Phonetic Accommodation to Hypernasal Speech

by

Monique Claire Tardif

A thesis submitted in conformity with the requirements

for the degree of Master of Science

Graduate Department of Rehabilitation Science

University of Toronto

Phonetic Accommodation to Hypernasal Speech

Monique Claire Tardif

Master of Science

Rehabilitation Science Institute

University of Toronto

2019

Abstract

Speech accommodation is a process in which an individual's speech becomes more (convergence) or less (divergence) similar to their interlocutor's speech. Individuals diverge to increase distance and reduce commonality between conversation partners (Shepard, Giles, Le Poire, 2001) based on social biases (Babel, 2010). Hypernasality is a negatively perceived (Blood & Hyman, 1997; Watterson, Mancini, Brancamp, & Lewis, 2013) speech disorder resulting from an excessive amount of acoustic energy emanating from the nasal cavity (Zajac & Vallino, 2017). Since it is theorized that individuals diverge in negative social contexts (Shepard et al., 2001) and that hypernasality is negatively perceived (Blood & Hyman, 1997; Watterson et al., 2013), we hypothesized that speakers would diverge from hypernasal speech. Speakers read sentences in response to hearing pre-recorded sentences with modelled hypernasal and typical speech. Results indicated that speakers inconsistently converged to typical levels of nasality and consistently diverged from hypernasal speech.

Keywords: Phonetic Accommodation, Communication, Nasality, Nasalance, Hypernasality, Speech Disorders

Acknowledgments

I wish to thank my primary supervisor Dr. Tim Bressmann for giving me the opportunity to work on this project, for his patience, and for the many edits. I wish to thank my co-supervisor Dr. Jennifer Pardo for her kind words and sharing her impressive knowledge of Rstudio and statistics.

Thank you to my committee members Dr. Pascal van Lieshout and Dr. Frank Russo for their long-lasting support throughout many meetings. Thank you to Dr. Yana Yunusova for her unwavering faith in me. And thank you to Loida, Antoinetta, and all the staff in the RSI administrative office who tirelessly work to help all of the students at RSI.

I would like to thank my coworkers, colleagues, and friends from the Voice and Resonance lab and UNESP: Dr. Gillian deBoer, Dr. Viviane Martino, Dr. Larissa Berti, Dr. Eliana Fabron, Dr. Gabriela Zuin, Charlene, Meghan, Mia, and all the other students, workers, and volunteers in the labs.

Thank you to my friends, volleyball partners, and life partner for offering love, kindness, and support. Last, but not least, I wish to thank my family for their unwavering support throughout this trying process. Thank you for the rock n' roll and roadtrips; they've made all the difference.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations, Symbols, and Nomenclature

ANOVA	: Analysis of Variance
CAT	: Communication Accommodation Theory
F	: Female
FLSD	: Fisher's Least Significant Difference
H (1-3)	: Hypothesis (1-3)
HONC	: Horii Oral-Nasal Coupling
Hyp	: Hypernasal
M	: Male
M	: Mean
N	: Number of Individuals
η_g^2	: Generalized Eta Squared
SD	: Standard Deviation
SLP	: Speech-Language Pathology / Pathologist
VC	: Vowel-Consonant
VN	: Vowel-Nasal

1. Introduction

Typical speech requires the propagation of air past the vocal folds and through the vocal tract (Hixon, Weismer, & Hoit, 2014). The air pressure from the lungs causes the adducted vocal folds to vibrate. The vocal tract and articulators then shape the air into speech sounds (Hixon et al., 2014). The vocal tract contains two resonant cavities: the mouth, which shapes the articulation of vowels and consonant sounds, and the nose, which adds resonance as a side channel. Typical non-disordered speech makes use of both the oral and nasal cavities (Kummer, 2008). The velopharyngeal port is the gate that separates the oral and nasal cavity. During oral sounds, the velopharyngeal port is closed (or mostly closed) to restrict airflow to the nasal cavity (e.g. during the vowel /a/). During nasal consonants (/m/, /n/, and /ŋ/), the velopharyngeal port is open to allow airflow to the nasal cavity.

In English, there are no phonologically nasalized vowels. However, during coarticulation of Vowel-Consonant (VC) syllables, English vowels adjacent to nasal sounds are usually nasalized (e.g. "on", "and", "lemon") (Chen, 1997; Hixon et al., 2014; Zellou et al., 2016). Chen's (1997) spectral analysis of nasalized vowels revealed a nasal peak between 250-300Hz (P0) below the first formant. This peak is often referred to as the nasal murmur (Kummer, 2008). Another peak occurs around 800-1000Hz (P1) between the first two formants (Chen, 1997; see also Fujimura & Lindqvist, 1971; Hixon et al., 2014). Nasalization during coarticulation changes the acoustic properties of the vowel (Chen, 1997; Hixon et al., 2014; Zellou et al., 2016). In a nasalized vowel, there is a reduction in the first formant's spectral peak amplitude compared to a non-nasalized vowel (Hixon et al., 2014; Kummer, 2008). These changes in the spectrum are regulated by changes in the opening of the velopharyngeal port, causing changes in the flow of air through the oral and nasal cavity. When the velopharyngeal port is damaged or is not functioning adequately, extra air may escape through the nasal cavity during non-nasal sounds: this results in hypernasality, an oral-nasal balance disorder (Hixon et al., 2014; Kummer, 2008).

1.1. Hypernasality

Hypernasality is a speech disorder that occurs when velopharyngeal closure is incomplete, resulting in excessive sound resonating in the nasal cavity (Zajac & Vallino, 2017). This is particularly noticeable during non-nasal sounds (Peterson-Falzone, Hardin-Jones, & Karnell, 2001). Hypernasality can have structural (e.g. cleft palate, oronasal fistulae, oral cancers) or neurological causes (e.g. dysarthria) (Kummer, 2008; Peterson-Falzone et al., 2001).

Hypernasality can be assessed through many different means. The primary method used for assessment of nasality disorders is auditory-perceptual judgments by a trained listener (Kuehn & Moller, 2000). This method remains relevant and valuable in detecting subtleties in nasality (Kuehn & Moller, 2000). However, auditory-perceptual assessments have been associated with low reliability (Whitehill & Lee, 2008).

As an alternative, Chen (1997) suggested an acoustic measure based on nasal coupling and formant peaks. The acoustic measure compares the amplitude of the first formant (A1) and a low frequency nasal peak (P0). The amplitude of the first formant is reduced during nasalization compared to oral productions (Chen, 1997). Thus, a nasalized vowel would have a lower A1-P0 ratio compared to an oral vowel. The accuracy of the A1-P0 measure is highest when analyzing the nasality of isolated vowels, as opposed to vowels in a word context (e.g. /a/ and /ã/ vs. 'cat' and 'can'). This measure is most helpful when analyzing speech collected using a single microphone. Alternative methods have been developed using more sophisticated equipment.

The Horii Oral-Nasal Coupling (HONC) score (Horii, 1983) is another method used to quantify nasality. This measure compares the vibration of the nose (quantified by an accelerometer mounted on the nose) to the combined oral and nasal acoustic output (quantified by a single microphone). These scores have been highly correlated with experts' auditory-perceptual judgments of nasality (Horii, 1983). The advantage of this technique is the small and lightweight measuring equipment. However, due to its complex technical setup and analysis the HONC score is not widely used in clinics (Horii, 1980, 1983; Laczi, Sussman, Stathopoulos, & Huber, 2005). A more clinically prevalent alternative to the HONC score is nasometry.

Nasometry is measured with two microphones (one for the oral channel and one for the nasal channel) separated by a plate; the plate is positioned on the face between the nose and mouth, and is often held in place with a headset. Nasalance scores consist of the sound pressure level of the nasal channel relative to the total amount of acoustic energy produced by the combined nasal and oral channels (Fletcher & Daly, 1976):

$$\text{Nasalance score (\%)} = \text{nasal} / (\text{nasal} + \text{oral}) \times 100$$

Many instruments exist to record nasalance scores: the Nasometer (KayPentax, Lincoln Park, NJ), the NasalView (Tiger Electronix, Seattle, WA), and the OroNasal System (Glottal Enterprises Inc., Syracuse, NY) (Kummer, 2008). All of these instruments collect relative acoustic pressure levels. However, studies by Bressmann (2005) and Lewis & Watterson (2003) indicate that the latter two nasometry systems (NasalView and OroNasal) provide different scores from one another and from the Nasometer. They are also less widely used in research and clinical settings compared to the Nasometer (Kummer, 2008; Peterson-Falzone et al., 2001). Hence, the present study utilized the Nasometer II 6450 to collect acoustic data and analyze nasalance scores.

1.1.1. Control of Nasality

Naïve listeners are able to perceive and identify nasality in speech (Brunnegård, Lohmander, & Van Doorn, 2009; Fletcher & Daly, 1976). With training, typical speakers are able to adjust their levels of nasality to become hypernasal (Lo et al., 2019; Wong, Tse, Ma, Whitehill, & Masters, 2013). A study by de Boer & Bressmann (2017) revealed that speakers are sensitive to the level of nasality in their own speech. In this study, speakers wore a Nasometer headset and a set of headphones. During the experiment, the participants' speech was recorded and fed back to them via headphones. Following a baseline period, the authors manually altered the auditory feedback received by the speakers either by increasing or decreasing the loudness of the nasal channel. De Boer & Bressmann (2017) found a compensatory response: nasalance scores during the baseline period ranged from 32-34%, during the manipulation period of maximal nasal loudness nasalance scores decreased to 28.30%, and during the manipulation period of minimal nasal loudness the scores increased to 35.73%. These results indicated that individuals, in response to altered auditory feedback and without explicit instructions, adjusted their level of nasality to compensate for their level of nasality that they perceived as incorrect (de Boer & Bressmann, 2017). These findings established that individuals can perceive and alter nasality according to an internal target (an individual's typical nasalance level). An extension of these results may examine whether an external speech signal with varying levels of nasality would evoke an analogous compensatory response.

1.1.2. Social Perception of Hypernasality

Hypernasality can affect intelligibility and acceptability of an individual's speech (Kummer, 2008). Blood and Hyman (1997) presented naïve typical speakers with portraits randomly paired with recordings of typical or hypernasal speech. They found that when asked to rate images paired with hypernasal speech or typical speech, individuals rated images paired with hypernasal speech lower on both personality and appearance. Similarly, Watterson, Mancini, Brancamp, & Lewis (2013) asked children to judge the level of hypernasality of speech samples. They then asked the same children to rate the social acceptability of randomized speech samples. Results indicated that as the perceived level of hypernasality increased, the perceived level of social acceptability decreased. For example, for the "really hypernasal" speech samples, 77% of children disagreed with the statement "this child would fit in with my friends" and 74% of children agreed with the statement "this child would be teased" (Watterson et al., 2013). These two studies provide evidence that naïve listeners make strong social judgements based solely on the presence of hypernasal speech.

From a young age, individuals make social judgments based on their perception of hypernasality in an interlocutor's speech (Watterson et al., 2013). Additionally, individuals compensate their level of nasalance in response to experimentally altered nasal signal levels (de Boer & Bressmann, 2017; Zellou et al., 2016). The literature on the perception of hypernasality indicates that individuals are sensitive to varying levels of nasality in their own speech and in an interlocutor's speech (Brunnegård et al., 2009; Watterson et al., 2013; de Boer & Bressmann, 2017). However, the influence of an interlocutor's speech, in the context of typical nasality and hypernasality on an individual's speech, is underexplored.

1.2. Communication Accommodation

Early research on accommodation focused on Giles and colleagues' Communication Accommodation Theory (CAT) as described in the first section below. CAT inspired other theories focused on external and social influences on accommodation, as described in the second section below. Other researchers extended these findings in order to understand internal/cognitive support for accommodation, relying on automatic perception-production linkages, as described in the third section below. A hybrid approach acknowledges the influence of both internal and external factors on accommodation in speech production.

1.2.1. Giles' Communication Accommodation Theory

Communication Accommodation Theory (CAT) proposes that an individual alters their speech in response to their interlocutor's speech (Shepard, Giles, & Le Poire, 2001). Accommodation occurs when an individual's speech becomes more (convergence) or less (divergence) similar to their interlocutor's speech. Accommodation is "a strategy whereby individuals adapt to each other's communicative behaviours in terms of a wide range of linguistic, prosodic, and non-verbal features" (Giles & Coupland, 1991, p.63). According to CAT, individuals converge to indicate interest and to increase commonality; individuals diverge to increase distance and reduce commonality (Shepard et al., 2001). Gregory and Webster (1996) analyzed interviews between Larry King (a radio and television show host) and interview partners of different social status. They found evidence that social status mediated accommodation between the guest and the host. The results of the study indicated that individuals with lower social status converged to individuals with higher social status (i.e., Larry King) in order to maintain or facilitate communication (Gregory & Webster, 1996). In a different study, individuals diverged to an interviewer (becoming less similar) when prompted by an insult (Bourhis & Giles, 1977). Although divergence is less commonly explored in the literature, it is an important consequence of negative social interactions.

CAT also extends to non-verbal communication, semantics, and phonetic speech features (Giles et al., 1991). Individuals accommodate to their interlocutors' social behaviours, such as nodding, hand gestures, and posture (Giles, et al., 1991). Imitation of actions increases familiarity, empathy, and reciprocal imitation (Shepard et al., 2001). Individuals also semantically accommodate to their conversation partners (Giles et al., 1991). Upon determining the subject of their conversation, individuals begin using similar terms to solidify concepts and facilitate the conversation (Brennan & Clark, 1996; Garrod & Doherty, 1994). Giles et al. (1991) postulated in CAT that phonetic accommodation (alongside non-verbal and semantic accommodation) can facilitate communication. This theory proposed that two individuals in a positive social situation would converge with each other's speech. More precisely, individuals will converge towards the phonetic traits or features (such as pronunciation, prosody, or fundamental frequency) that are most socially desirable. Thus, over time, both individuals will adhere to the most socially acceptable speech features for the given social setting. Conversely, individuals in a negative social situation will create more distance between each other, resulting in diverging speech features.

When an individual perceives a communication barrier, they may resort to overaccommodation (Coupland, Coupland, Giles & Henwood, 1988; Sheppard & Giles, 2001). When overaccommodating, individuals tend to speak loudly, slowly and with exaggerated intonation. This type of speech is intended to be clear and easily understood. However, overaccommodation is often based on the incorrect assumption that since an interlocutor cannot produce typical speech (e.g. due to a speech disorder) they cannot comprehend typical speech (Caporael, 1981; Caporael & Culbertson, 1986). Overacommodation has been described in the speech of staff in retirement homes (Caporael, 1981; Coupland, Coupland, Giles, & Henwood, 1988). Caregivers may use slow, loud, simplified speech in order to communicate with the elderly (called elderspeak). It is unclear whether training may help moderate overaccommodations in the aforementioned circumstances.

In sum, CAT describes and explains how social situations influence speech accommodation. Individuals may converge, or diverge based on the desire to reduce or increase social distance, respectively. CAT theory predominantly focuses on social circumstances motivating speech accommodation, whereas other approaches have focused on internal cognitive influences on accommodation.

1.2.2. Internal Cognitive Models of Accommodation.

Communication accommodation is a process that requires the perception of an interlocutor's speech, and the integration of the perception into an internal representation of speech. This results in a change in speech production. Speech production models such as the Directions Into Velocities of Articulators (DIVA) model (Tourville & Guenther, 2011) may be helpful in understanding the speech production component of accommodation. Though they do not directly address the topic of accommodation, having a framework to understand speech production is useful in also understanding speech production in accommodation. In brief, Tourville & Guenther (2011) postulated that there are two important directions for speech information to travel: through feedforward and feedback channels. Via a feedforward command, top down information from a sound map is sent to the articulators. In typical speech conditions, the information from this motor signal results in the coordinated movement of speech structures (i.e. jaw, tongue, larynx, etc.). As a consequence, auditory and somatosensory information about the produced speech is returned back via a feedback control loop. The feedback control loop sends information to the feedback control map that indicates whether the production speech contained any errors. The feedback information will in turn update the feedforward map, and repeat the previous steps (Tourville & Guenther, 2011).

Relating this model back to communication accommodation, we can use de Boer and Bressmann (2017) as a theoretical example. The participants were asked to produce speech, and their speech was artificially perturbed—nasal signal level was either increased or decreased. Extrapolating from the DIVA model, when the participant first started producing speech in typical feedback conditions the feedforward model was sending typical motor commands. However, when the feedback signal was manipulated and returned information to the speaker that their speech contained an error (increased nasal signal level), the feedforward system was modified to compensate for this perceived error, thus decreasing their level of nasality. This compensatory response is well documented for other speech features such as pitch, loudness, and vowel formants (e.g., Abur et al., 2018; Jones & Munhall, 2000; Munhall, MacDonald, Byrnes & Johnsrude, 2009; Purcell & Munhall, 2006a; Purcell & Munhall, 2006b). Generalizing from models of speech production, accommodation to an external signal can result information from other individuals' speech may impact a feedforward speech motor-command. The feedback is received and gradually adjusts the speech target in response to the interlocutor's speech. In sum, the DIVA model provides a framework to understand how speech is modified in response to altered feedback.

Other internal cognitive approaches emphasize the influence of memory on communication accommodation. Goldinger (1998) suggested that episodic memories of speech affect later perception and production of speech. Episodic memories are a collection of information about episodes and events with temporal-spatial specificity (Tulvin, 1972). Goldinger's (1998) theory proposed that voice-specific phonetic features get stored in an episodic memory and are retrieved later. During speech perception, these features are accessed and allow the interlocutor to have a direct memory of how the words were produced. However, it is important to note that when multiple memories are stored, the idiosyncrasies of individual speakers are washed out, and the system defaults to an average representation of the word and not of interlocutors' individual speech features. The individual's memory of how they perceived speech is most heavily weighted towards the most recent or prominent speaker stored in their episodic memory (Goldinger, 1998). Therefore, an individual will produce speech features similar to their most recent interlocutor, but mostly based on their general memory of how that word is typically produced. This model implies that memories of social and cultural speech norms could play an important role in accommodation.

Overall, the aforementioned theories suggest that internal representations and memories of speech may influence how individuals accommodate. In brief, an individual's speech may be influenced by internal factors, such as systems of speech motor control, and episodic memory systems. These internal factors will, in turn, influence the individual's speech. Following these internal models of communication accommodation, researchers have wanted to explore whether these processes were always modulated by social factors, or whether communication accommodation could be an automatic process.

1.2.3. Automaticity and Accommodation

Given the ubiquity of accommodation, some researchers have proposed that phonetic convergence, in particular, could be an automatic, non-volitional phenomenon. In the current context, automaticity is defined as 'without mediation or modulation by another factor (e.g. social biases or internal cognitive factors)', indicating a direct and mandatory connection between speech perception and production (Bargh & Ferguson, 2000).

Many theories utilize a similar automatic concept to describe the interactions between movement perception and production. On a general level, the common coding theory postulates that perception and production of movements share a common representational code (Prinz, 1990). In other words when perception and production share features (e.g. perceiving and producing speech), then the perception of an external stimulus will influence the planning and production of actions (Prinz, 1997). Based on this concept, Dijksterhuis & Bargh (2001) posited that we use the same representations for language comprehension and language production. "We have a tendency to imitate others because perception automatically elicits corresponding behavior." (Dijksterhuis & Bargh, 2001, p.8). Individuals are thus directly influenced by external contexts such as an interlocutor's speech resulting in changes in the individual's speech.

Pickering & Garrod (2013) proposed an interpersonal link between individuals during conversations based on an interactive loop. The loop includes a reciprocal adjustment between individuals: at each turn in the conversation an individual converges slightly more towards their interlocutor's speech. Pickering & Garrod (2013) posit that during a conversational interaction, an individual hears an interlocutor's speech, and generates an internal representation of that speech. The individual then uses this motor representation of their interlocutor's speech to produce their own speech. Hence, by internalizing a representation of their interlocutor's speech, and due to the common coding of perception and production, an individual will automatically converge to their interlocutor. When an individual perceives their interlocutor, the theory proposes that an individual automatically derives their motor plan based on the perceived motor plan of an interlocutor (Pickering & Garrod, 2004; Pickering & Garrod, 2013). Therefore, the individual—due to the common coding between perception and production—will accommodate their speech to their conversation partner's speech. In sum, when an individual perceives an interlocutor's speech, that perception is automatically linked with their resultant production of speech, indicating that individuals will automatically accommodate to their conversation partner.

In sum, the aforementioned automaticity theories suggest that external influences have a strong impact on speech accommodation. An individual's speech may be automatically, non-volitionally influenced by external factors, such as their conversation partner's speech features.

1.3. Accommodation Research

In early research on phonetic accommodation, it was common practice for researchers to record live conservations between two individuals (e.g. Giles, 1973). This method allowed for an ecologically valid experiment wherein two individuals were genuinely interacting face to face. However, to explore nuanced questions on phonetic aspects of accommodation, newer protocols control the speech exchanged between two individuals more closely. Studies like Pardo (2006) used a task-oriented conversation paradigm (e.g. the map task) to narrow speech content. This task allows researchers to ensure that both interlocutors' speech contain the same words.

Similarly, although not interactive, shadowing and exposure-response tasks can be used to offer greater control over participants' speech. In these tasks a participant is exposed to a model speaker and is asked to either repeat (shadowing) or read (exposure-response) a response out loud. These paradigms are helpful in controlling the exposure material and the responses produced (Goldinger, 1998). They allow the researcher to prescribe the exposure words, the model speakers, and the model speech. Both the shadowing and exposure-response tasks result in utterances that can be analyzed in different ways.

The two most common methods for examining accommodation are perceptual AXB judgments and acoustic measures of difference in distance (DID). In an AXB two-alternative forced choice task listeners are presented with three utterances and asked to judge which of the utterances (the first or the last) is most similar to middle utterance.

Based on a large number of listeners' judgments (e.g. Pardo, Urmanche, Wilman, and Wiener, 2017) the experimenters are able to determine whether or not individuals' speech perceptually converged to another talker's speech (Abel, 2013; Babel, McGuire, Walters, & Nicholls, 2014; Goldinger, 1998; Pardo, 2006).

DID scores quantify a physical change in the speaker's productions at baseline and following exposure compared to the model's speech. These scores provide a sensitive acoustic measure of accommodation and are widely used to quantify accommodation to acoustic features (Babel, 2010; Babel & Bulatov, 2012; Pardo, Jordan, Mallari, Scanlon, & Lewandowski, 2013; Zellou, Scarborough, & Nielsen, 2016). A positive DID score indicates that participants became more similar to models following exposure—speakers converged. Conversely, a negative DID score indicates that individuals became less similar to models following exposure—speakers diverged.

DID scores allow for quantification of the degree of accommodation to a variety of factors. Many factors may influence accommodation. In the following section we will discuss the role of speaker's sex on accommodation.

1.3.1. Sex Differences

Studies on accommodation have examined sex effects (Babel et al., 2014; Bilous & Krauss, 1988; Namy, Nygaard, & Sauerteig, 2002; Pardo, 2006), but the findings are inconclusive. Some studies provide evidence that females accommodate more than males (Babel et al., 2014; Namy et al., 2002). Other studies provide evidence that males accommodate more than females (Pardo, 2006). Overall, the evidence supporting an effect of speaker's sex on accommodation is inconsistent.

Similarly, the effect of model sex on accommodation is not well understood. Some studies have found an effect of model sex (e.g. Babel et al., 2014; Namy et al., 2002). However, because there are inconsistent effects of speaker and model sex on accommodation, studies need to include both female and male speakers and model speakers.

1.3.2. Accommodation to Disordered Speech

Recently, researchers have started assessing accommodation to atypical speech. Borrie & Liss (2014) investigated phonetic convergence to disordered speech. They investigated whether typical speakers would accommodate to speech features of dysarthric (ataxic and hypokinetic) speech. In a quasi-conservational paradigm, they asked individuals with typical speech to listen to a pre-recorded sentence, then read (aloud) a different response sentence. Individuals heard model sentences from individuals with dysarthria and from controls with typical speech. Their findings indicated that healthy individuals accommodated their speech rate and pitch variation to converge to individuals with dysarthria. Their speech converged but remained significantly different from dysarthric speech. The findings revealed that typical speakers accommodate to disordered speech (Borrie & Liss, 2014). The results highlight significant convergence in speaking rate and pitch variation. However, they did not explore other defining features of dysarthria, such as hypernasality. This research by Borrie & Liss (2014) lays the foundation for the current thesis to further explore the topic of accommodation to hypernasal speech.

In sum, the literature presents evidence that many features can influence an individual's speech. In the case of accommodation, an individual's speech is influenced by their conversation partner's speech features. We see this pattern of convergence even in response to atypical levels of rate of speech, and intonation. Nonetheless, many other features of atypical speech remain to be explored. The present studies aims to extend the study of accommodation to disordered speech by exploring the response to increased levels of nasality in speech

1.3.3. Accommodation to Nasality in Speech

Current research exploring accommodation to nasal speech is centered on typical levels of nasality. Zellou et al. (2016) investigated whether individuals would accommodate to increased or decreased nasality in a coarticulated vowel-nasal (VN) clusters. Due to coarticulation, pairing a vowel with a nasal results in a more nasalized vowel. A single male model produced the VN utterances. The level of nasality of the VN clusters was manipulated by altering the relative amplitude of A1-P0 peaks in the typical

VN utterance recordings. In this experiment, the authors posited that by increasing the level of coarticulated nasality in the vowel of the VN syllable, the nasal would be easier to predict, thus facilitating communication. In a shadowing task, results indicated that individuals converged to increased nasal coarticulation, but did not change in response to decreased nasal coarticulation in VN stimuli. These results suggest that individuals converged to an increased level of nasality given a context in which increased nasality facilitated communication (VN clusters) (Zellou et al., 2016).

Another study by Zellou, Dahan, & Embick (2017) extended the findings of Zellou et al., (2016). They recorded word utterances produced by a single male talker, and modulated the A1-P0 ratio using the same method as Zellou et al., (2016). In an exposure-response paradigm, they presented the speakers with naturally nasalized or hyper-nasalized word utterances, then asked the speakers to read a different printed word. They used A1-P0 DID scores to quantify changes in nasality following the exposure-response paradigm. Results indicated that speakers converged to both increased and decreased levels of nasality.

In sum, individuals are able to perceive varying levels of nasality in speech. Speakers, following a shadowing or exposure-response paradigm, adjusted the level of nasality in their speech to converge with the hyper-nasalized model speech. It is important to note that in both of these studies the hyper-nasalized speech could facilitate communication since it was occurring in a context typically nasalized (VN syllables). The current thesis explores the impact of pathological hypernasality on the interlocutor's speech, a context in which nasalization may impede communication.

1.3.4. Effects of Speech-Language Pathology Training

Professional training may play a role in speech accommodation. Speech-Language Pathologists (SLPs) are trained to actively listen to the concerns of the patients and empathize with their patients. During active empathetic listening, SLPs will notice elements of a patient's speech (such as hypernasality) (McNamara, 2014). It is not known whether, in an attempt to amenable and compassionate towards patients, SLPs may show patterns of overaccommodation, as described for other medical professionals (Caporael, 1981). Since SLPs are trained to listen more attentively, they will notice more elements in the patient's speech. Based on the automatic theories of accommodation, they may in turn be more likely to accommodate to those features. The question remains to be explored whether or not SLPs, who are trained to be sensitive to speech features, will accommodate differently to disordered speech than untrained controls.

1.3.5. Combining Internal and External Support for Accommodation

Neither external nor internal theories fully explain why individuals accommodate to disordered speech (Borrie & Liss, 2014; Späth et al., 2016). Speech disorders are not socially desirable (Blood & Hyman, 1977; Lass et al., 1995). However, individuals still converged to them. Their speech becomes more similar to the disordered speech, but does not match the level of severity (Borrie & Liss, 2014)

Alternatives to exclusively external or internal theories of accommodation have been proposed. Babel, McGuire, Walters, and Nicholls (2014) suggested a cognitive novelty theory. Cognitive novelty theory postulates that we converge more with novel words, speech features, and conversations partners. Newer and less familiar words are not well established in an individual's lexicon and are more open to new interpretation. Therefore, the novel speech may trigger a quasi-automatic convergence response. Nonetheless, in response to disordered speech, individuals do not accommodate to match the degree of severity present in disordered speech (Borrie & Liss, 2014), indicating that there may be cognitive control or auditory targets moderating speech productions.

An important factor to consider is the impact of social biases and cognitive heuristics on speech. Since social factors impact accommodation (Gregory & Webster, 1996). Pickering and Garrod (2013)'s model proposes an automatic perception-production loop. Automatic models of accommodation do not allow flexibility to incorporate non-automated cognitive processes.

Babel (2010) addresses the question of automaticity by providing evidence that individuals both converge and diverge based on social biases. One goal of the study was to establish the influence of social factors on phonetic accommodation across different dialects. For the study, authors recruited both male and female individuals from New Zealand. Individuals were asked to fill out a questionnaire assessing their biases towards Australians (to establish whether they were pro-Australian or not). Afterwards, the individuals participated in a word shadowing task, in which the model talker was Australian. DID scores quantified the level of accommodation. In brief, individuals with pro-Australian tendencies were more likely to accommodate to the Australian models compared to their peers with less affinity for Australians. These results indicate that social biases influence accommodation (Babel, 2010).

In summary, theories of accommodation began primarily as a socio-linguistic phenomenon (Shepard et al., 2001) evolved into more automatic phonetic processing models (Pickering & Garrod, 2013), then developed to include elements of both social and automatic integration of external stimuli (Babel, 2010). In the current thesis, we focused on phonetic accommodation, while keeping in mind the impact of both automatic phonetic processing and social bias. More concisely, the current thesis focused on phonetic accommodation to hypernasal speech.

1.4. Hypotheses

The aim of the current study was to determine the effect of hypernasality on a listener's speech. We aimed to test three hypotheses in this thesis:

H1: There is evidence that individuals converge to features of dysarthric speech (Borrie & Liss, 2014) and increased nasality in VN syllables (Zellou et al., 2016). Individuals also compensate to changes in auditory feedback of their own nasal signal levels (de Boer & Bressmann, 2017). Since hypernasality is negatively perceived (Blood & Hyman, 1977) and social biases influence accommodation (Babel, 2010), we hypothesized that speakers would converge to nasality in typical speech and diverge from hypernasal speech.

H2: There is conflicting evidence on effects of sex on accommodation. Some studies show evidence to support sex differences (Babel et al., 2014; Bilous & Krauss, 1988; Stephen D. Goldinger & Azuma, 2004; Miller, Sanchez, & Rosenblum, 2010; Namy et al., 2002; Pardo, 2006) whereas other studies report no effect of sex (Pardo et al., 2013, 2017). If females do accommodate more than males (Babel et al., 2014; Namy et al., 2002), then we hypothesized to that female speaker would show greater accommodation than male speakers and that there would be greater accommodation in response to female model talkers than to male model talkers.

H3: Since, SLPs are trained to be empathic and actively listen to their patients (McNamara, 2014)—thus resulting in a more detailed perception of nasality—we hypothesized that SLP students would accommodate more compared to typical untrained sex-matched controls.

1.5. Overview

We trained voice actors to produce hypernasal speech, and recorded them producing a set of 5 sentences (Appendix A, from Borrie & Liss, 2014). We also recorded control models producing the same sentences with their typical speech. Next, we presented these stimuli from the model talkers to the speakers in a quasi-conversational paradigm as per Borrie & Liss (2014). Speakers were asked to listen to the sentences, then read a different sentence out loud. We recorded and calculated their level of nasality with a Nasometer headset. From these nasalance scores, we calculated DID scores to quantify accommodation. The results provide insight into the influence of disordered speech on a conversation partner's typical speech.

2. Methods

This methods section has been divided into two sections. The first section will describe the data collection methods used to acquire the model speech. The second section will describe the main experimental methods.

2.1. Speech Model Stimuli

2.1.1. Stimuli Speakers (Controls and Actor Models)

To generate hypernasal speech samples, we recorded 4 professional voice actors (2 males, 2 females) with accents typical to southern Ontario. First, the actors were instructed to read the 5 sentences (Borrie & Liss, 2014) (Appendix A) with their typical voice. We used the typical-speech recordings as pacing cues for the hypernasal recordings. The voice actors were coached in simulating hypernasal speech, then asked to read the 5 sentences (Appendix A) with simulated hypernasality. We used these recordings as the hypernasal stimuli. The typical recordings from the voice actors were not used in the experiment to avoid confusing the speakers. We did not present both hypernasal and typical speech from the same actor since speakers may have become distrustful of the hypernasal speech. Therefore, we recorded four healthy typical voice models (2 males, 2 females) with accents typical to southern Ontario. The healthy voice models were asked to produce the 5 sentences from Borrie & Liss (2014) (Appendix A) using their typical voice. We used these recordings as the control stimuli. Voice actors and typical models wore a Nasometer headset to record nasalance scores. A condenser microphone collected the acoustic signals (see section 2.2.3 Randomization & Recording procedures below for more details).

2.1.2. Stimuli Results

Table 1 contains the models' mean nasalance score per condition and sex. We used the Number Cruncher Statistical Software Version 8 (NCSS LLC, Kaysville, UT) for the statistical analyses. We ran a two-way Analysis of Variance (ANOVA) which revealed a main effect of condition (F (1,7) = 262.04, p < .001, η^2 = .87). There was no main effect of model sex (F (1,7) = .01, p = .97, η^2 < .001) and no interactions between the two (F (1,7) = 1.71, p = .20, η^2 = .006). The main effect of condition indicates that the voice actors successfully portrayed distinct levels of hypernasality compared to the control actors (See Figure 1). According to a study by McHenry & Liss (2006), the mean nasalance score for 11 individuals with dysarthria was 68%; the models' simulated hypernasality was within a comparable range of severity.

Manipulation	Model Sex	N	Mean Nasalance score (%)	SD
Hypernasal	F	10	74.0	8.45
Hypernasal	M	10	70.8	8.34
Control	F	10	28.5	7.04
Control	M	10	32.1	8.95

Table 1: Mean percent nasalance scores of the experimental stimuli. N represents the number of utterances included in each group (five sentences per model; two models per line in the table).

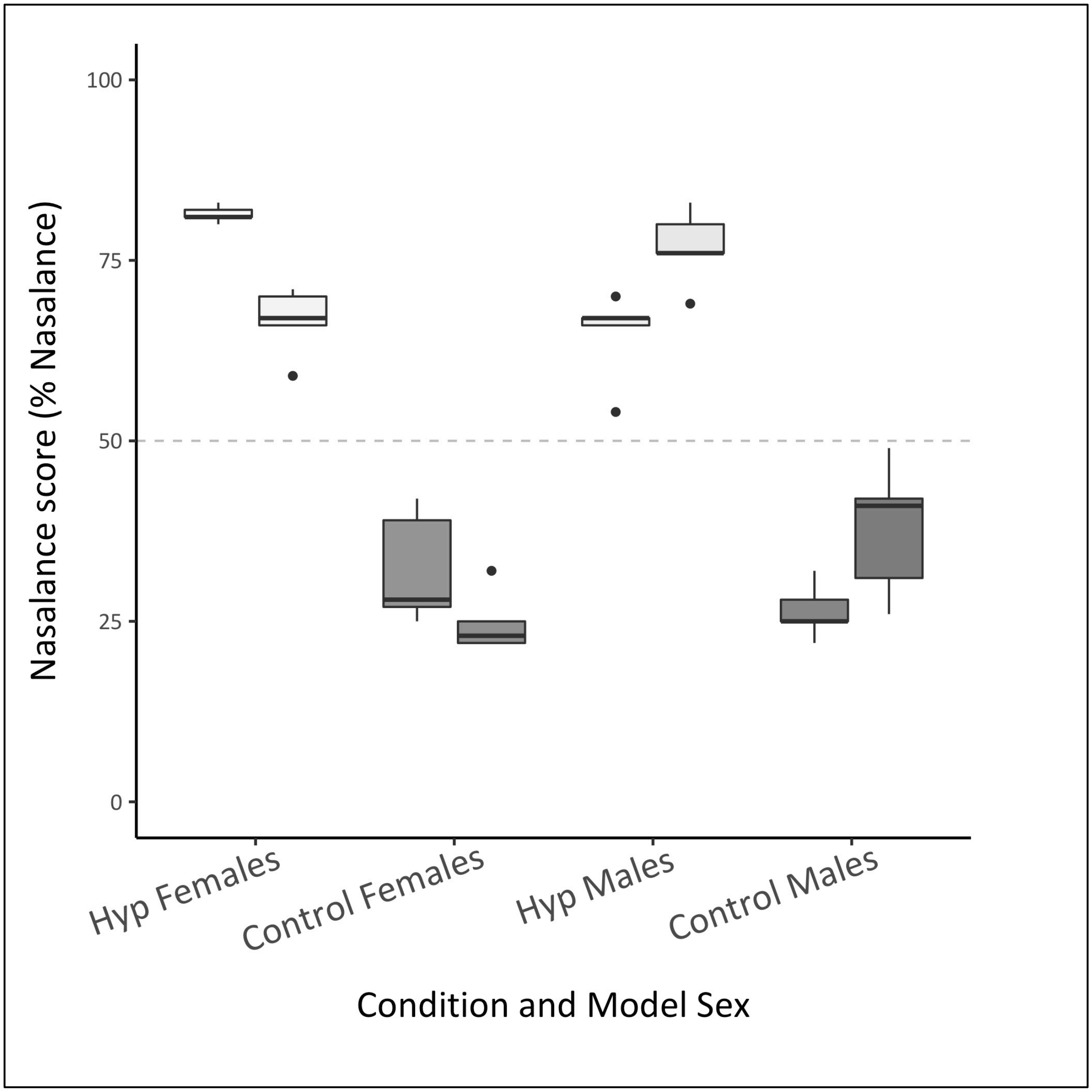

Figure 1- Box plots of mean percent nasalance scores distributed by actor and condition. Each shaded box represents the utterances of one model talker. The four boxes on the left represent the four female models' utterances; the four boxes on the right represent the four male models' utterances. The light boxes represent modelled hypernasal speech (Hyp); the darker boxes represent control speech.

2.2. Experimental Methods

2.2.1. Speakers

We recruited 30 speakers, 11 males and 19 females, including 9 female SLP students. The SLP students were enrolled in the Master's of Health Sciences graduate program at the University of Toronto. The SLP students had completed at least the first 8 months of study in the graduate training program. Speakers were all between 18 and 30 years old. All speakers signed a consent form and filled out a short demographics survey before participating in the experiment. Speakers were not made aware of the purpose of the study until the procedure was complete. Once debriefed on the purpose of the study, all speakers were given the opportunity to withdraw their data from the study or re-consent to their participation. All speakers chose to sign the re-consent form. Speakers were compensated $15 for their time. This study was approved by the research ethics board at the University of Toronto.

2.2.2. Procedures

Speakers wore the nasometer headset. They were seated at a comfortable distance from the computer and asked to hold their heads against a forehead rest. Speakers were told that during the experiment they would hear a sentence followed by a sentence presented on the screen that they would have to read out loud. Speakers were told that for the first ten trials they would not hear a sentence before being instructed to read the sentences out loud, but for subsequent trials they would hear a sentence first, then would be prompted to read the sentence on the screen out loud. To pass to the next trial speakers were instructed to press the space bar on the keyboard in front of them.

Stimuli and response sentences were pseudo-randomized to ensure that speakers did not directly repeat the sentence they had just heard. Four block randomizations orders (Table 2) were designed to control for presentation order. Speakers were evenly distributed across randomization orders. Blocks were presented three times to each speaker.

Trial #	Order 1	Order 2	Order 3	Order 4
1-5	F.1-Control	F.2-Hypernasal	M.1-Control	M.2-Hypernasal
6-10	F.2-Hypernasal	F.1-Control	M.2-Hypernasal	M.1-Control
11-15	F.3-Control	F.4-Hypernasal	M.3-Control	M.4-Hypernasal
16-20	F.4-Hypernasal	F.3-Control	M.4-Hypernasal	M.3-Control
21-25	M.1-Control	M.2-Hypernasal	F.1-Control	F.2-Hypernasal
26-30	M.2-Hypernasal	M.1-Control	F.2-Hypernasal	F.1-Control
31-35	M.3-Control	M.4-Hypernasal	F.3-Control	F.4-Hypernasal
36-40	M.4-Hypernasal	M.3-Control	F.4-Hypernasal	F.3-Control

Table 2: Intra-block randomization of stimuli presentation order. Female (F) and male (M) models were assigned a number, to differentiate them in this table. Presentation order was sequentially assigned to speakers ensuring equal numbers for all presentation orders.

Stimuli were presented via an ASUS laptop for the first half of speakers, then on an equivalent HP Laptop for the second set of speakers. Stimuli were presented over a set of SHL3060bk/28 Philips headphones (Philips, Amsterdam, Netherlands). The experiment was set up using the open source software OpenSesame 3 (Mathôt, Schreij, & Theeuwes, 2012). The experiment consisted of three parts: pre-experimental, the exposure-response paradigm, and post-experimental. The pre- and post- experimental trials (ten at the beginning for the pre-experimental, and ten at the end for the post-experimental trials) served as baselines. The middle part of the experiment was a quasi-conversational task. For this experimental portion, the model recordings were presented and speakers were asked to read (from the laptop's screen) a different sentence than the one they heard. Speakers heard three repetitions of every stimulus (three repetitions of the blocks outlined above) for a total of 120 exposure-response trials. The experiment took approximately 20 minutes.

Speakers' sentences were recorded with an APEX 435 gold diaphragm condenser microphone placed approximately 15cm perpendicular to the speaker's mouth. The signal from the microphone was recorded to a TASCAM Pocketstudio DP-008 (TEAC America, Inc., Montebello, CA, USA), then fed back into the laptop to be recorded through the OpenSesame3 software.

Additionally, the Nasometer 6450 (KayPentax, Lincoln Park, N.J., USA) was used to measure nasalance scores. We recorded the signal from Nasometry headset directly to a digital TASCAM linear PCM recorder DR-05 (TEAC America, Inc., Montebello, CA, USA). The signal was later segmented using SONY Sound Forge software (Sony Canada, Toronto, ON) and analyzed using the KayPentax software (KayPentax, Lincoln Park, N.J., USA).

2.2.3. Data analysis

Data were compiled and analyzed in Rstudio statistical software (Rstudio version 0.99.903). We ran a mixed-effects ANOVA with Nasalance DID scores as the outcome variable, Condition (hypernasal and control speech) and Model sex were repeated variables, and Speaker sex and SLP training were between participant variables. Fisher's Least Significant Differences (FLSD) confidence intervals were used to determine significance in post-hoc analyses.

To quantify the effects of an interlocutor's speech on a speaker's utterances, we calculated DID scores (Babel & Bulatov, 2012; Pardo et al., 2017). In order to calculate DID scores, we subtracted the model's nasalance score from the speakers' baseline utterances. This allowed us to determine how different the speaker was at baseline from the model's speech. Every sentence was subtracted independently. Next, we calculated the difference between the model's nasalance scores and the prompted utterances from the speakers. This difference allowed us to determine how different the speaker was from the model's speech after being exposed to their speech. Finally, we took the absolute values of these two differences and subtracted them using the formula: DID = |Baseline Difference| - |Exposure-Response Difference|. The difference between the baseline and response conditions resulted in the DID scores. If the DID score was positive then speakers were more similar to models after being exposed to their speech—indicating convergence. Conversely, if the DID score was negative then speakers were less similar to models after being exposed to their speech—indicating divergence.

3. Results

We organized the results into three different sections to assess accommodation. The first section contains the analysis including all of the predictor variables, including both conditions, both model talker sexes, across all three speaker groups (female non-SLPs, female SLPs, and male non-SLPs). In order to isolate the influence of speaker sex on the nasalance DID scores, we removed female SLPs from the second analysis. The second section contains an analysis across condition and model sexes, but pertaining only to speaker sex (including only female non-SLPs, and male non-SLPs). In order to isolate the influence of speaker training on the nasalance DID scores, we removed male non-SLPs from the third analysis. The third section contains an analysis across condition and model sexes, but pertaining only to speaker SLP training (included only female non-SLPs, and female SLPs). The outcome variables for all analyses were the nasalance DID scores in percent nasalance. A positive mean DID score indicated convergence, and a negative mean score indicated divergence. If the mean DID score is not different than zero this indicated that no accommodation occurred.

3.1. Complete Analysis

Appendix B consists of two tables containing the means and standard deviations of the nasalance DID scores per speaker group and condition. Figure 2 displays the means for males, females and SLP students (all females) grouped by condition, model sex, and speaker sex. The means of the nasalance DID scores were more extreme for female models in both the hypernasal (M = -9.33, SD = 7.03) and control (M = 3.50, SD = 8.98) conditions compared to males (hypernasal: M = -8.73, SD = 6.59; control: M = 1.59, SD =9.01).

Figure 2 displays the mean nasalance scores across conditions and models sex, and grouped by speakers' group. The error bars represent 95% confidence intervals based on Fisher's Least Significant Difference (FLSD). The overall trend in the hypernasal condition is that speaker's mean scores are all negative, indicating consistent divergence across all speaker groups. In the control condition most of the means are greater than 0, indicating convergence. However, according to the FLSD 95% confidence intervals when responding to the male models, the male non-SLPs are not different than 0, and the female non-SLPs are only slightly above 0, indicating inconsistent convergence across speaker groups in the control condition.

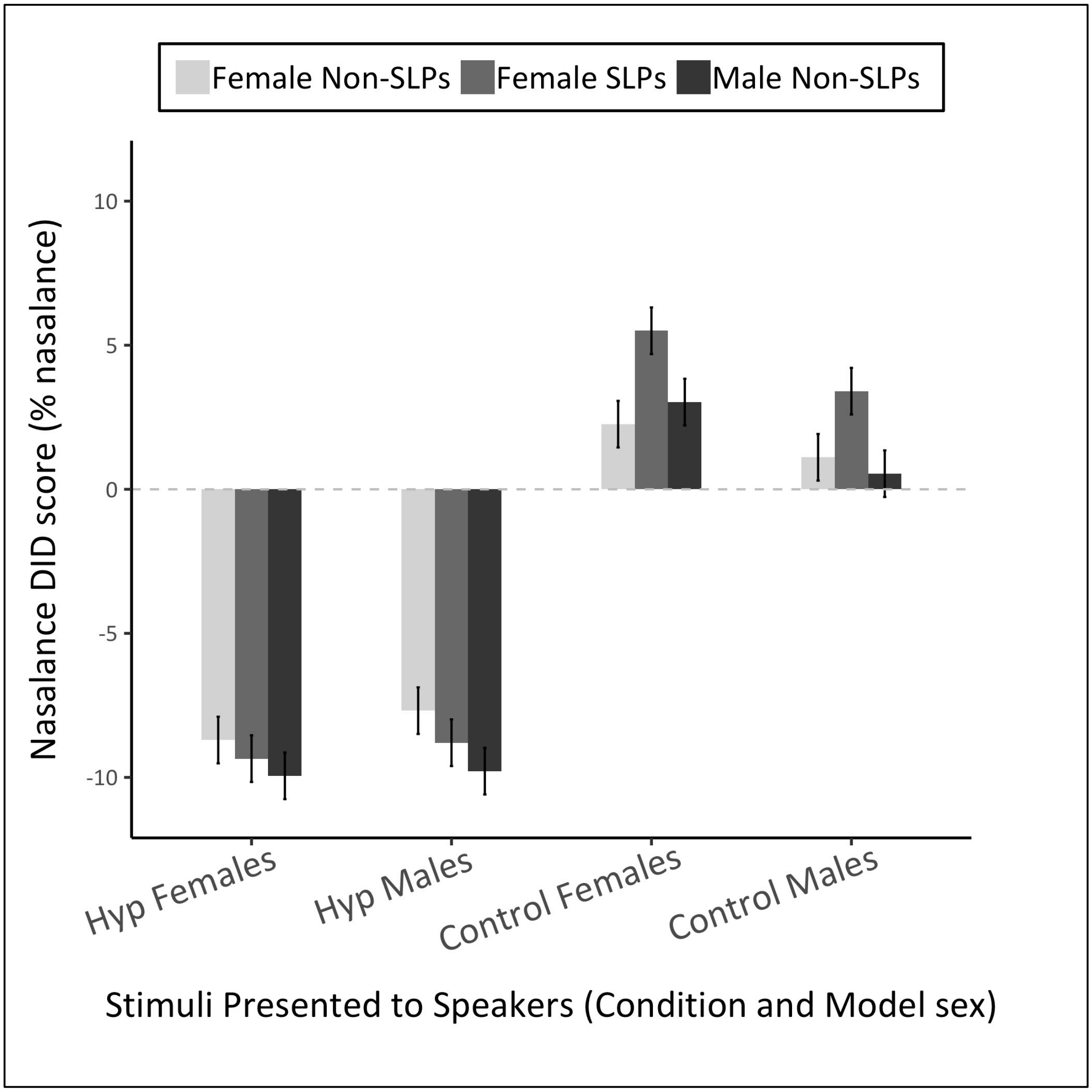

Figure 2. Nasalance DID Scores across conditions, model sex, and speaker groups. Above are displayed the mean nasalance DID score (in % nasalance) for the each stimuli type presented to the speakers: hypernasal (Hyp) and control for both male and female models. The shaded bars each represent a different speaker group: light-grey bars represent female non-SLPs, medium-grey bars represent female SLPs, and dark-grey bars represent Male non-SLPs. Error bars represent 95% confidence intervals from FLSD analysis (0.81).

To test our main hypothesis that speakers will diverge from hypernasal speech we used a mixed-effects ANOVA to investigate the trends in the nasalance DID scores. The outcome variable investigated was nasalance DID score. The within subject variables were condition (hypernasal and control) and model sex (male and female). The between subject variables were speaker group (female non-SLP, female SLP, and male non-SLP). We found a main effect of condition (F (1, 27) = 124, $p < .001$, $\eta_g^2 = .75$) and model sex (F (1, 27) = 10.4, $p = .003$, $\eta_g^2 = .01$), but no effect of speaker group (F (2, 27) = 1.98, $p = .16$, $\eta_g^2 = .04$). Additionally, we found an interaction between condition and model sex (F (1, 27) = 59.85, $p < .001$, $\eta_g^2 = .03$).

For the main effect of condition the mean nasalance DID scores were significantly lower in the hypernasal condition (M = -9.08, SD = 3.32) than in the control condition (M = 2.55, SD = 3.64). This result indicated that speakers diverged in the hypernasal condition, and converged in the control condition. Additionally, effect size reported in the ANOVA for this main effect ($\eta_g^2 = .75$) was large indicating a robust effect of condition.

For the main effect of model sex the mean nasalance DID scores were significantly lower in response to males (M = -3.62, SD = 2.26) than to females (M = -2.91, SD = 1.99). This result indicated that speakers diverged more to male speakers than to female speakers. Additionally, it is important to note that the effect size for this main effect was small ($\eta_g^2 = .01$).

Figure 3 reflects the means of nasalance DID score by condition and model's sex. Speakers accommodated more to female models than to male models. The means displayed in Figure 3 indicate that in the hypernasal conditions speakers diverged similarly from both female and male models. However, in the control condition speakers converged more to female models compared to male models. Again, the effect size for this interaction was small ($\eta_g^2 = .03$).

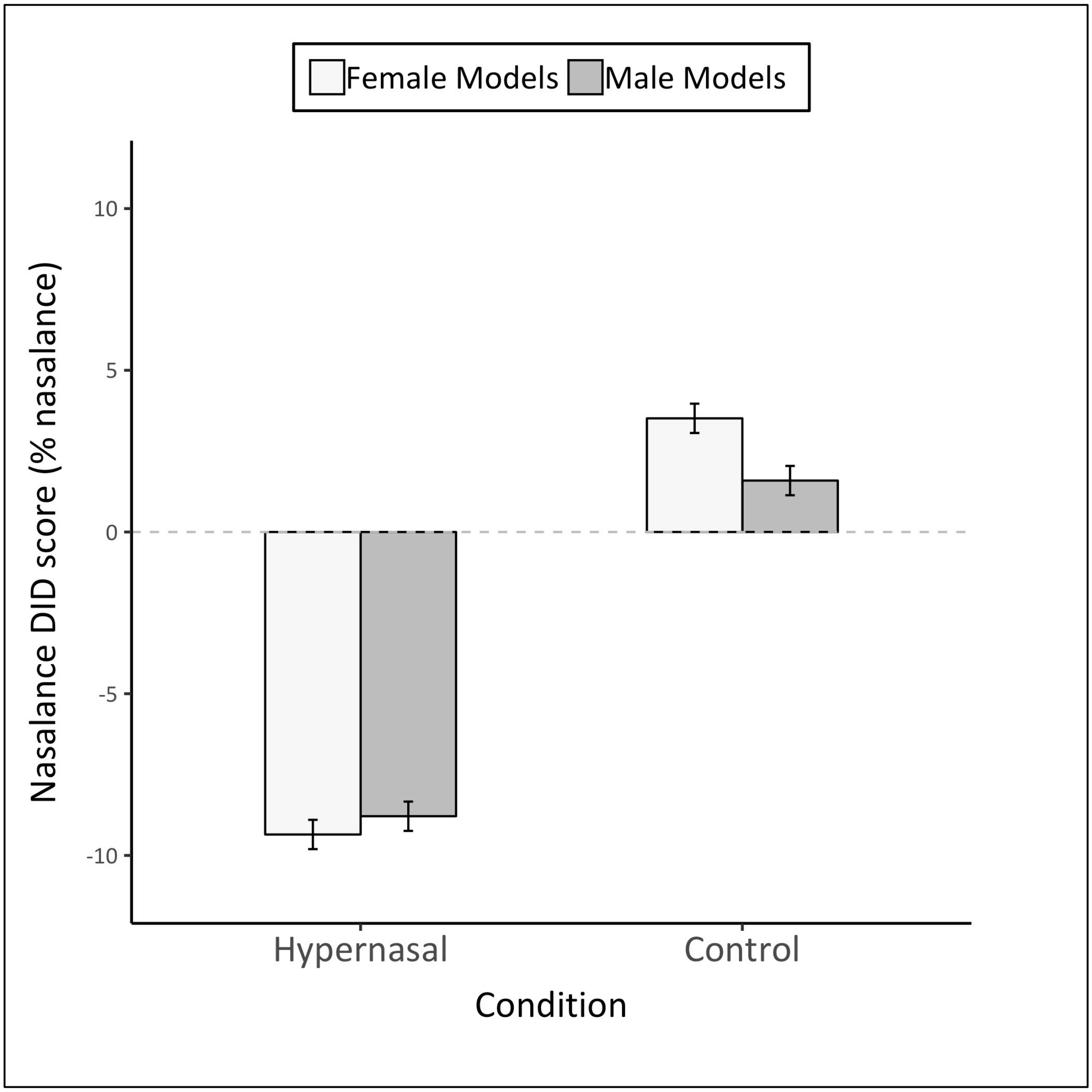

Figure 3. Mean Nasalance DID scores displayed by Condition and Model Sex. Above are displayed the mean nasalance DID scores (in % nasalance) across conditions (hypernasal and control). The shaded bars represent different represent models' sex: light-grey bars represent female models and medium-grey bars male models. Error bars represent 95% confidence intervals from FLSD analysis (0.45).

3.2. Effect of Speakers' Sex

Figure 4 displays the mean nasalance scores across conditions and models' sex, grouped by speakers' sex. These results do not include the female SLP group present in the previous analysis. In the hypernasal condition, speakers' mean scores were negative, indicating consistent divergence for both model sexes. The error bars represent 95% confidence intervals based on FLSD. The error bars indicate that in the hypernasal male model condition, male speakers may have diverged more than female speakers. In the control condition, speakers converged to female control speech. Whereas, according to the FLSD error bars, only the female speakers converged to the male control condition.

In order to better understand the effect of speaker's sex in accommodation to hypernasal speech we ran a second mixed effects ANOVA. To best analyze the effect of speaker sex, we removed the group of trained SLPs from this analysis (Figure 4). The outcome variable investigated was nasalance DID score. The within subject variables were condition (hypernasal and control) and model sex (male and female). The between subject variable was speaker sex (male and female). We did not find a main effect of speaker sex (F (1, 19) = .71, p = .41, η_g^2 = .01). All other main effects and interactions were repeated in this ANOVA: a main effect of model sex (F (1, 19) = 5.31, p = .03, η_g^2 = .007), a main effect of condition (F (1, 19) = 63.39, p < .001, η_g^2 = .69), and an interaction between model sex and condition (F (1, 19) = 33.72, p < .001, η_g^2 = .03).

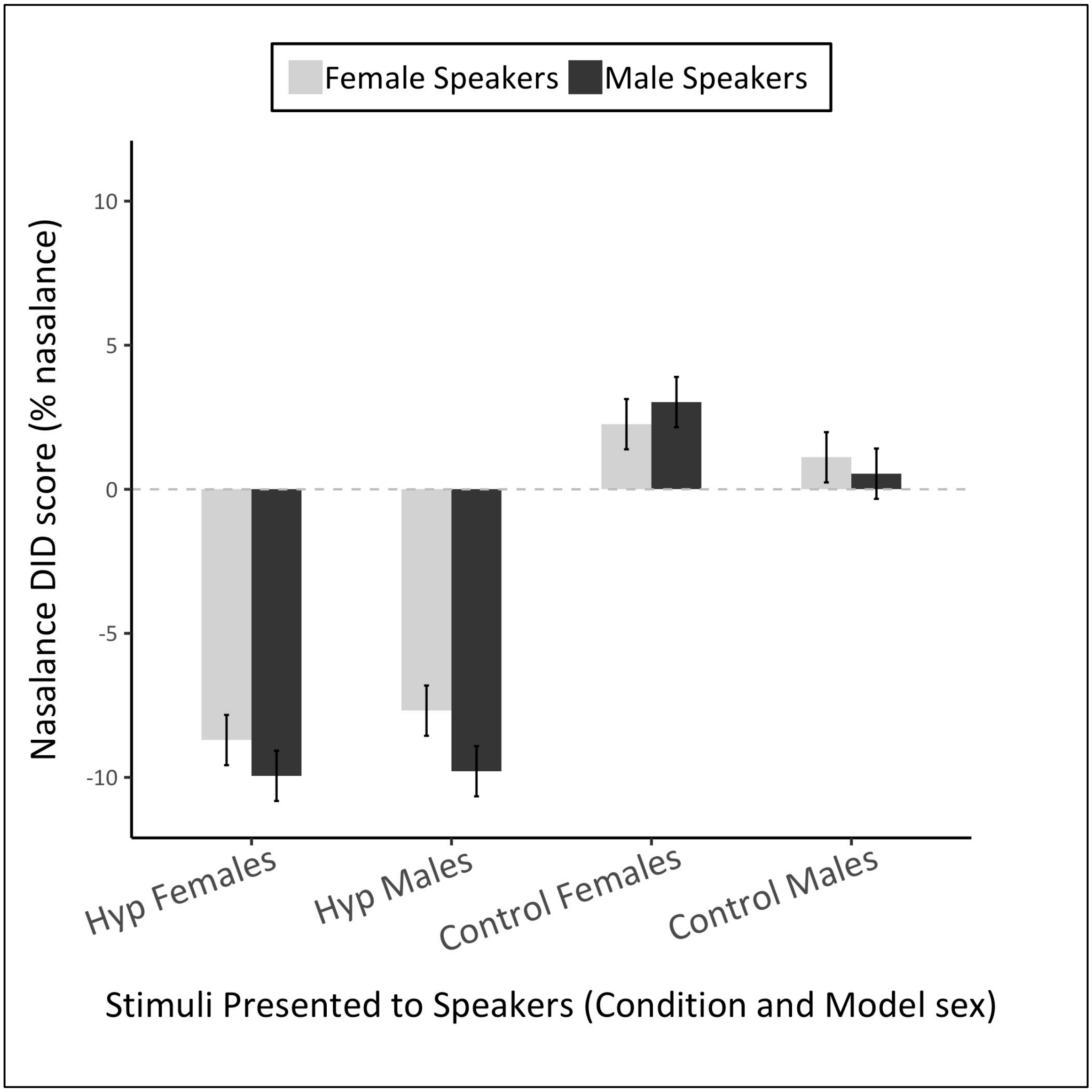

Figure 4. Mean Nasalance DID scores across conditions, model sex, and speaker sex. Above are displayed the mean nasalance DID score (in % nasalance) for the each stimuli type presented to the speakers: hypernasal (Hyp) and control for both male and female models. The shaded bars represent different represent speakers' sex: light-grey bars represent female non-SLPs, and dark-grey bars represent male non-SLPs. Error bars represent 95% confidence intervals from FLSD analysis (0.87).

3.3. Effect of Speakers' SLP status

Figure 5 displays the mean nasalance scores across conditions and models' sex, grouped by speakers' SLP training. These results do not include the male non-SLP group present in the previous analyses. The error bars represent 95% confidence intervals based on FLSD. The error bars indicate that in the hypernasal condition, speakers' mean scores were negative, indicating consistent divergence for both model sexes across both SLP statuses. In the control condition speakers converged to both males and female models. However, according to the FLSD error bars only the female SLPs appeared to converge more overall when compared to the female non-SLP group.

In order to better understand the effect of SLP training on accommodation to hypernasal speech we ran a third mixed-effects ANOVA. To best analyze the effect of SLP training, we removed the group of male non-SLPs from this analysis (Figure 5). The outcome variable investigated was nasalance DID score. The within subject variables were condition (hypernasal and control) and model sex (male and female). The between subject variable was speaker sex (male and female). We did not find a main effect of SLP training (F (1, 17) = 1.03, p = .32, η_g^2 = .02). We also did not find a main effect of model sex in this analysis (F (1, 17) = 2.02, p = .17, η_g^2 = .004). All other main effects and interactions were repeated in this ANOVA: a main effect of condition (F (1, 17) = 76.12, p < .001, η_g^2 = .74), and an interaction between model sex and condition (F (1, 17) = 33.49, p < .001, η_g^2 = .03).

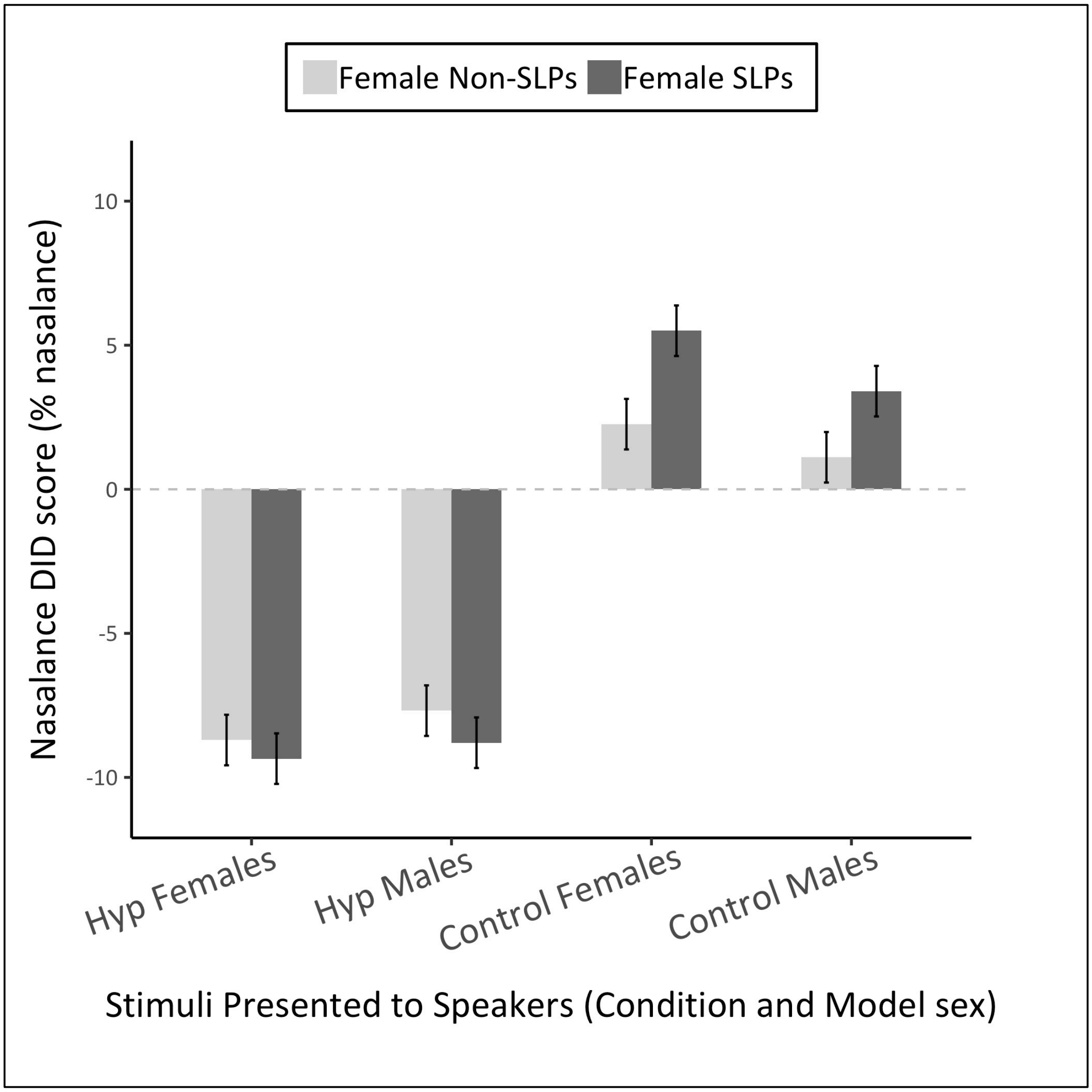

Figure 5. Mean Nasalance DID scores across conditions, model sex, and SLP training. Above are displayed the mean nasalance DID score (in % nasalance) for the each stimuli type presented to the speakers: hypernasal (Hyp) and control for both male and female models. The shaded bars represent different represent speakers' SLP status: light-grey bars represent female non-SLPs, and medium-grey bars represent female SLPs. Error bars represent 95% confidence intervals from FLSD analysis (0.88).

4. Discussion

The aim of the current study was to determine the effect of hypernasality on an interlocutor's speech. The first hypothesis was that speakers would converge to typical nasality, and diverge from hypernasality in speech. The second hypothesis explored the question of whether speaker or model sex had an influence of accommodation. The third hypothesis was that speaker SLP student training would influence speakers to accommodate more than untrained speakers. We explored these hypotheses by training and recording models to produce hypernasal and typical speech. Then we played those stimuli to speakers who were asked to listen and read a sentence in response. We recorded nasalance scores using the nasometer, and quantified the speakers' level of accommodation using DID scores.

4.1.Model Speech

The goal of this thesis was to test the hypothesis that speakers would diverge from hypernasal speech. In order to best control the model speech, we used voice actors. Specific simulations by trained actors allowed for better experimental control of the model speech. We trained actors to modulate their nasality while maintaining their other voice features constant. This controlled manipulation allowed us to clearly distinguish the hypernasal and control groups. Specific simulations by trained actors allowed for better experimental control of the model speech. The modelling of hypernasal speech was preferable in order to isolate hypernasality as a speech feature. In clinical speech disorders, hypernasality may co-occur with other features, such as voice or articulation disorders (Kummer, 2008). Co-occurrence of other speech disorders would have limited our understanding of the mechanisms specifically in response to hypernasal speech.

In the analysis of the stimuli, we found no difference between the two sexes, indicating that male and female models were both equally successful at manipulating their level of nasalance. The main effect of condition between hypernasal and control conditions indicated that the actors successfully manipulated their nasalance score. It is important to note the large difference of 41.9 points in nasalance score between the mean for the hypernasal and the control group. In manipulating nasality, we prompted actors to model severe hypernasality. In a study by McHenry & Liss (2006) looking at hypernasality in dysarthria, the mean nasalance scores (when reading an oral passage) was 68% nasalance, and the most severe case was 89% nasalance score. In the present study the mean nasalance score produced by voice actors in the modelled hypernasality condition was 72.4% (SD = 7.99). We used severe examples of hypernasality in order to elicit a strong response from the speakers.

4.2. Main Discussion

The results provided supporting evidence for our hypothesis that speakers would converge to typical nasality and diverge from hypernasality in speech. The main effect of condition and the following post-hoc tests indicated that individuals diverged consistently when exposed to hypernasal speech, and converged inconsistently when exposed to control "typical" speech. The differences in accommodation between the hypernasal and control speech indicated that individual speakers were able to perceive the two levels of nasality in the stimuli; they also responded by altering their own oral-nasalance balance.

The current results indicate a large change in nasalance score when exposed to hypernasal speech. De Boer and Bressmann (2017) found a compensatory response of approximately 5% nasalance score in response to maximally altered auditory feedback. The current results show a large difference of 11% between the mean scores in response to hypernasal and control speech. However, just like in de Boer & Bressmann (2017), individuals do not accommodate to an extreme, or by using their full possible range of nasality. Individuals' nasalance scores in response to the model's nasalance score did not match or mirror the exact severe score; their scores were more moderate.

Speakers are sensitive to hypernasal speech (Brunnegård et al., 2009; Fletcher & Daly, 1976), and thus in turn diverged from the hypernasal speech. In the current study, it is important to note that the contrast between the hypernasal and typical speech was not subtle. It may be due to the lack of gradual introduction of the hypernasality that divergence, rather than convergence occurred. Future studies should investigate the influence of a gradual increase of hypernasality in the model speech over the course of the experiment on a speaker's nasalance score.

The current results both complement, and contrast with, the findings of (Zellou et al., 2016), who found that speakers converged to increased nasal coarticulation in vowel-consonant pairs. The current results showed evidence that the speakers inconsistently converged with typical levels of nasalance in model speakers. However, in response to hypernasality, speakers diverged from the model speech, in contrast to Zellou et al., (2016) and Zellou et al., (2017). The aforementioned studies used hyper-nasalized speech—speech with increased nasality—but the presence of increased nasality in a VN context is typical, and therefore cannot be directly compared to pathological hypernasality—the presence of increased nasality in typically oral speech sounds. Since hyper-nasalized VN clusters do not impede communication and may even enhance it, according to CAT and internal models of accommodation, convergence may facilitate social cohesion. However, the stimuli used in the current study were modelled severe hypernasality. The level of nasality presented in the current study was not meant to facilitate communication, but rather help us understand phonetic accommodation in the presence of this oral-nasalance balance disorder. Hence, the difference between convergence to hyper-nasalized VN clusters, and divergence to modelled hypernasal speech may be due to the hindrance of hypernasal speech to communication.

In the current study, according to the DID scores, the degree of divergence was around twice as strong as the degree of convergence. It is possible that divergence to hypernasal speech is based on the social stigmatization and undesirable nature of speech disorders (Blood & Hyman, 1977). However, our experiment does not directly test this hypothesis. Further research quantifying speakers' reactions, social biases, and attitudes toward modelled speech is needed to explore that question. We can however argue that, according to CAT theory (Shepard et al., 2001), since hypernasal speech is not socially desirable (Lewis et al., 2003; Watterson et al., 2013), individuals should diverge from hypernasal speech. By diverging, speakers create distance between their own and an interlocutor's speech in order to strengthen the social distance between them (Shepard et al., 2001).

The current findings that speakers diverged from hypernasal speech differ from previous research showing only evidence for convergence to disordered speech. Borrie and Liss (2014)'s results presented evidence that individuals converge to fundamental frequency variability and rate of speech in dysarthric speech. However, individuals did not converge to the point of imitating disordered speech, their speech remained in a typical range (Borrie & Liss, 2014). The difference between the convergence presented in Borrei and Liss, (2014) and the divergence presented in the current results may be related to the speech feature explored. Future studies should explore accommodation to multiple speech features including rate or speech, pitch variability, and level of nasality, to determine whether divergence only occurs to oral-nasal balance disorders. The current study complements the Borrie & Liss (2014) findings by extending the study of accommodation to hypernasal speech.

The main effect of model sex and subsequent post hoc tests indicated that speakers accommodated more following female models' speech than to males'. However, the main effect of sex is probably best understood when interpreted in light of the interaction between condition and model sex. This indicated that, in the hypernasal condition, individuals diverged more from female models than from the males. In the control condition, individuals converged more to the female speakers. Thus, individuals accommodated to a greater degree to the female models. The literature is sparse in studies that looked at the effect of model on the speakers' response (Pardo et al., 2017). However, Babel et al. (2014) showed an interaction between model sex and shadower sex, which is echoed in the current study. Future studies should continue to include models of different sexes to control for varying levels of accommodation to different sexes.

4.3. Sex Differences

The current results do not support our second research question investigating the effect of speaker's sex on accommodation. We did not find an overall main effect of speaker sex. The accommodation literature is split in regards to the effect of speaker sex. Some studies have reported effects of sex (Babel et al., 2014; Bilous & Krauss, 1988; Goldinger & Azuma, 2004; Miller, Sanchez, & Rosenblum, 2010; Namy et al., 2002; Pardo, 2006). Other studies have shown no effects of sex (Pardo et al., 2013, 2017).

More broadly, some research reports findings of sex effects on the perception of stuttering (Burley & Rinaldi, 1986) and other speech disorders (Williams & Dietrich, 2001). However, the evidence seems to show no difference in the perception of speech disorders (St. Louis, 2012; Valente, St. Louis, Leahy, Hall, & Jesus, 2017) including dysarthria (Walshe, Miller, Leahy, & Murray, 2008) and simulated speech disorders (Allard & Williams, 2008). It appears that the perception of speech disorders is highly similar between the two sexes (Patterson & Pring, 1991). This is in line with the results of the current study.

4.4. Effects of SLP Training

Our third research hypothesis posited that SLP training may predispose individuals to accommodate more. Our current findings fail to show a main effect of SLP training. This may be due to the limited professional experience of the SLP students we recruited for the study. Since the students had not been exposed to a large number of patients, they may not have developed the acute listening skills or the increased empathy towards patients as described by McNamara (2014). It appears that SLP students who have not fully completed their training are not particularly likely to overaccommodate. Future studies should include experienced SLPs to examine the effect of training on accommodation to disordered speech.

4.5. General Discussion

Social factors may play an important role in the perception of hypernasality (Watterson et al., 2013). When explored in the context of accommodation as per CAT, the current results provide evidence that the hypernasality in a model talker leads to divergence in the speakers. The evidence suggests that the high level of hypernasality produced by the model talkers may have led the interlocutors to distance themselves from the model talkers.

Nonetheless, we also see evidence of automatic processing in the interlocutors. The current results show evidence that speakers converge with typical levels of nasality. Speakers were not instructed to converge or diverge. Nonetheless, both patterns appear to emerge in the current data.

The current data could be taken to support an integrated model of automatic processing with some top down social influence. However, due to the limited nature of the research task, it remains unclear to what degree social factors and other higher level processing may have influenced the accommodation reaction (Babel, 2010; Pardo et al., 2017).

4.6. Limitations

It is important to note that the present study had a number of limitations. The first was the use of modelled hypernasal speech, as opposed to naturally occurring hypernasal speech. However, this approach appeared preferable over synthesizing stimuli electronically. The voice actors were able to produce nasality levels within the range typical to severely hypernasal speech. Additionally, training model speakers allowed us to better control the level of hypernasality and remove any potentially co-occurring disorders (Kummer, 2008). Additionally, the present study was limited to only five sentences, as per the experimental methods of Borrie & Liss (2014). By limiting the speech materials, we were better able to determine the possible effects of hypernasality. Future studies are needed to investigate accommodation to hypernasality in a more ecologically valid conversational setting. With the current setup using a microphone and nasometer headset, the environment was not a typical conversation situation. Future studies will need to take this into consideration when using the nasometer to study nasalance in a conversational setting. Nonetheless, the current study provides additional evidence to motivate future studies about phonetic accommodation to disordered speech.

5. Conclusions

Consistent with previous studies (Zellou et al., 2017; Zellou et al., 2016) individuals converged to typical nasality in model speakers. Moreover, we extend previous findings of accommodation to disordered speech, by showing evidence of strong divergence to hypernasal speech. As for our second and third hypotheses, we did not find a significant effect of speaker sex or SLP student training on accommodation.

This is the first study to use the nasometer in a quasi-conversational task to study speech accommodation. The results present the groundwork to facilitate further research on the topic of accommodation to disordered speech. Future work may explore other features of disordered speech including rate of speech, prosody, or a combinations of these features with hypernasal speech. The study of individual features will further our understanding of the influence of specific features of disordered speech on an interlocutor's speech.

6. References

Abel, J. (2013). The effect of task difficulty on phonetic convergence. *The Journal of the Acoustical Society of America*, *134*(5), 4201–4201.

Abel, J., & Babel, M. (2017). Cognitive load reduces perceived Linguistic convergence between dyads. *Language and Speech*, *60*(3), 479–502.

Abur, D., Lester-Smith, R. A., Daliri, A., Lupiani, A. A., Guenther, F. H., & Stepp, C. E. (2018). Sensorimotor adaptation of voice fundamental frequency in Parkinson's disease. *PloS one*, *13*(1), 1-21.

Allard, E. R., & Williams, D. F. (2008). Listeners' perceptions of speech and language disorders. *Journal of Communication Disorders*, *41*(2), 108–123.

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, *39*(4), 437–456.

Babel, M., & Bulatov, D. (2012). The role of fundamental frequency in phonetic accommodation. *Language and Speech*, *55*(2), 231–248.

Babel, M., McGuire, G., Walters, S., & Nicholls, A. (2014). Novelty and social preference in phonetic accommodation. *Laboratory Phonology*, *5*(1), 123–150.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*(6), 925-945.

Bilous, F. R., & Krauss, R. M. (1988). Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads. *Language and Communication*, *8*(3-4), 183–194.

Blood, G. W., & Hyman, M. (1977). Children's perception of nasal resonance. *Journal of Speech and Hearing Disorders*, *42*(3), 446–448.

Borrie, S. A., & Liss, J. M. (2014). Rhythm as a coordinating device: Entrainment with disordered speech. *Journal of Speech, Language, and Hearing Research*, *57*(3), 815–824.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, Ethnicity, & Intergroup Relations* (pp. 119-135). London: Academic Press.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482-1493.

Brunnegård, K., Lohmander, A., & van Doorn, J. (2009). Untrained listeners' ratings of speech disorders in a group with cleft palate a comparison with speech and language pathologists ratings. *International Journal of Language and Communication Disorders*, *44*(5), 656–674.

Burley, P. M., & Rinaldi, W. (1986). Effects of sex of listener and of stutterer on ratings of stuttering speakers. *Journal of Fluency Disorders*, *11*(4), 329–333.

Caporael, L R. (1981). The paralanguage of caregiving: Baby talk to the institutionalized aged. *Journal of personality and social psychology*, *40*(5), 876–884.

Caporael, L. R., & Culbertson, G. H. (1986). Verbal response modes of baby talk and other speech at institutions for the aged. *Language and Communication*, *6*(1-2), 99–112.

Caporael, L. R., Lukaszewski, M. P., & Culbertson, G. H. (1983). Secondary baby talk: Judgments by institutionalized elderly and their caregivers. *Journal of Personality and Social Psychology*, *44*(4), 746–754.

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustic Society of America*, *102*(4), 2360–2370.

Cohen, G., & Faulkner, D. (1986). Does "elderspeak" work? The effect of intonation and stress on comprehension and recall of spoken discourse in old age. *Language and Communication*, *6*, 91–98.

de Boer, G., & Bressmann, T. (2017). Influence of Altered Auditory Feedback on Oral–Nasal Balance in Speech. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3135–3143.

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, *33*, 1–40.

Fletcher, S. G., & Daly, D. A. (1976). Nasalance in Utterances of hearing-impaired speakers, *Journal of communication disorders*, *9*(1), 63–73.

Fujimura, O., & Lindqvist, J. (1971). Sweep-Tone Measurements of Vocal-Tract Characteristics. *The Journal of the Acoustical Society of America*, *49*(2B), 541–558.

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, *53*(3), 181-215.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics, 15*, 87–105.

Giles, H., & Coupland, N. (1991). *Language: Contexts and consequences*. Thomson Brooks/Cole Publishing Co.

Coupland, N., Coupland, J., & Giles, H. (1991). *Language, society and the elderly: Discourse, identity and ageing*. Basil Blackwell.

Goldinger, S D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, *105*(2), 251–279.

Goldinger, S D., & Azuma, T. (2004). Episodic memory reflected in printed word naming. *Psychonomic Bulletin and Review*, *11*(4), 716–722.

Gregory, S. W., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of personality and social psychology*, *70*(6), 1231–1240.

Hixon, T. J., Weismer, G., & Hoit, J. D. (2014). *Preclinical Speech Science: anatomy, physiology, acoustics, and perception* (2nd ed.). San Diego.

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

Horii, Y. (1983). An accelerometric measure as a physical correlate of perceived hypernasality in speech. *Journal of Speech and Hearing Disorders*, *26*(3), 476–480.

Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America*, *108*(3), 1246-1251.

Kuehn, D. P., & Moller, K. T. (2000). Speech and Language Issues in the Cleft Palate Population: The State of the Art. *The Cleft Palate-Craniofacial Journal*, *37*(4), 348–348.

Kummer, A. W. (2008). *Clef palate and craniofacial anomalies: the effects on speech and resonance* (2nd ed.). Clifton Park: Delmar Cengage Learning.

Laczi, E., Sussman, J. E., Stathopoulos, E. T., & Huber, J. (2005). Perceptual evaluation of hypernasality compared to HONC measures: The role of experience. *Cleft Palate-Craniofacial Journal*, *42*(2), 202–211.

Lass, N. J., Ruscello, D. M., Pannbacker, M., Schmitt, J. F., Middleton, G. F., & Schweppenheiser, K. (1995). The perceptions of stutterers by people who stutter. *Folia Phoniatrica et Logopaedica*, *47*(5), 247-251.

Lewis, K. E., Watterson, T. L., & Houghton, S. M. (2003). The influence of listener experience and academic training on ratings of nasality. *Journal of Communication Disorders*, *36*(1), 49–58.

Lo, E. S.C., Wong, A. W.K., Tse, A. C.-Y., Ma, E. P.M., Whitehill, T. L., & Masters, R. (2019). Effects of error experience on learning to lower speech nasalance level. *American Journal of Speech-Language Pathology*, *28*(2), 1–8.

McHenry, M. A., & Liss, J. M. (2006). The impact of stimulated vocal loudness on nasalance in dysarthria. *Journal of Medical Speech-Language Pathology*, *14*(3), 197–205.

McNamara, K. M. (2014). Introduction to clinical methods in communication disorders. In R. Paul (Ed.), Introduction to clinical methods in communication disorders (3rd ed., pp. 203–235). Baltimore, MD.

Miller, R. M., Sanchez, K., & Rosenblum, L. D. (2010). Alignment to visual speech information. *Attention, Perception, & Psychophysics*, *72*(6), 1614–1625.

Munhall, K. G., MacDonald, E. N., Byrne, S. K., & Johnsrude, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *Journal of the Acoustical Society of America*, *125*(1), 384-390.

Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender Differences in Vocal Accommodation: The role of perception. *Journal of Language and Social Psychology*, *21*(4), 422–432.

Patterson, J., & Pring, T. (1991). Listeners attitudes to stuttering speakers. No evidence for a gender difference. *Journal of Fluency Disorders*, *16*(4), 201–205.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. The *Journal of the Acoustical Society of America*, *119*(4), 2382–2393.

Pardo, J. S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (2013). Phonetic convergence in Shadowed Speech: The Relation Between Acoustic and Perceptual Measures. *Journal of Memory & Language*, *69*(3), 183–195.

Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics*, *79*(2), 637–659.

Peterson-Falzone, S. J., Hardin-Jones, M. A., & Karnell, M. P. (2001). Cleft lip and palate (3rd ed.). St. Louis: Mosby.

Pickering, M J, & Garrod, S. (2004). The interactive-alignment model: Developments and refinements. *Behavioral and Brain Sciences*, *27*(2), 212–225.

Pickering, Martin J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences*, *36*(4), 329–47.

Printz, W. (1990). A common coding approach to perception and action. In Relationships between perception and action (pp. 167–201). Berlin, Heidelberg: Springer.

Prinz, W. (1997). Perception and Action planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154.

Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication accommodation theory. The New Handbook of Language and Social Psychology, 33–56.

Späth, M., Aichert, I., Ceballos-Baumann, A. O., Wagner-Sonntag, E., Miller, N., & Ziegler, W. (2016). Entraining with another person's speech rhythm: Evidence from healthy speakers and individuals with Parkinson's disease. *Clinical linguistics & phonetics*, *30*(1), 68-85.

St. Louis, K. O. (2012). Male versus female attitudes toward stuttering. *Journal of Communication Disorders*, *45*, 246–253.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and cognitive processes*, *26*, 952–981.

Valente, A. R. S., St. Louis, K. O., Leahy, M., Hall, A., & Jesus, L. M. T. (2017). A country-wide probability sample of public attitudes toward stuttering in Portugal. *Journal of Fluency Disorders*, *52*, 37–52.

Walshe, M., Miller, N., Leahy, M., & Murray, A. (2008). Intelligibility of dysarthric speech: Perceptions of speakers and listeners. *International Journal of Language and Communication Disorders*, *43,* 633–648.

Watterson, T., Mancini, M. C., Brancamp, T. U., & Lewis, K. E. (2013). Relationship between the perception of hypernasality and social judgments in school-aged children. *The Cleft Palate-Craniofacial Journal*, *50*, 498–502.

Whitehill, T. L., & Lee, A. S. Y. (2008). Instrumental analysis of resonance in speech impairment. *The Handbook of Clinical Linguistics*, (pp. 332–343).

Williams, D. F., & Dietrich, S. (1996). Effects of speech and language disorders on raters' perceptions. *Journal of Communication Disorders, 29*(1), 1–12.

Williams, D. F., & Dietrich, S. (2001). Perceptions of communicative disorders. Verification and specification of rater variables. *Journal of Communication Disorders*, *34*, 355–366.

Wong, A. W. K., Tse, A. C. Y., Ma, E. P. M., Whitehill, T. L., & Masters, R. S. W. (2013). Effects of error experience when learning to simulate hypernasality. *Journal of Speech, Language, and Hearing Research, 56*, 1764–1773.

Zajac, D. J., & Vallino, L. D. (2017). Evaluation and management of cleft lip and palate: A development perspective (Vol. 1). San Diego: Plural Publishing Inc.

Zellou, G., Dahan, D., & Embick, D. (2017). Imitation of coarticulatory vowel nasality across words and time. *Language, Cognition and Neuroscience*, 32(6), 776-791.

Zellou, G., Scarborough, R., & Nielsen, K. (2016). Phonetic imitation of coarticulatory vowel nasalization. *The Journal of the Acoustical Society of America*, *140*(5), 3560–3575.

7. Appendices

7.1. Appendix A: Speech Stimuli Sentences

Sentences used in the study by Borrie & Liss (2014).

Sentences:
The supermarket chain shut down because of poor management.
Much more money must be donated to make this department succeed.
In this famous coffee shop they serve the best doughnuts in town.
The chairman decided to pave over the shopping center garden.
The standards committee met this afternoon in an open meeting.

7.2. Appendix B: Summary Statistics of Nasalance DID scores

Speaker Group	Condition	N	Mean	SD
Female non-SLP	Hypernasal	10	-8.23	5.02
Female non-SLP	Control	10	1.69	3.43
Female SLP	Hypernasal	9	-9.05	2.04
Female SLP	Control	9	4.43	2.89
Male non-SLP	Hypernasal	11	-9.87	2.14
Male non-SLP	Control	11	1.78	4.07

Table 3: Summary statistics of nasalance DID scores grouped by speaker group and condition.

Speaker Group	Condition	Model Sex	N	Mean	SD
Female Non-SLP	Hypernasal	Female	10	-8.70	4.99
Female Non-SLP	Hypernasal	Male	10	-7.68	5.10
Female Non-SLP	Control	Female	10	2.26	4.15
Female Non-SLP	Control	Male	10	1.11	2.91
Female SLP	Hypernasal	Female	9	-9.35	2.22
Female SLP	Hypernasal	Male	9	-8.80	2.05
Female SLP	Control	Female	9	5.50	2.92
Female SLP	Control	Male	9	3.41	2.95
Male Non-SLP	Hypernasal	Female	11	-9.95	2.22
Male Non-SLP	Hypernasal	Male	11	-9.78	2.26
Male Non-SLP	Control	Female	11	3.03	4.09
Male Non-SLP	Control	Male	11	0.54	4.12

Table 4: Summary statistics of nasalance DID scores grouped by speaker group, condition, and model sex.