Complementary Approaches for Improving Nasality Assessment in Cleft Palate Patients: Exploring Nasal Signal-Based Acoustic Feature and Visual Feature Analysis in Stereo Signals

Fatemeh Abnavi

A thesis submitted to the University of Ottawa in partial fulfillment for the degree of Doctor of Philosophy in Rehabilitation Sciences

School of Rehabilitation Sciences
Faculty of Health Sciences
University of Ottawa

Dissertation Abstract

Background. Cleft Palate (CP) is a common congenital condition affecting approximately 500 infants born with orofacial clefts yearly in Canada (Public Health Infobase, 2017). It disrupts the velopharyngeal sphincter, leading to oral-nasal balance disorders such as hypernasality and hyponasality. Accurate detection and classification of oral–nasal balance disorders are essential for optimizing treatment outcomes. Although auditory–perceptual judgments remain the primary method for evaluating nasality, their inherent subjectivity limits reliability. Integrating complementary measures with auditory assessments can enhance diagnostic accuracy and support better treatment planning and follow-up.

Objective. This study aimed to improve the accuracy of auditory-perceptual assessments by incorporating acoustic and visual analyses as complementary tools. The research consisted of three phases. The first study focused on exploring effective acoustic parameters extracted solely from nasal signals to develop a multiparametric approach for automatically classifying simulated oral-nasal balance disorders using linear discriminant analysis (LDA). The second study applied the developed multiparametric acoustic diagnostic algorithm to a retrospective dataset of children with cleft palate to evaluate its clinical applicability. The third study conducted a pilot investigation into the use of visual pattern inspection in stereo signals for classifying oral-nasal balance disorders.

Methods. All three studies used a retrospective, observational design based on secondary analysis of pre-existing speech data. The first and third studies employed simulated voice samples, while the second study used both simulated and clinical samples from children with cleft palate.

Results. In Study 1, LDA showed 90.9% accuracy in classifying simulated conditions based on nasal channel acoustic energy. This model demonstrated potential as a cost-effective and quantitative alternative to tools like nasometry. Study 2 confirmed that nasal-signal-derived acoustic parameters could differentiate between normal, hypernasal, and hyponasal conditions but highlighted challenges in generalizing models across datasets. Study 3 achieved an overall classification accuracy of 85.15% using visual analysis of paired oscillogram images. The highest accuracy (100%) was observed for hyponasality classification, while the lowest accuracy (61%) was recorded for mixed nasality classification.

Conclusion. This thesis advances the field of nasality assessment by demonstrating the potential of acoustic analysis of the nasal signal in isolation, as well as a new visual method for detecting and classifying oral-nasal balance disorders. These approaches could be integrated with traditional auditory-perceptual evaluations to enhance diagnostic accuracy, guide treatment planning, and support follow-up care. This research contributes new practical diagnostic approaches that could become directly relevant to clinical populations such as individuals with cleft palate or neurological disorders. The present findings lay a foundation for future work to translate these research innovations into clinical practice.

Preface to This Dissertation

Doctoral student Fatemeh Abnavi was responsible for and actively involved in all components of this thesis, including the conceptualization, study design, data collection, data analysis and interpretation, and manuscript and thesis chapters writing.

All data from the first two studies were analyzed under the supervision of ethics arrangements at the University of Toronto. Approval for the secondary use of the speech dataset was obtained under Protocol #38903, titled "Re-analysis of data from previous protocol 'Nasalance and Acoustic Profile of Mixed Nasal Resonance" (original approval date: February 26, 2020; currently valid until February 25, 2026; REB Ref #27918; initial approval granted on August 17, 2012) (Appendix A). The second study in this thesis, presented in chapter 3, received ethical approval from The Hospital for Sick Children, protocol number 1000057703, and from the University of Toronto, protocol number 00035427. The third study in this thesis, presented in Chapter 4, received ethical approval from both the University of Ottawa Research Ethics Board (Protocol #: H-02-25-10552) (Appendix C) and the University of Toronto Research Ethics Board (Protocol #: 55853) (Appendix B).

Chapters 2, 3, and 4 are formatted according to the guidelines of the peer-reviewed journals to which the manuscripts are intended to be submitted. Table and figure numbers have been modified as necessary to ensure consistency and clarity throughout the thesis. Grammar and style revisions throughout this thesis were supported using Quill Bot, an AI-powered language editing tool. While Quill Bot was used to enhance sentence clarity and structure, no AI tools were used to generate original content. All ideas, analyses, and interpretations presented in this thesis are the author's own.

Acknowledgments

This doctoral journey has been one of the most transformative experiences of my life. It shaped how I think, deepened my resilience, and taught me the value of patience and perseverance. This dissertation is not only the result of academic work but also a reflection of the many people whose support made it possible.

I am deeply grateful to my supervisor, Dr. Heather Flowers at the University of Ottawa, whose strong expertise in communication sciences and research methodology has been invaluable. Her steady encouragement, insightful feedback, and thoughtful guidance have supported me through every stage of my doctoral studies. Her mentorship helped me clarify my ideas, strengthen my work, and stay focused on my goals. This project would not have been possible without her support.

I am also sincerely thankful to my co-supervisor, Dr. Tim Bressmann at the University of Toronto, whose guidance played a pivotal role in shaping this dissertation. His clinical and research expertise in speech disorders of patients with cleft lip and palate, as well as acoustic analysis of speech, brought clarity to crucial stages of my research. I especially appreciate his steady mentorship, insightful questions, and encouragement to think critically and independently. His consistent support, availability, and genuine interest in my progress made a lasting impact—I could not have asked for a more thoughtful and generous co-supervisor.

To my committee members, Dr. Hilmi Dajani and Dr. Suzy Ahn, both from the University of Ottawa, I would like to express my gratitude for your valuable input and support. Your interdisciplinary perspectives broadened the scope of this work and encouraged me to view my research through new lenses.

On a personal note, I owe immense gratitude to my family, without whom this would not have been possible. To my parents, who instilled in me the values of education, determination, and integrity, your love and sacrifices have shaped who I am today. A special thank you to my sister, whose unwavering support opened the door for me to pursue my studies in Canada. You not only helped me begin this journey, but you also walked beside me through every milestone. Your generosity, belief in me, and encouragement are gifts I will always treasure. Moreover, to my brothers, who cheered me on from afar—thank you for reminding me that I was never alone.
This research was made possible by the generous availability of recorded speech datasets used across all three studies. I sincerely thank the researchers and participants whose contributions formed the foundation for this work, enabling a more profound exploration in the field of cleft palate research. This thesis is the product of dedication, community, and love. It stands not only as an academic achievement but also as a testament to the extraordinary people who helped make it possible.

Funding

I am grateful to the funding sources that awarded me for my PhD project: the University of Ottawa Admission Scholarship (2021-2024) and the International Doctoral Scholarship (2020-2025).

Table of Contents

Chapter 2

Automatic Detection and Classification of Nasality Using Acoustic Analysis of Only the Nasal Speech Signal: a Pilot Study

Chapter 3

Clinical Application of Nasal Signal-based Acoustic Approach to Enhance Detection and Classification of Oral-Nasal Balance Disorders Related to Cleft Palate: A Retrospective/Secondary Analysis

Abbreviations

Abbreviation	Full Term
Aps	Acoustic Parameters
BWT	Bionic Wavelet Transform
CAPS-A	Cleft Audit Protocol for Speech–Augmented
CP	Cleft Palate
DME	Direct Magnitude Estimation
EAI	Equal Appearing Interval
F1	First Formant
F2	Second Formant
F3	Third Formant
GDAM	Group Delay Function-based Acoustic Measure
HAR	Harmonic Amplitude Ratio
HPF	High-Frequency Power
HSV	High-Speed Videoendoscopy
LDA	Linear Discriminant Analysis
LP	Linear Predictive
LPF	Low-Frequency Power
MFCC	Mel-Frequency Cepstral Coefficients
MDT	Maximum Duration Time
NAE	Nasal Air Emission
NAEM	Nasal Acceleration Energy Measure
NHA	Normalized Harmonic Amplitude
NSI	Nasality Severity Index
NORAM	Nasality Oral Ratio Meter
PCA	Principal Component Analysis
P0	First Nasal Peak
PHF	Prominent Harmonics Frequency
PMS	Velopharyngeal Mislearning
SLP	Speech-Language Pathologist

SPL	Sound Pressure Level
SVM	Support Vector Machine
TEO	Teager Energy Operator
UCLP	Unilateral Cleft Lip and Palate
VNE	Videonasoendoscopy
VAS	Visual Analogue Scaling
VLHR	Voice Low Tone to High Tone Ratio
VPD	Velopharyngeal Dysfunction
VPI	Velopharyngeal Incompetence
VPM	Velopharyngeal Mislearning
VPS	Velopharyngeal Sphincter

List of Tables

List of Figures

Chapter 1

General Introduction

Context and Statement of the Problem

Cleft palate (CP) is a common congenital anomaly affecting the palate, requiring lifelong multidisciplinary care from infancy to adulthood. Cleft palate classification is based on various anatomical criteria. Anatomically, palatal clefts are categorized as complete (front-back gap) or incomplete (gap in the hard or soft palate). In terms of severity, submucous clefts (gap hidden by tissue) and open clefts (visible gap) are distinguished (Jensen et al., 1988). Cleft palate can occur independently or as part of a syndrome (Singh et al., 2015). It can be combined with uni- or bilateral cleft lip and alveolus or occur in isolation. The degree of involvement can be complete (entire width) vs. incomplete (partial) clefts (Singh et al., 2015; Acharya et al., 2021).

Primary surgical intervention is typically conducted within the first year of age, and its success during the initial months after birth determines the outcomes for eating, speaking, hearing, dental development, and facial growth. The main aims of cleft palate repair are to restore the palate's structure and function and establish a functional velopharyngeal mechanism for normal feeding and speech development. Secondary surgeries, carried out as the child grows, address unresolved issues from the primary surgery. These procedures can target persistent speech difficulties after therapy, improve lip and nose appearance, and dental and orthodontic concerns (Nguyen et al., 2015). The cleft palate connects the oral and nasal passages, causing the nasal cavity to serve as an additional resonator. This structural change alters sound transmission through the vocal tract filter, resulting in abnormal speech production and leading to acoustic changes such as shifts in formant frequency peaks, bandwidth, and amplitude. Listeners perceive these acoustic changes as excess nasality, as variations in formant frequencies and other acoustic cues significantly contribute to the perception of changing speech quality. While acoustic analysis can provide data on the physical properties of nasality, auditory perception offers subjective information on how these acoustic properties directly influence the listener's perception of nasality through the auditory system (Kummer, 2011). An accurate evaluation of nasality is crucial for the therapeutic team to monitor speech development and determine appropriate interventions such as surgery, prosthodontics, or speech therapy (Kummer & Lee, 1996).

Speech Production

The classical source–filter theory, introduced by Fant 1960, is a significant model in speech production. It divides the process into two primary elements: the source and the filter. The source refers to the vibrating sound created by the vocal folds, while the filter represents the vocal tract's resonance properties that modify and shape the sound. Speech sounds are generated because of the combined functioning of the source and filter. The source's voice signal travels through the vocal tract, which serves as a filter, before exiting through the mouth or nose (Figure 1.1). The glottal wave produced by the source (vocal folds) is a complex waveform with several component frequencies (the fundamental and all the harmonics). The source signal is filtered by the vocal tract, amplifying specific harmonics based on the shape and size of the vocal tract. These resonant frequencies, known as formants, play a crucial role in the perception of speech. The filtering mechanism amplifies frequency components related to vocal tract resonances while decreasing others; this filtered sound results in distinct formant patterns for different speech sounds. In addition to the resonant filtering provided by the vocal tract cavities, the lips play an important role in speech resonance by modulating formant frequencies, such as lowering formant frequencies through lip rounding. Because this radiation effect occurs due to sound energy transmitted from a relatively small opening at the mouth, radiation efficiency becomes frequency dependent. Consequently, the final acoustic output reflects the combined effects of the glottal source, vocal tract resonances, and lip radiation characteristics (Fant, 1960; Titze, 2008).

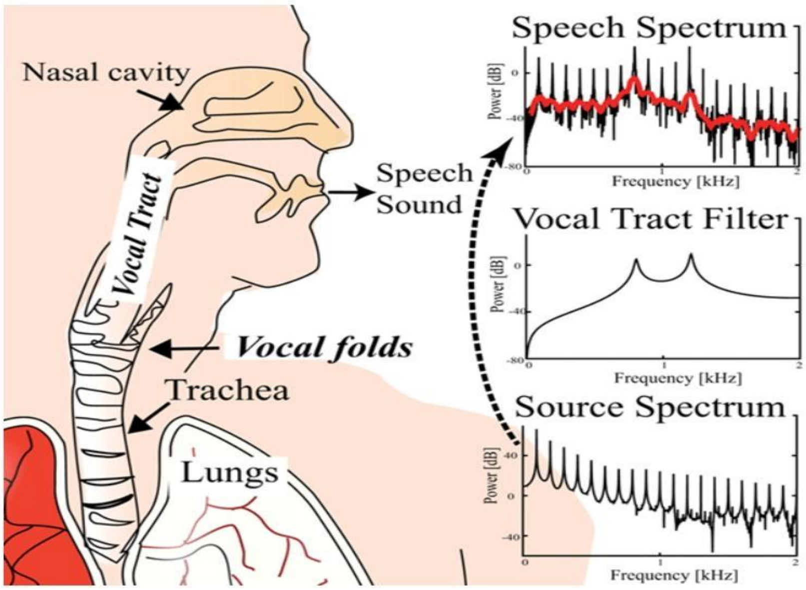


Figure 1.1. Schematic Representation of the Source–Filter Theory of Speech Production. Reproduced from Tokuda (2021) with permission.

The overall resonance of speech is influenced by the size and shape of the vocal tract cavities and by the movement of the velopharyngeal sphincter (VPS), which regulates coupling and/or isolation of the oral and nasal cavities relative to the source. The VPS closure directs the source signal into the oral cavity, effectively diverting acoustic energy away from the nasal cavity and thereby shaping oral-nasal airflow patterns and sound distribution. In velopharyngeal dysfunction (VPD) or other vocal tract obstructions, the normal direction of acoustic energy and the control of airflow can both be disrupted. Misdirected acoustic energy contributes to nasality in speech, whereas misdirected airflow results in audible nasal emission and turbulence. At a fundamental level, the classical source–filter model (Fant, 1960) remains a practical framework:

the vocal folds serve as the vibrating source, while the vocal tract acts as the filter that shapes the resulting speech sounds. However, the model is inherently a linear approximation that does not account for these nonlinear phenomena, particularly turbulence and emission arising from airflow misdirection or nasality arising from acoustic energy misdirection in VPD. Research on nonlinear source–filter coupling has demonstrated that the source and filter can interact dynamically, rather than behaving as fully independent components (Titze, 2008). Consequently, in conditions like VPD, the classical source–filter model may not fully capture the resulting airflow and acoustic behavior, and its predictions should be interpreted with caution.

Velopharyngeal Sphincter (VPS)

The velopharyngeal sphincter is a muscle valve that links the back of the hard palate to the back of the pharynx. It is also known as the velopharyngeal mechanism. This structure is made up of the velum (soft palate), lateral pharyngeal walls, and posterior pharyngeal walls (Figure 1.2). The velopharyngeal mechanism controls the opening and closure of the velopharyngeal port during speech production to direct airflow and sound into the proper cavity, whether oral or nasal. The velopharyngeal sphincter opens to variable degrees for nasal and nasalized speech sounds, whereas it closes completely for oral sounds (Marsh, 2004; Perry, 2011).

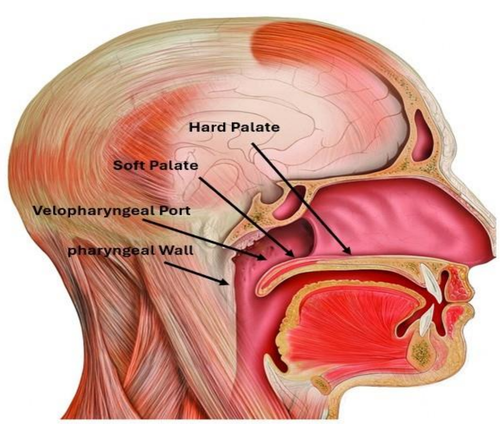


Figure 1.2. Sagittal view of the anatomy of the velopharyngeal sphincter (adapted from image by Patrick J. Lynch, medical illustrator [CC BY 2.5, http://creativecommons.org/licenses/by/2.5], via Wikimedia Commons)

From a functional standpoint, the velopharyngeal mechanism is closed by velum contraction and elevation, lateral pharyngeal wall movement toward the midline, and sometimes, posterior pharyngeal wall anterior movement. As a result, instead of passing through the nasal passage, sound and airflow are directed toward the mouth (Yanagisawa et al., 1990; Marsh, 2004; Dudas et al., 2006). Thus, the role of the VPS in separating and connecting the nasal and oral passages is an important physiological process for speech production. Oral speech sounds will only be formed with enough energy and precision by the pharynx and the oral cavity shape if the VPS is closed. Otherwise, the sound may also enter the nasal passage in part or in full if VPS closure is compromised (Kummer, 2002).

Velopharyngeal Dysfunction

Typical velopharyngeal function is based on normal anatomy, neurological function, and the process of speech learning. An issue in any of these systems can result in velopharyngeal dysfunction, which affects speech and other activities, such as blowing, swallowing, and singing (Kummer, 2009; Brunnegård et al., 2008; Rintala & Haapanen, 1995). Velopharyngeal dysfunction (VPD) is a comprehensive term encompassing instances where the VPS, responsible for separating the oral and nasal passages, experiences inconsistency or incomplete closure during the articulation of oral sounds (Figure 1.3). VPD has three main causes (Trost-Cardamone, 1989; Kummer & Lee, 1996; Kummer, 2009):

- VP Insufficiency (VPI): This type of VPD typically refers to anatomical or structural defects, such as a short velum, which is common after cleft palate repair.
- Velopharyngeal Incompetence (VPI): This type of VPD is caused by abnormal movement of velopharyngeal structures and is associated with neuromotor causes.

- Velopharyngeal mislearning (VPM): This type of VPD is caused by a failure to develop proper articulation patterns.

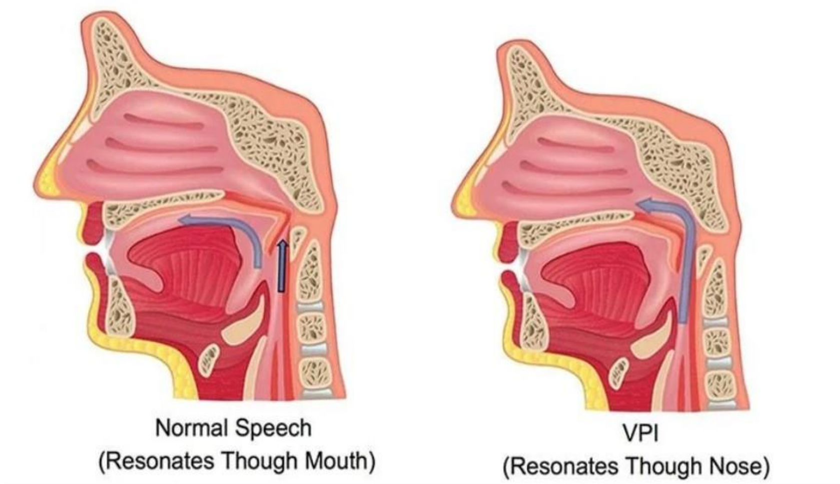


Figure 1.3. Schematic view of velopharyngeal dysfunction (VPD). Reproduced from Sanu, P. M. (2018) under the Creative Commons Attribution (CC BY 4.0) license.

The type and etiology of VPD determine the treatment process. VP insufficiency or incompetence generally necessitates surgical or prosthetic intervention. The most challenging aspect of VPI surgical care is tailoring the treatment procedure to the closure pattern and gap size of the VPS (Table 1.1). As a result, it is crucial to determine the type and degree of velopharyngeal dysfunction based on an accurate assessment (Kummer, 2009; Kummer, 2018). Cleft palate surgery aims to close the gap in the palate and restore its function. However, residual issues like oral-nasal imbalances, compensatory articulation, nasal airflow problems, altered voice quality, and nasal or facial grimaces can persist. Following surgery, speech therapy is usually still necessary to teach patients how to use their reconstructed VPS for accurate sound production (Nagarajan et al., 2009). Hence, a common approach to treating velopharyngeal dysfunction involves a combination of surgical intervention and speech therapy. However, if the speech problem stems from VPM, speech therapy alone is typically effective in addressing the speech difficulty. VPM occurs when a child learns incorrect ways to produce sounds. Some children develop unusual speech patterns that do not engage the VPS, even when it functions properly. Others adopt atypical speech patterns to compensate for structural issues in the mouth or throat, learning to produce sounds without the proper movement of the VPS. In such cases, speech therapy should focus on articulatory placement and oral airflow to facilitate the correct production of sounds. (de Stadler & Hersh, 2015; Kummer, 2009). Due to the diverse causes, severity, and impacts of VPD, successful therapy demands a comprehensive evaluation by the therapeutic team to formulate tailored treatment strategies and ensure the highest quality of care.

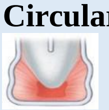
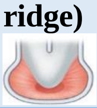
VP port closure pattern	VP gap size	Preferred surgical treatment
Sagittal 	Small and intermediate Large	Overlapping IVV with oral Z-plasty Superior-based pharyngeal flap
Coronal 	Small and intermediate Large	Sphincter pharyngoplasty Combined overlapping IVV with oral Z-plasty and sphincter pharyngoplasty
Circular (without  **Or with Passavant's ridge)** 	Small Intermediate or large	Overlapping IVV with oral Z-plasty Superior-based pharyngeal flap

Table1.1. Closure Pattern and Gap Size of Velopharyngeal Sphincter (Croft et al., 1981; Nam, 2018). Images sourced from Sweeney et al. (2015).
IVV = intravelar veloplasty (a surgical technique that reorients and reconstructs the levator veli palatini muscle to restore velopharyngeal function); Z-plasty = a plastic surgery technique in which Z-shaped incisions are used to lengthen or reorient tissue; sphincter pharyngoplasty = a procedure that creates a dynamic sphincter at the level of the velopharynx using flaps from the posterior faucial pillars; superior-based pharyngeal flap = a flap procedure where tissue from the posterior pharyngeal wall is attached to the soft palate to reduce the velopharyngeal gap (Nguyen et al., 2015).

Oral-Nasal Balance

Resonance is a physical phenomenon determined by the shape, size, and configuration of the vocal tract. The sound produced by the vocal folds functions as the input to this resonant system, which selectively amplifies specific frequencies based on its physical characteristics. In speech science, the term resonance refers to a physical phenomenon by which the vocal tract modifies the sound produced by the vocal folds, amplifying specific frequencies while dampening others. It involves the selective enhancement or suppression of specific harmonics of the sound wave as it travels through the oral and nasal cavities (Titze, 2008; Fant, 1970). The concept of resonance refers to how the vocal tract responds to the initial sound signal from the vocal folds. In contrast, speech-language pathologists primarily use the term "resonance disorder" to describe imbalances in oral and nasal sound transmission during speech (de Boer & Bressmann, 2015). According to de Boer and Bressmann (2015), the term "oral-nasal balance" better expresses the coupling or separation of sound transmission between the oral and nasal cavities. Oral-nasal balance is an acoustic feature of speech production that is determined by the coupling or separation of sound between the oral and nasal cavities (Santoni et al., 2020). When the velopharyngeal sphincter fails to function correctly, it causes abnormal sound transmission through the vocal tract filter, which is particularly common in children with cleft lip and/or palate (Yun & de Chalain, 2008).

Any factor that disrupts the normal oral and nasal balance during speech production can result in the following oral-nasal balance disorders (de Boer & Bressmann, 2015):

- Hypernasality: occurs when there is too much sound energy in the nasal cavity during speech due to velopharyngeal insufficiency/incompetence or an oronasal fistula (Kummer, 2018).
- Hyponasality: occurs when there is not enough nasal energy in the nasal cavity during the production of nasal sounds due to nasal cavity obstruction (deviated septum, enlarged

adenoids, nasal congestion, stenotic nares, nasal polyps, or maxillary retrusion, which restricts pharyngeal cavity space) (Kummer, 2018).

- Mixed nasality: Occurs when there is hypernasality on oral consonants and hyponasality on nasal consonants due to any nasopharyngeal blockage or velopharyngeal dysfunction (Kummer, 2018).

Other important features of vocal tract resonance that require assessment through auditory–perceptual judgment include cul-de-sac resonance, nasal turbulence, and nasal emission.

- Cul-de-sac Resonance: Occurs when a sound vibrates in a cavity (pharyngeal, oral, nasal) but cannot escape due to obstruction. Because the soft tissues partially absorb sound, the result is muffled, low-intensity sound. The placement of the obstruction dictates whether the cul-de-sac resonance is oral, nasal, or pharyngeal (Kummer, 2018; de Boer & Bressmann, 2015).
- Nasal Air Emission (NAE): NAE occurs when air escapes through the velopharyngeal valve due to a gap in the VP port or an oronasal fistula, resulting in extra sound sources with or without hypernasality. Nasal emission might be weak or even inaudible in a large VP gap due to low flow resistance (Cler et al., 2016).
- Nasal Rustle: A small VP opening frequently results in an inconsistent nasal rustling, also known as nasal turbulence. This turbulence often results in noisy or distorted speech sounds, which can impact the intelligibility of speech. Nasal turbulence and nasal emission are highly correlated with weak-pressure consonants because they can affect the production of high-pressure consonants (Zajac & Preisser, 2016; Kummer, 2018).

These VPD-related speech features can occur alone or in combination with other features, so the evaluator must analyze each of them and be able to detect them in the individual's speech tasks. Accurate and functional evaluation of these speech characteristics is crucial for the therapeutic team to determine whether patients require surgery, prosthodontics, and/or speech therapy, as well as to monitor speech development (Kummer & Lee, 1996).

Acoustic Features of Nasality

Typically produced nasal sounds can be detected by the presence of a nasal murmur. The nasal murmur’s acoustic energy is concentrated in a low-frequency range. Nasal formants have a lower amplitude than those of vowels. Furthermore, the oral cavity functions as a side-branching resonator to the main nasopharyngeal tube, resulting in oral anti-formants that absorb energy. It has been speculated that one of the reasons for the low overall amplitude of nasal consonants is the presence of anti-formants (Pruthi & Espy-Wilson, 2004).

Figure 1.4 illustrates these acoustic characteristics by showing a waveform (top) and spectrogram (bottom) of a nasalized segment. Note the concentration of energy in the low-frequency nasal formants (dark bands) and the weakened or missing higher-frequency formants caused by anti-formants. This provides a visual contrast to the well-defined formant structure of adjacent vowel segments.

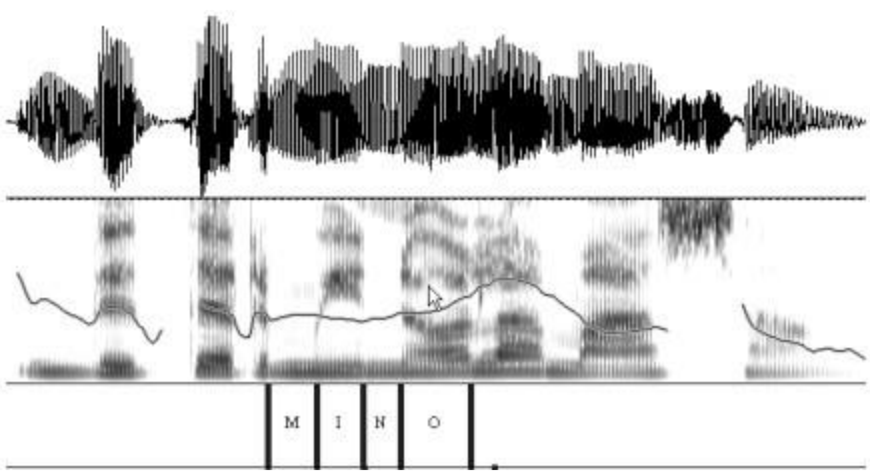


Figure 1.4. Waveform (top) and spectrogram (bottom) of a nasalized utterance. The formant tracks are plotted over the spectrogram. Reproduced from Ortega-Llebaria, M., & Prieto, P. (2008), with permission.

The nasal formant and anti-formant pairs that are added when the oral and nasal cavities are coupled have an impact on the vowel spectrum. The addition of extra-nasal formants weakens formants, reduces their amplitude, widens their bandwidth in the speech spectrum, shifts the location of the low-frequency spectral prominence, increases the amplitudes of the bands between the first formant (F1) and second formant (F2), decreases the amplitude of F2, and creates a dip between F2 and third formant (F3). The spectral characteristics of nasalization differ from one vowel to another. When low vowels are nasalized, the amplitude of F1 decreases because the first nasal zero appears in the frequency region of F1. When high vowels are nasalized, however, the first nasal zero appears in a higher frequency region than F1; therefore, the amplitude of F1 is not attenuated (Chen, 1995; Dodderi et al., 2016; Lee et al., 2003).

There are differences between phonetic and phonemic vowel nasality, as well as nasalization changes between different types of vowels. Phonetic nasalization refers to coarticulatory nasalization that occurs due to neighboring nasal consonants (e.g., in English). In contrast, phonemic nasalization is a contrastive feature (e.g., in French) where nasal vowels distinguish meaning from their oral counterparts. To quantify nasalization in vowels between nasal consonants for English speakers and nasal vowels for French speakers (which makes a linguistic distinction between oral and nasal vowels), Chen (1997) calculated the values for the acoustic correlates A1-P1 (A1 is the amplitude of the first formant and P1 is the amplitude of the extra peak between the first two formants) and A1-P0 (P0 is the amplitude of the extra peak below the first formant). The means of A1-P1 and A1-P0 for the nasalized vowels in French were smaller than for the nasal vowels in English. The fact that nasalization of English vowels was due to context, while nasalization of French vowels was due to contrast, may introduce variations in velopharyngeal opening size and help explain the smaller mean A1–P1 and A1–P0 values observed for French. This difference reflects the contrastive versus coarticulatory nature of nasalization: in English, nasalization typically involves a relatively consistent velopharyngeal opening for vowels adjacent to nasal consonants. In French, by contrast, nasal vowels are produced in a more controlled manner to keep them distinct from oral vowels, which can lead to greater variability in nasalization and, on average, a slightly smaller velopharyngeal port opening compared to the more automatic coarticulatory nasalization of English vowels. Furthermore, speaker differences or vowel environment affecting vowel breathiness may influence the values of A1-P0. These cross-linguistic differences show that A1–P1 and A1–P0 values can vary across languages, but the same acoustic analysis methods remain valid (Chen, 1997).

Velopharyngeal Dysfunction Management

Effective management of VPD involves a structured and systematic approach to evaluation and treatment, as depicted in the attached management algorithm (Figure 1.5). Accurate assessment is crucial to ensure accurate diagnoses, inform treatment strategies, and monitor the patient's progress over time (Marsh, 2004). The Assessment process begins with an auditory-perceptual evaluation performed by a skilled speech-language pathologist (SLP). This evaluation involves analyzing speech samples with increasing phonetic complexity to assess velopharyngeal function accurately (Jones, 1991; D'Antonio, 1992). If the patient demonstrates difficulty producing the oral consonants required for assessing velopharyngeal closure, speech therapy is initiated to correct articulatory errors. Once these targets are achieved, velopharyngeal function is reassessed to ensure accurate evaluation (Marsh, 2004).

For patients with intact articulation of VP closure consonants, the next step involves further auditory-perceptual evaluation. If no signs or symptoms of VPD are detected, therapy is directed toward non-VPD-related issues. However, in cases where inconsistent VPD features are present, therapy focusing on velopharyngeal function is provided. After therapy, the patient's progress is re-evaluated to determine whether the VPD has resolved or if additional assessment is required. For patients with persistent VPD features, instrumental evaluation is essential. Techniques such as nasendoscopy and videofluoroscopy are employed to obtain a detailed quantitative and qualitative analysis of VP closure and movement (Marsh, 2004). Based on instrumental findings, patients are categorized into three groups:

- VP Closes Consistently: No further VPD management is required.
- VP Closes Inconsistently: Therapy and/or management targeting the inconsistency are initiated.
- VP Never Closes: Comprehensive VPD management is undertaken to address the underlying dysfunction.

This systematic approach, as suggested by Marsh (2004), emphasizes the importance of tailoring interventions to the specific needs of the patient, as highlighted in the algorithm. By integrating dynamic diagnostic methods and individualized therapeutic strategies, this approach aims to promote developmental progress and optimize speech outcomes in children with VPD.

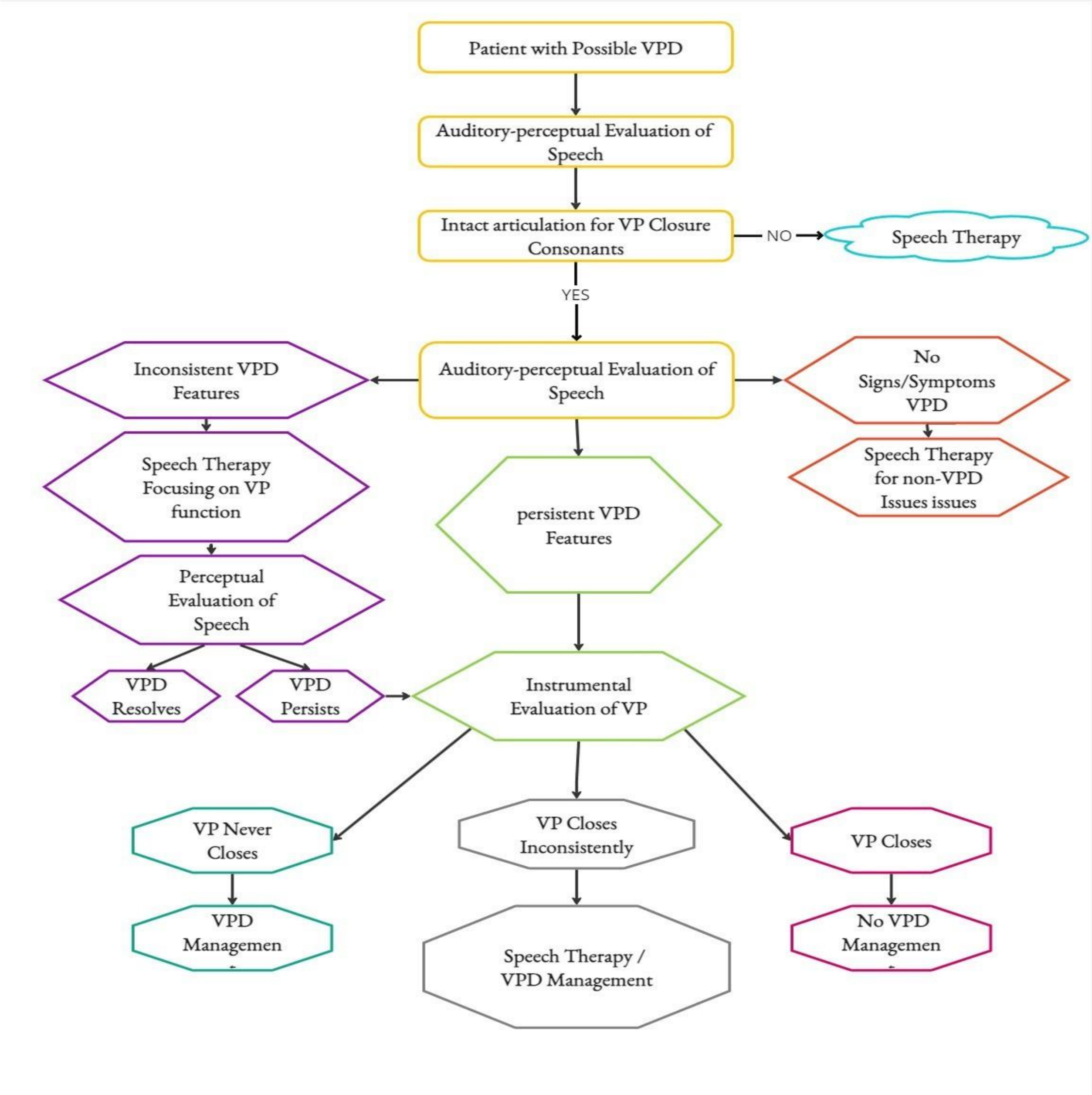


Figure. 1.5. Velopharyngeal management algorithm. Created based on Marsh (2004).

Review of the Literature

Assessment of Velopharyngeal Function and Nasality

An interdisciplinary team approach is the standard for evaluating and managing VPD in patients with cleft lip and palate (Marsh, 2009). VPD is evaluated using a comprehensive approach that includes a variety of dynamic diagnostic methods to evaluate velopharyngeal movement, structure, degree of closure, and timing, as well as the influence of VPD on airflow and sound (Shprintzen et al., 1979; Marsh & Louis, 2003; Ruscello, 2007).

Auditory-Perceptual Assessment

The foundation and standard method for clinical evaluation of nasality is auditory-perceptual testing performed by a trained clinician (Kummer & Lee, 1996). This includes listening to the patient's speech productions while monitoring for oral-nasal balance disorders, nasal rustle, nasal air emission, compensatory articulation errors, consonant strength/oral air pressure, and phonation in specific contexts (Karling et al., 1985; Kuehn & Moller, 2000).

Auditory-perceptual assessment of nasality is an important aspect of a comprehensive evaluation of speech in individuals with cleft palate. Indeed, Kuehn (1982) states that 'in a sense, a speech disorder does not exist until a listener perceives it'. Auditory-perceptual assessment not only evaluates speech nasality but also indirectly provides information about the velopharyngeal function (Kuehn & Moon, 1998; Sell et al., 2006; Kummer et al., 1992; Kummer et al., 2003). The relationship between perceived nasality and the size of the VP gap has been reported to be nonlinear because the degree of nasality is determined by a complex interaction of several factors, including different sizes of the oral, pharyngeal, and nasal cavities, respiratory effort, vocal intensity, vocal pitch, and the ratio of oral and nasal acoustic resistance (Kummer et al., 2003).

The auditory-perceptual analysis is a challenging process that is prone to errors and varies from person to person, even when performed by trained specialists. Several perceptual rating measures, specifically for VPD-related speech features, have been proposed to standardize auditory-perceptual examination among individual clinicians. These scales include the categorical scale (ordinal scale), the equal appearing interval (EAI) scale, and magnitude measures (ratio-based methods) (Castick et al., 2017; Yamashita et al., 2018).

An EAI scale uses whole numbers (between 1 and n) to divide the scale into equal intervals, whereas an ordinal scale uses terms like "normal," "mild," "moderate," and "severe," which are represented by 0, 1, 2, and 3, respectively. Ordinal scales, such as the most prevalent clinical auditory-perceptual evaluation procedures, Universal Parameters, and the Cleft Audit Protocol for Speech–Augmented (CAPS-A) tools, are widely used in cleft clinics for testing speech parameters (Castick et al., 2017; Yamashita et al., 2018).

Ratio scales, on the other hand, require listeners to assign numbers to stimuli based on their magnitude. A visual analogue scaling (VAS) is an undifferentiated horizontal line (usually 100 mm long), on which both extremes are marked, for example, normal and severe. A listener must indicate their rating on a VAS by making a mark somewhere along the line, and the score is determined by calculating the distance from the left-hand end to the mark. Similarly, direct magnitude estimation (DME) is not bound by fixed minimum/maximum values and enables a listener to scale individual speech samples relative to each other or to a standard stimulus (aka, the modulus), which is usually obtained from the middle of the range of stimuli and assigned a numerical value. According to several studies, ratio scales such as VAS or DME can be used to score nasality more accurately and consistently (Zraick & Liss, 2000). However, Castick et al. (2017) compared the reliability and suitability of using ordinal (CAPS-A) scaling and VAS ratio scaling for the auditory-perceptual judgment of hypernasality, hyponasality, nasal emission, nasal turbulence, understandability, and acceptability. They concluded that these cleft speech parameters can be measured equally well using both VAS and ordinal scaling. Despite efforts to develop standard perceptual rating scales, significant improvements in intra- and inter-rater reliability of nasality perceptual evaluation have not been achieved. (Cler et al., 2016; de Boer et al., 2020)

Another significant factor contributing to the differences in the nasality perceptual assessment's inter-rater reliability is probably listeners' various internal concepts. To minimize the influence of the listener's standards for evaluating speech signals and increase agreement and consistency among assessors, pre-assessment training and practice are found to be crucial (Santos et al., 2016; Lee et al., 2009). According to Lee et al. (2009), observers who received training had higher interobserver agreements on hypernasality than those who experienced only exposure to the samples. However, achieving a positive training effect for all target variables that are sustained over time remains a challenge (de Boer et al., 2020; John, 2006). The listeners' linguistic background can also affect the reliability of the auditory-perceptual evaluation. Individuals from various cultural groups could have varied ideas on what constitutes acceptable nasality (Yamashita et al., 2018). In Cantonese speech samples, Lee et al. (2008) confirmed that nonexpert Cantonese listeners rated hypernasality substantially higher than English listeners did, and they concluded that the listeners' linguistic background likely influenced this finding.

In a recent study, De Boer et al. (2020) aimed to improve perceptual evaluation agreement by using nasalance-based pre-classification as the foundation for auditory-perceptual categorization of oral-nasal balance disorders. Listeners were given a nasalance-based pre-classification before performing an auditory-perceptual evaluation. According to the findings, perceptual evaluation agreement was substantially higher when listeners were provided with nasalance-based pre-classification and severity. The findings suggested that to improve listener agreement, it may be helpful to use nasalance scores to pre-classify speakers' oral-nasal balance prior to the clinician's auditory-perceptual evaluation. However, further studies are needed to determine whether expert listeners will accept such pre-classifications as a supplement to their auditory-perceptual evaluation of oral-nasal balance disorders.

Despite the subjective character of auditory-perceptual evaluation, it remains a significant step in the clinical evaluation of nasality and velopharyngeal function, as well as a fundamental component of typical protocols for testing VPD. As a result, in clinical practice, instrumental measures are necessary to supplement or even guide perceptual findings, facilitating appropriate treatment decisions.

Instrumental Assessment

Over time, various methods have been employed to assess the velopharyngeal function and nasality of patients with cleft palate, both subjectively and objectively. While no instrumental approach can completely replace auditory-perceptual analysis, clinicians can use instrumental assessments in conjunction with auditory-perceptual evaluation to corroborate and enhance their findings. The correlation between clinical and instrumental evaluations is crucial in determining the actual condition of the velopharyngeal mechanism, particularly due to the limitations of auditory-perceptual evaluation alone. Instrumental evaluations can be classified into two categories: direct and indirect techniques.

Direct Assessment Techniques

Imaging technology that visualizes velopharyngeal function and structures is frequently used to quantify the extent of velopharyngeal dysfunction, select the appropriate therapeutic option, tailor surgical treatments, and monitor improvements following treatment (Sommerlad et al., 2002; Marsh, 2009; Armour et al., 2005). Common techniques include Videonasoendoscopy (VNE), Multiview Videofluoroscopy (MVF), and Dynamic Magnetic Resonance Imaging (MRI).

Videonasoendoscopy (VNE). VNE is one of the most widely used imaging procedures for assessing VPD in patients with cleft palate. This method is used to evaluate velopharyngeal movements, detect velopharyngeal gap size and location, and identify features such as palatal fistulae or submucous clefts (Pigott, 1969; Lipira et al., 2011).

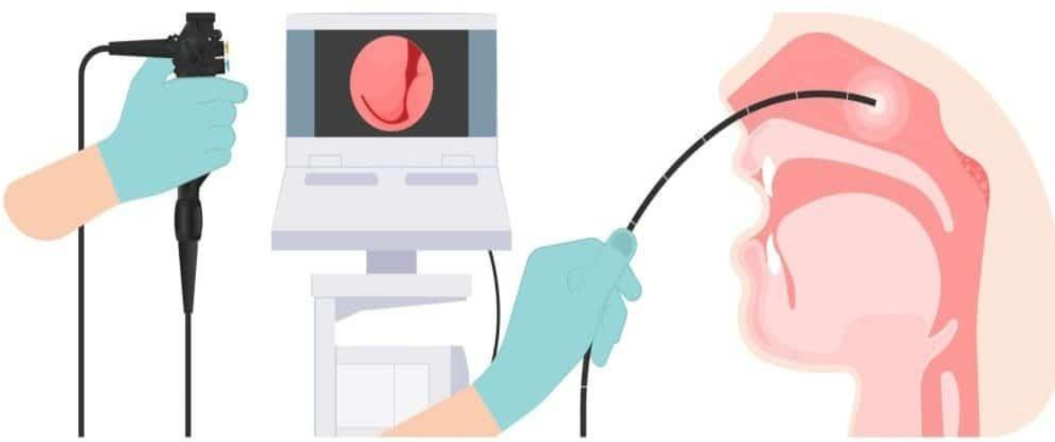


Figure 1.6. Videonasoendoscopy. Retrieved July 29, 2025, from https://teachmesurgery.com/consent/ent/consent-flexible-nasal-endoscopy/ and reproduced under the Creative Commons Attribution (CC BY 4.0) license.

VNE provides a wide range of clinical applications, including biofeedback training, diagnosis, postoperative follow-up, and prognosis (Paniagua et al., 2013).To get a birds-eye view of the nasal velar surface and velopharyngeal port at rest and during phonation, a flexible fiberoptic scope is inserted via the nose, into the nasopharynx, and positioned above the soft palate (Lam et al., 2006). Oren et al. (2020) recorded the dynamics of finer velum movements as well as gross motions using High-Speed Videoendoscopy (HSV). They stated that using HSV nasoendoscopy for research purposes has several advantages over traditional methods. HSV, in particular, enables the examiner to discriminate the velar function (closure rate, degree of velar extension, etc.) during specific phonemes. HSV nasoendoscopy in clinics is impractical due to the time required to examine the images, the lack of acceptable audio recording when the HSV data is played back in slow motion, and the high cost of high-speed cameras. In a clinical setting, HSV can only be used as an adjunct to and supplement of standard nasoendoscopy.

In the context of diagnosing VPD, one crucial criterion for determining the validity of a testing technique is how well the results correlate with perceptual assessments of nasality. Paniagua et al. (2013) evaluated the findings of the auditory-perceptual assessment and VNE (gap size) in persons with cleft lip and palate. Subjects with moderate/severe hypernasality showed significantly less velopharyngeal closure. The severity of hypernasality showed the strongest and most consistent association with velopharyngeal gap size observed during VNE, compared with other speech disorders, including compensatory and obligatory disorders. The association between the degree of hypernasality and the presence of other disorders increases the chance of a moderate to large gap. Lipira et al. (2011) investigated the relationship between perceptual speech assessment and velopharyngeal closure function as assessed by nasoendoscopy and videofluoroscopy. The findings revealed a correlation between nasoendoscopy closure assessment and hypernasal speech. Since nasoendoscopy correlated more closely with the outcomes of the perceptual speech evaluation, this study supports the use of nasoendoscopy over videofluoroscopy.

Intra-rater and inter-rater reliability are crucial to the quality of any diagnostic test, particularly for procedures that rely on visual–perceptual judgments. Research on nasoendoscopy has consistently demonstrated strong reliability in evaluating velopharyngeal function. D'Antonio et al. (1989) reported high agreement among experienced clinicians, while novice raters achieved lower but still meaningful consistency, emphasizing the importance of training. Yoon et al. (2006) similarly found that inter-rater reliability was higher than inter-rater reliability when using the standardized Golding-Kushner reporting system. Together, these findings demonstrate that nasoendoscopy yields reproducible results, particularly when supported by structured protocols and clinical expertise.

Each imaging method has advantages and disadvantages. The primary advantage of VNE is that the size, location, and cause of the opening can be visualized, enabling suitable intervention planning. Furthermore, it can be used as biofeedback for individuals with VPD induced by mislearning (Siegel-Sadewitz & Shprintzen, 1982; Neumann & Romonath, 2012). The primary disadvantage of VNE is its invasive nature, which may be difficult for young children to endure. Furthermore, the velopharyngeal function may be evaluated subjectively, but the quantitative analysis is challenging since three-dimensional anatomy is represented in two dimensions, restricting visual information to surface physiology from a two-dimensional viewpoint (Lam et al., 2006; Lipira et al., 2011). The challenges associated with using nasoendoscopic examination to measure lateral wall displacement have been documented in several studies (Lipira et al., 2011; Stringer & Witzel, 1989).

Multiview videofluoroscopy (MVF). MVF is a widely used radiographic technique for evaluating the structure (depth of pharynx, length of velum), movement, extent of closure, and timing of the velopharyngeal sphincter (Havstam et al., 2005). Because the view angle can be changed, MVF detects lateral wall movement and depth of velopharyngeal closure during phonation with greater precision than nasopharyngoscopy (Paniagua et al., 2013; Ysunza et al., 2016).

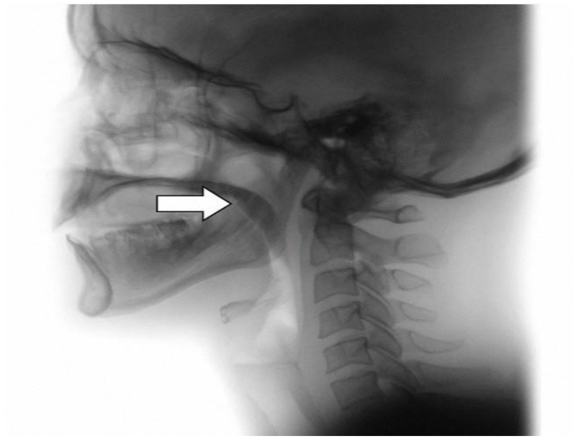


Figure 1.7. Lateral view videofluoroscopy showing the velum (soft palate) during speech. Reproduced from Ysunza et al. (2016), with permission.

During phonation, fluoroscopic images are recorded in the lateral, frontal, and base planes (Witt et al., 2000).It is possible to observe timing issues and unusual compensatory movements, such as tongue-backing. The lateral view makes it possible to observe abnormalities in the velum, posterior pharyngeal wall, tongue, Passavant ridges, velopharyngeal closure height, and tonsil size (Witt et al., 2000; Bettens et al., 2014). In addition, the amount of velar closure and the height of velopharyngeal closure may be clearly seen on lateral MVF. (Lam et al., 2006). The frontal view allows for the observation of lateral wall motions. Due to overlapping structures and difficulties in gathering sufficient information on sphincteric closure, a third view, the basal or Towne projection, is required. The basal view, like the nasoendoscopic view, reveals the link between the velum and the lateral-posterior parts of the pharyngeal wall (Witt et al., 2000).

According to Havstam et al. (2005), to lessen the burden of care for cleft patients, the first step in the clinical assessment of velopharyngeal function should be videofluoroscopy in lateral projection, combined with the auditory-perceptual examination. When further information is required or if a patient has a suspected or undiagnosed submucous cleft palate, a nasendoscopy should be performed.

Périco et al. (2014) examined the consistency between perceptual findings (hypernasality and air emission) and videofluoroscopy outcomes (consistent velopharyngeal closure, inconsistent velopharyngeal closure, and nonvelopharyngeal closure). The outcomes of this study showed a considerable degree of agreement between the perceptual tests and videofluoroscopy for consistent velopharyngeal closure and non-closure, but not for inconsistent closure. Another research study evaluated the effects of secondary palatal surgery for VPD correction using videofluoroscopy, nasoendoscopy, and perceptual evaluation. The results showed that videofluoroscopy supported the perceptual evaluation and that it was consistent with nasoendoscopy (Sommerlad et al., 2002). Sinclair et al. (1982) evaluated the lateral and basal reliability of videofluoroscopy and nasoendoscopy. They discovered that nasoendoscopy and lateral videofluoroscopy were both 80 percent reliable; however, basal videofluoroscopy was only 60 percent accurate. Birch et al. (1999) investigated videofluoroscopy measurement errors in healthy adults. They discovered that the measurement angular lifting of the soft palate above the hard palate, velopharyngeal distance, and soft palate extension at maximal closure can be made reliably and accurately. According to a systematic review of studies on intra-rater and inter-rater reliability for measurements in videofluoroscopy of swallowing, assessors should use well-defined guidelines for variable levels, follow pre-experimental training protocols, and achieve maximum agreement on the definition of measured variables to obtain accurate measurements in videofluoroscopy (Baijens et al., 2013).

Videofluoroscopy exposes the patient to radiation, albeit at a low dosage, which is important because patients with clefts usually receive recurrent radiographic exams at a young age. Images can be challenging to interpret when large adenoids are present, which is relatively frequent in children. During MVF, many projections are acquired; however, the three-dimensional architecture of the velopharynx is reduced to two-dimensional images, which may result in an overestimation of velopharyngeal closure (Havstam et al., 2005; Bettens et al., 2014).

Dynamic magnetic resonance imaging (MRI). MRI offers a high spatial resolution for imaging soft tissue, the velopharyngeal gap, internal components of the velopharyngeal mechanism, and assessment of submucous cleft palate. It is the only imaging technique that can accurately depict the location, shape, and function of muscles (Kuehn et al., 2001; Perry et al., 2014). Dynamic MRI images expand the capabilities of normal MRI into the temporal domain, exhibiting velopharyngeal components moving at rates of up to 15 or even 28 frames per second (Perry et al., 2014).

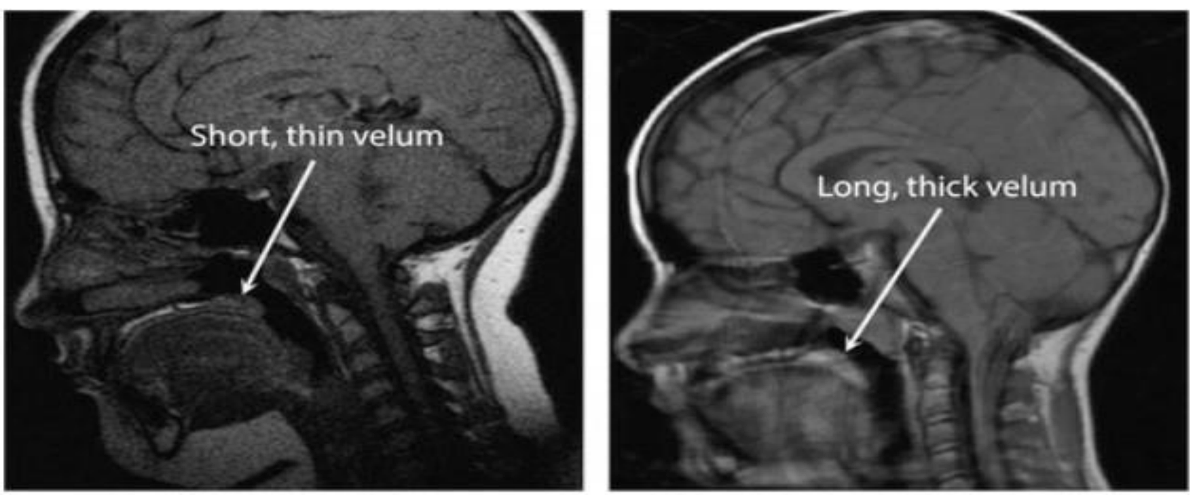


Figure 1.8. MRI Comparison of Velar Length and Thickness in Two Individuals. Reproduced from Kollara et al. (2017) under Creative Commons Attribution (CC BY 4.0) license.

MRI's multiplanar capability makes it possible to see and measure the velopharyngeal port in all three planes—sagittal, coronal, and axial. The mid-sagittal view reveals velum length and movement, and posterior pharyngeal wall forward movement during velopharyngeal closure. The coronal view illustrates the pharynx's width and the function of the lateral pharyngeal wall in velopharyngeal closure. Furthermore, the axial image at the hard palate level provides information on the type and degree of velopharyngeal closure (Perry et al., 2014; Perry et al., 2018).

Although MRI has been used to assess VPD since 1999, it is not being used routinely in clinical cleft care. The absence of a clear role for dynamic MRI in clinical decision-making compared to traditional imaging methods is one reason for the limited clinical application of MRI in cleft care. Studies in this field have suggested that MRI does not replace VNE or MVF in the evaluation of VPD, but it may be considered a complementary imaging modality in selected challenging cases (Beer et al., 2004; Sagar & Nimkin, 2015). The efficiency of dynamic MRI in evaluating velopharyngeal function and influencing treatment decisions, when combined with MVF and VNE, requires further study.

Other drawbacks of MRI include high costs, machine noise, claustrophobia, and the potential effects of gravity on the velum due to the supine posture during the examination, as well as the difficulties of testing young children. Furthermore, fixed intraoral metal appliances have a detrimental impact on image quality (Perry et al., 2018).

Indirect Assessment Techniques

Indirect assessment procedures, including acoustic and aerodynamic measurements, allow for inferences about the velopharyngeal mechanism based on sound, airflow, and air pressure.

Aerodynamic measurements. Nasal air leakage is commonly assessed as part of a comprehensive evaluation of velopharyngeal function. This evaluation can be conducted through various methods, including qualitative assessments involving auditory-perceptual evaluation of nasal emissions or observation of nasal air condensation on an examination mirror, or quantitative analyses using aerodynamic analysis. Aerodynamic measurements have been utilized in scientific studies of velopharyngeal function for decades. They address velopharyngeal function based on the premise that an incompletely closed velopharynx allows nasal air to escape (Warren, 1967). Nasal airflow, oral air pressure, nasal air pressure, and flow-pressure ratio are the most often studied parameters in aerodynamic measurements (Warren et al., 1964; Dotevall et al., 2001; Jones, 2000; Rollins & Oren, 2020). It is possible to assess the size of the velopharyngeal port while conducting speech production by putting these measures into the formula by Warren and DuBois (1964) as follows:

$$\text{Orifice area} = \frac{\text{Volume rate of airflow through the orifice}}{0.65\sqrt{\frac{2\times(\text{intraoral air pressure}-\text{nasal air pressure})}{\text{density of air}}}}$$

Historically, various means to assess nasal airflow included fogged mirror tests (Chow et al., 2015), the Pressure-Flow method (Dalston et al., 1986), technical instruments such as the aerophonoscope (Devani et al., 1999), pneumotachograph (Dotevall et al., 2001), and Nasal Oral Ratiometry System (SNORS) (Main et al., 1999). The common method of assessing nasal airflow involves placing a modified oxygen mask over the nose and mouth, connected to airflow sensors. In addition to detecting airflow, this mask may contain microphones to record a speech signal (Mclean et al., 1997). Instead of wearing a mask, intraoral and nasal air pressure can also be measured by inserting flexible catheters into the mouth and/or nose (Rollins & Oren, 2020); however, this approach may lack precision. Simultaneous airflow measurements using intraoral and nasal air pressure sensors can be combined to determine the size of the velopharyngeal valve and provide data on the timing of velopharyngeal function. The timing of velopharyngeal movement and closure during speech is important in the assessment of hypernasality (Jones, 2000; Leeper et al., 1998 Zajac & Mayo, 1996; Dotevall et al., 2001) as there is evidence that the degree of perceived hypernasality (Warren et al, 1993;1994; Jones, 2006; Ha & Kuehn, 2011) might relate more closely to the duration of velopharyngeal opening time than to the amount of nasal air and sound escape..

Aerodynamic tests are rarely used in clinical settings, despite their potential to accurately measure the amount of air passing through the VP port. This is most likely a result of the relatively high costs, the technical complexity, and the need for specialized equipment (Karnell, 2011; Kuehn & Moller, 2000), as well as the limited accuracy of these systems at the low levels of air leakage typical of speakers with mild nasality issues (Cler et al., 2016).

Acoustic measurements. Acoustic methods are especially appealing since they use the same airborne signal as the human auditory-perceptual system and constitute the approach that most closely correlates to auditory-perceptual processing (Kataoka et al., 2001). Research has focused on the acoustic evaluation of nasality as a method for indirectly assessing velopharyngeal function. Indicators of oral-nasal balance disorders are sometimes derived from measurements of the relationship between nasal and oral acoustic energy. Various techniques can be used to extract this information from the acoustic signal, including accelerometry, nasometry, and spectral analysis.

***Accelerometry*.** Because oral–nasal balance difficulties are caused by improper coupling of the oral and nasal cavities, researchers have also tried to quantify nasality using accelerometry, which detects vibrations associated with nasal versus oral resonance during voiced speech. An accelerometer may be used to assess the mechanical reaction of the tissues when air and sound are released through the VP port (Cler et al., 2016). To illustrate, the Horii's Oral Nasal Coupling (HONC) Index was developed by Horii (1980) as a non-invasive method of evaluating velopharyngeal function during sustained vowels or running speech by computing the ratio of nasal signal amplitude to laryngeal signal amplitude to reduce the influence of vocal intensity. Horii's approach involves attaching two accelerometers to the nose (at the midpoint of the ala) and the neck (above the sternal notch) (Horii, 1980). The formula used to calculate the index is as follows:

$$HONC = (Arms(n))/((k*Arms(v))$$

Here, Arms(n) signifies the root-mean-square amplitude of the nasal accelerometer signal, Arms(v) represents the root-mean-square amplitude of the vocal accelerometer signal, and k is a calibration constant chosen so that HONC equals one during a sustained phonation of /m/. The HONC index is reported on a scale from zero to one, where a value of zero indicates an entirely oral signal and a value of one signifies the signal produced during sustained phonation of the sound /m/. Alternatively, the HONC index can also be expressed in decibels (Horii, 1980). The HONC can distinguish between normal and hypernasal speech acquired from oral and nasal stimuli (Sussman, 1995; Mra et al., 1998), shows high interobserver consistency (Mra et al., 1998), and has a strong correlation with perceived nasality (Horii, 1980; Laczi et al., 2005). However, it is rarely employed in clinical or research areas, as it is not commercially available as a pre-assembled product (Laczi et al., 2005).

To monitor the relative durations of nasal and oral signals and compute their ratio, Karling et al. (1985) introduced the Nasality Oral Ratio Meter (NORAM). Conceptually, this device is a variation of the Horii Nasal Accelerometer system, extending its application from amplitude-based to time-based measures. Specifically, accelerometers are positioned at the nasal cavity and at the larynx, enabling the detection of signal activity at each site. The temporal measures derived from these signals include the duration of nasal vibration (Tn) and the duration of laryngeal vibration (Tl) during phonation. These measures are based on the time segments in which each accelerometer records sustained vibration above a defined threshold. Using the following formula, the percentage of nasality (n) is obtained:

$$\mathrm{n} \text{ ¿ } \frac{Tn}{Tl*100}$$

Although NORAM has the potential to assess nasality for treatment monitoring (Lohmander-Agerskov et al., 1996), the method’s low inter- and intra-rater reliability, as well as its inability to distinguish between normal resonance and hypernasality (Karling et al., 1985), limit its clinical and research applications. Nasal accelerometry measures only nasal air and sound emission. The suggested measure is based on the energy measured by a nasal accelerometer; consequently, it is known as the nasal acceleration energy measure (NAEM). The NAEM measures nasal airflow acceleration during the release period of stop consonants (here, /p/), allowing it to distinguish NAE from conventional nasalization or hypernasality, which are frequently evaluated during vowels or flowing speech. Furthermore, the NAEM gives a measure of NAE severity. Future research will be required to validate this metric further and determine its correlation with other measures of oral-balance dysfunction (Cler et al., 2016).

Nasometry. Nasometry involves instrumental assessment of nasality, whereby nasal and oral sound pressure levels are quantified to measure oral-nasal balance. To this end, a ratio of nasal acoustic energy to total acoustic energy during speech production is derived. Widely used instruments include the Nasometer II (Kay Elemetrics/KayPENTAX/PENTAX Medical, Lincoln Park, NJ, USA) and the NasalView system (Tiger Electronics, Seattle, WA, USA). Like the Nasometer II, other commercially available systems, such as ICSpeech nasometry and the Glottal Enterprises oral–nasal system, employ comparable dual-microphone configurations to separate oral and nasal acoustic signals. Unfortunately, resulting nasalance values are not directly interchangeable across devices due to proprietary differences in their construction and signal-processing algorithms.

How nasalance is determined remains comparable across these systems in that acoustic energy is captured separately from the oral and nasal cavities. This is achieved by positioning a separation plate attached to a head-mounted apparatus between the upper lip and the nose, enabling two microphones to independently record oral and nasal sound pressure levels. The recorded signals undergo frequency filtering and signal processing (e.g., time windowing and energy averaging), after which nasalance is calculated as the ratio of nasal acoustic energy to the total (oral plus nasal) acoustic energy, expressed as a percentage ranging from 0% to 100% (Fletcher, 1976).

$$\textbf{Nasalance} = \frac{\textit{Nasal Sound Pressure Level}}{\textit{Nasal Sound Pressure Level+Oral Sound Pressure Level}} * 100$$

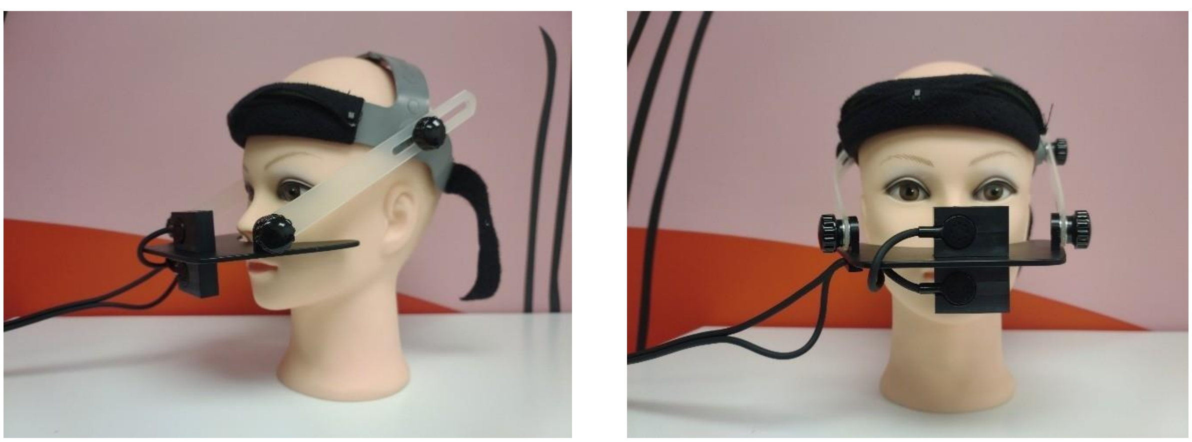


Figure 1.9. Nasometry System for Nasalance Measurement. Photograph taken by the author, June 16, 2025.

In cases where there is excessive nasal resonance, the scores for speech stimuli without nasal sounds tend to be higher than normal, indicating hypernasality. Conversely, when there is a lack of nasal resonance, the scores for speech stimuli containing nasal consonants are lower than usual, suggesting hyponasality (Dalston et al., 1991a; Kummer, 2008; de Boer & Bressmann, 2014). The Nasometer has the most published data on its reliability and normative values (Seaver et al., 1991; Dalston et al., 1991), its comparability to auditory-perceptual evaluation (Dalston et al., 1991; Bettens et al., 2018; Liu et al., 2022), and it is a widely used instrument in clinical practice (Kuehn & Moller, 2000; De Boer & Bressmann, 2015). There are several approaches to interpreting nasalance scores. Cut-off scores are commonly used in research to distinguish between normal, and sometimes hyponasal speech (Fukushiro et al., 2015) (Table 1.2).

	Zoo Passage	Nasal Sentences
Typical speakers	13.45 (SD, 5.94)	57.90 (SD, 6.69)
Mild hypernasality	17.68 (SD, 9.59)	40.94 (SD, 11.54)
Moderate to Severe hypernasality	34.06 (SD, 18.96)	51.00 (SD, 16.13)
Hyponasality		<50

Table 1.2. Mean Nasalance Scores and Standard Deviations from Nasometer Measurements for Typical and Hypernasality Conditions (Bressmann et al., 2006); Threshold Nasalance Score from Nasometer Measurements for Hyponasality (Dalston et al., 1991).

In addition to utilizing cut-off scores, Bressmann et al. (2000) introduced the concept of the nasalance distance to assess how effectively a speaker can distinguish between oral and nasal sounds. To compute the nasalance distance, two crucial measures are considered:

- Maximum Nasalance (nasal sentences): the mean nasalance obtained during nasal stimuli, typically yielding the highest scores.
- Minimum Nasalance (oral sentences): the mean nasalance obtained during oral (non-nasal) stimuli, typically yielding the lowest scores. The nasalance distance is then calculated by subtracting the minimum nasalance score from the maximum nasalance score as follows:

  Nasalance Distance = Maximum Nasalance Score – Minimum Nasalance Score

This distance measurement considers the individual's unique range of maximum and minimum nasalance values during the connected speech, acknowledging the inherent variations in speech patterns across different individuals. The Nasalance Distance has been shown to possess diagnostic value comparable to that of mean nasalance scores. (Bressmann et al., 2000; Bressmann et al., 2006).

Another alternative approach involves linear discriminant analysis (LDA), as employed by De Boer and Bressmann (2015) to develop a preliminary diagnosis formula based on nasalance scores. Nasalance values were calculated from normal speakers who simulated disorders of oral-nasal balance (hyponasal, hypernasal, and mixed nasality) based on oral and nasal speech stimuli. Nasalance-based LDA yielded formulas that accurately classified 88.6% of the four oral-nasal balance conditions. This research could lead to the development of statistical formulas for nasometric data, providing a quantitative classification of the main types of oral-nasal balance disorders that would complement the auditory-perceptual assessment.

Bettens et al. (2019) studied the clinical discriminatory power of the nasalance-based LDA formulas by using nasalance scores obtained from 55 Dutch-speaking children judged perceptually to have hypernasal, hyponasal, mixed, or normal nasality. The outcomes indicated that when using nasalance-based LDA to classify oral-nasal balance disorders, only 56% of the samples were accurately grouped into their respective diagnostic categories. Upon examining the data, it became evident that several samples were categorized as hyponasal, despite being perceived as normal or hypernasal. This discrepancy could be attributed to differences in language and the age of speakers in this study compared to a prior study by de Boer and Bressmann (2015). While adjustments and re-derived models improved accuracy to 80%, the persistent misclassifications suggest that nasalance-based LDA may not generalize reliably across different languages and age groups. This new approach for classifying oral-nasal balance disorders utilizing a nasalance-based LDA algorithm seems promising. However, it is important to note that further clinical research is necessary to refine the LDA functions and group centroids before practical clinical implementation can be considered.

Although computerized nasometry is the most commonly used acoustic measure for assessing oral-nasal balance disorders (Kuehn & Moller, 2000; De Boer & Bressmann, 2015), several limitations have been consistently reported in the literature. One major limitation is the high cost and limited accessibility of commercial nasometry systems, which require specialized equipment that may not be available in many clinical settings (Kılıç et al., 2021; Moreno-Torres et al., 2020). In addition to the constraints of cost and accessibility, Han et al. (2025) argue that traditional nasometer can be uncomfortable and difficult to adapt for a range of head shapes and sizes. Consequent adjustments may restrict natural lip movement and reduce suitability for prolonged or repeated use, particularly in pediatric populations (Han et al., 2025).

Recent studies have attempted to address these practical limitations by developing alternative or modified nasometry systems. For example, Kılıç et al. (2021) introduced a Praat-assisted nasalance meter as a low-cost alternative to commercial nasometers. Their system demonstrated that nasalance could be measured using readily available audio interfaces (external USB sound cards and portable audio recorders) and open-source software (Praat). This could substantially reduce financial barriers while still providing clinically meaningful data by producing nasalance measures that are comparable to those obtained with a commercial nasometer. The system was more cost-effective because it used non-proprietary hardware and open-source software, allowing for implementation with readily available recording devices. Despite cost reductions and a more flexible user interface, the authors emphasized that the system still required careful calibration, including matching the recording sensitivity (level) and frequency response of the oral and nasal microphones. Consequently, normative values remained device-specific. Similarly, Han and colleagues (2025) employed a pilot study design to evaluate a new device called Smart Naso™ with the intent to improve comfort, portability, and ease of use compared to the Nasometer II. Specifically, Smart Naso™ incorporates a lightweight, readily adjustable headgear with a flexible philtrum separator. Also, it operates via Bluetooth connection to a web-based application, thereby reducing reliance on traditional hardware systems. While this device demonstrated strong agreement with the Nasometer II and high test–retest reliability, it maintained prior nasalance computations requiring the nasal to total acoustic energy ratios, which are inherently dependent on the relative sensitivity and sound pressure level (SPL) balance between the oral and nasal microphones. Both systems may confer cost advantages and greater user interface flexibility compared to conventional nasometers, but neither is available on the open market.

Taken together, these studies demonstrate that while recent innovations have addressed some economic, comfort-, and portability-related limitations of traditional nasometry systems, they do not overcome the core methodological constraint of nasalance-based assessment. That is, all nasometry systems rely on dual-microphone signal separation and accurate microphone balance and therefore remain susceptible to variability arising from hardware characteristics, calibration procedures, and device-specific signal processing. Consequently, nasalance values obtained from different systems are not directly interchangeable (Bressmann, 2005; Bressmann, Klaiman & Fischbach, 2006), and careful interpretation requires adoption of device-specific normative data. There could be a window of opportunity to develop a new system with a simplified analysis algorithm that overcomes the limitations of the nasalance score. This motivated the present research study.

***Spectral analysis*.** Spectral analysis involves studying the frequency components of a signal to reveal how frequencies are distributed and their relative strengths within the signal. Several spectral characteristics of nasalized speech have been identified, such as pole-zero pairs(formant-anti formant) in the region of the first formant (Dodderi et al., 2016; Vijayalakshmi et al., 2007), a decrease in the amplitude of the first formant (Stevens, 1985), an increase in the bandwidth of the first and second formants (Fant, 1970), shifts in formant frequencies (Hawkins & Stevens, 1985), a rise in the amplitude between the first and second formants (Yoshida et al., 2000) and a decrease in the amplitude at or above the second formant (Lee et al., 2003). Figure 1.10 illustrates the vowel-dependent spectral effects of hypernasality for /a/, /i/, and /u/ in normal, mildly hypernasal, and moderately–severely hypernasal productions (Nikitha et al., 2017). For the vowel /a/ (panel a), hypernasality introduces an additional low-frequency resonance near the first formant, visible as a broad nasal peak associated with nasal coupling. For the vowel /i/ (panel b), hypernasality produces changes in the relative amplitudes of F1 and F2, accompanied by an additional nasal-related spectral pole between these formants. For the vowel /u/ (panel c), hypernasality results in an overall downward shift in formant locations compared with normal speech. Collectively, these patterns indicate that hypernasal speech exhibits alterations in formant structure, either through added nasal resonances or shifts in formant frequencies, as demonstrated by the linear predictive (LP) analysis.

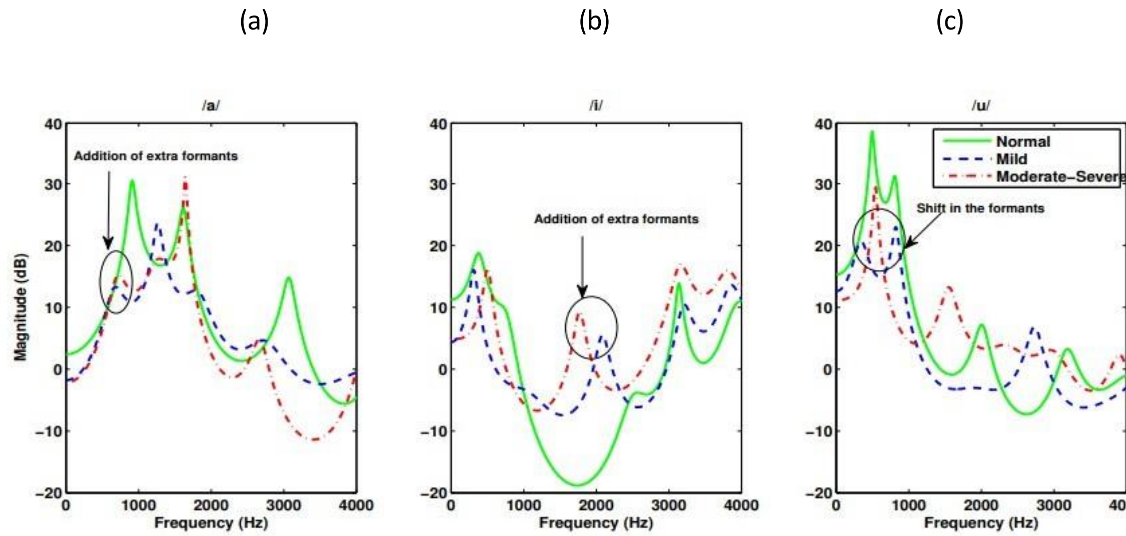


Figure 1.10. Spectra of vowels /a/ (panel a), /i/ (panel b), and /u/ (panel c) for normal, mild, and moderate–severe hypernasality. Panel (a) illustrates a low-frequency nasal peak with a broad bandwidth. Panel (b) illustrates the appearance of an extra nasal-related spectral pole between F1 and F2 and changes in the amplitude balance of these formants. Panel (c) demonstrates an overall downward shift in formant frequencies for /u/. Reproduced from Nikitha et al. (2017), with permission.

Two additional acoustic parameters (Chen, 1995, 1997), termed A1-P1 and A1-P0, can help quantify nasalization. A1 represents the amplitude of the first formant (F1), P1 represents the amplitude of the second nasal peak near F1, and P0 represents the amplitude of the first nasal peak. The first formant (F1) has the highest amplitude in the spectra of non-nasalized vowels, but for a nasalized vowel, the first nasal peak (P0), about 250 Hz to 400 Hz, can be as high as, or higher than, the first formant (Chen, 1997) (Figure 1.11). Chen (1995) studied the difference in amplitude of A1-P1 between the first and second formants of vowels produced by hypernasal hearing-impaired speakers. Except when the first formant and nasal peak (P1) frequencies were close, the difference in amplitude A1-P1 was demonstrated to be a helpful measure that correlates with listener evaluations of the degree of vowel nasality.

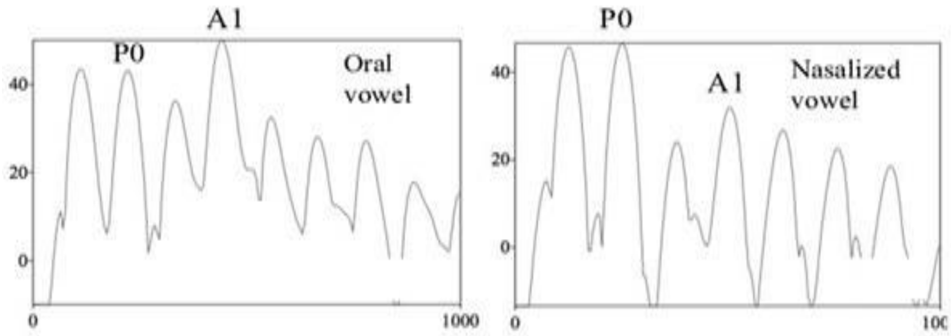


Figure 1.11. Spectra for an oral vowel (word "grade") and a nasalized vowel (word "grain") in a typical speaker. Reproduced from Tamminga, M., & Zellou, G. (2015), with permission.

Additionally, computer-automated algorithms can be used to distinguish between normal and nasalized speech based on spectral features of nasalized speech. The most common spectral characteristics include the One-third Octave Spectrum, Voice Low Tone to High Tone Ratio (VLHR), Time-frequency Analysis, and the Long-Term Average Spectrum (LTAS) (Bressmann & Abnavi, 2025).

One-third octave spectrum. The one-third octave analysis approach involves dividing each octave into three distinct frequency ranges and utilizing 16 digital band-pass filters in one-third octave intervals, with centre frequencies ranging from 15.6 Hz to 16 kHz, to filter the speech data. This technique aligns with the critical bandwidth of the ear's analyzing mechanism and possesses the ability to provide reliable quantitative measurements of nasality (Bakkum et al., 1995; Kataoka, 1988; Kataoka et al., 1996, 2001; Lee et al., 2003). Kataoka et al. (1996) used one-third octave bands to identify frequency bands that corresponded to the perceived hypernasality. They reported that during prolonged phonation of /i/, hypernasal productions had lower mean energy in the tenth, eleventh, twelfth, and thirteenth bands and greater mean energy in the fifth, sixth, and seventh bands. Kataoka et al. (2001) conducted a study focusing on the spectral analysis of differences in one-third octave frequency bands across three distinct groups: a normal resonance group, a mild to moderate hypernasal group, and a moderate to severe hypernasal group. The findings revealed significant differences in the intensities of these bands centred around frequencies of 630 Hz, 1000 Hz, 1600 Hz, 2000 Hz, and 3200 Hz. Specifically, the moderate to severe hypernasal group demonstrated significantly elevated amplitudes for the bands centred at 630 Hz, 1000 Hz, 1600 Hz, and 2000 Hz, while displaying significantly reduced amplitudes for the band centred at 3200 Hz.

Lee et al. (2003) employed one-third octave analysis to investigate the applicability of this method in assessing hypernasal speech in patients with dysarthria, maxillectomy, and cleft palate. The results revealed that individuals with hypernasality exhibited notably higher amplitudes in the frequency bands centred at 630 Hz, 800 Hz, and 1000 Hz. Conversely, individuals with hypernasal speech displayed significantly lower amplitudes in the frequency band centred at 2500 Hz compared to speakers with normal speech. Navya & Pushpavathi (2013) compared one-third octave spectral analyses to nasalance scores for vowels /a/ and /i/ to distinguish between hypernasal children with repaired cleft lip and palate and matched typical children. The results showed that the frequency region between 500Hz and 2663Hz had high sensitivity (0.87 to 0.75 for /a/ and 0.87 to 0.62 for /i/) and specificity (0.75 to 0.56 for /a/ and 0.87 to 0.81 for /i/) in differentiating between groups. Nevertheless, the outcomes of the conducted research show the challenges in accurately identifying particular frequency bands that consistently differentiate between hypernasal and normal resonance through one-third octave analysis. To address this, researchers have proposed new parameters that divide the spectrum into low and high-frequency components.

Voice low tone to high tone ratio (VLHR). VHHR is defined as a ratio of the sound power spectrum's low-frequency power (LPF) to high-frequency power (HPF) using a specific cutoff frequency (Kuo et al., 2003). In cleft palate patients with hypernasality, Lee et al. (2006) studied nasalance and VLHR analyses of the sustained nasal vowel /a/. When a cut-off frequency of 600

Hz was applied, correlations of VLHR with nasalance and hypernasality rating scores were shown to be significant for the normal and nasalized vowel /a/. The nasalized vowel /a/ increased the power ratio of the voice spectrum by producing nasal formants in the low frequency and anti-formants in the high frequency. Because of this power shift, VLHR increased in the nasalized vowel /a/. Lee et al. (2009) explored the use of VLHR for hypernasality detection by measuring the VLHR of six hypernasal patients' vowels. The authors investigated several cut-off frequencies, ranging from 200 Hz to 1600 Hz, and determined that the frequency range of 600 Hz to 800 Hz yielded the best results for six vowels. As a result, VLHR may be viewed as a method of objectively identifying hypernasality via prolonged vowels.

All the studies described used sustained vowels to determine spectral properties. Because mild and moderate hypernasality at the word or vowel level may be more challenging to identify, connected speech plays a crucial role in assessing oral-nasal balance for clinical diagnosis. Additionally, connected speech is one of the main tasks in the nasality evaluation procedure for individuals with cleft palate, as outlined in the Universal Parameters System and the Cleft Audit Protocol for Speech (CAPS), a standardized protocol developed to provide a reliable and consistent framework for assessing and reporting speech outcomes in cleft palate populations (John et al., 2006).

Long-term average spectrum (LTAS). LTAS analysis is an acoustic method used to examine the average frequency distribution of a speech signal, offering a detailed overview of the energy distribution across different frequency bands within the speech signal. It provides a comprehensive representation of the average spectral content of speech, encompassing both the spectral attributes of the vocal folds and the resonant characteristics of the vocal tract (Master et al, 2006; Tjaden et al, 2010). The term "long-term" in LTAS refers to the spectral analysis of speech signals over a more extended period, typically several seconds or minutes. According to Kitzing (1986) and Löfqvist (1986), when the analyzed signal is sufficiently long, typically ranging from 20 to 40 seconds, the resulting mean spectrum is less affected by differences in speech material, such as accent, articulation patterns, and individual characteristics.

De Boer and Bressmann (2016) studied the LTAS features in connected speech in different simulated oral-nasal balance disorders (hypernasality, hyponasality, and mixed nasality). In this study, the LDA method successfully categorized 80.7% of various oral-nasal balance disorders using long-term averaged spectra. The classification accuracy of 80.7% was roughly comparable to their earlier nasometry findings (88.6% in de Boer and Bressmann, 2015). They only computed discriminant functions from the acoustic spectra of normal women simulating disorders of oral-nasal balance; however, clinical participants with oral-nasal balance difficulties may have different discriminant functions.

Time-frequency analysis. Time–frequency analysis refers to a class of signal-processing techniques that are applied simultaneously to a speech signal in the time and frequency domains. Unlike traditional Fourier analysis, which provides only overall frequency content across time, time–frequency methods capture how spectral components change over time (Akan & Cura, 2021). This is particularly important for speech, where acoustic features vary rapidly within milliseconds. Common approaches include the Short-Time Fourier Transform (STFT), wavelet transforms, and Bionic Wavelet Transform (BWT). These methods divide the signal into short time windows and analyze the frequency content within each window, rather than averaging it across the entire signal, producing a dynamic representation that allows finer detection of transient features, such as short-duration increases in acoustic energy, that may be related to oral nasal balance disorders (Akan & Cura, 2021; Pradhan et al., 2023).

Golabbakhsh et al. (2017) investigated the non-linear automated recognition of hypernasality in connected speech of individuals with both normal speech and those with cleft lip and palate. Jitter, Mel-Frequency Cepstral Coefficients (MFCC), and shimmer were used individually in one-dimensional space, whereas BWT energy and MFCC, as well as BWT entropy and MFCC, were used in two-dimensional space. When MFCC was combined with BWT energy, the test's accuracy and sensitivity improved. Indeed, combining the criteria increases the likelihood of detecting hypernasality with better accuracy. Because both BWT and MFCC are based on the hearing system and derived from short time frames, they can quickly detect alterations in connected speech signals.

To date, research has largely focused on spectral measurements of hypernasality, with limited information available on the spectral characteristics of hyponasality and mixed nasality. However, because hypernasality and hyponasality can coexist in the same patient, several studies have emphasized the need to examine all forms of oral and nasal balance disorders as part of a comprehensive examination (Bressmann et al., 2000; de Boer & Bressmann, 2015).

Spectral analysis for determining nasality is non-invasive, quick, and objective, making the detection procedure simpler, less costly, and less subjective. Despite the data supporting the supposed efficiency of spectral measurement in the diagnosis of nasality, none of the spectral measurements are currently employed as a standard in clinical practice. Spectral studies have advanced our understanding of nasality's spectral properties, but face practical challenges (Bressmann & Abnavi, 2025). One such issue is the inability to draw practical conclusions from extensive spectral studies of nasality. Replication and broader application of techniques to diverse patient groups and speech samples are necessary for generalizability. Additionally, spectral patterns for different types of oral-nasal balance disorders are not well-established, with most data focusing on hypernasality. Spectral analysis is considered beneficial as it utilizes a single microphone, eliminating the requirement for specialized equipment to capture speech signals. However, using a single microphone, when placed at a distance from the mouth, presents difficulties in precisely extracting and quantifying nasality because it involves the merging of acoustic signals from both the oral and nasal cavities.

Methodological Considerations

Theoretical Framework of Research

The research paradigm is derived from the questions the researcher seeks to answer, based on ontological, epistemological, and methodological considerations (Scotland, 2012). This study aimed to enhance the accuracy and reliability of auditory–perceptual assessments of oral–nasal balance disorders by integrating acoustic and visual analyses as complementary tools. By incorporating acoustic measurements from the nasal signal alongside visual representations, such as oscillograms, the study sought to develop new measures to characterize normal, hypernasal, hyponasal, and mixed oral-nasal balance.

Ontology refers to the nature of reality and what is considered to exist (Scotland, 2012). The ontology of this research suggests that oral-nasal balance conditions have an objective existence, regardless of human interpretation or subjective perspectives. These conditions are assumed to be identified and classified based on observable and measurable acoustic features extracted from the nasal signal and oscillograms. The focus is on identifying the inherent characteristics and patterns that differentiate between normal, hypernasality, hyponasality, and mixed nasality through empirical investigation and the joint analysis of observable acoustic features and visual representations.

The epistemology in the context of this research question revolves around the belief that knowledge is generated through a combination of empirical observation and logical reasoning. The research aims to gather acoustic and visual data from the speech signal and oscillogram and use them to identify patterns and characteristics associated with different oral nasal balance conditions. This approach aligns with the belief that knowledge is built upon observations and evidence from both auditory and visual modalities. Once the empirical data is collected, this research will employ statistical techniques to develop the algorithm for classification. These analytical methods rely on logical reasoning and inference to identify the acoustic features and visual cues indicative of specific oral-nasal balance conditions.

The methodology of this research would involve employing a systematic and rigorous approach to gather empirical data, analyze it using objective methods, and draw reliable conclusions. Overall, the methodology of the current study emphasizes systematic data collection, objectivity, hypothesis testing, and statistical analysis to generate reliable and valid results in the development of the algorithm for the automatic detection and classification of oral-nasal balance conditions, considering both auditory and visual information.

Based on the theoretical perspective and research question, the post-positivist paradigm appears to be the most suitable approach for addressing the research objectives. Post-positivism is grounded in the belief that reality exists but can only be understood imperfectly, and that knowledge must be approached through systematic observation, empirical evidence, and critical reflection (Carpiano & Daley, 2006; Guba & Lincoln). This paradigm emphasizes objectivity and rigorous scientific inquiry while acknowledging the potential influence of researcher bias, measurement error, and contextual factors (Carpiano & Daley, 2006; Guba & Lincoln, 1994; Ryan, 2025). It encourages the use of multiple methods to allow for inquiry that can include multiple realities and perspectives as seen in qualitative research designs (Patton, 1999; Carter et al., 2014).

Post-positivism emerged from positivism, an approach to the study of science founded by Auguste Comte, a French philosopher. A positivist paradigm assumes that reality is fixed, singular, and fully knowable through objective measurement, hypothesis testing, and statistical analysis, and that researchers can observe phenomena without influencing them (Blumberg &

Feigl, 1931; Maksimovic & Evtimov, 2023; Verhaegh, 2024). While the current thesis adopts many positivist principles, including quantitative measurement, standardized procedures, and statistical analysis, a strictly positivist paradigm does not fully capture the philosophical stance needed for the line of inquiry. This is because treating findings as definitive or error-free would oversimplify the multidimensional nature of assessing oral–nasal balance disorders and the inherent limitations in the best known gold standard, being perceptual judgment by experienced clinicians. Although oral–nasal balance conditions are assumed to have an objective existence, they are difficult to classify and measure. Also, clinician and patient perspectives should play a central role in determining what might even constitute a disorder depending on the "ear of the beholder". That is, if an objective measure deemed nasality to be abnormal, this might not converge with the views of the speaker or a conversational partner. A post-positive approach can embrace research designs outside hypothesis-driven objective outcomes by incorporating the perspectives of the phenomenon in context with quantitative designs.

This could mean that work extending beyond the scope of the current thesis may include consideration of constructivist principles by adopting qualitative research approaches to complement the current quantitative findings. The constructivist paradigm holds that reality is socially constructed and that knowledge is subjective, context-dependent, and shaped by individuals' interpretations and lived experiences (Schwandt, 1994). Research within this paradigm typically emphasizes meaning-making and relies on qualitative methods, such as interviews or observations, to explore how individuals understand and interpret their experiences (Guba & Lincoln, 1994; Schwandt, 1994).

Although we aspire to consider multiple perspectives in future continued research, the current research studies attempt to apply standardized and controlled experimentation to support systematic pattern detection and classification of oral nasal balance disorders with intent to enhance clinical applicability of acoustic and visual analysis. In summary, the current thesis aligns best with post-positivist principles. Nevertheless, each of the three studies is grounded in empirical observation, where acoustic features are quantified and amplitude patterns are visualized, and statistical analysis lays the groundwork for interpreting the accuracy of classification findings. Because no single measurement or statistical approach can fully capture the complexity of velopharyngeal function and its resulting oral-nasal balance classifications, complementary sources of evidence (acoustic and visual) are explored to inform best practice. Although this study is framed within a post-positivist stance, in practice it remains primarily positivist, as it relies on quantitative acoustic and visual measures without incorporating a qualitative component. Thus, the present research represents an important step toward developing accurate and clinically feasible assessment methods. However, further qualitative inquiry, such as semi-structured interviews with speech-language pathologists and/or patients with CLP, would be valuable for understanding the phenomenon in context, corroborating the usefulness of the new practice tools, and understanding user-perceived classifications relative to traditional objective means. Ultimately, a mixed method line of inquiry will only enhance. Clinical practice and enable shared decision making for the benefit of those needing treatment for resonance disorders.

Knowledge Gaps

For decades, researchers have been looking for the best measuring method that can accurately evaluate the degree of velopharyngeal dysfunction and oral-nasal balance. There is a consensus in the literature that auditory-perceptual judgments remain the primary method for clinically evaluating nasality, despite criticism of their inherent subjectivity. To increase total assessment accuracy, acoustic assessments must be included in auditory-perceptual data (Kent, 1996; Marsh, 2009; Bettens et al., 2014, 2016; Paniagua et al., 2013). The acoustic measures have the potential to complete perceptual assessment and aid in the diagnosis, treatment, and follow-up of oral-nasal balance disorders.

Computerized nasometry, which utilizes a separation plate equipped with both nasal and oral microphones, is widely used for assessing nasality. However, its application in clinical settings is limited because it depends on costly technology and specialized equipment, which are often financially inaccessible to many clinicians and researchers. This limitation underscores the need for alternative methods that strike a balance between accessibility and accuracy in nasality assessment. In addressing this issue, researchers have used single-microphone measurements and spectrography as more affordable options. However, despite these efforts, the translation of research findings into clinical practice remains limited.

Given these limitations and in alignment with Kuehn's (1982) assertion that "a speech disorder does not exist until a listener perceives it," this study aimed to enhance the reliability of auditory–perceptual assessment. To this end, it explored the feasibility of incorporating two complementary approaches — acoustic and visual analyses — as tools to support and augment, rather than replace, auditory-perceptual evaluations. With respect to acoustic approaches, two recording methods exist. One involves positioning. A single microphone near the mouth to capture the combined acoustic output of the vocal tract. While this configuration is simple and widely accessible, nasal and oral acoustic components are mixed in the recorded signal and may mask key features of nasal speech. The second method is more widely adopted and makes use of two microphones, one for the oral and one for the nasal signal. Signal separation is achieved by placing a sound separation plate between the upper lip and the nose. This latter configuration is implemented in nasometry systems and is designed to minimize acoustic mixing between the oral and nasal signals (Dalston et al., 1991; Dalston & Seaver, 1992; Brunnegård et al., 2012; Bettens et al., 2018; Liu et al., 2022; Bressmann, 2015, 2016). In such systems, nasalance is calculated as the ratio of nasal acoustic energy to the total acoustic energy captured simultaneously (but separately) from the nasal and oral microphones. However, because nasalance is a ratio, it is inherently sensitive to factors affecting either of the two microphone channels, such as imperfect acoustic separation between the oral and nasal signals or variability in microphone placement. Previous research has demonstrated that relatively small variations in nasometer calibration values, specifically deviations in the relative level calibration between the oral and nasal microphone channels within the manufacturer's acceptable range (0.9–1.1), can systematically alter nasalance scores and increase test–retest variability (Hahm & Bressmann, 2022). In addition, microphone distance from the nose and mouth has been shown to significantly influence nasalance measurements, with changes in microphone position leading to systematic increases or decreases in nasalance values, particularly in the presence of unanticipated nasal emissions (Rollins & Oren, 2021). Together, these findings indicate that nasalance measurements are dependent on calibration and set-up factors such as microphone placement, which can alter measured values and potentially compromise classifications of oral-nasal balance disorders and/or test-retest outcomes.

In light of the challenges described above, the current research adopted two different approaches. Two studies attempted to classify oral-nasal balance disorders based only on the nasal speech signal. This still required separation of the nasal and oral speech signals, which was achieved with a standard nasometer headset (Figure 1.12). While this research relied on the existing technology, it is conceivable that future instruments could be developed with a separation plate housing only a nasal microphone. This would eliminate the need for exact calibration of the oral and nasal microphones. In the third study, oscillograms of both the oral and nasal signals were presented as visual information, so that raters could evaluate features associated with different oral-nasal balance disorders. Instructions for the visual analysis explicitly directed raters' attention to features associated with the nasal signal to maintain consistency with the nasal-focused analytical framework used throughout the thesis. As in the first two studies, standard signals from a nasometer headset were used. In future applications, it may be possible to present only the nasal signal for this evaluation method.

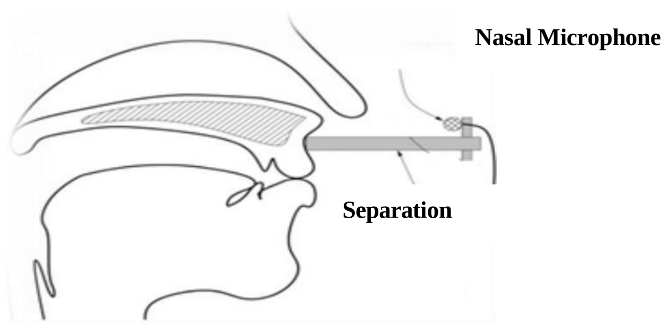


Figure 1.12. Schematic view of the separation plate and nasal and oral microphones of the nasometer. Adapted from Simpson, A. P. (2012), with permission.

For the acoustic investigations, an LDA approach was applied to discriminate among four nasality classifications. LDA is a statistical method used for dimensionality reduction and classification tasks, particularly when dealing with multiple features and distinct classes of data (Pohar et al., 2004). LDA operates by maximizing the ratio of the between-class scatter to the within-class scatter. The between-class scatter measures the distance between the means of different classes, while the within-class scatter measures the variability within each class. By maximizing this ratio, LDA finds a projection of the data that maximizes the separability between classes in the transformed feature space (Tharwat et al., 2017). In datasets with *d* features and K classes, LDA constructs a $(K-1)$-dimensional subspace where the data can be effectively classified using linear decision boundaries. This subspace is defined by the eigenvectors of $S_W^{-1}S_B$, where S_W represents the within-class scatter matrix, and S_B represents the between-class scatter matrix (Qu, L., & Pei, Y.,2024). In practical applications, LDA simplifies complex datasets by focusing on the most discriminative features, making it valuable for dimensionality reduction and classification (Tharwat et al., 2017).

LDA has proven effective in classifying oral-nasal balance conditions, as demonstrated in prior research. For example, Bressmann and colleagues used LDA in 2015 to develop diagnostic formulas based on nasalance scores, achieving an 88.6% correct classification rate with simulated data. A subsequent study by Bettens et al. (2019) further supported the use of nasalance-based LDA in clinical data classification. Additionally, Bressmann et al. (2016) later applied LDA to LTAS features, achieving an 80.7% accuracy rate in classifying oral-nasal balance disorders. These studies highlight LDA's efficacy in accurately categorizing these conditions, making it a promising choice for our study.

The second proposed approach in this study, visual analysis, focused on applying the oscillogram, which depicts sound waveforms, to classify oral–nasal balance conditions by identifying distinct waveform features associated with different conditions. Oscillograms are valuable because they provide a detailed graphical representation of speech signals, highlighting variations in amplitude over time. This visual representation enables a more thorough analysis of the temporal and intensity patterns of speech.

VPD, characterized by inadequate closure between the oropharynx and nasopharynx, affects speech intensity (Lundberg, 1991; Cairns et al., 1996). Variations in signal intensity can give rise to distinct patterns across different oral–nasal balance conditions. In nasometry, for example, higher nasalance values during non-nasal sentences are typically associated with hypernasality, whereas lower nasalance values during nasal sentences indicate hyponasality. Therefore, it is expected that oscillograms of stereo signals will also reflect these conditions through corresponding intensity pattern analysis that can be appreciated by visual inspection with the naked eye. For instance, hypernasality could be identified graphically by prominent and Noticeable intensity peaks (higher amplitude on the y-axis) in non-nasal sentences, whereas hyponasality would be characterized by the very low intensity in oscillograms of nasal sentences. In cases of mixed nasality, where both hypernasality and hyponasality coexist, energy peaks (again on the y-axis) during non-nasal sentences would indicate hypernasality, while decreased energy during nasal sentences would reflect hyponasality.

Thus, the goal of this second approach was to examine oscillograms of stereo signals, incorporating both oral and nasal components, to identify distinct intensity-based patterns that could improve the classification of oral–nasal balance conditions. The detailed visual analysis of intensity patterns in oscillograms of stereo signals is intended to complement auditory assessments by offering clear and trainable patterns of nasal and oral signal intensity.

Research Objectives

The primary objective of this research is to enhance diagnostic accuracy and clinical effectiveness in evaluating nasality among patients with cleft palates. This will be achieved through the development of two approaches:

Nasal Signal-Based Acoustic Feature Detection System. The study focused on creating a system for automatically detecting oral-nasal balance disorders using exclusively the acoustic features derived solely from the nasal signal. This approach introduced multiple acoustic features from only a nasal microphone to classify four oral-nasal balance conditions using LDA.

Visual Analysis of Stereo Signals. The second approach aimed to enhance the overall assessment process by integrating visual information into nasality assessments. This involved visually inspecting oscillograms of both oral and nasal stereo signals to classify four oral-nasal balance conditions.

Research Questions

I. How accurate is an LDA-derived automatic detection system, employing exploratory acoustic feature analyses derived solely from the nasal signal, in classifying oral-nasal balance patterns in simulated productions and subsequently among cleft palate patients?

II. How accurately do licensed speech-language pathologists (SLPs) and SLP students classify oral-nasal balance disorders using visual analysis of stereo acoustic signals?

Thesis Organization

This doctoral manuscript-based thesis is structured as a multi-method study presented across five chapters. Chapter 1 introduces the background, context, literature review, theoretical framework, and rationale for the study, and outlines the research questions and objectives. Chapter 2 presents the first study, which explores the effectiveness of acoustic parameters derived solely from nasal signals in classifying simulated oral-nasal balance disorders using linear discriminant analysis (LDA). Chapter 3 describes the second study, which applies the multiparametric acoustic diagnostic algorithm developed in Chapter 2 to a retrospective dataset of children with cleft palate to evaluate its clinical applicability and generalizability. Chapter 4 details a pilot study investigating the potential of visual pattern analysis using stereo acoustic signals as a complementary method for classifying oral-nasal balance disorders. Chapter 5 presents a general discussion and overall conclusion, summarizing the findings of each study and discussing their implications for clinical practice and future research. Chapter 6 outlines the contributions of the thesis collaborators. Finally, Chapter 7 includes the appendices, which provide supplementary materials related to the studies.

References

Akan, A., & Cura, O. K. (2021). Time–frequency signal processing: Today and future. *Digital Signal Processing*, *119*, 103216. https://doi.org/10.1016/j.dsp.2021.103216

Armour, A., Fischbach, S., Klaiman, P., & Fisher, D. M. (2005). Does velopharyngeal closure pattern affect the success of pharyngeal flap pharyngoplasty? *Plastic and Reconstructive Surgery*, *115(1)*. https://doi.org/10.1097/01.PRS.0000145635.87742.AF

Acharya, P., Sahana, B. A., Desai, A., Jalisatgi, R. R., & Kumar, N. (2021). Pre-operative and post-operative audiological assessment in cleft lip and palate patients-a prospective study. *Journal of Oral Biology and Craniofacial Research*, 11(2), 321-329.

Baijens, L., Barikroo, A., & Pilz, W. (2013). Intrarater and interrater reliability for measurements in videofluoroscopy of swallowing. *European Journal of Radiology*, *82(10*), 1683–1695. https://doi.org/10.1016/j.ejrad.2013.05.009

Bakkum, M. J., Plomp, R., & Pols, L. C. W. (1995). Objective analysis versus subjective assessment of vowels pronounced by deaf and normally hearing children. *Journal of the Acoustical Society of America*, 98, 745–762.

Beer, A. J., Hellerhoff, P., Zimmermann, A., Mady, K., Sader, R., Rummeny, E. J., & Hannig, C. (2004). Dynamic near-real-time magnetic resonance imaging for analyzing the velopharyngeal closure in comparison with videofluoroscopy. *Journal of Magnetic Resonance Imaging*, *20(5)*, 791–797. https://doi.org/10.1002/jmri.20197

Bettens, K., Wuyts, F. L., & van Lierde, K. M. (2014). Instrumental assessment of velopharyngeal function and resonance: A review. *Journal of Communication Disorders*, 52, 170–183. https://doi.org/10.1016/j.jcomdis.2014.05.004

Bettens, K., De Bodt, M., Maryn, Y., Luyten, A., Wuyts, F. L., & Van Lierde, K. M. (2016). The relationship between the Nasality Severity Index 2.0 and auditory-perceptual judgments of hypernasality. *Journal of Communication Disorders*, *62*, 67-81.

Bettens, K., Bruneel, L., Maryn, Y., De Bodt, M., Luyten, A., & Van Lierde, K. M. (2018). Perceptual evaluation of hypernasality, audible nasal airflow and speech understandability using ordinal and visual analogue scaling and their relation with nasalance scores. *Journal of Communication Disorders*, 76, 11-20.

Bettens, K., de Boer, G., Bressmann, T., Bruneel, L., & van Lierde, K. (2019). Clinical application of a new approach to identify oral–nasal balance disorders based on nasalance scores. *Cleft Palate-Craniofacial Journal, 56(5),* 628–638. https://doi.org/10.1177/1055665618796012

Blumberg, A. E., & Feigl, H. (1931). Logical positivism: A new movement in European philosophy. *The Journal of Philosophy*, 28(11), 281–296. https://doi.org/10.2307/2015437

Bressmann, T., Sader, R., Whitehill, T. L., Awan, S. N., Zeilhofer, H. F., & Horch, H. H. (2000). Nasalance distance and ratio: two new measures. *The Cleft Palate-Craniofacial Journal, 37(3)*, 248-256.

Bressmann, T. (2005). Comparison of nasalance scores obtained with the Nasometer, the NasalView, and the OroNasal System. *The Cleft Palate-Craniofacial Journal*, 42(4), 423–433.

Bressmann, T., Klaiman, P., & Fischbach, S. (2006). Same noses, different nasalance scores: Data from normal subjects and cleft palate speakers for three systems for nasalance analysis. *Clinical Linguistics and Phonetics, 20(2–3)*, 163–170. https://doi.org/10.1080/02699200500270689

Bressmann, T. (2019). *Nasometry*. In M. J. Ball (Ed.), *Manual of Clinical Phonetics* (Chap. 24). Routledge. https://ebrary.net/172346/health/nasometry (Retrieved April 20, 2025).

Bressmann, T., & Abnavi, F. (2025). Spectral speech analysis of hypernasality. In A. am Zehnhoff-Dinnen, J. Sopko, M.-C. Monfrans-Plauwadell, & K. Neumann (Eds.), *Phoniatrics II: Speech and speech fluency disorders—Literacy development disorders*. Springer. https://link.springer.com/book/9783031251757

Birch, M. J., Sommerlad, B. C., Fenn, C., & Butterworth, M. (1999). A study of the measurement errors associated with the analysis of velar movements assessed from lateral videofluoroscopic investigations. *The Cleft Palate-Craniofacial Journal*, *36*(6), 499-507.

Brunnegård, K. (2008). Evaluation of nasal speech: a study of assessments by speech-language pathologists, untrained listeners and nasometry (Doctoral dissertation, Klinisk vetenskap).

Carpiano, R. M., & Daley, D. M. (2006). A guide and glossary on postpositivist theory building for population health. *Journal of Epidemiology & Community Health, 60*(7), 564–570.

Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The Use of Triangulation in Qualitative Research. *Oncology Nursing Forum*, *41*(5), 545–547. https://doi.org/10.1188/14.ONF.545-547

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.

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

Chow, W., Brandt, M. G., Dworschak-Stokan, A., Doyle, P. C., Matic, D., & Husein, M. (2015). Validation of the mirror-fogging test as a screening tool for velopharyngeal insufficiency. *Open Otorhinolaryngol Journal*, *8*, 15–21.

Cler, G. J., Lien, Y. A. S., Braden, M. N., Mittelman, T., Downing, K., & Stepp, C. E. (2016). Objective measure of nasal air emission using nasal accelerometry. *Journal of Speech, Language, and Hearing Research*, 59(5), 1018–1024.

Cler, S., Knight, R. A., & Sell, D. (2017). Auditory-perceptual judgments of resonance, nasal airflow, understandability, and acceptability in speakers with cleft palate: Ordinal versus visual analogue scaling. *Cleft Palate-Craniofacial Journal, 54(1),* 19–31. https://doi.org/10.1597/15-164

Dalston, R. M., Warren, D. W., & Dalston, E. T. (1991a). A preliminary investigation concerning the use of nasometry in identifying patients with hyponasality and/or nasal airway impairment. *Journal of Speech, Language, and Hearing Research*, 34(1), 11–18.

Dalston, R. M., Warren, D. W., & Dalston, E. T. (1991b). Use of nasometry as a diagnostic tool for identifying patients with velopharyngeal impairment. *The Cleft Palate-Craniofacial Journal*, 28(2), 184–189.

Dalston, R. M., & Warren, D. W. (1986). Comparison of Tonar II, pressure-flow, and listener judgments of hypernasality in the assessment of velopharyngeal function. *The Cleft Palate Journal*, *23(2)*, 108–115.

D'Antonio, L. L., Marsh, J. L., Province, M. A., Muntz, H. R., & Phillips, C. J. (1989). Reliability of flexible fiberoptic nasopharyngoscopy for evaluation of velopharyngeal function in a clinical population. *Cleft Palate Journal, 26(3).*

D'Antonio, L. L. (1992). Evaluation and management of velopharyngeal dysfunction: a speech pathologist's viewpoint. *Problems in plastic and reconstructive surgery*, 2, 86–111.

de Almeida Santos, A. C., de Oliveira, F., Scarmagnani, R. H., Fukushiro, A. P., & Yamashita, R. P. (2016). The influence of listener training on the auditory-perceptual assessment of hypernasality. *Communication Disorders, Audiology and Swallowing Journal, 28(2),* 141–148. https://doi.org/10.1590/2317-1782/20162015163

de Boer, G., & Bressmann, T. (2014). Comparison of nasalance scores obtained with the nasometers 6200 and 6450. *Cleft Palate-Craniofacial Journal*, *51(1),* 90–97. https://doi.org/10.1597/12-202

de Boer, G., & Bressmann, T. (2015). Application of linear discriminant analysis to the nasometric assessment of resonance disorders: A pilot study. *Cleft Palate-Craniofacial Journal, 52(2)*, 173–182. https://doi.org/10.1597/13-109

de Boer, G., & Bressmann, T. (2016). Application of linear discriminant analysis to the long-term averaged spectra of simulated disorders of oral-nasal balance. *Cleft Palate-Craniofacial Journal, 53(5*), e163–e171. https://doi.org/10.1597/14-236

de Boer, G., Marino, V. C. de C., Dutka, J. de C. R., Pegoraro-Krook, M. I., & Bressmann, T. (2020). Nasalance-Based Preclassification of Oral–Nasal Balance Disorders Results in Higher Agreement of Expert Listeners' Auditory-Auditory-perceptual Assessments: Results of a Retrospective Listening Study. *Cleft Palate-Craniofacial Journal, 57(4),* 448–457. https://doi.org/10.1177/1055665619873506

de Stadler, M., & Hersh, C. (2015). Nasometry, videofluoroscopy, and the speech pathologist's evaluation and treatment. *Advances in Oto-Rhino-Laryngology*, *76,* 7–17. https://doi.org/10.1159/000368004

Denzin, N. K., & Lincoln, Y. S. (2005). *The SAGE handbook of qualitative research* (3rd ed.). SAGE Publications.

Devani, P., Watts, R., & Markus, A. E. (1999). Speech outcome in children with cleft palate: Aerophonoscope assessment of nasal emission. *Journal of Cranio-Maxillofacial Surgery (180).*

Dodderi, T., Narra, M., Varghese, S. M., & Deepak, D. T. (2016). Spectral analysis of hypernasality in cleft palate children: A pre-post surgery comparison. *Journal of Clinical and Diagnostic Research, 10(1),* MC01–MC03. https://doi.org/10.7860/JCDR/2016/15389.7055

Dotevall, H., Ejnell, H., & Bake, B. (2001). Nasal airflow patterns during the velopharyngeal closing phase in speech in children with and without cleft palate. *The Cleft Palate-Craniofacial Journal, 38(4),* 358–373.

Dudas, J. R., Deleyiannis, F. W. B., Ford, M. D., Jiang, S., & Losee, J. E. (2006). Diagnosis and treatment of velopharyngeal insufficiency: Clinical utility of speech evaluation and videofluoroscopy. *Annals of Plastic Surgery, 56 (5)*, 511–517. https:// *DOI:* 10.1097/01.sap.0000210628.18395.de

Fant, G. (1970). *Acoustic theory of speech production* (2nd ed.). Walter de Gruyter.

Fletcher, S. G. (1976). Nasalance vs. Listener Judgements of Nasality. *The Cleft Palate Journal, 13(1)*, 31-44.

Fukushiro, A. P., Ferlin, F., Yamashita, R. P., & Trindade, I. E. K. (2015). Influence of pharyngeal flap surgery on nasality and nasalance scores of nasal sounds production in individuals with cleft lip and palate. *Communication Disorders, Audiology and Swallowing Journal, 27(6*), 584–587. https://doi.org/10.1590/2317-1782/20152014088

Golabbakhsh, M., Abnavi, F., Kadkhodaei Elyaderani, M., Derakhshandeh, F., Khanlar, F., Rong, P., & Kuehn, D. P. (2017). Automatic identification of hypernasality in normal and cleft lip and palate patients with acoustic analysis of speech. *The Journal of the Acoustical Society of America, 141(2),* 929–935. https://doi.org/10.1121/1.4976056

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 195–220). Sage Publications.

Ha, S., & Kuehn, D. P. (2011). Temporal characteristics of nasalization in speakers with and without cleft palate. *The Cleft Palate-Craniofacial Journal*, 48(2), 134-144.

Hahm, J., & Bressmann, T. (2022). Effects of different calibration schedules on the test-retest differences of nasalance scores obtained with the Nasometer 6450. *Clinical linguistics & phonetics*, *36*(2-3), 292-300.

Han, K.-S., Kang, M.-S., & Ko, M.-H. (2025). Evaluation of a Newly Developed Device for Resonance Disorders: A Clinical Pilot Study. *Applied Sciences*, *15(10),* 5713. https://doi.org/10.3390/app15105713

Havstam, C., Lohmander, A., Persson, C., Dotevall, H., Lith, A., & Lilja, J. (2005). Evaluation of VPI-assessment with videofluoroscopy and nasoendoscopy. *British Journal of Plastic Surgery, 58(7),* 922–931. https://doi.org/10.1016/j.bjps.2005.02.012

Hawkins, S., & Stevens, K. N. (1985). Acoustic and auditory-perceptual correlates of the non-nasal–nasal distinction for vowels. *The Journal of the Acoustical Society of America, 77(4)*, 1560–1575.

Hixon, T. J., Weismer, G., & Hoit, J. D. (2008). *Preclinical speech science: Anatomy, physiology, acoustics, perception*. San Diego, CA: Plural Publishing.

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

Jensen, B. L., Kreiborg, S., Dahl, E., & Fogh-Andersen, P. (1988). Cleft lip and palate in Denmark, 1976–1981: Epidemiology, variability, and early somatic development. *The Cleft Palate Journal*, *25*(3), 258–269.

John, A., Sell, D., Sweeney, T., Harding-Bell, A., & Williams, A. (2006). The cleft audit protocol for speech—augmented: A validated and reliable measure for auditing cleft speech. *The Cleft palate-craniofacial journal, 43(3),* 272-288

Jones, D. L. (1991). Velopharyngeal function and dysfunction. *Clinics in Communication Disorders*, 1(3), 19–25.

Jones, D. L. (2000). The relationship between temporal aspects of oral-nasal balance and classification of velopharyngeal status in speakers with cleft palate. *The Cleft Palate-Craniofacial Journal, 37(4),* 363–369.

Jones, D. L. (2006). Patterns of oral-nasal balance in normal speakers with and without cleft palate. *Folia phoniatrica et logopaedica, 58(6),* 383–391.

Karling, J., Lohmander, A., de Serpa-Leitã;o, A., Galyas, K., & Larson, O. (1985). NORAM: calibration and operational advice for measuring nasality in cleft palate patients. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery, 19(3),* 261–267. https://doi.org/10.3109/02844318509074513

Karnell, M. P. (2011, May). Instrumental assessment of velopharyngeal closure for speech. In *Seminars in Speech and Language* (Vol. 32, No. 02, pp. 168–178). © Thieme Medical Publishers.Kataoka, R. (1988). Quantitative evaluation of hypernasality-Relation between spectral characteristics and perception of hypernasality. *Journal of the Japanese Cleft Palate Association*, 13(2), 204–216.

Kataoka, R., Michi, K. I., Okabe, K., Miura, T., & Yoshida, H. (1996). Spectral properties and quantitative evaluation of hypernasality in vowels. *Cleft Palate-Craniofacial Journal, 33(1).* https://doi.org/10.1597/1545-1569(1996)033<0043:SPAQEO>2.3.CO;2

Kataoka, R., Warren, D. W., Zajac, D. J., Mayo, R., & Lutz, R. W. (2001). The relationship between spectral characteristics and perceived hypernasality in children. *The Journal of the Acoustical Society of America, 109(5),* 2181–2189. https://doi.org/10.1121/1.1360717

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech and Language Pathology, 5(3),* 7–23.

Kitzing, P. (1986). LTAS criteria pertinent to the measurement of voice quality. *Journal of Phonetics, 14,* 477–482.

Kollara, L., Schenck, G., Jaskolka, M., & Perry, J. L. (2017). Examining a new method to studying velopharyngeal structures in a child with 22q11.2 deletion syndrome. *Journal of Speech, Language, and Hearing Research*, 60(4), 892–904. https://doi.org/10.1044/2016_JSLHR-S-16-0084

Kuehn, D. (1982). Assessment of resonance disorders. In N. Lass, L. McReynolds, J. Northern, & D. Yoder (Eds.), *Speech language and hearing: Vol. 11. Pathologies of speech and language* (pp. 499–524). Philadelphia, PA: W.B. Saunders.

Kuehn, D. P., & Moon, J. B. (1998). Velopharyngeal closure force and levator veli palatini activation levels in varying phonetic contexts. *Journal of Speech, Language, and Hearing Research, 41(1),* 51–62. https://doi.org/10.1044/jslhr.4101.51

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), 1–35.

Kuehn, D. P., Ettema, S. L., Goldwasser, M. S., Barkmeier, J. C., & Wachtel, J. M. (2001). Magnetic resonance imaging in the evaluation of occult submucous cleft palate. *The Cleft Palate-Craniofacial Journal, 38(5),* 421–431.

Kummer, A. W., Curtis, C., Wiggs, M., Lee, L., & Strife, J. L. (1992). Comparison of velopharyngeal gap size in patients with hypernasality, hypernasality and nasal emission, or nasal turbulence (rustle) as the primary speech characteristic. *Cleft Palate-Craniofacial Journal, 29(2).* https://doi.org/10.1597/1545-1569(1992)029<0152:COVGSI>2.3.CO;2

Kummer, A. W., & Lee, L. (1996). Evaluation and treatment of resonance disorders. *Language, Speech, and Hearing Services in Schools, 27*(3), 271–281. https://doi.org/10.1044/0161-1461.2703.271

Kummer, A. W. (2002). Velopharyngeal dysfunction: current thinking on the cause, effect, assessment, and treatment. *Current Opinion in Otolaryngology & Head and Neck Surgery, 10(6),* 455-459.

Kummer, A. W., Briggs, M., & Lee, L. (2003). The relationship between the characteristics of speech and velopharyngeal gap size. *The Cleft palate-craniofacial journal, 40(6),* 590-596.

Kummer, A. (2008). *Cleft palate and craniofacial anomalies: Effects on speech and resonance* (2nd ed.). New York, NY: Delmar Cengage Learning.

Kummer, A. W. (2009). Assessment of velopharyngeal function. *Comprehensive cleft care, 589-605.*

Kummer, A. W. (2011). Disorders of resonance and airflow secondary to cleft palate and/or velopharyngeal dysfunction. *Seminars in Speech and Language, 32(2),* 141–149. https://doi.org/10.1055/s-0031-1277716

Kummer, A. W. (2018). *Cleft palate and craniofacial conditions: A comprehensive guide to clinical management* (4th ed.). Burlington, MA: Jones & Bartlett Learning.

Laczi, E., Sussman, J. E., Stathopoulos, E. T., & Huber, J. (2005). Auditory-perceptual evaluation of hypernasality compared to HONC measures: The role of experience. *The Cleft palate craniofacial journal, 42(2),* 202-2

Lam, D. J., Starr, J. R., Perkins, J. A., Lewis, C. W., Eblen, L. E., Dunlap, J., & Sie, K. C. Y. (2006). A comparison of nasendoscopy and multiview videofluoroscopy in assessing velopharyngeal insufficiency. *Otolaryngology–Head and Neck Surgery, 134*(3), 394–402. https://doi.org/10.1016/j.otohns.2005.11.028

Lee, G., Yang, C. C., & Kuo, T. B. (2003). Voice low tone to high tone ratio-a new index for nasal airway assessment. *Chin J Physiol*, *46*(3), 123-27

Lee, A. S. Y., Ciocca, V., & Whitehill, T. L. (2003). Acoustic correlates of hypernasality. *Clinical Linguistics and Phonetics, 17(4–5),* 259–264. https://doi.org/10.1080/0269920031000080091

Lee, G. S., Wang, C. P., Yang, C. C. H., & Kuo, T. B. J. (2006). Voice low tone to high tone ratio: A potential quantitative index for vowel [a:] and its nasalization. IEEE Transactions on Biomedical Engineering, 53(7), 1437–1439 https://doi.org/10.1109/TBME.2006.873694

Lee, A., Brown, S., & Gibbon, F. E. (2008). Effect of listeners' linguistic background on auditory-perceptual judgements of hypernasality. *International Journal of Language and Communication Disorders, 43(5),* 487–498. https://doi.org/10.1080/13682820801890400

Lee, A., Whitehill, T. L., & Ciocca, V. (2009). Effect of listener training on perceptual judgement of hypernasality. *Clinical Linguistics and Phonetics, 23(5),* 319–334. https://doi.org/10.1080/02699200802688596

Leeper, H. A., Tissington, M. L., & Munhall, K. G. (1998). Temporal characteristics of velopharyngeal function in children. *The Cleft palate-craniofacial journal, 35(3),* 215-221.

Lipira, A. B., Grames, L. M., Molter, D., Govier, D., Kane, A. A., & Woo, A. S. (2011). Videofluoroscopic and nasendoscopic correlates of speech in velopharyngeal dysfunction. *Cleft Palate-Craniofacial Journal, 48(5),* 550–560. https://doi.org/10.1597/09-203

Liu, Y., Lee, S. A. S., & Chen, W. (2022). The correlation between perceptual ratings and nasalance scores in resonance disorders: A systematic review. *Journal of Speech, Language, and Hearing Research, 65(6),* 2215-2234.

Löfqvist, A. (1986). The long-time average spectrum as a tool in voice research. *Journal of Phonetics, 14*(3), 471–475.Lundberg, A. K. (1991). *Laryngeal compensation and linguistic stress in children with velopharyngeal dysfunction* (Undergraduate honors thesis, University of Northern Iowa). Presidential Scholars Theses (1990–2006), 107. https://scholarworks.uni.edu/pst/107

Main, A., Kelly, S., & Manley, G. (1999). Notes and discussion instrumental assessment and treatment of hypernasality, following maxillofacial surgery, using SNORS: a single case study. *International journal of language & communication disorders, 34(2)*, 223-238.

Maksimovic, J., & Evtimov, J. (2023). Positivism and post-positivism as the basis of quantitative research in pedagogy. Research in Pedagogy, 13(1), 208–218. https://doi.org/10.5937/IstrPed2301208M

Marsh, J. L., & Louis, S. (2003). Management of velopharyngeal dysfunction: differential diagnosis for differential management. *Journal of Craniofacial Surgery, 14(5)*, 621-628.

Marsh, J. L. (2004). The evaluation and management of velopharyngeal dysfunction. *In Clinics Plastic Surgery 31(2),* 261–269. https://doi.org/10.1016/S0094-1298(03)00124-X

Marsh, J. L. (2009). Velo-pharyngeal dysfunction: Evaluation and management. *Indian Journal of Plastic Surgery, 42(Suppl. 1),* S94–S103. https://doi.org/10.4103/0970-0358.57201

Master, S., Biase, N. D., Pedrosa, V., & Chiari, B. M. (2006). The long-term average spectrum in research and in the clinical practice of speech therapists. *Pró-Fono Revista de Atualização Científica, 18,* 111–120.

McLean, C. C., Kelly, S. W., & Manley, M. C. G. (1997). An instrument for the non-invasive objective assessment of velar function during speech. *Medical engineering & physics, 19(1)*, 7-14.

Mra, Z., Sussman, J. E., & Fenwick, J. (1998). HONC measures in 4-to 6-year-old children. *The Cleft Palate-Craniofacial Journal, 35*(5), 408–414.

Nagarajan, R., Savitha, V. H., & Subramaniyan, B. (2009). Communication disorders in individuals with cleft lip and palate: An overview. *Indian Journal of Plastic Surgery, 42(S 01*), S137-S143.

Navya, A., & Pushpavathi, M. (2013). One-third octave analysis: a diagnostic tool to measure nasality in conjunction with nasalance in children with repaired cleft lip and palate. *Journal of the All India Institute of Speech and Hearing, 32.*

Neumann, S., & Romonath, R. (2012). Effectiveness of nasopharyngoscopic biofeedback in clients with cleft palate speech systematic review. *Logopedics Phoniatrics Vocology, 37(3)*, 95–106. https://doi.org/10.3109/14015439.2011.638669

Nguyen, D. C., Patel, K. B., Skolnick, G. B., Skladman, R., Grames, L. M., Stahl, M. B., Marsh, J. L., & Woo, A. S. (2015). Progressive tightening of the levator veli palatini muscle improves velopharyngeal dysfunction in early outcomes of primary palatoplasty. *Plastic and Reconstructive Surgery, 136(1),* 131–141. https://doi.org/10.1097/PRS.0000000000001323

Nikitha, K., Kalita, S., Vikram, C. M., Pushpavathi, M., & Prasanna, S. M. (2017). Hypernasality severity analysis in cleft lip and palate speech using vowel space area. In *Proceedings of Interspeech 2017* (pp. 1829–1833). https://doi.org/10.21437/Interspeech.2017-1245

Oren, L., Rollins, M., Padakanti, S., Kummer, A., Gutmark, E., & Boyce, S. (2020). Using High-Speed Nasopharyngoscopy to Quantify the Bubbling Above the Velopharyngeal Valve in Cases of Nasal Rustle. *Cleft Palate-Craniofacial Journal, 57(5),* 637–645. https://doi.org/10.1177/1055665619894183

Ortega-Llebaria, M., & Prieto, P. (2008). Disentangling stress from accent in Spanish: Production patterns of the stress contrast in deaccented syllables. *In Segmental and prosodic issues in Romance phonology* (pp. 155-176). John Benjamins Publishing Company.

Paniagua, L. M., Signorini, A. V., da Costa, S. S., Martins Collares, M. V., & Dornelles, S. (2013). Comparison of videonasoendoscopy and auditory-auditory-perceptual evaluation of speech in individuals with cleft lip/palate. *International Archives of Otorhinolaryngology, 17(3),* 265–273. https://doi.org/10.7162/S1809-97772013000300006

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, *34*(5 Pt 2), 1189–1189.Périco, M. D. S., Dutka, J. D. C. R., Whitaker, M. E., Silva, A. F. R. D., Souza, O. M. V. D., & Pegoraro-Krook, M. I. (2014). Agreement between perceptual tests and videofluoroscopy in the diagnosis of velopharyngeal dysfunction. *Audiology-Communication Research*, *19*, 222–229.

Perry, J. L. (2011). Anatomy and physiology of the velopharyngeal mechanism. *Seminars in Speech and Language*, *32(2)*, 83–92. https://doi.org/10.1055/s-0031-1277712

Perry, J. L., Sutton, B. P., Kuehn, D. P., & Gamage, J. K. (2014). Using MRI for assessing velopharyngeal structures and function. *Cleft Palate-Craniofacial Journal*, *51(4)*, 476–485. https://doi.org/10.1597/12-083

Perry, J. L., Mason, K., Sutton, B. P., & Kuehn, D. P. (2018). Can dynamic MRI be used to accurately identify velopharyngeal closure patterns? *Cleft Palate-Craniofacial Journal*, *55(4)*, 499–507. https://doi.org/10.1177/1055665617735998

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

Pigott, R. W. (1969). The nasendoscopic appearance of the normal palatopharyngeal valve. *Plastic and Reconstructive Surgery*, *43(1*). https://doi.org/10.1097/00006534-196901000-00004

Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: a simulation study. *Metodoloski zvezki*, *1(1)*, 143.

Pradhan, B. K., Neelappu, B. C., Sivaraman, J., Kim, D., & Pal, K. (2023). A review on the applications of time–frequency methods in ECG analysis. *Journal of Healthcare Engineering*, *2023*(1), 3145483. https://doi.org/10.1155/2023/3145483

Pruthi, T., & Espy-Wilson, C. Y. (2004). Acoustic parameters for automatic detection of nasal manner. *Speech Communication*, *43*(3), 225-239.

Qu, L., & Pei, Y. (2024). A comprehensive review on discriminant analysis for addressing challenges of class-level limitations, small sample size, and robustness. Processes, 12(7), 1382.

Ryan, G. S. (2025). Postpositivist critical realism: philosophy, methodology and method for nursing research. Nurse researcher, 33(3). https://doi: 10.7748/nr 2019.e1598

Rintala, A. E., & Haapanen, M. L. (1995). The correlation between training and skill of the surgeon and reoperation rate for persistent cleft palate speech. *British Journal of Oral and Maxillofacial Surgery, 33(5),* 295-298.

Rollins, M., & Oren, L. (2020). Effects of nasal emission and microphone placement on nasalance score during /s/. *Proceedings of Meetings on Acoustics, 42(1).* https://doi.org/10.1121/2.0001353

Ruscello, D. (2007). Treatment of velopharyngeal closure for speech: Discussion and implications for management. *The Journal of Speech and Language Pathology – Applied Behavior Analysis, 2(1),* 55–75. https://doi.org/10.1037/h0100212

Sagar, P., & Nimkin, K. (2015). Feasibility study to assess clinical applications of 3-T cine MRI coupled with synchronous audio recording during speech in evaluation of velopharyngeal insufficiency in children. *Pediatric Radiology, 45(2),* 217–227. https://doi.org/10.1007/s00247-014-3141-7

Santoni, C., de Boer, G., Thaut, M., & Bressmann, T. (2020). Influence of voice focus adjustments on oral-nasal balance in speech and song. *Folia Phoniatrica et Logopaedica, 72*(5), 351–362. https://doi.org/10.1159/000501908

Sanu, P. M. (2018). Velopharyngeal insufficiency, https://drsanu.com/articles/velopharyngeal-insufficiency (Retrieved April 20, 2025)

Seaver, E. J., Dalston, R. M., Leeper, H. A., & Adams, L. E. (1991). A study of nasometric values for normal nasal resonance. *Journal of Speech, Language, and Hearing Research, 34(4),* 715-721.

Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 235–244). Sage Publications.

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching, 5(9),* 9–16. https://doi.org/10.5539/elt.v5n9p9

Siegel-Sadewitz, V. L., & Shprintzen, R. J. (1982). Nasopharyngoscopy of the normal velopharyngeal sphincter: an experiment of biofeedback. *The Cleft Palate Journal, 19(3),* 194-200.

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*(3), 477–490. https://doi.org/10.1016/j.wocn.2012.02.002

Sinclair, S. W., Davies, D. M., & Bracka, A. (1982). Comparative reliability of nasal pharyngoscopy and videofluorography in the assessment of velopharyngeal incompetence. *British Journal of Plastic Surgery, 35(2).* https://doi.org/10.1016/0007-1226(82)90146-1

Singh, D., Bastian, T. S., Kudva, S., Singh, M. K., & Sharma, P. (2015). Classification systems for orofacial clefts. *Oral and Maxillofacial Pathology Journal*, *6*(1).

Sommerlad, B. C., Mehendale, F. V., Birch, M. J., Sell, D., Hattee, C., & Harland, K. (2002). Palate re-repair revisited. *The Cleft palate-craniofacial journal, 39(3),* 295-307.

Stevens, K. N. (1985). Spectral prominences and phonetic distinctions in language. *Speech Communication journal, 4(1-3*), 137-144.

Stringer, D. A., & Witzel, M. A. (1989). Comparison of multi-view videofluoroscopy and nasopharyngoscopy in the assessment of velopharyngeal insufficiency. *Cleft Palate Journal, 26(2).*

Sussman, J. E. (1995). HONC measures in men and women: Validity and variability. *Cleft Palate-Craniofacial Journal, 32(1).* https://doi.org/10.1597/1545-1569_1995_032_0037_hmimaw_2.3.co_2

Tamminga, M., & Zellou, G. (2015). Cross-dialectal differences in nasal coarticulation in American English. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS).*

Tharwat, A., Gaber, T., Ibrahim, A., & Hassanien, A. E. (2017). Linear discriminant analysis: A detailed tutorial. *AI communications*, 30(2), 169-190. https://doi.org/10.3233/AIC-170729

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America, 123(5),* 2733–2749. https://doi.org/10.1121/1.2832337

Tjaden, K., Sussman, J. E., Liu, G., & Wilding, G. (2010). Long-term average spectral (LTAS) measures of dysarthria and their relationship to perceived severity. *Journal of Medical Speech-Language Pathology, 18(4*), 125.

Tokuda, I. (2021). The source–filter theory of speech. *Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.894

Trost-Cardamone, J. E. (1989). Coming to terms with VPI: a response to Loney and Bloem. *Cleft Palate Journal, 26*(1), 68-70.

Verhaegh, S. (2024). Logical positivism: The history of a "caricature." *Isis*, *115*(1), 46–66. https://doi.org/10.1086/728796

Vijayalakshmi, P., Reddy, M. R., & O'Shaughnessy, D. (2007). Acoustic analysis and detection of hypernasality using a group delay function. *IEEE Transactions on Biomedical Engineering, 54(4),* 621–629. https://doi.org/10.1109/TBME.2006.889191

Vyas, T., Gupta, P., Kumar, S., Gupta, R., Gupta, T., & Singh, H. P. (2020). Cleft of lip and palate: A review. *Journal of Family Medicine and Primary Care, 9*(6), 2621–2626.

Warren, D. W., & DuBois, A. B. (1964). A pressure-flow technique for measuring velopharyngeal orifice area during continuous speech. *Cleft Palate Journal, 1(1),* 52-71.

Warren, D. W. (1967). Nasal emission of air and velopharyngeal function. *The Cleft Palate Journal, 4(2),* 148-156.

Warren, D. W., Dalston, R. M., & Mayo, R. (1993). Aerodynamics of nasalization. In W. J. Hardcastle & J. Laver (Eds.), *Nasals, nasalization, and the velum* (pp. 119–146). Elsevier. https://doi.org/10.1016/B978-0-12-360380-7.50009-0

Warren, D. W., Dalston, R. M., & Mayo, R. (1994). Hypernasality and velopharyngeal impairment. *Cleft Palate-Craniofacial Journal, 31*(4), 257–262. https://doi.org/10.1597/1545-1569_1994_031_0257_havi_2.3.co_2

Witt, P. D., Marsh, J. L., McFarland, E. G., & Riski, J. E. (2000). The evolution of velopharyngeal imaging. *Annals of Plastic Surgery, 45*(6), 601–610. https://doi.org/10.1097/00000637-200045060-00019

Yamashita, R. P., Borg, E., Granqvist, S., & Lohmander, A. (2018). Reliability of hypernasality rating: Comparison of 3 different methods for auditory-perceptual assessment. *Cleft Palate-Craniofacial Journal*, *55*(8), 1060–1071. https://doi.org/10.1177/1055665618767116

Yanagisawa, E., Kmucha, S. T., & Estill, J. (1990). Role of the soft palate in laryngeal functions and selected voice qualities. Simultaneous velolaryngeal videoendoscopy. *The Annals of Otology, Rhinology, and Laryngology, 99(1), 18–28.*

Yoon, P. J., Starr, J. R., Perkins, J. A., Bloom, D., & Sie, K. C. (2006). Interrater and intrarater reliability in the evaluation of velopharyngeal insufficiency within a single institution. *Archives of Otolaryngology–Head & Neck Surgery, 132(9),* 947-951.

Ysunza, P. A., Bloom, D., Chaiyasate, K., Rontal, M., VanHulle, R., Shaheen, K., & Gibson, D. (2016). Velopharyngeal videofluoroscopy: Providing useful clinical information in the era of reduced dose radiation and safety. *International Journal of Pediatric Otorhinolaryngology*, *89*, 127–132. https://doi.org/10.1016/j.ijporl.2016.08.006

Yun, S. P., & de Chalain, T. (2008). Incidence of oronasal fistulae and velopharyngeal insufficiency after cleft palate repair: An audit of 211 children born between 1990 and 2004. *Cleft Palate-Craniofacial Journal*, *45*(2), 172–178. https://doi.org/10.1597/06-205.1

Yoshida, H., Furuya, Y., Shimodaira, K., Kanazawa, T., Kataoka, R., & Takahashi, K. (2000). Spectral characteristics of hypernasality in maxillectomy patients. *Journal of Oral Rehabilitation, 27,* 723–730.

Young, L. H., Zajac, D. J., Mayo, R., & Hooper, C. R. (2001). Effects of vowel height and vocal intensity on anticipatory nasal airflow in individuals with normal speech. *Journal of speech, language, and hearing research*, *44*(1), 52–60.

Ysunza, P. A., Bloom, D., Chaiyasate, K., Rontal, M., VanHulle, R., Shaheen, K., & Gibson, D. (2016). Velopharyngeal videofluoroscopy: Providing useful clinical information in the era of reduced dose radiation and safety. *International Journal of Pediatric Otorhinolaryngology*, *89*, 127–132. https://doi.org/10.1016/j.ijporl.2016.08.006

Zajac, D. J., & Mayo, R. (1996). Aerodynamic and temporal aspects of velopharyngeal function in normal speakers. Journal of Speech, Language, and Hearing Research, 39(6), 1199–1207. https://doi.org/10.1044/jshr.3906.1199

Zraick, R. I., & Liss, J. M. (2000). A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. Journal of Speech, Language, and Hearing Research, 43(4), 979–988.

https://doi.org/10.1044/jslhr.4304.979

Zajac, D. J., & Preisser, J. (2016). Age and phonetic influences on velar flutter as a component of nasal turbulence in children with repaired cleft palate. Cleft Palate-Craniofacial Journal, 53(6), 649–656. https://doi.org/10.1597/15-110

Chapter 2
Automatic Detection and Classification of Nasality Using Acoustic Analysis of Only the Nasal Speech Signal: A Pilot Study

Abstract

Objective: To develop an automatic classification algorithm to distinguish four oral-nasal balance conditions, normal, hypernasality, hyponasality, and mixed nasality—using acoustic analysis exclusively from nasal signals.

Methods: A retrospective dataset was analyzed, which included simulated oral-nasal balance conditions with both non-nasal and nasal stimuli. Exploratory acoustic feature extraction was performed on nasal signals, and the following parameters were calculated: mean and standard deviations of intensity, as well as the spectral amplitude of one-third octave frequency bands at 630Hz and 800Hz. Linear discriminant analysis was employed for classification.

Results: The automatic classification model achieved an accuracy of 90.9%, indicating the effectiveness of this multiparametric acoustic approach, which focuses on nasal signals.

Conclusion: This study demonstrated the potential of a multiparametric acoustic approach using nasal signals to automatically classify simulated oral-nasal balance conditions with high accuracy. By relying solely on nasal acoustic features, the method offers an alternative to traditional assessment techniques. Future research should focus on validating these findings in clinical populations to confirm the model's generalizability and practical applicability.

Introduction

Oral-nasal balance disorders disrupt the normal acoustic balance between the oral and nasal cavities during speech production. These disorders occur when the velopharyngeal valve fails to close properly, resulting in improper coupling or separation of sound between the oral and nasal cavities. This results in conditions such as hypernasality (excessive airflow through the nasal passages during speech), hyponasality (reduced airflow through the nasal passages due to blockage or narrowing), and mixed nasality (some of both hyper- and hyponasality) (De Boer & Bressmann, 2015; Santoni et al., 2020). These imbalances can significantly impact speech intelligibility and quality by altering the spectral patterns critical for clear and natural speech in patients with cleft palate (Hixon et al., 2008; Kummer, 2011).

Accurate evaluation and classification of oral-nasal balance disorders are essential for optimizing treatment in patients with cleft palate. Auditory-perceptual judgments are widely recognized as the primary clinical method for evaluating these disorders, in line with Kuehn's (1982) view that "a speech disorder does not exist until a listener perceives it." However, the subjective nature of these evaluations can lead to variability in scoring among speech-language pathologists (SLPs), particularly when training and experience differ. To enhance the accuracy of nasality assessment, objective methods can serve as valuable complements to auditory-perceptual data (Marsh, 2009; Bettens et al., 2018).

Acoustic methods are particularly appealing because they use the same airborne signals as the human auditory-perceptual system, providing a close correlation with auditory-perceptual processing. These techniques are non-invasive, rapid, and objective, generating data that can effectively support assessments of test-retest reliability and monitor interventional changes

(Kent, 1996; Kataoka et al., 2001; Paniagua et al., 2013; Bettens et al., 2018). Various studies have employed acoustic analysis techniques using different acoustic parameters, speech tasks, and classification methods to detect nasality, with a primary focus on distinguishing

Study	Speech Task	Classification	Classifier	Acoustic Parameters	Metric
Glass & Zue, 1985	Vowels	Normal, Hypernasality	Threshold	Average Spectra of Nasalized Vowel	Thresh. =74%
Cairns et al., 1996	Vowels	Normal, Hypernasality	Correlation	Teager Energy Operator (TEO)	Corr. = 93.3% for /a/ and 94.7% for /i/
Kataoka et al., 2001	Vowels	Normal, Hypernasality	Regression	One-third Octave Spectral Analysis	Reg.= 84%
Lee et al., 2003	Vowels	Normal, Hypernasality	Regression	One-third Octave Spectral Analysis	Reg.= 98%
Pruthi & Espy-Wilson, 2004	Sonorant Consonants	Normal, Hypernasality	SVR	Acoustic Parameters (APs)	Acc. = 89.53%
Vijayalakshmi et al., 2007	Vowels	Normal, Hypernasality	Threshold	Group Delay Function-based Acoustic Measure (GDAM)	Thresh.:100% for /a/, 88.78% for /i/, and 86.66% for /u/.
Lee et al., 2009	Vowels	Normal, Hypernasality	Correlation	Voice Low Tone to High Tone Ratio (VLHR)	Corr.= 62%
Akafi et al., 2012	Vowels	Normal, Hypernasality	K-means Clustering and Bayes Theorem	Cepstrum Coefficients	Accuracy: 81.12% (utterance-level), 97.14% (subject-level)
Tsai et al.,2012	Connected Speech	Normal, Hypernasality	Correlation	Voice Low Tone to High Tone Ratio (VLHR)	Corr.=76%
Navya & Pushpavathi., 2013	Vowels	Normal, Hypernasality	Receiver-Operating Characteristics (ROC)	One-third Octave Spectral Analysis	/a/: Sensitivity 0.87, Specificity 0.93, /i/: Sensitivity 1.00, Specificity 0.93,
He et al., 2014	Vowels	Four Hypernasality Levels	Gaussian Mixture Model (GMM)	Teager Energy Operator (TEO), First Formant, Number of Formants	Acc. = 83%
Nieto et al., 2014	Connected Speech, Vowels	Normal, Hypernasality	Threshold	Teager energy operator (TEO) Pitch, Jitter, Shimmer, Mel-Frequency Cepstral Coefficients (MFCC), and LPC	TEO=66.7% LPC and MFCC = 100%

hypernasality from normal speech (Table 2.1).

De Boer & Bressman., 2015	Connected Speech	Normal Hypernasality Hyponasalit Mixed nasality	LDA	Nasalance Scores	Acc. = 88.6%
Mirzaei & Vali, 2016	Vowels	Normal, Hypernasality	SVM	Group Delay Spectrum+ Wavelet Coefficients	Acc. = 94.1%
Orozco-Arroyave et al., 2016	Vowels Words	Normal, Hypernasality	Linear-Bayes	HNR in Cepstral Domain	Acc. = 94.6% (vowel), 90.8% (word)
De Boer et al., 2016	Connected Speech	Normal, Hypernasality Hyponasalit, mixed nasality	LDA	Long-Term Averaged Spectra (LTAS)	Acc. = 80.7%
Liu et al., 2016	Voiced Samples	Normal, Hypernasality	Back Propagation Neural Network	Homomorphic Spectrum Sequence	Acc. = 80%
Golabbakhsh et al., 2017	Connected Speech	Normal, Hypernasality	SVM	Mel-Frequency Cepstral Coefficients (MFCC), Jitter, Shimmer, Bionic Wavelet Transform Energy, Bionic Wavelet Transform Entropy	Acc. = 85%
Dubey et al., 2018	Vowels	Normal, Hypernasality	SVM	VSA + MFCC	Acc. = 91.70%
Dubey, et al., 2018	Vowels	Normal, Hypernasality	SVM	A pitch-adaptive, MFCC	Acc. = 83.45% for /a/ 88.04% for /i/ 85.58% for /u/
Wang et al., 2019	Vowels	Normal, Hypernasality	Deep Recurrent Neural Network (DRNN)	Vocal Tract Shape-Based Features, Formant-Based Features, Vocal Tract-Based Cepstral Features	Acc. = 93.35%
Dubey et al., 2020	Words	Normal, Hypernasality	SVM	Normalized Harmonic Amplitude (NHA), Harmonic Amplitude Ratio (HAR), Prominent Harmonics Frequency (PHF)	Acc. = 87.89

Table 2.1. Summary of Studies Conducting Acoustic Analysis for Nasality Detection.

Nasometry is a commonly used acoustic method for evaluating nasality by measuring nasalance, the percentage of acoustic energy released through the nasal cavity compared to the total energy from both the oral and nasal cavities during speech. Several approaches are used to interpret nasalance scores for diagnosing oral-nasal balance disorders. Cut-off scores are commonly used with specific thresholds for hypernasality and hyponasality (Dalston et al., 1991; Bressmann et al., 2006). The nasalance distance, introduced by Bressmann et al. (2000), calculates the difference between maximum nasalance (oral stimulus) and minimum nasalance (nasal stimulus), providing a diagnostic value comparable to mean nasalance scores. Nasalance-based linear discriminant analysis (LDA), developed by de Boer and Bressmann (2015), achieves an 88.6% classification accuracy for oral-nasal balance disorders.

The Nasometer, with extensive data on its reliability, normative values, and comparability to auditory-perceptual evaluation (Dalston et al., 1991; Bettens et al., 2018; Liu et al., 2022), has been primarily used for hypernasality assessment, with limited studies on hyponasality (Brunnegård et al., 2012; Dalston & Seaver, 1992) and mixed nasality (Bressmann, 2015, 2016). Despite being a non-invasive and objective tool, nasometry remains an expensive technology, requiring specialized equipment that may not be accessible to all clinicians (Han, Kang, & Ko, 2025).

As an alternative, researchers have applied spectral analysis to single-microphone recordings to identify hypernasality. Spectral analysis is a powerful method for investigating the frequency components of speech and their distribution, particularly in assessing changes caused by nasal-oral coupling in velopharyngeal dysfunction (VPD). Various spectral features of nasalized speech have been identified, including pole-zero pairs (formant-anti-formant) near the first formant (Dodderi et al., 2016), decreased first formant amplitude (Stevens, 1985), increased bandwidths of the first and second formants (Fant, 1970), shifts in formant frequencies (Hawkins & Stevens, 1985), and amplitude changes across formants (Lee et al., 2003). Specific parameters, such as A1-P1 and A1-P0, further quantify nasality by analyzing the amplitudes of formants and nasal peaks (Chen, 1995, 1997).

Several spectral analysis studies have identified objective and reliable methods for assessing hypernasality. The One-third octave spectrum divides frequencies into three bands per octave, aligning with the auditory system's natural perception and enabling precise measurement of hypernasality. Kataoka et al. (1996) identified specific frequency bands (fifth to seventh) associated with hypernasality during prolonged phonation of /i/. Subsequent studies by Kataoka et al. (2001) found significant amplitude differences in frequency bands centred at 630 Hz, 1000

Hz, and 1600 Hz in moderate to severe hypernasal speech. Similarly, Lee et al. (2003) observed consistent spectral patterns in individuals with hypernasal speech, characterized by heightened amplitudes at 630 Hz and 800 Hz, and diminished amplitudes at 2500 Hz. Navya and Pushpavathi (2013) demonstrated the diagnostic capability of one-third octave analysis, highlighting its sensitivity and specificity in distinguishing hypernasal children from typical peers.

The voice low-tone to high-tone ratio (VLHR), defined as the power ratio of a voice spectrum with a specific cutoff frequency, has demonstrated potential for detecting hypernasality. Lee et al. (2006, 2009) investigated VLHR as a quantitative index for hypernasality, finding significant correlations with nasalance and perceptual ratings of hypernasality. Their results suggest that VLHR is a reliable tool for evaluating hypernasality in sustained vowels. Tsai et al. (2012) reported a significant correlation between VLHR and nasalance in English-connected speech. Dodderi et al. (2016) examined VLHR in children with cleft palate before and after surgery, finding significantly lower postoperative VLHR values, which indicated a reduction in nasality.

Long-term average spectrum (LTAS) is a practical and effective measure for analyzing the frequency distribution of a speech signal by averaging its spectral characteristics over time. Research suggests that nasalization influences the speech signal differently across various frequency ranges. Chen (1997) found that nasalized speech exhibits increased energy around 250 Hz compared to normal speech, likely due to the first nasal formant amplifying lower harmonics. Similarly, Haapanen et al. (1996) observed that the mixed condition had a lower amplitude between 300 Hz and 700 Hz than the hypernasal condition. Furthermore, de Boer and Bressmann

(2016) demonstrated that LTAS can classify simulated oral-nasal balance disorders with an accuracy of 80.7%, highlighting its diagnostic potential.

Despite recent advancements, the clinical application of spectral analysis for nasality assessment still faces challenges that require further research. Most studies have primarily focused on hypernasality, with less attention given to hyponasality or mixed nasality. However, in patients with cleft palates, nasal obstruction can result in the simultaneous occurrence of both hypernasality and hyponasality (De Boer & Bressmann, 2015). Much of the existing research has relied on sustained vowels or syllables to detect hypernasality. Assessing nasality using vowels and syllables can be problematic because these may not accurately reflect a speaker's abilities or real communication, as conversations are based on connected speech. According to the universal parameters system and the cleft audit protocol for speech (CAPS) (John et al., 2006), for clinical diagnosis, connected speech plays a crucial role in identifying hypernasality, as mild to moderate hypernasality can be harder to detect in syllables or vowels (Lohmander et al., 2009; De Boer & Bressmann, 2015). A single microphone placed near the mouth captures both nasal and oral sounds simultaneously, making it challenging to accurately analyze and measure nasality using spectrographic data. The blending of these signals complicates the isolation of nasality. Additionally, dynamic changes in the articulators during speech production can significantly affect oral signals, making it challenging to differentiate VPD-related abnormalities from articulatory variation.

This study is motivated by the need to overcome challenges in identifying and classifying oral-nasal balance disorders while improving the accuracy of auditory-perceptual analysis through the integration of complementary approaches. To establish a reliable foundation, we conducted a pilot study to explore the most effective acoustic parameters derived from nasal signals for classifying simulated oral-nasal balance disorders, emphasizing the need to effectively separate oral and nasal signals in acoustic analysis. This approach uses a separation plate and a nasal microphone to isolate nasal signals, effectively minimizing the mixing or interference of oral and nasal sound signals. This isolation enables a more accurate and precise measurement of nasality in speech. The nasal microphone is positioned close to the nose with the specific purpose of capturing signals emanating from the nasal cavity. Oral–nasal balance disorders are associated with alterations in nasal acoustic energy; therefore, analyzing nasal signals may provide clearer insights into the acoustic features of nasality. By exploring a novel multivariate acoustic diagnostic technique derived from nasal signal recordings, this study sought to develop a multiparametric method for the automated detection and classification of various forms of oral-nasal balance disorders. Combining acoustic parameters in a multiparametric approach may more effectively distinguish the distinct and relevant features of oral-nasal balance conditions by overcoming the limitations of individual parameters (Van Lierde et al., 2007; Bettens et al., 2016; Bressmann et al., 2000; Golabbakhsh et al., 2017).

This study aimed to develop a multivariate acoustic approach utilizing exploratory acoustic feature analyses derived exclusively from nasal signals for detecting and classifying oral-nasal balance disorders in simulated conditions. Additionally, we evaluated the accuracy of the LDA-based multivariate acoustic diagnostic technique, which relies solely on nasal signal features, to classify oral-nasal balance patterns within simulated data. The central hypothesis of this exploratory study was that nasal signal–based acoustic features would demonstrate discrimination of simulated oral–nasal balance conditions using an LDA model with classification accuracy at least comparable to previously reported nasalance- and LTAS-based LDA approaches (Bressmann et al., 2015, 2016).

Materials and Methods

Study Design

This pilot study employed a retrospective, observational design involving the secondary analysis of a speech dataset previously collected and analyzed by de Boer and Bressmann (2015). Ethical approval for the secondary use of this dataset was obtained under Protocol #38903, titled "Re-analysis of data from previous protocol 'Nasalance and Acoustic Profile of Mixed Nasal Resonance' (original approval date: February 26, 2020; valid through February 25, 2026; REB Ref #27918).

Participants

Retrospective data were gathered from a speech database comprising eleven typical female speakers selected from the University of Toronto student population. The participants met specific criteria outlined in the dataset, including an age range of 22 to 30 years, native English speakers, normal hearing, and no history of cleft lip and palate, resonance issues, or nasal congestion (de Boer and Bressmann, 2015).

Speech Database and Recording Procedure

The speech database used in the study consisted of recordings of simulated oral and nasal balance disorders, which were collected as part of the research conducted by de Boer and Bressmann in 2015. Participants were instructed to simulate various oral-nasal balance conditions, including normal resonance, hypernasality (lowering the velum and nasalizing all speech sounds), hyponasality (closing one nostril with the index finger), and mixed nasality (speaking with a lowered velum and one closed nostril). Two trained listeners were involved in a validation process to ensure the accuracy of the participants' simulations of oral-nasal balance conditions. They independently evaluated and assessed the simulations and then reached a consensus decision through discussion (de Boer and Bressmann, 2015).

The Nasometer II 6450 (KayPentax, Montvale, NJ) was used for recording speech signals (Figure 1). This device utilizes two microphones separated by a plate pressed against the face, and the recordings were made at a sampling rate of 44.1 kHz with a 16-bit signal resolution. Only the signals captured by the nasal microphone were utilized to isolate the effects of nasality on the spectrographic nasal output. The recorded samples were then digitally transferred directly into Praat software (Version 6.4.27) (Boersma & Weenink, 1992–2022), where acoustic analysis was performed using Praat scripts.

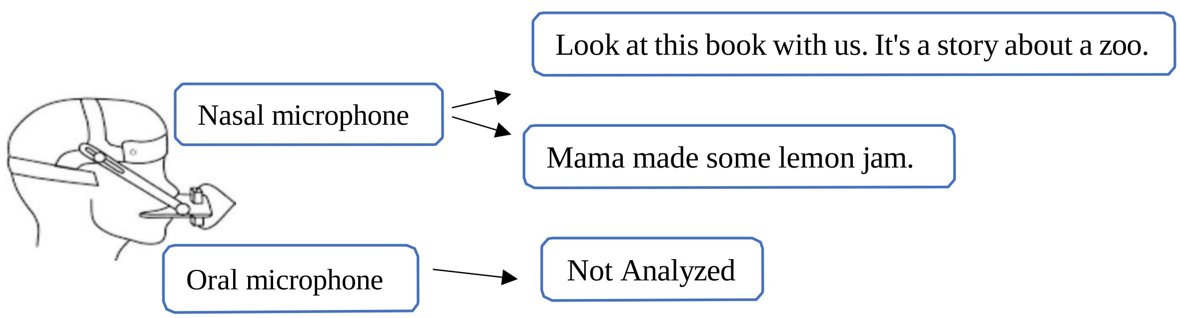


Figure 2.1. Nasometer Headset Set-up for Collecting Acoustic Data.

Stimuli

The study made use of oral and nasal stimuli derived from the first two sentences of the Zoo Passage ("Look at this book with us. It's a story about a zoo") and from the first sentence of the Nasal Sentences ("Mama made some lemon jam") (Fletcher, 1976). The stimuli for each resonance condition, including normal, hypernasal, hyponasal, and mixed nasal, were repeated twice in a randomized order (de Boer & Bressmann, 2015). Each repeated production was treated as a separate acoustic speech signal to support the exploratory aim of the study, which focused on signal-level acoustic feature analysis and algorithm development. Although repetitions were produced by the same speakers, they were not interpreted as independent speakers. Accordingly, the unit of analysis was the speech signal rather than the individual speaker, and repeated productions from the same speaker were not interpreted as independent observations. The recordings were segmented into individual sentences using Praat software and saved in *.wav format. In total, 88 speech samples were included in the dataset, with 22 samples available for each of four oral-nasal balance conditions.

Acoustic Analysis

The primary focus of the research was to develop a multivariate diagnostic approach using a nasal microphone to assess nasality by exploring a range of quantitative parameters for analyzing the acoustic characteristics of nasality. The study incorporated a comprehensive review of existing literature on the application of acoustic parameters in detecting oral-nasal balance conditions. Acoustic parameters were assessed based on criteria such as ease of implementation, interpretability, and their potential to accurately detect various forms of oral-nasal balance disorders. This led to specific assessments and analyses of parameters, including one-third octave spectrum frequency ranges (17 frequency bands in the centre frequency range of 80 to 4000 Hz), voice low tone high tone ratio (with cut-off frequencies ranging from 200 Hz to 1000 Hz and 2500 Hz), and vocal intensity measurements.

One-third octave spectral analysis. In the present study, the one-third-octave spectrum within the centre frequency range of 80 Hz to 4 kHz was used to measure the spectral amplitude of non-nasal and nasal stimuli collected via a nasal microphone. This frequency range encompasses the bands identified with high sensitivity to hypernasality, as reported in the literature (Kataoka et al., 2001; Lee, 2003). For each non-nasal and nasal stimulus, 17 frequency bands were identified, resulting in a total of 34 input variables for each participant.

Evidence suggests that the number of input variables used influences the performance of data-fit prediction algorithms, as a larger number may introduce complexity and noise, while reducing the number can simplify the model and improve its ability to predict new, unseen data. A simpler model focuses on the most relevant features and can enhance generalization, accuracy, and interpretability (Guyon & Elisseeff, 2003; Japkowicz & Shah, 2011).

Initially, principal component analysis (PCA) was utilized to decrease the number of input variables and prevent overfitting (Jolliffe, 1986). PCA reduced 34 input variables into eight principal components, resulting in a simpler model but diminishing data interpretability. PCA was ineffective in identifying the most important frequency bands within the one-third octave bands. Since the study aimed to both reduce variables and identify key frequency bands, PCA was unsuitable for this purpose. Consequently, the study shifted to an evidence-based approach, selecting key frequency bands within the one-third octave spectrum that have been shown to detect nasality effectively. Previous research (Kataoka et al., 1996, 2001; Lee et al., 2003; Navya et al., 2013) has highlighted the importance of frequency bands centred at 630 Hz, 800 Hz, and 1000 Hz in distinguishing normal speech from hypernasal speech. By narrowing the analysis to specific frequency bands, the study effectively reduced the feature set to six variables (3 variables for each nasal and non-nasal stimulus), while preserving the critical features needed for nasality detection.

De Boer and Bressmann (2015) suggested examining both non-nasal and nasal speech stimuli collectively in the diagnostic procedure to achieve a comprehensive classification of nasality. Therefore, to further minimize variables while optimizing the diagnostic process, this study developed the Oralance formula, which integrates the spectral characteristics of both non-nasal and nasal sentences. Three important frequency bands centred at 630 Hz, 800 Hz, and 1000 Hz were incorporated into the Oralance formula to enhance nasality measurements. The formula is defined as follows:

$$Oralance = \frac{Nasal\ SPL\ for\ the\ Non-Nasal\ Sentence}{Combined\ Nasal\ SPL\ for\ the\ Nasal + Non-Nasal\ Sentences} * 100$$

As a result, the variables Oralance630, Oralance800, and Oralance1000 were established to serve as final prediction variables for the one-third octave bands centred at 630 Hz, 800 Hz, and 1000 Hz, respectively.

Voice low tone to high tone ratio (VLHR). VLHR is an acoustic parameter identified for its effectiveness in detecting hypernasality and its strong correlations with nasalance and perceptual hypernasality ratings (Lee et al., 2003). This ratio is calculated by dividing the power in the low-frequency (LF) region by the power in the high-frequency (HF) region of the average spectrum, using a cutoff frequency of Fc Hz (Kuo et al., 2003; Lee et al., 2003). Two methods were used to select the final variables: testing cut-off frequencies ranging from 200 Hz to 1000 Hz and 2500 Hz, and considering existing evidence from previous studies (Lee et al., 2006, 2009; Tsai et al., 2012; Dodderi et al., 2016) to refine the selection of the most relevant variables for classification. Based on this evidence and the results obtained from testing various cut-off frequencies, final cut-off frequencies of 700 Hz, 1000 Hz, and 2500 Hz were chosen for calculating the VLHR. To incorporate both non-nasal and nasal spectral patterns and minimize variables, the Oralance formula was computed using cutoff frequencies of 700 Hz, 1000 Hz, and 2500 Hz. This approach included prediction variables for the VLHR as oralance-VLHR700, oralance-VLHR1000, and oralance-VLHR2500, enabling a comprehensive assessment of nasality.

Intensity. The intensity of a sound wave refers to the amount of energy transmitted through a unit area perpendicular to the direction of wave propagation per unit of time. In simpler terms, it measures the strength or power of a sound wave (Hixon et al., 2008). The intensity of speech signals can vary between nasal speech and normal speech due to velopharyngeal dysfunction and improper nasal-oral cavity coupling. Velopharyngeal dysfunction, characterized by inadequate closure between the oropharynx and nasopharynx, alters speech intensity by disrupting normal airflow and causing excessive air escape through the nasal cavity rather than being directed adequately through the oral cavity (Lundberg, 1991). Cairns et al. (1996) demonstrated that nasalization alters the energy distribution of speech by introducing anti-resonances and modifying the intensity of the first formant, leading to measurable changes in the speech signal. Energy-based signal processing approaches, such as the teager energy operator (TEO), have demonstrated viability in differentiating hypernasal speech from normal speech (Cairns et al., 1996; He et al., 2014; Dubey et al., 2020).

Additionally, studies have shown that vocal intensity can serve as a behavioral strategy to influence velopharyngeal function (Young et al., 2001). Consequently, intensity measures can serve as effective parameters for identifying nasality in speech. In this study, the mean and standard deviation of intensity during the production of nasal and non-nasal sentences were used as predictor variables. Specifically, these variables were denoted as mean nasal intensity, SD nasal intensity, mean non-nasal intensity, and SD non-nasal intensity.

Statistical Analysis

SAS statistics software version 25.0 (IBM Corp, Armonk, New York) was used for the statistical analysis of the data in this study. An ANOVA was conducted to examine the effect of oral–nasal balance condition on each input variable, followed by Tukey HSD post hoc tests to compare group means with control of the Type I error rate. To classify oral-nasal balance conditions, LDA was employed to identify the optimal combination of input variables for reliable and effective classification outcomes. The LDA approach, as used by de Boer and Bressmann (2015, 2016), was employed as a classifier to identify the optimal combination of input variables, automatically classify data, and identify key features by sequentially selecting the most significant ones. This approach enabled the development of a predictive model that highlights the distinguishing characteristics of all forms of oral-nasal balance disorders. N-fold stratified cross-validation was conducted. This method preserves the proportion of classes within each fold, providing a more reliable estimate of model performance. It also ensures that every sample is used for both training and testing, which is particularly valuable given our limited sample size.

Results

Various sets of acoustic parameters, including one-third octave spectrum frequency ranges, voice low-to-high tone ratios, and vocal intensity measurements, were used as input variables for LDA, and the accuracy of each set in classifying four oral nasal balance conditions was determined. Table 2.2 shows only those sets with an accuracy greater than 50%The most accurate and interpretable combination with fewer predictors was chosen, forming a multiparametric acoustic combination of measurements. Through iterative LDA, the most accurate and easily interpretable parameter combination was identified, consisting of six variables: Oralance630, Oralance800, mean nasal intensity, standard deviation (SD) of nasal intensity, mean non-nasal intensity, and SD of non-nasal intensity.

Input Variables	Number of Input Variables	Accuracy	Error Rate	Reason for Excluding
One-third octave spectrum with centre frequencies from 80 and 4000 Hz	34 Input Variables	97%	Normal: 0.04 Hyper:0.04 Hypo:0.00 Mixed:0.00	Large numbers of input variables
Applied PCA on one-third octave spectrum with centre frequencies ranging from 80 and 4000	8 principal components	87%	Normal: 0.18 Hyper:0.18 Hypo:0.09 Mixed:0.04	Impossibility of determining important frequency bands, Low Interpretability
VLHR with cut-off frequencies of 1000,2500,700	6 Input Variables	65%	Normal: 0.40 Hyper:0.18 Hypo:0.31 Mixed:0.45	Low accuracy
Oralance630, Oralance800, Oralance1000, Oralance-VLHR 1000, Oralance-VLHR 2500	5 Input Variables	79%	Normal:0.0 Hyper:0.50 Hypo:0.045 Mixed:0.27	The high error rate in detecting mixed nasality, Intermediate accuracy
Oralance630, Oralance800, Oralance1000, Oralance-VLHR 1000, Oralance-VLHR 2500, Oralance-VLHR 700	6 Input Variables	79%	Normal: 0.00 Hyper:0.45 Hypo:0.04 Mixed:0.31	Intermediate accuracy, complexities of outcomes interpretation
Oralance630, Oralance800, Mean Nasal Intensity, SD Nasal Intensity, Mean Non-Nasal Intensity, SD Non-Nasal Intensity.	6 Input Variables	90.9%	Normal:0.09 Hyper:0.09 Hypo:0.04 Mixed:0.13	

Table 2.2. Sets of acoustic parameters used as input variables for linear discriminant analysis

Repeated Measures ANOVA

Repeated-measures ANOVA tests revealed statistically significant differences for all variables across the four oral–nasal balance conditions ($P < 0.0001$; Table 2.3), indicating that each acoustic measure varied meaningfully among normal resonance, hypernasality, hyponasality, and mixed nasality.

VARIABLES	CONDITIONS				ANOVA		
	Normal M±SD	Hypernasality M±SD	Hyponasality M±SD	Mixed nasality M±SD	DF	F-Value	P-Value
Oralance 630	17.23±7.09	46.47±12.12	30.84±11.96	46.60±12.77	3	34.87	<.0001
Oralance 800	10.66±4.47	41.10±14.73	16.84±5.11	46.81±14.21	3	59.74	<.0001
Mean Nasal Intensity	67.03±3.12	70.45±3.50	56.63±5.87	62.03±6.41	3	32.70	<.0001
SD Nasal Intensity	17.23±1.98	18.03±3.82	12.89±2.27	13.99±2.27	3	18.78	<.0001
Mean Non-Nasal Intensity	49.10±2.78	67.32±6.16	44.76±3.08	60.04±5.97	3	102.01	<.0001
SD Non-Nasal Intensity	15.04±2.40	20.07±2.36	12.62±1.16	15.47±2.31	3	47.20	<.0001

Table 2.3. Comparison of acoustic measures for four oral-nasal balance conditions

Figure 2.2 presents the results of the post hoc pairwise comparisons for the oral-nasal balance conditions, corrected for a significance level of 0.05. Mean Non-Nasal Intensity is the most effective variable, distinctly separating all four conditions. Oralance630, Mean Nasal Intensity, and SD Non-Nasal Intensity show moderate effectiveness, differentiating three conditions while combining some others. Variables such as Oralance800 and SD Nasal Intensity are less effective, with considerable overlap, making them more suitable for complementary analyses.

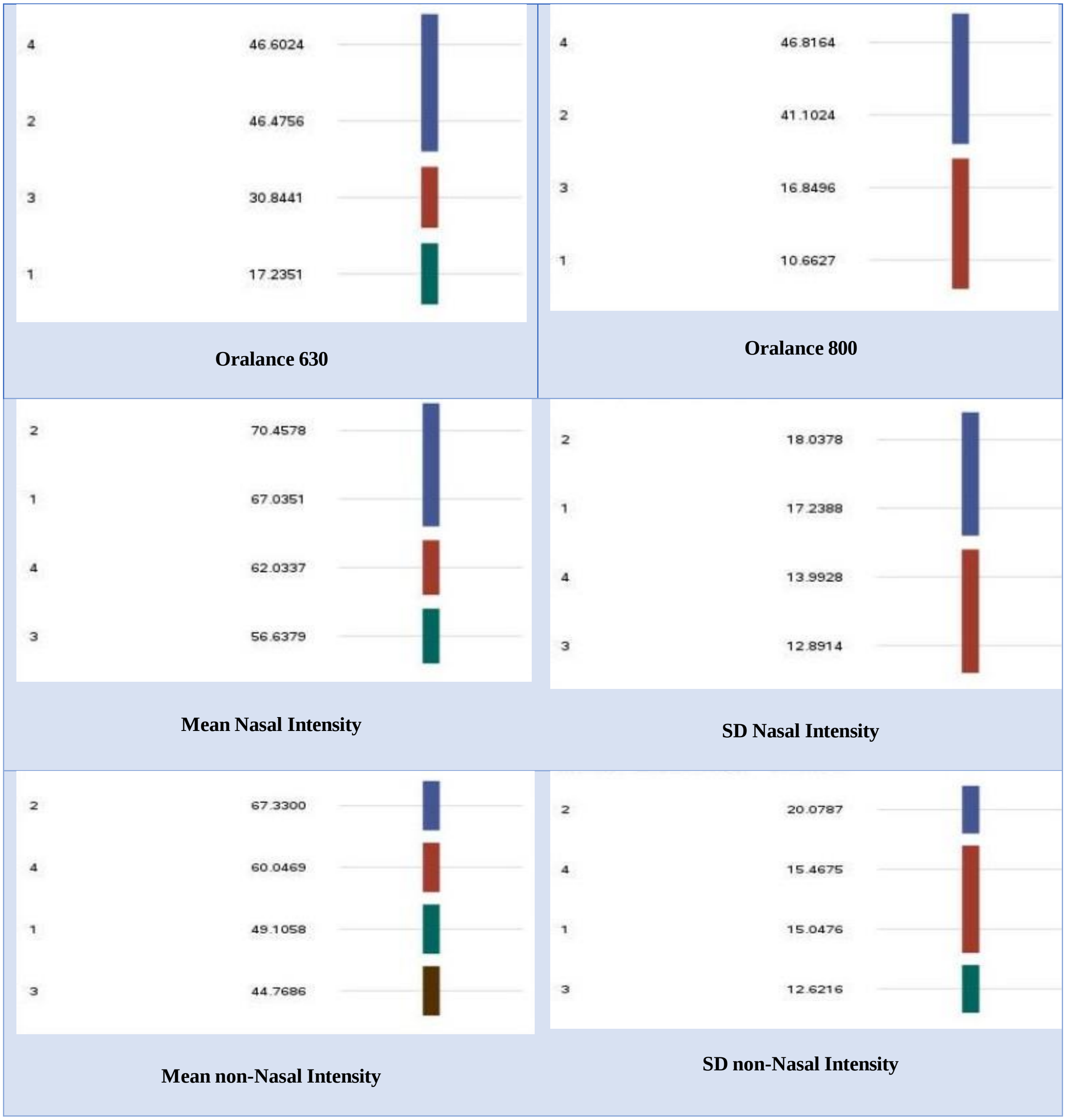
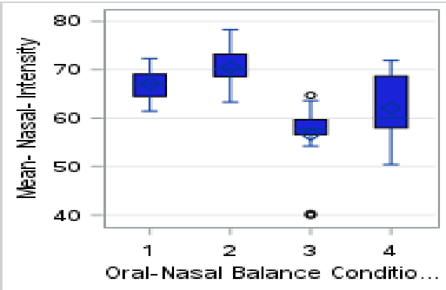
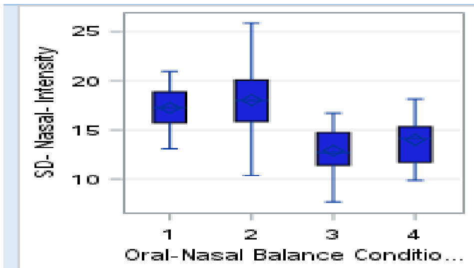


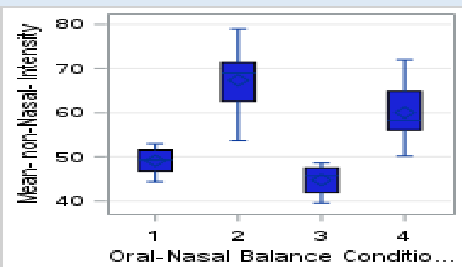
Figure 2.2. Input variables Tukey Grouping for Mean of Oral-Nasal Balance Conditions (1 = Normal, 2 = Hypernasality, 3 = Hyponasality, 4=Mixed nasality) (Means covered by the same bar are not significantly different). The Tukey HSD test adjusts for multiple comparisons, controlling the Type I error rate across all pairwise tests.



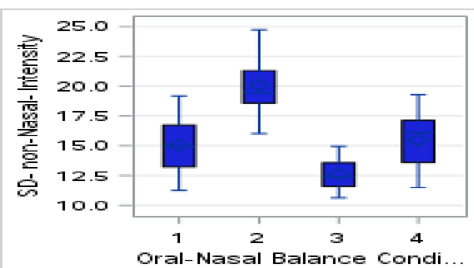
Mean Nasal Intensity distinguishes non-hyponasal groups (normal and hyper) from hyponasal groups (hypo and mixed) with higher intensity.



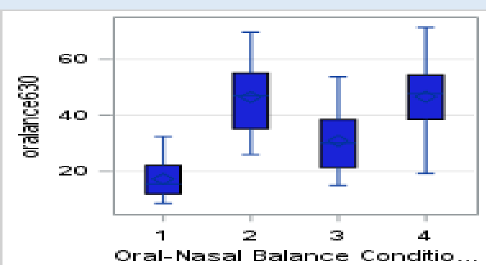
SD Nasal Intensity distinguishes non-hyponasal groups (normal and hyper) with higher variability from hyponasal groups (hypo and mixed).



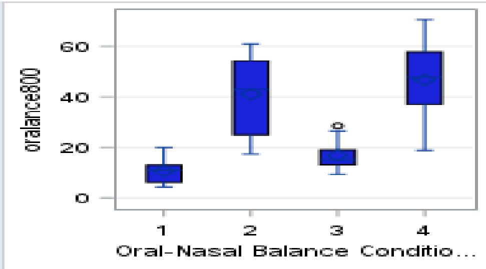
Mean Non-Nasal Intensity distinguishes non-hypernasal groups (normal and hypo) with lower intensity from hypernasal groups (hyper and mixed).



SD Non-Nasal Intensity distinguishes hypernasality with higher variability from other conditions (normal, hypo, and mixed)



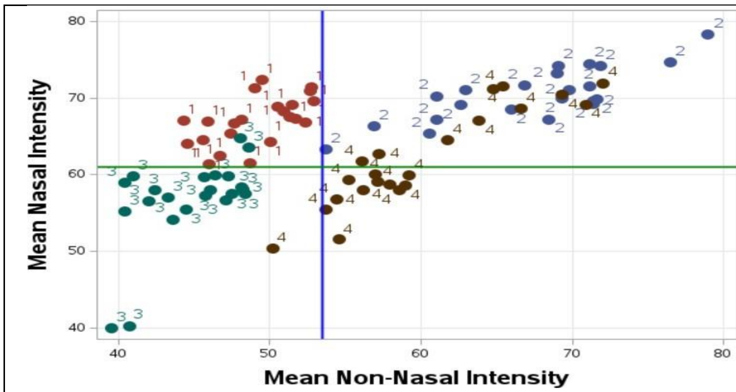
Oralance630 distinguishes normal conditions with lower values from pathological conditions (hyper, hypo, and mixed nasality)



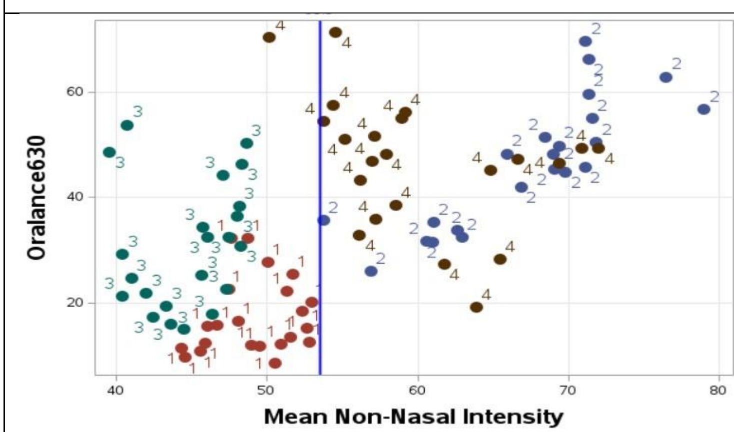
Oralance800 distinguishes non-hypernasal groups (normal and hypo) with lower values from hypernasal groups (hyper and mixed)

Figure 2.3. Boxplots of Variables for Oral-Nasal Balance Conditions (1 = Normal, 2 = Hypernasality, 3 = Hyponasality, 4=Mixed nasality). The horizontal line inside each box represents the median. The lower and upper edges of the box show the 25th and 75th percentiles (interquartile range, IQR). The whiskers (vertical lines extending from each box) indicate values within 1.5 × IQR from the lower and upper quartiles.

Figure 2.3 shows the boxplots for all variables, demonstrating that Mean Nasal Intensity, Mean Non-Nasal Intensity, and Oralance630 have stronger discriminatory power compared with the other variables. As a result, scatterplots were created by pairing these three variables. The scatterplots displayed distinct clustering patterns among the oral-nasal balance conditions (Figure 2.4).



Pairing **Mean Nasal Intensity and Mean Non-Nasal Intensity** effectively distinguishes hypernasality from hyponasality and normal, with some overlap in mixed nasality.



Pairing **Oralance630 and Mean Non-Nasal Intensity** differentiates hypernasal groups (hyper and mixed) from non-hypernasal groups (normal and hypo).

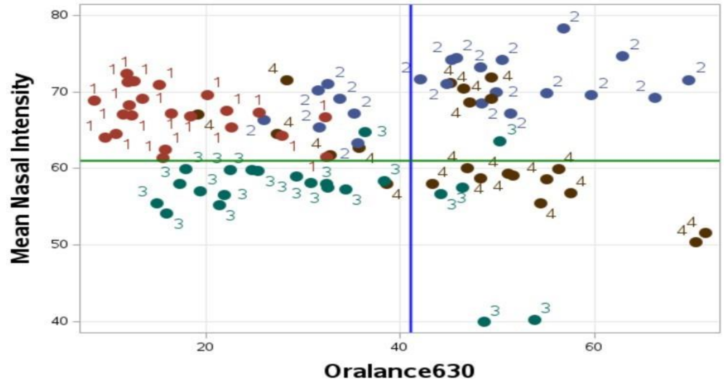
	Pairing **Oralance630 and Mean Nasal Intensity** separates hypernasality from hyponasality, with some overlap in mixed nasality and normal.

Figure 2.4. Scatter Plots of Variable Pairings Separating Oral-Nasal Balance Conditions (1 = Normal, 2 = Hypernasality, 3 = Hyponasality, 4=Mixed nasality)

Linear Discriminant Analysis

To assess the separation of the four oral nasal balance conditions when all variables are combined, a scatterplot matrix was used (Figure 2.5). A scatterplot matrix visualizes pairwise relationships between multiple variables in a dataset, helping to identify patterns and group separations. In this study, the scatterplot matrix was created using LDA-transformed variables, which are new features generated by LDA. These variables, known as linear discriminants, are linear combinations of the original variables that maximize the separation between the four oral-nasal balance conditions. The scatterplot matrix of the LDA Results for Six Variables illustrates the combined discriminatory effectiveness of the LDA-transformed variables, showcasing enhanced differentiation among the four conditions compared to individual or paired original variables.

Conducting an LDA using the final six selected acoustic variables resulted in a classification accuracy of 90.9% when the model was trained and tested on the full dataset (n = 88), consisting of equal numbers of simulated samples for normal resonance, hypernasality, hyponasality, and mixed nasality conditions (22 per class). The error analysis for each group revealed the following findings: Normal resonance detection achieved an accuracy of 90%, correctly identifying 20 out of 22 samples, where two samples were misclassified as hyponasal; hypernasality detection achieved an accuracy of 90%, correctly identifying 20 out of 22 samples, with one sample misclassified as hyponasal and one sample misclassified as mixed nasality; hyponasality detection achieved an accuracy of 95%, correctly identifying 21 out of 22 samples, while one sample was misclassified as normal; mixed nasality detection achieved an accuracy of 86%, correctly identifying 19 out of 22 samples, but three samples were misclassified as hypernasal.

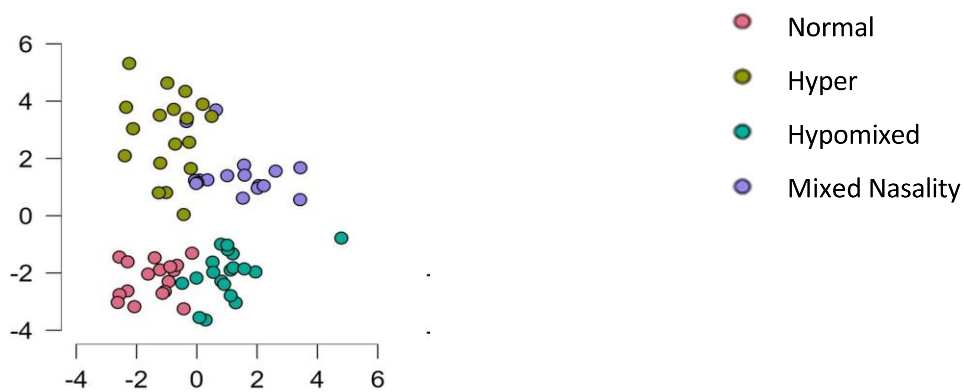


Figure 2.5. scatter plot of the linear discriminant functions (LD1 and LD2) from LDA of the six measured variables. Each point represents one observation projected into the two directions that maximize separation among the four oral-nasal balance conditions. The axes correspond to LD1 and LD2.

To assess the generalizability of the model, an 8-fold stratified cross-validation was conducted. In each fold, 77 samples were used for training and 11 samples for testing, with balanced representation of all four nasality conditions. This procedure was repeated across eight rounds, ensuring that each sample served once in the test set. The classification accuracy for each fold was recorded, and the final model performance was calculated as the mean accuracy across all eight folds, which was 82%.

LDA Functions

The analysis yielded four linear discriminant functions, each presenting a weighted and linear combination of six variables to optimize group separation. These linear discriminant functions can be described as follows:

- Function 1 = -122.50966 + 4.08411 (Mean Nasal Intensity) + 0.20058 (SD Nasal Intensity) -1.59821 (Mean Non-Nasal Intensity) + 2.10698 (SD Non-Nasal Intensity) + 0.87651 (Oralance-630) -0.05134 (Oralance-800)

- Function 2 = -155.64960 + 3.21510 (Mean Nasal Intensity) + 0.19519 (SD Nasal Intensity) -0.30140 (Mean Non-Nasal Intensity) + 3.01600 (SD Non-Nasal Intensity) + 0.88730 (Oralance-630) -0.00614 (Oralance -800)

- Function 3 = -95.80928 + 3.53946 (Mean Nasal Intensity) - 0.11157 (SD Nasal Intensity) -1.28034 (Mean Non-Nasal Intensity) + 1.86425 (SD Non-Nasal Intensity) + 0.92103 (Oralance-630) -0.12048 (Oralance-800)

- Function 4 = -119.77830 + 2.99493 (Mean Nasal Intensity) - 0.01578 (SD Nasal Intensity) -0.34895 (Mean Non-Nasal Intensity) + 2.23566 SD Non-Nasal Intensity) + 0.76010 (Oralance-630) + 0.10556 (Oralance-800)

Discussion

The study conducted an exploratory analysis using linear discriminant analysis to examine acoustic parameters specific to the nasal channel's acoustic energy and determine whether nasal signal acoustic features alone could effectively classify simulated oral-nasal balance conditions. This approach improves upon traditional acoustic methods that used an oral microphone, resulting in combined oral and nasal signals that introduced variability and obscured important features of nasality. The exploratory LDA results for nasal signal acoustic parameters demonstrated a 90.9% classification accuracy in simulated oral-nasal balance conditions, exceeding the findings of previous studies using the same database. De Boer and

Bressmann (2016) used LTAS-based LDA analysis, achieving an accuracy of 80.7% in classifying simulated oral-nasal balance disorders. In another study, De Boer and Bressmann (2015) employed a nasalance-based LDA approach, which correctly classified 88.6% of four simulated oral-nasal balance conditions.

To further evaluate model robustness and generalizability, a cross-validation procedure was subsequently applied, yielding a mean classification accuracy of 82%. The difference between full-dataset accuracy and cross-validated accuracy reflects the expected distinction between model fit and model generalization. Accuracy derived from the full dataset represents an upper-bound estimate of performance, whereas cross-validation provides a more conservative and realistic estimate of expected performance on unseen data (Pothuganti, 2018). Accordingly, full-dataset accuracy is reported to facilitate direct and methodologically consistent comparison with prior studies using the same dataset, which did not report cross-validation results (De Boer & Bressmann, 2015, 2016), while cross-validated accuracy offers a more informative estimate of the practical applicability and generalizability of the proposed classification approach. The high accuracy of the current study demonstrates the potential for using nasal channel acoustic energy to achieve more precise identification of oral-nasal balance disorders. By eliminating interference from oral signal artifacts, the proposed approach ensures a more accurate and specific analysis of nasal resonance patterns. Oral-nasal balance disorders are rooted in changes in nasal energy characteristics; however, traditional recording methods that combine oral and nasal signals often generate noise, potentially masking essential nasal-specific features. By focusing on nasal signals, the current approach minimizes oral articulatory variability and reduces interference, thereby potentially enhancing diagnostic accuracy.

A key objective was the development of a multiparametric diagnostic approach for classifying nasality with nasal signals alone. Through iterative LDA, six key variables were identified as the most accurate and interpretable combination, forming the foundation of a multiparametric tool designed to comprehensively evaluate all main types of oral-nasal balance disorders. The results, as shown by the boxplots and the statistically significant differences observed across conditions, highlight the diagnostic potential of these six acoustic parameters in distinguishing between different oral-nasal balance conditions. However, the overlap observed between conditions underscores the complexity and multidimensional nature of these classifications, as well as the limitations of relying on a single signal for accurate classification. These findings underscore the importance of a multiparametric approach that integrates complementary acoustic features, such as spectral energy distribution and intensity-based measures, to capture variability across nasality conditions and improve diagnostic accuracy more effectively.

Several previous studies have successfully implemented multiparametric approaches to assess nasality. Van Lierde et al. (2007) developed the Nasality Severity Index (NSI), a multidimensional tool that integrates nasalance scores, maximum duration time (MDT), and the Glätzel test, achieving 88% sensitivity and 95% specificity in distinguishing between hypernasality. Bettens et al. (2016) further enhanced this approach with NSI 2.0, incorporating the VLHR and spectral measures, which improved hypernasality assessment with 92% sensitivity and 100% specificity. Bressmann et al. (2000) introduced the nasalance distance and nasalance ratio, which analyzed variations across nasal and non-nasal sentences, refining oral-nasal balance evaluation with 89.6% sensitivity and 94.1% specificity. Golabbakhsh et al. (2017) demonstrated that integrating Mel-Frequency Cepstral Coefficients (MFCC) and Bionic Wavelet

Transform (BWT) energy significantly enhanced accuracy and sensitivity, highlighting the effectiveness of feature fusion in hypernasality detection. Although hypernasality has been extensively studied using multiparametric methods, other oral-nasal balance disorders remain underexplored. Given the potential coexistence of hypernasality and hyponasality (de Boer & Bressmann, 2015; Bressmann et al., 2000), the current study introduces a multiparametric tool that integrates diverse acoustic features to assess all types of oral-nasal balance disorders with high accuracy.

This study also introduced the Oralance formula, designed to ensure a more comprehensive analysis of oral-nasal balance by incorporating values from both oral and nasal stimuli. De Boer and Bressmann (2015) emphasized the importance of collectively analyzing nasalance values for oral and nasal stimuli to address the overlapping features of hypernasality and hyponasality. The "Oralance" formula combines nasal acoustic sound pressure levels (SPL) from both nasal and oral stimuli, which are captured using a nasal microphone. Two acoustic parameters, Oralance630 and Oralance800, were developed by integrating both oral and nasal spectral patterns of frequency bands centred at 630 Hz and 800 Hz into the Oralance formula. In contrast, the traditional nasalance formula evaluates the nasalance score separately for the nasal or oral stimuli. The Oralance formula combines spectral patterns from oral and nasal stimuli to effectively capture variations in oral-nasal balance, providing a more accurate method for calculating nasality.

The developed LDA algorithm demonstrated promising accuracy in classifying simulated oral nasal balance conditions using multiparametric acoustic analyses from the nasal signal. This investigation offers valuable insights into advancing oral-nasal balance assessment, laying the groundwork for practical applications and future research in the field of cleft palate. However, further validation is necessary to assess its effectiveness in real-world clinical settings. In this study, the discriminant functions are presented to illustrate the relative contribution of each acoustic variable to group separation within the simulated dataset and to support transparency of the modeling approach. The coefficients should be interpreted as exploratory and simulation-specific, as they were derived exclusively from simulated signals under controlled conditions. Consequently, the stability and generalizability of these coefficients to clinical speech data cannot be assumed. These functions are therefore not intended as finalized classification equations or clinical decision rules, but rather as a proof-of-concept demonstrating the feasibility of linear discriminant approaches prior to validation using clinical datasets.

Conclusion

The exploratory LDA results demonstrated high accuracy in classifying simulated oral-nasal balance conditions based solely on acoustic energy from the nasal channel. This LDA-based multiparametric model, centred on nasal signal analysis, shows potential as an objective and practical tool to supplement auditory-perceptual assessments, which can be subjective and variable. Additionally, it may offer a more accessible and cost-effective alternative to nasometry, as it requires a separation plate equipped with a nasal microphone, rather than specialized equipment like a nasometer. While our approach aims to improve precision in nasality assessment, further clinical studies are necessary to validate its diagnostic accuracy in individuals with hypernasality, hyponasality, and mixed nasality.

Limitations

The primary limitation of this study was the small sample size used. Larger sample sizes enable the identification of critical features and the distinction of meaningful patterns across groups more effectively. An additional important limitation is the use of simulated speech materials in this study. Although simulated productions allow precise control over acoustic variables and are commonly used in nasality research, they do not fully reflect the variability and complexity of real clinical speech. As a result, findings derived from simulated data may not entirely translate to clinical speech produced by individuals with resonance disorders. Future research should test the acoustic model using larger and more diverse clinical datasets, including different age groups such as children, adults, and older adults, to validate the robustness and clinical applicability of the proposed methods.

Next steps

The next phase of the research will focus on applying the multiparametric acoustic nasal signal analysis method to clinical data. This step aims to assess the method's adaptability and effectiveness in real-world clinical scenarios, ensuring its practical utility in diagnosing oral-nasal balance disorders.

References

Akafi, E., Vali, M., & Moradi, N. (2012). Detection of hypernasal speech in children with cleft palate. *In 2012, the 19th Iranian Conference of Biomedical Engineering (ICBME) (pp. 237–241). IEEE.*

Bettens, K., Bruneel, L., Maryn, Y., De Bodt, M., Luyten, A., & Van Lierde, K. M. (2018). Perceptual evaluation of hypernasality, audible nasal airflow and speech understandability using ordinal and visual analogue scaling and their relation with nasalance scores. *Journal of Communication Disorders, 76, 11–20.*

Bettens, K., De Bodt, M., Maryn, Y., Luyten, A., Wuyts, F. L., & Van Lierde, K. M. (2016). The relationship between the Nasality Severity Index 2.0 and auditory-perceptual judgments of hypernasality. *Journal of Communication Disorders, 62*, 67–81.

Boersma, P., & Weenink, D. (1992–2022). *Praat: Doing phonetics by computer* (Version 6.2.06) [Computer software]. https://www.praat.org

Bressmann, T., Klaiman, P., & Fischbach, S. (2006). Same noses, different nasalance scores: Data from normal subjects and cleft palate speakers for three systems for nasalance analysis. *Clinical Linguistics & Phonetics, 20*(2–3), 163–170. https://doi.org/10.1080/02699200500270689

Bressmann, T. (2005). Comparison of nasalance scores obtained with the Nasometer, the NasalView, and the OroNasal System. *The Cleft Palate-Craniofacial Journal, 42(4), 423–433.*

Bressmann, T., Sader, R., Whitehill, T. L., Awan, S. N., Zeilhofer, H. F., & Horch, H. H. (2000). Nasalance distance and ratio: Two new measures. *The Cleft Palate-Craniofacial Journal, 37(3), 248–256.*

Brunnegård, K., Lohmander, A., & van Doorn, J. (2012). Comparison between perceptual assessments of nasality and nasalance scores. *International Journal of Language & Communication Disorders, 47(5), 556–566.*

Cairns, D. A., Hansen, J. H., & Riski, J. E. (1996). A noninvasive technique for detecting hypernasal speech using a nonlinear operator. *IEEE Transactions on Biomedical Engineering*, 43(1), 35.

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

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

Dalston, R. M., Warren, D. W., & Dalston, E. T. (1991). A preliminary investigation concerning the use of nasometry in identifying patients with hyponasality and/or nasal airway impairment. *Journal of Speech, Language, and Hearing Research*, 34(1), 11-18.

de Boer, G., & Bressmann, T. (2015). Application of linear discriminant analysis to the nasometric assessment of resonance disorders: A pilot study. *Cleft Palate-Craniofacial Journal, 52*(2), 173–182. https://doi.org/10.1597/13-109

de Boer, G., & Bressmann, T. (2016). Application of linear discriminant analysis to the long-term averaged spectra of simulated disorders of oral-nasal balance. *Cleft Palate-Craniofacial Journal, 53(5), e163–e171*. https://doi.org/10.1597/14-236

Dodderi, T., Narra, M., Varghese, S. M., & Deepak, D. T. (2016). Spectral analysis of hypernasality in cleft palate children: A pre-post surgery comparison. *Journal of Clinical and Diagnostic Research, 10(1), MC01–MC03*. https://doi.org/10.7860/JCDR/2016/15389.7055

Dubey, A. K., Prasanna, S. M., & Dandapat, S. (2020). Sinusoidal model-based hypernasality detection in cleft palate speech using CVCV sequence. *Speech Communication, 124, 1-1*2. https://doi.org/10.1016/j.specom.2020.08.001

Dubey, A. K., Mahadeva Prasanna, S. R., & Dandapat, S. (2018). Pitch-adaptive front-end feature for hypernasality detection. In *Proceedings of Interspeech 2018* (pp. 372–376). https://doi.org/10.21437/Interspeech.2018-1251

Dubey, A. K., Tripathi, A., Prasanna, S. R. M., & Dandapat, S. (2018). Detection of hypernasality based on vowel space area. *The Journal of the Acoustical Society of America, 143*(5), EL412–EL417. https://doi.org/10.1121/1.5039718

Fant, G. (1960). *Acoustic theory of speech production*. The Hague, The Netherlands: Mouton & Co.

Glass, J., & Zue, V. (1985). Detection of nasalized vowels in American English. In *ICASSP '85: IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 10, pp. 1569–1572). IEEE.

Golabbakhsh, M., Abnavi, F., Kadkhodaei Elyaderani, M., Derakhshandeh, F., Khanlar, F., Rong, P., & Kuehn, D. P. (2017). Automatic identification of hypernasality in normal and cleft lip and palate patients with acoustic analysis of speech. *The Journal of the Acoustical Society of America, 141(2), 929–935.* https://doi.org/10.1121/1.4976056

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*, 1157–1182.

Haapanen, J., Liu, L., Hiltunen, T., Leinonen, L., & Karhunen, J. (1996). Cul-de-sac hypernasality test with pattern recognition of LPC indices. *Folia Phoniatrica et Logopaedica, 48(1), 35–43.*

Han, K.-S., Kang, M.-S., & Ko, M.-H. (2025). Evaluation of a Newly Developed Device for Resonance Disorders: A Clinical Pilot Study. *Applied Sciences*, *15(10)*, 5713. https://doi.org/10.3390/app15105713

Hawkins, S., & Stevens, K. N. (1985). Acoustic and auditory-perceptual correlates of the non-nasal–nasal distinction for vowels. *The Journal of the Acoustical Society of America, 77(4), 1560–1575.*

He, L., Zhang, J., Liu, Q., Yin, H., & Lech, M. (2014). Automatic evaluation of hypernasality and consonant misarticulation in cleft palate speech. *IEEE Signal Processing Letters*, *21*(10), 1298–1301.

Hixon, T. J., Weismer, G., & Hoit, J. D. (2008). *Preclinical speech science: Anatomy, physiology, acoustics, perception*. Plural Publishing.

Japkowicz, N., & Shah, M. (2011). *Evaluating learning algorithms: A classification perspective*. Cambridge University Press.

John, A., Sell, D., Sweeney, T., Harding-Bell, A., & Williams, A. (2006). The cleft audit protocol for speech—augmented: A validated and reliable measure for auditing cleft speech. *The Cleft Palate-Craniofacial Journal, 43(3), 272–288.*

Jolliffe, I. T. (1986). *Principal component analysis*. New York, NY: Springer-Verlag.

Kataoka, R., Michi, K., Okabe, K., Miura, T., & Yoshida, H. (1996). Spectral properties and quantitative evaluation of hypernasality in vowels. *The Cleft Palate–Craniofacial Journal, 33*(1), 43–50. https://doi.org/10.1597/1545-1569_1996_033_0043_spaqeo_2.3.co_2

Kataoka, R., Warren, D. W., Zajac, D. J., Mayo, R., & Lutz, R. W. (2001). The relationship between spectral characteristics and perceived hypernasality in children. *The Journal of the Acoustical Society of America*, *109*(5), 2181–2189. https://doi.org/10.1121/1.1360717

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, *5(3)*, 7–23.

Kuehn, D. P. (1982). Assessment of resonance disorders. In N. J. Lass, L. V. McReynolds, J. C. Northern, & D. E. Yoder (Eds.), *Speech, language and hearing* (Vol. II: Pathologies of speech and language). W. B. Saunders.

Kummer, A. W. (2011). Disorders of resonance and airflow secondary to cleft palate and/or velopharyngeal dysfunction. *Seminars in Speech and Language*, *32*(2), 141–149. https://doi.org/10.1055/s-0031-1277716

Lee, A., Whitehill, T. L., & Ciocca, V. (2009). Effect of listener training on perceptual judgement of hypernasality. *Clinical Linguistics & Phonetics*, 23(5), 319–334. https://doi.org/10.1080/02699200802688596

Lee, G. S., Wang, C. P., Yang, C. C. H., & Kuo, T. B. J. (2006). Voice low tone to high tone ratio: A potential quantitative index for vowel [a:] and its nasalization. *IEEE Transactions on Biomedical Engineering*, *53*(7), 1437–1439. https://doi.org/10.1109/TBME.2006.873694

Lee, G. S., Yang, C. C., Wang, C. P., & Kuo, T. B. (2005). Effect of nasal decongestion on voice spectrum of a nasal consonant–vowel. *Journal of Voice*, *19*(1), 71–77.

Lee, A. S. Y., Ciocca, V., & Whitehill, T. L. (2003). Acoustic correlates of hypernasality. *Clinical Linguistics & Phonetics*, 17(4–5), 259–264. https://doi.org/10.1080/0269920031000080091

Liu, Y., Lee, S. A. S., & Chen, W. (2022). The correlation between perceptual ratings and nasalance scores in resonance disorders: A systematic review. *Journal of Speech, Language, and Hearing Research*, *65*(6), 2215–2234.

Liu, Y., Wang, X., Hang, Y., He, L., Yin, H., & Liu, C. (2016). Hypernasality detection in cleft palate speech based on natural computation. In *2016, the 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* (pp. 523–528). IEEE.

Lohmander, A., Willadsen, E., Persson, C., Henningsson, G., Bowden, M., & Hutters, B. (2009). Methodology for speech assessment in the Scandcleft project—an international randomized clinical trial on palatal surgery: experiences from a pilot study. *The Cleft Palate-Craniofacial Journal*, 46(4), 347–362.

Lundberg, A. K. (1991). *Laryngeal compensation and linguistic stress in children with velopharyngeal dysfunction* (Presidential Scholars Thesis No. 107). University of Northern Iowa. https://scholarworks.uni.edu/pst/107

Mirzaei, A., & Vali, M. (2016). Detection of hypernasality from speech signal using group delay and wavelet transform. In *2016, the 6th International Conference on Computer and Knowledge Engineering (ICCKE)* (pp. 189–193). IEEE.

Navya, A., & Pushpavathi, M. (2013). One-third octave analysis: A diagnostic tool to measure nasality in conjunction with nasalance in children with repaired cleft lip and palate. *Journal of the All India Institute of Speech & Hearing, 32*.

Nieto, R. G., Marín-Hurtado, J. I., Capacho-Valbuena, L. M., Suarez, A. A., & Bolaños, E. A. B. (2014). Pattern recognition of hypernasality in voice of patients with cleft and lip palate. In *2014 XIX Symposium on Image, Signal Processing and Artificial Vision* (pp. 1–5). IEEE. https://doi.org/10.1109/STSIVA.2014.7010187

Orozco-Arroyave, J. R., Vargas-Bonilla, J. F., Vásquez-Correa, J. C., Castellanos-Domínguez, C. G., & Nöth, E. (2016). Automatic detection of hypernasal speech of children with cleft lip and palate from Spanish vowels and words using classical measures and nonlinear analysis. *Revista Facultad de Ingeniería Universidad de Antioquia, 80*, 109–123.

Paniagua, L. M., Signorini, A. V., da Costa, S. S., Martins Collares, M. V., & Dornelles, S. (2013). Comparison of videonasoendoscopy and auditory-perceptual evaluation of speech in individuals with cleft lip/palate. *International Archives of Otorhinolaryngology, 17*(3), 265–273. https://doi.org/10.7162/S1809-97772013000300006

Pothuganti, S. (2018). Review on over-fitting and under-fitting problems in machine learning and solutions. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 7*(7), 3692–3695.

Pruthi, T., & Espy-Wilson, C. Y. (2004). Acoustic parameters for automatic detection of nasal manner. *Speech Communication, 43*(3), 225–239.

Santoni, C., de Boer, G., Thaut, M., & Bressmann, T. (2020). Influence of voice focus adjustments on oral-nasal balance in speech and song. *Folia Phoniatrica et Logopaedica*, *72*(5), 351–362. https://doi.org/10.1159/000501908

Stevens, K. N. (1985). Spectral prominences and phonetic distinctions in language. *Speech Communication*, *4*(1–3), 137–144.

Tsai, Y. J., Wang, C. P., & Lee, G. S. (2012). Voice low tone to high tone ratio, nasalance, and nasality ratings in connected speech of native Mandarin speakers: A pilot study. *The Cleft Palate-Craniofacial Journal*, *49*(4), 437–446.

Van Lierde, K. M., Wuyts, F. L., Bonte, K., & Van Cauwenberge, P. (2007). The Nasality Severity Index: An objective measure of hypernasality based on a multiparameter approach. *Folia Phoniatrica et Logopaedica*, *59*(1), 31–38. https://doi.org/10.1159/000096548

Vijayalakshmi, P., Reddy, M. R., & O'Shaughnessy, D. (2007). Acoustic analysis and detection of hypernasality using a group delay function. *IEEE Transactions on Biomedical Engineering*, *54*(4), 621–629. https://doi.org/10.1109/TBME.2006.889191

Wang, X., Yang, S., Tang, M., Yin, H., Huang, H., & He, L. (2019). HypernasalityNet: Deep recurrent neural network for automatic hypernasality detection. *International Journal of Medical Informatics*, *129*, 1–12.

Young, L. H., Zajac, D. J., Mayo, R., & Hooper, C. R. (2001). Effects of vowel height and vocal intensity on anticipatory nasal airflow in individuals with normal speech. *Journal of Speech, Language, and Hearing Research*, *44*(1), 52–60.

Chapter 3

Clinical Application of Nasal Signal-Based Acoustic Approach to Enhance Detection and Classification of Oral-Nasal Balance Disorders Related to Cleft Palate: A Retrospective/Secondary Analysis

Abstract

Objective: This study evaluated the clinical applicability and generalizability of a multiparametric acoustic classification approach based on nasal signal recordings for detecting oral-nasal balance disorders in children with cleft palate.

Method: This retrospective study involved secondary analysis of clinical speech recordings from children with cleft palate. Six acoustic parameters, derived exclusively from the nasal signal, were used to classify clinical oral-nasal balance disorders (hypernasality and hyponasality) using linear discriminant analysis (LDA). Three classification models were tested: one trained on simulated data, one on clinical data, and one on a merged dataset combining clinical and simulated data.

Results: The model trained on simulated data achieved high accuracy within its dataset (97%) but demonstrated limited performance (32%) when applied to clinical data. Conversely, a model trained directly on clinical data achieved perfect classification accuracy (100%) within that set but only 63% when applied back to simulated data. A combined dataset approach (simulated + clinical) yielded a model with 71% accuracy, demonstrating improved generalizability across data sources.

Conclusion: Acoustic features derived from nasal signals can effectively classify oral-nasal balance disorders in clinical data. However, cross-application of models across different datasets reveals limitations in their ability to generalize to new datasets. Training models on combined datasets enhances robustness and underscores the importance of developing classification models using diverse and extensive clinical data sources to capture the full spectrum of variability present in oral-nasal balance disorders more accurately and improve generalizability.

Introduction

The source-filter model of speech production emphasizes the crucial role of vocal tract characteristics in identifying oral-nasal balance disorders in individuals with cleft palate (Vijayalakshmi et al., 2007; Rendón et al., 2011). Velopharyngeal dysfunction (VPD), a common condition associated with cleft palate, prevents adequate velopharyngeal closure, resulting in coupling between the oral and nasal cavities during speech. This coupling causes abnormal sound transmission through the vocal tract filter. Any factor that disrupts the normal balance between oral and nasal sound transmission can result in oral-nasal balance disorders including hypernasality (excessive airflow through the nasal passages during speech), hyponasality (reduced airflow through the nasal passages due to blockage or narrowing), and mixed nasality (some of both hyper- and hyponasality) (de Boer & Bressmann, 2015; Santoni et al., 2020). Accurate and functional evaluation of oral-nasal balance disorders is essential for the interdisciplinary team to determine the need for surgery, prosthodontics, or speech therapy (Shprintzen et al., 1979; Marsh & Louis, 2003; Ruscello, 2007; Marsh, 2009). Non-invasive acoustic analysis has proven to be an effective tool for diagnosing vocal tract pathologies, including disorders affecting oral-nasal balance. These assessments typically rely on two primary types of acoustic features: formant-based parameters and spectral parameters (Vijayalakshmi et al., 2007; Golabbakhsh et al., 2017).

Nasal speech is characterized by acoustic energy in the low-frequency range, making formant parameters essential for detecting nasality. Key formant-based features include formant-anti-formant pairs near the first formant (F1) (Dodderi et al., 2016; Vijayalakshmi et al., 2007), a reduction in F1 amplitude (Stevens, 1985, 1997), increased bandwidth of the first and second formants (Fant, 1970), shifts in formant frequencies (Hawkins & Stevens, 1985), and amplitude alterations across formants (Yoshida et al., 2000; Lee et al., 2003). Additional metrics, such as

A1-P1 and A1-P0, quantify nasalization by comparing the amplitudes of the first formant and nasal peaks, providing precise insights into hypernasality (Chen, 1995, 1997).

Spectral analysis has been used to assess hypernasality by examining frequency components that reflect changes caused by nasal-oral coupling due to VPD. Research on hypernasality detection has explored various spectral analysis methods and parameters. Kataoka et al. (1996, 2001) identified specific one-third octave frequency bands linked to hypernasality, noting increased energy in lower bands (630–2000 Hz) and decreased energy in higher bands (3200 Hz) for severe hypernasality. Lee et al. (2003) reported increased amplitudes at 630 Hz and 800 Hz, with decreased amplitudes at 2500 Hz in hypernasal speech. Navya and Pushpavathi (2013) demonstrated high sensitivity and specificity in distinguishing hypernasal children by analyzing frequency regions between 500–2500 Hz. To address the challenge of consistently identifying hypernasal-specific frequency bands, the voice low-tone to high-tone ratio (VLHR) was introduced to quantify the ratio of low-frequency to high-frequency power using a specific cutoff. Lee et al. (2006, 2009) found that VLHR, using a cutoff range of 600–800 Hz, effectively differentiated hypernasal vowels by capturing nasal formants in low frequencies and anti-formants in high frequencies. Tsai et al. (2012) confirmed VLHR as a reliable indicator of nasality in connected English speech.

Acoustic analyses have been used to detect nasality by calculating the correlation between a feature and a selected criterion (Lee et al., 2009; Tsai et al., 2012), setting a threshold (Glass & Zue, 1985; Vijayalakshmi et al., 2007), and applying various machine learning classifiers (He et al., 2014; Bressmann & de Boer, 2015, 2016; Liu et al., 2016; Golabbakhsh et al., 2017). Advancements in the automatic detection of nasality have integrated acoustic analysis and machine learning techniques to improve diagnostic precision and reliability. He et al. (2014)

introduced a two-step classification system that integrates cepstral and spectral features, utilizing a gaussian mixture model (GMM) classifier, which achieved an accuracy of 83% for detecting hypernasality levels. de Boer & Bressmann (2015) used nasometric analysis with LDA to classify oral-nasal balance conditions, achieving 88.6% accuracy in distinguishing normal, hypernasal, hyponasal, and mixed conditions based on nasalance scores. de Boer & Bressmann (2016) extended the application of LDA to the long-term average spectrum (LTAS), achieving 80.7% classification accuracy. Liu et al. (2016) introduced a neural network approach using homomorphic spectral analysis, achieving 80% accuracy in distinguishing hypernasality severity levels. Golabbakhsh et al. (2017) demonstrated that combining bionic wavelet transform (BWT) features with mel-frequency cepstral coefficients (MFCCs) enhanced hypernasality detection, achieving 85% accuracy using support vector machines (SVMs) as the classifier.

All existing studies on acoustic analysis for nasality detection have primarily focused on acoustic features extracted from single-microphone signals. Single microphone systems face challenges in accurately capturing and measuring nasality because they record blended acoustic signals from both the oral and nasal cavities. This overlap makes it difficult to isolate nasal-specific features required for detecting velopharyngeal dysfunction (VPD).

In the previous pilot study, which used simulated data (Chapter 2), a multiparametric method was developed based on exploratory analyses of acoustic features derived solely from nasal signals for the detection and classification of oral-nasal balance disorders. The exploratory LDA results for nasal signal acoustic parameters showed a 90.9% classification accuracy in simulated oral-nasal balance conditions. The observed high accuracy suggests that focusing on nasal channel acoustic energy may aid in the diagnostic identification of oral-nasal balance disorders. While statistical analyses on simulated signals demonstrated promising outcomes in classifying oral-nasal balance conditions, testing them on clinical data is essential to validate their practical application and ensure robustness and generalizability across the complexity and variability of real-world patient populations.

The challenge of applying diagnostic models trained on simulated data to clinical populations was highlighted by Bettens et al. (2019). Building on the work of De Boer and Bressmann (2015), they tested nasalance-based LDA functions—originally developed from simulated data—on speech recordings from 55 Dutch-speaking children with perceptually diagnosed oral-nasal balance disorders. Classification accuracy dropped significantly from 88.6% in simulated data to 56% in clinical data. Although adjustments and re-derived models improved accuracy to 80%, the study underscored the complexity of transferring models from simulated to clinical data. These findings highlight that, while simulated data are valuable for building initial diagnostic frameworks, such models cannot be assumed to generalize effectively to clinical populations without validation. Therefore, it is essential to test diagnostic algorithms on clinical data to assess their robustness, refine their performance, and ensure they account for the variability and complexity of real-world oral-nasal balance disorders.

Building on prior findings based on simulated data, this study tested the clinical applicability of the nasal-signal-based multiparametric approach by assessing its effectiveness in classifying clinical cases of oral-nasal balance disorders and evaluating whether acoustic features from the isolated nasal signal could provide an objective and reliable measure of nasality in a clinical population. Additionally, we evaluated the applicability of the previously identified parameters from our LDA algorithm (using the nasal signal) for the automatic detection and classification of oral-nasal balance disorders in clinical data. We hypothesized that an LDA-based multiparametric acoustic model derived from nasal signal features of simulated data would achieve comparable classification performance in clinical data as reported for nasalance-based LDA approaches in other clinical data (Bettens et al., 2019).

Materials and Methods

Study Design

This study employed a retrospective, observational design using secondary analysis of previously collected simulated and clinical speech data. This design was chosen because no experimental manipulation was required, and the analysis focused on naturally produced speech to develop and evaluate classification approaches.

Participants

Data were collected retrospectively from a database consisting of 12 children with typical oral-nasal balance and 13 patients with oral–nasal balance disorders, between the ages of 7.20 and 11.67 years, who had been diagnosed with non-syndromic unilateral cleft lip and palate (UCLP; Bertucci et al., 2022).

Speech Database and Recording Procedure

Each participant, including clinical patients and typically developing children, was asked to repeat a non-nasal sentence orally ("Look at this book with us - it is a story about a zoo") and a nasal sentence ("Mama made some lemon jam") (Fletcher, 1976). The speech recordings were conducted in a quiet office using a Nasometer 6450 (by Kay Pentax) headset (Figure 3.1). Only the signals captured by the nasal microphone were used to isolate the effects of nasality on the nasal output. Clinical data had been previously classified based on nasalance scores into hypernasality and hyponasality (Bertucci et al., 2022). The simulated speech database used in the first study was also used in this study to develop an LDA model based on three oral-nasal balance conditions—hypernasal, hyponasal, and normal—consistent with the clinical data.

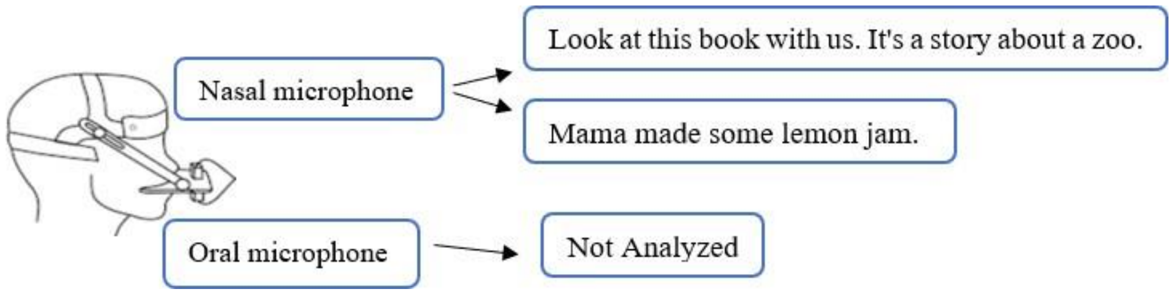


Figure 3.1. Nasometer Headset setup for Collecting Acoustic Data.

Acoustic Analysis

The current study obtained the final set of six variables derived from the first study (Chapter 2) using simulated data. The same variables —Oralance630, Oralance800, mean nasal intensity, SD nasal intensity, mean non-nasal intensity, and SD non-nasal intensity —were extracted from the nasal signals using Praat software scripts (Boersma & Weenink, 1992–2022).

Statistical Analysis

All analyses were conducted using IBM SPSS Statistics version 25.0 (IBM Corp., Armonk, NY, USA). To assess the effects of oral-nasal balance conditions on input variables, an analysis of variance (ANOVA) was conducted. Paired post hoc t-tests were then conducted to compare the mean variables across the oral-nasal balance conditions. To assess the discriminatory power of variables in classifying groups, LDA was applied. To obtain a robust estimate of classification performance and reduce reliance on a single partition, the LDA model was evaluated using 5-fold cross-validation on the merged dataset comprising simulated and clinical samples. This was in keeping with the same method used in the previous study, resulting in an accuracy of 83% for the simulated data. Visual inspection of the data supplemented the interpretation of statistical findings.

Results

Group Differences Across Oral–Nasal Balance Conditions

Repeated-measures ANOVA revealed significant condition differences for all acoustic parameters except SD nasal intensity (Table 3.1).

VARIABLES	CONDITIONS					
	Normal M±SD	Hypernasality M±SD	Hyponasality M±SD	DF	F-Value	P-Value
Oralance 630	37.53±12.17	66.25±6.09	23.67±13.13	2	23.17	<.0001
Oralance 800	38.80±20.80	56.45±11.58	29.32±17.86	2	3.31	0.0552
Mean Nasal Intensity	54.18±2.53	62.62±5.08	49.83±5.41	2	14.66	<.0001
SD Nasal Intensity	9.97±1.89	10.28±2.33	11.31±2.21	2	1.00	0.3830
Mean non-Nasal Intensity	45.44±2.29	55.41±3.36	39.81±3.06	2	48.74	<.0001
SD non-Nasal Intensity	8.86±1.38	10.07±2.65	7.61±1.36	2	3.39	0.0522

Table 3.1. Analysis of Variance (ANOVA) Results for three oral-nasal balance conditions.

Post-hoc analyses show that Oralance 630 and Mean Non-Nasal Intensity provide the strongest separation of the oral–nasal balance conditions. Mean Nasal Intensity demonstrates only moderate discriminatory ability, while Oralance 800, SD Non-Nasal Intensity, and especially SD Nasal Intensity show weak separation with considerable overlap (Figure 3.2).

Figure 3.3, represented as box plots, offers a visualization of the distribution and variability for each of the six variables across the four conditions. The boxplots indicate that mean nasal intensity, mean non-nasal intensity, and oralance630 exhibited greater discriminatory power compared to other variables.

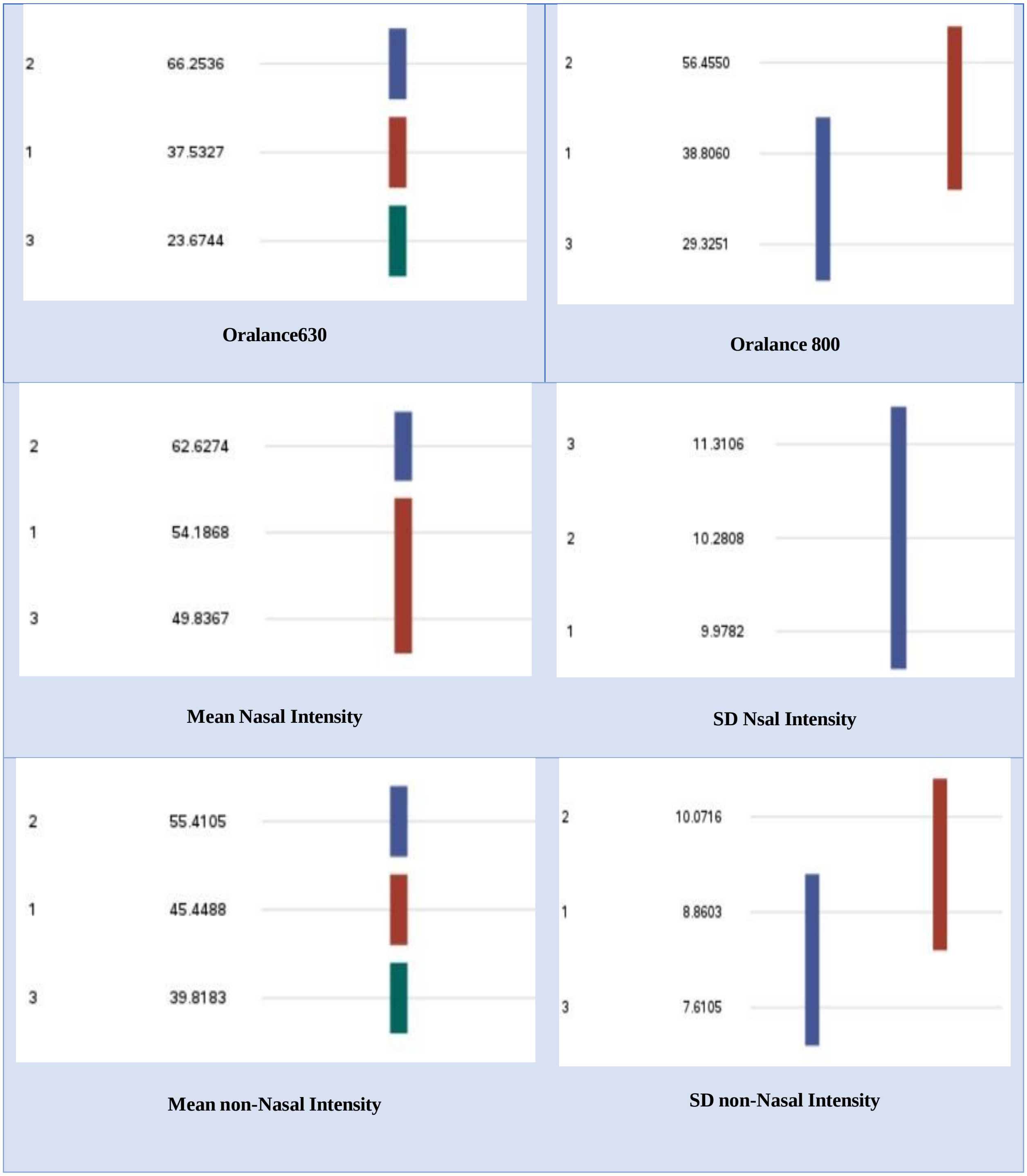


Figure 3.2. Input variables Tukey Grouping for Mean of Oral-Nasal Balance Conditions (1 = Normal, 2 = Hypernasality, 3 = Hyponasality) (Means covered by the same bar are not significantly different). The Tukey HSD test adjusts for multiple comparisons, controlling the Type I error rate across all pairwise tests.

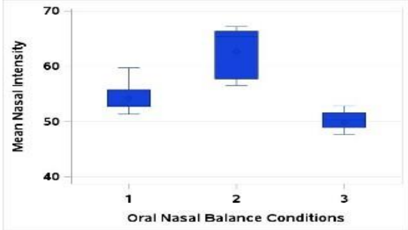
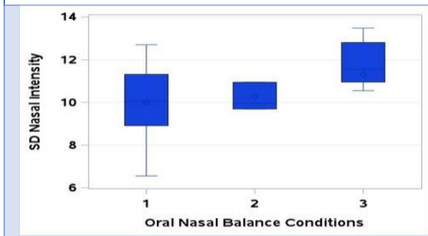
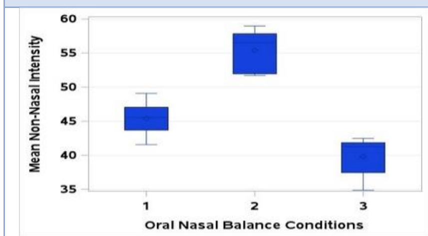
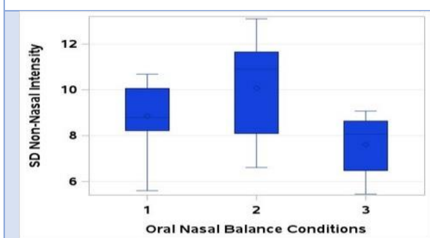
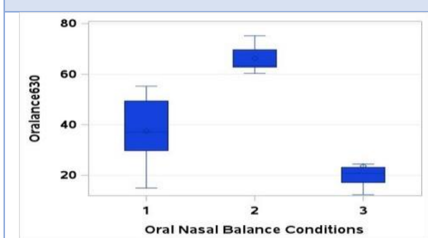
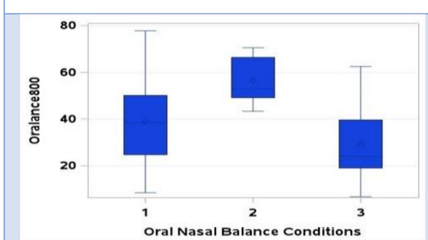
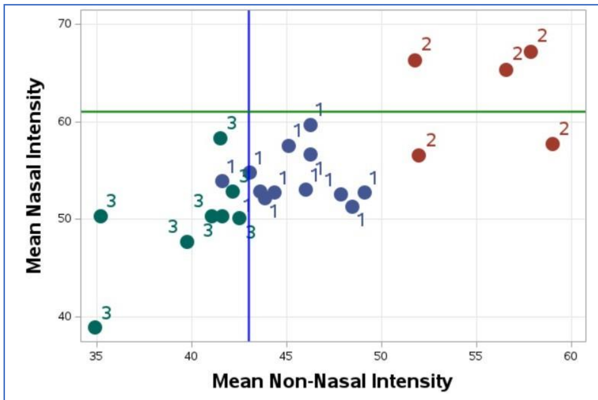
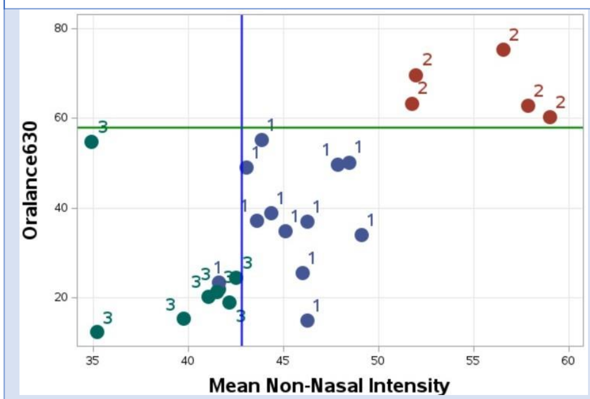
<table>
<tr><td>
</td><td>Mean Nasal Intensity demonstrates complete differentiation between hypernasality (box 2) and hyponasality (box 3).</td></tr>
<tr><td>
</td><td>SD Nasal Intensity shows overlapping variability across normal, hyper, and hypo conditions (boxes 1, 2, and 3).</td></tr>
<tr><td>
</td><td>Mean Non-Nasal Intensity demonstrates complete differentiation between hypernasality (box 2) and non-hypernasal groups (normal and hypo) (boxes 1 and 3).</td></tr>
<tr><td>
</td><td>SD non-Nasal Intensity shows significant overlapping variability across normal, hyper, and hypo conditions (boxes 1, 2, and 3).</td></tr>
<tr><td>
</td><td>Oralance630 demonstrates complete differentiation between hypernasality (box 2) and non-hypernasal groups (normal and hypo) (boxes 1 and 2).</td></tr>
<tr><td>
</td><td>Oralance800 partially distinguishes hypernasality (box 2) from hyponasality (box 3).</td></tr>
</table>

Figure 3.3. Boxplots of Variables for Clinical Conditions (1 = Normal, 2 = Hypernasality, 3 = Hyponasality). The ho) .ntal line inside each box represents the median. The lower and upper edges of the box show the 25th and 75th percentiles (interquartile range, IQR). The whiskers (vertical lines extending from each box) indicate values within 1.5 × IQR from the lower and upper quartiles; they do **not** represent 95 % confidence intervals.

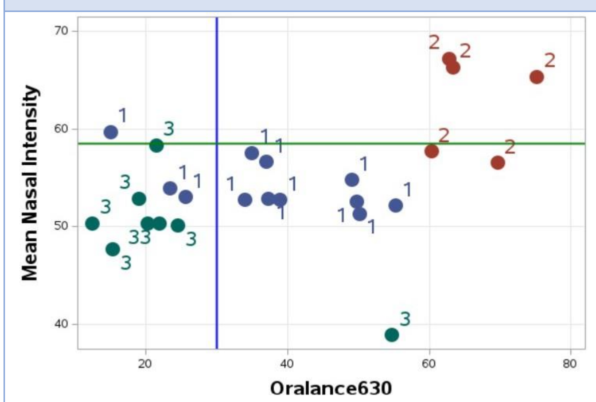
Based on the results of the boxplots, scatterplots were generated by pairing the variables with the strongest discriminatory power (mean nasal intensity, mean non-nasal intensity, and Oralance630), which revealed distinct clustering patterns across the oral-nasal balance conditions (Figure 3.4).



Pairing Mean Nasal Intensity and Mean Non-Nasal Intensity effectively distinguishes hypernasality from hyponasality, while the normal group falls in between with some overlap.



Pairing **Oralance630** and **Mean Non-Nasal Intensity** reveals a strong distinction between hypernasality and the other two groups, with some overlap between normal and hyponasal groups.



Pairing **Mean Nasal Intensity with Oralance630** distinguishes hypernasality from hyponasality, while the normal group falls in between with some overlap.

Figure 3.4. Scatter Plots of Variable Pairings Separating Clinical Conditions (1 = Normal, 2 = Hypernasality, 3 = Hyponasality).

Linear Discriminant Analysis

Applying the simulated data-based LDA formula to clinical data. In study 1, the LDA algorithm was developed using simulated data that included four oral–nasal balance conditions: hypernasality, hyponasality, mixed nasality, and normal speech. In contrast, the clinical data in the current study comprised only three conditions: hypernasality, hyponasality, and normal speech. Consequently, the original LDA formulas from Study 1 could not be directly applied, as the mixed nasality condition was not present in the clinical sample. To ensure consistency between datasets, the mixed nasality condition was removed from the simulated dataset, and the LDA formulas were recalculated using the same six nasal-signal–based classification variables: Mean Nasal Intensity, SD Nasal Intensity, Mean Non-Nasal Intensity, SD Non-Nasal Intensity, Oralance-630, and Oralance-800. This first LDA rendering achieved 97% accuracy in classifying hypernasality, hyponasality, and normal conditions within the simulated dataset (Table 3.2). However, when the customized LDA formula (based on three conditions from the simulated sample) was applied to the clinical dataset, the classification accuracy decreased to 32%. These latter LDA classification results revealed that all samples were assigned to the hyponasal group, although they were classified as normal or hypernasal based on nasalance scores (Table 3.3).

Oral-Nasal Balance Conditions	1	2	3	Total
1	21	0	1	22
	95.45	0.00	4.55	100.00
2	1	21	0	22
	4.55	95.45	0.00	100.00
3	0	0	22	22
	0.00	0.00	100.00	100.00

Table 3.2. LDA Classification Results for the Three Simulated Conditions (1= Normal, 2=Hypernasality, 3= Hyponasality) based on Sample Specific Formulas.

Oral-Nasal Balance Conditions	1	2	3	Total
1	0	0	12	12
	0.00	0.00	100.00	100.00
2	0	0	5	5
	0.00	0.00	100.00	100.00
3	0	0	8	8
	0.00	0.00	100.00	100.00

Table 3.3. LDA Classification Results for Applying Simulated Data-Derived LDA Formula to Clinical Data (1= Normal, 2=Hypernasality, 3= Hyponasality)

Rebuilding the LDA formula using clinical data. The LDA six-variable model was re-run on the clinical dataset and re-trained and tested on the entire dataset. Upon re-running the analysis, the model demonstrated 100% accuracy in classifying clinical cases of hypernasality, hyponasality, and normal nasality (Table 3.4).

Oral-Nasal Balance Conditions	1	2	3	Total
1	12	0	0	12
	100.00	0.00	0.00	100.00
2	0	5	0	5
	0.00	100.00	0.00	100.00
3	0	0	8	8
	0.00	0.00	100.00	100.00

Table 3.4. LDA Classification in the Clinical Data (1= Normal, 2=Hypernasality, 3= Hyponasality)

Applying the clinical data-based LDA formula to simulated data. Applying the LDA formula derived from the clinical dataset to the simulated dataset without mixed nasality yielded a 63% accuracy rate. The results of LDA classification demonstrated different levels of accuracy across different conditions (Table 3.5).

Oral-Nasal Balance Conditions	1	2	3	Total
1	16	0	6	22
	72.73	0.00	27.27	100.00
2	4	18	0	22
	18.18	81.82	0.00	100.00
3	13	1	8	22
	59.09	4.55	36.36	100.00

Table 3.5. LDA Classification Results for Applying Clinical Data-Derived LDA Formula to Simulated Data (1= Normal, 2=Hypernasality, 3= Hyponasality)

Combining simulated and clinical data for LDA. To address inconsistencies and enhance classification performance, the simulated dataset (66 samples) and clinical dataset (25 samples) were merged, creating a more comprehensive dataset of 91 samples that captured a wider range of conditions. Like in the previous study, a more robust estimate of classification performance for model evaluation was conducted using 5-fold cross-validation on the combined dataset (N = 91). The dataset was randomized and partitioned into five approximately equal folds. In each iteration, one fold served as the test set while the remaining four folds were used for training, ensuring that all samples contributed to both training and testing across folds. The mean classification accuracy across folds was 71%.

Discussion

The current study aimed to evaluate the clinical applicability of the previously developed LDA algorithm, derived initially from simulated data using a multiparametric acoustic approach based solely on nasal signals, for detecting and classifying hypernasality and hyponasality in children with cleft palate and in typically developing children. The findings support our primary hypothesis that the developed multiparametric acoustic approach, based solely on nasal signal recordings, can effectively classify oral–nasal balance disorders in clinical data, consistent with patterns observed in simulated data. As the nasal signal directly reflects airflow dynamics and oral-nasal balance patterns, it may serve as a robust source for extracting classification-relevant acoustic features. Consistent with the simulated dataset, mean nasal intensity, mean non-nasal intensity, and oralance630 demonstrated stronger discriminatory power in distinguishing hypernasality, hyponasality, and normal resonance. Although some overlap was observed in certain variables, particularly standard deviation measures, this underscores the importance of combining multiple acoustic measures. These observations are further supported by scatterplots that visually demonstrate separation between resonance conditions based on paired acoustic features (Van Lierde et al., 2007; Bettens et al., 2016).

However, challenges arose when applying LDA formulas across simulated and clinical datasets, resulting in a decline in classification accuracy. Running LDA on the modified simulated dataset (excluding the mixed nasality group to align with the clinical dataset) achieved a high classification accuracy of 97% for the three simulated conditions: hypernasality, hyponasality, and normal. However, applying this customized LDA formula to the clinical dataset considerably reduced classification accuracy to 32%. This result aligns with the findings of Bettens et al. (2019), who demonstrated that LDA models trained on simulated data underperform when applied to clinical datasets, revealing a substantial mismatch between simulated and clinical datasets.

In the current study, reconstructing the LDA model using clinical data significantly improved classification accuracy to 100%, consistent with findings by Bettens et al. (2019), who reported enhanced accuracy (up to 80%) after developing new LDA functions based on clinical speech data. However, when the clinically trained model was applied to the simulated dataset, its performance declined to 63%. This performance gap underscores the inherent differences between controlled, simulated data and real-world clinical data. Simulated data provides controlled, exaggerated representations of oral-nasal balance disorders, making it useful for hypothesis testing. However, it may lack the variability and subtle overlaps found in real-world speech, potentially oversimplifying the spectrum of oral-nasal balance conditions. In contrast, clinical data captures the diverse, complex, and multidimensional nature of oral-nasal balance disorders, offering a more accurate reflection of diagnostic variability. Training the LDA model on clinical data improved its generalizability, as it achieved a moderate accuracy of 63% when applied to the simulated dataset, whereas the simulated-data-trained model performed poorly on clinical data, with an accuracy of only 32%. This contrast suggests that models trained on clinical data generalize more effectively, although they still face structural differences when applied to simulated datasets. While simulated data is valuable for initial model development and controlled experimentation, clinical data remains essential for building classification models that are robust, generalizable, and applicable to real-world diagnostic settings.

To enhance generalizability, the LDA model was evaluated using a 5-fold cross-validation framework applied to the combined dataset of simulated and clinical speech samples. The resulting mean cross-validated classification accuracy of 71% reflects the expected estimated performance on unseen data for a newly trained model specific to the entire dataset. Merging the datasets improved the robustness of the LDA model by incorporating a broader range of acoustic variability, including both controlled conditions from simulated data and natural variability from real-world clinical speech. This integration also reduced overfitting to either the simulated or clinical dataset alone, enabling the classifier to learn patterns more representative of real-world diagnostic variability. The results highlight the importance of dataset characteristics, particularly sample size and diversity, in optimizing model performance. A larger dataset with a balanced representation of various age groups (e.g., children, adults, elderly) and a spectrum of severity levels (mild to severe) is crucial for improving both classification accuracy and the model's capacity to generalize to heterogeneous clinical populations. Expanding the dataset would not only minimize overfitting but also capture the complex variability of oral-nasal balance conditions, thereby improving the reliability and applicability of classification models in clinical practice.

While merging simulated and clinical datasets contributed to improved generalizability and helped mitigate some overfitting, concerns remain regarding the evaluation of model performance. Specifically, the absence of cross-validation raises the possibility that classification accuracy may have been overestimated. In our study, for example, the LDA model trained solely on clinical data achieved 100% accuracy; however, this result was based solely on the training set, without an independent validation procedure. Similarly, the LDA model trained on simulated data yielded 97% accuracy but dropped to 32% when tested on clinical samples, particularly those from children, indicating limited generalizability across subject populations. Overfitting occurs when a model becomes overly tailored to the training data, limiting its ability to perform well on prospective cases (Pothuganti, 2018). This issue is especially critical in clinical contexts, where diagnostic tools must be robust and reliable across diverse patient groups. In this study, the small clinical sample size made cross-validation impractical, as removing even a few samples would leave too few cases for stable training and could result in highly variable performance estimates. Moreover, the LDA analysis was primarily exploratory, aiming to assess whether the selected acoustic features could distinguish between classes rather than to generate a fully validated clinical classifier.

Conclusion

The successful classification of oral–nasal balance disorders using the previously developed multiparametric approach based solely on nasal signal analysis, integrated with LDA, demonstrates the potential of nasal-derived acoustic features as reliable diagnostic markers. However, the reduced performance observed in the cross-application of the LDA model between simulated and clinical datasets underscores the need for improved generalizability. Future development should focus on strategies such as robust cross-validation, balanced and diverse training samples, and model refinement to mitigate overfitting and ensure consistent diagnostic performance. These advancements are crucial for translating acoustic-based classification models into practical and accurate tools for the clinical assessment of resonance disorders.

Limitations

This study's limitations included a small sample size and limited dataset diversity, which hindered the model's ability to capture the full variability of oral-nasal balance disorders. The lack of strategies to enhance model generalizability and the reliance on a limited dataset restricted the model’s performance across different populations.

Next Steps

Future research should focus on using larger and more diverse datasets to better reflect the complexities of real-world clinical scenarios. Implementing cross-validation techniques will help validate the model’s performance, while breaking down the diagnosis into separate functions for specific oral-nasal balance disorders could enhance accuracy and flexibility, leading to more robust and reliable diagnostic tools for oral-nasal balance disorders

References

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistical Surveys, 4,* 40–79. https://doi.org/10.1214/09-SS054

Bertucci, V., Stevens, K., Sidhu, N., Suri, S., & Bressmann, T. (2022). The impact of fan-type rapid palatal expanders on speech in patients with unilateral cleft lip and palate. *The Cleft Palate–Craniofacial Journal*, *59*(9), 1070–1080. https://doi.org/10.1177/10556656221084541

Bettens, K., de Boer, G., Bressmann, T., Bruneel, L., & van Lierde, K. (2019). Clinical application of a new approach to identify oral–nasal balance disorders based on nasalance scores. *The Cleft Palate–Craniofacial Journal*, *56*(5), 628–638. https://doi.org/10.1177/1055665618796012

Bettens, K., De Bodt, M., Maryn, Y., Luyten, A., Wuyts, F. L., & Van Lierde, K. M. (2016). The relationship between the Nasality Severity Index 2.0 and auditory-perceptual judgments of hypernasality. *Journal of Communication Disorders*, *62*, 67–81. https://doi.org/10.1016/j.jcomdis.2016.06.002

Boersma, P., & Weenink, D. (1992–2022). *Praat: Doing phonetics by computer* (Version 6.2.06) [Computer software]. Retrieved January 23, 2024, from https://www.praat.org

Chen, M. Y. (1995). Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *The Journal of the Acoustical Society of America*, *98*(5), 2443–2453. https://doi.org/10.1121/1.413277

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, *102*(4), 2360–2370. https://doi.org/10.1121/1.419642

de Boer, G., & Bressmann, T. (2015). Application of linear discriminant analysis to the nasometric assessment of resonance disorders: A pilot study. *The Cleft Palate–Craniofacial Journal*, *52*(2), 173–182. https://doi.org/10.1597/13-109 de Boer, G., & Bressmann, T. (2016). Application of linear discriminant analysis to the long-term averaged spectra of simulated disorders of oral–nasal balance. *The Cleft Palate–Craniofacial Journal*, *53*(5), e163–e171. https://doi.org/10.1597/14-236

Dodderi, T., Narra, M., Varghese, S. M., & Deepak, D. T. (2016). Spectral analysis of hypernasality in cleft palate children: A pre-post surgery comparison. *Journal of Clinical and Diagnostic Research*, *10*(1), MC01–MC03. https://doi.org/10.7860/JCDR/2016/15389.7055

Fant, G. (1970). *Acoustic theory of speech production* (No. 2). Walter de Gruyter.

Fletcher, S. G. (1976). Nasalance vs. Listener Judgements of Nasality. *The Cleft Palate Journal*, *13(1)*, 31-44.

Glass, J., & Zue, V. (1985, April). Detection of nasalized vowels in American English. In *ICASSP'85: IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 10, pp. 1569–1572). IEEE. https://doi.org/10.1109/ICASSP.1985.1168562

Golabbakhsh, M., Abnavi, F., Kadkhodaei Elyaderani, M., Derakhshandeh, F., Khanlar, F., Rong, P., & Kuehn, D. P. (2017). Automatic identification of hypernasality in normal and cleft lip and palate patients with acoustic analysis of speech. *The Journal of the Acoustical Society of America*, *141*(2), 929–935. https://doi.org/10.1121/1.4976056

Hawkins, S., & Stevens, K. N. (1985). Acoustic and auditory-perceptual correlates of the non-nasal–nasal distinction for vowels. *The Journal of the Acoustical Society of America*, *77*(4), 1560–1575. https://doi.org/10.1121/1.391933

He, L., Zhang, J., Liu, Q., Yin, H., & Lech, M. (2014). Automatic evaluation of hypernasality and consonant misarticulation in cleft palate speech. *IEEE Signal Processing Letters*, *21*(10), 1298–1301. https://doi.org/10.1109/LSP.2014.2328133

Kataoka, R., Michi, K. I., Okabe, K., Miura, T., & Yoshida, H. (1996). Spectral properties and quantitative evaluation of hypernasality in vowels. *The Cleft Palate–Craniofacial Journal*, *33*(1), 43–50. https://doi.org/10.1597/1545-1569_1996_033_0043_spaqeo_2.3.co_2

Kataoka, R., Warren, D. W., Zajac, D. J., Mayo, R., & Lutz, R. W. (2001). The relationship between spectral characteristics and perceived hypernasality in children. *The Journal of the Acoustical Society of America*, *109*(5), 2181–2189. https://doi.org/10.1121/1.1360717

Lee, A. S. Y., Ciocca, V., & Whitehill, T. L. (2003). Acoustic correlates of hypernasality. *Clinical Linguistics and Phonetics*, *17*(4–5), 259–264. https://doi.org/10.1080/0269920031000080091

Lee, G. S., Wang, C. P., Yang, C. C. H., & Kuo, T. B. J. (2006). Voice low tone to high tone ratio: A potential quantitative index for vowel [a:] and its nasalization. *IEEE Transactions on Biomedical Engineering*, *53*(7), 1437–1439. https://doi.org/10.1109/TBME.2006.873694

Lee, G. S., Wang, C. P., & Fu, S. (2009). Evaluation of hypernasality in vowels using voice low tone to high tone ratio. *The Cleft Palate–Craniofacial Journal*, *46*(1), 47–52. https://doi.org/10.1597/07-184.1

Liu, Y., Wang, X., Hang, Y., He, L., Yin, H., & Liu, C. (2016). Hypernasality detection in cleft palate speech based on natural computation. In *2016 12th IEEE International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC–FSKD)* (pp. 523–528). IEEE. https://doi.org/10.1109/FSKD.2016.7603233

Marsh, J. L., & Louis, S. (2003). Management of velopharyngeal dysfunction: Differential diagnosis for differential management. *Journal of Craniofacial Surgery*, *14*(5), 621–628. https://doi.org/10.1097/00001665-200309000-00002

Marsh, J. L. (2009). Velo-pharyngeal dysfunction: Evaluation and management. *Indian Journal of Plastic Surgery*, *42*(Suppl. 1), S144–S154. https://doi.org/10.4103/0970-0358.57201

Navya, A., & Pushpavathi, M. (2013). One-third octave analysis: A diagnostic tool to measure nasality in conjunction with nasalance in children with repaired cleft lip and palate. *Journal of the All India Institute of Speech & Hearing, 32,* 54–62.

Pothuganti, S. (2018). Review on over-fitting and under-fitting problems in machine learning and solutions. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 7*(7), 3692–3695.

Qian, J., Fu, F., Liu, X., He, L., Yin, H., & Zhang, H. (2017). The analysis and detection of hypernasality based on a formant extraction algorithm. *Journal of Physics: Conference Series, 887*(1), 012082. IOP Publishing. https://doi.org/10.1088/1742-6596/887/1/012082

Rendón, S. M., Orozco Arroyave, J. R., Vargas Bonilla, J. F., Arias Londoño, J. D., & Castellanos Domínguez, C. G. (2011). Automatic detection of hypernasality in children. In J. Ferrández Vicente, J. R. Álvarez Sánchez, F. de la Paz López, F. J. Toledo Moreo, & H. Adeli (Eds.), *New challenges on bioinspired applications: 4th International Work-conference on the Interplay Between Natural and Artificial Computation, IWINAC 2011* (pp. 167–174). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21326-7_21

Ruscello, D. (2007). Treatment of velopharyngeal closure for speech: Discussion and implications for management. *The Journal of Speech and Language Pathology – Applied Behavior Analysis, 2*(1), 55–75. https://doi.org/10.1037/h0100212

Santoni, C., de Boer, G., Thaut, M., & Bressmann, T. (2020). Influence of voice focus adjustments on oral–nasal balance in speech and song. *Folia Phoniatrica et Logopaedica, 72*(5), 351–362. https://doi.org/10.1159/000501908

Shprintzen, R. J., Lewin, M. L., Croft, C. B., Daniller, A. I., Argamaso, R. V., Ship, A. G., & Strauch, B. (1979). A comprehensive study of pharyngeal flap surgery: Tailor-made flaps. *The Cleft Palate Journal, 16*(1), 46–55.

Tsai, Y. J., Wang, C. P., & Lee, G. S. (2012). Voice low tone to high tone ratio, nasalance, and nasality ratings in connected speech of native Mandarin speakers: A pilot study. *The Cleft Palate–Craniofacial Journal*, *49*(4), 437–446. https://doi.org/10.1597/10-162

Van Lierde, K. M., Wuyts, F. L., Bonte, K., & Van Cauwenberge, P. (2007). The Nasality Severity Index: An objective measure of hypernasality based on a multiparameter approach. *Folia Phoniatrica et Logopaedica*, *59*(1), 31–38. https://doi.org/10.1159/000096548

Vijayalakshmi, P., Reddy, M. R., & O'Shaughnessy, D. (2007). Acoustic analysis and detection of hypernasality using a group delay function. *IEEE Transactions on Biomedical Engineering*, *54*(4), 621–629. https://doi.org/10.1109/TBME.2006.889191

Yoshida, H., Furuya, Y., Shimodaira, K., Kanazawa, T., Kataoka, R., & Takahashi, K. (2000). Spectral characteristics of hypernasality in maxillectomy patients. *Journal of Oral Rehabilitation*, *27*, 723–730. https://doi.org/10.1046/j.1365-2842.2000.00561.x

Chapter 4

Improving the Assessment of Oral-Nasal Balance Disorders in Patients with Cleft Palate: A Pilot Study on Classifying Oral-Nasal Balance Disorders through Visual Feature Analysis in Stereo Signals

Abstract

Objective. To explore the feasibility of using visual feature analysis of paired oscillogram images for classifying oral-nasal balance disorders by licensed speech-language pathologists (SLPs) and SLP students.

Method. A two-phase study was conducted. In the first phase, the concept of visual analysis was outlined by developing two decision-making flowcharts and instructional guidelines for classifying oral-nasal balance conditions based on paired oscillogram patterns. In the second phase, a pilot study was conducted using an online questionnaire, which included a training module with instructional examples, a practice phase with immediate feedback, and an experimental classification task involving 20 paired oscillogram images representing simulated normal, hypernasal, hyponasal, and mixed nasality conditions. Classification accuracy was assessed descriptively.

Results. A total of 36 raters took part in the visual analysis study, including SLP students with and without training in cleft palate speech and licensed SLPs. Using visual feature analysis of paired oscillogram images, raters achieved an overall classification accuracy of 85.15% in identifying simulated oral-nasal balance conditions. Accuracy was highest among students who had completed the cleft lip and palate course (87.5%), followed by licensed SLPs (85.0%) and students training in cleft palate speech (82.94%). Hyponasality was classified most accurately

(100% by trained students and licensed SLPs), while mixed nasality showed the lowest accuracy, particularly among untrained students (61%).

Conclusion. Visual analysis of paired oscillogram images demonstrated promise as a low-tech, trainable, and accessible tool to support the clinical assessment of oral-nasal balance disorders.

Introduction

Speech-language pathologists (SLPs) commonly use auditory-perceptual analysis to assess oral-nasal balance disorders. This evaluation method plays a crucial role in guiding the therapeutic team's decisions regarding the need for additional assessments and the selection of suitable interventions. (Kummer & Lee, 1996; Kuehn & Moller, 2000; Bettens et al., 2016). However, due to the variability in experience and skill levels among SLPs, a gap remains concerning the consistency and reliability of auditory-perceptual analysis (Lee, Potts, & Bressmann, 2019). Therefore, exploring innovative approaches to improve reliability is crucial. A variety of instrumental methods are already available to assess nasality. SLPs can use indirect assessment techniques, such as acoustic measurement (nasometry and spectral analysis) and aerodynamic measurement, to complement auditory-perceptual findings. Aerodynamic tests are rarely used in clinical settings due to their relatively high costs, technical complexity, and the need for specialized equipment (Karnell, 2011; Kuehn & Moller, 2000). In current clinical practice, nasometry is the most frequently used acoustic measure of nasality; however, it is an expensive technology that may not be available to all clinicians (Han, Kang, & Ko, 2025). As an alternative, spectral analysis has been applied to identify hypernasality. Spectral analysis of speech signal has dramatically improved the understanding of nasality’s acoustic features, but faces several practical challenges. One major issue is the difficulty in drawing definitive clinical conclusions from existing research on the spectral analysis of speech signals (Bressmann &

Abnavi, 2025). To enhance generalizability, these techniques need to be replicated and tested on more diverse patient groups and speech samples. Furthermore, while most studies focus on hypernasality, spectral patterns for other oral-nasal balance disorders remain poorly defined. These challenges emphasize the need to explore simple, accurate, and cost-effective methods for classifying all types of oral-nasal balance disorders, which could complement or enhance auditory-perceptual analysis.

The velopharyngeal sphincter plays a vital role in regulating sound and airflow during speech production. It opens to varying degrees for nasal and nasalized speech sounds, directing sound and airflow toward the nasal cavity. In contrast, for non-nasal speech sounds, the sphincter closes, directing airflow and sound through the oral cavity (Marsh, 2004; Perry, 2011). As a result, separate acoustic signals from the nasal and oral channels exhibit distinct patterns during the production of nasal and non-nasal speech. In the normal condition, there is no acoustic energy in the nasal channel during non-nasal sentences, as the velopharyngeal sphincter is fully closed. In contrast, during nasal sentences, acoustic energy is present in the nasal channel as the sphincter opens to allow airflow through the nasal cavity (Kummer, 2002). In hypernasality, acoustic energy is present in the nasal channel during both nasal and non-nasal sentences. The excessive nasal acoustic energy during non-nasal speech indicates improper velopharyngeal sphincter closure. In hyponasality, there is reduced or absent acoustic energy in the nasal channel during the nasal sentence, typically due to a blockage or narrowing of the nasal cavity. However, the non-nasal sentence exhibits normal acoustic energy patterns. In mixed nasality, acoustic energy in the nasal channel is reduced during the nasal sentence (suggesting hyponasality) but is excessive during the non-nasal sentence (indicating hypernasality). (Kummer, 2018; Marsh, 2009; Dudas et al., 2006).

As shown in Figure 4.1, stereo signals can be captured using a dual-microphone headset that simultaneously records acoustic outputs from both the nasal and oral channels during speech.

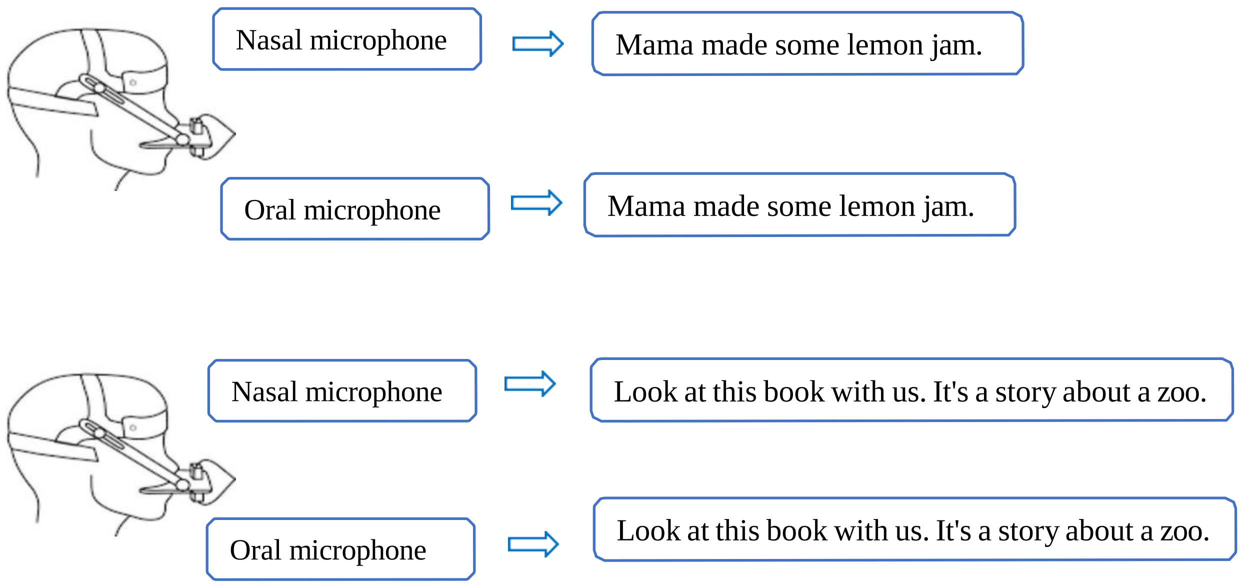


Figure 4.1. Headset Set-up for Collecting Stereo signals while producing nasal and non-nasal stimuli

This study explored whether oscillograms of stereo signals could reveal distinct visual patterns associated with four oral–nasal balance conditions. Oscillograms are valuable tools that provide detailed graphical representations of signal waveforms, highlighting variations in amplitude over time (Prasad et al., 2017; Auzou et al., 2000) (Figure 4.2). This visual representation facilitates the analysis of the temporal and intensity patterns of both oral and nasal signals, using oscillograms that may reveal distinct patterns observable through visual inspection.

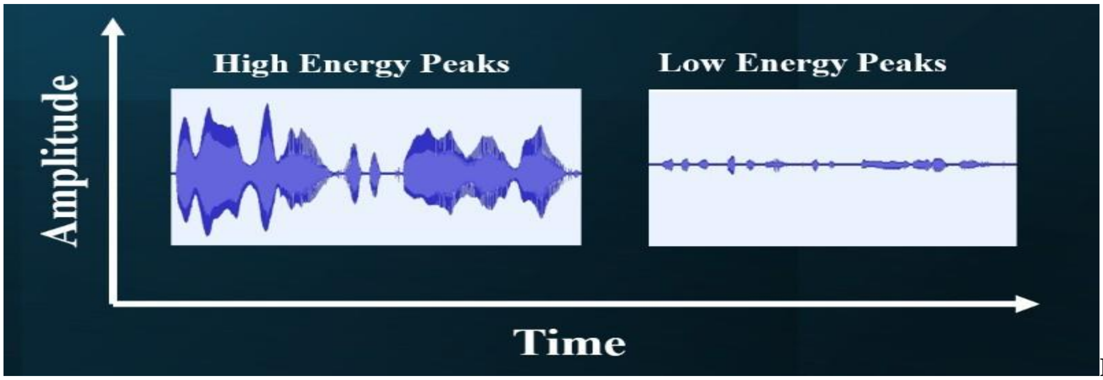


Figure 4.2. Oscillographic Representation: High vs. Low Energy Peaks

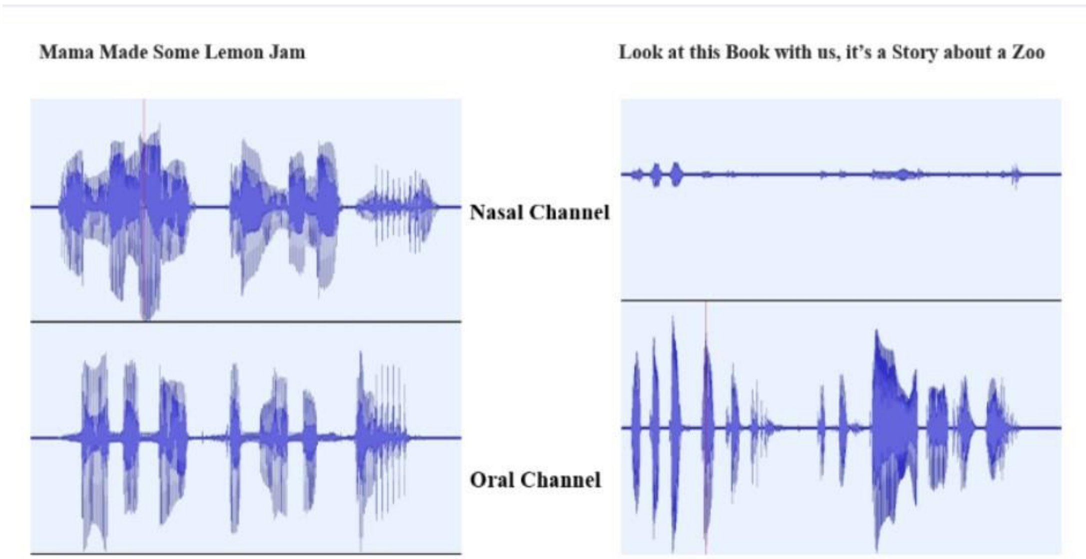


Figure 4.3. Pairing Oscillograms of Stereo Signals During Nasal and Non-Nasal Speech Stimuli.

Studies have demonstrated that visual representations offer an effective tool for analyzing complex data. Larkin and Simon (1987) explored the cognitive benefits of visual aids, showing that visual representations significantly reduce cognitive load and enhance pattern recognition. Their research found that visual tools help users identify relationships and anomalies in data that may be overlooked through textual or auditory formats, ultimately improving the accuracy of decision-making. Similarly, Harrell (2006) highlighted the value of visual diagrams, particularly argument diagrams, in fostering critical thinking and structured analysis. His study indicated that visual representations allow users to think systematically, guiding them through complex processes and improving their ability to classify and analyze information.

Building on these broader insights, recent advances in speech disorder research have also demonstrated the diagnostic power of visual signal analysis. For example, Aljarallah et al. (2025) demonstrated that Mel-spectrogram images, time–frequency visual representations of speech, can be classified using deep learning models with high accuracy to detect and categorize multiple speech disorders, such as dysarthria and dysphonia. Their study underscores how visual representations can capture clinically meaningful features that may be less apparent in raw acoustic signals, highlighting the broader applicability of visual data analysis in clinical contexts.

These findings underscore the power of visual analysis in facilitating precise and reliable decision-making, particularly when applied to tasks involving the interpretation of detailed or complex data. Indeed, pairing oscillograms representing stereo signals produced during both nasal and non-nasal sentences may offer a straightforward method for classifying oral-nasal balance disorders (Figure 4.3). This approach could not only support auditory-perceptual judgments, which can be subjective and variable, but could also provide an effective alternative to nasometry.

This study had two primary aims. The primary objective was to develop the concept of visual analysis as a method for classifying oral-nasal balance conditions based on stereo signal patterns in paired oscillograms. The second aim was to conduct a pilot study to assess how accurately licensed SLPs and SLP students classified oral-nasal balance disorders using visual analysis of paired oscillogram images. The exploratory intent was to demonstrate proof of concept whereby both licensed SLPs and SLP students could classify simulated oral–nasal balance conditions using paired oscillogram images based on visual feature analysis alone, and that the resulting classification performance would be comparable to that achieved by the previously developed multiparametric acoustic approach.

Materials and Methods

Study Design

This study involved two phases:

1. Development of the concept of visual analysis for classifying oral-nasal balance conditions
2. Pilot experimental study to assess the classification accuracy of oral-nasal balance conditions using training in visual analysis of stereo signals.

Concept development: visual analysis of stereo signals. To develop the concept of applying visual feature analysis for classifying four oral-nasal balance conditions, three steps were conducted:

1. Designing flowcharts
2. Providing instruction on visual pattern identification for oral–nasal balance classification
3. Applying visual pattern identification for classification of simulated oral-nasal balance conditions

Designing flowcharts: a decision-making tool for visual classification. We developed two flowcharts to facilitate the classification of oral-nasal balance conditions using visual features extracted from oscillogram images. Flowcharts provide a clear, structured, and efficient way to visualize complex processes, enabling easier decision-making and greater consistency in classification (Moody, 2011). The first flowchart served as an explanatory tool, outlining the steps for classifying conditions based on visual features (Figure 4.4), while the second flowchart directly classified conditions using oscillogram images and extracted features (Figure 4.5).

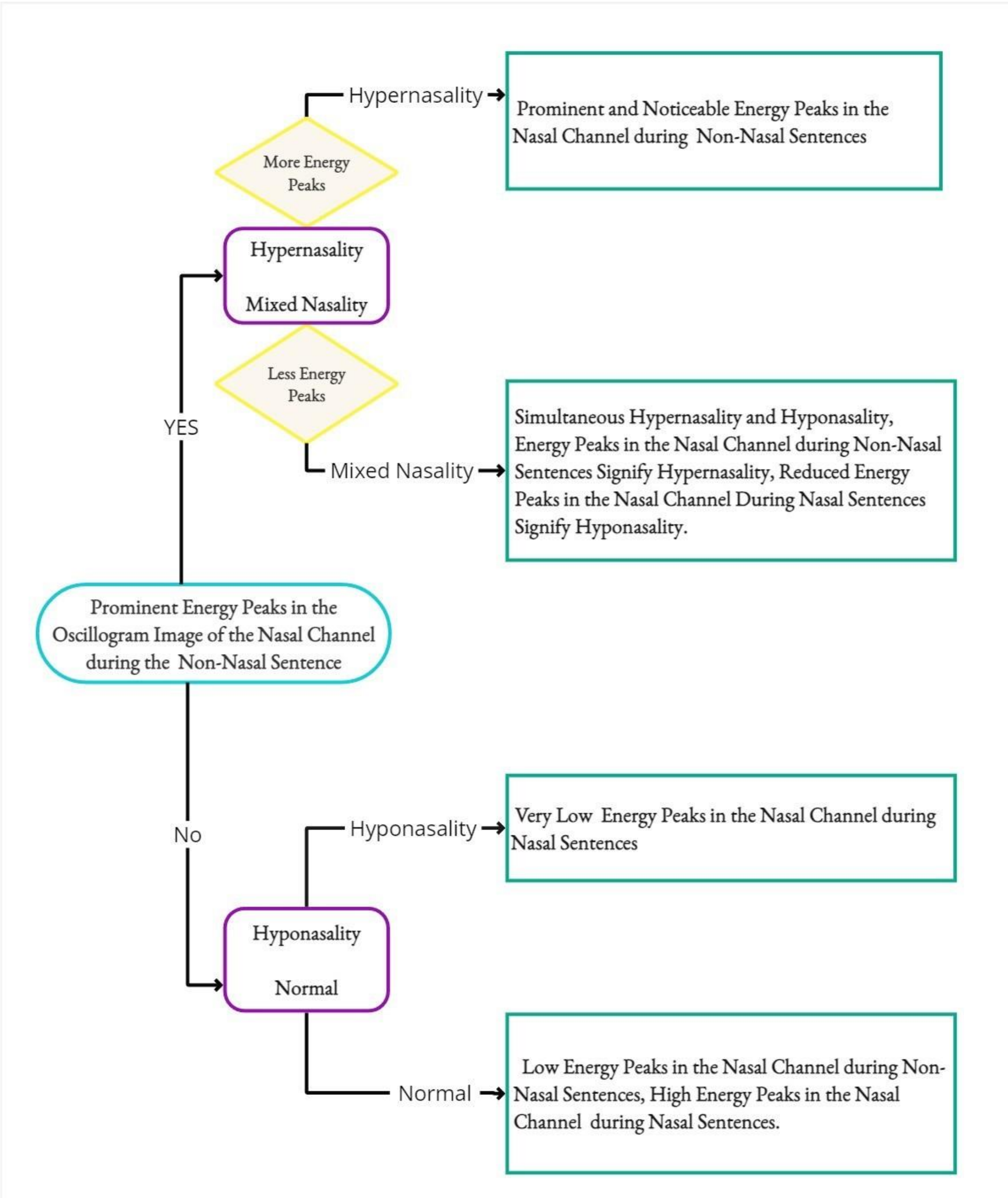


Figure 4.4. Flowchart: Visual Classification of Oral-Nasal Balance Conditions

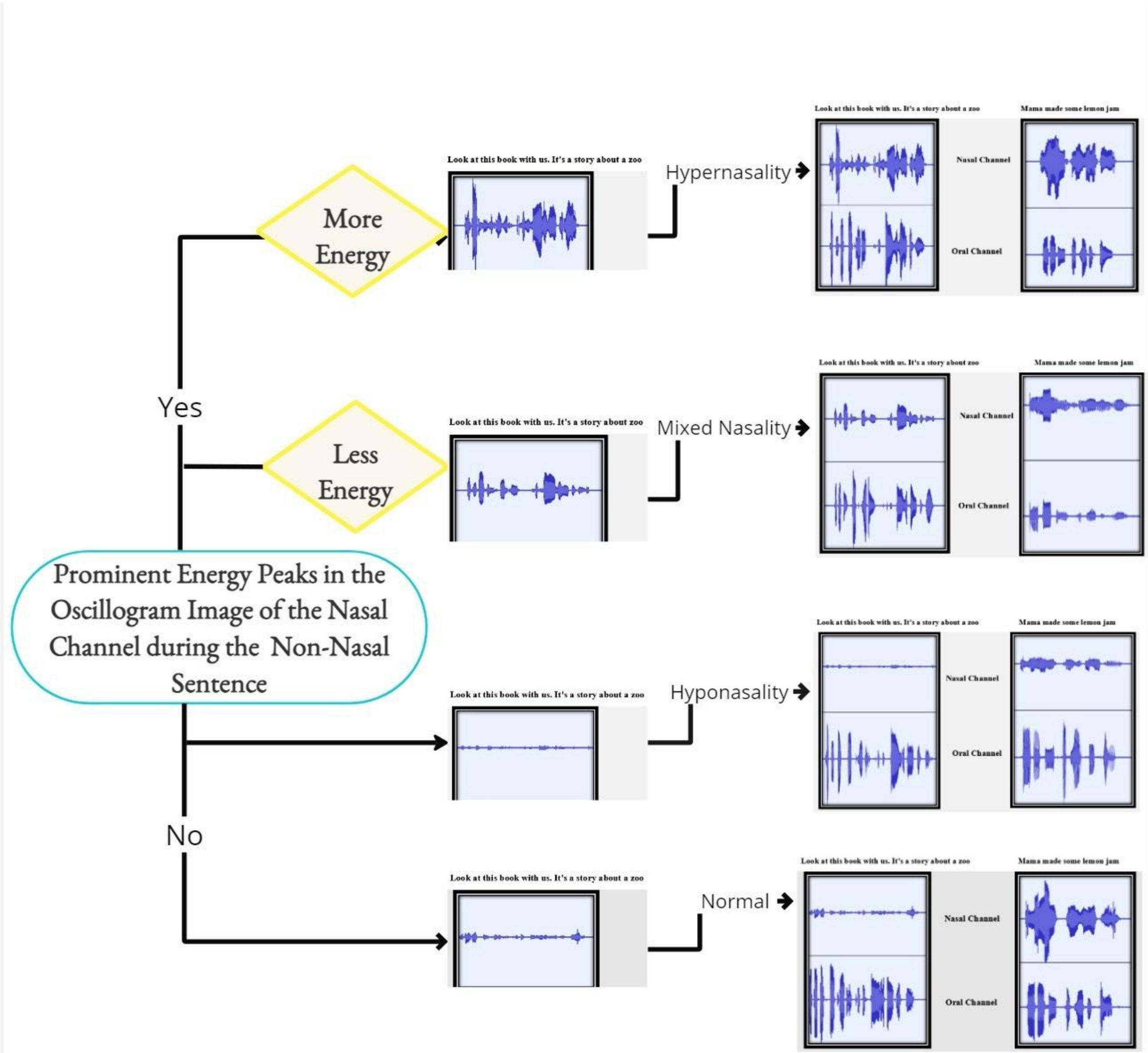


Figure 4.5. Flowchart: Visual Classification of Oral-Nasal Balance Conditions Based on Oscillogram Images

Providing instruction on visual pattern identification for oral–nasal balance classification. To develop this stage of the instructional materials, previously designed flowcharts were used, and paired oscillogram images were analyzed to extract the most salient visual cues associated with each oral-nasal balance condition. For the visual analysis, the oral signal was not included as a primary source of diagnostic information but rather served as contextual support to aid interpretation of the nasal signal. The reason behind this is that we anticipated that once proof of concept was demonstrated, future training protocols adopting analysis of the nasal signal alone could be developed.

In paired oscillograms of hypernasality, energy peaks are observed in the nasal channel during the nasal sentence, as expected for hypernasality. However, in the non-nasal sentence, excessive energy is also detected in the nasal channel, indicating the presence of excessive airflow through the nasal passages during speech (Figure 4.6).

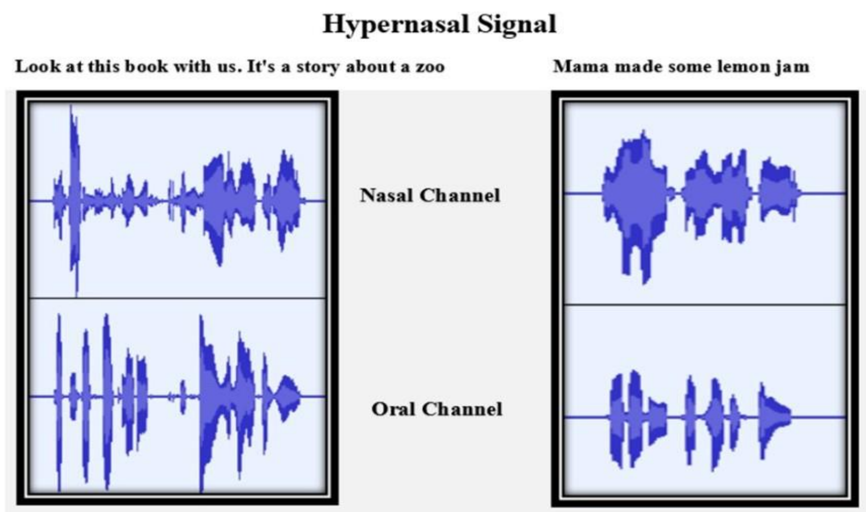


Figure 4.6. Paired oscillograms showing hypernasality during nasal and non-nasal speech stimuli.

In paired oscillograms demonstrating hyponasality, too little energy is observed in the nasal channel during the nasal sentence, suggesting a blockage in the nasal cavity. In contrast, the non-nasal sentence displays a normal energy pattern in the nasal channel (Figure 4.7).

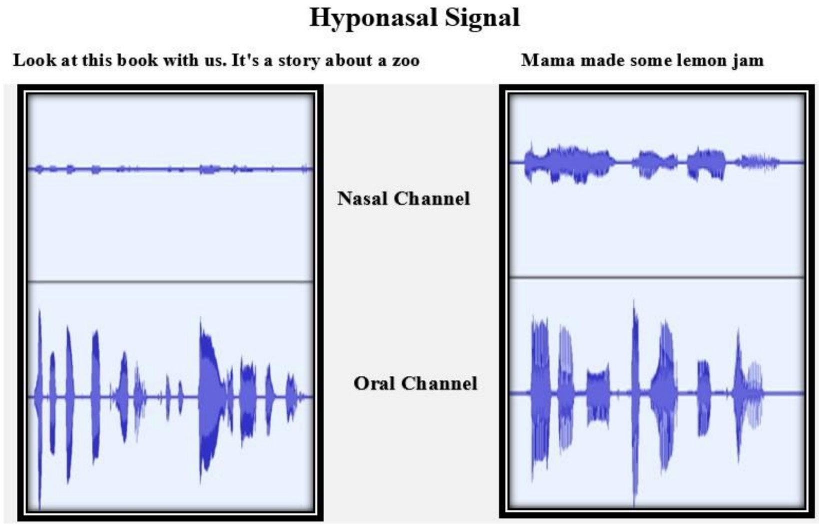


Figure 4.7. Paired oscillograms showing hyponasality during nasal and non-nasal speech stimuli.

In cases of mixed nasality, reduced energy peaks are observed in the nasal channel during the nasal sentence, indicating the presence of hyponasality. In contrast, excessive energy is detected in the nasal channel during the non-nasal sentence, reflecting characteristics of hypernasality. This pattern can be more challenging to distinguish (Figure 4.8).

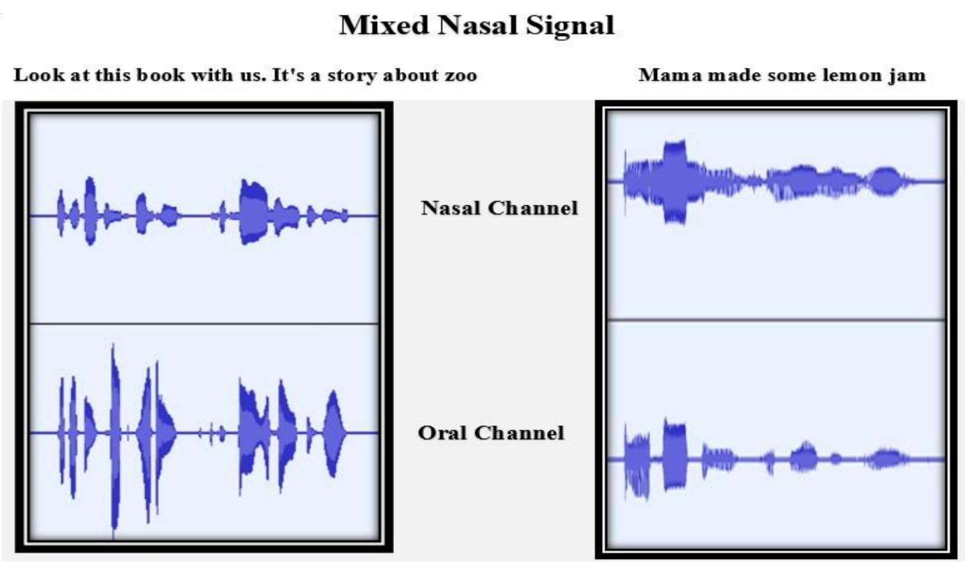


Figure 4.8. Paired oscillograms showing Mixed Nasality during nasal and non-nasal speech stimuli

Under normal condition, the nasal channel shows no energy during non-nasal sentences because the velopharyngeal sphincter remains tightly closed. In contrast, during nasal sentence production, the sphincter opens to direct airflow through the nasal cavity, resulting in detectable energy peaks in the nasal channel (Figure 4.9).

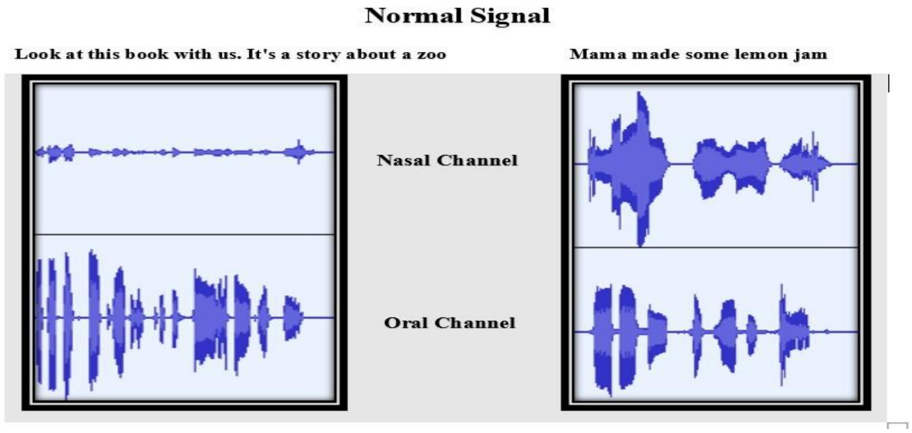


Figure 4.9. Paired oscillograms showing Mixed Nasality during nasal and non-nasal speech stimuli

Applying visual pattern identification for classification of simulated oral-nasal balance conditions. To verify the concept of visual pattern identification for classifying four oral-nasal balance conditions, an initial visual analysis was conducted on the simulated dataset by F.A., a member of the study team. A set of 16 paired oscillogram images was collected, with four paired images per condition (Figure 4.10). In the previous study (de Boer & Bressmann, 2015), these recordings were labeled based on auditory perceptual assessment. For visual classification, the previous labels of the images were first removed, and then the 16 paired oscillogram images were classified solely using visual pattern identification based on flowcharts. After classification, the results were compared with auditory perceptual conditions. A comparison between visual analysis-based labeling and the previous auditory perceptual assessment-based labeling revealed that all paired images were accurately classified. The results suggest that the concept may have potential for practical application in classifying oral-nasal balance disorders. However, further exploration through research is necessary to investigate this potential.

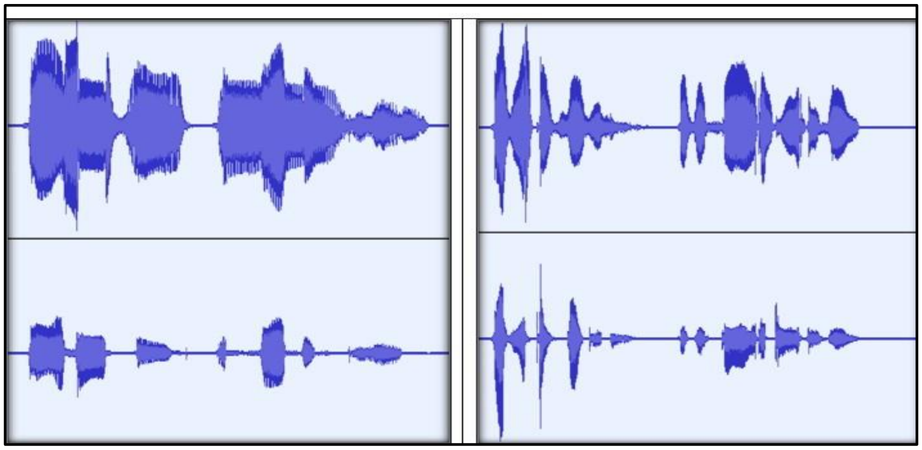


Figure 4.10 Sample Paired Oscillogram Images

Pilot Experimental Study

Participants and Recruitment

The study involved two groups of raters: licensed SLPs and SLP students from the University of Toronto. Licensed SLPs were recruited via email through professional networks, including the Canadian Cleft Palate Study Group. The recruitment email provided a brief overview of the study along with a link to an online questionnaire. Speech-language pathology students were recruited through email invitations distributed by the Department of Speech-Language Pathology at the University of Toronto. These invitations included a summary of the study and a link to the same online questionnaire (Appendix D). The student group was further divided into two subgroups based on prior coursework:

- SLP students who had already completed the Cleft Lip and Palate course
- SLP students who had not yet taken the course

All raters were informed that their participation was voluntary, and electronic informed consent was obtained before initiating the questionnaire.

Speech Database

The retrospective dataset consisted of the same speech recordings from study 1 involving eleven typical female speakers ranging in age from 22 to 30 years with native proficiency in English, normal hearing, and no history of cleft lip and palate, resonance disorders, or nasal congestion (De Boer and Bressmann, 2015). As previously described, using a nasometer headset, the speakers produced one nasal sentence ("Mama made some lemon jam") and one non-nasal sentence ("Look at this book with us. It is a story about a zoo"), with each sentence repeated twice. Each sentence was produced using the speakers' normal voice and then repeated while simulating hypernasality, hyponasality, and mixed nasality. Classification labels for this dataset were assigned based on auditory–perceptual assessment (see de Boer and Bressmann, 2015).

In the current study, each repeated production was treated as a separate individual speech signal for analysis, consistent with the signal-level focus of the study. Accordingly, the unit of analysis was the speech signal rather than the individual speaker, and repeated productions from the same speaker were not interpreted as independent observations. Using the Research Randomizer tool, 24 speech signals (6 per condition) were randomly selected from the pool of 88 available speech signals. Of these, four (1 per condition) were randomly assigned to the training phase, and the remaining 20 (5 per condition) were presented in the experimental phase.

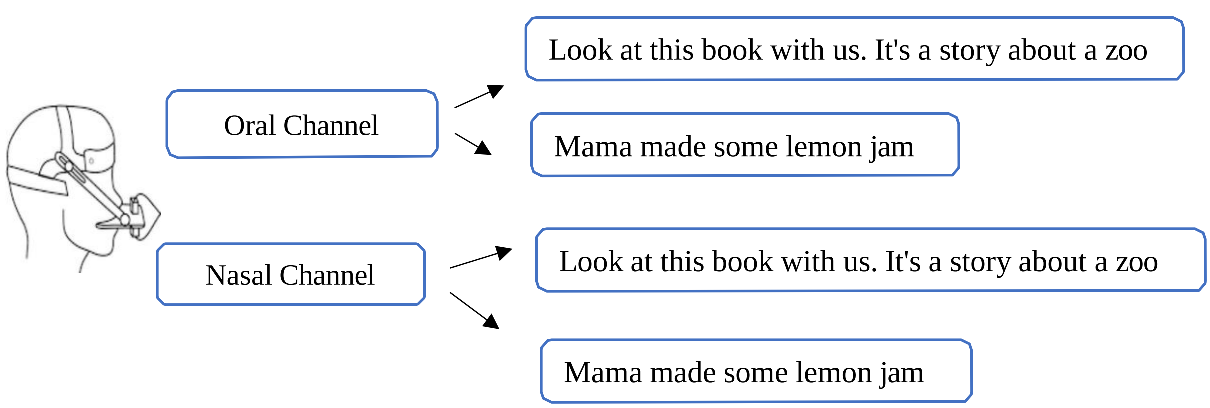


Figure 4.11. Nasometer Headset Set-up for Collecting Data.

Visualizing Stereo Signals

This study used oscillograms as visualization tools for stereo signals captured by oral and nasal microphones. The goal was to generate 24 paired oscillogram images—six paired images for each condition: normal, hypernasality, hyponasality, and mixed nasality—corresponding to both oral and nasal sentences. These images, with five paired images per condition allocated for the experimental phase and one paired image per condition used for the training phase, provided a visual representation for analyzing oral-nasal balance conditions.

Development of the Online Assessment Tool

An online questionnaire was developed based on a previously designed instructional resource on visual pattern identification for oral–nasal balance classification. Flowcharts were created to accompany the questionnaire, which was administered via Microsoft Forms (Appendix E), to evaluate how licensed SLPs and SLP students classify oral–nasal balance conditions through visual analysis. The questionnaire was divided into several sections. The first page presented raters with a consent form outlining the purpose of the study, the type of data being collected, and raters' right to withdraw at any time without consequence. After providing consent, raters submitted non-identifiable demographic information, including age, sex, and level of training. Raters then entered a training phase, during which they were instructed how to interpret paired oscillogram displays for oral–nasal balance classification relative to distinctive features (Figure 4.12). Each image pair consisted of two vertically aligned waveforms corresponding to the same speech stimulus. The upper waveform represented the nasal channel, and the lower waveform the oral channel. The two sentence types were shown side by side: the non-nasal sentence ("Look at this book with us. It's a story about a zoo") and the nasal sentence ("Mama made some lemon jam"). The training phase introduced basic oscillogram concepts (e.g., amplitude over time), explained the distinction between oral and nasal channels, and described characteristic visual patterns associated with normal resonance, hypernasality, hyponasality, and mixed nasality. Multiple labeled paired oscillograms were provided for each condition to direct attention to distinctive features in the nasal channel. The content of this training was based on the previously conducted visual analysis and flowchart development.

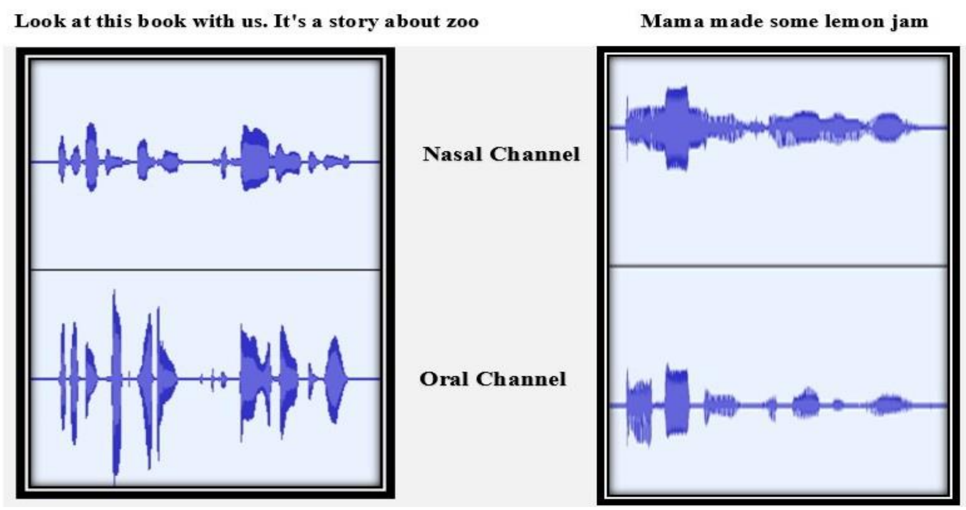
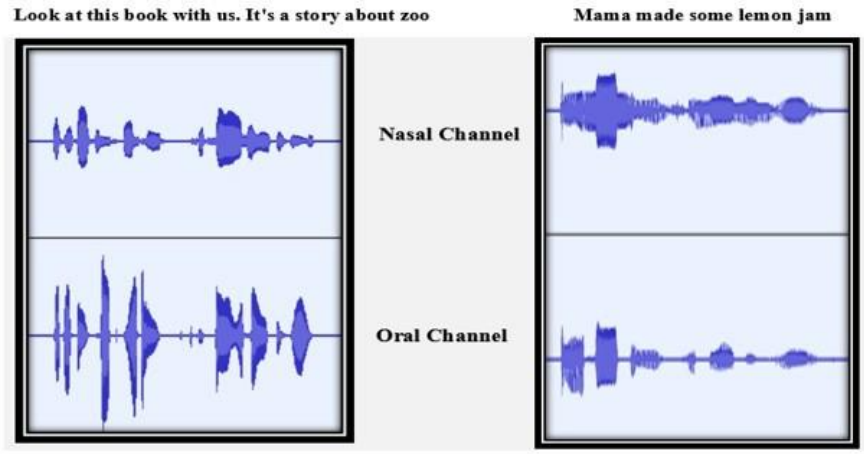


Figure 4.12. Example of a Paired Oscillogram Image Showing Mixed Nasality, Presented During the Training, Practice, and Experimental Phases for Distinctive Feature Inspection.

A practice phase followed, in which raters classified four additional paired oscillogram images—one for each of the four conditions. Immediate feedback was provided after each response, displaying the correct classification along with a brief explanation. The final part of the questionnaire was the experimental phase, in which raters were presented with 20 paired oscillogram images (five per condition) in randomized order. For each paired oscillogram display, raters were asked the question *"What condition does this picture represent?"* and prompted to select one multiple-choice response: normal, hypernasality, hyponasality, or mixed nasality. Brief definitions of each category were provided alongside the response options to ensure consistent interpretation of the classification criteria (Figure 4.13). This portion of the task was expected to be completed within approximately 10 to 15 minutes.

What condition does this picture represent?



- **Normal** (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- **Hypernasality** (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- **Hyponasality** (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- **Mixed Nasality** (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

Figure 4.13. Example of an Experimental-Phase Question Prompting Visual Inspection of Paired Oscillogram Image.

Data Collection

Data were collected and stored using Microsoft Forms, a secure survey tool integrated within the Microsoft Office 365 suite. As raters completed the online questionnaire, their responses were automatically saved and compiled into an Excel spreadsheet, which was stored within the Microsoft Forms platform. The research team downloaded the Excel spreadsheet from Microsoft Forms and used it for data analysis.

Statistical Analysis

Statistical analyses in this study were primarily descriptive in nature and aligned with the exploratory objectives of the research as a proof-of-concept approach. Classification performance was assessed using overall accuracy, group accuracy trends (licensed SLPs, SLP students with completed coursework in cleft palate, and SLP students without coursework in cleft palate), and condition-specific accuracy (normal, hypernasal, hyponasal, and mixed nasality). Misclassification patterns across conditions were examined using confusion matrices, while boxplots were used to visualize the variability of classification accuracy across rater groups. Inter-rater agreement for the visual classification task was quantified using Fleiss' kappa (κ) with the strength of agreement interpreted according to the benchmarks proposed by Landis and Koch (1977), where κ values of 0.41–0.60 indicate moderate agreement and values of 0.61–0.80 indicate high (substantial) agreement.

Results

A total of 36 raters took part in the visual analysis study, of whom 34 (94%) identified as female. The raters ranged in age from 22 to 62 years, with a mean age of 28.22 years. The sample included 12 SLP students who had already completed the cleft lip and palate course, 17 SLP students who had not yet taken the course, and 7 licensed SLPs.

Figure 4.13 shows the average classification accuracy for each rater group. SLP students who had completed the cleft lip and palate course demonstrated the highest accuracy at 87.50%, followed by licensed SLPs at 85.00%. SLP students who had not yet taken the course had a lower accuracy of 82.94%. The overall classification accuracy across all raters was 85.15%.

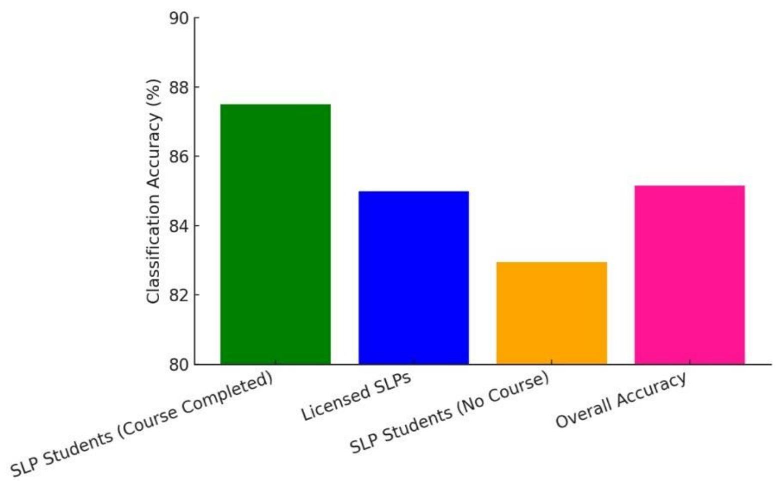


Figure 4.14. Classification Accuracy by SLP Students (with and without course), Licensed SLPs, and Overall Performance

Classification accuracy varied across both rater groups and nasality conditions. The highest accuracy was achieved in the classification of hyponasality, with both students who had completed the course and licensed SLPs reaching 100%. The lowest accuracy was observed in the classification of mixed nasality by students who had not taken the course (61%). Overall, accuracy across conditions was not consistent. Students who had completed a cleft lip and palate course and licensed SLPs performed better on all pathological conditions, but students lacking the course had the highest numerical accuracy for the normal condition. (Figure 4.15).

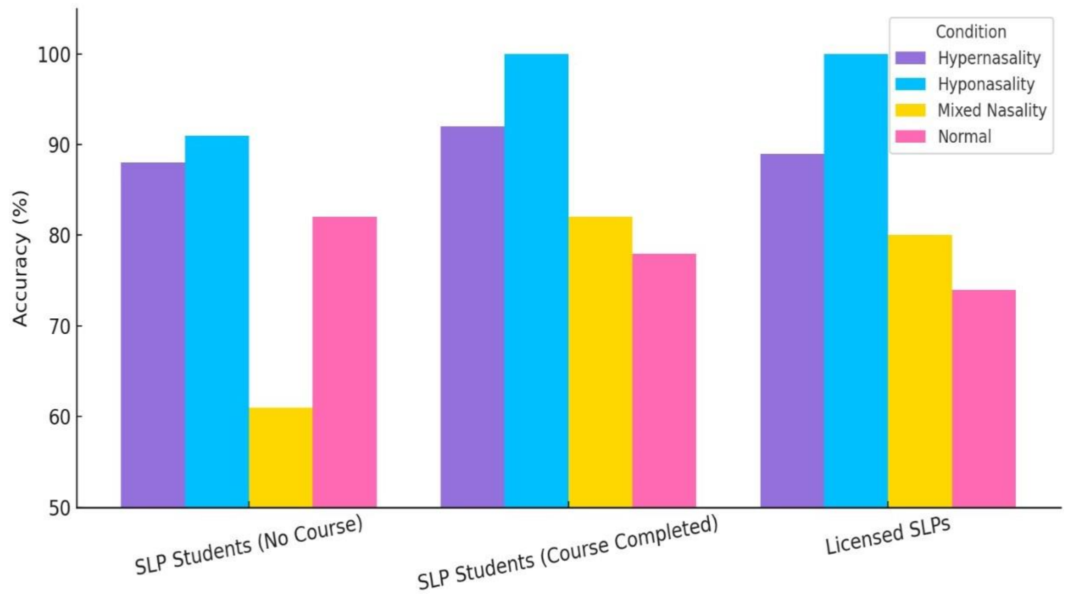


Figure 4.15. Classification accuracy across conditions (hypernasality, hyponasality, mixed nasality, and normal speech) by group.

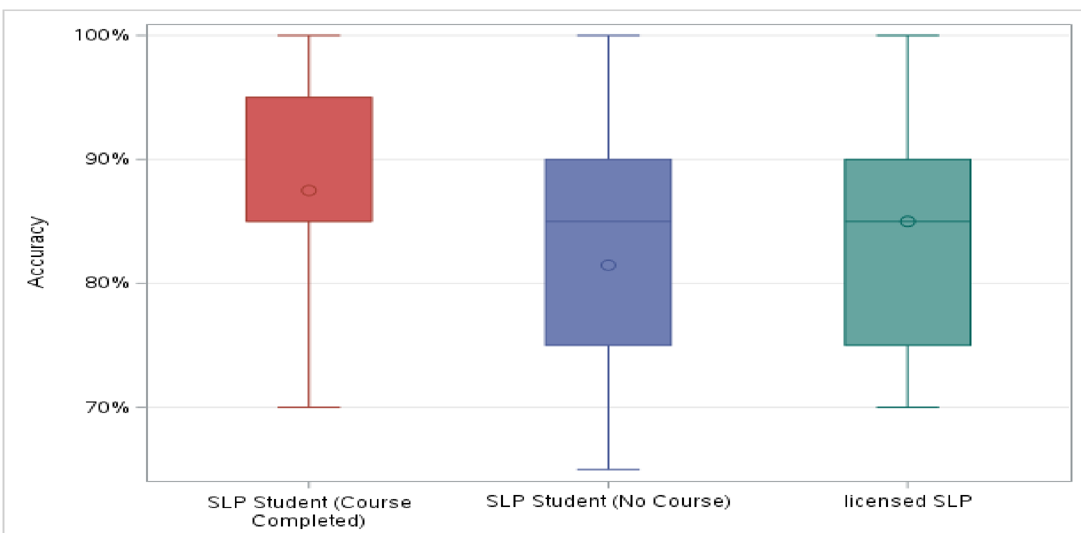


Figure 4.16. Boxplots of visual classification accuracy across rater groups. The horizontal line inside each box represents the median. The lower and upper edges of the box indicate the 25th and 75th percentiles (interquartile range, IQR). The whiskers (vertical lines extending from each box) represent values within 1.5 × IQR from the lower and upper quartiles.

Figure 4.16 illustrates the distribution of visual classification accuracy across groups using boxplots. SLP students who had completed the course showed the highest median accuracy and the narrowest interquartile range, indicating both highest overall performance and greatest consistency across raters. Error analysis based on combined responses across all rater groups revealed condition-specific patterns of misclassification (Table 4.1). Hyponasality achieved the highest classification accuracy at 95.6%, with misclassifications occurring equally as normal and mixed nasality. Hypernasality was classified with 89.4% accuracy, with all errors occurring as mixed nasality. Normal resonance was correctly identified in 79.4% of cases but was most often misclassified as hyponasality or hypernasality. Mixed nasality showed the lowest classification accuracy at 71.7%, with frequent misidentifications as hyponasality and hypernasality. These findings indicate that normal and mixed nasality were the most challenging conditions to classify accurately. Fleiss' kappa revealed high inter-rater agreement for SLP students who had completed the cleft lip and palate course (κ = 0.72) and moderate agreement for licensed SLPs (κ = 0.67), and for SLP students who had not yet taken the course (κ = 0.60).

Actual Oral-Nasal Balance Conditions	**Predicted Oral-Nasal Balance Conditions**					
	Normal	Hypernasality	Hyponasality	Mixed Nasality	Total	Accuracy (%)
Normal	143	13	15	9	180	79.4
Hypernasality	0	161	0	19	180	89.4
Hyponasality	4	0	172	4	180	95.6
Mixed Nasality	7	19	25	129	180	71.7

Table 4.1. Confusion Matrix Showing Classification Accuracy and Misclassification Patterns Across Oral-Nasal Balance Conditions

Discussion

This pilot study explored the potential of using visual feature analysis of paired oscillogram images for classifying oral-nasal balance conditions by licensed SLPs and SLP students. The results demonstrated that visual analysis may be a promising method, achieving an overall classification accuracy of 85.15%. Compared to previous acoustic-based approaches that used the same simulated dataset, visual analysis showed comparable performance. de Boer and Bressmann (2015) achieved 88.6% accuracy using nasometric analysis combined with LDA to classify simulated oral-nasal balance conditions. Later, de Boer and Bressmann (2016) applied LDA to long-term average spectrum (LTAS) features from the same dataset, obtaining an accuracy of 80.7%. Additionally, an exploratory LDA analysis focusing solely on nasal channel acoustic energy achieved a classification accuracy of 90.9% to classify simulated oral-nasal balance conditions (Chapter 2).

As indicated by the boxplots and inter-rater agreement analyses, both classification variability and inter-rater consistency differed across rater groups. Numerically greater consistency was observed among licensed SLPs and SLP students who had completed the cleft lip and palate course, whereas students without the course demonstrated lower agreement. This pattern suggests that the greater variability and reduced inter-rater agreement among untrained students may reflect less well-defined decision strategies and increased uncertainty when interpreting visually subtle or overlapping cues.

In the clinical setting, auditory-perceptual evaluation by trained clinicians is the primary method for assessing oral-nasal balance disorders (Kummer, 2009). However, due to its subjective nature, scoring reliability can vary among SLPs with different levels of training and experience (Lee et al., 2009). Auditory-perceptual assessment, while commonly used in clinical practice, is considered an imperfect reference standard for diagnosing nasality due to its inherent subjectivity and variability. In situations where no perfect gold standard exists, diagnostic accuracy can be improved by combining results from multiple imperfect measures—a strategy known as the construct reference standard method. One type of construct reference standard is the composite reference standard, which integrates the results of various tests into a unified diagnostic outcome and is often more discriminating than any single method alone. (Reitsma et al., 2009; Rutjes et al., 2007). Building on this idea, de Boer et al. (2020) proposed an approach to enhance the reliability of auditory-perceptual judgments by introducing an acoustic-based pre-classification step. In their study, nasalance scores were first used to categorize speech samples into oral-nasal balance groups (normal, hypernasality, hyponasality, or mixed nasality). Clinicians then performed auditory-perceptual evaluations based on these pre-classified samples. This process led to improved diagnostic accuracy and greater inter-rater agreement, demonstrating that combining objective measures with perceptual judgments can yield a more reliable reference standard. However, the use of nasometry requires specialized and often costly equipment that may not be readily accessible to all clinicians (Han, Kang, & Ko, 2025). As a more practical alternative, visual analysis of paired oscillogram images could serve as an accessible pre-classification tool. By visually highlighting differences in signal intensity between oral and nasal channels across various oral-nasal balance disorders, this method could function similarly to nasalance-based pre-classification. Specifically, visual analysis could be used as a triage step, a preliminary assessment conducted prior to auditory-perceptual evaluation, to help reduce variability in perceptual judgments and improve diagnostic agreement in clinical assessments of nasality.

Several key features characterize practical screening tools: they should be quick to administer, easy to use with minimal training, risk-free, and cost-effective. Screening tools are typically designed to prioritize efficiency and accessibility rather than providing definitive diagnoses, helping clinicians identify cases that require further detailed assessment (Iragorri & Spackman, 2018). The recorded signals can then be easily converted into oscillogram representations for visual inspection. This process does not require specialized equipment beyond basic recording devices already available in most clinical settings. Given its simplicity, low cost, and relatively high classification accuracy demonstrated in this study, visual analysis could serve as an effective screening tool for oral-nasal balance disorders, supporting clinical decision-making, particularly when access to advanced instrumental assessments such as nasometry is limited.

In the education setting, speech-language pathology students typically receive foundational training in the physics of sound and introductory acoustic analysis. However, most acoustic-based methods for detecting nasality, such as spectral analysis, require specialized signal processing expertise that falls outside the scope of standard coursework and clinical training. Visual analysis may offer a more accessible alternative that builds directly on students' existing knowledge of speech signals. An oscillogram displays amplitude along the vertical axis, reflecting sound energy, and time along the horizontal axis, representing speech over time. Through visual inspection of oscillograms, students can analyze oral and nasal airflow patterns without requiring complex technical skills. The results further supported this idea, as students who had not yet taken the cleft lip and palate course were still able to classify oral-nasal balance disorders with relatively high accuracy. Visual analysis could be readily integrated into educational modules, such as courses on cleft palate and resonance disorder, providing students with a practical and efficient method for understanding oral-nasal balance disorders and velopharyngeal function. Incorporating visual analysis into early clinical training could not only enhance students' theoretical understanding but also potentially equip them with a clinically applicable tool for screening, pre-classification, or complementing traditional auditory-perceptual assessments in practice.

In this study, SLP students who had completed a cleft palate course achieved slightly higher overall accuracy than licensed SLPs when using visual analysis of paired oscillograms. This may be because students' knowledge of acoustics is more recent, making them more responsive to the specific training protocol, while licensed SLPs may have moved on from this foundational knowledge, relying more on auditory-perceptual methods. Experienced clinicians often develop efficient, routine strategies based on auditory-perceptual evaluation, which may be less adaptable to a new technique such as visual analysis. In contrast, students with fewer established practice patterns may be more open to adopting novel diagnostic approaches, making them more flexible in integrating such methods into their work (Mattila, 2003; Eraut, 1994). Additionally, the licensed SLP group was relatively small, and inherent heterogeneity within this group may have contributed to greater variability in performance. Licensed SLPs may also approach classification tasks with an expectation of identifying pathology, which could bias interpretations toward abnormal categories.

Based on error analysis patterns, it appears that enhancing and expanding visual analysis training may be necessary to improve classification accuracy, particularly for normal and mixed nasality patterns, which demonstrated the highest rates of misclassification compared to other oral-nasal balance conditions. One likely explanation for the misclassification of normal resonance is that it does not exhibit the prominent signal deviations typically seen in disordered conditions such as hypernasality and hyponasality. As a result, oscillogram representations of normal speech may lack distinctive visual cues, making accurate identification more challenging. In addition, raters may have exhibited a bias toward over-identifying mild deviations, leading to a tendency to classify normal speech as disordered and thereby reducing accuracy for this condition. Such a bias may reflect a conservative clinical decision-making strategy, in which borderline patterns are more likely to be labeled as abnormal to avoid missing potential pathology, resulting in an increased rate of false-positive classifications for normal condition. In the case of mixed nasality, the lower classification accuracy may be explained by its complex nature. Mixed nasality includes characteristics of both hypernasality and hyponasality, which can appear within the same speech sample. Because of this overlap, the visual patterns in oscillograms may resemble either condition or shift between them, making it harder to identify mixed nasality as a distinct category. Without clear visual guidelines, raters may have found it challenging to recognize mixed nasality and instead labeled it as one of the more familiar single conditions. To address these challenges, training may have to place greater emphasis on the distinct visual characteristics of both normal and mixed nasality patterns. Providing more structured examples and clearer visual models for these conditions may help clinicians develop more refined interpretive strategies. Overall, these findings support the feasibility of visual analysis as a promising, accessible, and clinically applicable method for evaluating oral-nasal balance disorders, laying the groundwork for future research into improving and standardizing visual diagnostic approaches.

Conclusion

This pilot study demonstrated that visual analysis of paired oscillogram images was a feasible method for classifying oral-nasal balance disorders. The findings suggest that visual analysis has the potential to serve as a clinical tool for screening, assessment, and supporting and enhancing auditory-perceptual evaluations in the diagnosis of oral-nasal balance disorders.

Limitations

This pilot study has limitations that should be acknowledged. First, the sample size was relatively small, including only a limited number of licensed SLPs and SLP students, which may affect the generalizability of the findings. Second, the study relied on a small set of simulated speech samples to create the paired oscillogram images. The restricted number of stimuli may have reduced the range of visual patterns available to raters and may not have fully represented the diversity of clinical speech. Third, because the study relied on a simulated speech dataset rather than clinical recordings from individuals with diagnosed velopharyngeal dysfunction, the representations may not have captured the true variability and complexity encountered in real-world clinical populations (hence reduced validity).

Next Steps

Future work could further refine training protocols to rely exclusively on nasal signal visualizations, allowing raters to learn diagnostic patterns based solely on nasal acoustic activity and thereby enhancing consistency with the nasal-focused analytical framework of the present thesis. In addition, future research with more raters and a larger, more diverse set of clinically representative speech samples is needed to validate and expand upon these preliminary findings. With a sufficiently large and diverse set of oscillogram images, future work could leverage the systematic visual patterns observed in oscillogram representations to develop an automated classification system based on machine learning–based pattern recognition. Additionally, a focus group could be conducted with licensed SLPs specializing in cleft palate to gather feedback on the usability, effectiveness, and practical applicability of the visual analysis method in real-world clinical settings, as well as to explore clinicians’ perspectives on its strengths, limitations, and opportunities for further refinement and enhancement. Ultimately, the goal is to further develop visual analysis as a clinical tool for screening, assessment, support, and improvement of auditory-perceptual analysis in the evaluation of oral-nasal balance disorder.

References

Aljarallah, N. A., Dutta, A. K., & Sait, A. R. W. (2025). Image classification-driven speech disorder detection using deep learning technique. *SLAS technology*, *32*, 100261.

Auzou, P., Ozsancak, C., Morris, R. J., Jan, M., Eustache, F., & Hannequin, D. (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: A review. *Clinical Linguistics & Phonetics, 14*(2), 131–150.

Bettens, K., De Bodt, M., Maryn, Y., Luyten, A., Wuyts, F. L., & Van Lierde, K. M. (2016). The relationship between the Nasality Severity Index 2.0 and auditory-perceptual judgments of hypernasality. *Journal of Communication Disorders, 62*, 67-81.

Bressmann, T., & Abnavi, F. (2025). Spectral speech analysis of hypernasality. In A. am Zehnhoff-Dinnen, J. Sopko, M.-C. Monfrans-Plauwadell, & K. Neumann (Eds.), *Phoniatrics II: Speech and speech fluency disorders—Literacy development disorders*. Springer. https://link.springer.com/book/9783031251757

de Boer, G., & Bressmann, T. (2015). Application of linear discriminant analysis to the nasometric assessment of resonance disorders: A pilot study. *Cleft Palate-Craniofacial Journal, 52(2)*, 173–182. https://doi.org/10.1597/13-109

de Boer, G., & Bressmann, T. (2016). Application of linear discriminant analysis to the long-term averaged spectra of simulated disorders of oral-nasal balance. *Cleft Palate-Craniofacial Journal, 53(5*), e163–e171. https://doi.org/10.1597/14-236

de Boer, G., Marino, V. C. de C., Dutka, J. de C. R., Pegoraro-Krook, M. I., & Bressmann, T. (2020). Nasalance-based preclassification of oral–nasal balance disorders results in higher agreement of expert listeners' auditory–perceptual assessments: Results of a retrospective listening study. *The Cleft Palate–Craniofacial Journal*, *57*(4), 448–457. https://doi.org/10.1177/1055665619873506

Dudas, J. R., Deleyiannis, F. W. B., Ford, M. D., Jiang, S., & Losee, J. E. (2006). Diagnosis and treatment of velopharyngeal insufficiency: Clinical utility of speech evaluation and videofluoroscopy. *Annals of Plastic Surgery, 56*(5), 511–517. https://doi.org/10.1097/01.sap.0000210628.18395.de

Eraut, M. (1994). *Developing professional knowledge and competence.* Falmer Press. Han, K.-S., Kang, M.-S., & Ko, M.-H. (2025). Evaluation of a Newly Developed Device for Resonance Disorders: A Clinical Pilot Study. *Applied Sciences*, *15*(10), 5713. https://doi.org/10.3390/app15105713

Harrell, M. (2006). Diagrams that really are worth ten thousand words: Using argument diagrams to teach critical thinking skills. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (Vol. 28, pp. 244–249). Cognitive Science Society.

Iragorri, N., & Spackman, E. (2018). Assessing the value of screening tools: Reviewing the challenges and opportunities of cost-effectiveness analysis. *Public Health Reviews, 39,* 17. https://doi.org/10.1186/s40985-018-0093-8

Karnell, M. P. (2011). Instrumental assessment of velopharyngeal closure for speech. *Seminars in Speech and Language, 32, 168–178. Journal*, 20(3), 199-208.

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),* 1-35.

Kuehn, D. P. (1982). Assessment of resonance disorders. In N. J. Lass, L. V. McReynolds, J. C. Northern, & D. E. Yoder (Eds.), *Speech, language, and hearing: Vol. II. Pathologies of speech and language*. W. B. Saunders.

Kummer, A. W., & Lee, L. (1996). Evaluation and treatment of resonance disorders. *Language, Speech, and Hearing Services in Schools*, *27*(3), 271–281. https://doi.org/10.1044/0161-1461.2703.271

Kummer, A. W. (2002). Velopharyngeal dysfunction: current thinking on the cause, effect, assessment, and treatment. *Current Opinion in Otolaryngology & Head and Neck Surgery, 10(6)*, 455-459.

Kummer, A. W. (2009). Assessment of velopharyngeal function. In J. E. Losee & R. E. Kirschner (Eds.), *Comprehensive cleft care* (pp. 589–605). McGraw-Hill.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive science*, *11*(1), 65-100.

https://doi.org/10.1016/S0364-0213(87)80026-5

Lee, A., Whitehill, T. L., & Ciocca, V. (2009). Effect of listener training on perceptual judgement of hypernasality. *Clinical Linguistics and Phonetics, 23(5),* 319–334. https://doi.org/10.1080/02699200802688596

Lee, A., Potts, S., & Bressmann, T. (2019). Speech-language therapy students' auditory-perceptual judgements of simulated concurrent hypernasality and articulation disorders. *Clinical Linguistics & Phonetics*, *33(10–11)*, 985–1001. https://doi.org/10.1080/02699206.2019.1655666

Mattila, A. S. (2003). The impact of cognitive inertia on postconsumption evaluation processes. *Journal of the Academy of Marketing Science, 31*(3), 287–299. https://doi.org/10.1177/0092070303031003006

Marsh, J. L. (2004). The evaluation and management of velopharyngeal dysfunction. *Clinics in Plastic Surgery*, *31*(2), 261–269. https://doi.org/10.1016/S0094-1298(03)00124-X

Marsh, J. L. (2009). Velo-pharyngeal dysfunction: Evaluation and management. Indian Journal of Plastic Surgery, 42(Suppl. 1), S144–S154.

https://doi.org/10.4103/0970-0358.57201

Moody, D. L. (2011). Why a diagram is only sometimes worth a thousand words: An analysis of the bpmn 2.0 visual notation. *Acedido em*, *27*(07), 2017.

Perry, J. L. (2011). Anatomy and physiology of the velopharyngeal mechanism. *Seminars in Speech and Language, 32(2),* 83–92. https://doi.org/10.1055/s-0031-1277712

Prasad, K. S., Ramaiah, G. K., & Manjunatha, M. B. (2017). Speech features extraction techniques for robust emotional speech analysis/recognition. *Indian Journal of Science and Technology*, *10*(3), 1–9. https://doi.org/10.17485/ijst/2017/v10i3/110571

Reitsma, J. B., Rutjes, A. W. S., Khan, K. S., Coomarasamy, A., & Bossuyt, P. M. (2009). A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *Journal of Clinical Epidemiology*, *62*(8), 797–806. https://doi.org/10.1016/j.jclinepi.2009.02.005

Rutjes, A. W. S., Reitsma, J. B., Coomarasamy, A., Khan, K. S., & Bossuyt, P. M. (2007). Evaluation of diagnostic tests when there is no gold standard: A review of methods. Health Technology Assessment, 11(50), 1–172. https://doi.org/10.3310/hta11500

CHAPTER 5

Discussion and General Conclusions of the Dissertation

There is consensus in the literature that auditory-perceptual judgments are considered the gold standard for clinically evaluating nasality, particularly given that a speech disorder is not considered to exist until a listener perceives it. To enhance the accuracy of auditory-perceptual assessments, the studies in this thesis aimed to explore the feasibility of incorporating two additional approaches—acoustic and visual analysis—as complementary methods. These approaches were examined to support and potentially improve the reliability and objectivity of auditory-perceptual judgments. The first study highlighted the potential of nasal-signal-derived acoustic parameters in classifying simulated oral-nasal balance disorders, demonstrating how multiparametric acoustic features from the nasal signal can be used to enhance the classification process. The second study demonstrated that, although the LDA model, based on nasal-signal-derived acoustic features, yielded promising results on clinical data, the performance gap between simulated and clinical data highlighted the need for further refinement. This gap highlights the importance of enhancing the model's generalizability and preventing overfitting to ensure its effectiveness in clinical settings. The third study explored the potential of visual analysis as a low-tech, trainable, and accessible method for classifying oral-nasal balance disorders. The findings suggested that visual analysis holds significant promise as a supportive tool for both screening and diagnosing oral-nasal balance disorders, particularly as a complement to auditory-perceptual evaluation.

Taken together, these studies advance a multidimensional framework for the clinical evaluation of oral-nasal balance disorders, one that maintains the importance of expert perceptual judgment while integrating acoustic and visual data to enhance diagnostic reliability and accuracy. This framework supports the transition from subjective assessments toward reproducible approaches, which is particularly relevant for clinical populations, where precise evaluation of oral-nasal balance is essential for diagnosis and treatment planning. Moreover, it is especially valuable in clinical contexts where access to expert raters is limited or where consistency across evaluators is critical, as it provides a structured, evidence-based method that promotes fairness, comparability of results, and improved clinical decision-making across diverse healthcare settings.

Efficacy and evidence-based practice have received increasing attention in recent years. Clinical efficacy, as described by Karnell and Seaver, refers to "whether a procedure or instrument is appropriate for a specific task, given the constraints of patient tolerance, time, cost, space, training, and other practical considerations that are characteristic of the patient population and clinical facility." These considerations, taken together to minimize the burden of care on patients, underscore the importance of carefully selecting supplemental perceptual evaluation techniques. From the perspective of diagnostic accuracy, instrumental assessments of nasality should be strongly correlated with listeners' perceptual judgments of nasality, be repeatable and reliable, cost-effective, non-invasive, simple to administer, portable, and capable of providing rapid results. The methods developed in this dissertation address these criteria by being accessible, non-invasive, independent of highly specialized technology, and cost-effective, thereby increasing their applicability across diverse clinical settings. Additionally, the results demonstrated promise for classifying oral nasal balance conditions, with acoustic analysis achieving up to 90.9% accuracy for simulated data and up to 100% for clinical data, and visual analysis achieving an overall classification accuracy of 85.15%. While the accuracy appears excellent for the first two studies, it should be kept in mind that the high results may reflect overspecification of the models. Nevertheless, these findings underscore the potential of these approaches as complements to auditory-perceptual assessment. In the future, they could enhance diagnostic robustness and contribute to evidence-based clinical decision-making.

The broader implications of this work could extend to both clinical practice and research. Clinically, the integration of acoustic and visual tools into diagnostic workflows could enhance diagnostic accuracy, support more personalized intervention planning, and guide follow-up treatments, particularly for individuals with cleft palate. These strategies may also be valuable in evaluating oral-nasal balance disorders associated with neurological conditions, such as spastic dysarthria, Parkinson's disease, or amyotrophic lateral sclerosis (ALS), thereby broadening their applicability beyond cases related to clefts. From a research standpoint, this work underscores the importance of validating diagnostic models in diverse populations and settings. It highlights the need for interdisciplinary collaboration among clinicians, engineers, and data scientists to develop tools that are both accurate and feasible in real-world clinical settings.

Importantly, the present thesis highlights the potential value of focusing on the nasal signal alone in classifying oral nasal balance conditions, which may offer a more feasible and cost-effective alternative to traditional dual-microphone or ratio-based methods such as nasometry. By prioritizing the nasal signal, detection and classification of oral–nasal balance disorders can be implemented with reduced reliance on precise oral–nasal microphone balance and calibration procedures, thereby increasing feasibility and accessibility in real-world clinical settings. This reduction in technical complexity could have direct cost implications, as a nasal-focused approach would reduce hardware requirements and minimize setup constraints.

In the third study, visual representations of both the oral and nasal signals were provided for analysis. However, the instructions explicitly directed raters' attention to features associated with the nasal signal to maintain consistency with the nasal-focused analytical framework adopted throughout the thesis. Future work should refine the visual analysis protocol to rely exclusively on nasal-signal visualizations, enabling clinicians to differentiate diagnostic patterns based solely on the nasal signal.

These findings validate the theoretical assumption that oral–nasal balance conditions exist as measurable phenomena that can be quantified and analyzed using multiple modalities. The results indicate that each modality captures meaningful dimensions of oral–nasal balance: acoustic features provide precise, quantifiable indices of oral–nasal balance patterns, while visual analyses of oscillograms offer a dynamic, visual representation of these patterns, enabling observers to detect subtle temporal and intensity variations. These outcomes further underscore the value of the post-positivist theoretical framework, highlighting how systematic, empirical observation and logical inference strengthen both methodological rigor and the interpretive power of findings. By adopting this framework, the research relied on objective, measurable evidence and rigorously evaluated each modality independently. This approach facilitates critical reflection on the limitations and strengths of each method, supporting the development of conclusions that are robust, generalizable, and interpretable within a broader scientific and clinical context. Importantly, these results provided direct answers to the thesis's research questions. The evaluation of the LDA-derived algorithm demonstrated that acoustic features from the nasal signal can reliably classify oral-nasal balance patterns, addressing the first research question. Similarly, the use of visual analyses of oscillograms provided evidence that clinicians and students can accurately detect oral-nasal balance disorders through visual inspection, supporting the second research question. By grounding the interpretation of these findings in a post-positivist theoretical framework, this dissertation demonstrates how systematic measurement and logical reasoning across different modalities yield complementary insights, reinforcing the validity and applicability of multimodal assessment approaches for oral-nasal balance disorders.

While the present dissertation demonstrates the promise of quantitative acoustic and visual methods, it also highlights the need to expand inquiry beyond measurement alone. From a post-positivist perspective, future research should complement these findings with qualitative inquiry to explore how these methods are perceived in clinical practice. A phenomenological qualitative approach would be particularly well-suited to extending the findings of this thesis, as it would allow for an in-depth exploration of clinicians' experiences when engaging with acoustic and visual diagnostic tools. While the present work focuses on classification accuracy, phenomenology could examine how speech-language pathologists interpret, trust, and integrate nasal-signal–based acoustic and visual approaches into their clinical decision-making. Semi-structured interviews could be used to capture clinicians' perceptions of clarity, cognitive load, and confidence in decision-making. Such insights would help identify potential barriers and facilitators to clinical adoption that are not detectable through quantitative studies.

In conclusion, this dissertation contributes to a growing body of evidence supporting multimodal assessment approaches for oral-nasal balance disorders. It lays the foundation for future research aimed at developing and validating standardized, user-friendly, and accessible diagnostic systems that complement traditional clinical expertise, with the ultimate goal of enhancing treatment outcomes and optimizing overall patient care.

Chapter 6

Contribution of Collaborators

Research Team

This doctoral dissertation was conceptualized and developed through the collaborative support of the supervisory and thesis committee members. Fatemeh Abnavi (FA), a PhD candidate in the School of Rehabilitation Sciences at the University of Ottawa, served as the student researcher and was responsible for the complete conceptualization, implementation, analysis, and writing of all research activities presented in this dissertation. FA accepts full responsibility for the final manuscripts and the dissertation. The research team includes FA's supervisor, Dr. Heather Flowers (HF), SLP, PhD, and co-supervisor, Dr. Tim Bressmann (TB), SLP, PhD. In addition, the thesis committee includes Dr. Hilmi R. Dajani (HRD), PhD in Biomedical Engineering, and Dr. Suzy Ahn (SA), PhD in Linguistics.

Primary Research Team. FA is a speech-language pathologist with expertise in acoustic analysis of speech and acquired language disorders, including aphasia and dementia. During her doctoral studies, FA worked as a clinical research assistant in the tDCS Clinical Research Lab at Baycrest's Rotman Research Institute. HF is an Associate Professor in the School of Rehabilitation Sciences at the University of Ottawa and a registered speech-language pathologist with over 20 years of clinical experience across various healthcare settings. She earned both her Master of Health Science and PhD in Speech-Language Pathology from the University of Toronto. Her research focuses on post-stroke swallowing and communication disorders, recovery patterns following stroke, and intervention strategies for speech and swallowing difficulties in individuals with neurological disease. TB is an Associate Professor in the Department of Speech-

Language Pathology at the University of Toronto, with a cross-appointment to the Faculty of Dentistry. He holds a PhD in Phonetics from the University of Munich and has a clinical background in speech therapy within neurological rehabilitation and maxillofacial surgery. His research focuses on structurally related speech and voice disorders, particularly in individuals with head and neck cancer and craniofacial syndromes. He utilizes two- and three-dimensional ultrasound imaging to study tongue function and nasometry to evaluate nasality in speech. Each member of the primary research team contributed to the design, development, and approval of the research plan and provided essential methodological and content expertise throughout the doctoral process.

Thesis Committee. HRD is a Full Professor in the School of Electrical Engineering and Computer Science at the University of Ottawa. He holds a PhD in Electrical and Computer Engineering with a specialization in Biomedical Engineering from the University of Toronto. His research focuses on developing advanced signal analysis methods and instrumentation for evaluating the function of the cardiovascular and auditory systems. His work includes applications in blood pressure monitoring, PPG and ECG signal processing, and auditory evoked potentials, integrating machine learning and physiological modeling. SA is an Assistant Professor in the Department of Linguistics at the University of Ottawa. She earned her PhD in Linguistics from New York University in 2018 and previously served as a Lecturer at the University of California, Los Angeles. Her research interests include phonetics, phonology, bilingualism, and multilingualism. Dr. Ahn’s work combines theoretical and experimental approaches to the study of speech sound systems and language structure, with a focus on how these are processed and represented across different languages.

Each committee member provided specialized guidance and feedback within their areas of expertise—Dr. Dajani contributed insight into signal processing and machine learning applications, while Dr. Ahn offered expertise in linguistics and acoustic analysis of speech. Their input during the comprehensive examination, proposal development, data interpretation, and writing stages enhanced the overall rigor and interdisciplinary depth of the dissertation.

Chapter 7

Appendices

Appendix A – The University of Toronto Research Ethics Board approval of the Secondary Use of the Speech Datasets (Chapters 2, 3,4)

Appendix B – The University of Toronto Research Ethics Board approval of Visual Feature Analysis Study (Chapter 4)

Appendix C – The University of Ottawa Research Ethics Board approval of the Visual Feature Analysis Study (Chapter 4)

Appendix D – Recruitment Emails for Raters (Licensed SLPS and Student SLPS) -Visual Feature Analysis Study (Chapter 2)

Appendix E – Online Questionnaire and Consent Form for Visual Feature Analysis Study (Chapter 2)

Appendix E - Figure Permissions Summary (Chapter 1)

Appendix A – The University of Toronto Research Ethics Board approval of the Secondary Use of the Speech Datasets (Chapter 2, 3,4)

UNIVERSITY OF TORONTO

OFFICE OF THE VICE-PRESIDENT, RESEARCH AND INNOVATION

RIS Protocol Number: 38903

Approval Date: 11-Feb-25

PI Name: Dr Tim Bressmann

Division Name:

Dear Dr Tim Bressmann:

Re: Your research protocol application entitled, "Re-analysis of data from previous protocol "Nasalance and Acoustic Profile of Mixed Nasal Resonance", approved August 17, 2012, Ref 27918"

The Health Sciences REB has conducted a Delegated review of your application and has granted approval to the attached protocol for the period 2025-02-11 to 2026-02-25.

This approval covers the ethical acceptability of the human research activity; please ensure that all other approvals required to conduct your research are obtained prior to commencing the activity.

Please be reminded of the following points:

- An **Amendment** must be submitted to the REB for any proposed changes to the approved protocol. The amended protocol must be reviewed and approved by the REB prior to implementation of the changes.
- An annual **Renewal** must be submitted for ongoing research. Renewals should be submitted between 15 and 30 days prior to the current expiry date.
- A **Protocol Deviation Report** (PDR) should be submitted when there is any departure from the REB-approved ethics review application form that has occurred without prior approval from the REB (e.g., changes to the study procedures, consent process, data protection measures). The submission of this form does not necessarily indicate wrong-doing; however follow-up procedures may be required.
- An **Adverse Events Report (AER)** must be submitted when adverse or unanticipated events occur to participants in the course of the research process.
- A **Protocol Completion Report** (PCR) is required when research using the protocol has been completed.
- If your research is funded by a third party, please contact the assigned Research Funding Officer in Research Services to ensure that your funds are released.

Best wishes for the successful completion of your research.

		Protocol #:55853			
Status:Delegated Review App	Version:0001	Sub Version:0000	Approved On:11-Feb-25	Expires On:25-Feb-26	Page 8 of 8

OFFICE OF RESEARCH ETHICS
McMurrich Building, 12 Queen's Park Crescent West, 2nd Floor, Toronto, ON M5S 1S8 Canada
Tel: +1 416 946-3273 • Fax: +1 416 946-5763 • ethics.review@utoronto.ca • http://www.research.utoronto.ca/for-researchers-administrators/ethics

Appendix B – The University of Toronto Research Ethics Board approval of Visual Feature Analysis Study (Chapter 4)

UNIVERSITY OF TORONTO

OFFICE OF THE VICE-PRESIDENT, RESEARCH AND INNOVATION

RIS Protocol Number: 47781

Approval Date: 11-Feb-25

PI Name: Tim Bressmann

Division Name:

Dear Tim Bressmann:

Re: Your research protocol application entitled, "Improving the Assessment of Oral-Nasal Balance Disorders in Patients with Cleft Palate: A Pilot Study on Classifying Oral-Nasal Balance Disorders through Visual Feature Analysis in Stereo Signals"

The Health Sciences REB has conducted a Delegated review of your application and has granted approval to the attached protocol for the period 2025-02-11 to 2026-02-10.

This approval covers the ethical acceptability of the human research activity; please ensure that all other approvals required to conduct your research are obtained prior to commencing the activity.

Please be reminded of the following points:

- An **Amendment** must be submitted to the REB for any proposed changes to the approved protocol. The amended protocol must be reviewed and approved by the REB prior to implementation of the changes.
- An annual **Renewal** must be submitted for ongoing research. Renewals should be submitted between 15 and 30 days prior to the current expiry date.
- A **Protocol Deviation Report** (PDR) should be submitted when there is any departure from the REB-approved ethics review application form that has occurred without prior approval from the REB (e.g., changes to the study procedures, consent process, data protection measures). The submission of this form does not necessarily indicate wrong-doing; however follow-up procedures may be required.
- An **Adverse Events Report (AER)** must be submitted when adverse or unanticipated events occur to participants in the course of the research process.
- A **Protocol Completion Report** (PCR) is required when research using the protocol has been completed.
- If your research is funded by a third party, please contact the assigned Research Funding Officer in Research Services to ensure that your funds are released.

Best wishes for the successful completion of your research.

		Protocol #:52740			
Status:Delegated Review App	Version:0002	Sub Version:0000	Approved On:11-Feb-25	Expires On:10-Feb-26	Page 9 of 9

OFFICE OF RESEARCH ETHICS
McMurrich Building, 12 Queen's Park Crescent West, 2nd Floor, Toronto, ON M5S 1S8 Canada
Tel: +1 416 946-3273 • Fax: +1 416 946-5763 • ethics.review@utoronto.ca • http://www.research.utoronto.ca/for-researchers-administrators/ethics

Appendix C – The University of Ottawa Research Ethics Board approval of Visual Feature Analysis Study (Chapter 4)

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

H-02-25-10552 - OTH-10552 - Certificat d'approbation éthique / Certificate of Ethics Approval

(English message follows)

Cher/Chère Fatemeh Abnavi,

Veuillez trouver ci-joint le certificat d'approbation éthique pour le projet intitulé «Improving the Assessment of Oral-Nasal Balance Disorders in Patients with Cleft Palate: A Pilot Study on Classifying Oral-Nasal Balance Disorders through Visual Feature Analysis in Stereo Signals».

Le certificat est valide jusqu'au : 25-02-2026

The expiry date is matched with the one on the ethics certificate from University of Toronto Research Ethics Board.

I wish you pleasant and fruitful research activities.

Recherche financée : veuillez faire suivre une copie du certificat au Service de gestion de la recherche.

Si vous avez des questions, n'hésitez pas à communiquer avec le Bureau d'éthique à ethique@uottawa.ca ou en composant le 613-562-5387.

Vous pouvez voir votre demande en vous connectant à votre compte eReviews.

Cordialement,

Germain Zongo
Responsable d'éthique en recherche

Ceci est une réponse automatisée, merci de ne pas répondre à ce courriel.

Dear Fatemeh Abnavi,

Please find attached the certificate of ethics approval for your research project titled "Improving the Assessment of Oral-Nasal Balance Disorders in Patients with Cleft Palate: A Pilot Study on Classifying Oral-Nasal Balance Disorders through Visual Feature Analysis in Stereo Signals".

This certificate is valid until: 25-02-2026

The expiry date is matched with the one on the ethics certificate from University of Toronto Research Ethics Board.

I wish you pleasant and fruitful research activities.

Funded research: A reminder that you must provide a copy of this certificate to Research Management Services.

If you have any questions, please contact the Ethics Office at ethics@uottawa.ca or by telephone at 613-562-5387.

You can view your project at any time by logging into eReviews.

Best regards,

Germain Zongo
Protocol Officer

550, rue Cumberland, pièce 154 | 550 Cumberland Street, Room 154
Ottawa (Ontario) K1N 6N5 Canada | Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics

Appendix D – Recruitment Emails for raters (Licensed SLPS and Student SLPS) -Visual Feature Analysis Study (Chapter 4)

E-mail to professional Speech-Language Pathologists

Subject: Invitation to Participate in a Brief Online Experiment on Oral-Nasal Balance Disorders

Hello,

My name is Fatemeh Abnavi, and I am a PhD student in Rehabilitation Sciences. We are conducting a joint research study with the University of Toronto and the University of Ottawa and would like to invite you, a professional speech-language pathologist specializing in the care of patients with cleft lip and palate, to participate in a brief online experiment. This study aims to explore how well speech-language pathologists can classify oral-nasal balance disorders based on visual analysis of stereo sound signals, to determine whether visual analysis can complement auditory-perceptual assessments.

The experiment includes consent, a short demographics survey, a training phase, a practice phase involving hands-on classification of four disorders with feedback, and an experimental phase where you will analyze 20 pairs of images. The entire process takes approximately 10–15 minutes.

Participation is voluntary, and you may withdraw at any time. You will not receive any reimbursement for your participation.

If you have any questions, please feel free to contact me at You can also contact the study supervisors, Dr. Tim Bressmann and Dr. Heather Flowers

Your participation would be greatly appreciated.

Best regards,
Fatemeh Abnavi
PhD Student, University of Ottawa

E-mail to students of Speech-Language Pathology

Subject: Invitation to Participate in a Brief Online Experiment on Oral-Nasal Balance Disorders

Hello,

My name is Fatemeh Abnavi, and I am a PhD student in Rehabilitation Sciences. We are conducting a joint research study with the University of Toronto and the University of Ottawa and would like to invite you, a student of Speech-Language Pathology, to participate in a brief online experiment. This study aims to explore how well students of Speech-Language Pathology can classify oral-nasal balance disorders based on visual analysis of stereo sound signals, to determine whether visual analysis can complement auditory-perceptual assessments.

The experiment includes consent, a short demographics survey, a training phase, a practice phase involving hands-on classification of four disorders with feedback, and an experimental phase where you will analyze 20 pairs of images. The entire process takes approximately 10–15 minutes.

Participation is voluntary, and you may withdraw at any time. You will not receive any reimbursement for your participation.

If you have any questions, please feel free to contact me at You can also contact the study supervisors, Dr. Tim Bressmann and Dr. Heather Flowers

Your participation would be greatly appreciated.

Best regards,
Fatemeh Abnavi
PhD Student, University of Ottawa

Reminder e-mail to professional Speech-Language Pathologists

Subject: Reminder: Participation in Research Study

Hello,

I'm following up on my previous email regarding our research study on how SLPs classify oral-nasal balance disorders using visual analysis of stereo images. If you haven't yet participated, we kindly encourage you to complete the brief 10–15-minute experiment.

Your participation would be greatly appreciated. If you have any questions, feel free to contact me at or the study supervisors, Dr. Tim Bressmann and Dr. Heather Flowers

Thank you!

Best regards,
Fatemeh Abnavi
PhD Student, University of Ottawa

Appendix E – Online Questionnaire and Consent Form for Visual Feature Analysis Study (Chapter 4)

A pilot study on classifying ōral-nasal balance disorders through visual analysis

We invite you to participate in a short online experiment. This experiment is designed to evaluate the effectiveness of visual analysis for the classification of oral-nasal balance disorders. Your help is greatly appreciated.

Study Purpose: This study, conducted by the University of Ottawa and the University of Toronto, investigates how speech-language pathologists (SLPs) and SLP students classify oral-nasal balance disorders (hypernasality, hyponasality, mixed nasality, and normal) based only on visual analysis. This study will tell us whether visual analysis of stereo signals could serve as a complement to auditory-perceptual assessments of oral-nasal balance disorders.

Experiment Format :The experiment includes the following sections:

1. **Consent**
2. **Demographics**
3. **Training Phase**
4. **Practice Phase**
5. **Experimental Phase**

Confidentiality: Participation in this experiment is voluntary. Your responses will remain anonymous and confidential, with no identifiable information collected. This experiment is online. No conceivable harm is anticipated from participation. If you have started the experiment and you do not wish to continue, you can stop at any time.

Time Commitment:

- **Training and practice**: 5 minutes.
- **Experimental phase**: 5-10 minutes.

Contact Information:

If you have any questions or concerns, please contact:

- **Fatemeh Abnavi**, PhD Student, University of Ottawa,
- **Dr. Tim Bressmann**, Study Supervisor, University of Toronto,
- **Dr. Heather Flowers**, Study Supervisor, University of Ottawa,

For questions about your rights as a participant, please contact the **Office of Research Ethics** at ethics.review@utoronto.ca or call 416-946-3273.

Important: Please ensure you complete the experiment in a single session. Your progress will not be saved and you cannot finish an incomplete experiment later.

Required

Consent

1

I have read and understood the study details and voluntarily consent to participate in this research. *

◯ I have read this section

2

By clicking 'yes' below, I agree to participate in this experiment. *

◯ Yes, I want to participate.

Demographics

Please read the following information and check the appropriate boxes:

3

What is your age? *

4

What is your gender identity? *

- ◯ Female
- ◯ Male
- ◯ Other identity
- ◯ Prefer not to say

5

Are you an SLP student or are you a licensed SLP? *

- ◯ Registered SLP
- ◯ SLP Student (I have not taken a course on cleft lip and palate, yet)
- ◯ SLP Student (I have already taken a course on cleft lip and palate)

Demographics

Please read the following information and check the appropriate boxes:

3

What is your age? *

4

What is your gender identity? *

- ◯ Female
- ◯ Male
- ◯ Other identity
- ◯ Prefer not to say

5

Are you an SLP student or are you a licensed SLP? *

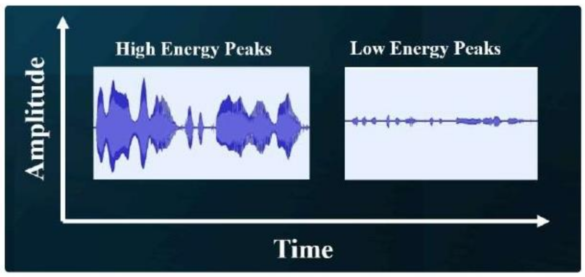
- ◯ Registered SLP
- ◯ SLP Student (I have not taken a course on cleft lip and palate, yet)
- ◯ SLP Student (I have already taken a course on cleft lip and palate)

Training Phase

When speech is recorded, it is shown as an **oscillogram** on the screen. Higher peaks in the oscillogram indicate a louder signal. Lower peaks show a quieter signal. Let's look at an example:

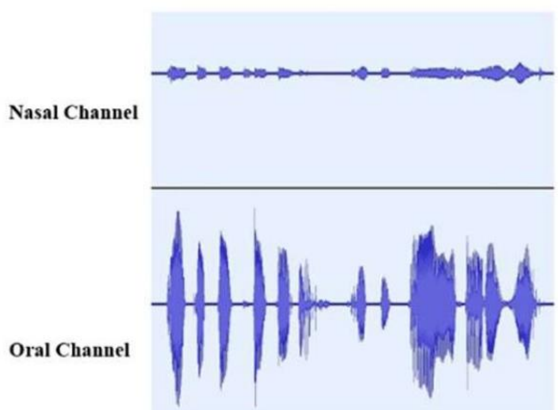
6

*



◯ I have read this section

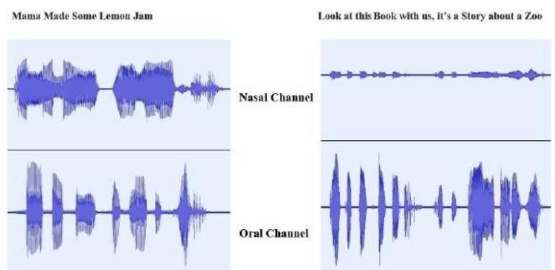
We can use a headset with a sound separation palate to record the oral and nasal parts of the speech signal separately. In the images below, the upper channel show the **nasal speech signal** and the lower channel shows the **oral speech signal**. *



◯ I have read this section

8

When a typical speaker says a sentence with many nasal sounds, we see a loud signal in the **nasal channel**. For a non-nasal sentence, we see only a very quiet signal in the **nasal channel**. This is how the two sentences should look. *

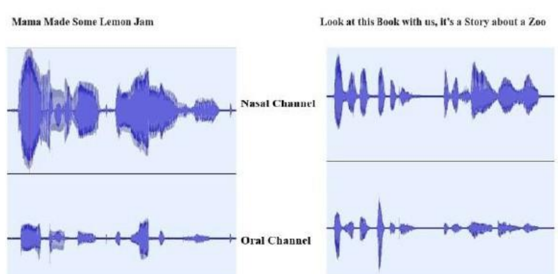


◯ I have read this section

9

Now, let's discuss how we can use oscillograms to detect different types of nasality disorders:

In hypernasality, you will see **energy peaks** in the nasal channel for the **nasal sentence** (as expected). But if you look at the **non-nasal sentence**, there is **too much energy** in the **nasal channel**. This shows excessive nasality. *



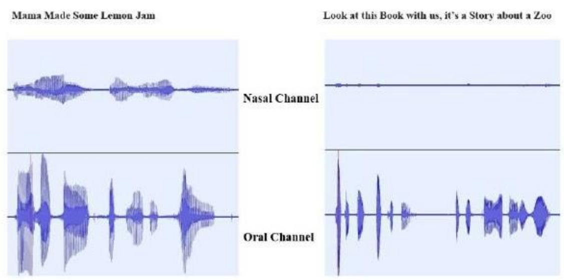
◯ I have read this section

Practice Phase

Please practice classifying oral-nasal balance conditions using visual analysis. There are 4 conditions: **hypernasality, hyponasality, mixed nasality, or normal**. For each image, select the condition that you think best matches it. During this practice phase, you will receive feedback whether your answer was correct, along with a brief explanation.

12

What condition does this picture represent? *



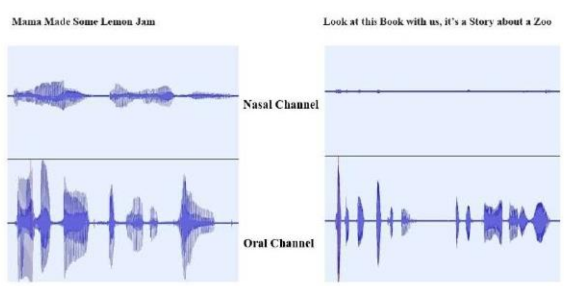
◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**

◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**

◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**

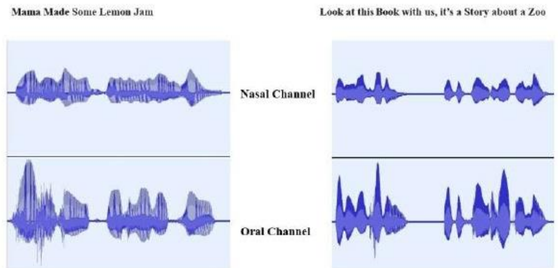
◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

Correct Answer: Hyponasality
In this image, there is little energy in the nasal channel for the nasal sentence, indicating that the nose is blocked. The non-nasal sentence looks normal in the nasal channel. *



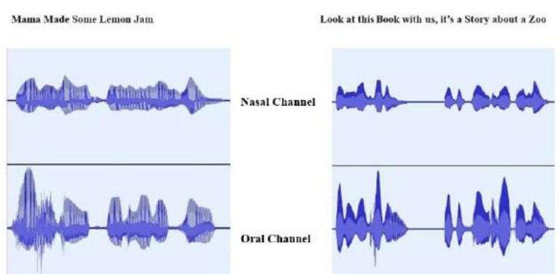
◯ I have read this section.

What condition does this picture represent? *



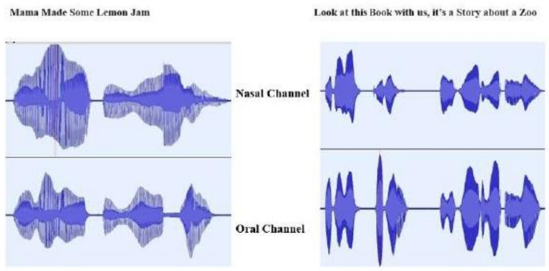
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

Correct Answer: Mixed Nasality In this image, reduced energy peaks in the nasal channel for the nasal sentence indicate some hyponasality. However, for the non-nasal sentence, excessive energy in the nasal channel suggests hypernasality. This combination demonstrates mixed nasality. *



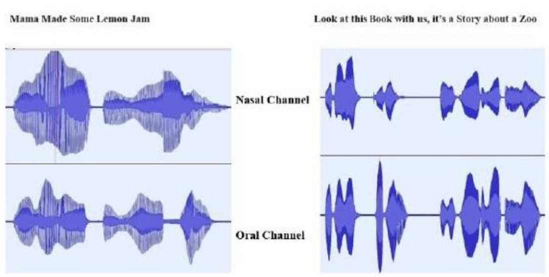
- ◯ I have read this section

What condition does this picture represent? *



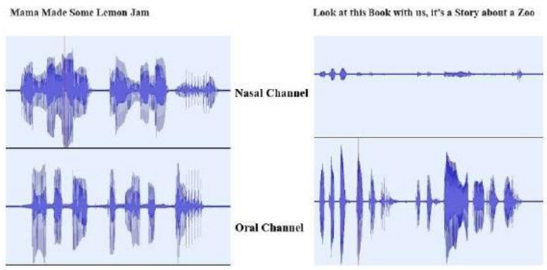
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

Correct Answer: Hypernasality
In this image, you can see energy peaks in the nasal channel for the nasal sentence (as expected). But if you look at the non-nasal sentence, there is too much energy in the nasal channel. This shows excessive nasality. *



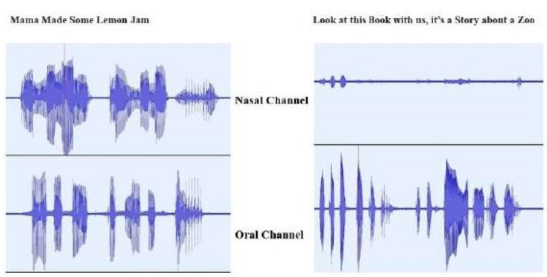
◯ I have read this section

What condition does this picture represent? *



- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

Correct Answer: Normal
In this image, you can see a loud signal for the nasal sentence in the nasal channel. For the non-nasal sentence, there is only a very quiet signal in the nasal channel. *



◯ I have read this section

20

Were you able to spot the different nasality conditions in the examples above? If not, have another look. When you are ready, please start the experiment in the next section. *

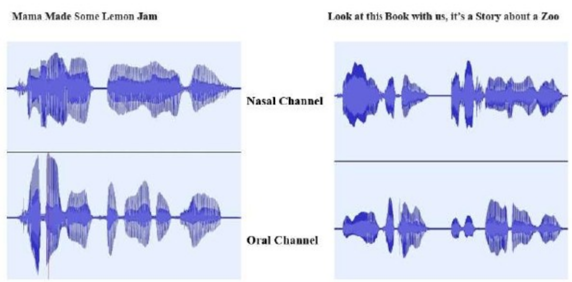
◯ I have read this section

Experimental Phase

Welcome to the experimental phase! You will see 20 image pairs. For each image, please select the answer that best matches what you observe.

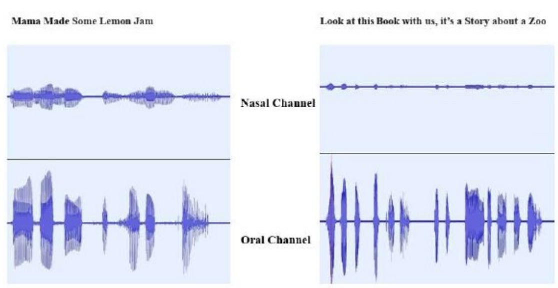
21

What condition does this picture represent? *

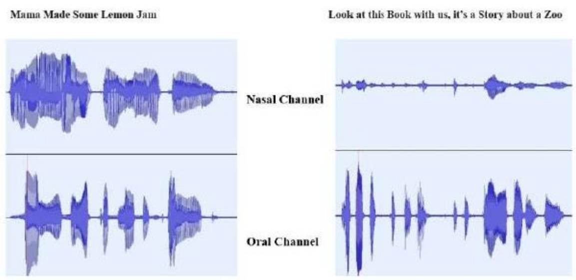


- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *

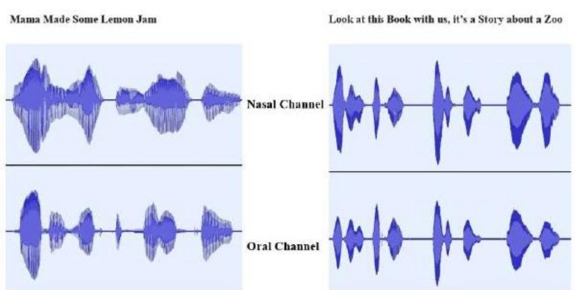


- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**



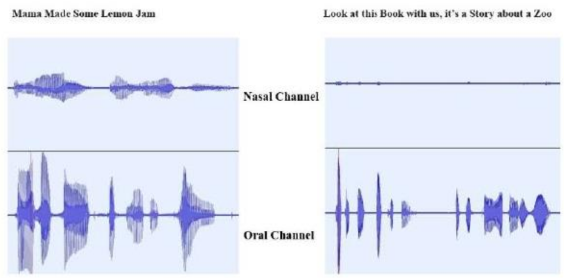
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *



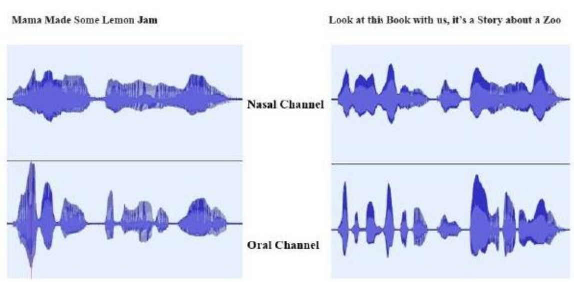
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *



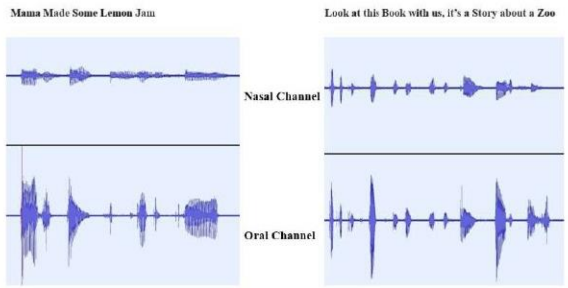
- Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

What condition does this picture represent? *



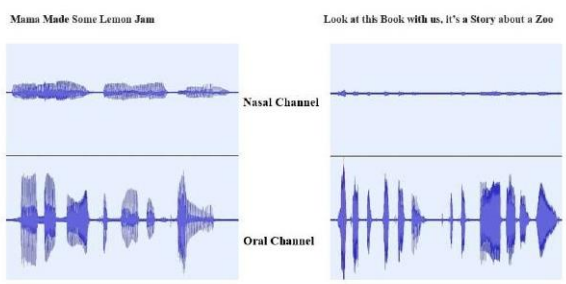
- Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

What condition does this picture represent? *



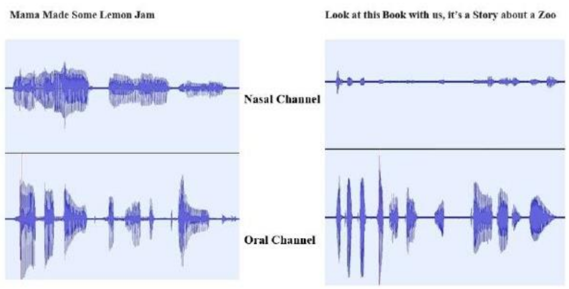
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *



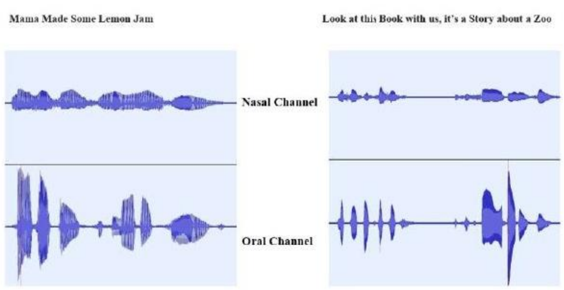
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *



- Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

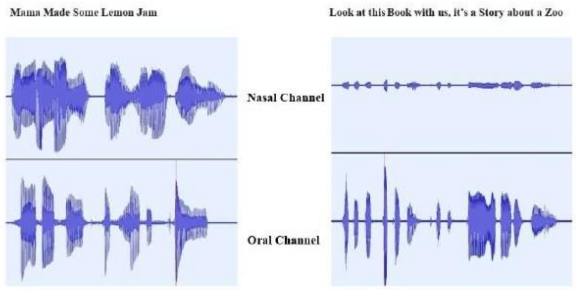
What condition does this picture represent? *



- Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

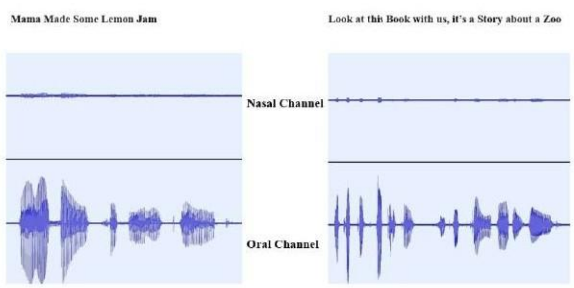
31

What condition does this picture represent? *



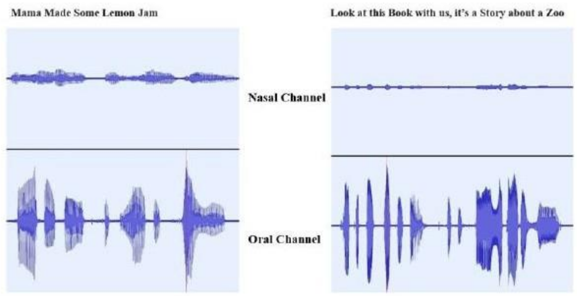
- ◯ Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- ◯ Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- ◯ Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- ◯ Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

What condition does this picture represent? *



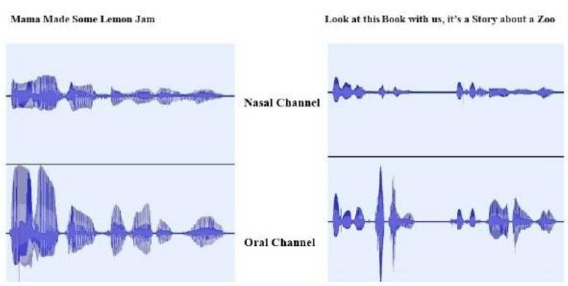
- ◯ Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- ◯ Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- ◯ Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- ◯ Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

What condition does this picture represent? *



- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

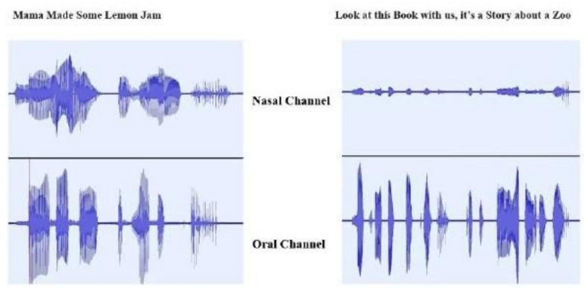
What condition does this picture represent? *



- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

35

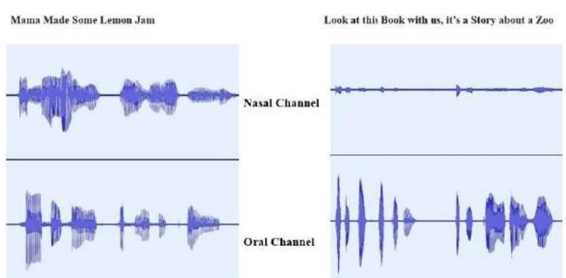
What condition does this picture represent? *



- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

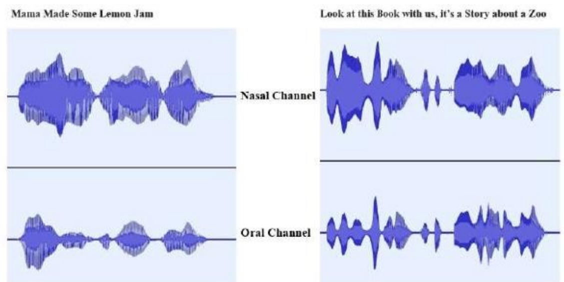
36

What condition does this picture represent? *



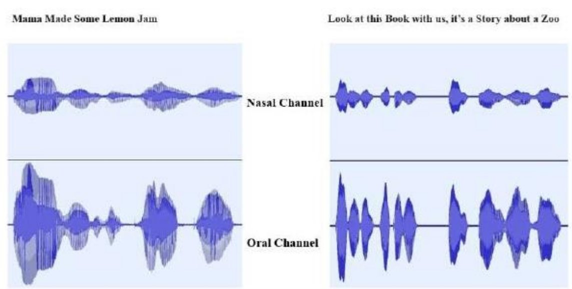
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *



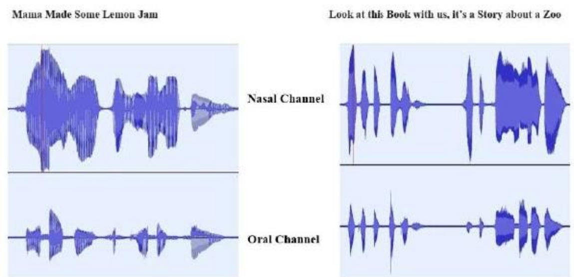
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *



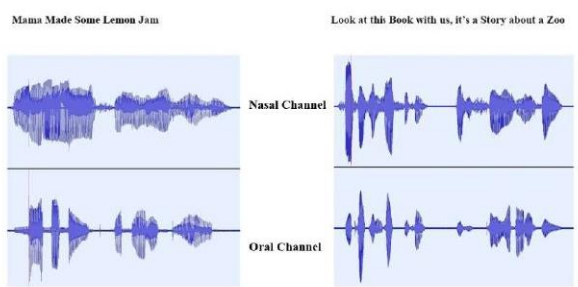
- ◯ **Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)**
- ◯ **Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)**
- ◯ **Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)**

What condition does this picture represent? *



- Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

What condition does this picture represent? *



- Normal (there is nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Hypernasality (there is nasality for the nasal sentence, and there is too much nasality for the non-nasal sentence)
- Hyponasality (there is too little nasality for the nasal sentence, and there is no nasality for the non-nasal sentence)
- Mixed Nasality (there is reduced nasality for the nasal sentence, and there is increased nasality for the non-nasal sentence)

Appendix E - Figure Permissions Summary

Figure Number	Figure Title	Source	Permission Type
Figure 1.1	Schematic Representation of the Source–Filter Theory of Speech Production	Tokuda. (2021)	Reproduced with permission from the author and Oxford University Press
Figure 1.2	Sagittal view of the anatomy of the velopharyngeal sphincter	Anderson et al. (2019)	Adapted under CC BY 2.5, http://creativecommons.org/licenses/by/2.5], via Wikimedia Commons
Figure 1.3	Schematic view of velopharyngeal dysfunction (VPD)	Sanu, P. M. (2018)	Reproduced under Creative Commons Attribution (CC BY 4.0) license
Figure 1.4	Waveform (top) and spectrogram (bottom) of a nasalized utterance	Ortega-Llebaria, M., & Prieto, P. (2008)	Reproduced with permission from the author
Figure 1.6	Videonasoendoscopy (VNE) procedure	https://teachmesurgery.com/consent/ent/consent-flexible-nasal-endoscopy/	Reproduced under the Creative Commons Attribution (CC BY 4.0) license
Figure 1.7	Lateral view videofluoroscopy showing the velum (soft palate) during speech	Ysunza et al. (2016)	Reproduced with permission from the author
Figure 1.8	MRI Comparison of Velar Length and Thickness in Two Individuals	Kollara et al. (2017)	Reproduced under Creative Commons Attribution (CC BY 4.0) license
Figure 1.10	Spectra of vowels /a/, /i/, and /u/ for normal, mild, and moderate–severe hypernasality signals	Nikitha et al. (2017)	Reproduced with permission from the author
Figure 1.11	Spectra for an oral vowel (word “grade”) and a nasalized vowel (word “grain”) in a typical speaker	Tamminga & Zellou (2015)	Reproduced with permission from the author

Figure 1.12	Schematic view of the separation plate and nasal and oral microphones of the nasometer	Simpson, A. P. (2012).	Adapted with permission from the author and the Journal of Phonetics (Elsevier)

Figure 1.1. Oxford University Press License Agreement

PLS Clear

PARTIES:

1. **Oxford Publishing Limited** (Licensor); and
2. **Fatemeh Abnavi** (Licensee).

Thank you for your recent permission request. Some permission requests for use of material published by the Licensor, such as this one, are now being facilitated by PLSclear.

Set out in this licence cover sheet (the **Licence Cover Sheet**) are the principal commercial terms under which Licensor has agreed to license certain Licensed Material (as defined below) to Licensee. The terms in this Licence Cover Sheet are subject to the attached General Terms and Conditions, which together with this Licence Cover Sheet constitute the licence agreement (the **Licence**) between Licensor and Licensee as regards the Licensed Material. The terms set out in this Licence Cover Sheet take precedence over any conflicting provision in the General Terms and Conditions.

Licence Terms

Licence Date:	13/06/2025
PLSclear Ref No:	106801

The Licensor

Company name:	Oxford Publishing Limited
Address:	Rights Department Great Clarendon Street Oxford OX2 6DP GB

The Licensee

Licensee Contact Name:	Fatemeh Abnavi
Licensee Address:	[redacted]

Licensed Material

title:	The Source-Filter Theory of Speech
ISBN/ISSN:	10.1093/acrefore/9780199384655.013.894
publisher:	Oxford Publishing Limited

Issue Date	29 November 2021
Are you requesting permission to reuse the cover of the publication?	No
Article Title	The Source–Filter Theory of Speech
Article URL	https://doi.org/10.1093/acrefore/9780199384655.013.894
Figure number & title	Figure 1. Concept of the source-filter theory. Airflow from the lung induces vocal fold vibrations, where glottal source sound is created. The vocal tract filter shapes the spectral structure of the source sound. The filtered speech sound is finally
Are you the author of the content that you are requesting to reuse?	No
Is the content you are requesting available through an Open Access Licence?	Yes
Will you be changing or editing the image?	No

For Use In Licensee's Publication(s)

usage type	Book, Journal, Magazine or Academic Paper-Thesis or Dissertation
Will your dissertation be placed in an online repository?	Yes
Author	Fatemeh Abnavi
Language	English
Title of dissertation/thesis	Complementary Approaches for Improving Nasality Assessment in Cleft Palate Patients: Exploring Nasal Signal-Based Acoustic Feature and Visual Feature Analysis in Stereo Signals
University or institution	University of Ottawa
Unlimited circulation?	No

Rights Granted

Exclusivity:	Non-Exclusive
Format:	Thesis/Dissertation
Language:	English
Territory:	World
Duration:	Lifetime of Licensee's edition
Maximum Circulation:	Maximum print circulation: 1 copies
Additional Terms:	- If at some future date your thesis is published it will be necessary to re-clear this permission. Please also note that if the material to be used is acknowledged to any other source, you will need to clear permission with the rights holder, and for any electronic version the © line must appear on the same page as the OUP material and the OUP material should not be included under a Creative Commons license, or any other open-access license allowing onward reuse.

Payment Details

Fee Payable:	£0.00 [+ VAT if applicable]
Payment Terms:	Strictly 30 days from date of Licence

GENERAL TERMS AND CONDITIONS

1. Definitions and Interpretation

1.1 Capitalised words and expressions in these General Terms and Conditions have the meanings given to them in the Licence Cover Sheet.

1.2 In this Licence any references (express or implied) to statutes or provisions are references to those statutes or provisions as amended or re-enacted from time to time. The term **including** will be construed as illustrative, without limiting the sense or scope of the words preceding it. A reference to in **writing** or **written** includes faxes and email. The singular includes the plural and vice versa.

2. Grant of Rights

2.1 Subject to payment by Licensee of the Licence Fee in accordance with paragraph 3 below, Licensor grants to Licensee the non-exclusive right to use the Licensed Material as specified in the Licence Cover Sheet.

2.2 The rights licensed to Licensee under this Licence do not include the right to use any third party copyright material incorporated in the Licensed Material. Licensee should check the Licensed Material carefully and seek permission for the use of any such third party copyright material from the relevant copyright owner(s).

2.3 Unless otherwise stated in the Licence Cover Sheet, the Licensed Material may be:

2.3.1 subjected to minor editing, including for the purposes of creating alternative formats to provide access for a beneficiary person (provided that any such editing does not amount to derogatory treatment); and/or

2.3.2 used for incidental promotional use (such as online retail providers' search facilities).

2.4 Save as expressly permitted in this Licence or as otherwise permitted by law, no use or modification of the Licensed Material may be made by Licensee without Licensor's prior written permission.

3. Payment

3.1 Licensee must pay to Licensor the Licence Fee by means of either credit card or on receipt of an invoice, as selected by Licensee during the licence application process via the PLSclear service.

3.2 If payment is by invoice, Licensee agrees to pay the Licence Fee in full by no later than the payment date specified in the relevant invoice.

4. Copyright Notice and Acknowledgement

4.1 Licensee must ensure that the following notices and acknowledgements are reproduced prominently alongside each reproduction by Licensee of the Licensed Material:

4.1.1 the title and author of the Licensed Material;

4.1.2 the copyright notice included in the Licensed Material; and

4.1.3 the statement "Reproduced with permission of the Licensor through PLSclear."

5. Reversion of Rights

5.1 The rights licensed to Licensee under this Licence will terminate immediately and automatically upon the earliest of the following events to occur:

5.1.1 the Licence Fee not being received by Licensor in full by the payment date specified in the relevant invoice;

5.1.2 the Licensed Material not being used by Licensee within 18 months of the Licence Date;

5.1.3 expiry of the Licence Duration; or

5.1.4 the Maximum Circulation being reached.

6. Miscellaneous

6.1 By using the Licensed Material, Licensee will be deemed to have accepted all the terms and conditions contained in this Licence.

6.2 This Licence contains the entire understanding and agreement of the parties relating to its subject matter and supersedes in all respects any previous or other existing arrangements, agreements or understandings between the parties whether oral or written in relation to its subject matter.

6.3 Licensee may not assign this Licence or any of its rights or obligations hereunder to any third party without Licensor's prior written consent.

6.4 This Licence is governed by and shall be construed in accordance with the laws of England and Wales and the parties hereby irrevocably submit to the non-exclusive jurisdiction of the Courts of England and Wales as regards any claim, dispute or matter arising under or in relation to this Licence.

Figure 1.10. Elsevier License Agreement

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jun 12, 2025

This Agreement between Fatemeh Abnavi ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	6046690963274
License date	Jun 12, 2025
Licensed Content Publisher	Elsevier
Licensed Content Publication	Journal of Phonetics
Licensed Content Title	The first and second harmonics should not be used to measure breathiness in male and female voices
Licensed Content Author	Adrian P. Simpson
Licensed Content Date	May 1, 2012
Licensed Content Volume	40
Licensed Content Issue	3
Licensed Content Pages	14
Start Page	477
End Page	490

PLS Clear

PARTIES:
1. **Oxford Publishing Limited** (Licensor); and
2. **Fatemeh Abnavi** (Licensee).

Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of new work	Complementary Approaches for Improving Nasality Assessment in Cleft Palate Patients: Exploring Nasal Signal-Based Acoustic Feature and Visual Feature Analysis in Stereo Signals
Institution name	University of Ottawa
Expected presentation date	Sep 2025
Portions	Fig. 3
The Requesting Person / Organization to Appear on the License	Fatemeh Abnavi
Requestor Location	Ms. Fatemeh Abnavi [redacted] Toronto, [redacted] Canada [redacted]

Total 0.00 CAD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your RightsLink account and that are available at any time at https://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given. The material may not be reproduced or used in any other way, including use in combination with an artificial intelligence tool (including to train an algorithm, test, process, analyse, generate output and/or develop any form of artificial intelligence tool), or to create any derivative work and/or service (including resulting from the use of artificial intelligence tools).

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk here). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional

translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use

    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please

reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.10

Questions? customercare@copyright.com.