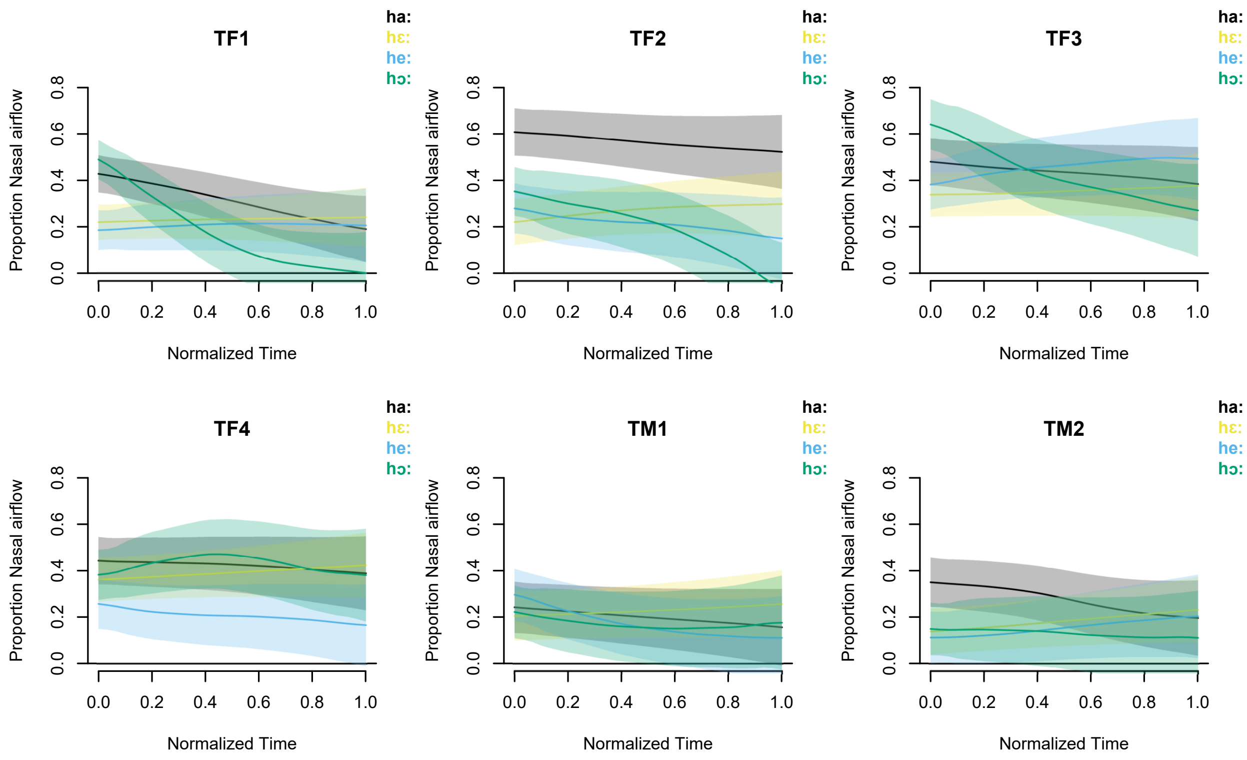


Figure 4.23: GAMM-predicted proportional nasal airflow during vowels after /h/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error.

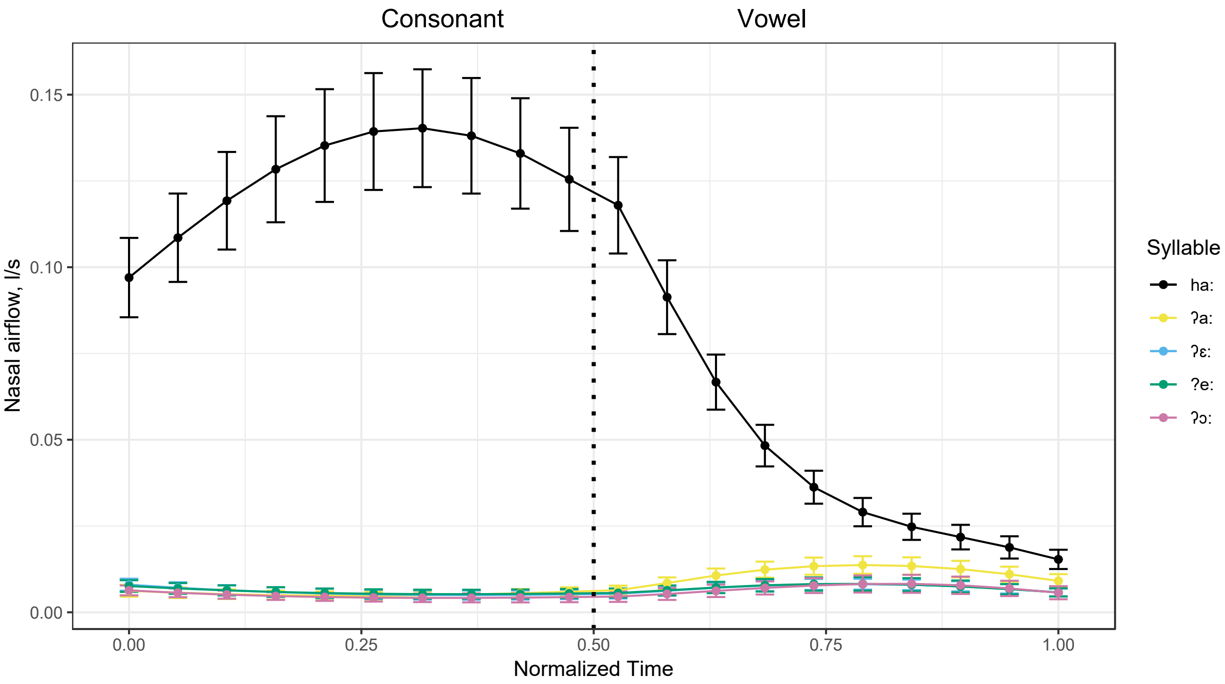


Figure 4.24: Mean nasal airflow of all syllables with the onset consonant /ʔ/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined.

than /haː/ but similar to /daː/ and /t^haː/ (Figure 4.30). The GAMM model of predicted proportional nasal airflow shows a similar pattern (Figure 4.32).

A plot of mean nasal airflow during syllables beginning with /t^h/ shows a similar trend as with syllables beginning with /ʔ/ (Figure 4.33). Nasal airflow is low or close to zero during both the onset consonant and following vowel. The GAMM predicts nasal airflow during syllables with /t^h/ that closely resembles mean nasal airflow (Figure 4.34). Visualization of individual speakers' mean nasal airflow (Figure 4.35) and GAMM-predicted nasal airflow (Figure 4.34) show similar low nasal airflow during both the consonant and vowel.

An analysis of oral airflow for syllables beginning with /t^h/ shows that oral airflow is highest during the end of the consonant, 0.1-0.2 l/s higher than /h/, for all productions of /t^h/. Oral airflow is similarly high at the start of the following vowel and then sharply falls to be similar to /h/ about halfway through the vowel. This high oral airflow is likely associated with aspiration.

A plot of mean proportional nasal airflow during the vowel of syllables beginning with

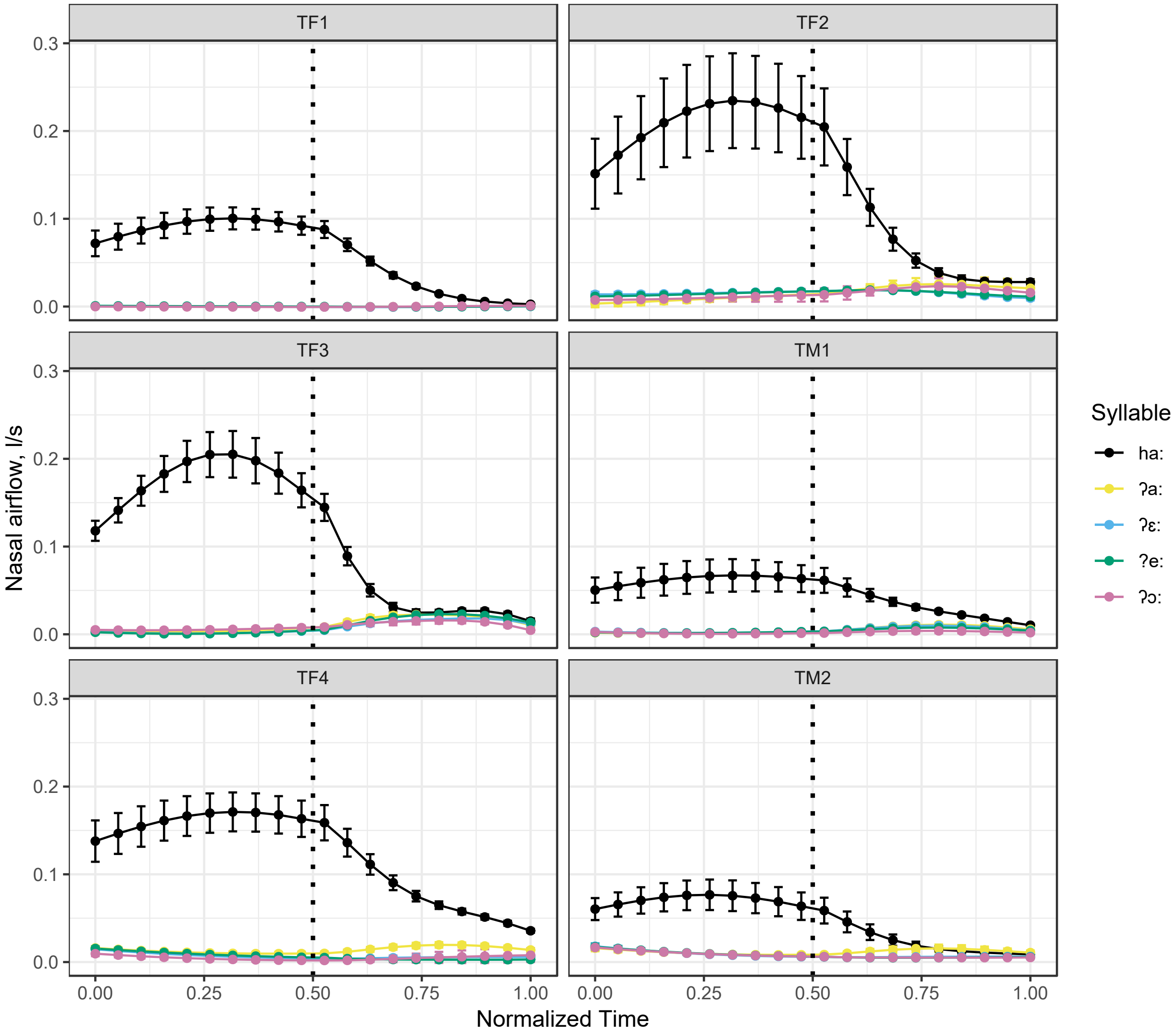


Figure 4.25: Mean nasal airflow of all syllables with the onset consonant /ʔ/. Confidence interval bars are based on the standard error.

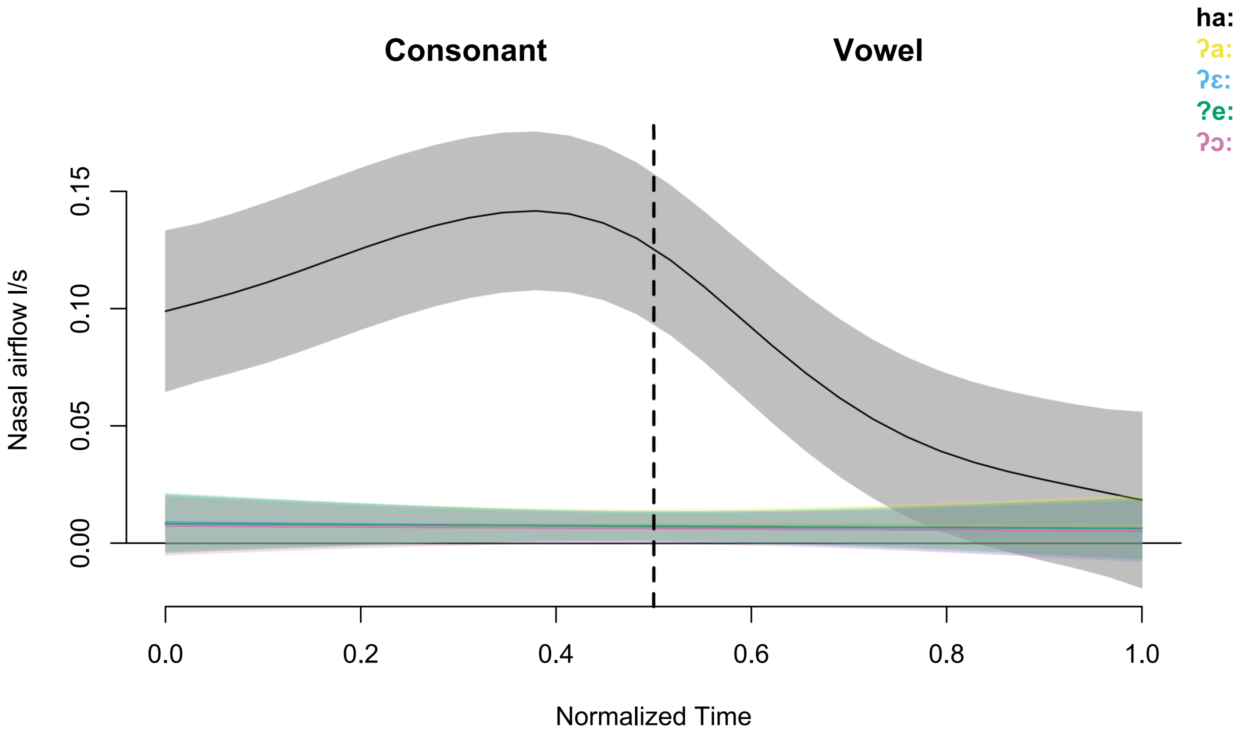


Figure 4.26: GAMM-predicted nasal airflow of all syllables with the onset consonant /ʔ/, liters/second. All speakers' data are combined.

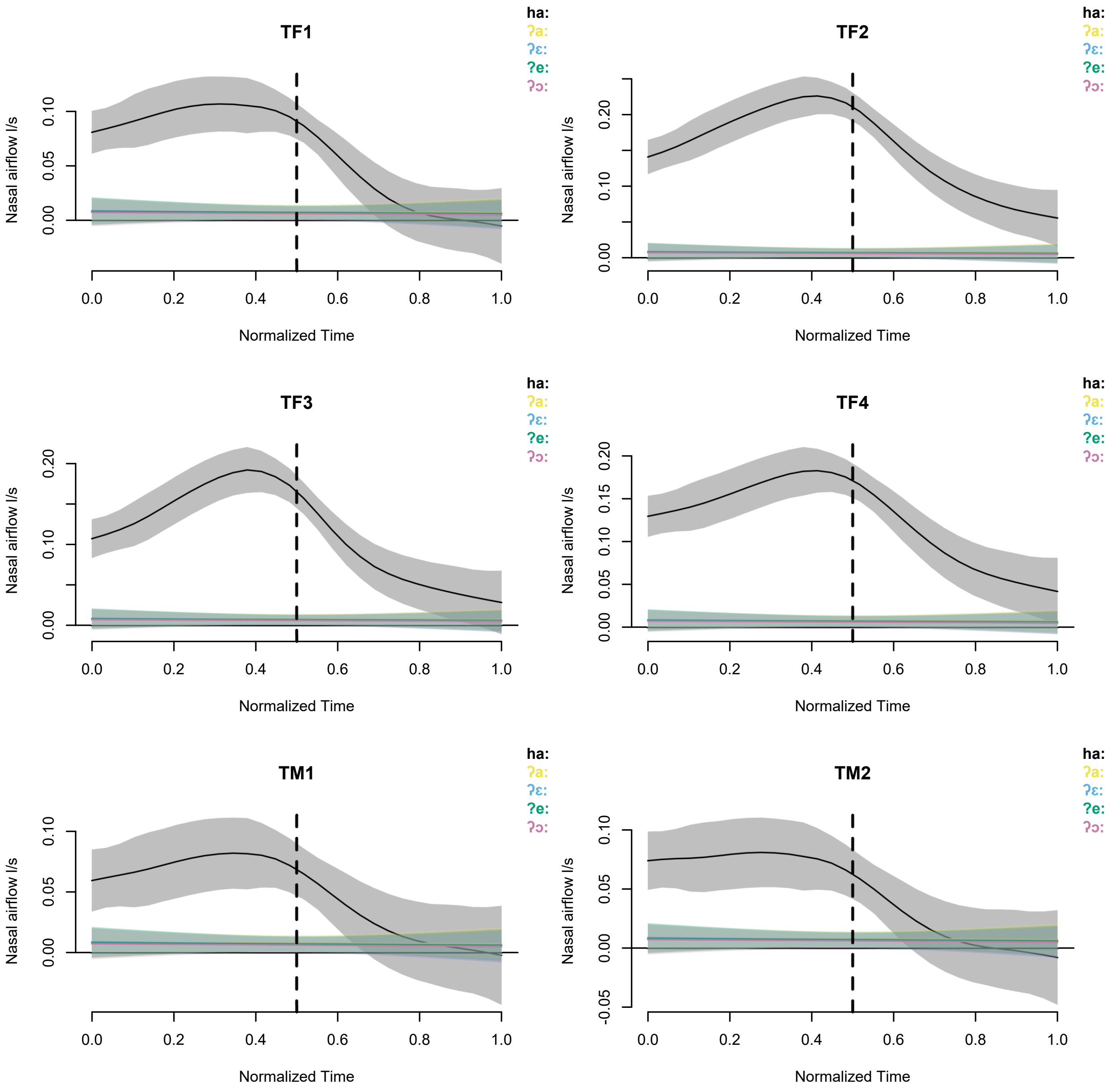


Figure 4.27: GAMM-predicted nasal airflow of all syllables with the onset consonant /ʔ/, liters/second. /haː/ is included for comparison.

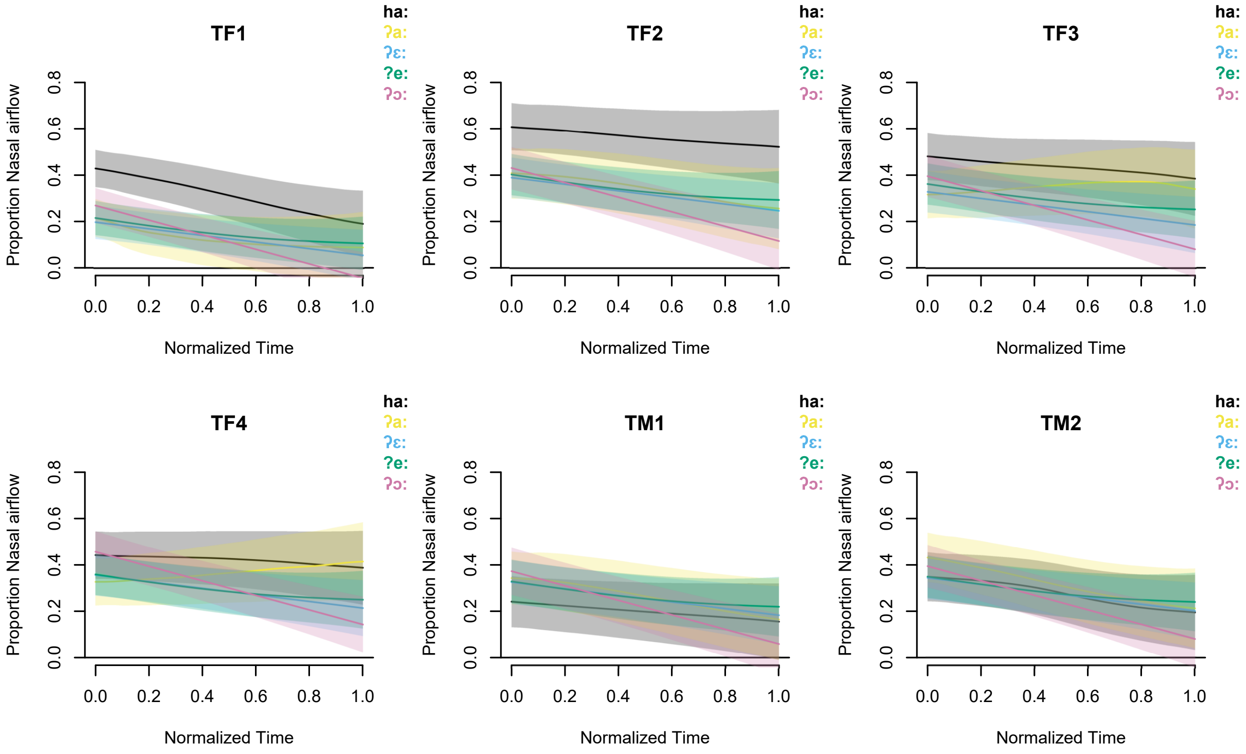


Figure 4.28: GAMM-predicted proportional nasal airflow during vowels after /ʔ/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined

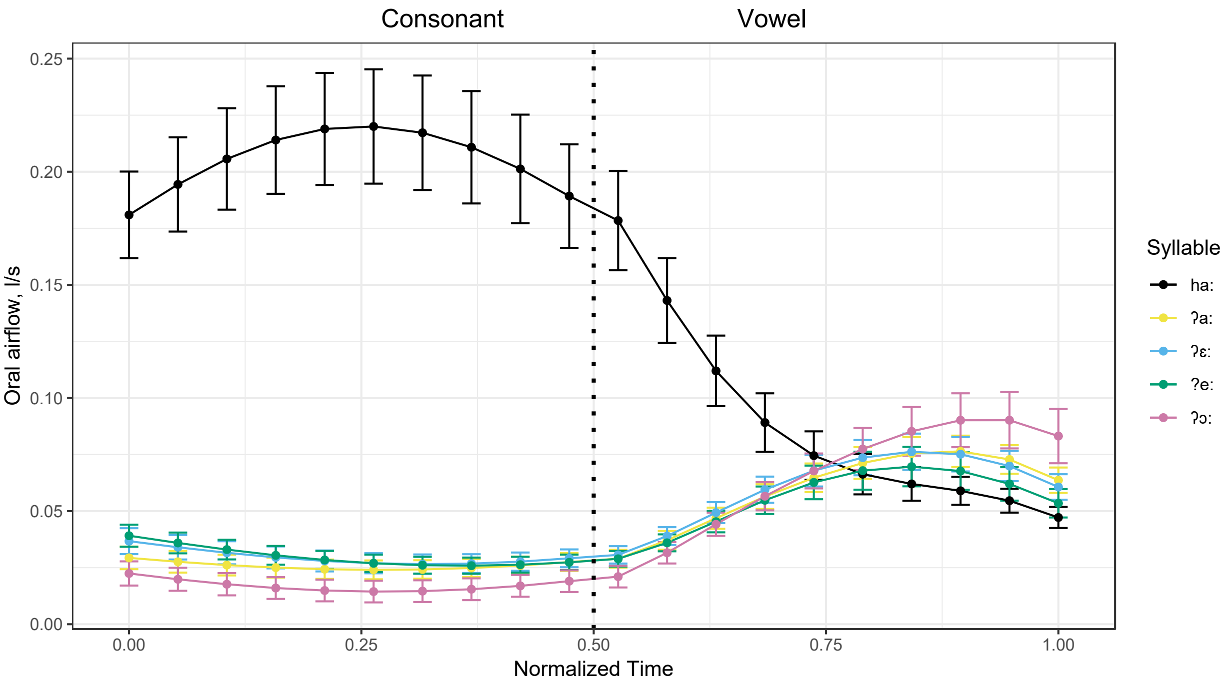


Figure 4.29: Mean oral airflow of all syllables with the onset consonant /ʔ/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined

/t^h/ can be found in Figure 4.38. We observe that proportional nasal airflow is low during the start of these vowels, lower than any other syllable. The proportional nasal airflow gradually rises towards the middle of the syllable, but remains lower than during /haː/ and /naː/. GAMM-predicted proportional nasal airflow for these vowels can be found in Figure 4.40. The predicted values closely resemble the displays of mean proportional nasal airflow.

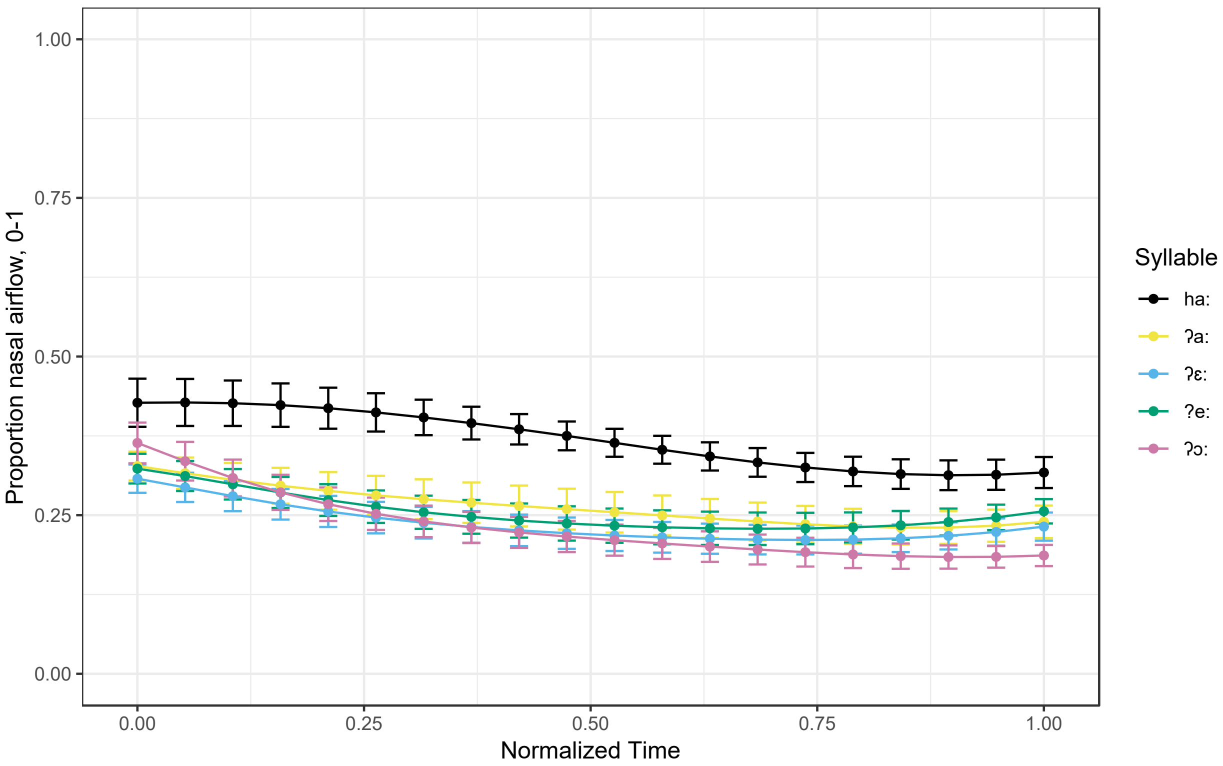


Figure 4.30: Mean proportional nasal airflow during vowels after /ʔ/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral)Confidence interval bars are based on the standard error. All speakers' data are combined

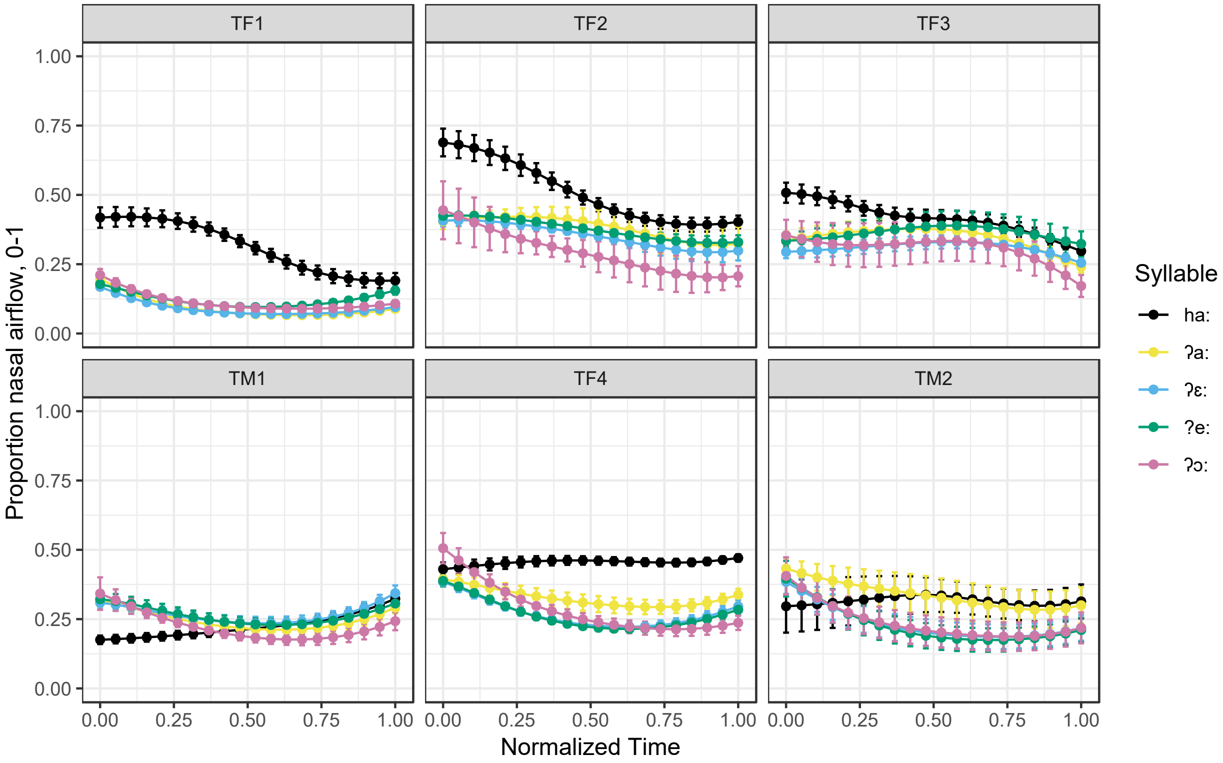


Figure 4.31: Mean proportional nasal airflow during vowels after /ʔ/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral)Confidence interval bars are based on the standard error.

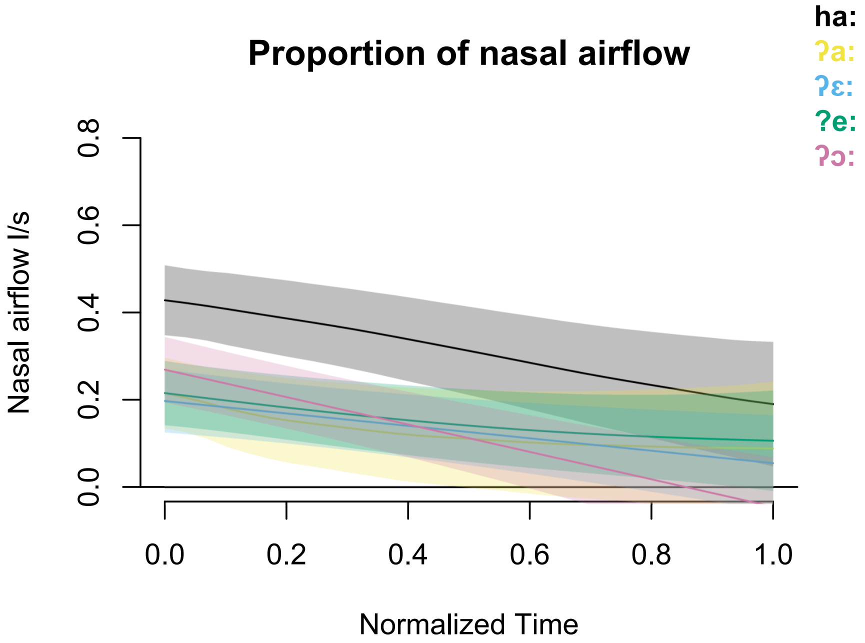


Figure 4.32: GAMM-predicted proportional nasal airflow during vowels after /ʔ/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined

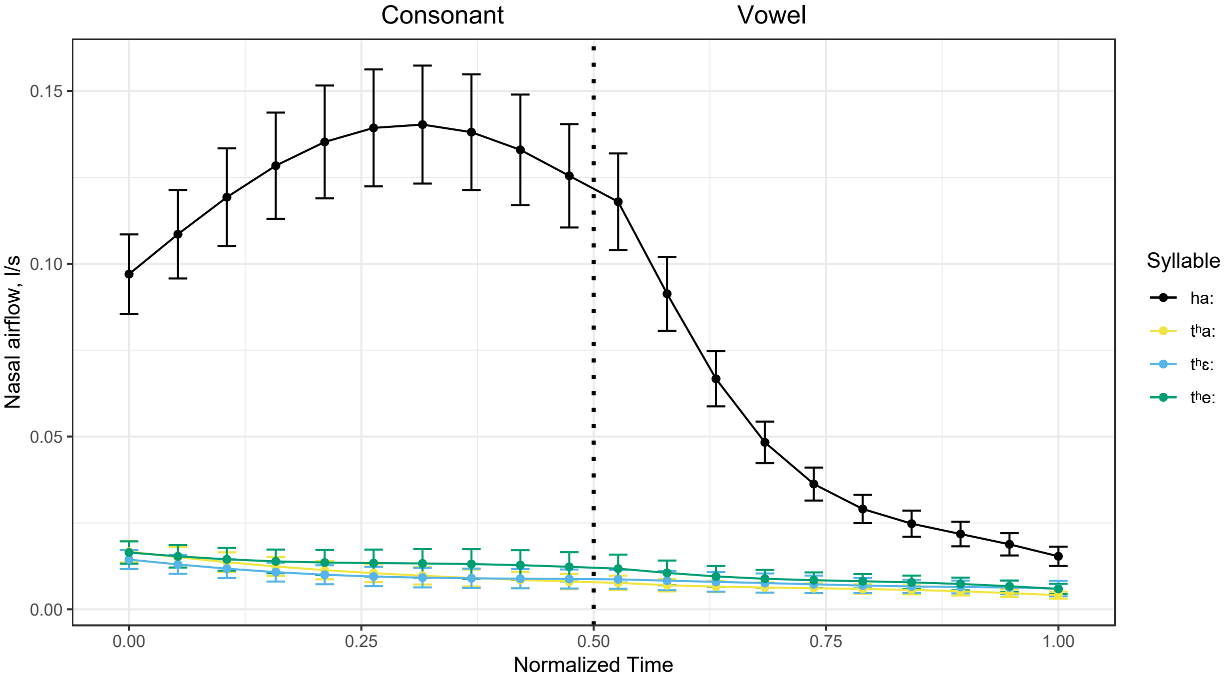


Figure 4.33: Mean nasal airflow of all syllables with the onset consonant /t^h/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined

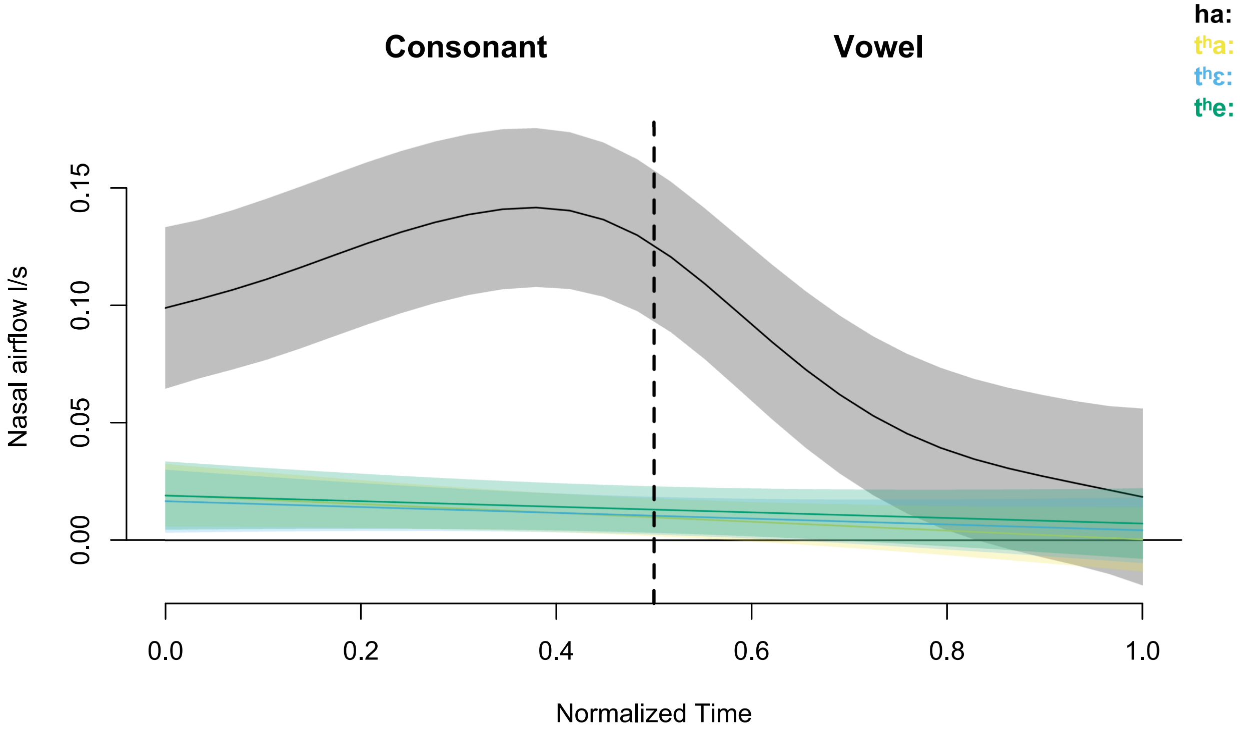


Figure 4.34: GAMM-predicted nasal airflow of all syllables with the onset consonant /t^h/, liters/second. The graph includes /haː/ for comparison. All speakers' data are combined.

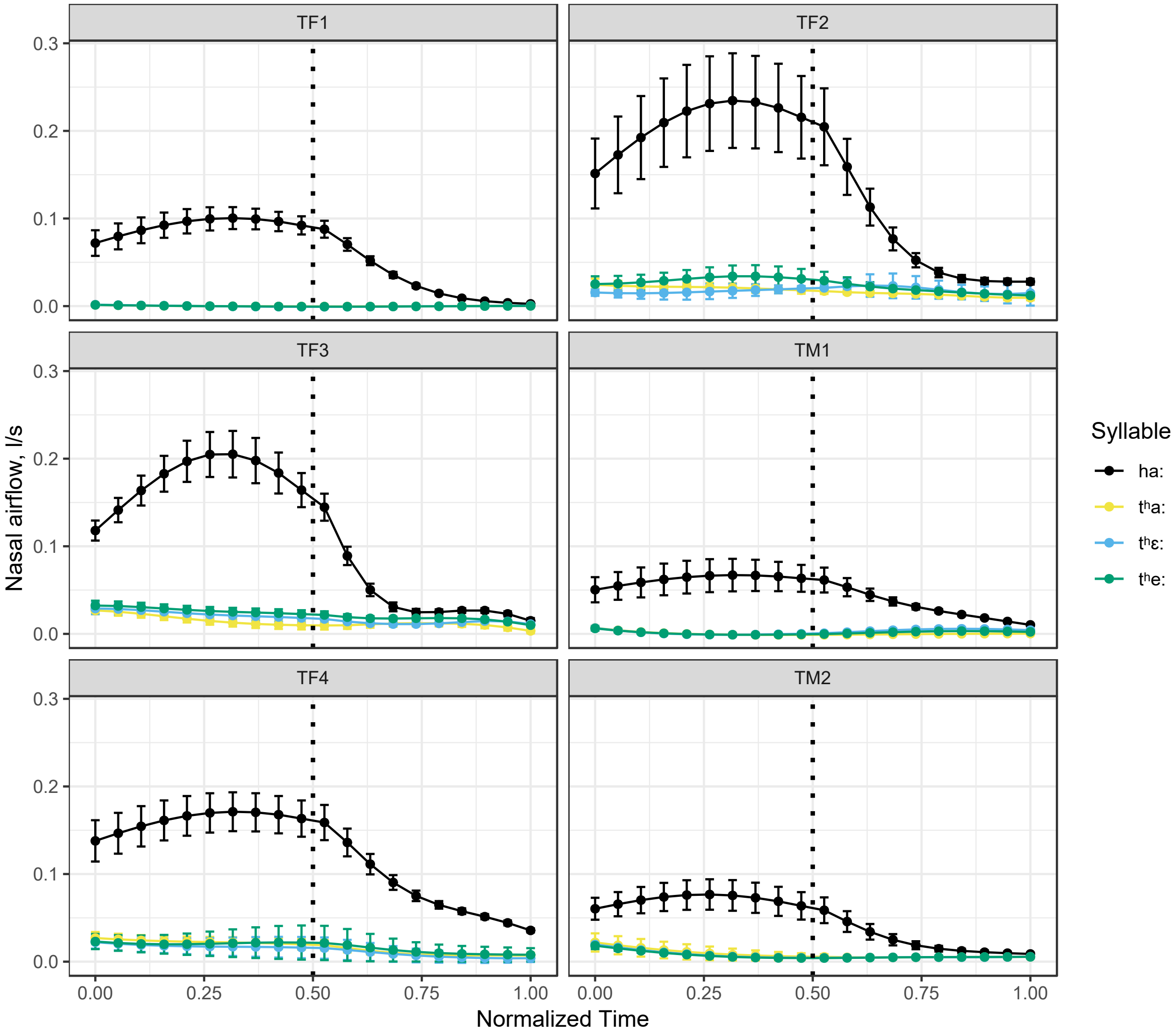


Figure 4.35: Mean nasal airflow of all syllables with the onset consonant /$\mathrm{t^h}$/. Confidence interval bars are based on the standard error.

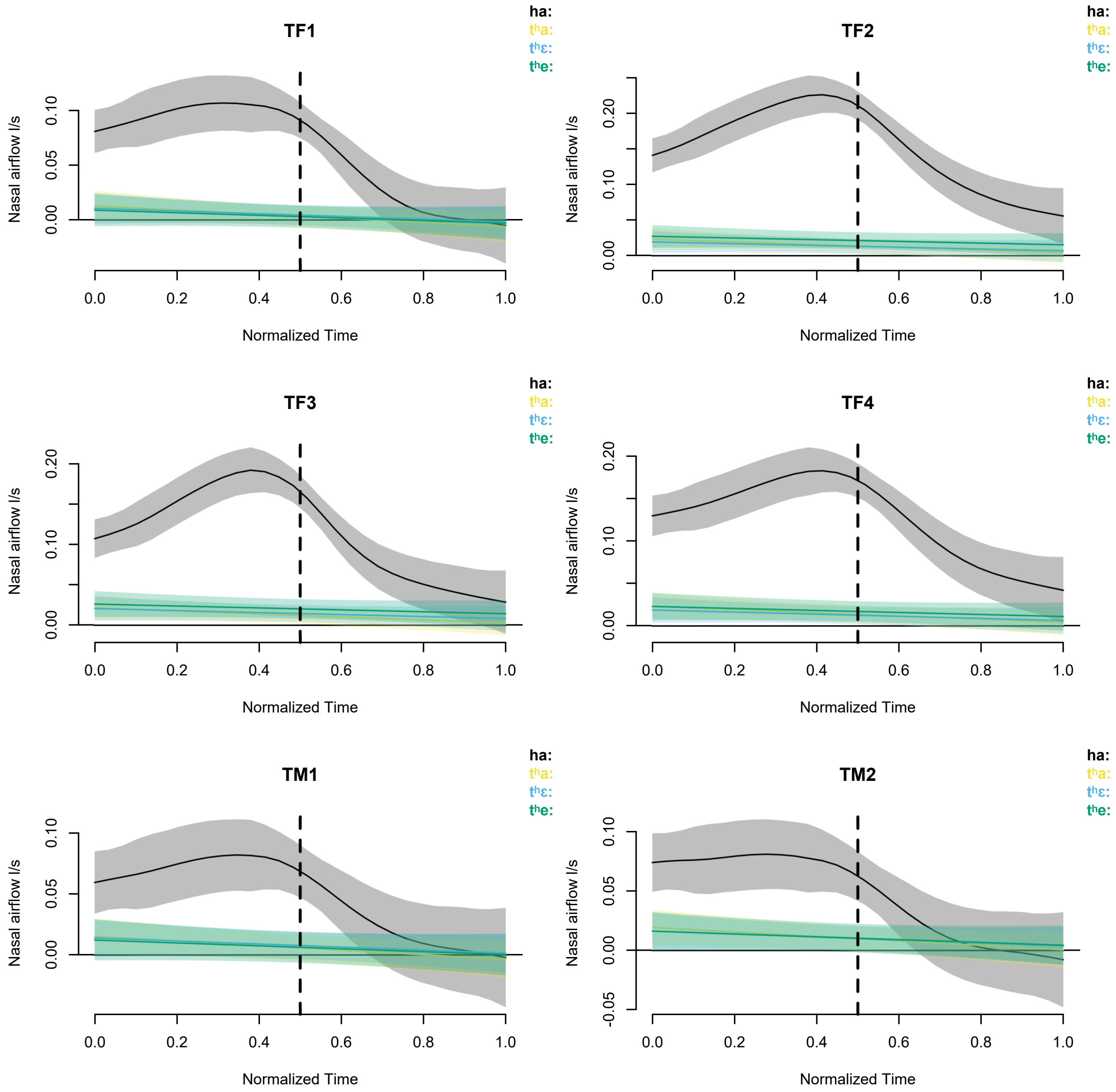


Figure 4.36: GAMM-predicted nasal airflow of all syllables with the onset consonant /tʰ/, liters/second. /haː is included for comparison.

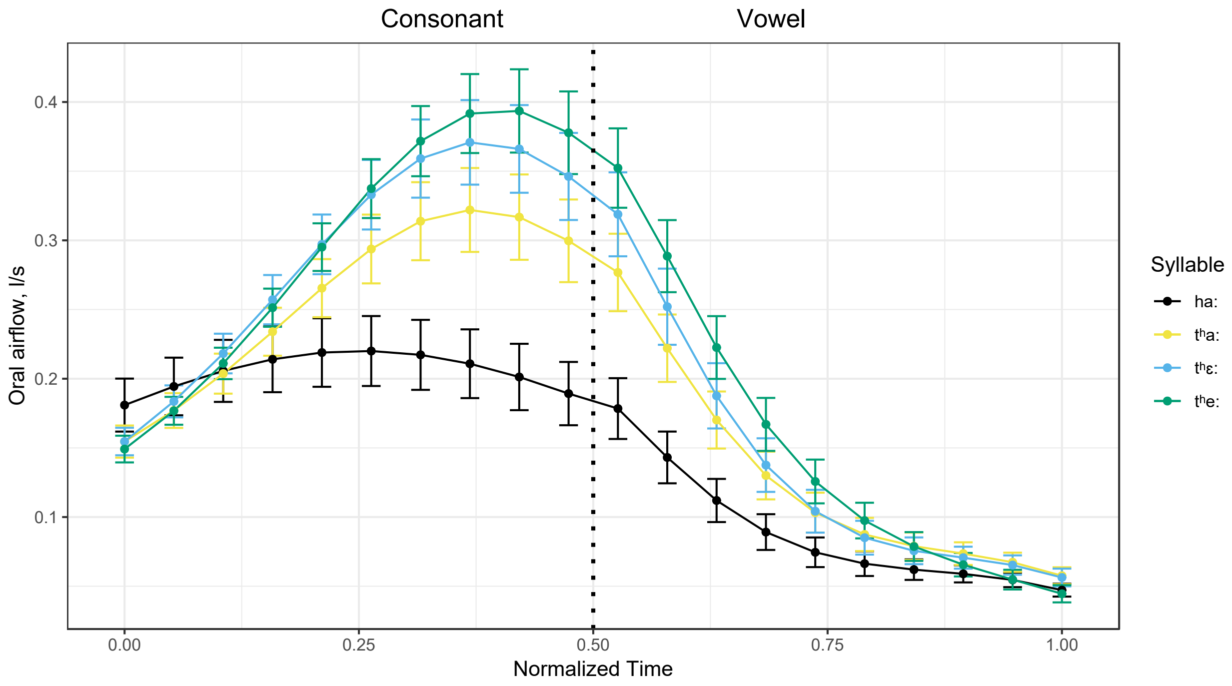


Figure 4.37: Mean oral airflow of all syllables with the onset consonant /t^h/, liters/second. Confidence interval bars are based on the standard error. All speakers' data are combined

4.4.2 Integrated nasal airflow

Nasal airflow was normalized to 20 samples per syllable, 10 during the onset consonant and 10 during the following vowel, so that summed integrated nasal airflow could be compared in the onset and following vowel. The integral was then taken of all samples within the consonant and the vowel separately. This yielded a single measure of cumulative nasal airflow during both the consonant and vowel for each token. The integrated flow during consonant and vowel were compared token-by-token and in each case it was registered which segment (consonant or vowel) had the greater flow value.

The data were analyzed with a linear mixed effects model using the *lmerTest* package in R. Integrated nasal airflow was the dependent variable with fixed effects including syllable, segment affiliation (consonant or vowel) of maximum integrated flow (MaxLoC), and segment type (consonant or vowel). Speaker was included as a random effect.

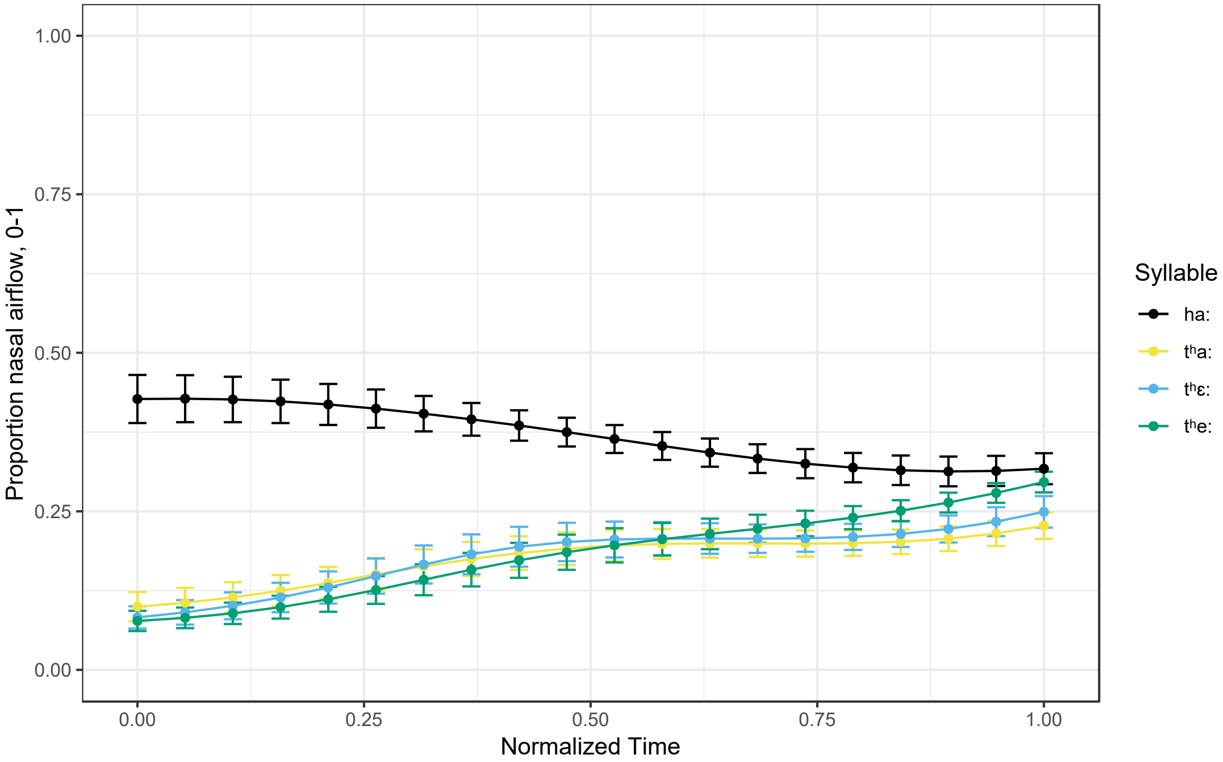


Figure 4.38: Mean proportional nasal airflow during vowels after /t^h/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral)Confidence interval bars are based on the standard error. All speakers' data are combined.

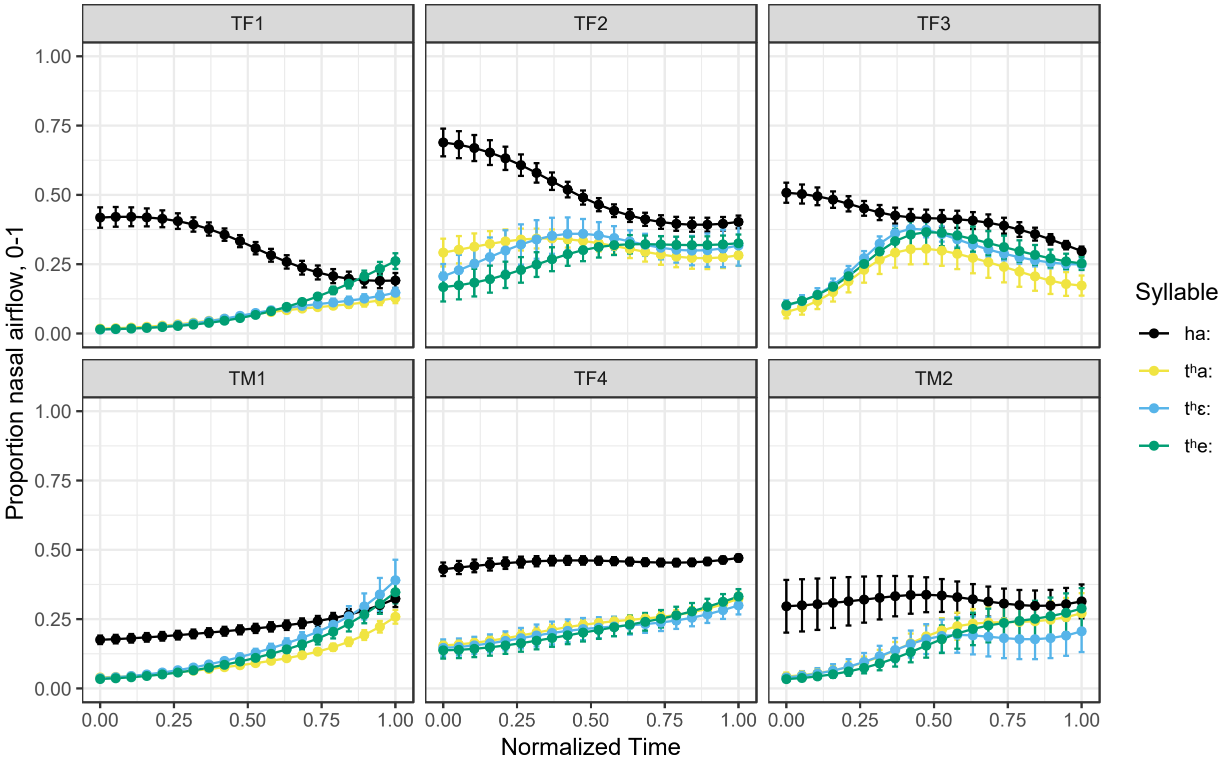


Figure 4.39: Mean proportional nasal airflow during vowels after /t^h/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral)Confidence interval bars are based on the standard error.

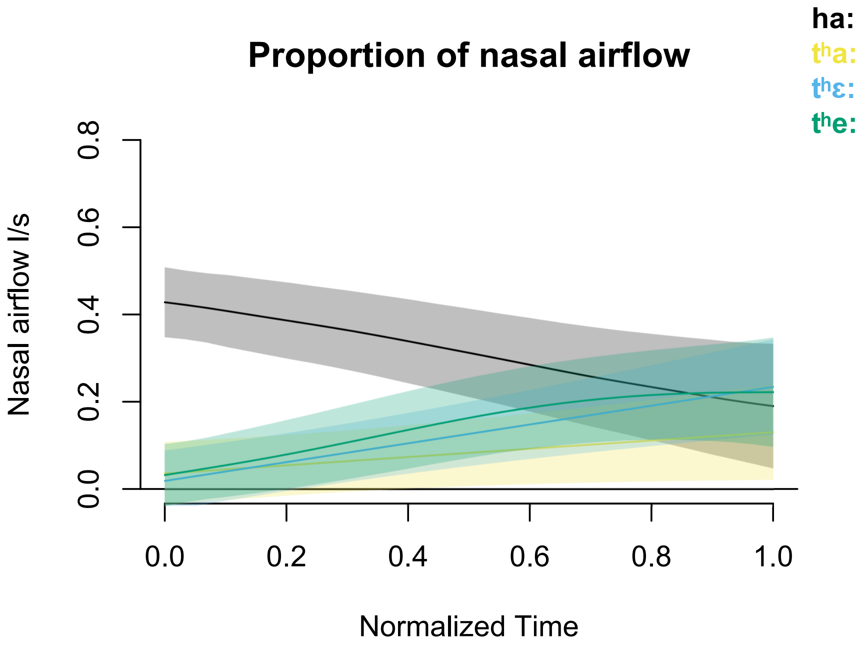


Figure 4.40: GAMM-predicted proportional nasal airflow during vowels after /t^h/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined

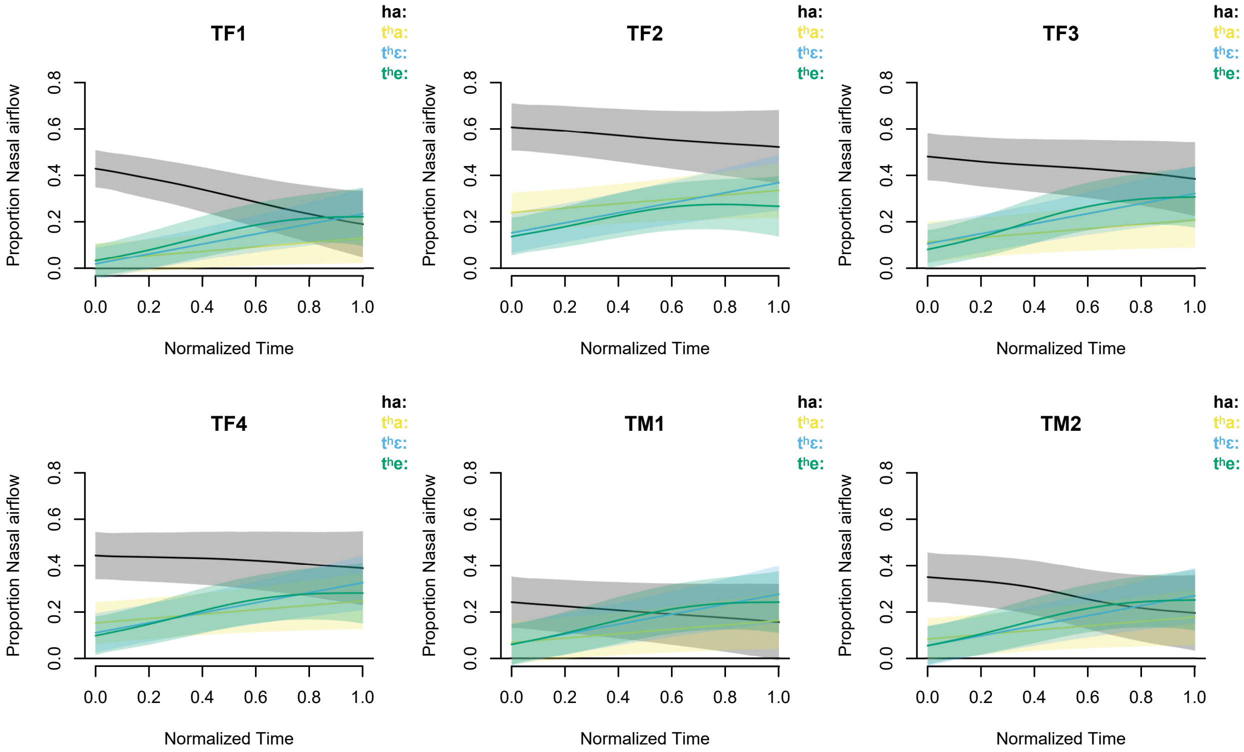
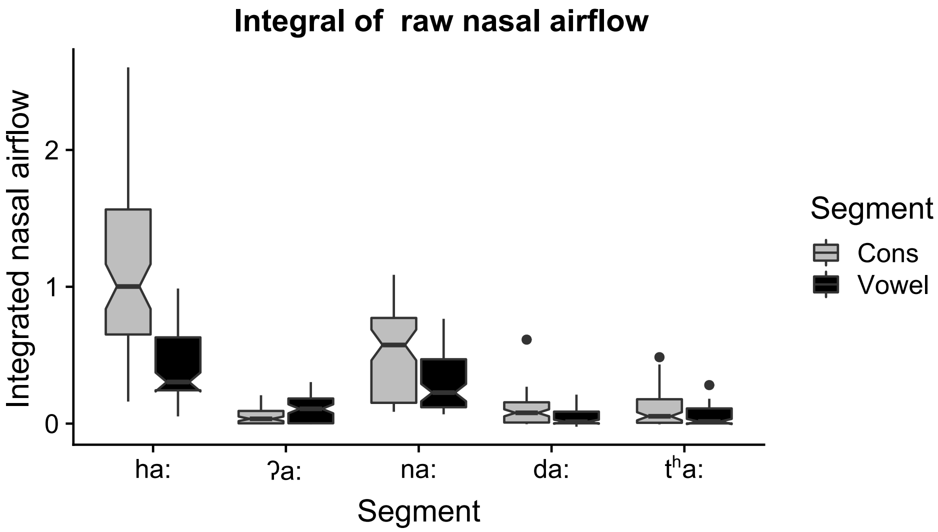


Figure 4.41: GAMM-predicted proportional nasal airflow during vowels after /t^{h}/. The proportion of nasal airflow was calculated by dividing the nasal airflow by the sum of oral and nasal airflow (nasal / nasal + oral). Confidence interval bars are based on the standard error. All speakers' data are combined

Airflow maximum analysis

Figure 4.42: Integrated nasal airflow for syllables with /a/.



Results of a linear mixed effects model (Table 4.4) show significant effects between the location of maximum nasal airflow (consonant or vowel) and nasal airflow. Furthermore, whether the segment is a consonant or a vowel also effects nasal airflow. Finally, the integrated nasal airflow of /haː/ is significantly different, i.e. greater, than any other syllable.

Figure 4.42 shows boxplots of integrated nasal airflow for all syllables with the vowel /aː/. Two post-hoc analyses of the lme model were performed by calculating estimated marginal means with a Tukey multiplicity adjustment (Lenth, 2018). The first test was performed on integrated nasal airflow by Syllable contrast (Tables 4.6 and 4.7) while the second test was performed on integrated nasal airflow by Segment type (consonant or vowel) within the same syllables (Table 4.5). These tests revealed many differences in integrated nasal airflow among segments. The consonant /h/ has the greatest integrated nasal airflow of any other syllable (Table 4.6). /n/ exhibits less integrated nasal airflow than the consonant of /haː/, but similar nasal airflow to other syllables with /h/. Both /h/ and /n/ are produced with significantly more integrated nasal airflow than during the following vowel (Table 4.5). All

	Model 1
(Intercept)	$1.18(0.07)^{***}$
Syllable /hɛː/	$-0.54(0.04)^{***}$
Syllable /heː/	$-0.63(0.04)^{***}$
Syllable /hɔː/	$-0.38(0.04)^{***}$
Syllable /ʔaː/	$-1.05(0.04)^{***}$
Syllable /ʔɛː/	$-1.08(0.04)^{***}$
Syllable /ʔeː/	$-1.08(0.04)^{***}$
Syllable /ʔɔː/	$-1.08(0.04)^{***}$
Syllable /naː/	$-0.65(0.04)^{***}$
Syllable /daː/	$-1.07(0.04)^{***}$
Syllable /tʰaː/	$-1.05(0.04)^{***}$
MaxLocV	$-0.08(0.02)^{***}$
SegmentVowel	$-0.76(0.04)^{***}$
Syllable /hɛː/:SegmentVowel	$0.40(0.06)^{***}$
Syllable /heː/:SegmentVowel	$0.43(0.06)^{***}$
Syllable /hɔː/:SegmentVowel	$0.27(0.05)^{***}$
Syllable /ʔaː/:SegmentVowel	$0.81(0.06)^{***}$
Syllable /ʔɛː/:SegmentVowel	$0.77(0.05)^{***}$
Syllable /ʔeː/:SegmentVowel	$0.77(0.05)^{***}$
Syllable /ʔɔː/:SegmentVowel	$0.78(0.06)^{***}$
Syllable /naː/:SegmentVowel	$0.53(0.06)^{***}$
Syllable /daː/:SegmentVowel	$0.71(0.05)^{***}$
Syllable /tʰaː/:SegmentVowel	$0.71(0.05)^{***}$
AIC	173.80
BIC	309.85
Log Likelihood	-61.90
Num. obs.	1706
Num. groups: Speaker	6
Var: Speaker (Intercept)	0.02
Var: Residual	0.06

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4.4: Linear mixed effects model. NasalAir ∼ Syllable + Segment + MaxLoc + Syllable*Segment + (1—Speaker)

other syllables are produced with similar integrated nasal airflow during the consonant and following vowel.

Figure 4.43: Integrated nasal airflow for vowels after glottal consonants

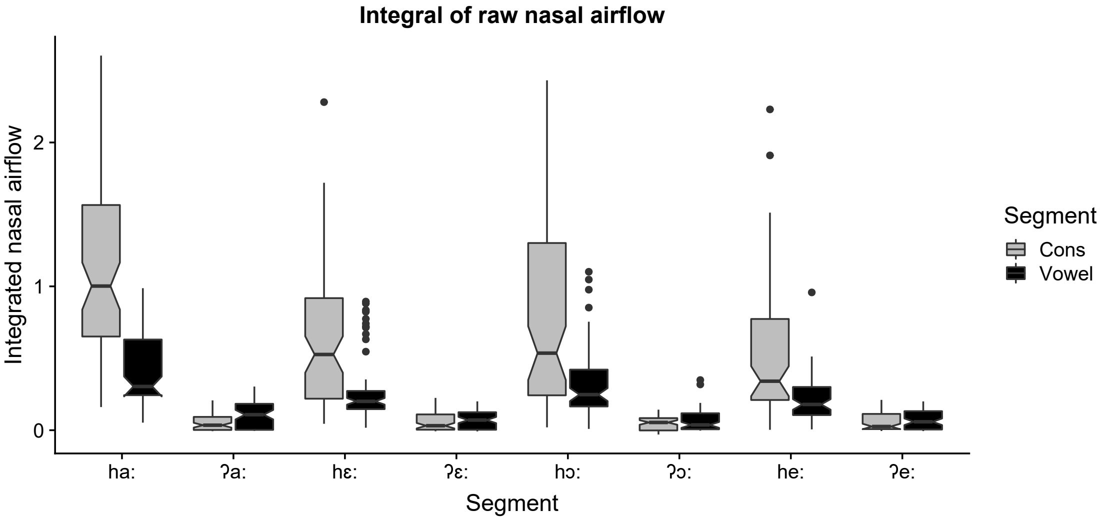


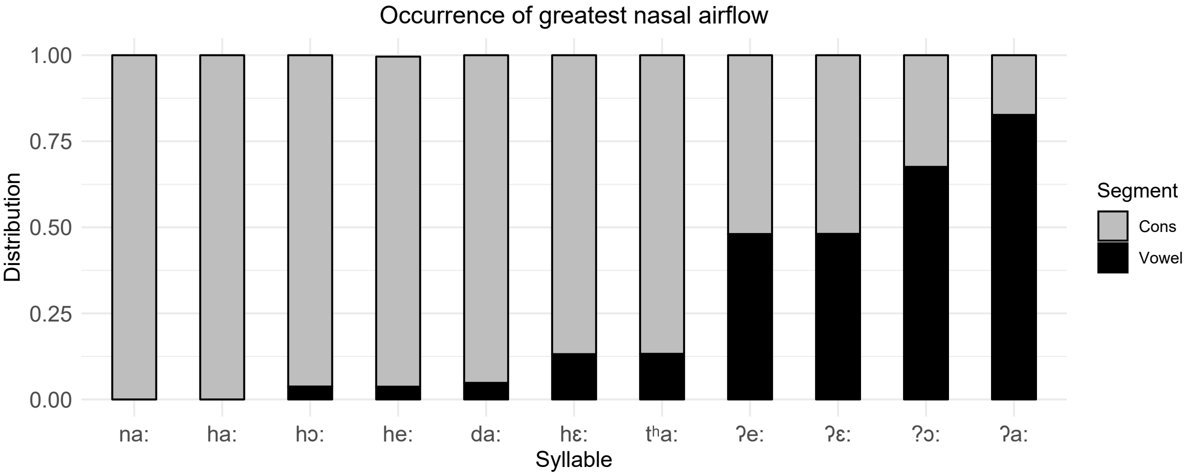
Figure 4.43 shows boxplots of integrated nasal airflow for all vowels after /h/ and /ʔ/. Again, we notice that the consonant of /haː/ has greater nasal airflow than all other consonants and vowels, including other syllables with /h/ (Table 4.6). Every consonant /h/ is produced with greater integrated nasal airflow than every /ʔ/. All vowels after /h/ are also produced with greater nasal airflow than vowels after /ʔ/, except for the contrasts /hɛː/ vs /ʔaː/, /hɔː/ vs /ʔaː/ and /heː/ vs any syllable beginning with /ʔ/ (Table 4.7). Furthermore, while nasal airflow is greater for every /h/ compared to the nasal airflow during the immediately following vowel, nasal airflow is similar during every /ʔ/ and its following vowel (Table 4.5). In all, vowels after /ʔ/ exhibit similar nasal airflow as predictably non-nasal vowels of the syllables /daː/ and /t^{h}aː/, with the exception for the vowel of /ʔaː/ compared to /hɛː/ and /hɔː/.

The distribution of maximum nasal airflow is shown in Figure 4.44. For syllables /naː/ and /haː/, the maximum integrated nasal airflow occurs during the onset consonant 100% of the time. Maximum nasal airflow occurs during the consonant most of the time for syllables

contrast	estimate	SE	df	t.ratio	p.value
Syllable = /haː/					
Cons - Vowel	0.7578	0.0387	1678.00	19.561	<.0001
Syllable = /hɛː/					
Cons - Vowel	0.3597	0.0392	1678.00	9.165	<.0001
Syllable = /heː/					
Cons - Vowel	0.3282	0.0403	1678.01	8.138	<.0001
Syllable = /hɔː/					
Cons - Vowel	0.4880	0.0383	1678.00	12.757	<.0001
Syllable = /ʔaː/					
Cons - Vowel	-0.0562	0.0395	1678.00	-1.424	0.1548
Syllable = /ʔɛː/					
Cons - Vowel	-0.0128	0.0385	1678.00	-0.332	0.7398
Syllable = /ʔeː/					
Cons - Vowel	-0.0124	0.0390	1678.00	-0.317	0.7512
Syllable = /ʔɔː/					
Cons - Vowel	-0.0203	0.0398	1678.00	-0.510	0.6103
Syllable = /naː/					
Cons - Vowel	0.2321	0.0392	1678.00	5.913	<.0001
Syllable = /daː/					
Cons - Vowel	0.0488	0.0376	1678.00	1.301	0.1935
Syllable = /tʰaː/					
Cons - Vowel	0.0508	0.0376	1678.00	1.352	0.1766

Table 4.5: Post-hoc analysis of integrated nasal airflow compared between the consonant and vowel of the same syllable using estimated marginal means with a Tukey multiplicity adjustment (Lenth, 2018).

Figure 4.44: Temporal location of greatest nasal airflow: consonant vs. vowel



/hɔː/, /heː/, /daː/, and /tʰaː/. The location of maximum nasal airflow is more likely to occur during the vowel for all syllables beginning with /ʔ/.

contrast	estimate	SE	df	t.ratio	p.value
Segment = Cons					
/haː/ - /hɛː/	0.5438	0.0391	1678.04	13.915	<.0001
/haː/ - /heː/	0.6307	0.0394	1678.01	16.006	<.0001
/haː/ - /hɔː/	0.3779	0.0385	1678.01	9.814	<.0001
/haː/ - /ʔaː/	1.0527	0.0422	1678.77	24.942	<.0001
/haː/ - /ʔɛː/	1.0786	0.0397	1678.27	27.181	<.0001
/haː/ - /ʔeː/	1.0786	0.0399	1678.28	27.013	<.0001
/haː/ - /ʔɔː/	1.0780	0.0414	1678.58	26.067	<.0001
/haː/ - /naː/	0.6469	0.0390	1678.01	16.589	<.0001
/haː/ - /daː/	1.0744	0.0382	1678.01	28.153	<.0001
/haː/ - /tʰaː/	1.0543	0.0382	1678.02	27.577	<.0001
/hɛː/ - /heː/	0.0870	0.0397	1678.02	2.191	0.5113
/hɛː/ - /hɔː/	-0.1659	0.0388	1678.03	-4.275	0.0010
/hɛː/ - /ʔaː/	0.5090	0.0416	1678.54	12.249	<.0001
/hɛː/ - /ʔɛː/	0.5348	0.0394	1678.13	13.566	<.0001
/hɛː/ - /ʔeː/	0.5348	0.0397	1678.13	13.481	<.0001
/hɛː/ - /ʔɔː/	0.5343	0.0409	1678.37	13.074	<.0001
/hɛː/ - /naː/	0.1031	0.0393	1678.02	2.622	0.2379
/hɛː/ - /daː/	0.5306	0.0385	1678.03	13.800	<.0001
/hɛː/ - /tʰaː/	0.5105	0.0384	1678.01	13.291	<.0001
/heː/ - /hɔː/	-0.2529	0.0392	1678.01	-6.457	<.0001
/heː/ - /ʔaː/	0.4220	0.0425	1678.69	9.923	<.0001
/heː/ - /ʔɛː/	0.4478	0.0402	1678.22	11.153	<.0001
/heː/ - /ʔeː/	0.4478	0.0404	1678.23	11.085	<.0001
/heː/ - /ʔɔː/	0.4473	0.0417	1678.51	10.716	<.0001
/heː/ - /naː/	0.0162	0.0397	1678.01	0.407	1.0000
/heː/ - /daː/	0.4436	0.0388	1678.01	11.427	<.0001
/heː/ - /tʰaː/	0.4236	0.0389	1678.02	10.900	<.0001
/hɔː/ - /ʔaː/	0.6749	0.0417	1678.76	16.166	<.0001
/hɔː/ - /ʔɛː/	0.7007	0.0393	1678.26	17.831	<.0001
/hɔː/ - /ʔeː/	0.7007	0.0395	1678.26	17.718	<.0001
/hɔː/ - /ʔɔː/	0.7001	0.0409	1678.57	17.107	<.0001
/hɔː/ - /naː/	0.2690	0.0388	1678.01	6.941	<.0001
/hɔː/ - /daː/	0.6965	0.0379	1678.01	18.374	<.0001
/hɔː/ - /tʰaː/	0.6764	0.0379	1678.02	17.825	<.0001
/ʔaː/ - /ʔɛː/	0.0259	0.0396	1678.19	0.653	0.9999
/ʔaː/ - /ʔeː/	0.0258	0.0398	1678.18	0.648	0.9999
/ʔaː/ - /ʔɔː/	0.0253	0.0397	1678.03	0.636	0.9999
/ʔaː/ - /naː/	-0.4058	0.0424	1678.70	-9.571	<.0001
/ʔaː/ - /daː/	0.0216	0.0414	1678.77	0.523	1.0000
/ʔaː/ - /tʰaː/	0.0016	0.0408	1678.64	0.039	1.0000
/ʔɛː/ - /ʔeː/	-0.0000	0.0387	1678.01	-0.001	1.0000
/ʔɛː/ - /ʔɔː/	-0.0006	0.0393	1678.09	-0.014	1.0000
/ʔɛː/ - /naː/	-0.4317	0.0399	1678.23	-10.817	<.0001
/ʔɛː/ - /daː/	-0.0042	0.0389	1678.26	-0.108	1.0000
/ʔɛː/ - /tʰaː/	-0.0243	0.0386	1678.17	-0.629	0.9999
/ʔeː/ - /ʔɔː/	-0.0005	0.0396	1678.08	-0.013	1.0000
/ʔeː/ - /naː/	-0.4317	0.0402	1678.23	-10.751	<.0001
/ʔeː/ - /daː/	-0.0042	0.0392	1678.27	-0.106	1.0000
/ʔeː/ - /tʰaː/	-0.0242	0.0389	1678.18	-0.624	0.9999
/ʔɔː/ - /naː/	-0.4311	0.0416	1678.51	-10.373	<.0001
/ʔɔː/ - /daː/	-0.0036	0.0405	1678.57	-0.090	1.0000
/ʔɔː/ - /tʰaː/	-0.0237	0.0401	1678.45	-0.591	1.0000
/naː/ - /daː/	0.4275	0.0384	1678.01	11.127	<.0001
/naː/ - /tʰaː/	0.4074	0.0385	1678.01	10.586	<.0001
da - /tʰaː/	-0.0201	0.0376	1678.02	-0.534	1.0000

Table 4.6: Post-hoc analysis of integrated nasal airflow during the onset consonant compared across syllables using estimated marginal means with a Tukey multiplicity adjustment (Lenth, 2018).

contrast	estimate	SE	df	t.ratio	p.value
Segment = Vowel					
/haː/ - /hɛː/	0.1456	0.0391	1678.04	3.727	0.0092
/haː/ - /heː/	0.2011	0.0397	1678.01	5.067	<.0001
/haː/ - /hɔː/	0.1081	0.0385	1678.01	2.807	0.1562
/haː/ - /ʔaː/	0.2387	0.0422	1678.77	5.655	<.0001
/haː/ - /ʔɛː/	0.3080	0.0397	1678.27	7.762	<.0001
/haː/ - /ʔeː/	0.3084	0.0399	1678.28	7.724	<.0001
/haː/ - /ʔɔː/	0.3000	0.0414	1678.58	7.253	<.0001
/haː/ - /naː/	0.1211	0.0390	1678.01	3.107	0.0705
/haː/ - /daː/	0.3654	0.0382	1678.01	9.576	<.0001
/haː/ - /tʰaː/	0.3473	0.0382	1678.02	9.084	<.0001
/hɛː/ - /heː/	0.0555	0.0400	1678.02	1.388	0.9515
/hɛː/ - /hɔː/	-0.0376	0.0388	1678.03	-0.968	0.9968
/hɛː/ - /ʔaː/	0.0931	0.0416	1678.54	2.240	0.4768
/hɛː/ - /ʔɛː/	0.1624	0.0394	1678.13	4.118	0.0020
/hɛː/ - /ʔeː/	0.1627	0.0397	1678.13	4.103	0.0021
/hɛː/ - /ʔɔː/	0.1543	0.0409	1678.37	3.776	0.0076
/hɛː/ - /naː/	-0.0245	0.0393	1678.02	-0.623	0.9999
/hɛː/ - /daː/	0.2198	0.0385	1678.03	5.716	<.0001
/hɛː/ - /tʰaː/	0.2016	0.0384	1678.01	5.249	<.0001
/heː/ - /hɔː/	-0.0930	0.0394	1678.01	-2.359	0.3942
/heː/ - /ʔaː/	0.0376	0.0428	1678.68	0.878	0.9986
/heː/ - /ʔɛː/	0.1069	0.0404	1678.22	2.644	0.2272
/heː/ - /ʔeː/	0.1073	0.0407	1678.23	2.638	0.2302
/heː/ - /ʔɔː/	0.0988	0.0420	1678.50	2.353	0.3980
/heː/ - /naː/	-0.0800	0.0399	1678.01	-2.002	0.6474
/heː/ - /daː/	0.1643	0.0391	1678.01	4.201	0.0014
/heː/ - /tʰaː/	0.1462	0.0391	1678.01	3.734	0.0089
/hɔː/ - /ʔaː/	0.1306	0.0417	1678.76	3.129	0.0661
/hɔː/ - /ʔɛː/	0.1999	0.0393	1678.26	5.088	<.0001
/hɔː/ - /ʔeː/	0.2003	0.0395	1678.26	5.066	<.0001
/hɔː/ - /ʔɔː/	0.1919	0.0409	1678.57	4.688	0.0002
/hɔː/ - /naː/	0.0131	0.0388	1678.01	0.337	1.0000
/hɔː/ - /daː/	0.2574	0.0379	1678.01	6.789	<.0001
/hɔː/ - /tʰaː/	0.2392	0.0379	1678.02	6.304	<.0001
/ʔaː/ - /ʔɛː/	0.0693	0.0396	1678.19	1.751	0.8085
/ʔaː/ - /ʔeː/	0.0697	0.0398	1678.18	1.750	0.8089
/ʔaː/ - /ʔɔː/	0.0613	0.0397	1678.03	1.541	0.9057
/ʔaː/ - /naː/	-0.1176	0.0424	1678.70	-2.772	0.1697
/ʔaː/ - /daː/	0.1267	0.0414	1678.77	3.064	0.0796
/ʔaː/ - /tʰaː/	0.1086	0.0408	1678.64	2.661	0.2187
/ʔɛː/ - /ʔeː/	0.0004	0.0387	1678.01	0.010	1.0000
/ʔɛː/ - /ʔɔː/	-0.0080	0.0393	1678.09	-0.205	1.0000
/ʔɛː/ - /naː/	-0.1869	0.0399	1678.23	-4.682	0.0002
/ʔɛː/ - /daː/	0.0574	0.0389	1678.26	1.476	0.9278
/ʔɛː/ - /tʰaː/	0.0393	0.0386	1678.17	1.017	0.9952
/ʔeː/ - /ʔɔː/	-0.0084	0.0396	1678.08	-0.213	1.0000
/ʔeː/ - /naː/	-0.1872	0.0402	1678.23	-4.663	0.0002
/ʔeː/ - /daː/	0.0570	0.0392	1678.27	1.456	0.9336
/ʔeː/ - /tʰaː/	0.0389	0.0389	1678.18	1.001	0.9958
/ʔɔː/ - /naː/	-0.1788	0.0416	1678.51	-4.302	0.0009
/ʔɔː/ - /daː/	0.0655	0.0405	1678.57	1.615	0.8762
/ʔɔː/ - /tʰaː/	0.0473	0.0401	1678.45	1.180	0.9846
/naː/ - /daː/	0.2443	0.0384	1678.01	6.358	<.0001
/naː/ - /tʰaː/	0.2261	0.0385	1678.01	5.876	<.0001
/daː/ - /tʰaː/	-0.0182	0.0376	1678.02	-0.483	1.0000

Table 4.7: Post-hoc analysis of integrated nasal airflow during the vowel compared across syllables using estimated marginal means with a Tukey multiplicity adjustment (Lenth, 2018).

Summary

A summary of airflow results can be found in Table 4.8. Overall, we observe that /haː/ is produced with the greatest nasal airflow during both the consonant and during some of the following vowel compared to any other syllable. The consonants of all syllables beginning with /h/ are produced with nasal airflow during the consonant similar to or higher than /n/. /hɔː/, /hɛː/, and /heː/ are produced with nasal airflow similar to /naː/ or higher during the vowel but that gradually lessens. Syllables beginning with the onset /ʔ/, /tʰ/, and /d/ are produced with minimal nasal airflow.

Contrast	Consonant airflow	Vowel airflow	Proportional nasal airflow
Nucleus /a/	haː>naː>ʔaː/tʰaː/daː	haː>naː*>ʔaː>tʰaː/daː	naː>haː*>ʔaː>daː*>tʰaː*
Onset /h/	haː>hɔː/hɛː/heː	haː>hɔː/hɛː/heː*	haː/hɔː>hɛː/heː*
Onset /ʔ/	n.s.	n.s.	n.s.
Onset /tʰ/	n.s.	n.s.	n.s.

Table 4.8: Airflow results summary. A star indicates that the difference recorded between the syllable and the previous syllable only occur for a portion of the segment rather than the entire duration of the segment. Non-significantly different airflow is indicated by n.s.

The analysis of proportional nasal airflow shows that, given oral airflow, /naː/ is produced with the greatest proportional nasal airflow for at least part of the syllables, followed by /haː/. Of syllables beginning with /h/, those containing low and mid-low back vowels /haː/ and /hɔː/ are produced with higher proportional nasal airflow. Syllables beginning with /ʔ/, /tʰ/, and /d/ are produced with low proportional nasal airflow.

4.5 Open quotient

As described in Section 4, open quotient was calculated using the DEGG method of measuring opening and closing glottal phases (Michaud *et al.*, 2017). When a glottal cycle lacks a distinct opening phase, as is often the case with extreme creaky phonation (Figure 4.5 from Section4.2), open quotient is unmeasurable using the DEGG method. After manual inspection, it was determined that 255 out of 16320 glottal cycles (1.5% of the data) were unmeasurable and were excluded from further analysis. Of these excluded glottal cycles, all originated from tokens containing the onset /ʔ/ and all primarily occurred during the initial 15% of the duration of the vowel. It is likely that these excluded glottal cycles exhibit coarticulatory creaky phonation from the preceding glottal stop. See Table 4.10 for a list of how many glottal cycles were omitted from analysis per speaker.

Figure 4.45 shows mean open quotient of all syllables with an /aː/ nucleus across all speakers' data combined. These syllables include /haː/, /ʔaː/, /naː/, /daː/, and /tʰaː/. In Figure 4.45, confidence interval bars are based on the standard error. We observe that open quotient is largely similar throughout most of the vowel in all syllables containing the nucleus /aː/. During the initial 10% of the vowel, the open quotient of /haː/ and /tʰaː/ is significantly higher than other syllables in that the confidence intervals do not overlap.

While a GAMM of open quotient (AR-1 = 0.54, $r^2 = 0.4$) reveals no overall predicted significant differences among syllables, Time is a significant factor in predicting open quotient for every syllable (Table 4.9). A plot of predicted open quotient values for all speakers' combined data is shown in Figure 4.46. GAMM-predicted open quotient is similar to the observed mean open quotient in that every syllable overlaps during most of the duration of the vowel. During the initial 10% of the vowel, open quotient during /haː/ and /tʰaː/ is predicted to be larger than other syllables except for /naː/. In summary, open quotient is high as the beginning of vowels after /h/ and /tʰ/ and falls during the vowel. This increased breathiness is likely coarticulation from the preceding voiceless frication of /h/ and aspiration that quickly transitions into modal voice.

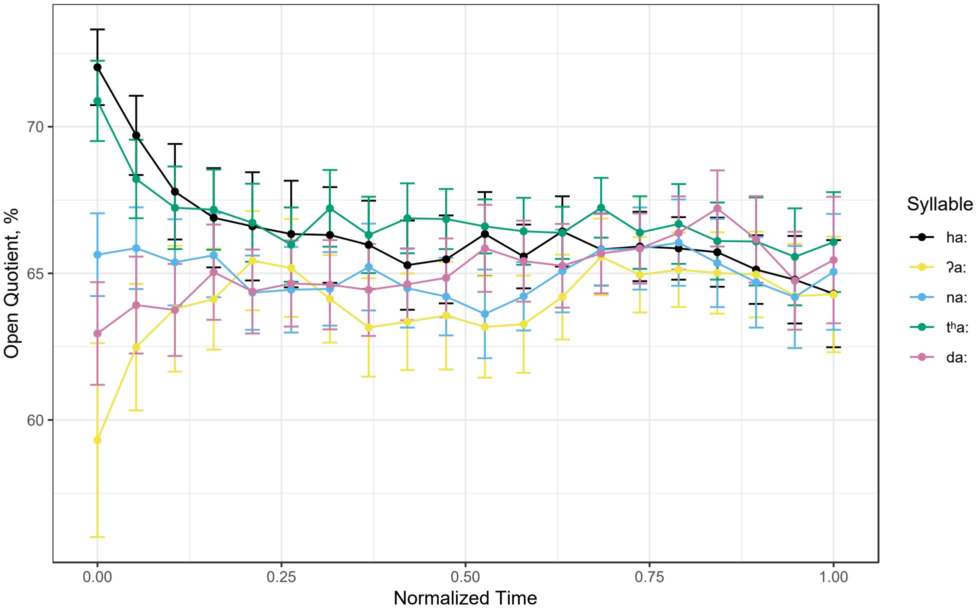


Figure 4.45: Mean open quotient during all vowels /a/. Confidence interval bars are based on the standard error. All speakers' data are combined.

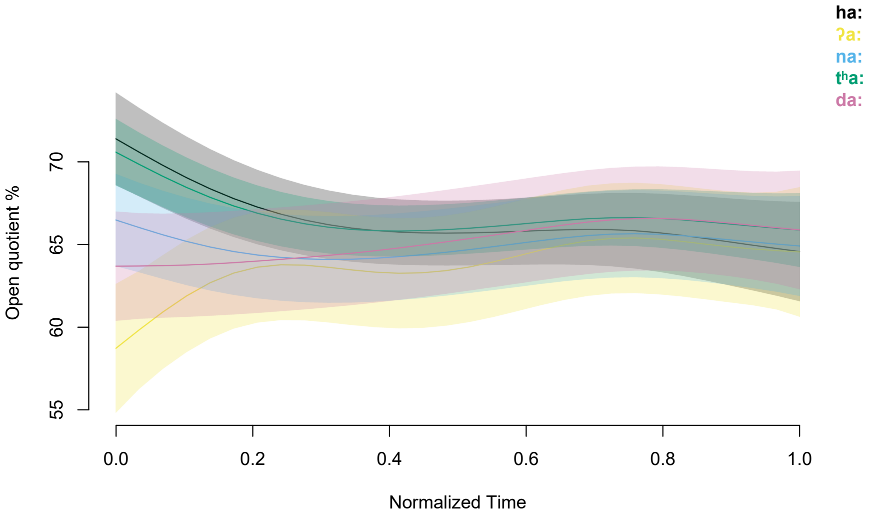


Figure 4.46: GAMM-predicted open quotient during all vowels /a/. Confidence interval bars are based on the standard error. All speakers' data are combined.

	Open quotient
(Intercept)	66.49(0.94)***
Syllable /hɛː/	1.86(1.73)
Syllable /hɔː/	−0.48(1.35)
Syllable /heː/	−1.72(1.87)
Syllable /ʔaː/	−2.80(1.48)
Syllable /ʔɛː/	−1.83(1.92)
Syllable /ʔɔː/	−2.78(1.50)
Syllable /ʔeː/	−2.95(1.68)
Syllable /naː/	−1.49(1.56)
Syllable /t^{h}aː/	0.26(1.17)
Syllable /t^{h}ɛː/	−0.94(2.04)
Syllable /t^{h}eː/	−0.53(1.61)
Syllable /daː/	−1.31(1.73)
EDF: s(Time)	4.98(5.97)***
EDF: s(Time):Syllable /haː/	1.00(1.00)***
EDF: s(Time):Syllable /hɛː/	1.05(1.07)***
EDF: s(Time):Syllable /hɔː/	1.36(1.53)***
EDF: s(Time):Syllable /heː/	2.95(3.48)***
EDF: s(Time):Syllable /ʔaː/	4.44(5.25)***
EDF: s(Time):Syllable /ʔɛː/	3.87(4.62)***
EDF: s(Time):Syllable /ʔɔː/	4.60(5.56)***
EDF: s(Time):Syllable /ʔeː/	4.36(5.19)***
EDF: s(Time):Syllable /naː/	1.97(2.41)***
EDF: s(Time):Syllable /t^{h}aː/	1.00(1.00)***
EDF: s(Time):Syllable /t^{h}ɛː/	1.00(1.00)***
EDF: s(Time):Syllable /t^{h}eː/	1.00(1.00)***
EDF: s(Time):Syllable /daː/	2.86(3.48)***
EDF: s(Speaker,Time):Syllable /haː/	17.25(53.00)***
EDF: s(Speaker,Time):Syllable /hɛː/	20.60(53.00)***
EDF: s(Speaker,Time):Syllable /hɔː/	15.41(53.00)***
EDF: s(Speaker,Time):Syllable /heː/	20.41(53.00)***
EDF: s(Speaker,Time):Syllable /ʔaː/	20.42(53.00)***
EDF: s(Speaker,Time):Syllable /ʔɛː/	18.63(53.00)***
EDF: s(Speaker,Time):Syllable /ʔɔː/	20.32(53.00)***
EDF: s(Speaker,Time):Syllable /ʔeː/	19.56(53.00)***
EDF: s(Speaker,Time):Syllable /naː/	8.12(53.00)***
EDF: s(Speaker,Time):Syllable /t^{h}aː/	10.41(53.00)***
EDF: s(Speaker,Time):Syllable /t^{h}ɛː/	8.47(53.00)***
EDF: s(Speaker,Time):Syllable /t^{h}eː/	19.63(53.00)***
EDF: s(Speaker,Time):Syllable /daː/	11.89(53.00)***
AIC	101726.20
BIC	103865.29
Log Likelihood	-50584.73
Deviance	421328.15
Deviance explained	0.41
Dispersion	44.84
R^2	0.40
GCV score	50941.72
Num. obs.	16065
Num. smooth terms	27

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4.9: GAMM of open quotient. The model was constructed using the algorithm: Open quotient ~s(Time) + Syllable + s(Time, by = Syllable) + s(Speaker, Time,by=Syllable, bs = ”fs”, m=1, na.action=na.gam.replace).

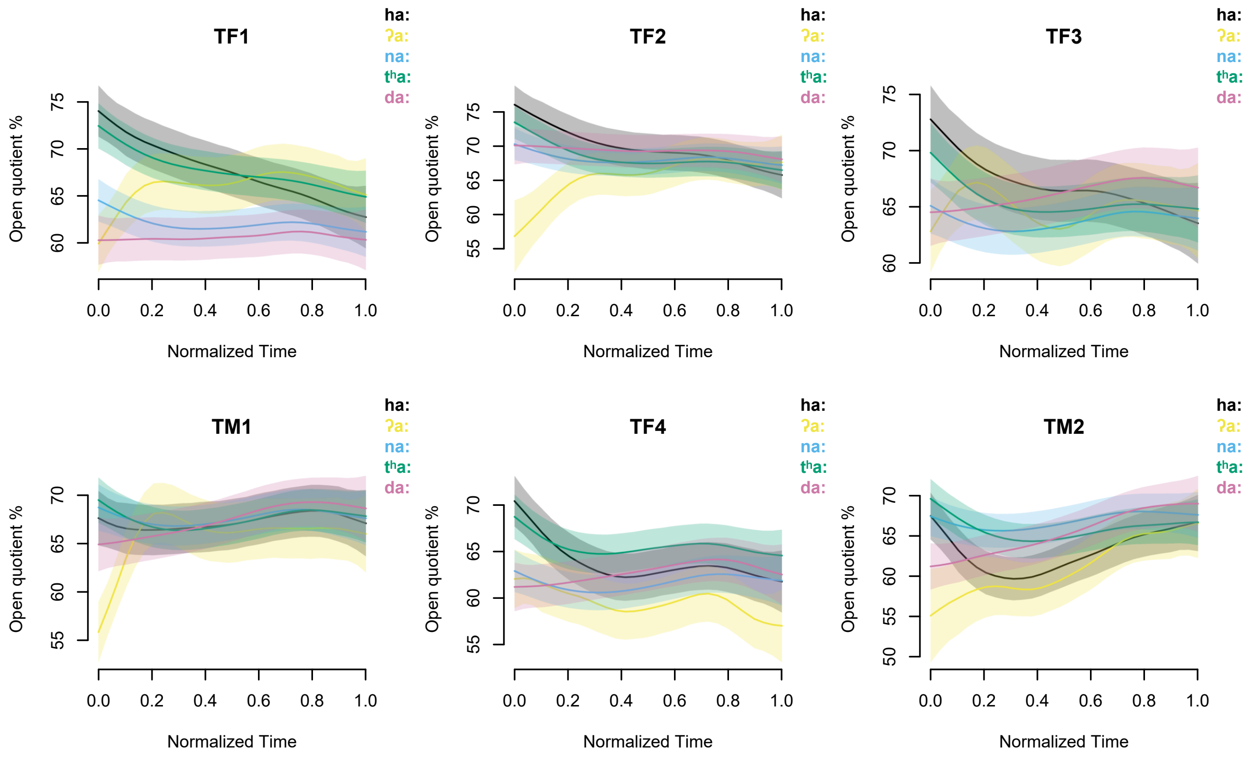


Figure 4.47: GAMM-predicted open quotient during all vowels /a/. Confidence interval bars are based on the standard error.

A plot of individual predicted open quotient can be found in Figure 4.47. We observe some individual variability among speakers. TM1 and TM4 are predicted to produce /haː/ and /tʰaː/ with greater open quotient during the initial 10% of the vowel, while other speakers are predicted to produce syllables with overlapping open quotient. Speakers TF2 and TM1 are predicted to produce /ʔaː/ with lower open quotient than all other syllables at the start of the vowel.

Figure 4.48 shows mean open quotient during the vowels of all syllables beginning with the onset /h/. Open quotient is similar across all syllables with an onset /h/. The GAMM-predicted open quotient (Figure 4.49) reveals a similar pattern as the observed mean values. Open quotient is predicted to be similar during all syllables beginning with the consonant /h/, regardless of vowel height. A plot of GAMM-predicted open quotient for individual speakers also shows that open quotient is predicted to be similar across all /h/-onset syllables,

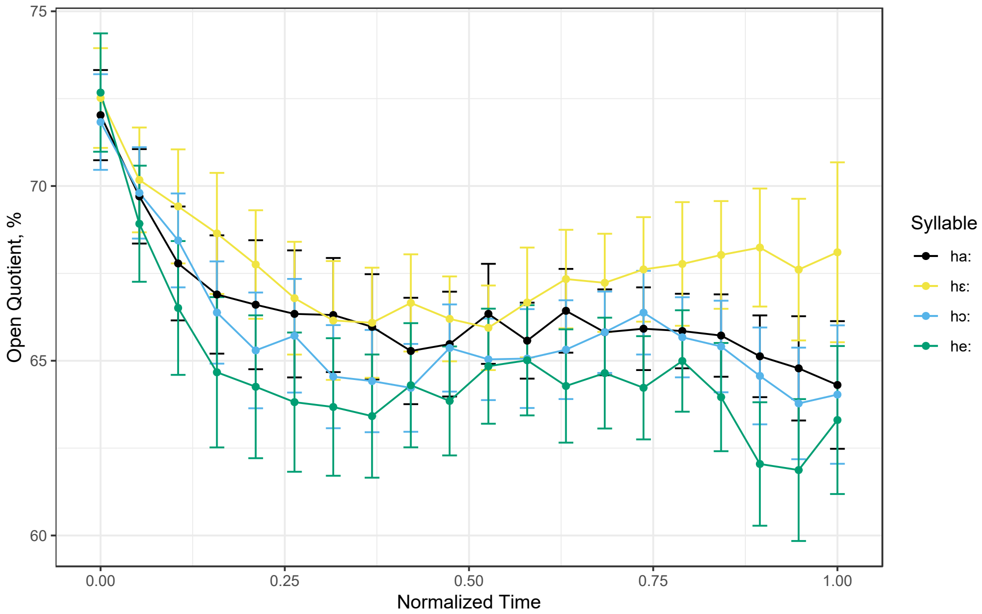


Figure 4.48: Mean open quotient during all vowels after /h/. Confidence interval bars are based on the standard error. All speakers' data are combined.

although TM2 is predicted to produce /hɛː/ with somewhat greater open quotient than other syllables at the end of the vowel (Figure 4.50).

Figure 4.51 shows mean open quotient for all syllables with the onset consonant /ʔ/; /haː/ is also included for comparison. Results show that open quotient begins low during vowels following /ʔ/ relative to open quotient after /h/. Open quotient quickly rises by approximately 10% during the first 10% of the vowel and remains steady throughout the rest of the vowel. Furthermore, open quotient is similar during all vowels after /ʔ/.

GAMM-predicted open quotient after /ʔ/ is similar to the observed mean of the data (Figure 4.52). Open quotient after /ʔ/ begins relatively low, then rises to similar values as /haː/ about 10% into the duration of the vowel. Plots of individual GAMM-predicted open quotient show similar open quotient trajectories as the combined data, although TM1 is predicted to produce /ʔɔː/ with slightly lower open quotient during the initial 10% of the vowel compared to other syllables.

Finally, Figure 4.54 shows open quotient during all vowels following the onset /t^h/; /haː/ is also shown for comparison. We observe that all syllables with a /t^h/ onset exhibit similar

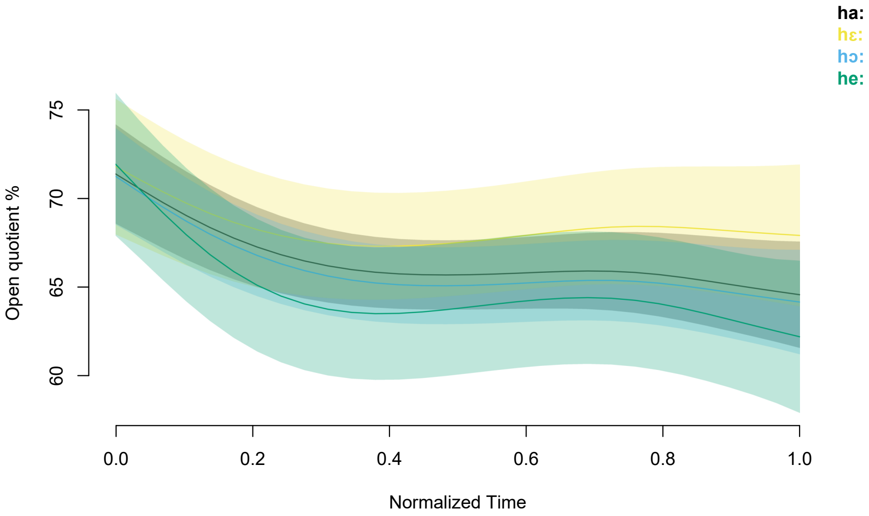


Figure 4.49: GAMM-predicted open quotient during all vowels after /h/. Confidence interval bars are based on the standard error. All speakers' data are combined.

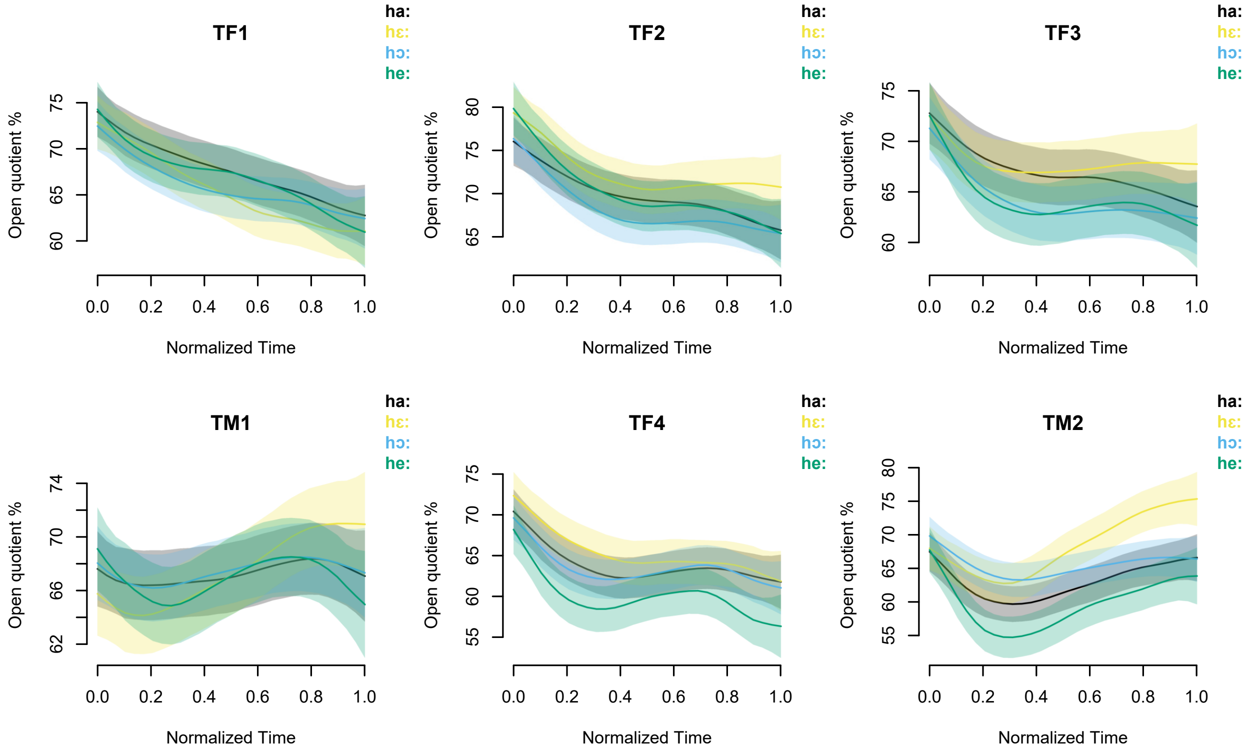


Figure 4.50: GAMM-predicted open quotient during all vowels after /h/. Confidence interval bars are based on the standard error.

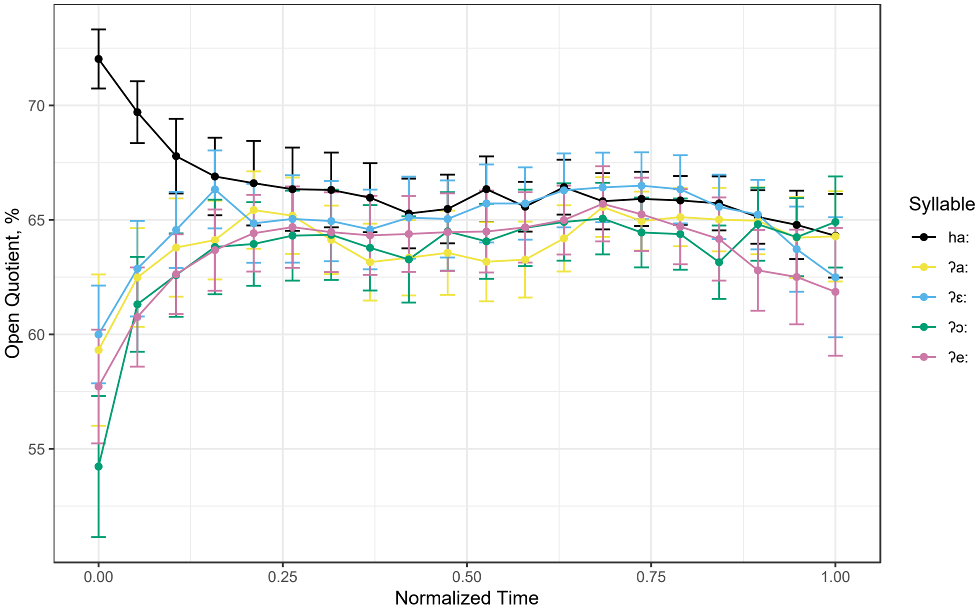


Figure 4.51: Mean open quotient during all vowels after /ʔ/. Confidence interval bars are based on the standard error. All speakers' data are combined.

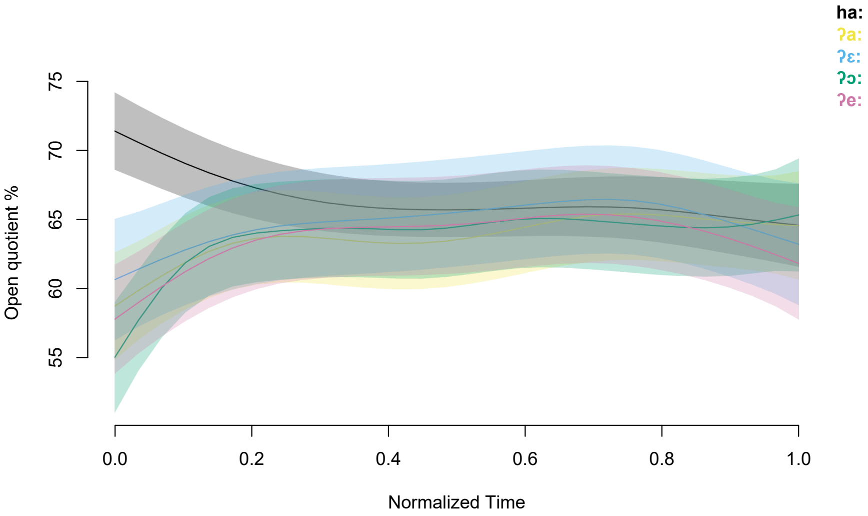


Figure 4.52: GAMM-predicted open quotient during all vowels after /ʔ/. Confidence interval bars are based on the standard error. All speakers' data are combined.

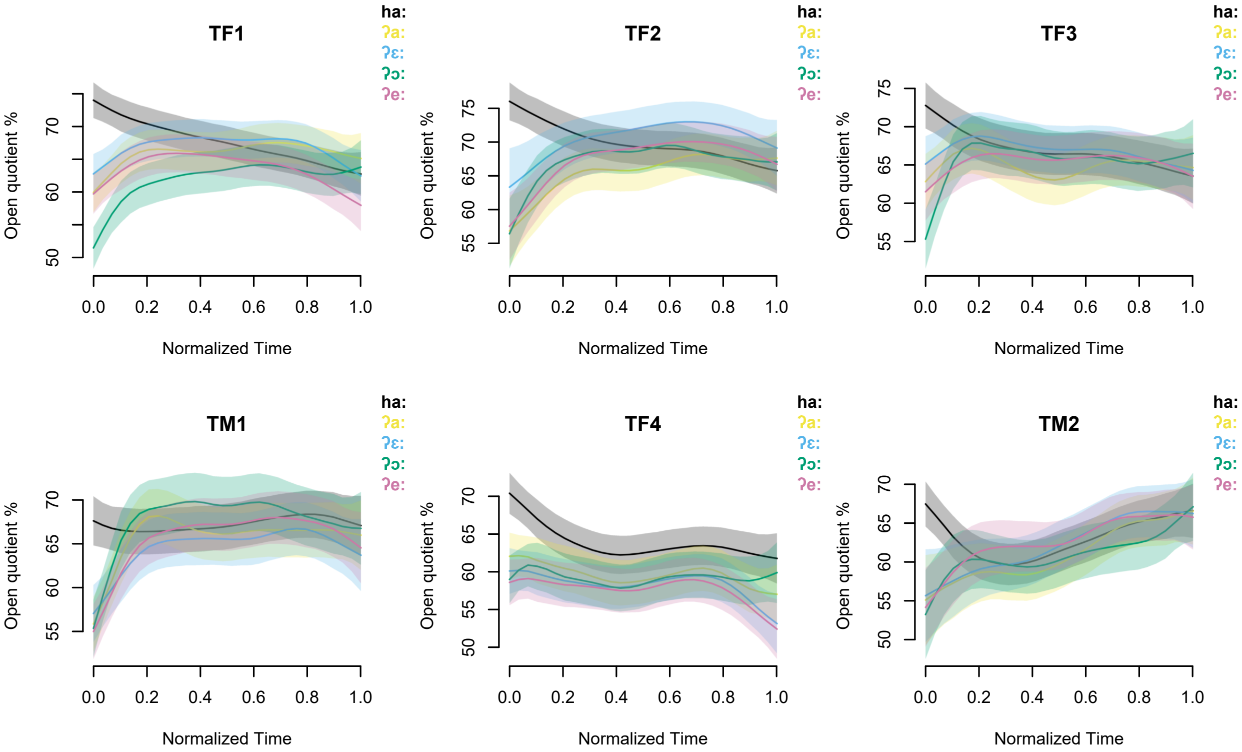


Figure 4.53: GAMM-predicted open quotient during all vowels after /ʔ/. Confidence interval bars are based on the standard error.

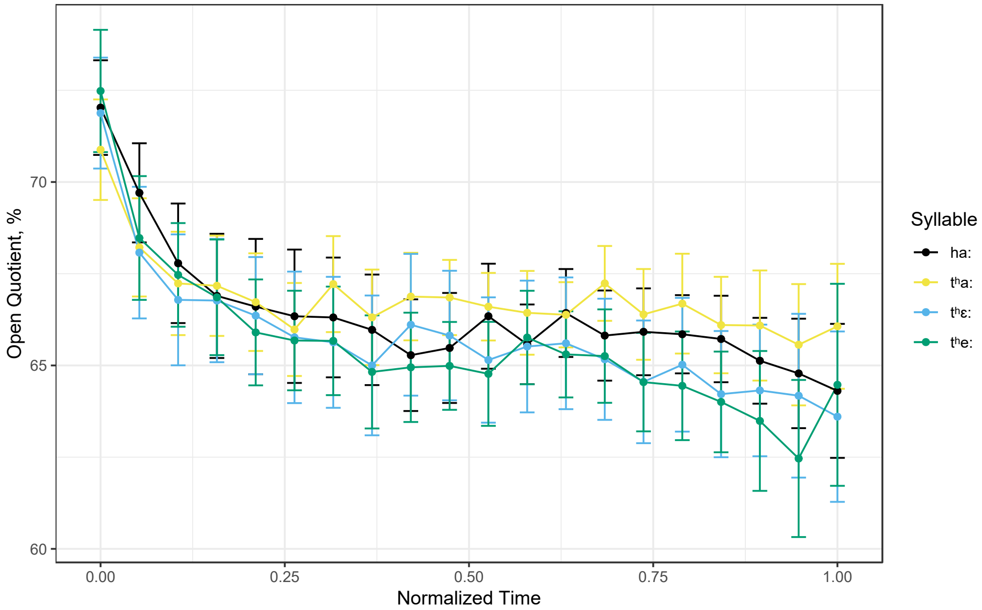


Figure 4.54: Mean open quotient during all vowels after /tʰ/. Confidence interval bars are based on the standard error. All speakers' data are combined.

open quotient as /haː/ throughout the entire duration of the vowel. A plot of GAMM-predicted open quotient yields a similar result (Figure 4.55). Plots of individual GAMM-predicted open quotient are similar to the combined plot, with the exception of TM2 who is predicted to produce /tʰɛː/ with lower open quotient during the latter half of the vowel compared to other syllables.

Speaker	No. Omissions	%
TF1	0	0%
TF2	108	3.9%
TF3	9	0.4%
TF4	0	0%
TM1	3	0.1%
TM1	135	5.6%

Table 4.10: Number of EGG data omissions per speaker. EGG/DEGG signals were hand-checked to determine amenability to analysis (see Section 4). A total of 1.5% of all EGG data was determined to be not amenable to analysis due to creaky phonation. The only syllables with unanalyzable values were those with the onset /ʔ/ during the initial 15% of the duration of the vowel.

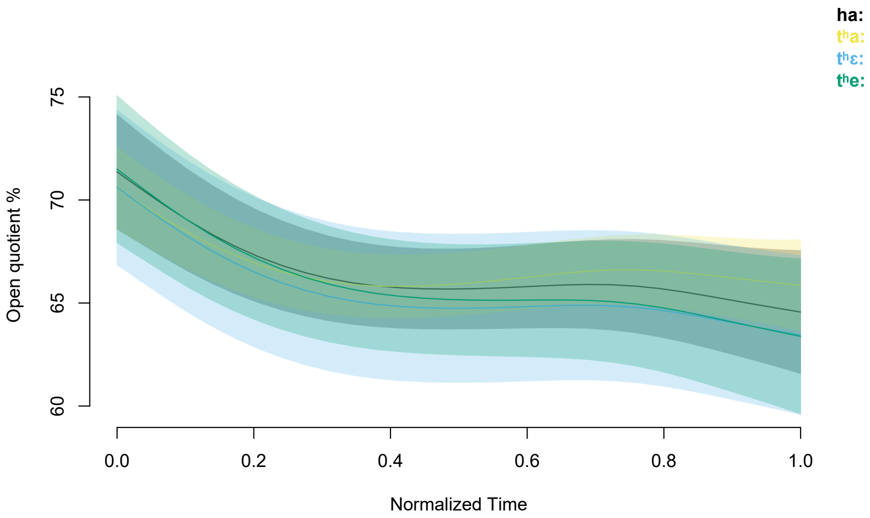


Figure 4.55: GAMM-predicted open quotient during all vowels after /tʰ/. /haː/ is also shown for comparison. Confidence interval bars are based on the standard error. All speakers' data are combined.

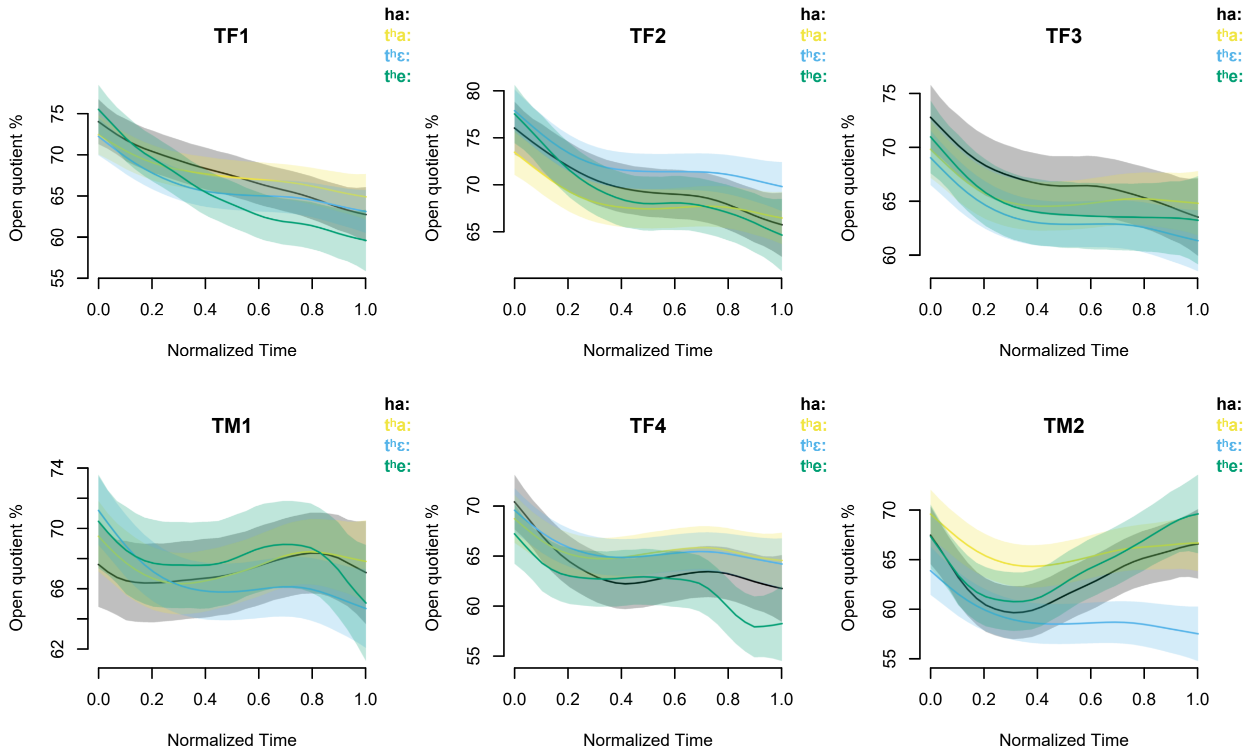


Figure 4.56: GAMM-predicted open quotient during all vowels after /tʰ/. /haː/ is also shown for comparison. Confidence interval bars are based on the standard error.

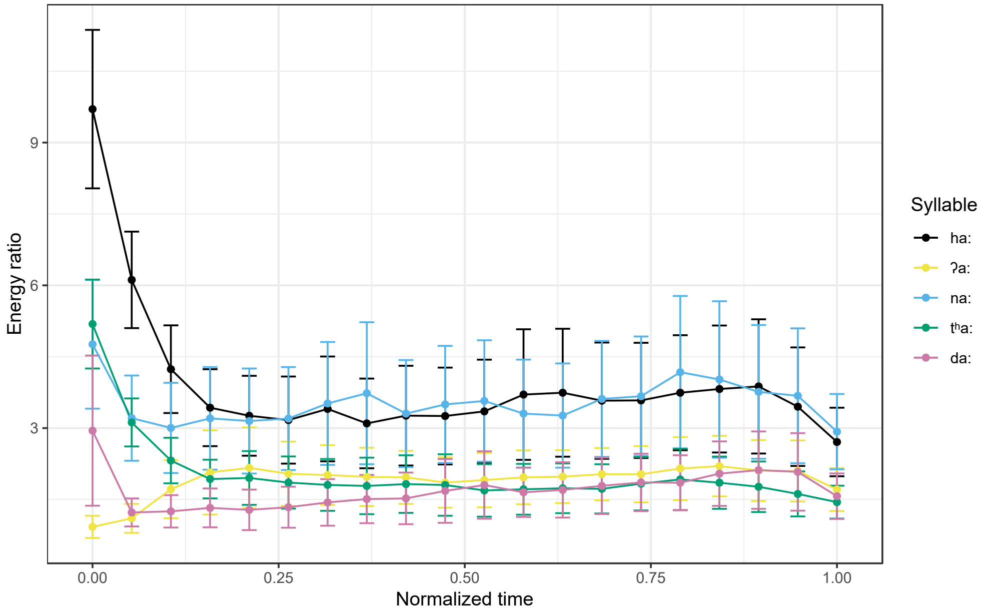


Figure 4.57: Mean energy ratio measured during all vowels /a/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

4.5.1 Energy ratio

Spectral tilt was measured by calculating the energy ratio of 0–320 Hz by 320–5360 Hz (Carignan, 2017; Pruthi and Espy-Wilson, 2005). Recall that this measure of energy ratio has been applied to measure nasal murmur during consonants. The purpose of this measure is to detect the presence of increased energy in low frequencies (<320 Hz). While this measure was specifically defined to detect nasal coupling during nasal consonants, increased energy in low frequency bands can also be associated with vowel nasalization or breathiness. Therefore, because the efficacy of this measure for nasalized and breathy vowels has not yet been assessed, the following analysis is exploratory.

Mean energy ratio of all speakers' data combined can be found in Figure 4.57. Results show that energy ratio is largely similar during all /a/ vowels; the primary difference being that the initial 10% of the vowel of /haː/ is produced with a higher energy ratio. Plots of individual speakers' data (Figure 4.58) show that /haː/ is generally produced with a relatively high energy ratio in all speakers during the initial 10% of the vowel. For speakers

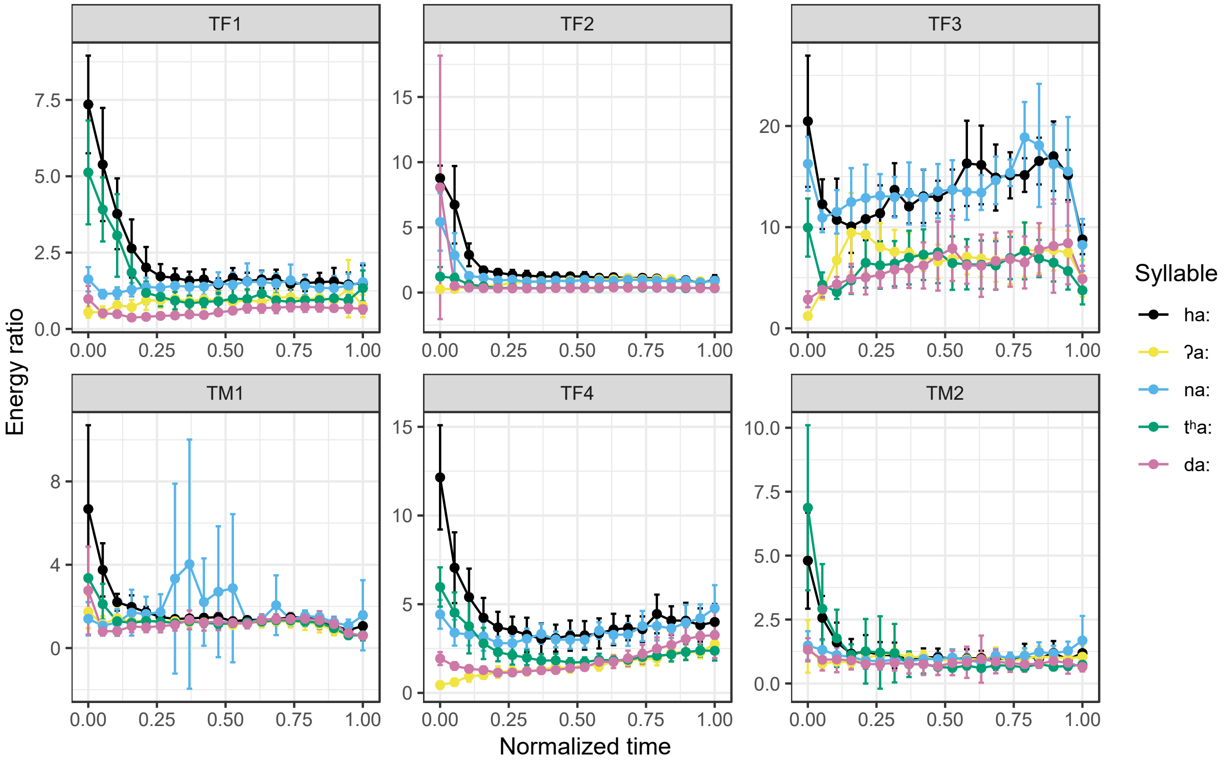


Figure 4.58: Mean energy ratio measured during all vowels /a/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

TF1 and TM2, /tʰaː/ is also produced with an initially high energy ratio while speaker TF3 also produces /naː/ with initially high energy ratio. For most speakers, energy ratio is identical across all syllables for the remainder of the vowel. Only for TF3 are /haː/ and /naː/ produced with higher energy ratio during most of the duration of the vowel.

Results of a GAMM predicting energy ratio (AR-1 = 0.48, R2 = 0.78) can be found in Table 4.11. Results show no overall predicted differences in energy ratio across syllables. However energy ratio over Time across syllables is significant for all syllables beginning with /h/ and /ʔ/. A plot of predicted energy ratio across all speakers during the vowel /aː/ is shown in Figure 4.59. Overall, we observe no predicted difference in energy ratio across syllables during any part of the vowel. Plots of predicted individual speaker energy ratio reveals some variability. Every speaker is predicted to produce /haː/ with relatively higher energy ratio during the initial 10% of the vowel. Speakers TF1 and TM2 are also predicted

	Energy ratio
(Intercept)	4.25(1.94)*
Syllable /hɛː/	0.12(2.54)
Syllable /heː/	−0.10(2.72)
Syllable /hɔː/	−2.23(2.12)
Syllable /ʔaː/	−2.24(2.05)
Syllable /ʔɛː/	−1.96(2.11)
Syllable /ʔeː/	−2.17(2.06)
Syllable /ʔɔː/	−3.19(1.98)
Syllable /naː/	−0.64(2.76)
Syllable /tʰaː/	−1.98(2.12)
Syllable /tʰɛː/	−1.68(2.21)
Syllable /tʰeː/	−0.97(2.20)
Syllable /daː/	−2.55(2.06)
EDF: s(Time)	8.25(8.68)***
EDF: s(Time):Syllable /haː/	5.63(6.49)***
EDF: s(Time):Syllable /hɛː/	4.44(5.32)**
EDF: s(Time):Syllable /heː/	6.69(7.71)***
EDF: s(Time):Syllable /hɔː/	5.77(6.88)***
EDF: s(Time):Syllable /ʔaː/	5.68(6.59)***
EDF: s(Time):Syllable /ʔɛː/	5.98(6.97)***
EDF: s(Time):Syllable /ʔeː/	5.57(6.51)***
EDF: s(Time):Syllable /ʔɔː/	5.46(6.51)***
EDF: s(Time):Syllable /naː/	2.21(2.46)
EDF: s(Time):Syllable /tʰaː/	1.00(1.00)
EDF: s(Time):Syllable /tʰɛː/	3.26(3.81)
EDF: s(Time):Syllable /tʰeː/	1.42(1.67)
EDF: s(Time):Syllable /daː/	1.89(2.12)
EDF: s(Speaker,Time):Syllable /haː/	26.92(53.00)***
EDF: s(Speaker,Time):Syllable /hɛː/	17.17(53.00)***
EDF: s(Speaker,Time):Syllable /heː/	17.62(53.00)***
EDF: s(Speaker,Time):Syllable /hɔː/	9.13(53.00)***
EDF: s(Speaker,Time):Syllable /ʔaː/	24.30(53.00)***
EDF: s(Speaker,Time):Syllable /ʔɛː/	20.85(53.00)***
EDF: s(Speaker,Time):Syllable /ʔeː/	22.63(53.00)***
EDF: s(Speaker,Time):Syllable /ʔɔː/	13.33(53.00)***
EDF: s(Speaker,Time):Syllable /naː/	29.80(53.00)***
EDF: s(Speaker,Time):Syllable /tʰaː/	16.02(53.00)***
EDF: s(Speaker,Time):Syllable /tʰɛː/	23.43(53.00)***
EDF: s(Speaker,Time):Syllable /tʰeː/	33.97(53.00)***
EDF: s(Speaker,Time):Syllable /daː/:	26.97(53.00)***
AIC	64321.03
BIC	67212.29
Log Likelihood	-31785.04
Deviance	45501.28
Deviance explained	0.78
Dispersion	3.79
R^2	0.78
GCV score	32419.79
Num. obs.	16320
Num. smooth terms	27

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4.11: GAMM of spectral energy ratio (0–320 Hz / 320–5360 Hz). The model was constructed using the algorithm: Energy ratio ˜s(Time) + Syllable + s(Time, by = Syllable) + s(Speaker, Time,by=Syllable, bs = ”fs”, m=1).

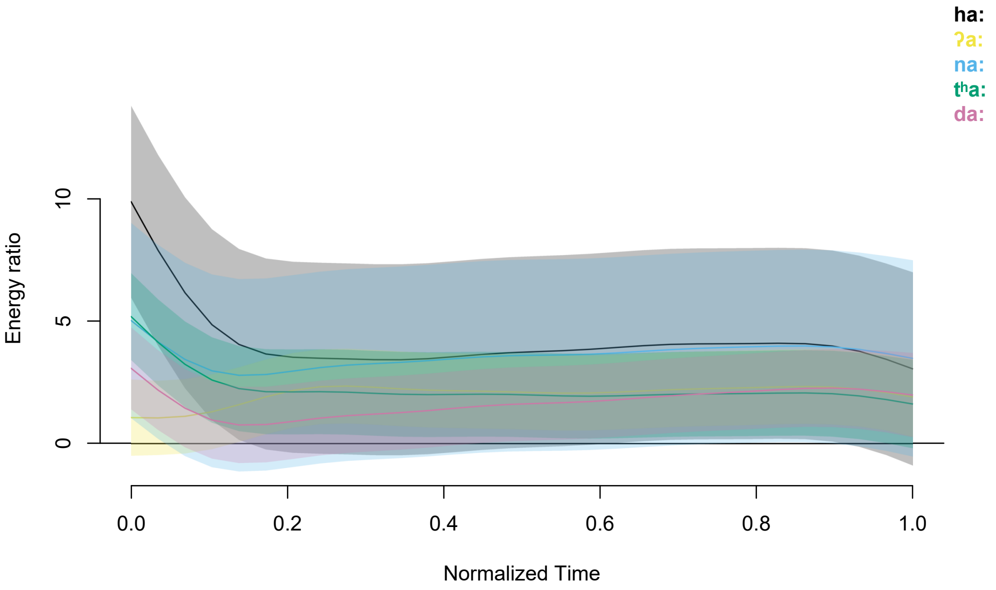


Figure 4.59: GAMM-predicted energy ratio measured during all vowels /a/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

to produce /tʰaː/ with similarly high energy ratio during the initial 10% of the vowel, while TF2 and TF3 are predicted to produce /naː/ with initially high energy ratio as well. Only TF3 maintains higher energy ratio for /haː/ and /naː/ throughout the entire vowel compared to other syllables.

Figure 4.61 shows mean energy ratio during the vowel of all /ɛː/-nucleus syllables. Results show that /hɛː/ and /tʰɛː/ are produced with greater energy ratio during the initial 10% of the vowel compared to /ʔɛː/. During the remainder of the vowel all energy ratio values are similar across syllables. Plots of individual speaker data all show /hɛː/ and /tʰɛː/ produced with greater initial energy ratio compared to /ʔɛː/ (Figure 4.62).

Plots of GAMM-predicted energy ratio show that /hɛː/ and /tʰɛː/ are predicted to be produced with greater initial energy ratio than /ʔɛː/ (Figure 4.63). Similarly, individual predictions show the same trend in Figure 4.64, although TF3 maintains a relatively higher energy ratio during /hɛː/ throughout the duration of the vowel.

Figure 4.65 shows mean energy ratio during the vowels of syllables with /ɔ/. We observe

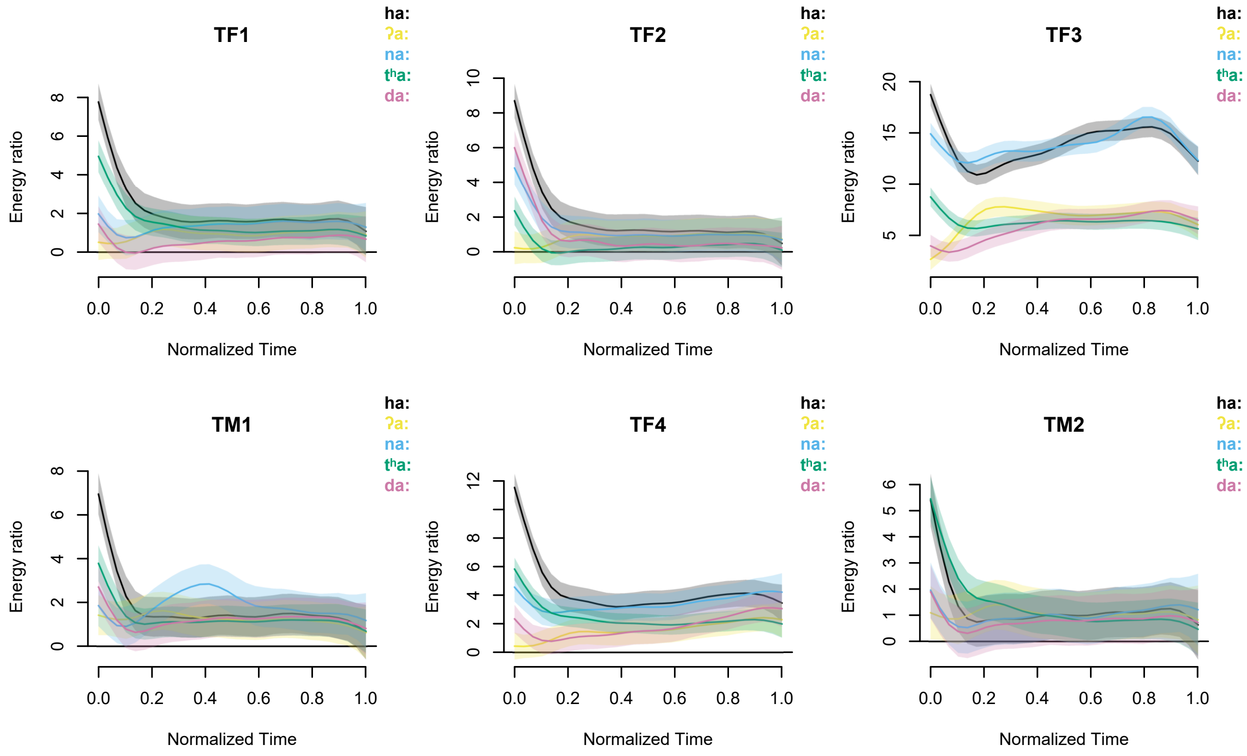


Figure 4.60: GAMM-predicted energy ratio measured during all vowels /ɛ/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

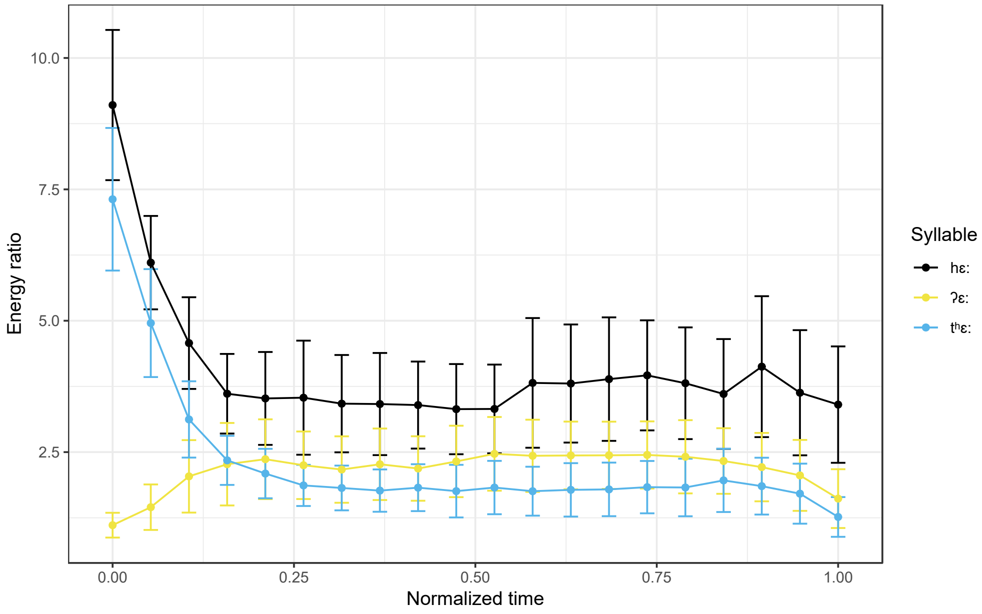


Figure 4.61: Mean energy ratio measured during all vowels /ɛ/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

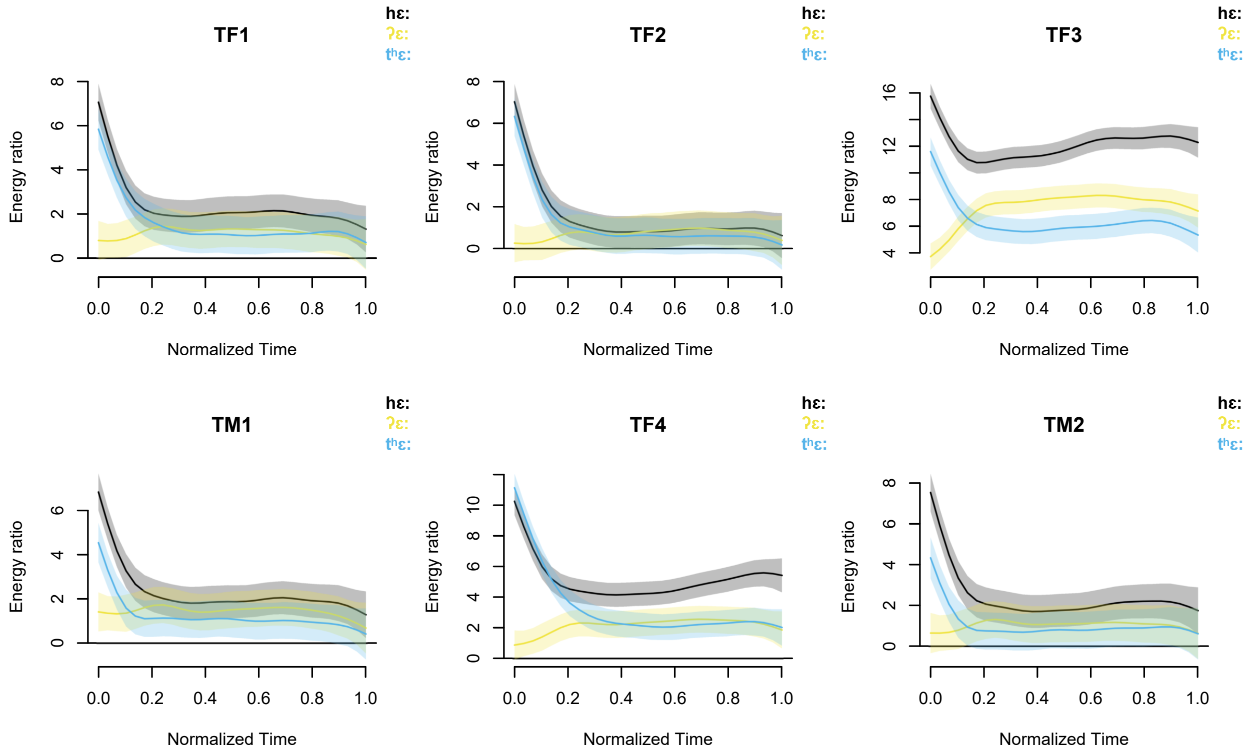


Figure 4.62: Mean energy ratio measured during all vowels /ɛ/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

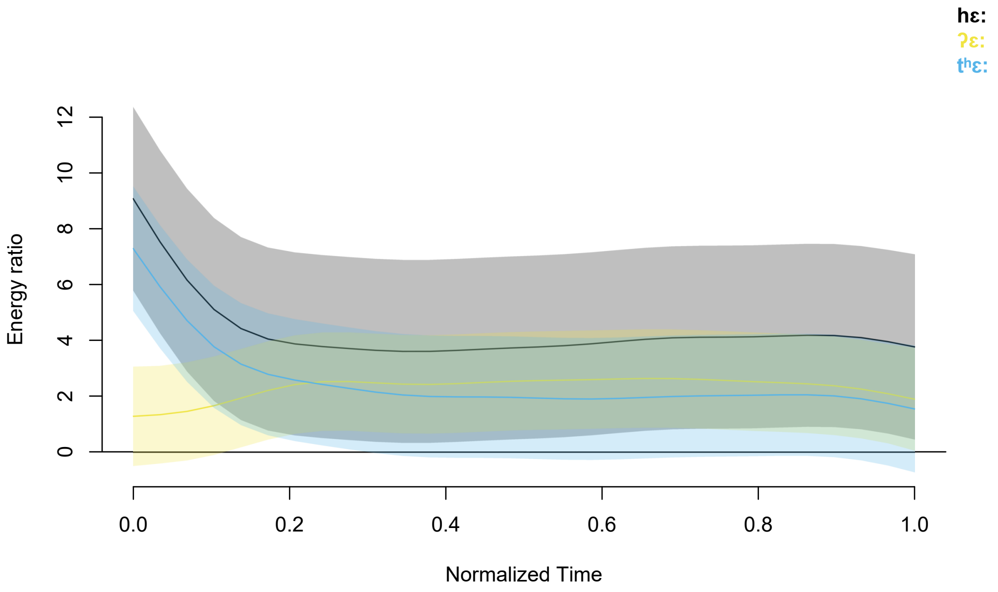


Figure 4.63: GAMM-predicted energy ratio measured during all vowels /ɛ/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

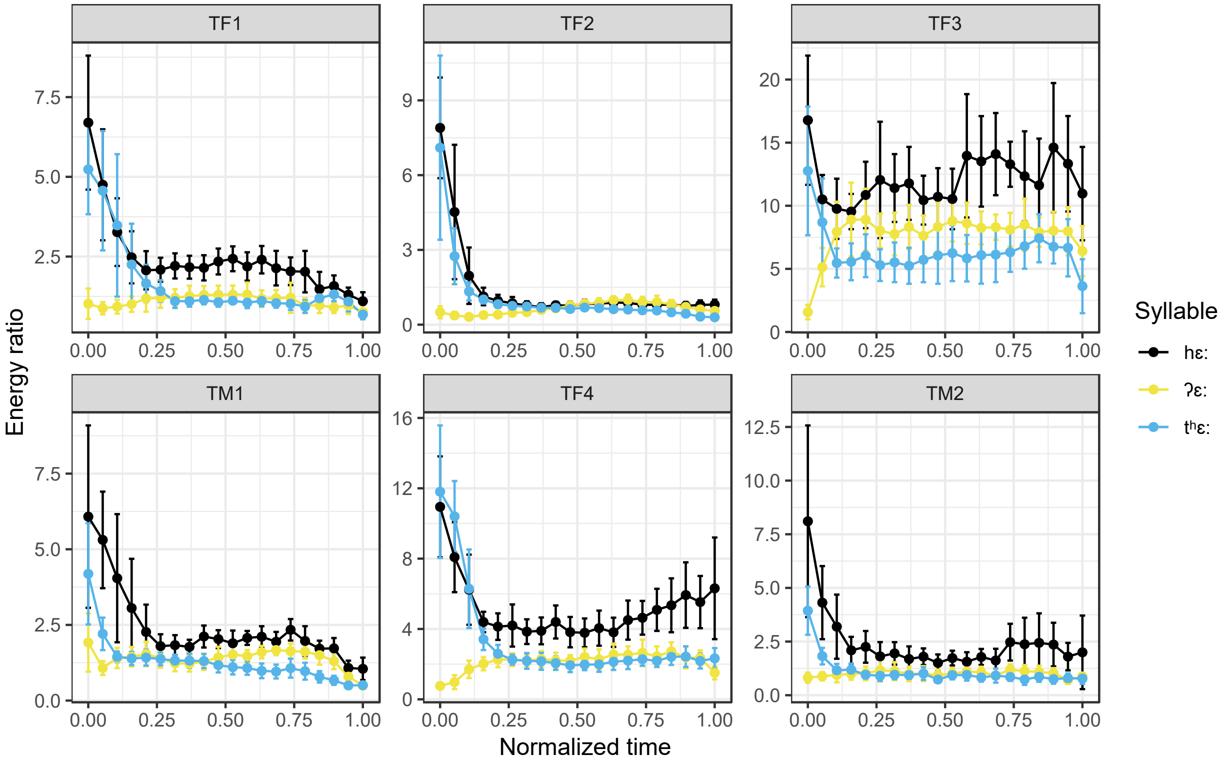


Figure 4.64: GAMM-predicted energy ratio measured during all vowels /ɛ/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

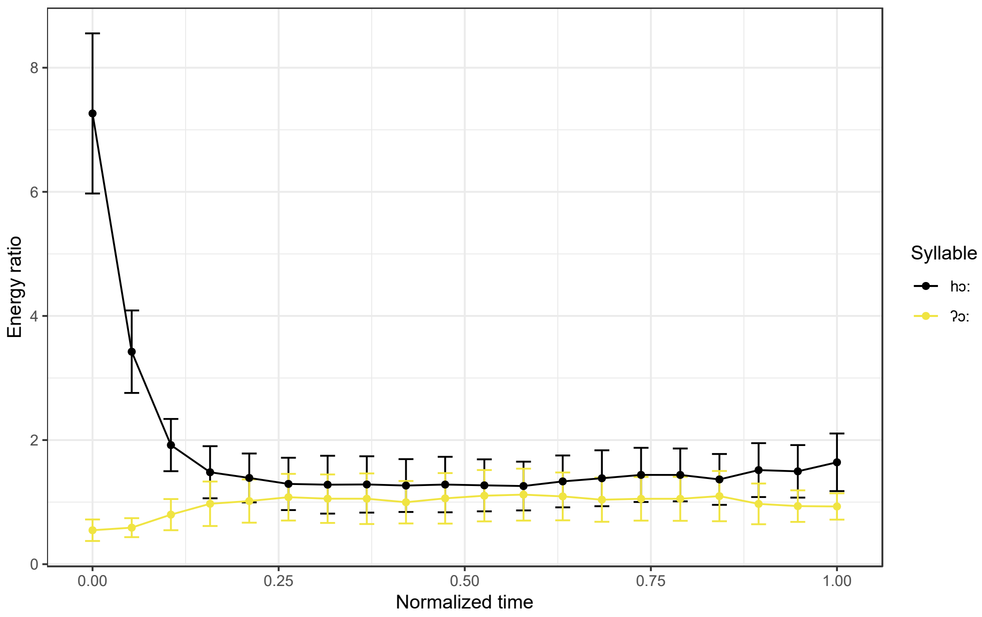


Figure 4.65: Mean energy ratio measured during all vowels /ɔ/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

that /hɔː/ is produced with greater energy ratio during the initial 10% of the vowel compared to /ʔɔː/, and that energy ratio is similar between the two vowels thereafter. A plot of individual speaker data mirrors this observation in every speaker (Figure 4.66).

Similarly, plots of GAMM-predicted energy ratio show that /hɔː/ is predicted to be produced with higher energy ratio than /ʔɔː/ during the initial 10% of the vowel before converging (Figure 4.67). Plots of individual GAMM-predicted values also mirror this observation in every speaker (Figure 4.68).

Finally, Figure 4.69 shows mean energy ratio during all vowels /eː/. We observe again that vowels preceded by /h/ and /t^{h}/ exhibit energy ratios that are higher than the vowel preceded by /ʔ/ during the initial 10% of the vowel. Individual speaker plots generally mirror this observation, although the difference between the vowels of /t^{h}eː/ and /ʔeː/ are diminished in speakers TM1 and TM2.

Similarly, GAMM-predicted energy ratio shows that /heː/ and /t^{h}eː/ are predicted to exhibit higher energy ratio during the initial 10% of the vowel (Figure 4.71). The GAMM-

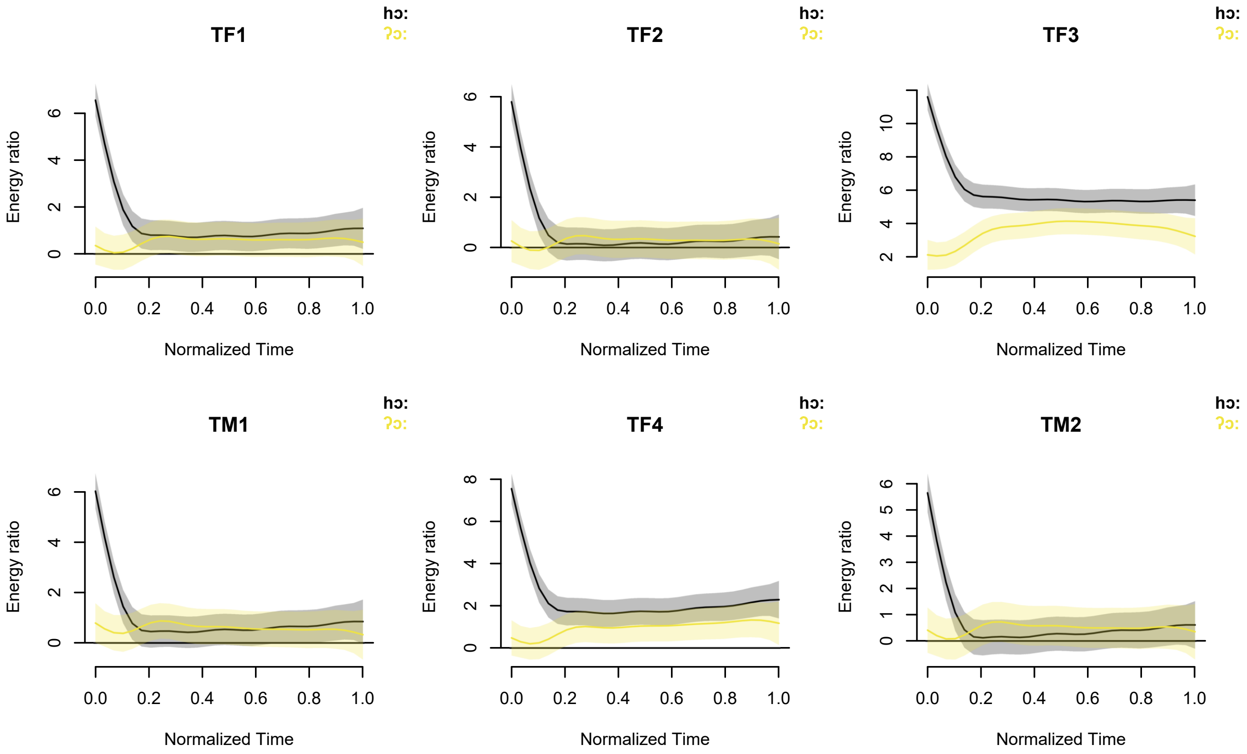


Figure 4.66: Mean energy ratio measured during all vowels /ɔ/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

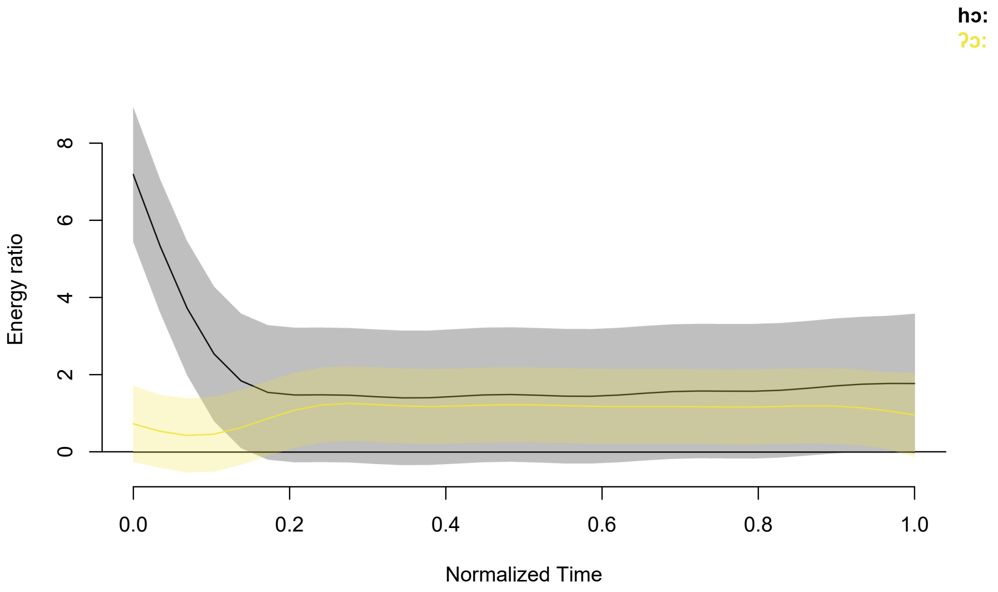


Figure 4.67: GAMM-predicted energy ratio measured during all vowels /ɔ/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

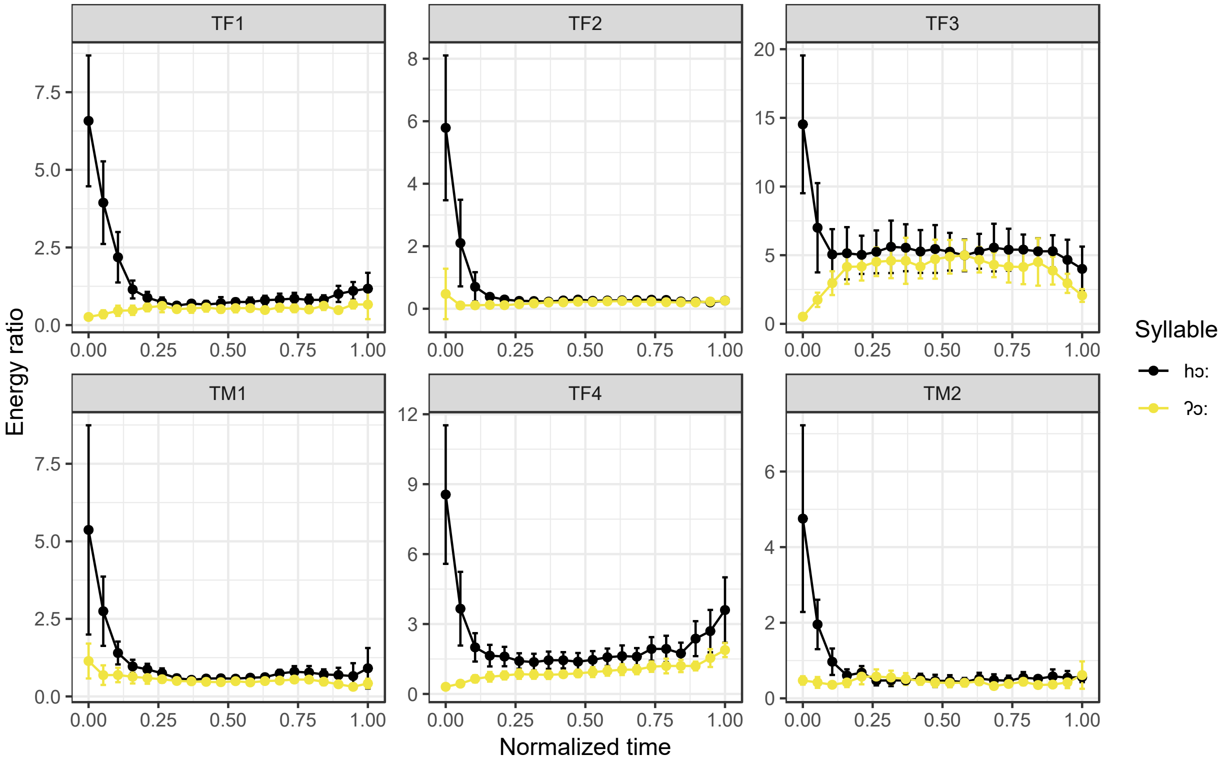


Figure 4.68: GAMM-predicted energy ratio measured during all vowels /ɔ/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

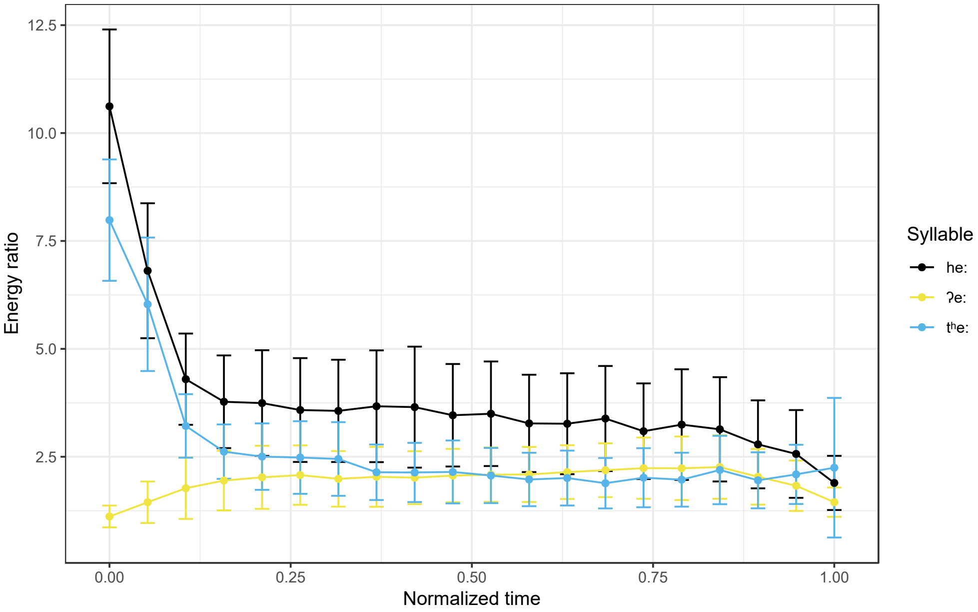


Figure 4.69: Mean energy ratio measured during all vowels /e/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

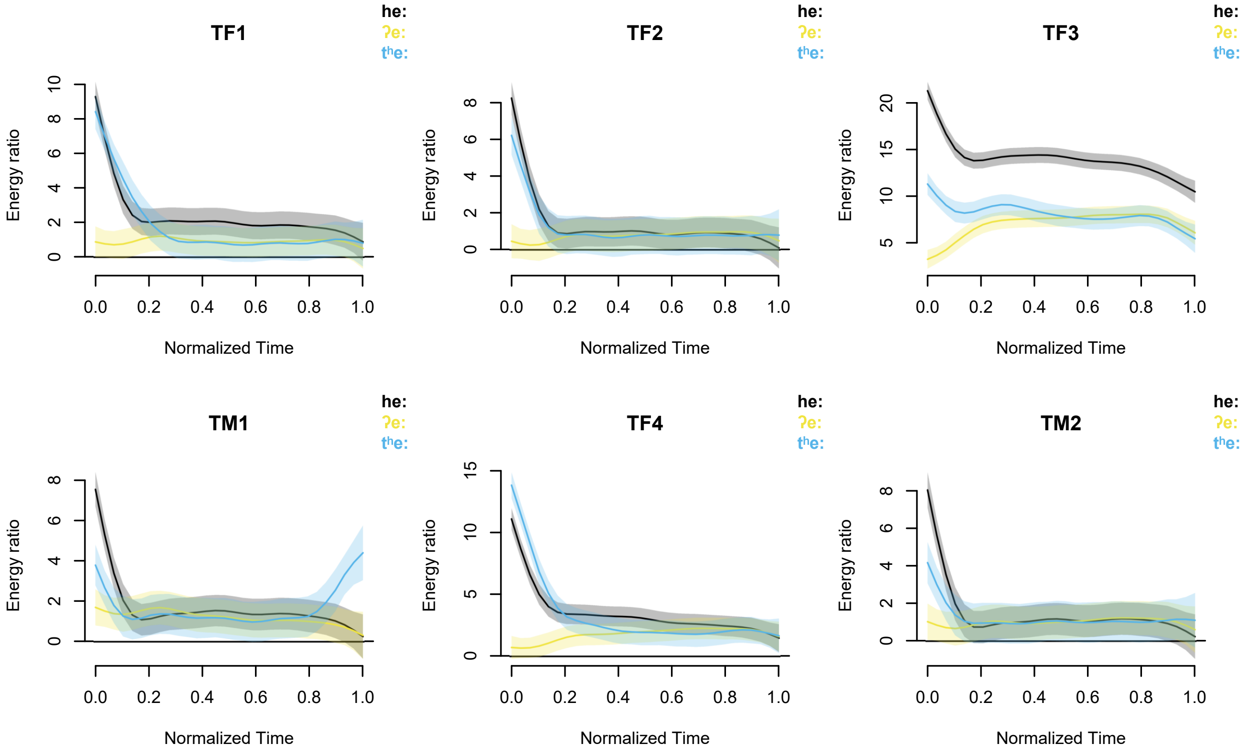


Figure 4.70: Mean energy ratio measured during all vowels /e/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

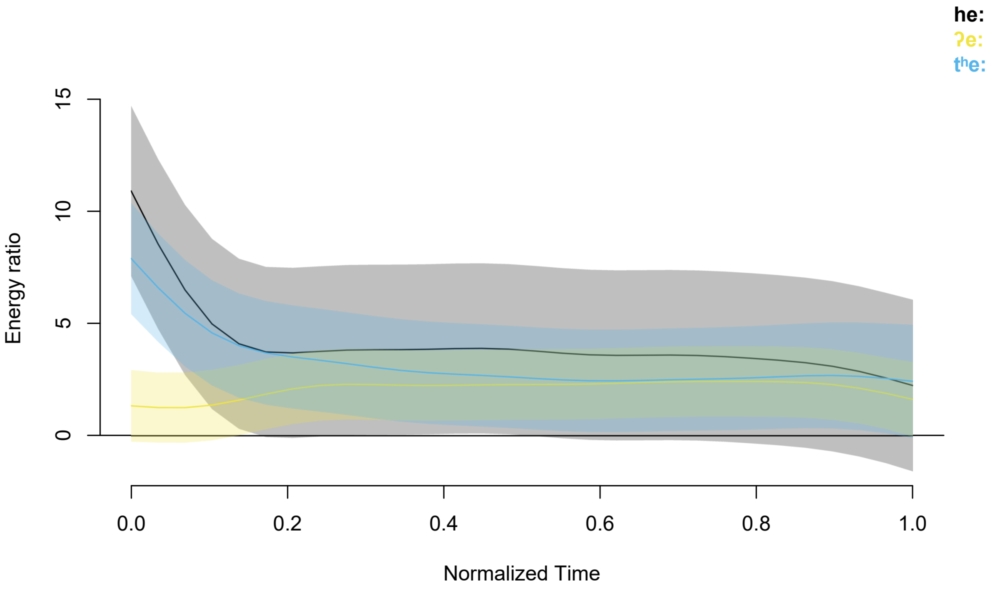


Figure 4.71: GAMM-predicted energy ratio measured during all vowels /e/. Confidence interval bars are based on the standard error. All speakers' data are combined. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

predicted energy ratio plots of individual speakers generally mirrors this observation as well, although the productions of /t^{h}eː/ and /ʔeː/ are predicted to overlap in TM1.

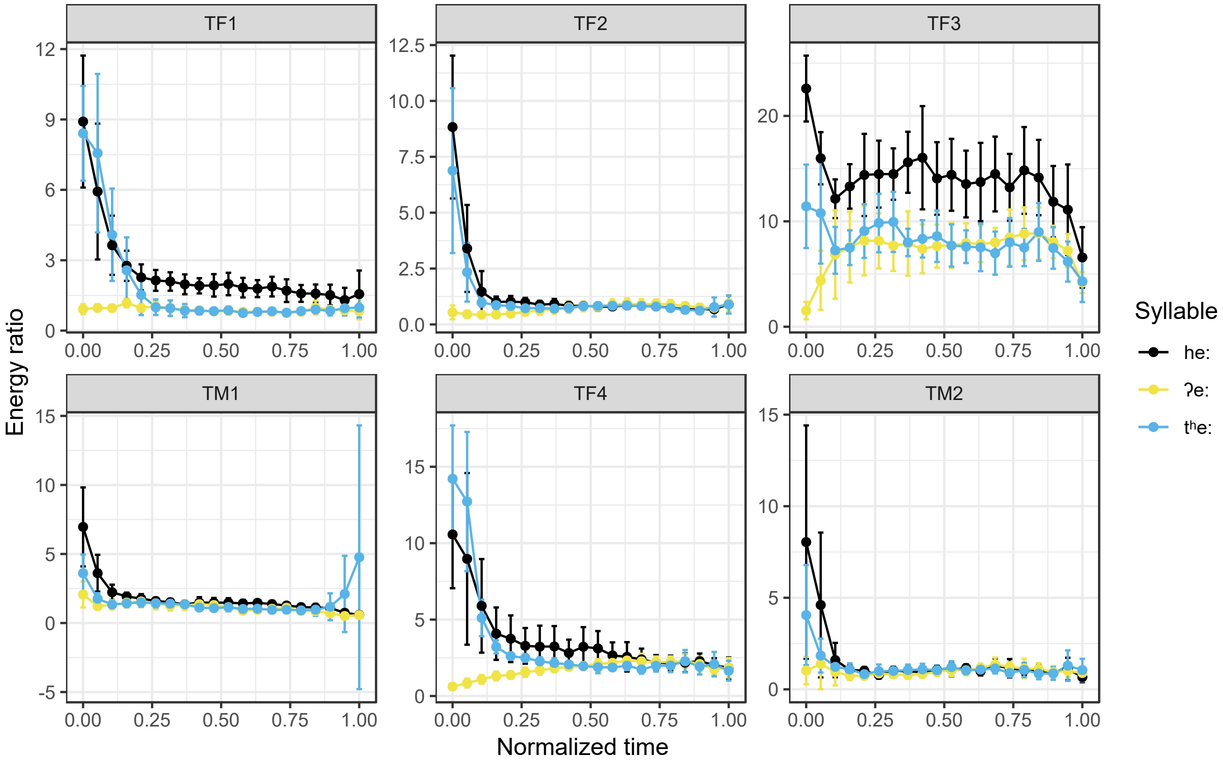


Figure 4.72: GAMM-predicted energy ratio measured during all vowels /e/. Confidence interval bars are based on the standard error. Energy ratio was calculated by dividing low and high spectral energy (0–320 Hz / 320–5360 Hz).

Summary

A summary of EGG results can be found in Table 4.12. Results show that overall, vowels that are preceded by onsets /h/ and /tʰ/ are produced with larger open quotient than vowels after /d/ and /ʔ/ during the initial 10% of the vowel. This briefly elevated open quotient is likely attributed to coarticulatory breathy phonation during the transition between glottal frication/aspiration of the preceding consonant and the following modally-voiced vowel. We also observe no difference in open quotient based on vowel height, i.e. EGG does not vary between low and mid-high vowels. We observe some individual differences in productions of vowels following /ʔ/. Some speakers produce the initial portion of vowels after /ʔ/ with significantly low open quotient compared to other syllables, while other speakers produce the entire vowel with similar open quotient as /daː/ and /naː/.

Contrast	EGG
Nucleus /a/	haː/tʰaː > naː/ʔaː/daː*
Onset /h/	n.s.
Onset /ʔ/	haː > ʔaː/ʔɛː/ʔɔː/ʔeː*
Onset /tʰ/	n.s.

Table 4.12: EGG results summary. A star indicates that the difference recorded between the syllable and the previous syllable only occur for a portion of the segment rather than the entire duration of the segment. Non-significantly different values is indicated by n.s.

A summary of energy ratio results can be found in Table 4.13. In general, the initial 10% of the vowel of /haː/ exhibits higher energy ratio than other syllables with the nucleus /a/. For all other vowel nuclei /ɛ/, /ɔ/, and /e/, vowels following /h/ and /tʰ/ are produced with similarly high energy ratio compared to vowels after /ʔ/.

Contrast	Energy ratio
Nucleus /a/	haː > naː/ʔaː/tʰaː/daː*†
Nucleus /ɛ/	hɛː/tʰɛː > ʔɛː
Nucleus /ɔ/	hɔː > ʔɔː
Nucleus /e/	heː/tʰɛeː > ʔeː

Table 4.13: Energy ratio results summary. A star indicates that the difference recorded between the syllable and the previous syllable only occur for a portion of the segment rather than the entire duration of the segment. Non-significantly different values is indicated by n.s. A †indicates that a discrepancy exists between the mean/confidence interval and GAMM-predicted data results.

4.6 Discussion

In this study we analyzed both nasal airflow and laryngeal open quotient in order to characterize degree of nasalization and phonation quality during syllables predicted to undergo spontaneous nasalization in Thai. In the overview below, we summarize findings for nasal airflow and EGG during the same speech conditions.

Overall, we observed high nasal airflow during all /h/ consonants that continued into the initial half of the following vowel. This nasal airflow was highest during the syllable /haː/, although an analysis of proportional nasal airflow revealed /haː/ to have similar or slightly lower proportional nasal airflow as /naː/. Furthermore all consonants /h/ were produced with greater nasal airflow then their following vowels. Nasal airflow generally did not vary based on vowel height, with the exception that /haː/ was produced with slightly higher mean nasal airflow during the consonant and vowel compared to /hɛː/, /hɔː/, and /heː/. Syllables beginning with /ʔ/ were produced with little to no nasal airflow during the consonant and little nasal airflow during the following vowel. It was predicted that /ʔ/ would be produced with minimal nasal airflow because glottal closure is expected to result in minimal supraglottal airflow. However, it is surprising that we observed minimal nasal airflow during the vowel after /ʔ/ because of reports that low and mid-low vowels after /ʔ/ are nasal (Cooke, 1989; Matisoff, 1975; Noss, 1964). The six speakers in the current airflow study did not produce vowels after /ʔ/ with any appreciable nasalization.

Maximum integrated nasal airflow occurs during the consonant for all syllables, regardless of degree of raw nasal airflow, except for syllables with /ʔ/. Although the maximum of nasal airflow occurs during the consonant of /daː/ and /tʰaː/ most of the time, recall that nasal airflow is near zero during these contexts. Increased nasal airflow compared to the following vowel may be a property of the plosive. In the case of /d/, the speaker may lower the velum somewhat as compensation to maintain voicing during the stop. Speakers have been known to perform compensatory articulations to expand the oropharyngeal cavity and maintain a pressure differential across the glottis (Ladefoged and Maddieson, 1996). For /tʰ/, increased nasal flow compared to the following vowel may be a result of the aspirated burst; the velum may begin to lower in anticipation for the following vowel, thus allowing an escape of some nasal flow. Further measures of VPO are needed to verify this possibility. For syllables with /ʔ/, maximum nasal airflow is more likely to occur during the vowel compared to the consonant. Given lowered supraglottal air flow due to an adducted glottis, this result is expected.

According to the the VPU explanation for spontaneous nasalization, the source of vowel nasalization in Thai is the consonant. Our results support this hypothesis. There is usually greater nasal airflow during /h/ within syllable /hV/ syllables. Based on our findings regarding greater nasal airflow during /h/ compared to the following vowel for all syllables, we reason that the onset consonant is likely the primary locus of nasalization that spreads to the following vowel. While we cannot assess the possibility of this pattern for syllables with /ʔ/, our results set the stage for an analysis that directly measures VPO during /ʔ/ and the following vowel.

We conclude that the the initial locus of spontaneous nasalization in Thai is the consonant for syllables beginning with /h/. VPU does not induce spontaneous vowel nasalization per se. Rather, it induces spontaneous glottal consonant nasalization that spreads. Given the relatively large degrees of nasal airflow during the consonant and following vowel of syllables beginning with /h/, we can conclude that the entire syllable /hV/ has undergone spontaneous nasalization. See Chapter 5 for a discussion of the perceptibility of nasalization during /hV/ syllables.

We observed relatively high open quotient (compared to /daː/ and /ʔaː/) during the initial 10% of all vowels following an /h/ onset. We observed similar open quotient during vowels following /h/ and /t^h/ that did not vary by vowel height. However, /t^h/ and the following vowels were produced with null or low nasal airflow. The low degree and short duration of the breathiness during vowels just after /h/ and /t^h/, both segments produced with voiceless turbulent frication, suggests that this breathiness is likely coarticulatory from the preceding consonant.

The results detailed above suggest that both the consonant and initial half of the vowel are nasalized for all syllables studied that begin with /h/, regardless of vowel height. Furthermore, these same syllables are produced with brief breathy phonation just after consonant release. However, this brief breathy transition is not greater than or longer in duration than other syllables with an aspirated consonant, /t^h/. One exploratory question we initially asked was whether or not Thai speakers take advantage of the acoustic similarities between breathiness and nasalization to produce spontaneously nasalized vowels with increased breathiness or breathiness that lasts longer in duration. These results show that speakers do not produce spontaneously nasalized vowels after /h/ with greater breathiness than other non-nasalized, contextually partially breathy vowels, i.e. those following /t^h/. In Chapter 5 Section 5.2 we discuss the implications of coarticulatory breathiness after /h/ and the minor role it may play in Thai spontaneous nasalization.

Our analysis of energy ratio (0–320 Hz / 320–5360 Hz) revealed elevated values during the initial 10% of the vowel for all syllable beginning with /h/ and /t^h/. We also observed increased energy ratio during the vowel after /n/ onset for three speakers. These results suggest that the increased energy ratio observed in the current data set is likely attributed to coarticulatory breathiness from /h/ and /t^h/ at the start of the vowel. The fact that energy ratio is elevated briefly after an /n/ onset also suggests that the presence of nasalization also influences this measure for three speakers. Recall that while the energy ratio measure was originally proposed to detect nasal coupling during consonants (Carignan, 2017; Pruthi *et al.*, 2007), it is theoretically possible for this measure to detect both nasalization and breathiness; this measure assesses the relative energy at low frequencies compared to high frequencies. Both nasalization and breathiness are produced with high spectral tilt, i.e. higher energy at low frequencies (Stevens, 2000; Chen, 1995; Styler, 2017; Garellek, 2014; Garellek and Keating, 2011; Wayland and Jongman, 2003). The fact that energy ratio was not high for all speakers' productions of /naː/ while nasal airflow was observed to be high for these speakers during and earlier recording session (Section 4.4.1) suggests that this measure did not successfully detect vowel nasalization in all speakers. One reason for this failure may be the influence of F1 on spectral energy below 320 Hz. While we only compared vowels of the same height in order to control for variation in F1, individual differences in F1 position or bandwidth may effect the relative energy at low frequencies. More research is needed to assess the efficacy of a full-spectrum measure of spectral tilt like the one used in the present study.

Chapter 5

GENERAL DISCUSSION

This dissertation undertook a detailed investigation of the production of spontaneous nasalization in Thai using ultra-fast MRI, aerodynamics, and EGG. In these studies we analyzed velopharyngeal opening, nasal airflow, and laryngeal open quotient in order to characterize degree of physiological nasalization and breathiness during syllables predicted to undergo spontaneous nasalization in Thai. Below we review and integrate the major findings of these studies and discuss implications for the role of both the larynx and velum in spontaneous nasalization. We situate these findings within the literature of spontaneous nasalization in Thai and other languages.

5.1 The velum

Major findings

See Table 5.1 for a list of major findings from the MRI and aerodynamic studies. The first major finding from the MRI and aerodynamic studies is that syllables with /h/ onset are produced with greater physiological and aerodynamic nasalization than syllables with /ʔ/ onset. In the MRI study we observed that all speakers produced vowels after /h/ with greater VPO than predictably non-nasal conditions, whereas half of the speakers also produced vowels after /ʔ/ with increased VPO (Table 5.1). Furthermore, vowels after /ʔ/ that were produced with increased VPO exhibited lower VPO than (nasalized) vowels after /h/. In the aerodynamic study, we observed increased nasal airflow during the /h/ onset and the initial half of the following vowel. Not only was nasal airflow higher at any given time point during the onset consonant, the integrated sum of nasal airflow during the onset /h/ was almost always greater than during the following vowel. This finding suggests that the /h/ onset is not only associated with spontaneous vowel nasalization, it is the locus of nasalization. The vowel nuclei of syllables with /h/ onset undergo coarticulatory nasalization as a result of their proximity to /h/. Properties of /h/ that may lead to this phenomenon are discussed in further detail below and in Section 5.2.

	Prediction	VPO (4 speakers)	Nasal airflow (6 speakers)
Vowel	Low > mid-low > mid-high/high	Variable	Variable
Glottal	Glottal > non-glottal	/h/ > non-glottal (4/4 speakers); /ʔ/ > non-glottal (2/4 speakers)	only /h/* > non-glottal
/h/ vs. /ʔ/	/h/ > /ʔ/	/h/ > /ʔ/	/h/ > /ʔ/*

Table 5.1: Major findings of the MRI and nasal airflow experiments. A star* indicates that the difference between syllables only occurs for a portion of the segment rather than the entire duration of the segment.

We observed no nasal airflow during the consonant /ʔ/. This finding was expected because /ʔ/ is produced with diminished or zero transglottal airflow due to glottal closure. We also observed minimal nasal airflow during the vowels following /ʔ/; these vowels exhibited similarly low nasal airflow as predictably non-nasal vowels such as /t^{h} aː/ and /daː/. While two out of four speakers of the MRI study produced vowels after /ʔ/ with increased VPO, our airflow data suggest that no speakers from the aerodynamic study produced vowels after /ʔ/ with increased VPO. This suggests that spontaneous nasalization after /ʔ/ in Thai is rare. Earlier literature on spontaneous nasalization in Thai reported that low/mid-low vowels sounded nasal after both /h/ and /ʔ/ (Matisoff, 1975; Cooke, 1989). We reasoned that the cause of nasalization was velopharyngeal underspecification of glottal consonants (Section 2). While we observe that the VPO is underspecified for /h/, our data suggests that VPO is in fact specified for /ʔ/.

This finding has an important implication for spontaneous nasalization in Thai. Low and mid-low vowels after /h/ usually nasalize, while vowels after /ʔ/ rarely nasalize. We must therefore modify our initial criteria for spontaneous nasalization in Thai. Initially we broadly predicted that vowels after glottal consonants nasalize to some degree due to the fact that the velum need not be elevated, as during buccal obstruents like /t/. However, we now understand that an onset with a glottal place of articulation (paired with a low/mid-low vowel in the nucleus) is not sufficient to induce nasalization. Obviously, the primary difference between /h/ and /ʔ/ is that the former is a fricative while the latter is a stop. Voiceless frication, in addition to glottal place of articulation, may be the optimal environment for spontaneous nasalization in Thai. Here we define frication as voiceless aerodynamic turbulence channeled through the vocal tract; it is associated with segments including fricatives and aspiration (see Section 2 for a review). In the following paragraphs we step through articulatory evidence associated with this problem and speculate why hV (glottal fricative followed by a low/mid-low vowel) is more conducive to spontaneous nasalization than /ʔ/ in onset position.

We observed that nasal airflow was highest during the consonant /h/, even higher than the nasal stop /n/. The vowel just after /h/ was also produced with the greatest nasal airflow, followed by the vowel after /n/. However, /naː/ was produced with greater proportional nasal airflow than /haː/. This elevated airflow associated with /h/ is indicative of large VPO. Our findings also show that nasal airflow is larger during the consonant /h/ than following vowel. It is important to note that while elevated nasal airflow during /h/ suggests large VPO, it is unlikely that this voiceless nasalized frication contributes to the percept of nasality. That is, the primary perceptual correlates of nasalization involve the introduction of nasal resonances and antiresonances and thus require voicing. Nasal fricatives are typologically rare

(Ohala *et al.*, 1998; Shosted, 2006). The reason is that fricatives are generally poor targets for nasalization both aerodynamically and perceptually. Most fricatives are aerodynamically antagonistic to nasalization because air leakage through an open velopharyngeal port inhibits pressure build up at the place of fricative articulation (Ohala, 1975). The acoustics of buccal fricatives are are usually adversely affected by nasal coupling because of nasal air leakage. However, pharyngeal and glottal fricatives may not be affected in the same way as buccal fricatives because their place of primary constriction is posterior to the velum. Ohala (1975) writes that "[N]oise produced by voiceless glottal and pharyngeal obstruents is so diffuse, so low in intensity, and with higher frequencies dominating in the spectrum that oral-nasal coupling would have little acoustic effect on it" (pg. 301). Shosted (2006) argues that "while pharyngeal and glottal nasalization are physiological possibilities, these are not likely to be adopted in any language due to problems with perceptibility" (pg. 20). Therefore, it is unlikely that physiological nasalization affects the perception of the glottal fricative /h/; rather the lowered velar position continues into the following vowel, causing the vowel to reportedly sound nasal. Indeed, previous reports of spontaneous nasalization in Thai observe that vowels sound nasal nasalized after glottal consonants; no researchers report that the glottal consonant also sounds nasal (Cooke, 1989; Matisoff, 1975; Noss, 1964).

We analyzed productions of two consonants that are accompanied with voiceless turbulence through frication or aspiration in the present study: /h/ and /t^h/. These two consonants exhibit turbulent aspiration at the release of the consonant (/t^h/) or during the consonant (/h/), which we predicted to trigger coarticulatory breathiness. Both were produced with high oral airflow and similar degrees of slightly elevated open quotient just after consonant release (see Section 5.2); but only /h/ (not the aspiration at the release of /t^h/) was produced with increased nasal airflow during and after the consonant. If both /t^h/ and /h/ have some phonetic similarity, i.e. voiceless turbulence at the glottis that induces breathiness, why is only /h/ nasalized? During the voiceless aspirated stop, aspiration noise follows the frication noise burst (Klatt, 1973). Abramson and Whalen (2017) describe /h/

as “...essentially a vowel or diphthong excited for part of its length by turbulence through a glottal opening yielding a voicing lag.” They further note that the acoustic distinction between fricatives and aspiration is difficult to determine due to their acoustic and aerodynamic similarity. This makes detection of the boundary between fricative and aspiration during aspirated fricatives such as /s^h/ difficult to perform (Abramson and Whalen, 2017). Despite the phonetic similarities between /h/ and aspiration, we would not prediction an aspirated stop like /t^h/ to nasalize. Because /t^h/ is a buccal stop, we would expect that VPO is specified and that the velum is raised for the [t] closure. While velum position may be unspecified for aspiration /$^\text{h}$/ as it is for /h/, it would likely be raised because of the [t] closure. Therefore, while both segments exhibit similar degrees of coarticulatory breathiness that is similar to nasalization acoustically, only the glottal consonant /h/ may be nasalized because it is underspecified for velum position. Whether or not breathiness after /h/ contributes to the perception of nasalization in Thai is matter that requires further consideration; we discuss this topic in more detail below in Section 5.2.

Another major finding is differences in VPO and nasal airflow based on vowel height. In the MRI study we observed variable VPO that sometimes followed the pattern low > mid-low > mid-high/high (approximately half of observations), although we also observed a great degree of interspeaker variability. In the aerodynamic study, we observed little difference in nasal airflow, in both proportional and nasal airflow data, based on vowel height after glottal consonants. We also observed no variation in nasal airflow for non-nasalized vowels of different heights after /t^h/. We only observed some variability in nasal airflow at different vowel heights after /h/. /haː/ is produced with greater raw and proportional nasal airflow than other vowels during the initial half of the vowel in four out of six speakers. Some speakers also vary in producing /hɛː/ and /hɔː/ with similar nasal airflow to /haː/. We often observe elevated nasal airflow during /heː/ compared to predictably non-nasal syllables with /t^h/ and /d/ onsets. We conclude that VPO minimally varies based on vowel height in Thai, and that any variation is subject to individual speaker variability. Matisoff (1975) and

Cooke (1989) previously reported that only the low/mid-low vowels of Thai underwent vowel nasalization after glottal consonants. Our findings suggest that potentially all vowels after /h/ are capable of undergoing some vowel nasalization in central Thai. This suggests that, in terms of vowel height-based VPO differences, spontaneous vowel nasalization in Central Thai may now be similar to Northeastern Thai, a dialect where all vowels were reported to nasalize after glottal consonants (Matisoff, 1975).

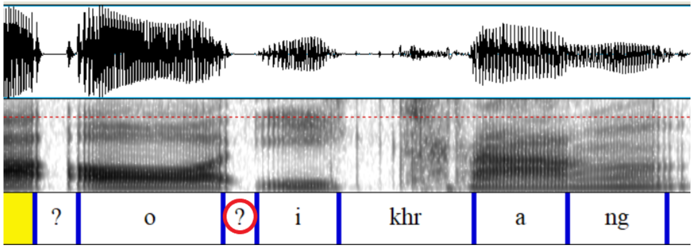
Other considerations

Returning to our comparison of MRI and aerodynamic data, there is an important divergence between temporal trends observed in these data sets. In the MRI study, when a speaker produced a vowel with elevated VPO, the speaker typically produced the entire vowel with mostly level, elevated VPO. It does not usually fall or rise during the vowel. Indeed, the general additive mixed models revealed no or minimal influence of the factor normalized time on VPO. However, normalized time is an important factor during nasal airflow productions. When a syllable was produced with high nasal airflow (/h/ and /n/-onset syllables), the level of nasal airflow changed over time. The nasal airflow remained high during the onset consonant and then fell dramatically during the following vowel, until nasal airflow was similar to non-nasal vowels during the latter half of the vowel. If a perfect correlation existed between VPO and nasal airflow, we would expect that if VPO were consistently large and level in every speaker in the MRI study, that nasal airflow would be similarly consistently high and level. However, we observe falling nasal airflow.

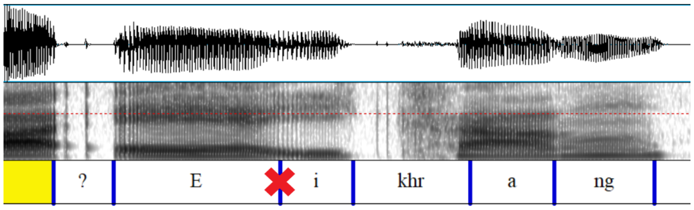
One possibility is that the observed differences in timing are simply due to the fact that different speakers were recorded in both studies. It is possible that, if invited to return for an MRI study, the six aerodynamic study speakers would produce VPO that correlated with their productions during the aerodynamic study, i.e. VPO would begin high during vowels after /h/ and gradually fall. However, given the consistent finding in VPO and nasal airflow of vowels after /h/ for all speakers, this possibility is unlikely. Rather, the explanation for this difference in timing is likely articulatory and aerodynamic in origin. During the production of a phrase, amplitude, pitch, and air pressure generally fall over time (Hixon *et al.*, 2014). It is possible that as airflow decreases during the production of the target syllable, the recorded nasal and oral air pressure decreases over time, despite exhibiting large, stable VPO.

Unlike nasal airflow, VPO remains level from the start of the vowel to the start of the following syllable. One reason for this might be that the word immediately following the target syllable is /iːk˧kʰraŋ˥/ 'again'. It is possible that this word begins with a /ʔ/ that is underspecified for VPO, as [ʔiːk˧kʰraŋ˥]. Glottal stop is usually predictable in Thai; it is present before vocalic onsets that are in stressed/prominent positions but absent in unstressed positions (Abramson, 1962; Harris, 2001; Noss, 1964). Prominence is assigned in Thai based on status as a content word and semantic importance; stress within polysyllabic words typically falls on the last syllable (Anawin, 1998). The most prominent word in the test sentence, /pʰuːt˥˩ kʰam˧ waː˥˩ "__" ʔiːk˧ kʰraŋ˥ / "Say the word '__' again", is the target word that varies across repetitions. Indeed, this word sounds the most prominent. The word /iːk˧kʰraŋ˥/ is not in the most prominent position of the sentence, and the second syllable /kʰraŋ˥/ is more prominent following Thai stress rules. A post-hoc inspection of the Thai speakers' acoustic data revealed that all speakers intermittently produce /iːk˧kʰraŋ˥/ as [ʔiːk˧kʰraŋ˥]. However, we observe much intraspeaker variability across tokens; sometimes the glottal stop is dropped and only brief creakiness is present, while other tokens are produced with a glottal stop at the beginning of the word. See Figures 5.1b and 5.1a for examples of two repetitions, one where /ʔ/ is produced during /iːk˧kʰraŋ˥/ and one where it is not produced in the same speaker.

Because the syllable /iːk˧/ may begin with a glottal stop, it may be underspecified for VPO. Note that we do not observe a sharp increase in VPO for syllables that are not already produced with large VPO into the following glottal stop. This reveals a crucial characteristic of velopharyngeal underspecification: VPO is optional, therefore unpredictable. When the velum is already lowered, it can continue without reversal if the following syllable begins with



(a) Example of /ʔ/ during the onset of the word /(ʔ)iːk˨kʰraŋ˥/ ‘again’; from the token /pʰuːt˥˩ kʰam˧ waː˥˩ ‘ ʔɔː ’ ʔiːk˨ kʰraŋ˥/ “Say the word ‘crowd together’ again”. This token was produced by TF2. The circle indicates the glottal stop that was produced at the start of /(ʔ)iːk˨kʰraŋ˥/.



(b) Example of absent /ʔ/ at the onset of the word /(ʔ)iːk˨kʰraŋ˥/ ‘again’; from the token /pʰuːt˥˩ kʰam˧ waː˥˩ ‘ ʔɛː ’ ʔiːk˨ kʰraŋ˥/ “Say the word ‘small/young’ again”. This token was produced by TF2. The X indicates the spot where the glottal stop was omitted.

Figure 5.1: Glottal stop production and elision

a glottal stop. It is also possible that a vocalic onset is also not specified for VPO in Thai. Contrast this observation with productions of the target syllables that begin with a /ʔ/ in the MRI study. Recall that the target syllable is in a stressed phrase-medial position in the carrier phrase. If the stressed syllable begins with a /ʔ/, speakers will likely not produce the consonant or following vowel with increased VPO. However, if the prominent syllable begins with /h/ and the syllable undergoes spontaneous nasalization, the MRI data show that VPO will remain high into the start of the following unstressed syllable /(ʔ)iːk˧/. If /ʔ/ is usually not nasalized in Thai, why would it nasalize in a context following a spontaneously nasalized /h/-onset syllable? Stress/prominence may have an effect on nasalization in this case. Cho *et al.* (2017) found that vowels in phrase-final, non-focused words were produced with more coarticulatory nasalization than in phrase-initial, focused words in English. It is possible that prosodic context may also affect spontaneous nasalization in Thai. The glottal stop of the unstressed syllable [ʔiːk˧/] may undergo nasalization because it is in an unstressed phrasal position and is influenced by the previous /h/-onset syllable. Researchers have found that glottal consonants are often transparent to nasalization processes (Walker, 2000; Blevins and Garrett, 1992). Our findings suggests that /ʔ/ is usually specified for VPO in stressed position but is not specified in unstressed position and may nasalize when the surrounding context is already nasal. Further study is needed to assess the influence of prosody on spontaneous nasalization in Thai.

5.2 The larynx

Major findings

The first major finding from our analysis of open quotient is that all vowels after /h/ are produced with increased open quotient (physiological breathiness) for a very short duration just after consonant release compared to most syllables, including /ʔ/. See Table 5.2 for an overview of major EGG findings. The breathy interval lasts approximately 20–40 ms. The degree and duration of physiological breathiness does not vary based on vowel height. The syllables /haː, hɛː, hɔː, heː/ are produced with similar degrees of open quotient. Furthermore, the degree and duration of breathiness after /h/ is similar to the transient breathiness produced after /$\mathrm{t^h}$/ at all vowel heights. These results suggest that the breathiness observed after /h/ is likely attributed to coarticulation from the preceding consonant. If speakers used breathiness as an enhancement strategy to increase the percept of nasalization, we would expect to find a higher physiological breathiness than the non-nasalized context /$\mathrm{t^h}$/ or breathiness that lasted longer. In the Yi languages and Southern French, researchers found increased breathiness during nasalized vowels. They speculate that speakers augment nasalized vowels with increased breathiness as an enhancement strategy to make the vowel sound more nasal or through a process of misperception (Carignan, 2017; Garellek *et al.*, 2016). Because physiological breathiness in /$\mathrm{t^hV}$/ (non-nasalized) and /hV/ (nasalized) syllables is similar in Thai, our results show that Thai speakers likely do not augment nasalized vowels with increased physiological breathiness. Rather a brief duration of breathiness already exists during nasalized /h/-onset syllables just after consonant release due to coarticulation.

Whether this brief duration of breathy phonation after /h/ has any influence on the percept of nasalization in Thai needs further study. It is possible that the brief 20–40 ms of breathiness after /h/ may cause the vowel to sound more nasal, since breathiness and nasalization have some spectral similarity. This may help explain previous reports that vowels after /h/ sound more nasal than vowels after /ʔ/ in Thai (Cooke, 1989; Matisoff, 1975). However, given the result that vowels after /h/ are produced with more articulatory nasalization (VPO and nasal airflow) than vowels after /ʔ/, if the presence of breathiness after /h/ is a contributing factor to the perception of nasalization at all, it is a minor one.

	Open quotient prediction	EGG (6 speakers)
/h/ vs /ʔ/	/h/ > /ʔ/	/h/ > /ʔ/* (4/6)
/h/ vs /tʰ/	/h/ > /tʰ/ (exploratory)	n.s.
Nasalized glottal vs non-nasalized glottal	/haː/, /hɛː/, /hɔː/ > /heː/ (exploratory)	n.s.

Table 5.2: EGG major findings. A star* indicates that the difference between syllables only occurs for a portion of the segment rather than the entire duration of the segment. Non-significantly different values is indicated by n.s.

5.3 Implications and future research

The differences we observe in VPO and nasal airflow, along with the presence of breathiness after /h/, raise important theoretical questions about the definition of nasality. Is nasality increased velopharyngeal opening or is it the abstract percept of nasalization induced by certain spectral features? In the latter case, the presence of breathiness, similar to nasalization acoustically, may mistakenly lead to the percept of nasality. Some scholars view speech perception as the ultimate target of speech production, such that articulation culminates into a single acoustic product that is perceived and interpreted by the listener (Diehl *et al.*, 1990; Kingston, 1991; Ohala, 1996). This means that some articulatory variability can be tolerated across speakers, as long as the acoustic target is similar in the population. Recent work on nasality has found that many different articulations, including manipulation of the velopharyngeal port, the tongue, lips, and the pharynx, contribute to the acoustic output of nasal and nasalized vowels (Barlaz *et al.*, 2015, 2018; Carignan *et al.*, 2011, 2015; Bothorel *et al.*, 1986; Engwall *et al.*, 2006; Shosted *et al.*, 2012a; Shosted, 2015). The larynx has recently been added to the aforementioned list, as evidence has emerged that nasalized vowels are sometimes produced with increased breathiness (Garellek *et al.*, 2016; Carignan, 2017).

Our results show increased physiological nasalization and, to a less degree, breathiness during spontaneously nasalized /h/-onset syllables in Thai. Increased physiological nasalization and breathiness each have the potential to induce the percept of nasalization during Thai vowels after /h/. How do these articulatory factors fit into models of sound change? Ohala's phonetic typology of sound change includes hypocorrection, hypercorrection, and misperception-based mechanisms (Garrett and Johnson, 2013; Ohala, 1993a). Hypocorrection occurs when a listener undercorrects a coarticulated production and reinterprets it as phonologically intentional. Examples include umlaut, syncope, and final devoicing. These processes begin with coarticulation or other contextual variation based on syllable structure or stress. Listeners misinterpret these contextual differences as phonologically intended and sound change often follows. This thesis shows that in Thai, glottal /h/ is underspecified for VPO; VPU facilitates a strong tendency to consistently lower the velum during /h/. This lowered velum continues into the following vowel, causing it to reportedly sound nasal (Matisoff, 1975; Cooke, 1989). The critical question that remains is whether Thai listeners have come to interpret this spontaneous nasalization as phonological or as an integral part of the acoustic profile of /h/-onset syllables.

One way to test this would be to design a perception experiment where Thai listeners are presented with [hãː] and [haː], either synthesized or produced by a trained phonetician, with background noise applied. This paradigm can be used to test whether the addition of nasalization has an effect on Thai listeners' accuracy rate in identifying the word /haː˧/ 'to guffaw'. If Thai listeners are more accurate in identifying the word /haː/ during the nasalized production, we can conclude that the addition of nasalization is an important feature in the acoustic profile of /h/-onset syllables. This would mean that /h/-onset syllables have undergone a hypocorrection-type nasal sound change: VPU of /h/ induced VPO that remained during the following vowel, eventually becoming an important acoustic feature for /h/-onset syllables.

The third mechanism of Ohala's sound change typology is misperception-based changes.

With this type of change, listeners misinterpret an acoustic feature as belonging to a different articulatory source that has a similar acoustic output; listeners reinterpret the sound and produce it according to this reinterpretation. A common example of this kind of sound change is /θ/ variation with /f/ (Garrett and Johnson, 2013; Blevins, 2004; Ohala, 1993a). Both segments are acoustically similar; for example, many languages have been known to change /θ/ to /f/. In Thai, we observe increased physiological breathiness during vowels after all /h/ consonants. It is possible that listeners misinterpret this breathiness as nasalization due to the acoustic similarities between these two productions. If they misinterpreted the breathiness as nasalization, they may produced vowels after /h/ with increased VPO because they reinterpret nasalization as an important acoustic features of /h/-onset syllables. The results of this dissertation find large degrees of physiological nasalization during and after /h/; this suggests that a misperception-based model is likely insufficient to completely explain impressions of nasality during vowels after /h/. Rather, it is possible that misperception is a contributing factor that induces greater VPO after /h/ compared to other syllables such as those beginning with /ʔ/. Further perception study, such as a paradigm that asks listeners to rank nasality during Thai syllables (that vary based on onset consonant) would be required to assess the role that breathiness may play in the percept of nasalization.

Chapter 6

CONCLUSION

Spontaneous nasalization has been an historically difficult topic of study because its analysis requires a combination of articulatory data that measures both nasalization and phonation quality. This dissertation integrated measures from state-of-the-art ultra-fast MRI, aerodynamics, and electroglottography to approach the problem of spontaneous nasalization. This dissertation presents a detailed study of spontaneous nasalization production in Thai. The speech of ten speakers of Thai was assessed using various tools to measure speech physiology. Four speakers were recorded recorded using ultra-fast MRI. This method allowed us to measure velopharyngeal opening at 25 frames per second. Six different speakers were recorded using a nasal and oral airflow mask. Simultaneous electroglottography and acoustic data were recorded from these six speakers separately. These methods allowed us to measure nasal airflow and open quotient in order to approximate degree of nasalization and phonation quality. Speakers produced syllables that varied by onset consonant and vowel height within the same carrier phrase. Syllables with an onset glottal consonant /h, ʔ/ were varied at four different vowel heights; other syllables with onset consonants /n, t^h, d/ were included to present contrasts of nasal (/n/), aspirated/breathy non-nasal (/t^h/), and non-aspirated/breathy non-nasal (/d/) consonant environments.

Spontaneous nasalization is characterized by nasalization that develops in contexts lacking an etymological nasal. Triggers of spontaneous nasalization include glottal consonants (rhinoglottophilia), low vowels (rhinochthamalophilia), and consonants with voiceless turbulence (rhinosyrigmatophilia). In Thai, low and mid-low vowels were reported to nasalize after glottal consonants /h/ and /ʔ/ (Cooke, 1989; Matisoff, 1975). Explanations for spontaneous nasalization in these contexts involve underspecification for VPO during glottal consonants, modulation in VPO based on tongue height, and acoustic similarities between nasalization and breathiness Blevins and Garrett (1992); Matisoff (1975); Ohala (1974).

The results of this dissertation deepen our understanding of the production of spontaneous nasalization and contexts that facilitate its realization in Thai. The major finding of this dissertation is that /h/-onset syllables are produced with greater physiological nasalization than /ʔ/-onset syllables. This finding suggests that, counter to previous claims that glottal consonants are generally underspecified for VPO in Thai, /ʔ/ is in fact specified, at least in a prosodically prominent context. Furthermore, we observed increased breathiness during vowels just after /h/, but not after /ʔ/, suggesting a potential relationship between nasalization and breathiness in Thai /h/-onset syllables. The voiceless turbulence of /h/ induces coarticulatory breathiness during the following vowel; this breathiness may induce the perception of nasalization during the following vowel. This may explain why vowels after /h/ are reported to sound more nasal than vowels after /ʔ/ (Cooke, 1989; Matisoff, 1975). Further perceptual testing is needed to assess this possibility.

Previous accounts of spontaneous nasalization in Central Thai report that only low and mid-low vowels are nasalized after glottal consonants (Cooke, 1989; Matisoff, 1975; Noss, 1964). However, we observed inconsistent and often minimal variation in nasal airflow during vowels of varying height in both nasalized non-nasal context. While did not observe large VPO during the mid-high vowel of /heː˧/ ‘to swiftly flock’ in four speakers, we did observe increased nasal airflow during this vowel in six speakers. This suggests that for some Central Thai speakers spontaneous vowel nasalization may now be similar to Northeastern Thai, a dialect where all vowels were reported to nasalize after glottal consonants, not just low and mid-low vowels (Matisoff, 1975).

Our results demonstrate that spontaneous nasalization in Thai primarily results from rhinoglottophilia: CV syllables beginning with a glottal fricative /h/ are produced with increased physiological nasalization during the onset consonant and vowel. This nasalization is likely primarily attributed to underspecification of the onset consonant /h/. The presence of coarticulatory breathiness after /h/ may potentially play a role in spontaneous nasalization in Thai. Nasal coupling and breathiness may be integrated into a single acoustic object that the listener perceives as nasal. This possibility presents an intriguing example of the many-to-one problem that requires further perceptual testing (Maeda, 1990).

Appendix A

APPENDIX

A.1 MRI: number of tokens per speaker

Figures A.1 and A.2 show the number of repetitions produced for each syllable for each speaker.

Word	TF1	TM1	TM2	TF2
/haː˧/	48	100	102	80
/hɛː˧/	45	88	96	78
/heː˧/	48	98	104	80
/ʔaː˧/	50	102	94	86
/ʔɛː˧/	51	108	98	96
/ʔeː˧/	50	90	92	86
/naː˧/	51	110	96	96
/tʰaː˧/	46	98	102	78
/daː˧/	45	94	96	82
/diː˧/	46	96	94	86

Table A.1: Number of repetitions per speaker – Velopharyngeal scans

A.2 MRI: VPO based on sample length

In Figures A.2 and A.3 we observe that token duration is similar across all syllables, for both the MRI and acoustic only data. In the Figure A.1 we plot mean VPO across all time samples by token length. First, the VPO of all time samples of all tokens was averaged. Then the samples were binned by the number of samples of each token. Bins were distributed at 100, 500, 1000, and 3000 samples, indicated by color. We observe that mean VPO does not

Word	TF1	TM1	TM2	TF1
/haː˧/	68	60	60	68
/hɛː˧/	68	59	66	62
/heː˧/	68	59	57	65
/ʔaː˧/	68	61	60	72
/ʔɛː˧/	41	61	59	76
/ʔeː˧/	68	62	62	72
/naː˧/	72	60	59	75
/t^haː˧/	70	59	62	65
/daː˧/	64	59	64	61

Table A.2: Number of repetitions per speaker – Audio recordings

appear to be affected by token duration (operationalized by number of samples). Figures A.2 and A.3 show vowel duration by syllables type. For the MRI tokens, vowel duration is similar across all syllables. For the acoustic only data, we notice that the vowel of /daː/ is longer than the vowel of /t^haː/ and that the vowel /ɛː/ tends to be slightly longer than other vowels.

A.3 Oral airflow data

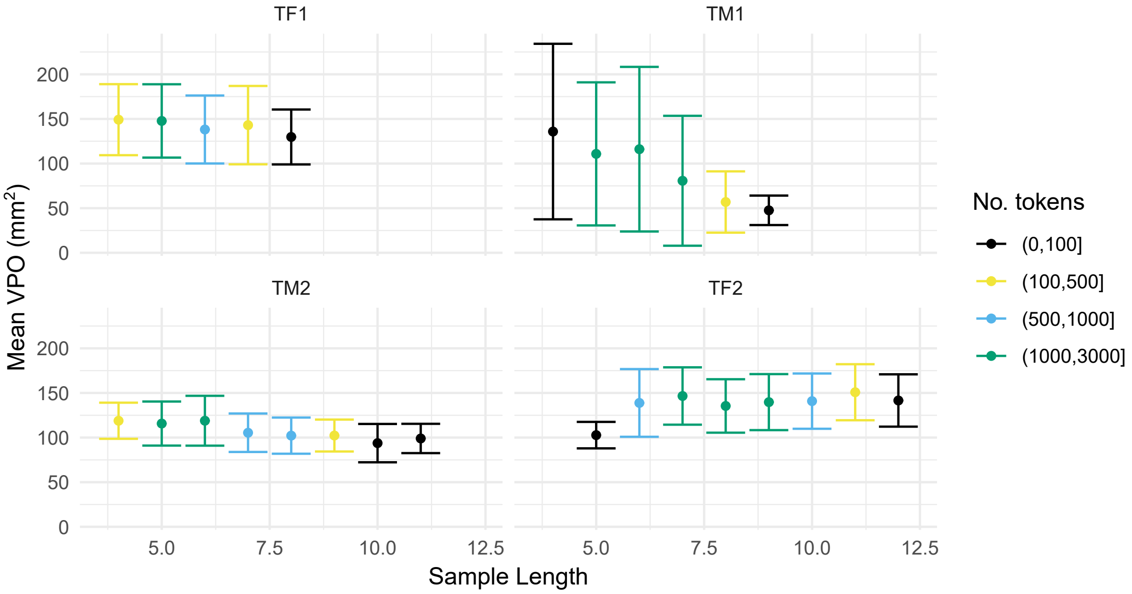


Figure A.1: Plot of mean VPO based on sample length. Samples were binned into distributions of 100, 500, 1000, and 3000. Mean VPO does not appear to be affected by token duration (sample length).

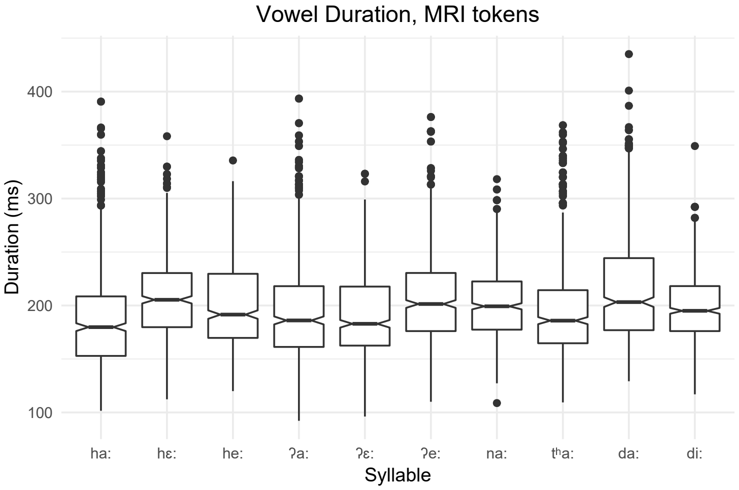


Figure A.2: Plot of vowel duration by syllable of speech data collected in MRI session.

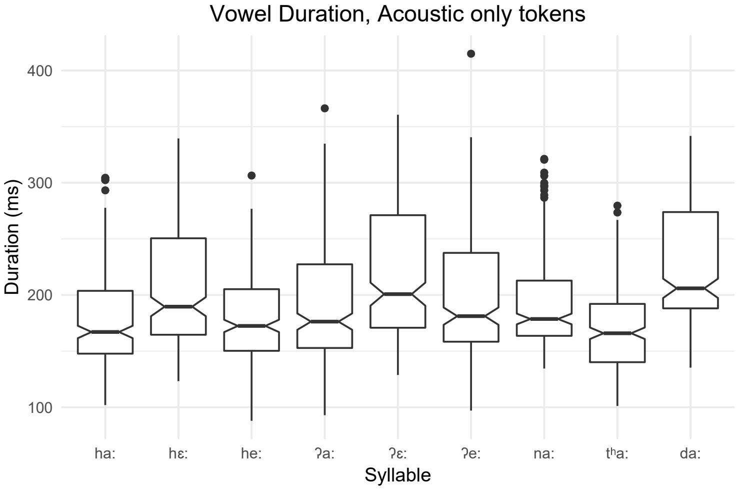


Figure A.3: Plot of vowel duration by syllable of speech data collected in acoustic only recording session.

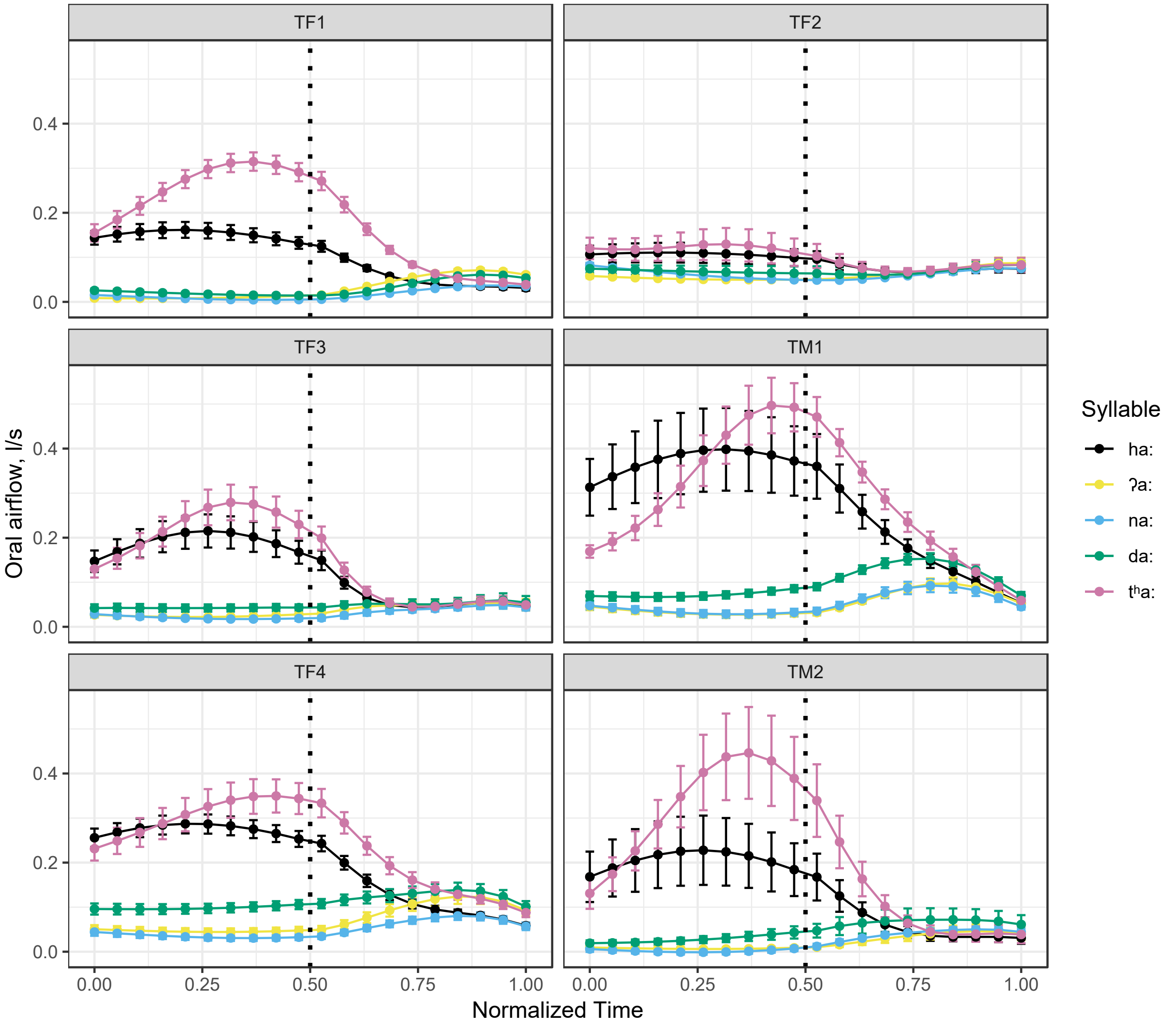


Figure A.4: Mean oral airflow of all syllables with the vowel /a/, liters/second. Confidence interval bars are based on the standard error.

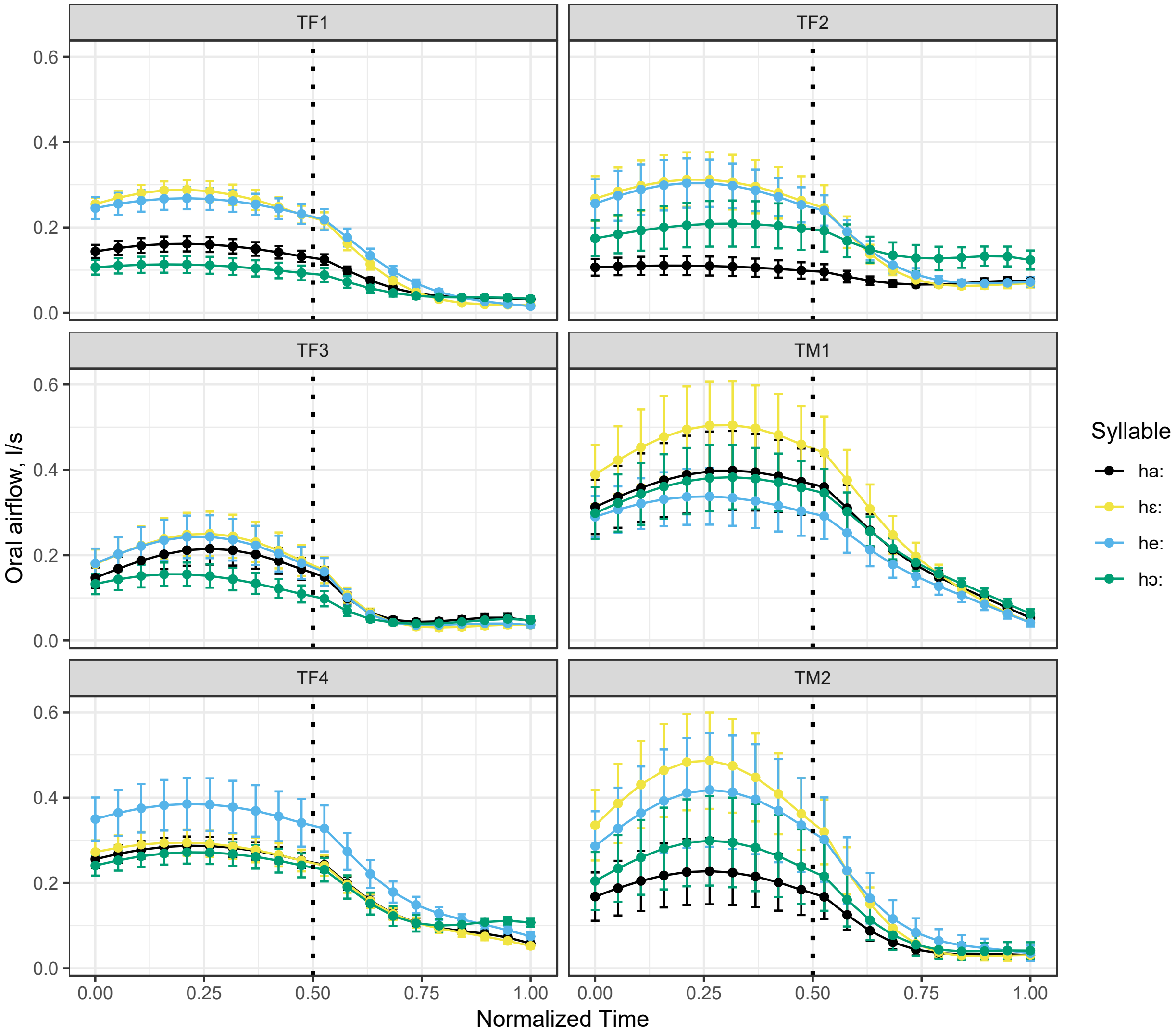


Figure A.5: Mean oral airflow of all syllables with the onset consonant /h/, liters/second. Confidence interval bars are based on the standard error.

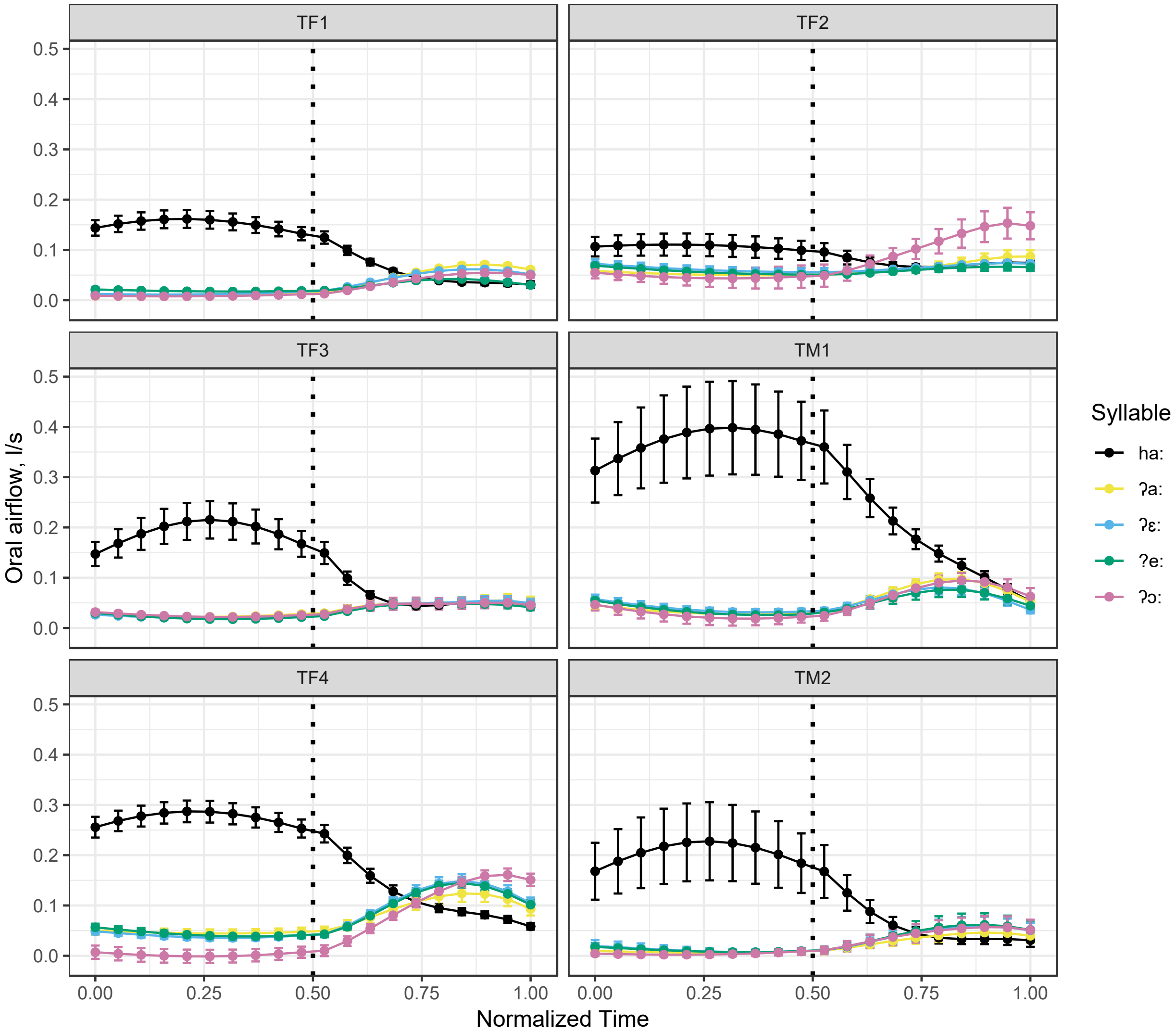


Figure A.6: Mean oral airflow of all syllables with the onset consonant /ʔ/, liters/second. Confidence interval bars are based on the standard error.

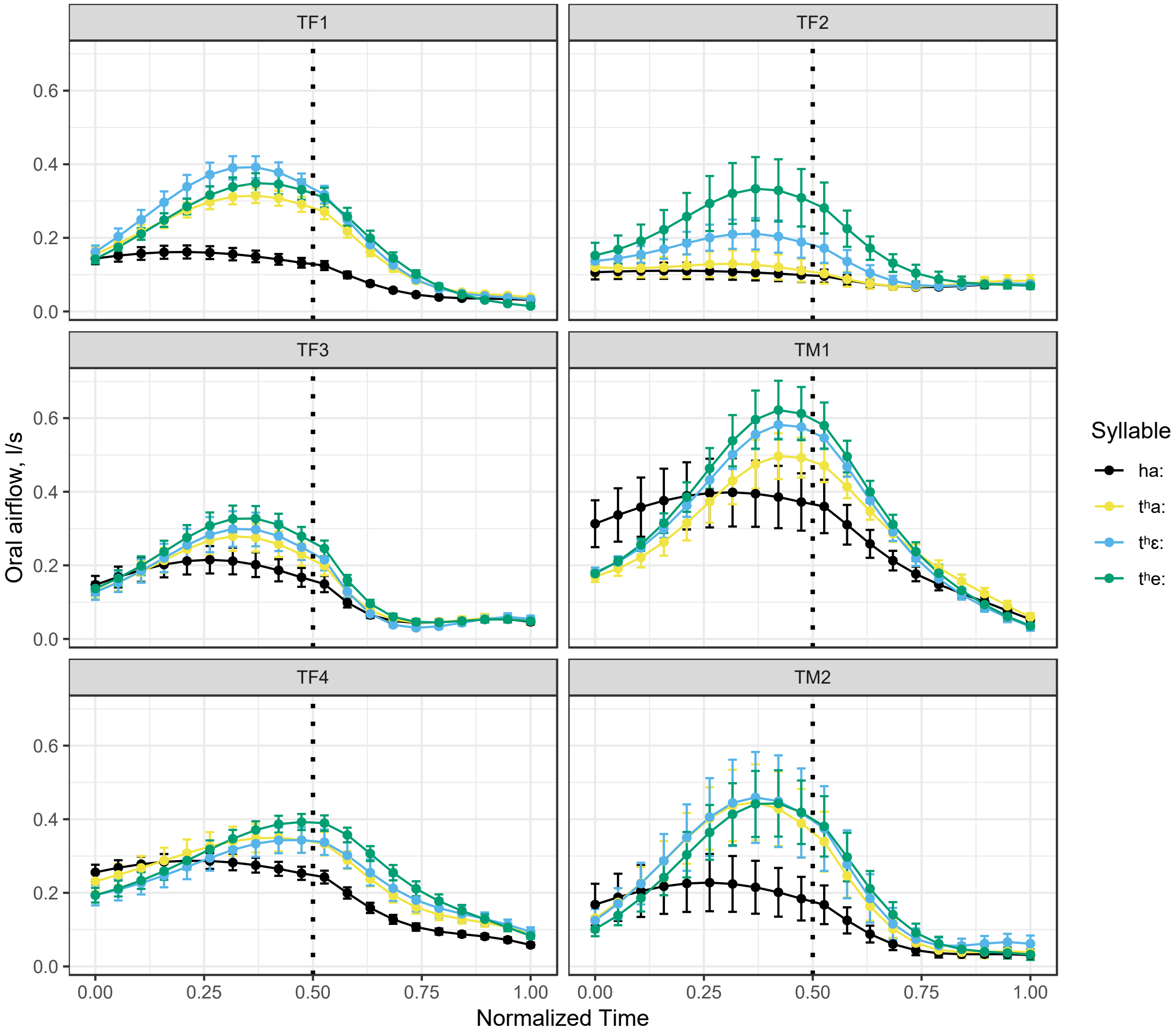


Figure A.7: Mean oral airflow of all syllables with the onset consonant /t^{h}/, liters/second. Confidence interval bars are based on the standard error.

BIBLIOGRAPHY

Abramson, A. (1962). The vowels and tones of standard thai: Acoustical measurements and experiments. *International Journal of American Linguistics*, **28**(2).

Abramson, A. S. and Whalen, D. H. (2017). Voice onset time (vot) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of phonetics*, **63**, 75–86.

Abramson, A. S., Tiede, M. K., and Luangthongkum, T. (2015). Voice register in mon: Acoustics and electroglottography. *Phonetica*, **72**(4), 237–256.

Al-Bamerni, A. (1983). *Oral, velic, and laryngeal coarticulation across languages*. Ph.D. thesis, University of Oxford.

Anawin, L. (1998). Intonation in thai. In D. Hirst and A. Cristo, editors, *Intonation Systems: A Survey of Twenty Language*, pages 376–394. Cambridge University Press.

Baayen, R. H., van Rij, J., de Cat, C., and Wood, S. (2018). Autocorrelated errors in experimental data in the language sciences: Some solutions offered by generalized additive mixed models. In *Mixed-Effects Regression Models in Linguistics*, pages 49–69. Springer.

Bae, Y., Perry, J. L., and Kuehn, D. P. (2014). Videofluoroscopic investigation of body position on articulatory positioning. *Journal of Speech, Language, and Hearing Research*, **57**(4), 1135–1147.

Baken, R. J. and Orlikoff, R. F. (2000). *Clinical Measurement of Speech and Voice*. Singular Publishing Group, San Diego, CA, 2nd edition.

Barlaz, M., Fu, M., Dubin, J., Shosted, R., Liang, Z. P., and Sutton, B. (2015). Lingual differences in brazilian portuguese oral and nasal vowels: An mri study. In *the 18th Meeting of the International Congress of the Phonetic Sciences*.

Barlaz, M., Liang, Z. P., and Sutton, B. (2017). Articulatory correlates of phonemic and coarticulatory nasalization. In *the 174th Meeting of the Acoustical Society of America*.

Barlaz, M., Shosted, R., Fu, M., and Sutton, B. (2018). Oropharygneal articulation of phonemic and phonetic nasalization in Brazilian Portuguese. *Journal of Phonetics*, **71**, 81–97.

Beddor, P. S. (1993). The perception of nasal vowels. In M. K. Huffman and R. A. Krakow, editors, *Nasals, Nasalization, and the Velum*, volume 5 of *Phonetics and Phonology*, pages 171–196. Academic Press, San Diego, CA.

Beddor, P. S. and Krakow, R. A. (1999). Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *The Journal of the Acoustical Society of America*, **106**(5), 2868–2887.

Bell-Berti, F. (1993). Understanding velic motor control: Studies of segmental context. In M. K. Huffman and R. A. Krakow, editors, *Nasals, Nasalization, and the Velum*, volume 5 of *Phonetics and Phonology*, pages 63–86. Academic Press, San Diego, CA.

Bell-Berti, F., Baer, T., Harris, K. S., and Niimi, S. (1979). Coarticulatory effects of vowel quality on velar function. *Phonetica*, **36**, 187–193.

Bickley, C. (1982a). Acoustic analysis and perception of breathy vowels. *MIT speech communication working papers*, **1**, 71–81.

Bickley, C. (1982b). Acoustic analysis and perception of breathy vowels. *Speech communication group working papers*, **1**, 71–81.

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, **30**(2), 163–191.

Blevins, J. (2004). *Evolutionary Phonology: The emergence of sound patterns*, chapter Section 6.1.1. Cambridge University Press, Cambridge.

Blevins, J. and Garrett, A. (1992). Ponapeic nasal substitution: New evidence for rhinoglottophilia. *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Place of Morphology in a Grammar*, pages 2–21.

Blevins, J. and Garrett, A. (1993). The evolution of Ponapeic nasal substitution. *Oceanic Linguistics*, pages 199–236.

Boersma, P. and Weenink, D. (2001–2018). Praat: Doing phonetics by computer. `http://www.praat.org/`.

Bothorel, A., Simon, P., Wioland, F., and Zerling, J.-P. (1986). Cinéradiographie des voyelles et consonnes du français [Cinéradiography of French vowels and consonants]. *Travaux de l'Institut de Phonétique de Strasbourg*, **18**, 1–296.

Carignan, C. (2017). Covariation of nasalization, tongue height, and breathiness in the realization of F1 of Southern French nasal vowels. *Journal of Phonetics*, **63**, 87–105.

Carignan, C., Shosted, R. K., Shih, C., and Rong, P. (2011). Articulatory compensation for nasality: An EMA study of lingual positin during nasalized vowels. *Journal of Phonetics*, **39**, 668–682.

Carignan, C., Shosted, R. K., Fu, M., Liang, Z. P., and Sutton, B. P. (2015). A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French. *Journal of Phonetics*, **50**, 34–51.

Chasaide, A. N. and Gobl, C. (1997). Voice source variation. In W. J. Hardcastle and J. Laver, editors, *The Handbook of Phonetic Sciences*, pages 427–461. Blackwell, Oxford.

Chen, M. Y. (1995). Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *Journal of the Acoustical Society of America*, **98**, 2443–2453.

Chi, X. and Sonderegger, M. (2007). Subglottal coupling and its influence on vowel formants. *Journal of the Acoustical Society of America*, **122**(3), 1735–1745.

Childers, D., Moore, G., Naik, J., Larar, J., and Krishnamurthy, A. (1983). Assessment of laryngeal function by simultaneous, synchronized measurement of speech, electroglottography and ultra-high speed film. In *Transcripts of the eleventh symposium: Care of the professional voice*, pages 234–44.

Childers, D., Hicks, D., Moore, G., and Alsaka, Y. (1986). A model for vocal fold vibratory motion, contact area, and the electroglottogram. *The Journal of the Acoustical Society of America*, **80**(5), 1309–1320.

Childers, D., Hicks, D., Moore, G., Eskenazi, L., and Lalwani, A. (1990). Electroglottography and vocal fold physiology. *Journal of Speech, Language, and Hearing Research*, **33**(2), 245–254.

Childers, D. G. (1983). Electroglottography, speech, and ultra-high speed cinematography. *Vocal fold physiology: Biomechanics, acoustics and phonatory control*, pages 202–220.

Cho, T., Kim, D., and Kim, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in english. *Journal of Phonetics*, **64**, 71–89.

Clumeck, H. (1973). *An experiment using nasograph*. Ph.D. thesis.

Clumeck, H. (1975). A cross-linguistic investigation of vowel nasalization: An instrumental study. In C. Ferguson, L. Hyman, and J. J. Ohala, editors, *Nasálfest: Papers from a symposium on nasals and nasalization*, pages 133–154. Stanford University Language Universals Project, Palo Alto, CA.

Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, **4**(4), 337–351.

Colarusso, J. (1988). *The Northwestern Caucasian languages: A phonological survey*. Routledge, New York.

Cooke, J. R. (1989). Thai nasalized vowels. *Papers in South–East Asian linguistics*, **80**, 169–174.

Cranen, B. and Boves, L. (1987). On subglottal formant analysis. *Journal of the Acoustical Society of America*, **81**, 734–746.

Desmeules-Trudel, F. and Brunelle, M. (2018). Phonotactic restrictions condition the realization of vowel nasality and nasal coarticulation: Duration and airflow measurements in Québécois French and Brazilian Portuguese. *Journal of Phonetics*, **69**, 43–61.

Diehl, R. L., Kluender, K. R., and Walsh, M. A. (1990). Some auditory bases of speech perception and production. In W. A. Ainsworth, editor, *Advances in speech, hearing and language processing*, pages 243–268. JAI Press, London.

Dietsch, A. M., Cirstea, C. M., and Searl, J. P. (2013). Effects of body position and sex group on tongue pressure generation. *International Journal of Orofacial Myology*, **39**, 12–22.

Diller, A. (2004). Resources for thai language research. In *The Tai-Kadai Languages*, pages 47–98. Routledge.

Dingemanse, M., Torreira, F., and Enfield, N. J. (2013). Is "huh?" a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *PloS one*, **8**(11), e78273.

Engwall, O., Delvaux, V., and Metens, T. (2006). Interspeaker variation in the articulation of nasal vowels. In H. Yehia, D. Demolin, and R. Laboissière, editors, *Proceedings of the 7th International Seminar on Speech Production*, pages 3–10, Ubatuba, São Paulo. Centro de Estudos da Fala, Acústica, Linguagem e Música.

Esling, J. H. and Harris, J. G. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In W. J. Hardcastle and J. M. Beck, editors, *A Figure of Speech: A Festschrift for John Laver*, pages 347–384. Lawrence Erlbaum Associates, Inc, New York.

Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton, The Hague.

Fant, G., Ishizaka, K., Lindqvist, J., and Sundberg, J. (1972). Subglottal formants. *Speech Transmission Laboratory Quarterly Progress and Status Report*, **1**, 79–84.

Fu, M., Barlaz, M., Shosted, R., Liang, Z. P., and Sutton, B. P. (2015a). High–resolution dynamic speech imaging with deformation estimation. In *Proceedings of the 37th International Conference of the IEEE Engineering in Medicine and Biology Society*.

Fu, M., Zhao, B., Carignan, C., Shosted, R. K., Perry, J. L., Kuehn, D. P., Liang, Z.-P., and Sutton, B. P. (2015b). High-resolution dynamic speech imaging with joint low-rank and sparsity constraints. *Magnetic resonance in medicine*, **73**(5), 1820–1832.

Fu, M., Barlaz, M. S., Holtrop, J. L., Perry, J. L., Kuehn, D. P., Shosted, R. K., Liang, Z.-P., and Sutton, B. P. (2017). High-frame-rate full-vocal-tract 3d dynamic speech imaging. *Magnetic resonance in medicine*, **77**(4), 1619–1629.

Fujimura, O. and Lindqvist, J. (1971). Sweep-tone measurements of vocal-tract characteristics. *Journal of the Acoustical Society of America*, **49**, 541–558.

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, **45**, 106–113.

Garellek, M. (2019). The phonetics of voice. In W. F. Katz and P. F. Assmann, editors, *The Routledge Handbook of Phonetics*, chapter 4. Routlege, New York.

Garellek, M. and Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association*, **41**, 185–205.

Garellek, M., Ritchart, A., and Kuang, J. (2016). Breathy voice during nasality: A cross-linguistic study. *Journal of Phonetics*, **59**, 110–121.

Garrett, A. and Johnson, K. (2013). Phonetic bias in sound change. In C. Alan, editor, *Origins of sound change: Approaches to phonologization*, pages 51–97. OUP Oxford.

Gordon, M. and Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, **29**, 383–406.

Gordon, R. G. (2005). *Ethnologue: Languages of the World*. SIL International, Dallas, TX, 15th edition. Online version: http://www.ethnologue.com/.

Grierson, G. A. (1922). Spontaneous nasalization in the indo-aryan languages. *Journal of the Royal Asiatic Society*, **54**(3), 381–388.

Hajek, J. and Maeda, S. (2000). Investigating universals of sound change: the effect of vowel height and duration on the development of distinctive nasalization. In M. Broe and J. Pierrehumbert, editors, *Papers in Laboratory Phonology V*. Cambridge University Press, Cambridge.

Hanson, H. M. (1995). *Glottal characteristics of female speakers*. Ph.d. dissertation, Harvard University, Cambridge MA.

Harris, J. G. (2001). States of the glottis of thai voiceless stops and affricates. In M. Kalaya and A. S. Abramson, editors, *Essays in Tai Linguistics*, pages 3–12. Chulalongkorn University Press Bangkok.

Hawks, J. W. and Miller, J. D. (1995). A formant bandwidth estimation procedure for vowel synthesis [43.72. ja]. *The Journal of the Acoustical Society of America*, **97**(2), 1343–1344.

Henrich, N., d'Alessandro, C., Doval, B., and Castellengo, M. (2004). On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *The Journal of the Acoustical Society of America*, **115**(3), 1321–1332.

Hetzron, R. (1969). Two notes on semitic laryngeals in east gurage. *Phonetica*, **19**(2), 69–81.

Hillenbrand, J. and Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearning Resarch*, **39**, 311–321.

Hiroto, I., Hirano, M., and Umeno, M. (1964). A cineradiographic study on the movement of the soft palate during phonation of speech sound. *Studia Phonologica*, **3**, 35–46.

Hixon, T. and Hoit, J. (2005). *Evaluation and management of speech breathing disorders: Principles and methods*. Redington Brown, Tuscon, AZ.

Hixon, T., Weismer, G., and Hoit, J. (2014). *Preclinical speech science: Anatomy, physiology, acoustics, perception*. Plural, San Diego, CA, 2 edition.

Hoit, J. and Weismer, G. (2018). *Foundations of Speech and Hearing: Anatomy and Physiology*. Plural Publishing, Inc, San Diego, CA.

Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (1988). Glottal airflow and pressure measurements for soft, normal and loud voice by male and female speakers. *Journal of the Acoustical Society of America*, **84**, 511–529.

Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., and Goldman, S. L. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research*, **38**(6), 1212–1223.

House, A. S. and Stevens, K. N. (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders*, **21**, 218–232.

Hsieh, F. Y., Goldstein, L., Byrd, D., and Narayanan, S. (2013). Truncation of pharyngeal gesture in english diphthong [aI]. In *Proceedings of INTERSPEECH*, pages 968–972.

Iseli, M. and Alwan, A. (2004). An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. In *2004 IEEE international conference on acoustics, speech, and signal processing*, volume 1, pages I–669. IEEE.

Ishizaka, K. and Matsudaira, M. (1972). *Fluid mechanical considerations of vocal cord vibration*. Speech Communication Research Laboratory, Santa Barbara, CA. Monograph No. 8.

Johnson, K. (1997). *Acoustic and Auditory Phonetics*. Blackwell, Oxford.

Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, **5**(2), 275–292.

Kingston, J. (1991). Integrating articulations in the perception of vowel height. *Phonetica*, **48**, 149–179.

Klatt, D. H. (1973). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, **18**(4), 686–706.

Klatt, D. H. and Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, **87**, 820–857.

Ko, V., Wieling, M., Wit, E., Nerbonne, J., and Krijnen, W. (2014). Social, geographical, and lexical influences on dutch dialect pronunciations. *Computational Linguistics in the Netherlands Journal*, **4**, 29–38.

Kollara, L. and Perry, J. L. (2014). Effects of gravity on the velopharyngeal structures in children using upright magnetic resonance imaging. *The Cleft Palate–Craniofacial Journal*, **51**(6), 669–676.

Kreiman, J., Gerratt, B. R., and Antoñanzas-Barroso, N. (2007). Measures of the glottal source spectrum. *Journal of speech, language, and hearing research*, **50**(3), 595–610.

Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Blackwell, Oxford.

Lenth, R. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R Package Version 1.3.2*.

Liang, Z. P. (2007). Spatiotemporal imaging with partially separable functions. In *Proceedings from 4th IEEE International Symposium on Biomedical Imaging: from nano to macro, (ISBI 2007)*, pages 988–991.

Lintz, L. B. and Sherman, D. (1961). Phonetic elements and perception of nasality. *Journal of Speech and Hearing Research*, **4**, 381–396.

Lulich, S. M., Alwan, A., Arsikere, H., Morton, J. R., and Sommers, M. S. (2011). Resonances and wave propagation velocity in the subglottal airways. *Journal of the Acoustical Soceity of America*, **130**(4), 2108–2115.

Maeda, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In *Speech production and speech modelling*, pages 131–149. Springer, Netherlands.

Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In M. K. Huffman and R. A. Krakow, editors, *Nasals, Nasalization, and the Velum*, volume 5 of *Phonetics and Phonology*, pages 147–170. Academic Press, San Diego, CA.

Matisoff, J. A. (1970). Glottal dissimilation and the lahu high–rising tone: a tonogenetic case–study. *Journal of the American Oriental Society*, **90**(1), 13–44.

Matisoff, J. A. (1975). Rhinoglottophilia: the mysterious connection between nasality and glottality. Nasálfest: Papers from a symposium on nasals and nasalization.

Michaud, A., Hendrich, N., and Gendrot, C. (2017). Tools for electroglottographic analysis: Software, documentation and databases.

Miller, A. L. (2007). Guttural vowels and guttural co-articulation in Ju|'hoansi. *Journal of Phonetics*, **35**(1), 56–84.

Moon, J. B. and Canady, J. (1995). effects of gravity on velopharyngeal muscle activity during speech. *Cleft Palate–Craniofacial Journal*, **32**, 371–375.

Nixon, J. S. and Best, C. T. (2018). Acoustic cue variability affects eye movement behaviour during non-native speech perception: a gamm model. In *Proceedings of Speech Prosody*, volume 9.

Noss, R. B. (1964). *Thai reference grammar*. Foreign Service Institute, Washington, D.C.

Ohala, J. J. (1971). Monitoring soft-palate movements in speech. *Journal of the Acoustical Society of America*, **50**, 140.

Ohala, J. J. (1974). Experimental historical phonology. In J. M. Anderson and C. Johns, editors, *Historical linguistics II. Theory and description in phonology*, pages 353–389. Proceedings of the 1st International Conference on Historical Linguistics, Amsterdam: North Holland.

Ohala, J. J. (1975). Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, and J. J. Ohala, editors, *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, pages 289–316. Stanford University Language Universals Project, Palo Alto, CA.

Ohala, J. J. (1980). The application of phonological universals in speech pathology. *Speech and Language: Advances in basic research and practice*, **3**, 75–97.

Ohala, J. J. (1993a). The phonetics of sound change. In C. Jones, editor, *Historical linguistics: Problems and perspectives*, pages 237–278. Routledge.

Ohala, J. J. (1993b). Sound change as nature's speech perception experiment. *Speech Communication*, **13**, 155–161.

Ohala, J. J. (1996). Speech perception is hearing sounds, not tongues. *Journal of the Acoustical Society of America*, **99**, 1718–1725.

Ohala, J. J. and Amador, M. (1981). Spontaneous nasalization. *The Journal of the Acoustical Society of America*, **69**, S54.

Ohala, J. J. and Ohala, M. (1992). Nasal and nasalization in Hindi. Paper presented at 3rd International Symposium on Language and Linguistics: Pan-Asiatic Linguistics. Chulalongkorn Univ., Bangkok, Thailand. 8-10 Jan 92.

Ohala, J. J., Solé, M.-J., and Ying, G. (1998). Do nasalized fricatives exist? *Journal of the Acoustical Society of America*, **103**(5), 3085.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, **9**(1), 62–66.

Pruthi, T. and Espy-Wilson, C. Y. (2005). Acoustic parameters for automatic detection of nasal manner. *Speech Communication*, **43**(3), 225–239.

Pruthi, T., Espy-Wilson, C. Y., and Story, B. H. (2007). Simulation and analysis of nasalized vowels based on magnetic resonance imaging data. *Journal of the Acoustical Society of America*, **121**, 3858–3873.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ruhlen, M. (1973). Nasal vowels. *Working Papers on Language Universals*, **12**, 1–36.

Schadeberg, T. C. (1989). The velar nasal of Nyole. *Annales Aequatoria*, **10**, 169–179.

Schwyzer, E. (1934). Dissimilatorische geminatenauflösung als folge yon übersteigerung, zunächst im neugriechischen und im spätaltgriechischen. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen*, **61**(3./4. H), 222–252.

Shosted, R. K. (2006). *The aeroacoustics of nasalized fricatives*. Ph.D. thesis, University of California Berkeley.

Shosted, R. K. (2015). Nasal vowels are not [+ nasal] oral vowels. In *Romance Linguistics 2012: Selected papers from the 42nd Linguistic Symposium on Romance Languages (LSRL)*, Utah. John Benjamins Publishing Company.

Shosted, R. K., Carignan, C., and Rong, P. (2012a). Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *Journal of the Acoustical Soceity of America*, **1**, 455.

Shosted, R. K., Carignan, C., and Rong, P. (2012b). Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *Journal of the Acoustical Society of America*, **131**(1), 455–465.

Shue, Y.-L., Keating, P., Vicenik, C., and Yu, K. (2011). Voicesauce: A program for voice analysis. *Proceedings of the ICPhS XVII*, pages 1846–1849.

Simpson, A. P. (2012). The first and second harmonics should not be used to measure breathiness in male and female voices. *Journal of Phonetics*, **40**, 477–490.

Smith, C., Proctor, M. I., Iskarous, K., Goldstein, L., and Narayanan, S. (2013). Stable articulatory tasks and their variable formation: Tamil retroflex consonants. In *Proceedings of INTERSPEECH*, pages 2006–2009, Lyon, France.

Sóskuthy, M. (2017). Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. *arXiv preprint arXiv:1703.05339*.

Stevens, K. and Hanson, H. (1994). Classification of glottal vibration from acoustic measurements. *Paper presented at the 8th vocal fold physiology conference, Kurume, Japan.*

Stevens, K. N. (2000). *Acoustic Phonetics*. MIT Press, Cambridge.

Stevens, K. N. and Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, **38**(1), 10–19.

Stone, M., Stock, G., Bunin, K., Kumar, K., and Epstein, M. (2007). Comparison of speech production in upright and supine position. *Journal of the Acoustical Society of America*, **122**, 532.

Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *Journal of the Acoustical Society of America*, **142**(4), 2469–2482.

Turner, R. L. (1921). Gujarati phonology. *Journal of the Royal Asiatic Society*, **53**(4), 505–544.

van Rij, J., Wieling, M., Baayen, R. H., and van Rijn, H. (2016). itsadug: Interpreting time series and autocorrelated data using gamms. *R package version 2.2*.

Vorperian, H. K., Kurtzweil, S. L., Fourakis, M., Kent, R. D., Tillman, K. K., and Austin, D. (2014). Effect of body position on vocal tract acoustics: Acoustic pharyngometry and vowel formants. *The Journal of the Acoustical Society of America*, **138**(2), 833–845.

Walker, R. (2000). *Nasalization, Neutral Segments, and Opacity Effects*. Garland, New York.

Wayland, R. and Jongman, A. (2003). Acoustic correlates of breathy and clear vowels: the case of khmer. *Journal of Phonetics*, **31**(2), 181–201.

Whalen, D. H. (1990). Intrinsic velar height in supine vowels. *Journal of the Acoustical Soceity of America*, **88**, S54.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between l1 and l2 speakers of English. *Journal of Phonetics*, **70**, 86–116.

Wieling, M., Tomaschek, F., Arnold, D., Tiede, M., Broker, F., Thiele, S., Wood, S. N., and Baayen, R. H. (2016). Investigating dialectal differences using articulography. *Journal of Phonetics*, **59**, 122–143.

Wood, S. and Wood, M. S. (2009). mgcv (R package). `https://cran.r-project.org/web/packages/mgcv/mgcv.pdf`.

Wood, S. N. (2006). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Zemlin, W. R. (1997). *Speech and Hearing Science: Anatomy and Physiology*. Allyn & Bacon, Boston, MA, 4th edition.