



**HAL**  
open science

# Dynamic modeling of the velopharyngeal trajectory in Brazilian Portuguese nasal diphthongs.

Rita de Cassia Demasi

► **To cite this version:**

Rita de Cassia Demasi. Dynamic modeling of the velopharyngeal trajectory in Brazilian Portuguese nasal diphthongs.. Linguistics. Université de la Sorbonne nouvelle - Paris III, 2023. English. NNT : 2023PA030024 . tel-04582825

**HAL Id: tel-04582825**

**<https://theses.hal.science/tel-04582825v1>**

Submitted on 22 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Sorbonne Nouvelle University – Paris 3

Doctoral School 622 – « Language Science »  
Phonetics and Phonology Laboratory  
(UMR7018, Sorbonne-Nouvelle/CNRS)

*Doctoral thesis in Phonetics, Phonology, and Speech Sciences  
to obtain the degree of Ph.D.*

Rita De Cassia DEMASI

---

## Dynamic Modeling of the Velopharyngeal Trajectory in Brazilian Portuguese Nasal Diphthong

---

Under the supervision of M. Professor Ph.D. Didier DEMOLIN

Presented and defended on [April 24<sup>th</sup>, 2023]

**Committee in Charge:**

M. Professor Ph.D. Didier DEMOLIN, Université Sorbonne Nouvelle, Paris 3, (Adviser)

M. Professor Ph.D. Rudolph SOCK, Université de Strasbourg (Rapporteur)

Ms. Professor Ph.D. Eleonora ALBANO, University Public of Campinas (Rapporteur)

Ms. Ph.D. Angélique AMELOT, Université Sorbonne Nouvelle, Paris 3, (Examiner)

Ms. Ph.D. Solange ROSATO, Université Grenoble Alpes, (Examiner)

Ms. Professor Ph.D. Alice TURK, University of Edinburgh, (Examiner)

---



Université Sorbonne Nouvelle – Paris 3

École Doctorale 622 – « Sciences du Langage »  
Laboratoire de Phonétique et Phonologie  
(UMR7018, Sorbonne-Nouvelle/CNRS)

Thèse de Doctorat en Phonétique, Phonologie et Sciences de  
la Parole pour obtenir le grade de Docteur

Rita De Cassia DEMASI

---

La modélisation dynamique de la trajectoire vélo-pharyngée  
dans les diphtongues nasales du Portugais Brésilien

---

Sous la direction de M. le Professeur Dr. Didier DEMOLIN

Présentée et soutenue publiquement le [24 avril 2023]

**Composition du Jury :**

M. le Professeur Dr. Didier DEMOLIN, Université Sorbonne Nouvelle, Paris 3, (Directeur)

M. le Professeur Dr. Rudolph SOCK, Université de Strasbourg (Rapporteur)

Mme. la Professeur Dr. Eleonora ALBANO, Université Publique de Campinas (Rapporteuse)

Mme. Dr. Angélique AMELOT, Université Sorbonne Nouvelle, Paris 3, (Examinatrice)

Mme. Dr. Solange ROSATO, Université Grenoble Alpes, (Examinatrice)

Mme. la Professeur Dr. Alice TURK, Université de Edinburgh, (Examinatrice)

---



Universidade Sorbonne Nouvelle – Paris 3

Escola de Doutorado 622– « Ciências da Linguagem »  
Laboratório de Fonética e Fonologia  
(UMR7018, Sorbonne-Nouvelle/CNRS)

Tese em Fonética, Fonologia e Ciências da Fala submetida à  
obtenção do título de Doutora

Rita De Cassia DEMASI

---

A MODELIZAÇÃO DINÂMICA DA TRAJETÓRIA  
VELOFARÍNGEA NOS DITONGOS NASAIS DO  
PORTUGUÊS BRASILEIRO

---

Sob supervisão do Sr. Professor Dr. Didier DEMOLIN

Apresentada e defendida em [24 de abril de 2023]

**Júri:**

Sr. Professor Dr. Didier DEMOLIN, Universidade Sorbonne Nouvelle, Paris 3, (Orientador)  
Sr. Professor Dr. Rudolph SOCK, Universidade de Strasbourg (Parecerista)  
Sra. Professora Dra. Eleonora ALBANO, Universidade Estadual de Campinas (Parecerista)  
Sra. Dra. Angélique AMELOT, Universidade Sorbonne Nouvelle, Paris 3, (Examinadora)  
Sra. Dra. Solange ROSATO, Universidade Grenoble Alpes, (Examinadora)  
Sra. Professora Dra. Alice TURK, Universidade de Edinburgh, (Examinadora)

---

## ABSTRACT

---

The objective of this Ph.D. research is to understand the characteristics of nasal diphthong and front nasal vowel production in the Paulistano dialect of Brazilian Portuguese through experimental phonetic and phonological analysis. In this particular case, the nasalization of speech sounds involves a phonetic process known as nasal diphthongization. This allophonic process alters the quality of vowels and nasal diphthongs. To illustrate this phenomenon, we compare the pronunciation of the diphthong 'ão' [ãw̃] with its oral equivalent 'au' [aw]. Similarly, we examine the production of the diphthongized nasal vowel /ẽ/, represented as [ẽ̃], and the production of the oral diphthong [ej]. The study comprises an aerodynamic analysis (pneumotachograph), a 2D electro-articulography (electromagnetic midsagittal articulography) study, and a nasal fiberscope study. A corpus of 20 words repeated three times by native speakers of the São Paulo dialect was created. The general results suggest that nasal diphthongs are a co-articulatory phenomenon involving gestural assimilation, leading to the emergence of unspecified nasal consonants called nasal glides. This phenomenon is also associated with the formation of nasal appendixes. The articulatory process alters the quality of resonance influenced by tongue movements associated with the opening and closing of the velum. As a result, the synchronization of tongue movement with soft palate movement causes nasal gliding, which propagates its resonance and creates nasal appendixes. Depending on the degree of co-articulation of the nasal vowels, either velar or palatal constriction may occur. Consequently, we conclude that nasal diphthongization modifies the geometry of the vocal tract, not only changing the shape of the vocal tract but also the position and type of tongue articulation. This complex articulation leads to the emergence of narrowed vowel segments that sound more consonant-like. Based on these findings, we can consider nasal diphthongization as a transitional phenomenon that plays a phonetic role in enhancing perception.

**KEYWORDS:** Brazilian Portuguese, Nasal Diphthongization, Experimental Phonetics and Phonology and Speech Timing.

## RÉSUMÉ

---

L'objectif de cette recherche doctorale est de comprendre, d'un point de vue expérimental, les caractéristiques de la production de la diphtongue nasale postérieure et de la voyelle antérieure nasale en portugais brésilien à l'aide d'analyses phonétiques et phonologiques. La nasalisation de la voix est affectée par un processus phonétique appelé diphtongaison nasale. Ce processus allophonique est un phénomène qui modifie la qualité des voyelles et des diphtongues nasales. Pour démontrer ce phénomène, nous comparerons la prononciation de la diphtongue « ão » [ãõ] avec celle du phonème équivalent oral « au » [aw]. De même, nous comparerons la production de la voyelle nasale diphtonguée / ã /, réalisée comme [ẽ̃], avec celle de la diphtongue orale [ej]. Nous avons créé un corpus de 20 mots répétés trois fois par des locuteurs natifs du dialecte de São Paulo. Afin de présenter les caractéristiques générales de la production des voyelles nasales et des diphtongues nasales ainsi que leurs mécanismes articulatoires, nous avons réalisé des études aérodynamiques, d'imagerie et articulatoires. Cette étude comprend une étude aérodynamique (pneumotachographe), une étude d'électro-articulographe 2D (articulographe électromagnétique sagittal médian) et une étude de fibroscope nasal. Les résultats généraux suggèrent que la diphtongaison nasale est un phénomène consonantique coarticulatoire d'assimilation gestuelle. Ceci est interprété comme une occurrence de consonnes nasales non spécifiées, appelées semi-voyelles/semi-consonnes nasales, résultant du contexte de l'émergence d'appendices nasaux. Ce processus altère la qualité de la résonance en raison de l'influence des mouvements de la langue associés à la descente et à la fermeture du voile du palais. Ainsi, la synchronisation des mouvements de la langue avec ceux du voile du palais se traduit par des semi-voyelles/semi-consonnes nasales, dont les résonances se propagent, générant l'appendice nasal et le processus de nasalisation. Selon le degré d'articulation des voyelles nasales, il peut y avoir une constriction vélaire ou palatine. De cela, nous concluons que la diphtongaison nasale modifie la géométrie du tube vocal, la position et le lieu d'articulation de la langue. Cette articulation complexe conduit à l'apparition de segments vocaliques. L'élévation des voyelles produit une transition vers une semi-voyelle/semi-consonne, avec l'émergence d'appendices nasaux homorganiques, rendant le segment sonore plus consonantique. À partir de ces résultats, nous pouvons considérer que la diphtongaison nasale est un phénomène transitionnel qui joue un rôle phonétique pour renforcer la perception.

**MOTS CLÉS : Portugais Brésilien, Diphtongaison nasale, Phonétique et Phonologie Expérimentale et Synchronisation de la parole.**

## RESUMO

---

O objetivo desta pesquisa de doutorado é o uso da análise Fonética e Fonológica para compreender as características de produção dos ditongos nasais e da vogal anterior nasal no português brasileiro, a partir de um viés experimental. No dialeto Paulistano, a nasalização vocálica passa por um processo fonológico denominado ditongação nasal. Esse processo alofônico é um fenômeno que pode alterar a qualidade das vogais e dos ditongos nasais. Para demonstrar esse fenômeno, compararemos a produção do ditongo “ão” [ãõ] e sua contraparte oral “au” [aw]. Da mesma forma, compararemos a produção da vogal nasal ditongada / ã /, produzida como [ẽ], com o ditongo oral [ej]. Montamos um corpus com vinte palavras que foram repetidas três vezes, por falantes nativos do dialeto Paulistano. Para apresentarmos as características gerais da produção da vogal nasal e do ditongo nasal, bem como seus mecanismos articulatórios, estabelecemos um estudo aerodinâmico, um estudo de imagem e um estudo articulatório. Esta pesquisa contém um estudo aerodinâmico (pneumatógrafo), um estudo eletro-articulógrafo 2D (articulógrafo eletromagnético mediano sagital) e um estudo com um naso-fibrocópio. Os resultados gerais sugerem que a ditongação nasal é um fenômeno co-articulatório de assimilação gestual resultante do surgimento de uma consoante nasal não especificada, chamada de glide nasal, associada ao contexto e ao surgimento do apêndice nasal. Esse processo altera a qualidade das ressonâncias, devido à influência do movimento da língua associado ao abaixamento e fechamento do velum. Consequentemente, o movimento da língua e sua sincronização com o movimento do véu palatino resultam no surgimento da glide nasal que espalha suas ressonâncias gerando o apêndice nasal. Dependendo do grau de coarticulação da vogal nasal, uma constrição velar ou palatina pode surgir. Assim, concluímos que a ditongação nasal altera a geometria do trato, o ponto e o modo de articulação da língua. Essa articulação complexa resulta no surgimento de um segmento vocálico constricto e o alçamento vocálico produz a propagação da glide, bem como o surgimento do apêndice nasal homorgânico, tornando o som mais consonantal. A partir desses resultados, podemos considerar que a ditongação nasal é um fenômeno de transição que desempenha um papel fonético para intensificar a percepção.

**PALAVRAS-CHAVE: Português Brasileiro, Ditongação Nasal, Fonética e Fonologia Experimental e Tempo de fala.**

*With love to Clarice, Otavio and Tom.*



## ACKNOWLEDGEMENTS

---

I would like to thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Process Number: BEX 0419-12-5), a Brazilian government institution, for financing 34 months of my research.

I am grateful to my adviser, Mr. Didier Demolin, for dedicating his time to teaching me throughout my entire post-graduate study; he has helped me to navigate the tricky waters of academic life since 2007. Thank you for sharing your valuable time and research, and for teaching me and developing my skills.

I would like to express my deepest gratitude to all the speakers who participated in the different experimental designs, with a special mention to Gabriela for her exceptional contributions. Their involvement was essential to the completion of this work.

I would also like to extend my thanks to all my colleagues and friends at Gipsa-Lab and Stendhal University. A heartfelt thanks goes out to Rosario for our valued friendship. To Yan Lu and Manon, I am grateful for everything that we shared in our office, and I would also like to thank Hien, Giovanni, and Sylvia for their support and assistance.

I would also like to express my gratitude to Nathalie Vallée and Solange Rossato for their warm welcome when I first visited Grenoble in 2009. Their kindness and generosity were greatly appreciated, and I am deeply grateful for their hospitality. I would like to express my gratitude to Gang Feng for introducing me to the fascinating world of acoustic engineering, as well as to Christophe Savariaux, the resident engineer at Gipsa-lab, for his guidance during the recording of the articulatory data that formed the foundation of this research. Without these crucial data, this thesis would not have been possible.

I am deeply thankful to the city of Grenoble and Stendhal University for the invaluable life experiences and opportunities that I have had during my time there. My gratitude knows no bounds.

would like to express my gratitude to the ILPGA at Université Sorbonne Nouvelle-Paris 3 for their warm welcome when I transferred to their new laboratory in the middle of my research. Their hospitality has been greatly appreciated, and I am grateful for the opportunity to complete my academic journey in this new environment.

Special thanks are due to Angélique Amelot, the resident engineer from ILPGA, for her invaluable help with the recording of the fiberscope data that were analyzed in this thesis. Her time and expertise were crucial to the success of this experiment. I would also like to extend my thanks to Shinji Maeda and Lise Crevier-Buchman for their contributions to that data collection.

I would also like to extend my thanks to Ms. Alexandra DURANT for her administrative assistance throughout this process. Her help has been greatly appreciated.

I would like to express my gratitude to my family, including my mother Clarice, my father Otavio, my son Tom, my husband Adriano, and my brothers João Otavio and João Paulo. Their unwavering emotional and financial support has been crucial to my success throughout the years of dedication to this study. I am forever grateful for their love, encouragement, and understanding during the ups and downs of this academic journey.

Finally, I am grateful to my friend Yaru Wu for her warm welcome, as well as to my colleagues and the ILPGA staff for the enriching experiences we shared.

I would like to express my gratitude to all the members of the jury, including my advisor, Mr. Didier DEMOLIN, Ms. Angélique AMELOT, Mr. Rudolph SOCK, Ms. Eleonora ALBANO, Ms. Alice TURK, and Ms. Solange ROSSATO. Their valuable insights and constructive feedback were instrumental in shaping my research and thesis.

***Je serai toujours reconnaissant envers la France.  
Vive la France !***

# CONTENTS

---

ABSTRACT.....	iv
RÉSUMÉ.....	v
RESUMO.....	vi
ACKNOWLEDGEMENTS.....	viii
CONTENTS.....	x
LIST OF FIGURES.....	xiii
LIST OF TABLES.....	xviii
LIST OF CHARTS.....	xx
LIST OF ABBREVIATIONS AND SYMBOLS.....	xxii
CHAPTER 1: BACKGROUNDS.....	26
1. INTRODUCTION.....	26
1.1 Objectives.....	51
1.2 Research Questions.....	52
1.3 Theoretical Framework.....	54
1.3.1 Phonetics and Phonology.....	59
1.3.2 Articulatory Phonology.....	70
1.3.3 Timing, Control Hypothesis, and Implementation.....	83
1.3.4 Outline of BP.....	92
2. Outline of Nasalization.....	97
2.1.1 BP: Phonetics and Phonology.....	113
2.1.2 Oral and Nasal Vowels.....	119
2.1.3 Oral and Nasal Diphthongs.....	130
2.1.4 Anatomy and Physiology.....	142
2.1.4.1 Nasal Cavity.....	145
2.1.4.2 Soft Palate.....	154
CHAPTER 2: AERODYNAMIC ANALYSIS.....	159
3. Introduction: Aerodynamic Analysis.....	159
3.1 Objectives.....	164
3.2 Research Questions.....	164
3.3 Theoretical Framework: Overview of Aerodynamics.....	167
3.3.1 Methodology.....	172
3.3.2 Material.....	173
3.3.3 Equipment.....	174
3.3.4 Corpus.....	175
3.3.5 Speakers.....	179
3.3.6 Data Processing.....	180

3.3.7	<i>Segmentation</i> .....	182
3.3.8	<i>Parameters</i> .....	191
3.4	<i>Results</i> .....	193
3.4.1	<i>Mask Resonance</i> .....	198
3.4.2	<i>The Negative Nasal Airflow</i> .....	201
3.4.3	<i>Back Diphthong</i> .....	203
3.4.4	<i>Back-Nasal Diphthong</i> .....	206
3.4.5	<i>Front Diphthong</i> .....	213
3.4.6	<i>Front Nasal Vowel</i> .....	218
3.4.7	<i>Nasal Appendix</i> .....	225
3.4.8	<i>Duration and Nasality Rate</i> .....	228
3.5	<i>Partial Conclusion</i> .....	242
CHAPTER 3: FIBERSCOPE ANALYSIS .....		250
4.	<i>Introduction to Fiberscope Analyses</i> .....	250
4.1	<i>Objectives</i> .....	260
4.3	<i>Endoscopy</i> .....	264
4.3.1	<i>PNG</i> .....	268
4.4	<i>Theoretical Framework</i> .....	271
4.4.1	<i>Velopharyngeal Sphincter</i> .....	280
4.5	<i>Methodology</i> .....	287
4.5.1	<i>Equipment</i> .....	287
4.5.2	<i>Corpus</i> .....	290
4.5.3	<i>Speaker</i> .....	292
4.5.4	<i>Data Processing</i> .....	293
4.5.5	<i>Segmentation</i> .....	294
4.5.6	<i>Parameters</i> .....	303
4.6	<i>Results</i> .....	304
4.7	<i>Partial Conclusion</i> .....	323
CHAPTER 4: ARTICULATORY ANALYSIS .....		329
5.	<i>Introduction: Articulatory Analysis</i> .....	329
5.4	<i>Methodology</i> .....	341
5.4.1	<i>Material</i> .....	341
5.4.2	<i>Equipment</i> .....	343
5.4.3	<i>Corpus</i> .....	345
5.4.4	<i>Speaker</i> .....	347
5.4.5	<i>Data Processing</i> .....	347
5.4.6	<i>Segmentation</i> .....	351
5.4.7	<i>Parameters</i> .....	358
5.5	<i>Results</i> .....	359
5.5.1	<i>Velum Movement</i> .....	364
5.5.2	<i>Soft Palate Spatial Trajectory</i> .....	367
5.5.2.1	<i>Back Segments</i> .....	368
5.5.3	<i>Tongue Movement</i> .....	380

5.5.3.1	<i>Back Diphthongs</i> .....	380
5.5.3.2	<i>Front Segments</i> .....	393
5.6	<i>Partial Conclusion</i> .....	410
CHAPTER 5: CONCLUSION.....		415
REFERENCES.....		426

## LIST OF FIGURES

---

Figure 1: Map of Brazil. The portion in grey indicates the São Paulo State and the black dot is the location of São Paulo City.....	95
Figure 2: Model of speech production of the human vocal tract with the two tubes in parallel (nasal and oral cavities).....	107
Figure 3: An example of overlapping graphics of poles and zeros from an oral vowel spectrum [a] in the blue line and the nasalized spectrum of [ã] in the red line. Adapted from Pickett (1980) and Styger & Keller (1994, p. 117). .....	110
Figure 4: Adaptation of the vocal tract from Kent (1997, p. 64). Articulatory points of nasal consonants. ....	117
Figure 5: Acoustic articulatory adaptation from BP vowels (RUSSO & BEHLAU, 1993, p. 36) and its vowel diagram (BARBOSA & ALBANO, 2004, p. 229). .....	120
Figure 6: Diphthong chart diagram a) oral falling diphthongs and (b) oral rising diphthongs. Acoustic articulatory adaptation of nasal diphthongs diagram (CAGLIARI, 1970, p. 13-14). The arrows indicate the direction of tongue motion during speech production.....	134
Figure 7: Diphthong chart diagram. Acoustic articulatory adaptation of nasal diphthongs diagram (CAGLIARI, 1970, p. 13-14). The arrows indicate the direction of tongue motion during speech production.....	138
Figure 8: Illustration of inhalation and exhalation process. Adapted from Mayo Clinic website: <a href="https://www.mayoclinic.org/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873">https://www.mayoclinic.org/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873</a> . © Mayo Foundation for Medical Education and Research. All rights reserved. ....	143
Figure 9: Illustration of Boyle's law. Time-temperature-transformation (TTT) is the relationship between pressure and volume when temperature and mass are constant. ....	143
Figure 10: Illustration of the respiratory system adaptation (STEVENS, 1998). Midsagittal section of the vocal tract, vocal folds, and main cavities in the respiratory system. ....	144
Figure 11: Illustration of open and closed vocal folds. Adapted from Mayo Clinic website: <a href="https://www.mayoclinic.org/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873">https://www.mayoclinic.org/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873</a> . © Mayo Foundation for Medical Education and Research. All rights reserved. ....	144
Figure 12: Sagittal mid-section of the oral and nasal cavities. Illustration of the places of articulation and active and passive articulators. Adaptation of the picture of the vocal tract in STEVENS (1998). .....	146
Figure 13: Lateral walls of the nasal cavity (75%). Medial view of the right side of the nasal septum (paramedian section) of the dry skull. The individual parts of bones are represented by different colors. Adapted from WOLF-HEIDDEGGER (2006, p. 70), NETTER (2000), and ARTMED (2000). .....	147
Figure 14: Lateral walls of the nasal cavity (75%). Medial view of bones of the nasal cavity, meatus, and nasal concha. Posterior view. The individual parts of bones are represented in different colors. Adaptation (WOLF-HEIDDEGGER, 2006, p. 70), NETTER (2000), and ARTMED (2000). .....	148
Figure 15: Frontal view of the nasal septum (60%). Bony components are shown in tan and cartilage in blue. Adapted from (WOLF-HEIDDEGGER, 2006, p. 68), NETTER (2000), and ARTMED (2000). .....	149
Figure 16: Anterolateral view of the nasal septum and the cartilage in the nasal septum. Bony components are shown in tan and cartilage in blue. Adapted from WOLF-HEIDDEGGER (2006, p. 28), NETTER (2000), and ARTMED (2000). .....	150
Figure 17: Inferior view. The base of the nose represents the paired alar cartilages and the tipping point of the nose. Adapted from WOLF-HEIDDEGGER (2006, p. 68), NETTER (2000) and ARTMED (2000).....	151

Figure 18: Lateral wall of the nasal cavity and the paranasal sinuses (75%). The paramedian cut the nasal septum. Paranasal sinus projection on the face. Adapted from WOLF-HEIDDEGGER (2006, p. 70). .....	152
Figure 19: Nasal cavity and paranasal sinuses. The Paranasal sinus projection is in the anterior cranial fossa. Adapted from WOLF-HEIDDEGGER (2006, p. 74). .....	153
Figure 20: Midsagittal section of the vocal tract. Vocal tract and velopharyngeal sphincter movement representation. Adapted from Kent (1997, p. 189) and Shprintzen & Bardach (1995). In (a), the velum is opened, allowing the airstream to pass through the nasal cavity. In (b), the velum is closed in the default position. ....	155
Figure 21: Adapted image of the lateral view of the muscles of the pharynx. All rights reserved to the website Teach Anatomy. Adapted from <a href="https://teachmeanatomy.info/neck/viscera/pharynx/">https://teachmeanatomy.info/neck/viscera/pharynx/</a> .....	156
Figure 22: Soft Palate muscles: (1) Tensor Veli Palatini. (2) Levator Veli Palatini. (3) Palatoglossus. (4) Musculus Uvulae. (5) Palatopharyngeus. Adapted from Kent's illustration (1997, p. 190).....	156
Figure 23: Lateral view of the palatoglossal arches, anterior and posterior pillars. Arrows indicate muscle movement. The picture was adapted from Kent's illustration (1997, p.191). .....	157
Figure 24: Sensor position for the EVA data recording. Image is taken from Ghio & Teston (2004). .....	175
Figure 25: This figure displays a spectrogram (top), oscillogram (middle), and graphs depicting oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'pau'. The speaker associated with the data is identified as BH. The number one represents the oral vowel, while number two represents the oral glide.....	183
Figure 26: This figure illustrates a spectrogram (top), oscillogram (middle), and graphs representing oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'pão'. The data corresponds to the speaker identified as BH. In the figure, the number one represents the oral vowel, and the number two represents the oral glide.....	185
Figure 27: This figure displays a spectrogram (top), oscillogram (middle), and graphs depicting oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'sei'. The data corresponds to the speaker identified as LC. In the figure, the number one represents the oral vowel, and the number two represents the oral glide. The letters represent the vowel portion (V) and the glide (G) portion. ....	187
Figure 28: This figure presents a spectrogram (top), oscillogram (middle), and graphs illustrating oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'sem'. The data corresponds to the speaker identified as LC. In the figure, the number one represents the oral vowel, and the number two represents the oral glide. The letters represent the vowel portion (V), the glide (G) portion and the nasal appendix (AP)	188
Figure 29: The figure consists of three panels. The top panel displays a spectrogram, the middle panel shows an oscillogram, and the bottom panel presents two-line graphs. The blue line represents the oral airflow, and the red line represents the nasal airflow. The keyword analyzed in this figure is 'lei'. The speaker for the data is identified as BH.....	197
Figure 30: The figure displays the overlap of FFT (black line) and LPC spectra with the oral front vowel [e] (green line) and the glide [j] (pink line) within a 25 ms window. The analyzed word in this figure is 'frei'. The speaker for the data is identified as OD. Dashed circles are used to mark the positions of F1, F2, and F3 in the spectra. Additionally, an arrow represents the mask resonance (MR). .....	199
Figure 31: The figure illustrates the overlap of FFT (black line) and LPC spectra of the oral glide [w] (red line) and the nasal glide [w̃] (green line) within a 25 ms window. The analyzed words in this figure are 'sal' and 'são'. The speaker for the data is identified as LC. Dashed circles are used to mark the positions of F1 and F2 in the spectra. Additionally, an arrow represents the mask resonance (MR). .....	200
Figure 32: This figure includes a spectrogram (top), oscillogram (middle), and graphs showing oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'tem'. The data corresponds to the speaker identified as LC. The arrow in the figure indicates the negative airflow.....	202

Figure 33: Illustration of mid-sagittal section of the oral cavity. Illustration of target positions for each phoneme [a] and [w]. Adapted from Kent (1997, p. 66). .....	203
Figure 34: This figure displays a spectrogram (top), oscillogram (middle), and graphs showing oral airflow (blue) and nasal airflow (red line). The keyword examined in this figure is 'sal'. The data pertains to the speaker identified as BH.....	204
Figure 35: This figure presents a spectrogram (top), oscillogram (middle), and graphical representations of oral airflow (blue) and nasal airflow (red line). The analyzed keyword in this figure is 'mão'. The data corresponds to the speaker identified as LC.....	209
Figure 36: The FFT (black line) and LPC spectrum overlap of the oral central vowel [a] (pink line) and the nasal vowel [ã] (blue line) in the 25 ms window. Words 'sal' a 'são'. Speaker PL. Dashed circles mark F1, F2, and FN3. The arrow represents MR.....	210
Figure 37: The FFT (black line) and LPC spectrum overlap of the oral central vowel [w] (green line) and the nasal vowel [w̃] (green line) in the 25 ms window. Words 'sal' and 'são'. Speaker LC. Dashed circles mark F1 and F2. The arrow represents MR.....	212
Figure 38: Illustration of the mid-sagittal section of the oral cavity. Illustration of target positions for each phoneme [e] and [j]. Adapted from Kent (1997, p. 66). .....	213
Figure 39: This figure presents a spectrogram (top), oscillogram (middle), and graphics depicting the oral airflow (blue) and nasal airflow (red line). The keyword used for analysis is 'dei', and the speaker for this data is LC.	216
Figure 40: The FFT (black line) and LPC spectrum overlap of the oral central vowel [e] (pink line) and the nasal vowel [j] (green line) in the 25 ms window. Words 'frei' and 'sem'. Speaker: OD. Dashed circles mark F1, F2, and F3. The arrow represents MR.....	218
Figure 41: This figure displays a spectrogram (top), oscillogram (middle), and graphics representing the oral airflow (blue) and nasal airflow (red line). The arrow indicates the direction of positive nasal airflow. The keyword analyzed in this context is 'tem', and the speaker for this data is OD. ....	219
Figure 42: This figure presents the overlap of the FFT (black line) and LPC spectrum of the oral vowel [e] (pink line) and the nasal vowel [ẽ] (blue line) in a 25 ms window. The analyzed words are 'sei' and 'sem', spoken by the speaker OD. Dashed circles indicate the positions of F1, F2, and F3. The arrow represents the mask resonance (MR).....	222
Figure 43: This figure shows the overlap of the FFT (black line) and LPC spectrum of the oral vowel [j] (blue line) and the nasal vowel [j̃] (red line) in a 25 ms window. The analyzed words are 'sei' and 'sem', spoken by the speaker OD. Dashed circles indicate the positions of F1, F2, and F3. The arrow represents the mask resonance (MR).....	224
Figure 44: This figure illustrates the negative airflow indicated by the arrow. It includes a spectrogram (top), oscillogram (middle), and graphs showing the oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'cão', spoken by the speaker OD. ....	226
Figure 45: Nasal Airflow Shape. Representation of the AFn pattern for our nasal segment data for both groups of speakers.....	245
Figure 46: <i>Der Lichtleiter</i> or 'The Light Conductor' (MORGENSTERN, 2005). .....	265
Figure 47: Antoine Jean Desormeaux's endoscope (VERGER-KUHNKE, REUTER & BECCARIA, 2007)..	266
Figure 48: Illustration of Nitze's cystoscope (GRUNERT & OERTEL, 2011). .....	267
Figure 49: Image of Rudolph Schindler's endoscope (DE GROEN, 2017). .....	267
Figure 50: Image of the modern video endoscope (GAAB, 2013). .....	268
Figure 51: Image of the flexible endoscope (DI IEVA, TAM, TSCHABITSCHER and CUSIMANO, 2014). .....	268



Figure 52: Measurements of opening and closing movements of the velum (SHPRINTZEN, 1995). Velopharyngeal patterns. (A) Coronal pattern. (B) Sagittal pattern. (C) Circular pattern. (D) Circular with Passavant's ridge pattern (SHPRINTZEN, 1995). .....	285
Figure 53: Nasofiberscope video-endoscopy position to collect data (VAISSIÈRE, HONDA, AMELOT, MAEDA & CREVIER-BUCHMAN, 2010). .....	288
Figure 54: Illustration of the light diffusion inside the cavity (VAISSIÈRE, HONDA, AMELOT, MAEDA & CREVIER-BUCHMAN, 2010). .....	288
Figure 55: Nasal cavity anterior. Images of the velum's movement during the production of the word 'cão'. The velum's opening movement to the velum's lowering movement are shown from left to right. Speaker RD.....	295
Figure 56: Nasal cavity anterior. A frame showing the reference point (RF). Adapted from Amelot's (2004) methodology. Speaker RD .....	296
Figure 57: Nasal cavity anterior. Example of velar measurement. The frames show the velum's lower and higher points. Adaptation of Amelot's (2004) methodology. Speaker RD.....	297
Figure 58: Anterior view of the nasal cavity. Image frames and the oscillogram. The alignment of the nasofiberscope frames and the acoustic output data. Target word: 'pão'. Adapted from Amelot's (2004) methodology. Speaker RD. ....	297
Figure 59: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'tão'. Speaker RD. [N] is the nasal appendix notation. The stars in blue are the fiberscope measurements for each frame. The red arrows indicate the acoustical segmentation. ..	300
Figure 60: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'tem'. Speaker RD. [N] is the nasal appendix notation. The stars in blue are the fiberscope measurements for each frame. The red arrows indicate the acoustical segmentation. ..	301
Figure 61: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'nei'. Speaker RD. The stars in blue are the fiberscope measurements for each frame. The red arrows indicate the acoustical segmentation. ....	302
Figure 63: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'pão'. Speaker: RD. The stars in blue are the fiberscope measurements for each frame. ....	306
Figure 63: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'bem'. The stars in blue are the fiberscope measurements for each frame. ....	308
Figure 64: Illustration of the velum's behavior. The red lines represent the beginning and end of the velum displacement. Adapted from Amelot's (2004) work. ....	313
Figure 66: Illustration of the temporal velum trajectory. Adapted from Lovatto (2007). ....	325
Figure 66: Speaker RD wearing the helmet used to calibrate the 2D EMMA system.....	344
Figure 67: Speaker GV wearing the helmet before the recording session. ....	344
Figure 68: Position of pellets within the oral cavity in a) RD, b) GV, and (c) illustration of the pellets in the articulators. Midsagittal cut. Speakers pictured with the sensors glued to the articulatory points before starting the collection session. ....	349
Figure 69: Pellets in rest position in the oral cavity. RD. Midsagittal view of the oral tract.....	350
Figure 70: a) Image of the electromagnetic pellets used for the collection of articulatory data, manufactured by the company Carstens. b) Illustration of acrylic plate bitten by the participant at the end of data collection. ...	351
Figure 71: Illustrative image of the phonatory apparatus and the medium sagittal cut placed on the axis of the x-y coordinates. Illustration of the positioning of the target nasal glides. ....	352

Figure 72: Raw EMMA data. Pellet displacement using the x-y coordinate plane (cm). Targets and articulatory movements [kãw̃]. RD. ....	354
Figure 73: EMMA data. Pellets displacement using the x-y coordinate plane (cm). Targets and articulatory movements [ej]. Speaker: RD. ....	356
Figure 74: EMMA data. Pellet displacement using the x-y coordinate plane (cm) Targets and articulatory movements [ẽj]. Speaker: RD. ....	356
Figure 75: Graphic of x-y coordinates in cm. Spatial displacement of the pellets. Midsagittal view of the oral tract. Stimulus [aw]. Speaker: GV. ....	362
Figure 76: Graphic of x-y coordinates in cm. Spatial displacement of the pellets. Midsagittal view of the oral tract. Stimulus [ãw̃]. Speaker: GV. ....	363
Figure 77: Graphic of x-y coordinates in cm. Spatial displacement of the pellets. Midsagittal view of the oral tract. Stimulus [ej]. Speaker: GV. ....	363
Figure 78: Graphic of x-y coordinates in cm. Spatial displacement of the pellets. Midsagittal view of the oral tract. Stimulus [ẽj]. Speaker: GV. ....	364
Figure 79: Articulatory trajectory of the velum and the tongue back on the x-axis in cm. Speaker: RD. ....	365
Figure 80: Articulatory trajectory of the velum and the tongue back on the x-y-axis in cm, and acoustic signal. Speaker: RD. ....	366
Figure 81: EMMA data. Pellets displacement on the x-y coordinate plane (cm). Targets and articulatory movements [aw]. Speaker: RD. ....	381
Figure 82: EMMA data. Pellets displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [maw]. Midsagittal view of oral cavity. Speaker: RD. The explanation for the numbers 1-4 will be given in the text. ....	383
Figure 83: EMMA data. Pellet displacement using the x-y coordinate plane (cm). Target segment [ãw̃]. Speaker: RD. ....	384
Figure 84: EMMA data. Pellet displacement on the the x-y coordinate plane (cm). Tongue targets and articulatory movements of [pãw̃]. Midsagittal view of oral cavity. Speaker: RD. The explanation for the numbers 1-4 will be given in the text. ....	385
Figure 85: EMMA data. Pellet displacement on the the x-y coordinate plane (cm). Tongue targets and articulatory movements of [paw]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-4 will be given in the text. ....	386
Figure 86: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [pãw̃]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-4 will be given in the text. ....	388
Figure 87: EMMA data. Pellets displacement using the x-y coordinate plane (cm). Target segment [ej]. Speaker: RD. ....	394
Figure 88: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [sej]. Midsagittal view of oral cavity. Speaker: RD. The explanation for the numbers 1-2 will be given in the text. ....	395
Figure 89: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [sẽj]. Midsagittal view of oral cavity. Speaker: RD. The explanation for the numbers 1-4 will be given in the text. ....	396
Figure 90: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Target segment [sẽj]. Speaker: RD. ....	397

Figure 91: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [sej]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-2 will be given in the text.....	400
Figure 92: Emma data. Pellets displacement in x-y coordinate plane (cm). Tongue targets and articulatory movements of [sěj]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-2 will be given in the text.....	401
Figure 93: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: “pão”. Speaker: RD.....	407
Figure 94: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: “pão”. Speaker: GV.....	407
Figure 95: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: “bem”. Speaker: RD. ....	408
Figure 96: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: “bem”. Speaker: GV.....	409
Figure 97: Illustration of the tongue-displacement model. ....	411
Figure 98: Illustration of the tongue-displacement model for RD and GV. ....	412
Figure 99: Illustration of velum-lowering model for RD and GV.....	413
Figure 100: Gesture guideline comparing the phases of gestures in the production of oral diphthongs and nasal diphthongs. The solid lines show the motions that are phase-coupled, and the dashed lines represent the gestures’ phase.....	422
Figure 101: Dynamic oscillatory modeling (Tilsen, 2014) illustrates the coordination planning from the activation of tongue height and soft palate trajectories. The tongue gestures are co-selected and coordinated, while the nasality gesture is competitively selected in relation to the oral gesture. ....	423

## LIST OF TABLES

---

Table 1: IPA Chart of BP Consonant Phonemes. ....	114
Table 2: Set of words with front oral diphthongs and front nasal vowels. ....	178
Table 3: Set of words with front oral diphthongs and front nasal vowels. ....	178
Table 4: Speakers identified via their initials. The table is divided according to gender. ....	180
Table 5: Total values of the average (A) and the standard deviation (SD) in the resonance generated by the vibration of the oral mask during the production of the oral and nasal segments. N = 36. ....	200
Table 6: Absolute acoustic duration values for each segment. All gender groups. ....	229
Table 7: Segment acoustic duration values for the nasal segment: AdVn, AdGn and AdApn. All gender groups. ....	233
Table 8: Duration of the nasal airflow shape. Values for the nasal segment: Total duration of the velic trajectory (Tdv <sub>t</sub> ) (ms). Velic opening duration (Vod) (ms). Duration of glide constriction (Dgc) (ms). Velic rising duration (Vrd) (ms) in the male and female groups. ....	237
Table 9: Volume variation in nasal airflow shape. Values for the nasal segment in high NP (dm <sup>3</sup> /s). ....	239
Table 10: Set of Words with Back Oral and Nasal Diphthongs. ....	292
Table 11: Set of Words with Front Oral Diphthongs and Front Nasal Vowels. ....	292
Table 12: Soft palate target values extracted from the fiberscope data. Lateral pharyngeal wall values extracted from the PNG data. Back and front nasal segments. Average (Standard Deviation) (px) (N =10). ....	309
Table 13: Trajectory of the soft palate and the VELOPHARYNGEAL PORT cycle. Fiberscope data and VELOPHARYNGEAL PORT data for the front and back nasal segments. Average (SD) ms. (N =8). ....	314
Table 14: Delta of the temporal coordination. Opening: nasal vowel onset and velum lowering (ms). Average (SD). Speaker RD (n=20), ....	316
Table 15: Delta of the temporal coordination (ms). Carry over: Appendix duration and the velum offset. Average (SD). Speaker RD (n =20). ....	319
Table 16: Acoustic duration (ms) of front and back oral diphthongs. Average (SD). Speaker: RD. (N =10). ..	321
Table 17: Acoustic duration (ms) of front and back oral and nasal diphthongs: Average (SD). Speaker: RD. (N =10). ....	322
Table 18: Set of words for back oral and nasal diphthongs. ....	345
Table 19: Set of words for front oral diphthong and front nasal vowels. ....	346
Table 20: Soft palate spatial displacement. Average and standard deviation. Production [ãw̃]. Coordinates x-y (cm). Speakers: RD and GV (N =25). ....	369
Table 21: Soft palate delta's spatial variation. Averages and standard deviations. Production [ãw̃]. Coordinates x-y (cm). Speakers: RD and GV (N =25). ....	371
Table 22: Soft palate spatial displacement. Absolute duration (opening and closing movements [ms]). Production [ãw̃]. Speakers: RD and GV. (N =25). ....	373
Table 23: Soft palate spatial displacement. Averages and standard deviations. Production [ẽj̃]. Coordinates x-y (cm). Speakers: RD (N =25) and GV (N =15). ....	375
Table 24: Averages and standard deviations of the spatial delta of soft palate displacement. Front nasal vowel [ẽj̃]. RD (N =25) and GV (N =15) on the x-y axis (cm). ....	377
Table 25: Soft palate spatial displacement. Absolute duration (opening and closing movements [ms]). Production [ẽj̃]. Speakers: RD (N =25) and GV (N =15). ....	379
Table 26: Averages and standard deviations of tongue-displacement trajectory: post-dorsal. Oral and back nasal diphthong: [aw] and [ãw̃]. RD (N =25) (cm). ....	386

Table 27: Averages and standard deviations of tongue-displacement trajectory: post-dorsal. Oral and back nasal diphthong: [aw] and [ãw̃]. Speaker: GV (N =5) (cm).....	388
Table 28: Averages and standard deviations. Delta from tongue-displacement trajectory and velum trajectory. Back nasal diphthong [ãw̃]. Speaker: RD. (N =25) (cm).....	390
Table 29: Averages and standard deviations. Delta from tongue-displacement trajectory and velum trajectory. Back nasal diphthong [ãw̃]. Speaker: GV. (N =20) (cm),.....	391
Table 30: Averages and standard deviations of tongue-displacement trajectory: tongue body. Oral and front nasal segments: [ej] and [ẽj̃]. Speaker: RD. (N =25) (cm).....	398
Table 31: Average and Standard Deviation of Tongue Displacement Trajectory: Tongue Body. Oral and Nasal Front Segments: [ej] and [ẽj̃]. GV (N =25) (cm). ....	401
Table 32: Averages and standard deviations of temporal acoustic duration of velum trajectory. Delta variation between nasal-vowel onset and velum lowering. Delta variation between velum closure and stop VOT. Speakers: RD (N =25) and GV. (N =15) (ms).....	403

## LIST OF CHARTS

---

Chart 1: Absolute acoustic duration (ms). Comparison between the target segment's production and the male group (N=36).....	230
Chart 2: Absolute acoustic duration (ms). Comparison between the target segment's production and the female group (N=36).....	230
Chart 3: Absolute acoustic duration (ms). Comparison between the target segment's production and both target's productions (N=36). ....	231
Chart 4: Average of the segment acoustic duration (ms). Comparison between the target segment's production in the male and female groups (N=36). Values for the nasal segment: AdVn, AdGn, and AdApn.....	234
Chart 5: Average of segment acoustic duration (ms). Comparison between the target segment's production in the male and female groups (N=36). Values for the nasal segment: AdVn, AdGn, and AdApn.....	236
Chart 6: Average duration of the nasal airflow shape. Values for the nasal segment: Velic opening duration (Vod) (ms). Duration of glide constriction (Dgc) (ms). Velic rising duration (Vrd) (ms) for the male and female groups. N= 36.....	238
Chart 7: Average duration of the nasal airflow shape. Values for nasal segment NP (dm <sup>3</sup> /s) for the male and female groups.....	240
Chart 8: Occurrences of NP. The 3-phase standard for AFn (N=120). ....	241
Chart 9: Spatial velum movement in the fiberscope data and the PNG data for soft palate displacement. Front and nasal diphthongs: Average (SD) px (N=10).....	309
Chart 10: Spatial velum movement in the fiberscope data and the PNG data for soft palate displacement. Front and nasal diphthongs: Average (SD) px (N=10).....	310
Chart 11: Maximum velum opening. Fiberscope data for the soft palate displacement. Front and nasal diphthongs: Average (SD) px (N=10).....	311
Chart 12: Spatial velum movement in the fiberscope data and PNG data for soft palate displacement. Front and nasal diphthongs: Average (SD) px (N=20). ....	313
Chart 13: Trajectory of the soft palate and the velopharyngeal port cycle. Fiberscope data and PNG data for the front and back nasal segments. Average (SD). Values in pixels. (N=8).....	315
Chart 14: Soft palate and the velopharyngeal port cycle. Fiberscope data and PNG data for the front and back nasal segments. Average (SD) ms. (N=20).....	319
Chart 15: Average of the Duration of Acoustic Nasal Segments: Average (SD) px (N=10). ....	322
Chart 16: Position of soft palate spatial displacement. Production [ãw̃]. Coordinates x-y (cm). Speaker: RD. (N=25). ....	369
Chart 17: Position of soft palate spatial displacement. Production [ãw̃]. Coordinates x-y (cm). Speaker: GV. (N=25). ....	370
Chart 18: Soft palate spatial displacement. Delta between onset and velum target. Production [ãw̃]. Coordinates x-y (cm). Speaker: RD and GV. (N=25).....	371
Chart 19: Soft palate trajectory. Average duration to opening trajectory and the closure of the soft palate. Total duration (ms). Production [ãw̃]. RD and GV (N=25). ....	374
Chart 20: Position of soft palate spatial displacement. Production [[ễ]]. Coordinates x-y (cm). Speaker: RD. (N=25). ....	376
Chart 21: Position of soft palate spatial displacement. Production [[ễ]]. Coordinates x-y (cm). Speaker: GV. (N=15).....	376

Chart 22: Soft palate spatial displacement. Lowering trajectory. Production [ẽ̃]. Coordinates x-y (cm). Speaker: GV. (N =15).....	377
Chart 23: Soft palate trajectory. Average duration of opening trajectory and the closing of the soft palate. Total duration (ms). Production [ẽ̃]. Speakers: RD (N =25) and GV (N =15).....	379
Chart 24: Tongue position: post-dorsal. Back oral diphthong [aw] and nasal [ãw̃]. RD (N =25) (cm). ....	387
Chart 25: Tongue Position: post-dorsal. Oral and nasal back diphthong: [aw] and [ãw̃]. Speaker: GV. (N =25) (cm). ....	389
Chart 26: Tongue Position trajectory: post-dorsal. Oral and back nasal diphthong: [ãw̃]. Static value of the pellet. RD. (N =25) (cm).....	391
Chart 27: Tongue-displacement trajectory: post-dorsal. Oral and back nasal diphthong: [ãw̃]. Static value of the pellet. Speaker: GV. (N =25) (cm).....	392
Chart 28: Tongue-displacement position: tongue body. Oral and front nasal diphthong: [ej] and [ẽ̃]. Static value of the pellet. RD (N =25) (cm).....	399
Chart 29: Tongue-displacement position: tongue body. Front oral diphthong [ej] and nasal [ẽ̃]. Static value of the pellet. Speaker: GV. (N =25) (cm).....	402

## LIST OF ABBREVIATIONS AND SYMBOLS

---

A	Average
AdAPn	Acoustic Temporal Duration of The Nasal Appendix
AdGn	Acoustic Temporal Duration of The Nasal Glide
AdVn	Acoustic Temporal Duration of The Nasalized Vowel
AdVs	Acoustic Temporal Duration of The Oral Segment
AFn	Nasal Airflow
AFo	Oral Airflow
BC	Before Christ
BCE	Before Current Era
BH	Male Speaker
BP	Brazilian Portuguese
CN	Nasal Consonant
CVG	Consonant Vowel Glide
CṼG	Consonant Nasal Vowel Nasal Glide
dB	Decibel
dm <sup>3</sup> /s	Cubic decimeter per second
Dgc	Duration of Glide Constriction
EMMA	Electromagnetic Midsagittal Articulography
EP	European Portuguese
EVA	Système d'Évaluation Vocale Assistée
F0	Fundamental Frequency
F1	First Formant
F2	Second Formant
F3	Third Formant
FFT	Fast Fourier Transform
FN	Nasal Formant
FN1	First Nasal Formant
FN2	Second Nasal Formant
FN3	Third Nasal Formant
G	Glide
GV	Glide Vowel
Hz	Hertz
int	Intensity
IPA	International Phonetic Alphabet
JAW	Lower Incisors
k	Constant
kHz	Kilohertz
LC	Female Speaker
LpITt	Nasal inlet pressure microphone
LPC	Linear Predictive Coding
LZ	Male Speaker
MAt	Oral airflow microphone
MATLAB	Matrix Laboratory



MR	Mask Frequency Resonance
ms	<i>Millisecond</i>
N	Number of tokens
<i>naf</i>	Nasal Airflow
NegAFn	Nasal Airflow Negative Duration
NP	Nasalization Peak
NS	Nose
∅	Empty Element
oaf	Oral Airflow
OD	Male Speaker
p	Pressure
PL	Female Speaker
PNC	Primary Nasal Consonant
PNG	Photo-nasography
px	Pixels
r	Pearson correlation coefficient
RD	Female Speaker
SFs	Surface Forms
SD	Standard Deviation
SG	Singular
SV	Semi-voyelle
TBCK	Tongue Back
TBL	Tongue Blade
Tdvt	Velum Total Duration - Trajectory
TPD	Tongue post-dorsum
TRAJ TNG-BACK	Tongue Back Trajectory
TRAJ TNG-DOS	Tongue Dorsum Trajectory
TRAJ VEL F	Velum Closing Position
TRAJ VEL I	Velum Initial Trajectory
TRAJ VEL M	Velum Maximum Opening Position
TRAJ VELUM	Velum Trajectory
TRAJ_TNG_DOS	Tongue Dorsum/Body Trajectory
TRAJ_TNG_PDOS)	Tongue Post-dorsal/Back Trajectory
TRAP	Traitement des Signaux de Parole
TT	Tongue Tip
TTT	Time-temperature-transformation
UI	Upper Incisors
URs	Underlying Forms
V	Vowel
v	Volume
VEL	Velum
VEL OP_D	Soft Palate_Opening Movement Duration
VEL_CL_D	Soft Palate_Closing Movement Duration
VEL_TD	Soft Palate Total Duration
VGN	Nasal Vowel Glide
VIT VELUM	Velum Speed
VN	Nasal Vowel
Vod	Velic Opening Duration

VOT	Voice Onset Time
VPhS	Velopharyngeal Sphincter
Vrd	Velic Rising Duration
VSP_TRAJ	Velum Motion
XT/3C	Phonology-Extrinsic-Timing-Based Three-Component
wa1	Wave 1

# CHAPTER 1: BACKGROUNDS

---

## Résumé (Summation)

In this chapter, we bring a theoretical overview that will serve as a foundation for the development of this thesis. We will address the general theoretical framework of phonetics and phonology, articulatory phonology, and the articulatory control theory to understand nasalization and nasal diphthongization through a spatiotemporal approach. We will briefly revisit the phonetics and phonology of Brazilian Portuguese to detail the current analyses and characterizations of the phonetic and phonological behavior of nasal diphthongization. To gain a better understanding of how the articulation and emergence of nasal diphthongs occur, we will also study the anatomy and physiology of the nasal cavity and soft palate.

## 1. Introduction

The purpose of this thesis is a descriptive and analytic study of nasal diphthongization<sup>1</sup> in the Paulistano dialect of Brazilian Portuguese (BP) to determine the precise coordination of velum and tongue activity, using different experimental methods from a phonetic perspective.

Nasal diphthongization is a typical phenomenon in the *Paulistano* dialect. We chose to analyze the nasal diphthongs produced in the city of São Paulo because the target phenomenon is a characteristic of the higher social classes, while nasalization is even reduced in other syllable positions or other Brazilian dialects.

We will describe the global articulation of nasalization in the realization of [ãũ] and [ẽĩ], demonstrating through articulatory, aerodynamic, acoustic, and image data, the correlation between the velum and the tongue movement during its production.

---

<sup>1</sup> The concept of diphthongization is a well-established linguistic phenomenon that has been extensively examined and documented in the literature of linguistics (GOLDSMITH, 1990; LADEFOGED & JOHNSON, 2014). Diphthongization is a linguistic process in which two vowel sounds merge into a single, more complex sound, referred to as a diphthong. This process occurs when one vowel sound transitions into another vowel sound through a sliding or gliding movement within the span of a single syllable.

Our goal is to show that there are spatial and temporal patterns, at the phonetic level, that do not change the phonological symbolic meaning of the language. We will delve into the phonetic process that involves the interaction of articulators during speech production. This process gives rise to phenomena such as nasal diphthongization, where nasalization affects the vowel quality of diphthongs.

Nasal diphthongization refers to the process by which a vowel or diphthong undergoes a change in its articulatory configuration, resulting in a glide-like movement or transition between two vowel-like elements within the same segment. This hypothesis suggests that nasal vowels and nasal diphthongs suffer a process of diphthongization. This phenomenon changes the characteristics of the segments that have both vowel and glide-like elements.

Additionally, we will investigate the mechanisms and gestures individually from tongue and velum employed in the production of nasalized sounds. We will examine how the velum and tongue play crucial roles in modulating the airflow through both the oral and nasal cavities during speech. Furthermore, we will explore the interplay between results from nasalization, such as assimilation and coarticulation, shedding light on the relationships between nasal<sup>2</sup> sounds and their surrounding context.

The central purpose of this research is to understand the articulatory mechanism of nasality<sup>3</sup> production and its properties, particularly focusing on the role of tongue movement and the function of the velopharyngeal during speech production.

We expect that the complexity<sup>4</sup> of nasal diphthongization production will provide insights into the importance of tongue movement and the function of the velum

---

<sup>2</sup> The term *nasal* refers to a fully nasal phoneme, and the term *nasalized* refers to oral sound partially nasal phonemes as allophones.

<sup>3</sup> Nasalization and nasality are terms related to sound production through nasal cavities, but they are different concepts in phonetics. Nasalization refers to the process by which an oral sound is produced with the addition of nasal resonance. This occurs when air is directed into the nasal cavity at the same time as it passes through the mouth. On the other hand, nasality refers to the property of a sound being nasal or not. Therefore, nasalization is a process that adds nasal resonance to an oral sound, while nasality is a property of a sound being nasal or not.

<sup>4</sup> Complex phonetic events can be understood and analyzed by breaking them down into sequences, which are combinations of more elementary units.

synchronicity during nasalization speech production. Consequently, this thesis is an original study on nasal diphthongization in the Paulistano dialect from BP and its realization was only possible due to the increased development of and access to new technologies.

The investigation of how nasality is produced, and its various properties helps to uncover the intricate articulatory patterns involved in the movement of the soft palate and the coordination between the soft palate and the tongue.

In our study, our objective is to demonstrate anticipatory and postpone nasalization, showcasing how it extends and affects the neighboring sounds. Furthermore, we seek to highlight the influence of nasalization on the duration and spectral characteristics of sounds, specifically observing longer durations in nasalized sounds and distinct acoustic features such as lowered formant frequencies and heightened energy in lower frequency regions.

Through detailed analysis of articulatory data, we aim to emphasize the role played by the velum and tongue in nasalization. These articulatory mechanisms regulate the airflow, facilitating its passage into the nasal cavity while simultaneously creating a barrier in the oral cavity.

In our investigation, we draw upon the relevant thesis of Cagliari (1977)<sup>5</sup>, his research has made significant contributions to the understanding of nasality in speech production, providing valuable insights that base our study. From an experimental point of view, this thesis is a continuation of the experimental steps of Cagliari (1977).

According to Cagliari (1997, p.8), “(...) *the phonological treatment of phonetic diphthongs is one of the most controversial areas of the language (...)*”. He analyzed that some monophthongs may be realized as phonetic diphthongs in the following context /i, e, o, u/, when followed by /N/. In his thesis, Cagliari (1997, p.34) shows that:

---

<sup>5</sup> Cagliari thesis includes various experiments exploring the acoustics, aerodynamics, perception, and linguistic significance of nasality, utilizing techniques such as fluorography, pneumotachography, electrokymography, laryngography, spectrography, and palatography using phonetic data from Paulistano dialect of Brazilian Portuguese.

“The phonetic nature of the palatal and velar nasals as exponents of /n/ in wordfinal position before pause, needs to be considered. They are unreleased nasals in the sense that towards the end of their articulation, when there is still an articulatory closure in the mouth, the vocal cords stop vibrating and airflow pressure decreases rapidly behind the oral blockage so that when the closure is removed, no sound is produced. At that point, there is usually a well-increased level of nasal airflow. When these nasals occur as unreleased nasals at the end of words, principally before a pause, they have an auditory effect similar to hum without noticeable differentiation between them in terms of place of articulation (...).”

Cagliari (1977) showed the phonetic emergence and the nature of unreleased velar and palatal nasals in the word-final position before a pause. These findings explained the specific characteristics of nasal diphthongization, which is the focal point of our study.

By understanding how nasal palatalization and nasal velarization occur through experimental evidence, we can establish a solid empirical basis for our research and contribute to the understanding of these phonetic processes.

As we said, Cagliari's work helped support the framework and techniques of our study as a reference work of nasality, reinforcing the importance of investigating nasal diphthongization systematically and empirically.

Cagliari's (1977) study examines nasalization in Brazilian Portuguese, focusing on acoustic output, velopharyngeal movements, and aerodynamics associated with nasal sounds. The research analyzes nasal and oral airflow levels and their distribution, providing insights into nasalization patterns in connected speech. The study reveals that vowels before nasal consonants can be fully nasalized, partially nasalized, or not nasalized at all.

Nasals generally exhibit higher peaks on the nasal inlet pressure microphone (LpITt) compared to nasalized vowels. Additionally, voiced stops and nasal consonants show weak vibrations on the oral airflow microphone (MA<sub>t</sub>), indicating residual oral resonances despite complete airflow blockage. Rarely does a nasal consonant follow a nasalized vowel and precede an oral consonant, especially when

they share the same place of articulation. Before a pause, the nasal archiphoneme<sup>6</sup> is typically realized as an unreleased nasal sound.

Following a nasal consonant, the initial half of a vowel's duration undergoes partial nasalization. Nasal consonants in liaison are infrequent, and in cases of hiatus, the first vowel tends to be predominantly oral while the second vowel is primarily nasalized. The onset of nasal airflow coincides with vocal cord vibrations observed on the laryngogram. Cagliari's study enhances our understanding of nasalization in speech production.

One of the main purposes of our study is to show how tongue adjustment can contribute to the production of nasality. The production of nasality is a complex process involving various articulators, including the tongue, velum, and lips. The movement of these articulators plays a crucial role in determining the degree of nasality in the sound. In particular, the position and shape of the tongue are crucial in adjusting the nasal resonance of the vocal tract.

Previous studies have highlighted the importance of tongue adjustment in the production of nasality. For example, Cagliari (1977) found that the position of the tongue is a major factor in determining the degree of nasality in Brazilian Portuguese. Similarly, Shosted (2006) noted that the adjustment of the tongue is essential in producing nasal resonance in English.

In this study, we aim to further investigate the contribution of tongue and velum adjustment to the production of nasality in the context of the Paulistano dialect of Brazilian Portuguese. Specifically, we will compare the physical movements of the tongue and velum during the production of the front nasal vowel and back-nasal diphthong in this dialect.

We will investigate the differences in the gestures involved in the oral and nasal

---

<sup>6</sup> Archiphonemes are abstract representations that encompass multiple phonemes that are not distinguished phonemically in a particular context. The variability refers to the different ways in which these archiphonemes are realized or manifested in speech.

segments<sup>7</sup> of the pronunciation of the diphthong 'ão' [ãũ] and its oral counterpart 'au' [aw] and the production of the diphthongized nasal vowel /ẽ/, which is produced as [ẽ̃], and the production of the oral diphthong [ej]. We will compare the nature of these gestures, including their timing, degree of movement, and coordination.

By focusing on the contribution of the tongue to the production of nasality, we hope to shed further light on the mechanisms underlying the production of nasal vowels and diphthongs in Brazilian Portuguese. This will not only provide a better understanding of the phonetics of this dialect but also contribute to our broader understanding of the role of articulators in speech production.

This study will provide crucial descriptive data about how these sounds are articulated, to increase the understanding of the complexity of the synchronization mechanism between the tongue and the movement of the velum during the production of nasal diphthongs.

We hope that new insights into the aerodynamic and articulatory functions could enrich the field, providing new insights into the aerodynamic and articulatory functions involved in nasality and contributing to the broader understanding and refinement of phonetic and phonological theories related to nasality in this language variety. Lastly, we hope that the different approaches in this thesis contribute to improving the discussion of phonetics and phonology about nasality in BP.

Several phonetic and phonological studies of nasality have classified sound production (phones, phonemes, and allophones) and perception in different languages; however, most of these studies have an element in common, namely the difficulty of describing the characteristics of nasalization during the production of vocalic sounds.

The origin of nasal vowels in Brazilian Portuguese dates to the colonial period when European Portuguese (EP) was brought by colonizers to the territory that is now

---

<sup>7</sup> The term "segment" refers to a discrete unit of speech sound, which can be an individual sound or a combination of sounds. In our research, we analyze the properties of nasal/oral vowels and oral/nasal glides in order to understand the phonetic structure underlying their coarticulatory patterns and how these patterns contribute to the production of nasal diphthongization.



Brazil. The presence of nasal vowels in Brazilian Portuguese is the result of both remnants from Old Portuguese and influences from indigenous and African languages.

In Old Portuguese, nasal vowels emerged through the nasalization of oral vowels before nasal consonants. For example, the vowel /a/ followed by the nasal consonant /n/ became nasalized, giving rise to the nasal vowel /ã/. This process occurred similarly with other oral vowels, resulting in the nasal vowels /ã/, /õ/, and /ĩ/.

Additionally, the indigenous and African languages spoken by native people and enslaved individuals also influenced the formation of Brazilian Portuguese. These languages had phonetic characteristics, including the use of nasal vowels, which were incorporated into the Portuguese language spoken in Brazil.

Thus, the nasal vowel in Brazilian Portuguese is multifaceted, resulting from both internal phonetic processes within the language and the influences of other languages present in the country's historical context.

Most studies focus on the diachronic change from VN to  $\tilde{V}$ , intending to approach how the source segment (VN) starts to be interpreted as vowel nasalization (BEDDOR, 2009).

Lacerda & Rossi's (1958) investigation compared nasalization in Lisbon and Carioca dialect, they concluded that BP exhibits stronger nasalization compared to European Portuguese. Additionally, Mioni (1973) conducted a comparative study of EP and BP, asserting that the nasalization of a vowel followed by a nasal at the beginning of the next syllable in words is more pronounced in BP than in EP.

Head (1964) employed an acoustic criterion to evaluate the varying degrees of nasalization in EP and BP, identifying the relative darkness of nasal formants in Carioca nasal vowels and diphthongs as evidence for stronger nasalization in BP. However, it is worth noting that the perception of nasality does not consistently correspond to the darkness of formants in spectrograms, as nasality affects the acoustic structure of vowels and consonants differently and influences different vowel types in distinct ways.

Viana (1892), Sweet (1913), Heffner (1950), and Delattre (1969) have examined the differences between Portuguese and French nasality, with all of them contending that Portuguese nasality is comparatively less intense than French nasality.

Viana specifically highlights the presence of two types of nasality in Portuguese, namely the "nasal vowels of first degree" found in the South, and the "nasal vowels of second degree" heard in the North, which can be extended over the vowels through gutturalization.

In contrast, French nasal vowels, particularly those from the North, exhibit similar characteristics. These comparative studies contribute valuable insights into the specific phonetic specifications of nasality in BP.

Stevens (1954), in his study of EP, identified two degrees of nasalization. He referred to the stronger degree as "the second degree," which represents phonologically significant vocalic nasalization, while the weaker degree occurs when nasalization is not phonologically significant.

Stevens (1954) explains that the second degree of nasal resonance requires maximum nasality production for a given loudness, indicating a lowered soft palate to its fullest extent. Further investigations conducted by Stevens & Lacerda (1956) reaffirm the existence of these two distinct degrees of nasality in EP, where the stronger degree corresponds to nasal vowels and the weaker degree to nasalized vowels.

However, few studies have examined the co-articulatory characteristics of nasal diphthongization, especially in nasal diphthongs. Few studies have focused on the allophonic realization of nasal diphthongs or their nature. So, our study is important because we are interested in how the "*articulatory covariation of duration and temporal effects on N (source)*" (BEDDOR, 2009) and affects the vowel and the diphthong.

Moreover, our study offers the opportunity to investigate the intricate relationship between the nasal source and the vowel in nasal diphthongs. The articulatory covariation of duration and temporal effects on the nasal source can reveal valuable

information about the co-articulatory patterns and coordination between the nasal cavity and the vocal tract.

By examining these aspects, we can gain a deeper understanding of the dynamic interaction between nasalization and diphthongization, providing insights into the articulatory and acoustic characteristics of nasal diphthongs. However, our approach combines elements from laboratory phonology, articulatory phonology, and control theory to understand the production and diversity of nasal vowels and nasal diphthongs, all these different approaches can help us understand the mechanisms that influence changes in production and symbolic phonological representations (TURK & SHATTUCK-HUFNAGEL, 2020).

A contemporary theoretical proposal by Turk and Shattuck-Hufnagel (2020) suggests the production of speech based on Articulatory Phonology, Speech Control, spatiotemporal representations, and symbolic phonology. They discuss the key features of existing articulatory models of speech production concerning timing, considering their advantages and disadvantages. Also, they aim to evaluate these features by examining evidence from both speech and nonspeech motor control literature.

The XT/3C (Phonology-Extrinsic-Timing-Based Three-Component) approach proposed by Turk & Shattuck-Hufnagel offers a framework for understanding speech production and timing. It consists of three components: Phonological Planning, Phonetic Planning, and Motor-Sensory Implementation.

In this approach, the Phonological Planning Component sets the goals for the utterance based on abstract symbolic representations, including lexical and prosodic structures, as well as other speech requirements like rate and style. It selects qualitative acoustic cues that correspond to distinctive features of the words in the planned utterance. These cues are specific to each context within the utterance and are not idealized spatial locations.

The Phonetic Planning Component translates the goals set by the Phonological Planning Component into quantitative acoustic targets and timing characteristics. It

specifies how these goals will be achieved in the planned utterance and plans coordinated articulatory movements that generate the desired acoustic cues with appropriate values.

The Motor-Sensory Implementation Component tracks and adjusts the planned movements to ensure that the goals defined in the Phonological Planning Component and specified in the Phonetic Planning Component are met at the appropriate times. It ensures that the desired acoustic and articulatory goals are achieved during speech production.

The XT/3C framework emphasizes the use of general-purpose timing mechanisms that are extrinsic to phonology. It also highlights the coordination between the endpoint of movements (e.g., articulatory targets) and surface timing, giving priority to these endpoints over other movement parts. This approach provides a plausible architecture that can account for observed timing behaviors and systematic timing variability in speech.

The three components of the XT/3C approach align with similar stages in previous production models. However, the XT/3C framework offers a more comprehensive treatment of surface timing and provides a promising alternative to the traditional Articulatory Phonology/Task Dynamics model.

Overall, the XT/3C approach offers a detailed framework that integrates phonological, phonetic, and motor-sensory aspects of speech production and timing. It provides a way to link abstract symbolic representations with specific acoustic and articulatory targets, considering the contextual requirements of the utterance.

Their book provides an enriching overview of the architecture of speech production planning and discusses the cognitive representations of symbolic elements and their relationships, focusing on motor control processes and spatiotemporal relationships of speech.

One important aspect of the evidence discussed is the greater timing precision observed at movement endpoints compared to other parts of movements. This observation suggests separate control of endpoint timing, challenging models that

assume all parts of a movement trajectory are controlled by the same equation of motion.

Instead, the evidence supports models in which abstract, symbolic phonological representations map onto spatial and temporal characteristics of the movement, particularly those related to the goal of producing specific acoustic cues for phonological contrasts. The book presents three types of evidence for phonology-extrinsic timing, indicating that surface duration requirements are represented during speech production.

This finding is reinforced by greater timing variability observed for longer intervals, attributed to noise in a general-purpose timekeeping process that is independent of phonology. These pieces of evidence contradict models that have a unified Phonology/Phonetics Component, do not represent surface timing of phonetic events, and lack general-purpose timekeeping mechanisms to represent and track timing.

Turk & Shattuck-Hufnagel (2020) take different approaches to the same spatiotemporal aspects of speech production, as we do. Both of our works have convergent views on the subject. However, Turk & Shattuck-Hufnagel make a more generalized review of existing data from another research.

While Turk & Shattuck-Hufnagel (2020) offer a comprehensive exploration of the architecture of speech production planning and the cognitive representations of symbolic elements, our thesis takes a more specific approach by analyzing data on nasalization and nasal diphthongization. Turk & Shattuck-Hufnagel's (2020) work provides an alternative approach to previous proposals made in the field.

Whereas in this thesis there is an analysis of unpublished data on nasal diphthongization, from four different experimental approaches. Our study delves into the spatiotemporal aspects of speech production, examining the co-articulatory characteristics and the impact of articulatory covariation on the vowel and diphthong in the context of nasalization.

By narrowing our focus to nasalization and nasal diphthongization, we can provide a detailed analysis of the spatiotemporal effects on the vowel and diphthong during the nasalization processes. Our approach allows for a deeper understanding of the co-articulatory patterns, durations, and temporal relationships associated with nasal diphthongization, which may not have been extensively explored in the broader context of speech production planning.

This theory is important to help us to understand nasal diphthongization because it provides insights into the underlying mechanisms and processes involved in speech production, particularly about timing. Nasal diphthongization is a complex phenomenon that involves the coordination of various articulatory movements and the precise timing of these movements.

By discussing the key features of existing articulatory models of speech production and evaluating their approaches to timing, we gain a better understanding of the theoretical frameworks used to explain speech production processes. This understanding helps us assess the advantages and disadvantages of different models and identify their relevance to nasal diphthongization.

The evidence presented, such as the greater timing precision at movement endpoints and the presence of phonology-extrinsic timing, provides empirical support for specific aspects of speech production related to nasal diphthongization. May the nasal diphthongization be not solely governed by phonological representations or spatiotemporal constraints but involves the coordination of motor control processes and the integration of timing cues.

The discussion of phonology-extrinsic timing mechanisms and the representation of surface timing requirements in understanding nasal diphthongization is significant. It emphasizes the need to consider both phonological and phonetic factors when investigating this phenomenon.

By examining how specific parts of the movement trajectory are accurately implemented to the phonological goals of producing nasal diphthongs, we gain insights into the prioritization of certain aspects of speech production.

In this thesis, our focus is to parameterize the systematic temporal variations observed in the production of nasal diphthongization. We aim to establish a link between these temporal variations and symbolic phonology, while also connecting them to the spatiotemporal properties of speech production and symbolic phonology.

Our analysis will specifically focus on vowel rising resulting from nasal diphthongization in the Paulistano dialect of Brazilian Portuguese. By examining the articulatory output and its relation to spatiotemporal properties. Diphthongization is a phonetic change that occurs in the internal structure of the syllable, in which a vowel splits into two elements: a vowel and a semivowel/semi-consonant. This sound change typically occurs because of diachronic evolution.

There are two possible outcomes: either the nasal vowel disappears, or the nasal consonantal element moves to the preceding vowel. This observation suggests the presence of a nasal archiphoneme (/N/) that functions as a nasal glide, like a nasal consonant.

This idea supports the hypothesis that nasalized vowels in Brazilian Portuguese are derived from oral vowels. The re-syllabication of nasal vowels, where the nasal element moves the glide. In this view, the nasal archiphoneme /N/ is responsible for the nasalization of the preceding vowel that changes vowel position and functions as a distinct phonological unit.

In our data, we observe two segments that undergo this phenomenon of nasal diphthongization: the front nasal vowel, which is diphthongized by adding a palatal glide, and the back nasal diphthong, which is the result of a transformation of the syllabic vowel into a velar nasal semivowel/semi-consonant.

The soft palate movement and the forward movement of the tongue, associated with the soft palate movement and the nasal segment in the tonic position, are related to nasal diphthongization.

For us, laboratory or experimental phonology is the methodological model that can encompass all the symbolic and spatiotemporal aspects of speech, bringing phonetics and phonology together, without any kind of gap.

This approach considers the abstract structure of the speech production system, as well as its articulatory and acoustic properties, and the spatiotemporal variations involved. It acknowledges that speech production is inherently variable, and there are gradient boundaries between gestures that can vary based on factors such as prosody, biomechanics, and, within certain limits, articulatory variation.

As mentioned earlier, our discussion will focus on two contemporary theories from the field of experimental phonetics and symbolic phonological representations: the theory of speech control developed by Kingston & Diehl in 1994, and the theory of articulatory phonology proposed by Browman & Goldstein in 1992.

The integration of these theories offers significant potential for gaining valuable insights into the functioning of the speech-motor system and its underlying architecture. Turk & Shattuck-Hufnagel (2020) particularly emphasize the importance of exploring the association between these theories to enhance our understanding of speech production processes.

Articulatory phonology is based on the mapping of the temporal trajectory and co-articulatory effects during speech production. The theory of speech control specifically focuses on the planning aspects of discrete symbolic phonological representations. It delves into how speech production is organized and coordinated at the level of phonological units, such as phonemes or syllables.

This theory seeks to understand how these symbolic representations are translated into motor commands and executed by the speech-motor system. By studying the planning processes involved in speech production, the theory of speech control provides valuable insights into the cognitive and motor aspects of producing linguistic units during communication.

In the context of articulatory phonology, the concept of “mapping” is essential for understanding various aspects of speech production, including nasal diphthongization. The mapping of the temporal trajectory helps us comprehend how the timing and coordination of articulatory gestures correspond to nasal diphthongization, exploring how the temporal organization of articulatory movements



aligns with the hierarchical structure of linguistic units, such as phonemes or syllables, during the production of nasalized diphthongs.

By examining the relationship between the temporal trajectory, coarticulation, and linguistic representations, we gain insights into the intricate mechanisms underlying nasal diphthongization.

Considering these theories as complementary, they can help us understand the mechanisms that influence changes in spatiotemporal variations in phonological representations. We believe that these theories help describe and analyze nasal diphthongs.

We selected these theories for their methodological approach, as they both originate from a phonetic and phonological perspective that considers the abstract and physical aspects of speech production. These theories are generally concerned with coordinated gestures, aiming to identify and define the phonemes present in languages worldwide.

Our choice of these theories allows us to gather evidence and investigate various aspects of nasal diphthongization in speech production. Although the theories themselves do not possess a specific methodological approach, they have been developed and utilized within well-defined methodological frameworks to explore the phenomenon of nasal diphthongization.

The methodological approaches associated with these theories enable us to design experiments, collect data, analyze findings, and test hypotheses systematically. This systematic approach allows us to examine the mechanisms and factors involved in nasal diphthongization. By employing these methodologies, we deepen our understanding of how nasal diphthongs are produced and shed light on the underlying processes and patterns associated with nasalization in speech.

The theoretical approach of this thesis begins with a discussion of phonetics and phonology; that is, a debate about articulatory phonology (BROWMAN & GOLDSTEIN, 1992) and the theory of speech control (KINGSTON & DIEHL, 1994). We also present an overview of BP, including its phonetics and phonology.

We structured this thesis to encompass a theoretical framework and practical analysis. The primary objective of our aerodynamic-acoustic and articulatory experiments was to explore the relationships between time segments, which refer to specific temporal units or intervals within the process of speech production.

We focused on investigating the timing and duration of phonemes, aiming to understand how these temporal intervals relate to various aspects of aerodynamics, acoustics, and articulation. By examining these relationships, we aimed to gain insights into the temporal organization of speech production.

Evaluation of Aerodynamic EVA2. In the articulatory study, the device used was Midsagittal Electromagnetic Articulography (EMMA), it was used to identify some patterns that classified n-like vocalic sounds as a phoneme. We also established a corpus to compare oral and nasal diphthongs to determine the patterns in the distribution of sounds.

The first step was to analyze the aerodynamic and acoustic data; the results allowed us to infer the position of the tongue and the velum from the variations in the airflow. Nevertheless, as we aimed to clarify certain articulatory characteristics, it was necessary to design an EMMA experiment.

Due to the constellation of articulatory gestures, we assumed that a velar or palatal constriction occurred during the production of nasal diphthongization; that is, total or partial blockage of the oral cavity occurred at the end of the nasal glide due to the oronasal coupling.

We hypothesized that the tongue was raised in the palatal or velar regions when the soft palate was open. Thus, these inferences about the articulatory behavior of the tongue and velum articulators reflected the gestural coarticulatory configuration based on assumptions about the aerodynamic consequences with regard to the shape, duration, volume of the oral, and nasal airflows.

The experimental design proposed for this work consisted of three types of empirical analyses, namely acoustic, aerodynamic, and articulatory. The original aim of this thesis was to create a dialogue prescriptively among the three methods of the

study mentioned above to obtain an articulatory view of the behavior of the phenomenon in the phonation process.

The diverse experimental methodologies employed in this thesis aimed to create a consistent set of occurrences of the phonetic phenomenon observed during the production of the specific phonemes mentioned. Through experimental research, which is a methodology used to test models and theories (DEMOLIN, 2012),

Our articulatory and imaging studies will provide valuable insights into the movement of the velopharyngeal sphincter during speech production, specifically focusing on the gestures of the velum and tongue. These studies offer direct information about how these gestures shape speech sounds. An important aspect of our analysis is the adoption of an intrinsic temporal perspective, which involves examining the characteristics and dynamics of these gestures concerning time and the inherent temporal properties of speech production.

By adopting an intrinsic temporal perspective, we aim to understand the timing, duration, sequencing, and coordination of the velum and tongue movements during speech production. This perspective allows us to uncover the temporal patterns and organization that govern these articulatory gestures.

It is important to note that our analysis is influenced by an articulatory bias, as we prioritize the articulatory aspects of speech production. However, by focusing on the intrinsic temporal perspective, we gain a deeper understanding of how these gestures unfold over time and contribute to the overall temporal structure of speech.

Through our articulatory studies, we will examine the intricate temporal relationships between the velum, tongue movements, and the production of speech sounds. By doing so, we aim to shed light on the temporal organization and dynamics of speech production, providing a comprehensive understanding of the role played by the velopharyngeal sphincter in shaping speech sounds.

In our study, various experimental approaches were employed to provide a comprehensive description of synchronous nasal production. Through this, we aim to

illustrate the general properties of nasality and identify the characteristic features of nasal diphthongization.

Our experimental investigation entailed the production of the back-nasal diphthong [ãw̃] and the front nasal vowel [ẽ̃]. Most of the target words were monosyllables in which the target phenomena occur in the offset position. We only discuss these nasalized segments to generalize some features pertaining to the nature of these sounds.

To achieve this, a corpus of words that could be used in the four types of experimental analyses, which constitute this dissertation, was proposed. The word is in the coda position in the target words analyzed, thus making a boundary between voiceless consonants. The internal particle of the articulatory gesture is composed of a group articulator.

Our goal is to show that both the nasal vowel and the nasal diphthong underwent the phonetic process of nasal diphthongization. It suggests that these specific speech sounds exhibit characteristics or patterns associated with nasal diphthongization. Nasal diphthongization refers to a phonetic process in which a nasal vowel or a nasalized vowel transitions into a nasal diphthong, typically involving a change in the quality or articulation of the sound. This process may occur due to various factors, such as coarticulation effects or phonetic assimilation.

As mentioned previously, nasal vowels have been studied more often than diphthongs, including from an experimental perspective. Investigating and analyzing nasal vowels and nasal diphthongs in our study, we provide evidence or observations that support the occurrence of nasal diphthongization in both speech sound categories. This can contribute to our understanding of how nasalization affects vowel and diphthong production and shed light on the phonetic processes involved in nasal diphthongization.

Nasal diphthongs exhibit complex temporal dynamics and transitions between two vowel qualities, making them an intriguing phenomenon to study from a temporal perspective. Investigating the timing, duration, and coordination of articulatory

movements during diphthongs can provide valuable insights into the temporal organization and dynamics of speech production. This knowledge contributes to our understanding of the underlying mechanisms of speech-motor control and the coordination of articulatory gestures.

Furthermore, studying diphthongs contributes to our understanding of coarticulation and phonological processes. Diphthongs often exhibit coarticulatory effects, where the gestures involved in the transition between the vowel qualities influence each other. By examining these coarticulatory effects, we can better comprehend how speech sounds are produced in context and how phonetic and phonological factors interact in shaping the acoustic properties of speech.

Therefore, it is necessary to characterize nasal diphthongs and nasal diphthongization in terms of their phonetic status. Additionally, we believe that describing and analyzing their properties are crucial for the field and the characterization and parameterization of natural languages.

To achieve a comprehensive understanding of nasality coarticulation, our study will consider essential topics such as the gestural nature, temporal aspects, speech dynamics, phonetic control, spatial variation, and airflow variation. We believe that these parameters will encompass the typology of nasalization and nasal diphthongization.

Consequently, we examined the nasal details using physical representations (articulatory and aerodynamic measures) and fiberscope observations to analyze nasal diphthongs as dynamic segments. This perspective can enhance our understanding of speech control (KINGSTON & DIEHL, 1994), gestural coordination, and the production of nasal diphthongs and nasal vowel diphthongization.

We address several parameters of the chosen phenomenon based on the theory of speech control to identify the control of variables in the phonatory system, which contributes to the perception of speech production. Motor programming governs the system's variables, while the neuromuscular mechanism activates them during motor tasks (HUBER et al., 2004). Examining the timing, sequencing, and articulatory

movements involved in producing oral and nasal sounds, with a specific focus on coordinating the movements of the tongue and velum, is crucial for understanding how airflow control distinguishes between oral and nasal sounds.

Furthermore, variations in airflow can significantly impact speech clarity and precision, highlighting the significance of consistent and appropriate airflow volume for accurate and intelligible speech production. By exploring the implications of airflow variations and articulation, this research aims to illuminate the fundamental mechanisms underlying speech production and contribute to our understanding of gestural control in the context of speech.

Given the complexity of the phenomena under investigation, we rely on laboratory phonology, control theory, and articulatory phonology to provide a foundation for analyzing our data and interpreting the results of airflow variations and articulatory gestures.

The underlying phonological representation assumed for nasal vowels and nasal diphthongs is based on the concept of nasalization. Nasalization refers to the phonetic process by which a vowel sound becomes nasalized due to the influence of adjacent nasal consonants. In the context of classifying nasal diphthongs, the underlying representation assumes that nasalization affects the vowel portion of the diphthong, resulting in a nasalized vowel.

The specific underlying representation of nasal vowels and nasal diphthongs may vary depending on the theoretical framework and linguistic analysis being employed. However, generally, it is assumed that nasalization is phonologically represented by the presence of a nasal feature or nasal place of articulation associated with the vowel segment. This representation indicates that the vowel is produced with a lowered velum, allowing air to pass through the nasal cavity during its articulation.

Following a brief review of the phonological literature on BP, the underlying phonological representation of nasal vowels was used for the classification of nasal diphthongs.

The most widely accepted thesis is that proposed by Camara Jr. (1964), in which the nasal vowels contrasted with the oral vowels by adding the [+ nasal]. The vowel assimilates the [+ nasal] of the homorganic nasal consonant, and the nasal diphthongs also follow the same phonological rule.

Most of the existing studies have classified nasal/nasalized vowels as a binary segment; that is, [ $\pm$ nasal]. The oral and nasal opposition is described as a binary feature in the world's languages. This theory, associated with the work of linguists such as Roman Jakobson and Morris Halle, proposes that nasalization is represented by the presence of a distinctive feature [+nasal] in the underlying representation of nasal vowels and nasal diphthongs. The nasal feature distinguishes them from their oral counterparts.

However, the fundamental articulatory property of nasalized sounds has not yet been clarified (HAWKINS & STEVENS, 1985).

The core characteristic or mechanism involved in the production of nasalized sounds is not fully understood or clearly defined. There is still uncertainty or lack of consensus regarding crucial aspects of nasalized sounds.

In this thesis, we will not conduct a linear analysis of our data, following structuralism. While binary distinctive features, such as [+nasal], are commonly used to represent nasalization in phonological analyses. The features may not capture the complete range of articulatory details and variations involved in producing nasalized sounds. We light the velum lowering as a primary mechanism associated with nasalization.

Some aspects of nasalization that are not fully understood, such as the precise coordination and timing of articulatory movements, the degree of velum lowering, the role of nasal cavity resonance, and articulatory gesture interactions will be lighted in this research.

As mentioned, Cagliari (1977) proposed an experimental hypothesis regarding temporal and dynamic variations in speech production. According to this hypothesis, nasal vowels and diphthongs can exhibit different degrees of nasalization, and both segments are susceptible to diphthongization.

Cagliari suggests that nasal vowels and diphthongs can exhibit diphthongization to different extents, implying that they may have transitional or glide-like elements within their articulatory patterns.

The purpose of our study is not to test Cagliari's proposal regarding the articulation of the velum and its relationship to nasalization. But it is employing different experimental methods as the author did, to study aerodynamic and articulatory functions relating to the timing of speech, and to gain insights into the mechanisms of variations in nasalization.

Cagliari's proposal might have been driven by the need to better understand the role of the velum and its articulatory functions in nasalization. By investigating different degrees of nasalization, Cagliari likely aimed to uncover the variability and nuances in the nasalization process and how they manifest in nasal vowels and diphthongs. Different degrees of nasalization could result from variations in the position and movement of the velum, airflow patterns, and other articulatory factors, leading to distinct acoustic and perceptual characteristics.

By examining and analyzing these aspects, it can contribute to our understanding of the mechanisms underlying nasalization, which is important for the broader field of phonetics and phonology. Understanding the reasons for different degrees of nasalization can provide insights into the nature of nasal sounds and their linguistic functions within a given language system.

Therefore, it is essential to describe the articulation of the velum via different experimental approaches such as the aerodynamic and articulatory functions (CAGLIARI, 1977). Accordingly, the theme of this research is vital for the field of phonetic articulation to elucidate the articulatory and biomechanical aspects of nasal diphthongs that have not yet been studied.



According to studies by Shosted (2006) and Lipski (1975), the nasal appendix in the Carioca dialect of Brazilian Portuguese exhibits articulatory properties like a velar occlusive consonant /ŋ/.

However, Lipski suggests that the nasal consonant may have a degree of critical constriction rather than a full specification. Investigating these hypotheses can provide insights into the nature of the nasal appendix and its role in nasalization. Analyzing words with nasal appendices, such as "*pente*" [ˈpẽ̃ntɨ], can provide empirical evidence supporting the hypotheses proposed by Shosted (2006) and Lipski (1975).

The presence of an underlying nasal consonant in an offset position leads to the emergence of the nasal appendix, contributing to nasal diphthongization. Various studies (e.g., MEDEIROS ET AL., 2008; SOUSA, 1994; JESUS, 2002; SEARA, 2000) have focused on investigating the production and articulatory behaviors of the nasal appendix, particularly in the Carioca dialect.

Nasalization in Brazilian Portuguese can occur through nasalization by coda or nasalization by onset. In nasalization by coda, the nasal appendix is interpreted as a full nasal consonant [n], leading to its re-syllabification from the coda to an onset position.

Nasalization by onset involves the spreading of the nasal feature from a nasal consonant to the vowel, either regressively to the left (heterosyllabic) or to the right (tautosyllabic) within the syllable. Nasal spreading is optional, and the underlying form of the words does not contain the nasal feature.

The status of nasal vowels in Brazilian Portuguese is a topic of debate. Some argue that the nasality in nasal vowels is coarticulatory, attributed to the presence of the nasal appendix. This perspective suggests that nasal vowels should not be considered distinct phonemes. However, if vowel nasality is inherent and not coarticulatory, nasal vowels would be phonemic, with the nasal appendix being an integral part of the vowel (MARQUES & SCARBOROUGH, 2020).

Nasal diphthongization is a phonetic process that involves the coupling of the oral and nasal resonating cavities, along with tongue movement. It results from

synchronized and controlled gestural displacement, generated by multiple articulatory gestures.

These gestures exhibit different degrees of dependence, influenced by individual spatial-temporal properties of phonemes, prosodic factors, and speech rate. Understanding these gestural associations is essential for comprehending the nuanced meaning of nasal sounds.

In future research, our data will enable us to replicate our generalizations in other similar segments, such as nasal vowels occurring in the nasal tonic or coda position: “**tanta**” [ˈtã.te] (much) or ˈkã.ter “**cantar**” (to sing), “**tentar**” [ˈtẽ.ter] (to try), “**pintar**” [ˈpĩ.ter] (to paint), “**tonta**” [ˈtõw̃.te] (dizzy), [ˈʒũw̃.ter] “**juntar**” (to joint) or “**tom**” [tõw̃] (tone), “**bom**” [bõw̃] (good) and “**Rum**” [hũw̃] (Rum).

We also aim to question, test, confirm, or refute existing scientific theories in the field, particularly understanding the relationship between nasalization and vowel pronunciation. One theory might propose that nasalization affects the quality or duration of vowels in certain linguistic contexts. To test this, we will examine the acoustic properties of vowels in nasal and non-nasal contexts and compare them.

Furthermore, we will analyze the articulatory mechanisms of nasalization by describing the specific articulatory adjustments during the production of nasalized vowels and diphthongs. Techniques such as Electromagnetic Midsagittal Articulography or imaging techniques like fiberscope and photo-nasography will be employed to examine the articulatory patterns associated with nasalization.

Through a comparative analysis of similar segments, we can delve into the intricate details of nasalization and nasal diphthongization, shedding light on the phonetic and phonological mechanisms underlying these phenomena. Systematic examination of the acoustic properties, duration, and spectral characteristics of nasal vowels and nasal diphthongs, along with their corresponding oral counterparts, as well as considering articulatory factors like onset, target, offset, and airflow variation, will facilitate a comprehensive understanding.

Analyzing the acoustic properties allows for a quantitative assessment of the distinct features exhibited by nasal vowels and nasal diphthongs compared to their oral counterparts. We will explore changes in formant frequencies and intensity levels, unveiling patterns associated with nasalization. These acoustic characteristics serve as markers for differentiating nasal and oral sounds, contributing to an overall understanding of the effects of nasalization on vowel pronunciation.

Examining the articulatory aspects, including onset, target, and offset, will provide valuable insights into the specific gestures and movements involved in producing nasal vowels and nasal diphthongs. By investigating the coordination of articulatory gestures during these segments, we will uncover the precise mechanisms that govern nasalization and nasal diphthongization. This knowledge contributes to a deeper comprehension of how nasalization impacts vowel production.

Moreover, analyzing airflow variation during the production of nasal vowels and nasal diphthongs offers further insights into the phonetic and phonological processes at play. By studying the airflow patterns, we will uncover the influence of nasalization on speech production and gain a more nuanced understanding of the aerodynamic principles involved in nasal vowel and nasal diphthong production.

Conducting this comprehensive comparative analysis will allow us to identify commonalities and differences between nasal and oral segments, establish generalizations regarding the effects of nasalization on vowel pronunciation, and uncover the underlying principles that govern these phenomena. The findings from these investigations will provide valuable insights into the relationship between nasalization and the phonetic and phonological mechanisms of speech production.

We used the Pearson correlation coefficient to quantitatively analyze the data. This is a statistical test that measures the strength of the relationship between two continuous variables. Since we don't have an extensive database, we applied a descriptive test called Pearson Correlation to some fiberscope and articulatory data.

To accurately represent the relationship between elements linear. The Pearson correlation coefficient, which ranges from +1 to -1, is used to assess this relationship.

A value of zero indicates no correlation between the variables, while a positive value indicates a positive correlation, meaning that as one variable increases, the other also tends to increase.

Performing a Pearson's correlation requires that both variables are measured on an interval or ratio scale. This means that the data points should be spread out across the range of values rather than being heavily skewed or clustered in specific areas. Additionally, the association between the variables should be linear to ensure accurate interpretation of the correlation analysis.

### **1.1 Objectives**

The scope of this research is to understand how the articulatory structures, such as the velum and nasal passages, function in producing nasal diphthongization of the sounds involved in opening and closing the velopharyngeal port and their precise movements and coordination. Our objectives are:

- To describe the aerodynamic, acoustic, and articulatory characteristics of the front nasal vowel and the back-nasal diphthong produced in the *Paulistano* dialect of BP.
- To compare the nature of the gestures in the oral and nasal segments.
- To establish a connection between the physical act of producing speech and more abstract linguistic concepts.

In this study, we aim to investigate the differences in the physical movements of the tongue and velum during the production of the front nasal vowel and back-nasal diphthong in the Paulistano dialect of BP. We will focus on the gestures involved in the oral and nasal segments of these sounds, comparing their timing, degree of movement, and coordination with other articulators. The oral segment is assumed to involve only movements of the tongue and lips, while the nasal segment involves the movement of the velum and redirection of airflow through the nasal cavity.

The study of nasal diphthongization requires investigating the coordination and timing of oral and nasal gestures to gain a better understanding of how abstract concepts such as phonemes and syllables are realized in concrete speech production. To achieve this, we will be comparing the nature of these gestures, with the "nasal segments" referring to the portion of the sound where the velum is open, and airflow is directed through the nasal cavity.

By linking the physiological and articulatory features of speech to abstract concepts, such as phonetic parameters and phonological models, we can create a bridge between the physical and abstract aspects of speech, enabling a more complete understanding of the mechanisms underlying speech production. Through the investigation of the details of speech production and the comparison of the oral and nasal gestures involved in these sounds, we aim to gain insight into the mechanisms of nasal vowel production in BP.

This introduction covers our research questions and objectives too; the theoretical framework covers the anatomy, the velopharyngeal port, and the physiological structures involved during nasalization. The following chapters describe the experimental aerodynamic, fiberscope, and articulatory analyses with each experimental approach being explained in a dedicated chapter.

## **1.2 Research Questions**

The research questions were developed based on gaps in the existing literature and the need to understand the mechanisms involved in the production of nasal diphthongs in Brazilian Portuguese. Each question addresses a particular aspect of the production of nasal diphthongs, aiming to provide a more detailed understanding of the physical and cognitive processes involved.

- How is the role of tongue adjustment involved in the production of nasality?
- Is the tongue movement equal during the production of the oral and nasal segments?

- What is the relationship between nasal airflow variation and tongue movement?
- What are the mechanisms the speaker uses to specify space and time during speech production?
- How and why does motor synchronization generate emergent properties during speech production?
- How is the velum displaced?
- What is the degree and the size of the velum opening?
- How does velarization or palatalization occur during speech production?
- Are the nasal features adaptive or cognitive?

For instance, the question about the role of tongue adjustment in nasality explores the contribution of the physical movements of the tongue to nasality. The question about the mechanisms for specifying space and time during speech production investigates the cognitive processes underlying the production of complex sounds like nasal diphthongs.

Overall, the research questions aim to provide a comprehensive understanding of the production of nasal diphthongs in Brazilian Portuguese and contribute to the broader field of phonetics and phonology.

In addition to the above, the question about the degree and size of the velum opening investigates the physical mechanisms involved in the production of nasality, specifically the movement of the velum. The question about soft palate movement is related to this and seeks to understand how the soft palate contributes to the production of nasal sounds.

Similarly, the question about the relationship between nasal airflow variation and tongue movement aims to explore the coordination between these two mechanisms in the production of nasal diphthongs. The question about the adaptive or cognitive nature of nasal features aims to understand how language acquisition and cognitive processes may affect the production of nasal sounds.

By answering these research questions, the study aims to deepen the understanding of the mechanisms involved in the production of nasal diphthongs in Brazilian Portuguese and contribute to the broader field of speech production research.

### **1.3 Theoretical Framework**

This overview aims to provide a brief outline of the theories in articulatory phonology, control speech, and timing in speech that are essential for understanding nasal diphthongization and speech production. These theories are relevant for understanding the mechanisms involved in nasal diphthongization and speech production in general. While they do not specifically focus on sound change, they provide insights into the articulatory gestures, coordination, and movement of the articulators involved in producing nasal diphthongs.

By examining these theories, we can gain a better understanding of how nasal diphthongs are formed, how the nasal appendix interacts with the surrounding articulatory gestures, and have insights into how the timing of gestures influences the production of nasal diphthongs. This knowledge contributes to our understanding of the phonetics and linguistics of nasal diphthongs, including their potential role in sound change processes within Brazilian Portuguese.

While the focus of the study may not be on sound change directly, understanding the underlying mechanisms and articulatory processes involved in nasal diphthongization can have implications for broader phonetic and phonological analyses, including the potential effects of nasalization on vowel pronunciation and the phonemic status of nasal vowels in Brazilian Portuguese.

Articulatory phonology theories focus on the physical aspects of speech production and explain how articulators, such as the tongue, lips, and velum, are used in producing speech sounds. These theories provide insights into the coordination and movement of these articulators during the production of nasal diphthongs. They help us understand the role of the tongue and velum in shaping the oral and nasal cavities,

controlling airflow, and producing specific articulatory gestures required for nasal diphthongization.

Control speech theories address how speakers regulate and control their speech production. They explore the mechanisms involved in coordinating different articulatory gestures and timing them accurately. These theories can help us understand how speakers adjust the timing and coordination of tongue and velum movements to achieve the desired nasal diphthongization patterns.

Timing in speech theories focus on the temporal organization of speech production. They examine the precise timing and sequencing of articulatory gestures and how they contribute to the production of specific speech sounds. Understanding the timing aspects of nasal diphthongization can provide insights into the temporal coordination between the oral and nasal segments and the emergence of glide constriction as a property associated with nasal diphthongs.

By applying these theories to the study, we can gain insights into the underlying mechanisms involved in their production. This includes understanding the coordination between the oral and nasal cavities, the role of the velum and tongue in regulating nasal airflow, and the articulatory strategies used to produce complex nasal sounds. These theories help us analyze and explain the intricate processes and interactions that contribute to nasal diphthongization.

Therefore, an understanding of relevant theories in articulatory phonology, control speech, and timing in speech is crucial for gaining a comprehensive understanding of nasal diphthongization and speech production. It provides the necessary theoretical framework and analytical tools to investigate the mechanisms and patterns involved and to interpret the findings of this study context.

Nasal diphthongs, a relatively uncommon phenomenon in world languages, consist of a unique and inseparable combination of a nasal vowel and a nasal glide, resulting in a complex syllable nucleus. The nasal diphthong encompasses two distinct elements: the nasal vowel, characterized by the passage of air through the nasal cavity during its production, and the nasal glide, a semivowel/semi-consonant



created by the upward movement of the tongue and its positioning towards the palatal/velar region of the oral cavity.

In the context of speech production, the nasal diphthong presents an intriguing case due to the intricate interplay between the oral and nasal articulatory mechanisms. The nasal vowel requires the coordination of the velum and the vocal tract to allow the passage of air through the nose, while the nasal glide involves specific tongue adjustments to create the desired acoustic output.

Understanding the precise mechanisms involved in the production of nasal diphthongs is crucial for advancing our knowledge of speech production and phonetics. It involves exploring how the velum and the tongue interact and adapt during the production of these complex sounds. The role of the velum is particularly significant, as it regulates the balance between oral and nasal airflow, ensuring the appropriate resonance and articulation of the nasal diphthong.

The nasal glide acts as a consonant-like element in the syllable, closing off the airflow and shaping the resonance of the sound. The pronunciation of nasal diphthongs varies across languages and dialects, influenced by factors such as the degree of nasalization, tongue position, and airflow through the nose.

In the context of laboratory phonology and articulatory phonology, our goal is to demonstrate that the coordination between oral and nasal gestures in the production of nasal diphthongs is a result of a gradient overlap. This means that the gestures involved in producing the nasal vowel and nasal glide blend gradually rather than abruptly.

By investigating the details of this overlap and the associated aerodynamic and articulatory properties, we hope to gain a deeper understanding of the mechanisms underlying nasal diphthongization that involves the co-production of consecutive segments with varying degrees of overlap.

We consider the timing and coordination of oral and nasal gestures during speech production, as the duration of an oral vowel may affect the timing and degree of nasalization. Therefore, our analysis takes into consideration the spatiotemporal

properties, such as the physical characteristics of the speech signal, such as duration and amplitude.

Regarding prosodic factors and speech velocity (FOWLER, 1977, 1980) it is important to note that they can also influence nasalization patterns. Prosodic features such as stress, intonation, and phrasing may have an impact on the degree of nasalization in certain contexts. However, in this research study, our focus is primarily on the spatiotemporal aspects and physical characteristics of nasalization, and we do not extensively explore the influence of prosodic factors. This decision is made to maintain the scope and clarity of our analysis.

Abstract concepts in the context of speech production and phonetics involve theoretical constructs used to describe and analyze speech sounds. They encompass phonetic parameters, which represent properties like place of articulation, manner of articulation, and voicing.

These concepts also include exploring the interaction of nasalization with vowel quality, consonant assimilation, and syllable structure. The goal of connecting concrete speech production with abstract concepts is to understand physical articulatory gestures and the linguistic structures and patterns that form the basis of language sounds.

By examining these connections, insights can be gained into how the concrete aspects of speech production contribute to abstract representations and language production processes.

To connect the concrete production of speech with abstract concepts, we aim to study the physiological and articulatory features of speech. This approach allows us to represent significant variations in phonetic parameters, analyze phonological models, and abstract representations.

Different theories offer varying insights into the nature and organization of nasalization in language. In our study, we have adopted different frameworks to better understand nasal diphthongization and its associated processes.

Generative Phonology primarily focuses on abstract phonological representations and rules, which may limit its ability to capture the detailed articulatory information involved in nasal diphthongization. This complex articulatory process requires a more fine-grained analysis of the coordination between oral and nasal gestures, which may not be fully accommodated within the Generative Phonology framework.

In Generative Phonology, speech sounds are represented as underlying forms (URs) that undergo transformation into surface forms (SFs) through a set of phonological rules. These rules account for observed variations in both phonetics and phonology. Nasalization rules are employed to introduce nasal features or nasalize specific segments within appropriate phonological contexts, determining the timing and conditions of nasalization.

However, the representation of nasal diphthongs poses a challenge within the discrete segment-based framework of Generative Phonology. The simultaneous nasalization and glide movements involved in nasal diphthongs are not easily captured by this approach.

Furthermore, nasal diphthongization exhibits cross-linguistic variation, with different languages and dialects employing various strategies to realize nasalized diphthongs depending on the surrounding context. Generative Phonology may struggle to account for these language-specific variations, as its general rules and representations may not readily accommodate the specific phonetic and phonological patterns observed in particular cases, such as the *Paulistano* dialect.

Nasal diphthongization, being a complex articulatory process, requires a more fine-grained analysis of the coordination between oral and nasal gestures, which may not be fully captured within the Generative Phonology framework.

The background to this thesis is the main studies of nasality in the vowel segments as a complex subject with extensive research based on speech production (LADEFOGED, 1974). In a comprehensive literature review, it is important to examine

and discuss relevant studies that have explored nasality in vowel segments, considering both speech production and perception.

By reviewing the works of Cagliari (1977), Krakow & Huffman (1993), Dang, Honda, & Suzuki (1994), Feng and Castelli (1996), Teixeira, Moutinho, & Coimbra (2001), Delvaux (2003), Delvaux et al. (2002), Rossato, Badin, & Bouaouni (2003), Amelot (2004), Serrurier & Badin (2005), Shosted (2006), Medeiros, D'imperio, & Espesser (2008), Medeiros & Demolin (2006), and Turk & Shattuck-Hufnagel (2020), I tried to identify the key findings, methodologies, and theories proposed by these researchers.

In our literature review, we discuss how these studies have contributed to understanding nasality in vowel segments, and their methodologies for investigating speech production. We identified gaps in the existing literature in analyzing the production of nasal diphthongization.

Our review aims to examine the anatomical and physiological aspects of the phenomenon of diphthongization by presenting phonological and phonetic perspectives pertaining to nasal vowels, nasal diphthongs, and the essential elements of nasal diphthongization in the Paulistano dialect of BP.

### **1.3.1 Phonetics and Phonology**

The current understanding of Phonetics and Phonology acknowledges the importance of laboratory data in providing reliable descriptions and explanations of phonetic-phonological phenomena. It is now essential to have a well-equipped phonetics laboratory for capturing and analyzing speech sounds, enabling scholars to acquire knowledge and conduct research (RAPOSO DE MEDEIROS & DEMASI, 2006).

We intend to maintain a close relationship between speech sounds and Phonology, as well as language in general. We aim to describe language as it is produced by speakers, emphasizing its characteristics of speech (CAGLIARI, 2006).

Phonetics is concerned with oral language and the language system, which leads to a reinterpretation from a phonological perspective (CAGLIARI, 2006).

The discussion of phonetics and phonology is directly relevant to our study on nasal diphthongization in the Paulistano dialect because the distinction between these two sub-fields of linguistics had general relevance to the study of speech production and sounds in natural languages.

Since the Prague Circle<sup>8</sup>, there has been a complete separation between the phonetic and phonological fields. The distinction between Phonetics and Phonology was evident. The separation of these domains was based on the opposition between the concrete and the abstract (TRUBETZKOY, 1981 [1933]).

There was a tension between the fields of phonetics and phonology. The distinction between them has become the main exemplification of the opposition between language and speech. Trubetzkoy's proposal highlighted the need for separating these areas as an idea, even if not fully realized in practice.

There was a tendency across these disciplines to shift from atomism to structuralism and from individualism to universalism. Phonology is part of this broader scientific movement, and it was hoped that other areas of linguistics, such as morphology, syntax, lexicology, and semantics could be similarly merged into phonology (TRUBETZKOY, 1981 [1933]).

Phonology emphasized the fundamental difference that existed between itself and phonetics. Phonetics focused on studying the physical aspects of human speech sounds, such as air vibrations and articulatory movements. In contrast, phonology studied phonemes, which were the immaterial constituent elements of the linguistic signifier (TRUBETZKOY, 1981 [1933]).

Phonetics and phonology have been different fields, although both focus on the same object, namely the sounds in natural languages. In the context of our study, the

---

<sup>8</sup> Prague Circle was a group of linguists active during the early 20th century in Prague, Czechoslovakia (now the Czech Republic). The circle was formed around the influential linguist Vilém Mathesius, Roman Jakobson and Nikolai Trubetzkoy. They were known for their contributions to the development of structuralist and functionalist approaches to linguistics.

concepts of phonetics and phonology are indirectly relevant. Understanding the phonetic properties of nasal diphthongs, such as the articulatory gestures and physical aspects of their production, falls within the realm of phonetics. On the other hand, analyzing the role of nasal diphthongs within the phonological system of Brazilian Portuguese, including their potential phonemic status and effects on vowel pronunciation, aligns with the domain of phonology.

Phonology studies the speech system using a purely linguistic method that considers the study of speech production through the lens of a logical and functional object structure, while phonetics studies the concrete act of speech using experimental tools in the natural sciences to analyze physical aspects of the processing, production, and perception of sound signals (TROUBETZKOY, 1949).

Trubetzkoy's studies (1981) outlined the methodological differences between phonetics and phonology; according to the author, phonetics is 'founded on physiology and physics' (TRUBETZKOY, 1981, p. 16), as sound is a psychophysical phenomenon, while phonology is based on phonic images with linguistic features. Both fields went their separate ways.

According to Trubetzkoy (1981), the distinction between sound and phonemes is a linguistic one, and this point of view created a gap between the linguistic (phonological) and the non-linguistic (phonetic) disciplines.

Nonetheless, it is often difficult to determine which of the disciplines explains the origin of some phonological phenomena clearly (TRUBETZKOY, 1981, p. 18). Structuralism has placed phonetics in a borderline position within the linguistic field, at least in theory.

Phonetics, the study of speech sounds, is concerned with representing these sounds. However, this raises the question of what constitutes a speech sound and, consequently, what speech itself entails (LADEFOGED, 2004). The sound wave carries a wealth of information, not all of which is language related. relationship between phonetics, which deals with the actual sounds, and phonology, which examines how these sounds form significant linguistic distinctions.

Phonetics deals with the actual sounds of speech, while phonology examines how these sounds form meaningful linguistic distinctions. Ladefoged argues that if phonology is limited to studying aspects of speech that convey information through a formal system of contrasts, then the relationship between phonological and phonetic units becomes more straightforward.

According to the author, phonologists are no longer obligated to describe the specific phonetic details that distinguish one speaker, dialect, or language from another. Instead, phonological analysis can focus on the formal contrasts within a language. This view differs from previous perspectives, including those of Ladefoged (1971), Chomsky & Halle (1968), and Schane (1965), which aimed to characterize both the internal oppositions within a language and the differences between languages.

Studying the general property of human language is needed for an overall phonetic framework comparing sounds among languages and the identification of phonological universals. This framework would require a set of possible sounds and criteria for determining when a sound in one language is equivalent to a sound in another.

We also believe that structuralist theories led to the development of acoustic and articulatory phonetic knowledge and the refinement of phonological methodology. Phonetics developed via Feature Theory (JAKOBSON, FANT, & HALLE, 1952), binarity in Generative Theory (CHOMSKY & HALLE, 1968), Articulatory Phonology (BROWMAN & GOLDSTEIN, 1992), and Quantal Theories of speech production (STEVENS, 1972).

Even the International Phonetic Alphabet (IPA) was established to provide a standardized set of symbols representing speech sounds, aiming to equate sounds found in multiple languages. Both the IPA and standard phonological feature theory assume a limited set of phonetic categories.

However, the classification of sounds within these frameworks often involves ad hoc descriptions, introducing new categories, or modifying existing ones when

encountering previously unobserved contrasts. This raises concerns about the theoretical basis of the phonetic framework and its ability to account for the full range of phonetic capabilities of human language.

Auto-segmental phonology<sup>9</sup> assumed that phonological representations were essentially a linear sequence. Auto-segmental phonology introduces the concept of autosegments, which are independent units of phonological representation that can be associated with different segments.

This theory proposed departures from strictly linear “beads-on-a-string” kinds of representations — crucially tones could be associated with multiple segments and multiple segments could be associated with single tones.

In this framework, phonological processes are represented as interactions between these autosegments rather than operations on individual segments. This allows for a more flexible and accurate description of various phonological phenomena, including tone, nasalization, vowel harmony, and reduplication.

One of the key insights of auto-segmental phonology is that different features or phonological elements can be associated with distinct tiers, which are parallel representations. Consequently, speech was represented by segments with abstract representations, but without any temporal correlation with the representation of speech, except linear temporality. That is, each segment follows the other element.

Language is not only primarily a formal system in the mind of an individual speaker. Instead, language should be considered an observable social institution with its principles and self-organizing properties. Language involves various components and processes, resulting in a complex and multifaceted nature.

Language is not solely a product of individual minds but also emerges from collective interactions and shared conventions. It is governed by principles of communicative efficiency and group identification. From this perspective, language is shaped by societal factors and exhibits self-organizing properties.

---

<sup>9</sup> Auto-segmental phonology is a linguistic framework that addresses some of the limitations of traditional segmental approaches to phonology. It was developed in the 1970s as an extension of generative phonology, primarily by linguists like John Goldsmith and Andrew Nevins.



In Turk and Shattuck-Hufnagel's theory of speech production, the planning and execution of speech involve the coordination of spatial and temporal characteristics. The planning stage involves organizing the intended sequence of speech sounds, considering factors such as the sound context, speed, and prosody. During the execution stage, the planned parameters are adjusted to the specific context of the speech production.

According to their theory, there are no default values for the parameters that need to be adjusted. The specific values of the parameters depend on the speech context and the intended linguistic message. Phonetic details, such as specific articulatory gestures or acoustic properties, are concrete aspects of speech production and do not necessarily have symbolic or representational value on their own.

The theory emphasizes the dynamic nature of speech production, where adjustments are made to accommodate the specific requirements of the speech context. The planning and execution processes work together to ensure the production of intelligible and context-appropriate speech. Thus, we would need a separate phonetic and phonological plan to map a word's concrete and symbolic features. A small number of features defines a phonological class.

The features in phonology are defined as abstract or distinctive, and they serve to differentiate one sound from another. On the other hand, the features in phonetics are concrete, measurable aspects of speech production, such as tongue position or air pressure. While there is a relationship between the two domains, the linguistic structure and the articulators' concrete, physiological structure.

In Turk & Shattuck-Hufnagel's theory, phonological categories are defined as abstract representations that lack specific details on how they should be produced. The phonetic plan, on the other hand, is responsible for specifying the concrete realization of those abstract phonological categories.

The phonetic plan maps the abstract phonological categories onto the physiological properties of speech production, such as the movements of the

articulators and the acoustic properties of the speech signal. In this way, the phonetic plan bridges the gap between abstract phonology and concrete speech production.

The abstract phonological categories are realized in concrete form through the phonetic plan, which specifies the physiological properties of speech production. The combination of abstract elements can define phonological categories, but these categories only have concrete realization through the phonetic plan. Therefore, phonetics describes linguistic functions specific to each language and its phonological universals (KEATING, 1985, 1990).

Phonetics and phonology are closely interconnected fields in the study of language. Phonetics deals with the physical aspects of speech sounds, such as their production, transmission, and perception, while phonology focuses on the abstract, rule-based organization of sounds in a particular language.

Traditionally, phonetics and phonology have been treated as separate components, where phonological representations are translated into concrete phonetic realizations. However, alternative views, such as the one proposed by Browman & Goldstein in 1992, challenge this division.

Browman & Goldstein argued that phonological representations already possess spatiotemporal characteristics, and that phonetic and phonological planning occurs concurrently. They proposed a model where the abstract representations and the physiological properties of speech production are tightly integrated. In this perspective, there is no need for a separate phonetic component to translate symbolic phonological representations into concrete speech.

Furthermore, the relationship between phonetics and phonology is not a one-way path. Phonological concepts can be influenced by the behavior of their concrete instances, as demonstrated by studies by Pierrehumbert in 1990 and Huber, Stathopoulos, and Sussman in 2004. This means that the interaction between phonetics and phonology is cyclical, with each domain influencing and shaping the other.

Browman & Goldstein (1992) presented an alternative view that challenged the traditional division between phonetics and phonology. They argued that phonological representations already have a spatiotemporal nature. So, phonetic and phonological planning occur concurrently. This perspective suggests that there is no need for a separate phonetic component to implement a symbolic phonological representation (TURK & SHATTUCK-HUFNAGEL, 2020).

Browman & Goldstein's perspective on the relationship between phonetics and phonology highlights the complexity of language as a biological system, with macro and microstructures that are interdependent. Internal relationships involve multiple reciprocity or mutual influences, and there is a linguistically relevant relationship.

It is worth noting that the timing of speech-motor planning is distinct at the phonetic level and not at the phonological level. The time is not specified in symbolic models, suggesting that separate components are responsible for planning the speech-motor time at the phonetic level.

In articulatory phonology, time is viewed as an inherent property of oscillators, which are adjusted in terms of their activation intervals based on speech planning (such as syllables, feet, and phrases), contextual variability, speech rate, and prosody. This perspective considers the spatiotemporal interaction of the articulatory components and adapts them accordingly.

Research conducted by Turk & Shattuck-Hufnagel in 2020, as well as studies by Byrd & Saltzman in 2003 and Saltzman et al. in 2008, have demonstrated that changes in the attractors (the stable states of the oscillators) do not affect the parameters of individual speech sounds. Therefore, this approach suggests that the timing of the speech is planned through separate components at the phonetic level rather than at the phonological level.

The phonetic function is to provide information to phonology about the sounds of a language, by specifying the articulatory, acoustic, and perceptual properties of speech sounds. These properties are necessary to produce and distinguish between

different sounds in a language and are used to activate or limit the production of specific sounds by the phonatory apparatus.

The phonological system relies on qualitative properties to initiate selection and switching operations. This allows for a comparison between the selection process and the functioning of the vocal apparatus, as discussed by Ohala (1983), and the switching process with the structuring of vowel systems.

The influence of the phonological system on phonetic realization is bidirectional, with feedback occurring from both phonetics to phonology and from phonology to phonetics (KEATING, 2011).

The phonological system imposes contrasts in phonetic phenomena and guides the phonetic realization of sounds, while the phonetic output provides feedback that can refine and adjust phonological representations. This bidirectional relationship helps to delineate the phonetic realization of sounds and the phonological system of a given language (LADEFOGED, 1982).

The phonological system is composed of contrasting units, which inform the phonetic realization of sounds and establish paradigmatic relationships that help to structure the system and generate phonetic output (GOLDSMITH, 2011).

Ladefoged (1982) demonstrated the influence of the phonological system on the phonetic realization by comparing the production of different natural languages that had contrasting phonological inventories.

By analyzing the presence or absence of specific contrasts in these languages, Ladefoged showed how the phonological system shapes the phonetic realization of sounds. This demonstrates the bidirectional relationship between phonology and phonetics, as phonetic output can also refine and adjust phonological representations.

Phonological inventories refer to the set of sounds (consonants, vowels, and tones) that are used in a language. The phonetic and phonological properties of these sounds are shaped by the constraints and rules of the language's phonological system.

The presence or absence of specific sounds or sound contrasts in a language's inventory can have significant implications for its phonetic and phonological specifications. Therefore, phonological inventories are seen as limiting the parameters of the phonetic and phonological specifications of a linguistic Community (LADEFOGED, 1982).

Ohala (1990) considered there to be a complete interaction of phonetics and phonology, while Kingston & Diehl (1994) posited that phonetics and phonology were two separate modules. Nonetheless, it is the interface between the two that ensures language functioning by providing the concrete production of speech and generating a mechanism that includes the phonetic realization of phonological representations.

Some previous scholars, such as Öhman (1966), have pointed out that vowels and consonants have a certain degree of independence in spoken language production. The independence of vowels and consonants in spoken language production allows them to influence each other discontinuously, meaning that changes in the pronunciation of one sound can influence the pronunciation of another sound in the same word.

This interaction between vowels and consonants is an important aspect of phonetics and phonology because it helps to maintain the unity between phonetics and phonology. By allowing vowels and consonants to influence each other discontinuously, the independence of these sounds helps to bridge the gap between the physical properties of speech sounds and the abstract rules that govern their use in the language.

Accordingly, many phonological phenomena have been described as features of earlier phonetic processes as a product of vowel-to-vowel assimilation (OHALA, 1994). Vowel-to-vowel assimilation is a type of sound change in which the vowel sound in one syllable changes to become like the vowel sound in the neighboring syllable. This process can occur due to the influence of the surrounding sounds, and it is often seen to make speech production more efficient.

Many phonological phenomena, such as vowel harmony or nasalization, have been described as features of earlier phonetic processes like vowel-to-vowel assimilation. This means that phonetic changes in the way sounds are produced can lead to changes in the patterns of how sounds are used in language.

Articulatory phonetics studies speech production from a biomechanical perspective. It treats speech production as a complex system of articulators (such as the tongue, lips, and jaw) and their movements that work together to produce speech sounds, and it analyzes how physical forces act on the articulators to understand how and where sounds are produced in the case of phonetics.

Articulatory phonetics also studies the functions of the organs that constitute the speech apparatus and how the speech itself takes shape. Thus, the object of study can include three systems, which are the respiratory, articulatory, and phonatory systems.

It is essential to understand that the organs that work together to achieve speech production are responsible for vital activities such as breathing and eating and consist of a subsystem that underwent an adaptive evolution (MATEUS et al., 2005).

Articulatory analyses allow us to observe how the organs (articulators) behave during the production of an acoustic signal. Therefore, it is possible to confirm the behavior of a given articulator within the phonatory system, as well as the speaker's anatomical and physiological variations.

The nature of each phoneme refers to the specific characteristics and properties of a particular phoneme, such as its place and manner of articulation, voicing, length, and any other distinctive features.

By analyzing the degree of constriction, the point of articulation, and any secondary articulations involved in producing a given phoneme, articulatory phonetics provides a detailed understanding of the physical mechanisms involved in speech production. Therefore, the general approach of articulatory phonetics is to use concrete language analysis as a methodology (CAVALIERE, 2005).

Therefore, we aim to link phonetics and phonology as complementary areas. It is important to note that while phonetics and phonology are complementary areas. They are not interchangeable. Phonetics deals with the physical properties and production of speech sounds, while phonology deals with the abstract system of sound patterns and rules governing their use in language.

Among the post-structuralist theories, gestural phonology was proposed by Browman & Goldstein (1988; 1989a; 1989b; 1990; 1991; 1992; 1995) as articulatory phonology. This theory is considered the most suitable theoretical framework for our purposes since it bridges the gap between phonetics and phonology.

Gestural phonology emphasizes the relationship between physical movements involved in producing speech sounds and the abstract system of sound patterns and rules in language. Therefore, it can explain and analyze phenomena related to speech production and sound change, considering both categorical and gradient features of speech production.

We believe that gestural phonology allows for a more detailed analysis of the physical properties of nasal diphthongization and its relationship to abstract sound patterns in language. This is crucial for understanding the complexities of its production in Brazilian Portuguese.

### **1.3.2 Articulatory Phonology**

Articulatory phonology was initially implemented and developed by Browman & Goldstein (1992). These authors aimed to implement a theoretical model that covered the analysis of the articulatory information of speech production from a dynamic space-time perspective. They propose a framework in which speech production is described in terms of articulatory gestures, which are dynamic and continuous movements of the speech articulators (e.g., tongue, lips, jaw).

In Articulatory Phonology, these articulatory gestures are not seen as symbolic representations, but rather as physical movements that directly relate to the phonological processes of a language's grammar.

The gestures are thought to be coordinated in time and space to produce the specific sounds of a language. This approach emphasizes the connection between the articulatory properties of speech production and the phonological patterns and structures of a language.

Consequently, gesture information and overlapping<sup>10</sup> were not seen as binary features, but as articulatory relationships that maintain a direct and dynamic relationship with the phonological processes of a language's grammar.

Gestures in the context of Articulatory Phonology are not coarticulatory events that occur during speech production. Instead, they are abstract representations that capture the intended articulatory movements associated with speech sounds. In Articulatory Phonology, gestures are defined as the discrete, invariant units of articulation that are organized in time to produce speech.

Each gesture represents a specific articulatory target or action, such as the movement of the tongue, lips, or jaw. These gestures are not directly tied to the physical events of coarticulation that occur during speech production but rather provide an abstract representation of the intended articulatory movements.

This means that a set of articulators must develop a pre-planned trajectory to reach the target, which is referred to as task dynamics. These tasks are dynamic equations that characterize the tract variables as the rest position, stiffness, and damping (BROWMAN & GOLDSTEIN, 1992, p. 23; HAWKINS, 1992).

Importantly, the realization of gestures can vary depending on contextual factors such as the starting position and the overlap with other gestures. This variation in realization is not captured by the gestures themselves but is accounted for by the dynamics of their coordination and interaction (SALTZMAN, 1986; SALTZMAN & KELSO, 1987; SALTZMAN & MUNHALL, 1989).

---

<sup>10</sup> The term "gesture information" refers to the specific articulatory movements or configurations that are involved in producing individual speech sounds. The term "overlapping" refers to the fact that speech sounds often involve the simultaneous coordination of multiple articulatory gestures.



The movement of articulators in speech production is dynamic and occurs in coordinated groups. The concept of gestural constellations goes beyond the individual articulators' motion and encompasses organized and structured tasks.

This means that gestures are not only associated with specific articulatory actions but also involve the coordination of multiple articulators. As a result, gestures can be considered as tract variables that contribute to the overall production of speech. These ideas have been supported by research conducted by Fowler, Rubin, Remez, & Turvey (1980) and Turvey (1977).

In Articulatory Phonology, the variables involved in speech production are the constrictions or releases that occur in the vocal tract. These variables are controlled by the actions of the articulators, which contribute to the process of producing speech sounds. There is a systematic regulation of these variables and their values during speech production.

The parameter values have specific settings or configurations that determine how the articulators move and interact to achieve a target constriction or release. These parameter values specify the dynamic equation that governs the movement of the articulators and their coordination during the production of speech sounds. Browman & Goldstein (1992, p. 24) provided an example of tract variables and associated articulators. The tract variables are:

- (a) Lip protrusion;
- (b) lip aperture;
- (c) tongue tip constriction location;
- (d) tongue tip constriction degree;
- (e) tongue body constriction location;
- (f) tongue body constriction degree;
- (g) velic aperture; and
- (h) glottal aperture.

The articulators are:

- (a) Upper, lower lips, and jaw;
- (b) upper, lower lips, and jaw;
- (c) tongue tip, tongue body, and jaw;
- (d) tongue tip, tongue body, and jaw;
- (e) tongue body and jaw;
- (f) tongue body and jaw;
- (g) velum; and
- (h) glottis.

The example provided demonstrates how different articulatory actions or gestures contribute to the production of speech sounds. Each tract variable represents a specific aspect of the vocal tract configuration, while the associated articulators are the physical structures involved in achieving those configurations.

For instance, to perform lip aperture, the upper and lower lips, as well as the jaw, need to coordinate their movements. Similarly, other articulatory gestures involve the coordinated actions of different articulators, such as the tongue tip, tongue body, jaw, velum, and glottis (BROWMAN & GOLDSTEIN, 1992).

In Articulatory Phonology, the articulatory gesture is considered the smallest unit of analysis, capturing the coordinated actions of multiple articulators. The theory emphasizes the dynamic nature of speech production and explores the processes of gestural transformation, including temporal displacement and coarticulatory effects (ALBANO, 2001).

Gestures are abstract representations that encompass timing, duration, and coordination, and their realization can vary depending on contextual factors such as starting positions, overlap with other gestures, and the influence of neighboring sounds. The integration of gestural representations and phonetic planning allows for a comprehensive understanding of the relationship between phonetics and phonology.

Gestural analysis in Articulatory Phonology provides insights into the spatiotemporal characteristics of speech production, including the coordination, timing, and dynamic properties of articulatory movements. By examining the gestural organization and dynamics, we gain a deeper understanding of how articulatory gestures contribute to the production and perception of speech sounds, as well as their influence on coarticulation effects.

The dynamic changes that occur in articulatory gestures involve modifications in the temporal organization and coordination of articulatory movements. While the abstract gestures themselves remain constant, their realization in terms of actual articulatory trajectories can vary depending on the specific context of speech production. This includes factors such as changes in starting positions, overlap with other gestures, and the influence of neighboring sounds.

Gestural overlap occurs because gestures involve physical events with their internal duration and trajectory, allowing different gestures to co-occur in space and time. This means that different gestures can occur in space and time during the production of a specific sound. The internal duration and trajectory of each gesture contribute to their coordination and can result in overlapping movements. These overlapping gestures can affect each other, and the resulting articulatory movements may exhibit gradient edges from the speaker to the gestural overlap.

The recognition of gestural constellations and overlap aids in analyzing allophonic processes, which involve variations in sound realization based on phonetic context. For example, the pronunciation of a phoneme may be influenced by neighboring sounds, resulting in allophonic variations (BROWMAN & GOLDSTEIN, 1992).

However, it's important to note that the concept of gestural constellations and overlap does not negate the role of phonetic planning. Phonetic planning, which involves the specification of articulatory movements and coordination, is still an integral part of the speech production process. The integration of gestural representations and phonetic planning allows for a more comprehensive

understanding of between phonetics and phonology: phonetic detail, phonological patterns, and the spatiotemporal aspects of speech production.

In this framework, the representation of gestures and their dynamic interactions is seen as more suitable for capturing the fine-grained details of allophonic processes and the coarticulatory effects observed in speech production. It provides a way to account for the context-dependent variations in articulatory gestures and their influence on the acoustic output.

Rather than viewing gestures as static entities, Articulatory Phonology recognizes that their realization is influenced by coarticulatory processes and temporal adjustments (BROWMAN & GOLDSTEIN, 1992). The precise articulatory movements needed to achieve a particular gestural target can be shaped and transformed by these contextual factors.

During speech production, discrete elements, represented as articulatory gestures, are activated with specific temporal and spatial values (Albano, 2001). These gestures encompass a dynamic task with a predetermined trajectory, indicating that the movement of articulators follows a pre-planned course to achieve the desired constriction or release of the vocal tract.

This pre-determined trajectory refers to the predefined patterns of movement that govern the coordination and sequencing of articulatory actions (BROWMAN & GOLDSTEIN, 1992). As a result, gestural overlap can occur, where different gestures co-occur in space and time, potentially influencing the overall tract geometry.

These variations can be observed in factors such as timing, amplitude, or duration of gestures, representing the smooth transitions and nuanced changes in the articulatory process (ALBANO, 2001). Articulatory Phonology offers a framework for understanding phonetic detail and phonological patterns (BROWMAN & GOLDSTEIN, 1992). It allows for the investigation of how articulatory gestures contribute to the production and perception of speech sounds, as well as their influence on coarticulation effects.

Within the framework of Articulatory Phonology, gestures, and their edges also account for the diverse realizations of sounds within a language. Allophonic variations arise due to contextual influences, where the production of a sound is shaped by neighboring sounds, speaking rate, or other phonetic factors (BROWMAN & GOLDSTEIN, 1992).

Gestural edges serve as the demarcations or transitions between gestures, marking the boundaries within the articulatory sequence. Furthermore, gestures and their edges exhibit gradience, encompassing continuous and gradual changes in articulatory movements (BROWMAN & GOLDSTEIN, 1992).

These variations demonstrate the flexibility and context-dependent nature of speech production. This framework allows for a comprehensive understanding of the relationship between phonetic detail, phonological patterns, and the various factors that influence speech production, including allophonic and physiological variations. Additionally, physiological variations are observed in the articulatory patterns of individuals, stemming from anatomical or physiological differences. These variations reflect the unique characteristics of each speaker's articulatory system and contribute to the rich diversity of speech production patterns (BROWMAN & GOLDSTEIN, 1992).

In Articulatory Phonology, gestures are considered dynamic, gradient, cognitive, and physical entities. They capture the spatial-temporal variations involved in speech production. This perspective diverges from the traditional view of phonemes as static, linear, and binary units. Rather than assuming a fixed relationship between phonemes and gestures, Articulatory Phonology allows for flexibility and freedom in the mapping between them.

In our analysis, we used gestures as dynamic articulatory units with characteristic physics that realize and organize the characteristics of the language. Turk & Shattuck-Hufnagel's (2020) discussion sheds light on the reasons behind the absence of symbolic representations in Articulatory Phonology.

This perspective challenges the traditional notion of phonemes as static, discrete units with fixed symbolic representations. Instead, Articulatory Phonology focuses on the physical properties and dynamics of articulatory gestures in speech production.

One of the key arguments put forth by Turk & Shattuck-Hufnagel (2020) is that the relationship between phonemes and gestures is not arbitrary or symbolic. They emphasize that sounds belonging to the same phonological category can be realized by different gestures, which contradicts the notion of one-to-one symbolic mapping.

For example, in languages like Dutch, the uvular and alveolar /r/ sounds are produced using different articulatory gestures. Similarly, in British English, there are variations in glottal stop variants and /t/ phonemes with tongue raising, each associated with different gestures. These examples demonstrate the non-arbitrary mapping between phonological categories and gestures in Articulatory Phonology.

Another reason discussed by Turk & Shattuck-Hufnagel (2020) is the dynamic and gradient nature of gestures. They argue that gestures exhibit continuous variations in their spatial-temporal characteristics, making them unsuitable for rigid symbolic representations. Unlike symbolic representations, which imply discrete and categorical distinctions, gestures capture the gradual and nuanced changes involved in speech production. This dynamic and gradient nature of gestures aligns with the empirical findings from articulatory studies that highlight the variability and flexibility of articulatory movements.

Furthermore, Turk & Shattuck-Hufnagel (2020) emphasize the importance of considering the physical and physiological aspects of articulatory gestures. They argue that gestures are grounded in the anatomical and physiological properties of the vocal tract and its articulators. By focusing on the physical nature of gestures, Articulatory Phonology provides a more comprehensive understanding of the coordination and execution of articulatory movements in speech production.

Overall, Turk & Shattuck-Hufnagel's (2020) discussion highlights the limitations of symbolic representations in capturing the complex and dynamic nature of

articulatory gestures. Articulatory Phonology offers an alternative framework that embraces the physical, dynamic, and gradient properties of gestures, allowing for a more nuanced understanding of speech production

In Articulatory Phonology, gestures do not have a default status that represents the onset of a trajectory. Unlike traditional views, Articulatory Phonology does not specify the starting position of a gesture in its representation. This lack of specification allows for different movement trajectories to be generated by the same gesture. A single gesture can give rise to a variety of movement patterns, depending on its context and coarticulatory influences.

The displacement of a gesture occurs along its trajectory, indicating the movement of articulators away from their resting or default positions. However, it is crucial to note that gestures do not have a predefined target position or a specific offset where they return to a default position. The offset of a gesture is not fixed but rather influenced by the surrounding context and the subsequent gestures in the speech sequence.

By allowing flexibility in the starting position and the lack of a predetermined target position, Articulatory Phonology captures the dynamic and context-dependent nature of gestural movements. This approach enables the modeling of a wide range of speech patterns and the exploration of how gestures interact and influence one another within the production of speech sounds.

In the context of Articulatory Phonology, time is indeed considered an essential parameter for representing the articulatory chain of speech production. The timing and temporal organization of articulatory gestures are crucial factors in shaping the production of speech sounds.

The overlapping of gestures during the speech, as observed in coarticulation, is viewed as a mechanical and adaptive representation of information. Coarticulation refers to the phenomenon where the articulatory movements of different gestures overlap or influence each other in the production of speech sounds. This overlapping

allows for efficient and fluent production by anticipating and preparing for upcoming articulatory configurations.

The representation of overlapping gestures and the incorporation of temporal variables in the analysis of nasal diphthongs within Articulatory Phonology challenge the traditional binary model proposed by Chomsky & Halle in 1968. The binary model postulates a discrete and sequential representation of speech sounds, where each phoneme is represented as a combination of distinctive features. In this model, there is no explicit consideration of temporal or coarticulatory effects.

By incorporating time as a parameter and acknowledging the overlapping nature of gestures, Articulatory Phonology diverges from the binary model. It recognizes that speech production involves a continuous and dynamic interplay of articulatory movements, with gestures influencing and shaping each other in a time-dependent manner. This departure from the binary model allows for a more detailed and accurate account of the complexities involved in the production of nasal diphthongs and other speech sounds. According to Generative Phonology (CHOMSKY & HALLE, 1968) perspective, the timing mechanisms are not related to phonology. This assumption is because physical manifestations of speech, such as utterances, are not in the symbolic representation.

Thus, a phoneme is binary and can only have binary features as [- nasal] or [+ nasal] traces. Therefore, the units only become linear sequences of a self-segment, and the environmental factors (the changes in the vocal tract geometry) are not considered. Consequently, we lose the possibility of understanding the sound segments from the perspective of the gestural coordination effects.

When the overlapping of dynamic units surpasses the notion of linear and static unity, a conceptual change occurs. Later models consider the articulatory movements and phonetic variability of a given surface, but they do not provide a complete description of the time component. (KEATING 1990; GUENTHER 1995; GUENTHER, GHOSH, & TOURVILLE 2006; GUENTHER 2016; FUJIMURA 1992, 2000 ET



AL). Thus, auto-segments are no longer static and are defined as dynamic gestures that change state over time (GAFOS, 2002).

In the planning component, task requirements are specified, and acoustic cues and surface phonetic variation contribute to the process of changing perspectives. The speech context, including dialect, prosody, stylistic factors, and speech rate, can compress or expand the articulatory gestures, changing their relative time duration (TILSEN, 2014). Due to the influence of the speech context, the nature of articulatory gestures becomes dynamic and discrete. This means that the gestures exhibit continuous variation and can be finely graded along a continuum. They also acquire intrinsic temporal control, meaning that the timing and duration of gestures are influenced by inherent factors within the speech production system.

These components are an essential part of the speech's characteristics and Articulatory Phonology focuses on the physical and dynamic aspects of speech production, without employing symbolic representations as found in other linguistic frameworks.

The formation of the gestural constellation attributes specific characteristics to a sound and generates their overlap. Therefore, it creates a hierarchy of gestures, in which one has dominant status. At the same time, they create a contrasting element. Thus, the hierarchically weaker gesture can suffer overlap, reflected as a specific perception of a sound attribute.

Gestures are spatiotemporal units that represent constrictions occurring somewhere in the vocal tract. In the context of Articulatory Phonology, a gesture can be understood as a spatiotemporal unit that represents a constriction or a specific articulatory movement within the vocal tract. In Articulatory Phonology, the C-center, release, onset, target, and offset are terms used to describe the internal temporal structure of gestures (BROWMAN & GOLDSTEIN, 1992).

- TC-center: This term refers to the temporal midpoint of the constriction within a gesture. It represents the point at which the constriction reaches its maximum degree.

- Release: This term does not denote the distance between the target and the beginning of the end trajectory. Instead, it signifies the phase of the gesture where the articulatory movement transitions from the constriction position (target) to a more neutral or relaxed position. It marks the release of the constriction and the initiation of the subsequent movement.
- Onset: This term represents the beginning of the movement of a gesture. It indicates the starting point of the articulatory action.
- Target: This term refers to the desired or intended constriction that the gesture aims to achieve. It is not an indication representation but an inherent component within the gestural equation of motion. The target influences the trajectory and movement of the gesture, but it is not represented separately as an independent entity.
- Offset: This term indicates the end of the gesture's movement. However, it does not necessarily correspond to the end of the end trajectory. Instead, it represents the point in time when the articulatory gesture returns to a default or resting position.

The gestural framework considers the coordinated actions and movements of articulators, incorporating both spatial and temporal aspects. By understanding these terms within the context of the gestural theory, we can better analyze and describe the dynamics of speech production from the nasalization aspects.

In linear phonology, nasalization is typically represented as a binary feature, where a segment is either classified as [+nasal] or [-nasal]. It is considered a distinct phonological feature that can be present or absent in a segment. Nasalization is often treated as a phonemic contrast, meaning that it can differentiate words and convey meaning in certain languages. Thus, the form of linear representation is not as relevant to the notion of temporality in phonology (GOLDSMITH, 1976).

On the other hand, in Articulatory Phonology, nasalization is analyzed in terms of articulatory gestures and their temporal properties. Articulatory Phonology views nasalization as a dynamic phenomenon that involves coordinated actions of the

articulators in producing nasal sounds. It emphasizes the role of gestures and their temporal characteristics in speech production.

Articulatory Phonology proposes that nasalization is realized through specific gestural actions, such as lowering the velum to allow airflow through the nasal cavity while maintaining an oral constriction. The coordination of these gestures and their temporal sequencing contribute to the production of nasal sounds.

Unlike linear phonology, which treats nasalization as a discrete feature, Articulatory Phonology focuses on the continuous and dynamic nature of nasalization. It considers nasalization as a result of the coordination and timing of specific articulatory gestures rather than a binary distinction. This perspective allows for a more detailed analysis of the temporal and spatial aspects of nasalization and its interaction with other articulatory gestures in speech production.

The gestures are defined by spatiotemporal parameters that are intrinsic to each gesture, not separable temporal and spatial parameters. The parameters pertain to the specific characteristics of the gesture itself, such as its trajectory, timing, and coordination with other gestures.

In Articulatory Phonology, gestures are connected to consecutive segments and are co-produced by the articulatory system. This means that gestures are not independent entities but rather part of a coordinated sequence of movements that contribute to the production of speech sounds. The articulatory system works in synergy to produce these gestures, emphasizing the interconnectedness and interdependence of the articulators.

However, it is worth noting that aerodynamic modulation, although not currently incorporated in gestural theory, is a parameter that should be considered. Aerodynamic factors, such as airflow and air pressure, can influence the execution and characteristics of articulatory gestures. Incorporating aerodynamic modulation into gestural theory would provide a more comprehensive understanding of the mechanisms involved in speech production.

We used airflow as a means of understanding the behavior of the displacement of the velum and tongue during the production of nasal diphthongs. The nasal airflow will support inferences about the beginning and the end of the gesture of the movement of the soft palate. The maximum volume of the nasality rate will be related to the target of the velar constriction.

Unlike traditional approaches that focus on discrete segments, Articulatory Phonology emphasizes the continuous coordination and co-production of articulatory gestures. By considering the dynamic nature of gestures and their coordination within a sequence, Articulatory Phonology provides a more nuanced and detailed account of speech production processes. This perspective highlights the importance of capturing the interconnectedness and temporal aspects of articulatory gestures for a comprehensive understanding of speech sound production.

### **1.3.3 Timing, Control Hypothesis, and Implementation**

Coarticulation is the phenomenon where speech sounds influence each other in production, which is a fundamental aspect of human speech. It involves the coordination of articulatory movements and control mechanisms that shape the production of speech sounds.

Extrinsic timing theory posits that coarticulatory effects arise from the temporal coordination of articulatory movements (FOWLER, 1980; KELSO, TULLER, & HARRIS, 1983). According to this theory, speech production is influenced by the surrounding context, leading to anticipatory or carryover effects on articulatory gestures. For instance, the articulatory positioning for a consonant may begin before the preceding vowel sound is completed, ensuring smooth transitions between sounds and facilitating faster speech production.

Articulatory control theory emphasizes the role of biomechanical constraints in coarticulation (BROWMAN & GOLDSTEIN, 1986; SALTZMAN & MUNHALL, 1989). This theory suggests that speech production is governed by a system of coupled articulators, where the movements of one articulator affect the movements of others.

Coordination between articulators is achieved through control parameters that regulate the timing and coordination of movement patterns. Coarticulation is viewed as a result of optimizing articulatory movements to minimize effort and maximize efficiency.

The control hypothesis proposed by Kingston & Diehl (1994) suggests that speakers create a phonetic behavior model to identify implicit acts as adaptive strategies. The model emphasizes the speaker's control over articulations during speech production, which influences the accurate perception of sounds by listeners. It highlights the role of production in shaping perception (KINGSTON & DIEHL, 1994). According to Kingston & Diehl (1994), speakers possess extensive phonetic knowledge, and their phonetic behavior is not restricted by fixed parameters. Each speech act is controlled through a mechanism grounded in this phonetic knowledge (KINGSTON & DIEHL, 1994).

Phonetic behavior refers to the observable actions and patterns exhibited by speakers during speech production, including aspects of articulation, timing, and acoustic properties. It is influenced by the speaker's phonetic knowledge, which encompasses the understanding of phonetic properties and processes involved in producing speech sounds (KINGSTON & DIEHL, 1994).

The control mechanism described by Kingston & Diehl (1994) is an active cognitive mechanism that occurs subconsciously, allowing speakers to improve their phonetic behavior by minimizing articulatory effort and increasing auditory distinction (KINGSTON & DIEHL, 1994).

Phonetic implementation in speech involves various control mechanisms dictating how speech sounds are produced. Different perspectives on phonetic implementation have been proposed, including literal and inflexible phonetics, flexible but automatic phonetics, and controlled phonetics.

Literal and inflexible phonetics perceive speech production as a direct and fixed mapping from linguistic representations to articulatory movements. Flexible but

automatic phonetics acknowledges variability in speech production but attributes it to automatic and subconscious processes.

Controlled phonetics proposes that speakers have significant control over articulatory movements, allowing for variability and optimization based on contextual factors. Articulatory co-variation, involving coordinated articulatory movements, is a control strategy used by speakers. By integrating multiple acoustic properties such as pitch, intensity, duration, and spectral characteristics, speakers create a reinforced auditory impression, enhancing the perception of speech sounds (KINGSTON & DIEHL, 1994).

An example of integrating several acoustic properties can be observed in the production of vowel sounds. The speaker adjusts the positions of articulators, such as the tongue and lips, to create specific formant patterns. This coordination of articulatory movements, combined with the acoustic properties associated with vowel production, strengthens the resulting auditory impression (KINGSTON & DIEHL, 1994).

The traditional model in which phonological representations automatically implement phonetics assumes that underlying phonological representations determine the phonetic realization of speech sounds without significant involvement from the speaker's control.

In contrast, Kingston & Diehl (1994) propose a model where speech sound realizations are activated and controlled by the speaker, relying on their phonetic knowledge. The speaker actively adjusts articulations based on this knowledge, resulting in specific realizations of speech sounds (KINGSTON & DIEHL, 1994).

The control hypothesis suggests that speakers develop a phonetic behavior model, guided by phonetic knowledge, to adaptively control their speech production. The speaker's control of articulatory movements influences perception, and integrating acoustic properties enhances auditory impressions. This model challenges the view of automatic phonetic implementation from phonological representations, emphasizing the speaker's active role in shaping speech sound realizations.

These principles are comparable to the proposal by Lindblom (1983, 1990). Consequently, the control mechanism is defined as an active cognitive mechanism, although it occurs subconsciously.

The model proposed by Kingston & Diehl (1994) acknowledges a broader range of possible variations in the production of speech sounds, allowing for more significant variability within the system that defines phonetic contrasts. This means that speakers have more flexibility and freedom in how they articulate speech sounds, leading to individual differences and linguistic diversity. However, there are limits to this variability.

These limits serve as boundaries or constraints on how much variation is considered acceptable within the system of phonetic contrasts. The boundaries ensure that communication remains effective and mutually intelligible between the speaker and the listener.

The relationship between the limits of variability and communication effectiveness lies in finding a balance. While some degree of variability is tolerated and can contribute to linguistic diversity and individual speech characteristics, excessive variation can hinder mutual understanding between the speaker and the listener.

Effective communication relies on a shared understanding of speech sounds, and if the variability exceeds certain limits, it can impede successful communication. The model proposed by Kingston & Diehl allows for greater variability in the production of speech sounds. Balancing flexibility and variation with the need for mutual understanding is crucial for successful communication.

Browman et al. (1992) argue that the separation of the phonetic and phonological systems is considered undesirable from an empirical standpoint. This is because maintaining a clear distinction between these two areas presents challenges in developing a comprehensive theoretical perspective, given their close interconnection in speech production.

The processing or analysis of the auditory signal refers to the way it is perceived by the listener. It suggests the existence of a mechanism that selects and differentiates relevant information from background noise during the perception of speech sounds. On the other hand, Ohala's proposal focused on a specific perceptual process of speech that emphasized phonetic changes, which was not the primary focus of Kingston & Diehl's theory.

In our analysis, there is evidence to challenge the notion that speech sounds are solely determined by their phonological specifications or constraints imposed by the speech production or perception apparatus. Instead, we argue that the implementation of phonological contrasts requires a complex intermediate device that controls sets of articulations.

By employing this control mechanism, the implementation of phonological distinctions is achieved, leading to the emergence of distinct phonetic categories rather than continuous variations along phonetic dimensions.

This perspective challenges the view that speech production is a straightforward mapping from phonological representations to articulatory gestures or acoustic properties. Instead, it emphasizes the need for an intermediate control mechanism that shapes the realization of speech sounds, ensuring the emergence of distinct phonetic categories.

The significance of controlling in the implementation of phonological contrasts. These devices play a crucial role in regulating articulatory gestures, limiting energy expenditures, and generating distinct phonetic categories. By challenging the notion of a direct mapping between phonological specifications and speech realization, we shed light on the intricate and complex nature of speech production.

Multiple independent correlates and their sub-properties play a crucial role in the perception and production of phonological contrasts. They collectively enhance the auditory distinctiveness of phonological distinctions and contribute to the phonetic realization of these contrasts.



Therefore, the realization of speech sounds is not solely determined by their phonological features or the constraints imposed by the speech production and perception systems. Instead, it proposes the existence of an intermediate device that controls articulatory behaviors to optimize the recognition of phonological contrasts. This control mechanism may involve limiting energy expenditures or creating mutually enhancing acoustic effects.

Speakers possess knowledge of the perceptual mechanisms employed by listeners in recognizing speech sounds. This knowledge guides speakers to reorganize their articulatory behaviors to take advantage of these mechanisms. This implies that speakers actively adapt their speech production based on their understanding of how listeners perceive and distinguish sounds.

The result of such reorganizations is the generation of distinct phonetic categories rather than continuous variations along phonetic dimensions. This suggests that the control mechanism plays a crucial role in shaping the phonetic realization of phonological contrasts.

The coarticulation in speech production is influenced by timing, control, and implementation mechanisms. The coordination of articulatory movements and the influence of the surrounding phonetic environment shape the production of speech sounds. The theories of extrinsic timing and articulatory control provide valuable insights into the mechanisms underlying coarticulation.

The control mechanism proposed by Kingston & Diehl highlights the active role of speakers in shaping their speech production. By integrating multiple acoustic properties and adjusting articulatory movements, speakers optimize the perception of speech sounds. This challenges the view of a direct mapping between phonological representations and speech realization and emphasizes the need for an intermediate control mechanism.

The control mechanism allows for greater variability in the production of speech sounds, but certain boundaries ensure effective communication. Balancing flexibility

and variation with the need for mutual understanding is crucial for successful communication (STEVENS, 1989, 2002, 2005).

These ideas challenge the traditional view that speech sounds are solely determined by their phonological specifications or constraints imposed by the speech production or perception apparatus. Instead, the implementation of phonological contrasts involves complex control mechanisms that shape articulatory behaviors, generate distinct phonetic categories, and optimize the perception of speech sounds.

Kingston & Diehl (1994) posited an automatic mechanism that establishes an early stage of perception, where a filter selects and differentiates information from noise. However, this occurs at a level where the auditory signal undergoes processing. The authors did not specifically address speech perception, as proposed by Ohala (1981, 1986), who focused on phonetic changes.

The speaker does not desire the elimination of noise in the signal because it affects the coarticulation of phonemes at the source, allowing the speaker to reinterpret the information present in the signal. However, in terms of the process, all procedures in the speaker's phonetic knowledge are essential for the phonological language system if each part is related to the detailed understanding of the phonetic aspects of speech.

Kingston and Diehl's theory (1994, p. 419-454) assumes a robust and transparent relationship between symbolic knowledge and perceptual properties. According to this view, phonological features are considered primitive, and the relationships between the features and gestures are indirect and more complex.

The purpose of speech production is sound perception. However, the theory of motor speech perception proposed by Liberman & Mattingly (1985) presents a contrasting view, suggesting that perception focuses on gestures. Both the motor theory and the auditory theory of speech perception strive to achieve a balance between the influences of the two interlocutors involved in the communication process. One interlocutor may make efforts to adapt their representations to restore this balance.

These theories address the challenge of invariance in speech perception by defining discrete linguistic units that capture the unchanging aspects of the acoustic signal. These units can be represented by dashes, allophones, phonemes, and/or syllables.

The motricity theory forms the foundation of articulatory phonology (BROWMAN & GOLDSTEIN, 1992). Gestures are regarded as specific physical units that interact with the articulatory domain. These gestures are composed of abstract units and phonological contrasts, providing a framework for understanding and distinguishing different gestures.

Speakers can regulate acoustic variability to some extent to standardize the speech signal. However, perception does not solely rely on the signal's invariance, as it also considers various acoustic cues and contextual factors. Thus, from an auditory perspective, speakers may produce statements that appear inconsistent when analyzed solely from an acoustic standpoint, but these variations may still be perceived as equivalent when considered from an auditory perspective.

The acoustic signal in speech is a complex representation that captures both spatial and temporal aspects of the constrictions occurring at different points along the vocal tract. Therefore, to obtain a comprehensive phonetic-phonological description, it is crucial to consider the temporal dimension and the coordinated timing of articulatory gestures.

Sound patterns in speech result from the coordinated temporal organization of the different organs involved in the production of the speech apparatus, as well as the spatiotemporal behavior of the articulatory gestures. These gestures encompass the coordinated movements and positioning of the speech organs during the production of the speech signal.

Oral gestures primarily involve movements performed by the articulatory organs, with a specific focus on the tongue. These gestures play a crucial role in the phonetic control and automation of the speech system, which undergoes intensive learning processes.

The stems represent individual control strategies that organize the articulators into distinct phonetic categories. For example, allophones are not mere continuous variations along specific phonetic dimensions but are strategically organized.

Kingston & Diehl (1994) proposed a phonetic analysis model called phonetic automation. While this model may not account for all the variability and contrasts in different languages and contexts, it can provide specifications for the performance of various phonetic universals.

The hypotheses put forth by Pierrehumbert (1990) and Huber et al. (2004) suggest that control occurs outside of phonology. Consequently, phonetic phenomena, which are rooted in cognitive processes, are qualitatively recognized, thereby constituting a discrete system category.

This hypothesis originates from the cognitive sciences; however, there is a lack of experimental tests to characterize the active yet unconscious control mechanism of the speech system.

From a linguistic perspective, control is inferred by analyzing the phonetic variations in achieving contrasts and the co-variation of articulatory movements and their acoustic consequences. When multiple co-variant articulations exist for a phonetic realization or phonological specification, we attempt to determine the specific type of co-variation present.

Automatic production of speech actions can be linked to various phonetic contrasts, such as coarticulation between segments, which depend on the physical relationships between articulators and aerodynamic constraints. Improved production and perception of sounds may be indicative of controlled reorganization of the articulatory system.

Regarding sound perception, speakers can modify phonetic components to anticipate the effects of phonetic contrasts on the production of certain sounds. Additionally, speakers can co-vary articulatory movements based on emphasis to create acoustic outputs that ensure optimal hearing comprehension (DELVAUX, 2003).

Linguistic knowledge of phonological representation is not something possessed by the representations themselves but is rather implicit knowledge in humans. This knowledge is based on a refined understanding that encompasses all the phonetic parameters influencing the production and perception of speech sounds.

Phonetics characterizes each sound segment in the speech chain based on both its articulatory and acoustic characteristics. These articulatory characteristics are associated with linguistic attributes that reflect the acoustic resonances produced by the movements of articulatory gestures.

The hypothesis proposed by Stevens (1989, 2002, 2005) aims to explain why certain articulatory and acoustic dimensions favor specific distinctive features in languages. It seeks to identify an articulatory continuum associated with one or more discontinuous acoustic characteristics, defining an interval along this continuum that corresponds to a relatively stable position in terms of acoustic output.

This proposition suggests that linguistic universals, which are patterns or characteristics found across languages, are not arbitrary but arise from interactions between the articulatory parameters of speech and their acoustic effects.

#### **1.3.4 Outline of BP**

This section presents a brief description of the history of BP. Understanding the historical background of BP can provide insights into its development, influences, and unique characteristics of the language, which may be relevant to our studying. The historical context can help establish a foundation for our further analysis, comparisons, and exploration of nasalization aspects related to BP.

Portuguese is a Western Romance language that originated from Vulgar Latin, which colonists spoke in the third century BCE. (GUIMARAES, 2005). Portuguese originated in the north-western part of the Iberian Peninsula and is an Indo-European language like French, Italian, Spanish, and Catalan.

After the decline of the Roman Empire, the archaic Portuguese language developed into the Galician-Portuguese dialect in the fifth century BC. In 1290,

Galician-Portuguese was adopted as the official language of Portugal (GUIMARAES, 2005). The sound changes in that language were different from those in other Iberian Romance languages.

Until 1500, before the arrival of Pedro Álvares Cabral, the Brazilian territory was inhabited by several ethnicities of Native Americans who were speakers of several of the native languages in the region. Pedro Álvares Cabral's arrival is historically significant because it marked the beginning of Portuguese colonization in Brazil. This event led to the introduction of the Portuguese language to have contact with other languages that already existed in the region, and this had a profound impact on the linguistic and cultural landscape of Brazil.

In 1516, during the Renaissance, modern Portuguese became an official language following the publication of the *Cancioneiro Geral* by Garcia Resende. This language standardization was marked by the first grammar, which was written by Fernão de Oliveira and João de Barros in 1536. Portuguese became the official language of several countries due to the colonization of various areas in the sixteenth century (GUIMARAES, 2005). Portuguese became the official language of Portugal's colonies, including Brazil, Angola, Mozambique, Cape Verde, Guinea-Bissau, São Tomé and Príncipe, and East Timor, among others.

During the colonization of Brazil, Tupinambá was the most spoken language on the Atlantic coast. In 1758, the Marquis of Pombal prohibited the Tupi language in Brazil and this language was replaced by Portuguese.

Tupi-Guarani and Tupinambá are distinct branches of the Tupi language family spoken in Brazil. Tupi-Guarani encompasses a broader group of related languages found across South America, while Tupinambá specifically refers to a variant spoken by the Tupinambá people in the northeastern coastal regions of Brazil.

Tupinambá was primarily spoken by the Tupinambá people in the northeastern coastal regions of Brazil, specifically in the area that is now known as the state of Bahia. The Tupinambá people inhabited a stretch of coastline that extended from present-day Salvador to the south of Bahia.

Tupi-Guarani historically had more speakers and a wider geographic distribution. The Tupinambá people had interactions with Portuguese settlers during colonization, whereas Tupi-Guarani includes various indigenous groups with their cultural contexts. Each language variant within Tupi-Guarani has unique linguistic features. Both Tupi-Guarani and Tupinambá, like many indigenous languages, are endangered today, prompting efforts for preservation.

The settlers incorporated this language, and it was taught to the natives at the missions. This language became known as the General Language or Nheengatu. Due to the exchanges between the indigenous peoples and the settlers, Tupi-Guarani influenced the Portuguese that was spoken in Brazil.

Tupi-Guarani and Tupinambá are both branches of the Tupi language family spoken by indigenous peoples in Brazil. While both languages belong to the same language family, they are distinct languages with their characteristics and variations.

At least 274 dialects and indigenous languages are currently cataloged in Brazil, with Tupi and Macro-Jê being the most prominent language groups. There are still 19 isolated languages that are not like each other that have been grouped as a larger linguistic group. Brazilian indigenous languages are among the most endangered in the world<sup>11</sup>.

Portuguese was declared the official language in Brazil in 1758. Over time, BP absorbed linguistic expressions from Tupi, Yoruba, and Kimbundu (African languages). BP also was enriched by the vocabularies of the European and Asian emigrants. Thus, due to linguistic evolution, European Portuguese and Brazilian Portuguese have come to be considered two linguistic varieties of Portuguese.

The Portuguese language, which is spoken in 20 countries and is the official language of 15 countries, has approximately 356 thousand lexical units. Portuguese is ranked the seventh most spoken language in terms of number of speakers in the

---

<sup>11</sup> The fact that Brazilian indigenous languages are among the most endangered in the world highlights the importance of documenting, preserving, and studying these languages. It also emphasizes the significance of studying the impact of colonization and language contact on the linguistic diversity of Brazil, including the influence of indigenous languages on the development of BP.

world. BP is a natural language that is spoken by over 200 million people, both inside and outside of its territorial borders.

In the Lusophone countries of the world, BP is the variant that is the most widely read, written, spoken, and studied variety of Portuguese in the world due to the geopolitical importance of Brazil.

BP is the official language of Brazil; this country is composed of 26 states, and the state of São Paulo is the most densely populated. Over 42 million people inhabit the state<sup>12</sup>, and over 12 million people inhabit São Paulo, the wealthiest city in Brazil. The figure below shows the Brazilian map; the black dot represents the region of São Paulo city.



Figure 1: Map of Brazil. The portion in grey indicates the São Paulo State and the black dot is the location of São Paulo City.

In the Brazilian territory, 800 thousand people consider themselves to be indigenous, and there are about 60 indigenous group families. At present, 305 indigenous ethnic groups speak 274 different languages.

The national distribution of indigenous languages has been cataloged, and there the four trees have expanded into 18 different linguistic branches, nine linguistics families are composed of 23 natural languages, and from these languages, nine are isolated. Most of these languages have not yet been described, although there are linguistic inventories.

Macro-Tupi is a family tree with about ten linguistic families, 46 natural languages, and 14 dialects; one of these languages is extinct, and two are almost

---

<sup>12</sup>The dialect spoken in the São Paulo State won't be studied in this thesis.



extinct. Macro-Gê is a family tree that has about nine linguistic families and 14 languages, one of which is extinct.

Jê-Kaingang is a family tree with 13 natural languages, 25 dialects, and three extinct languages. Arawak has four linguistic families and 17 languages. The Caribbean is a linguistic branch that consists of 15 languages, and it can be found in the north of Brazil (Ikpeng language) and Xingu Basin (three languages).

The presence of a significant indigenous population in Brazil is relevant to the study of BP due to the potential linguistic and cultural interactions between Portuguese and indigenous languages. Exploring the impact of these interactions on the development of BP can provide insights into language contact phenomena and the formation of linguistic varieties.

The Paulista Geral, also known as the Southern Geral, was a lingua franca and creole language that emerged in the 16th century in the Captaincy of São Vicente, Brazil (NAVARRO, 2013. p. 537). It became the most widely spoken language in the southern region of Brazil during colonial times.

The language originated from the Tupi language spoken by the Tupinambá indigenous people in the São Paulo region. It spread rapidly through the actions of Bandeirantes (colonial explorers), influencing the way Brazilians speak today.

The Paulista Geral language was formed through a process of interethnic unions between Portuguese men and indigenous Tupi women, resulting in a mixed population known as "*caboclos*."

The language spoken by the *caboclos* gradually differentiated from the original Tupi and became known as the Paulista Geral language. It was heavily influenced by the Spanish language and the Guarani language, especially during conflicts between Portuguese Bandeirantes and Spanish Jesuit missions.

During the 19th century, while the Portuguese language became dominant in São Paulo, remnants of the Paulista Geral language could still be heard among the older generation. However, by the beginning of the 20th century, the Paulista Geral language was considered extinct. Today, it is of historical interest and has left a legacy

in Brazilian toponymy, with many current Brazilian place names derived from the language.

In 2014, a new source for studying the Paulista Geral language was identified. A document titled "*Vocabulário Elementar da Língua Geral Brasileira*" (OLIVEIRA, 1936) was discovered, which provides valuable insights into the language. Despite its title mentioning the Brasília language (Old Tupi), the vocabulary authored by José Joaquim Machado de Oliveira is considered one of the sources for the Paulista Geral language.

Overall, the Paulista Geral language represents an important linguistic and cultural heritage in the history of São Paulo and the southern region of Brazil, showcasing the intermixing of indigenous and Portuguese influences during the colonial era (SCHADEN, 1958).

BP and European Portuguese exhibit differences in phonetics, phonology, syntax, and prosody, despite sharing a similar written standard. One distinction is that European Portuguese is characterized by the process of unstressed syllable deletion, making it an accent language.

In contrast, Brazilian Portuguese features an accentual syllabic structure in its prosody. This distinction is influenced by various non-linguistic factors, including speech rate, gender, social stratification, and regional dialects.

The prominence or stress patterns of syllables in BP). It means that in BP, syllables can have different levels of prominence or stress, and this stress placement plays a significant role in the phonetic and prosodic patterns of the language.

On the other hand, "an accent language" refers to European Portuguese, where the process of unstressed syllable deletion occurs. In European Portuguese, unstressed syllables tend to be reduced or even deleted in certain phonetic contexts, which affects the overall rhythm and prosody of the language.

## **2. Outline of Nasalization**

Nasal vowels are often subjects of phonetic and phonological analyses due to their inherent complexity. Nasal sound patterns have been studied since ancient times<sup>13</sup>, with mentions in Panini's work in the 5th century BC (DELVAUX, AMELOT & ROSSATO, 2007).

The complexity stems from several factors, including the involvement of the nasal cavity in their production, which adds resonance to the vocal tract, and the potential for coarticulatory effects with surrounding sounds.

These factors introduce unique challenges in understanding the acoustic and articulatory characteristics of nasal vowels. Researchers in the fields of phonetics and phonology often study nasal vowels to unravel their intricate acoustic properties and investigate their patterns and behaviors within language systems.

Although nasality has been studied widely, some questions about the specific acoustic cues are related to how to distinguish nasal sounds from oral sounds and how these cues vary across different languages or dialects.

Nasalization influences the adjacent sounds during speech production and nasal sounds interact with neighboring vowels and consonants in terms of nasal coarticulation effects.

The concept of the nasality pattern proposed by Ferguson (1975) in his search for linguistic universals in natural languages makes a distinction between the nasal consonants /m/ and /n/. This concept is widely recognized and remains relatively constant across languages. However, the presence and characteristics of other nasal segments in a language's phonological system can vary depending on their frequency and whether they are distinctive.

---

<sup>13</sup> The nasal sounds played a role in experimental phonetics investigating using inventive instrumentation, such as Rousselot's work (1897-1901). In the 20th century, the principles of the acoustics and perception of nasalization were established through studies by Delattre (1954), House & Stevens (1956), Fant (1970), and others. Phonetic studies in the 60s and 70s focused on nasal sound production using various techniques like radiography, electromyography, and aerodynamic data (DELVAUX, AMELOT & ROSSATO, 2007).

Five distinctive segments have nasal features<sup>14</sup> in the world's languages, namely nasal consonants, nasal vowels, nasal continuums (nasalized semivowels), nasal glides<sup>15</sup> and nasal clicks.

The production of nasalization depends on multiple factors; 75% of natural languages contain diphthongs [aj], and 65% of them contain [aw] (LADEFOGED, 2003). However, only 20% of the world's languages have nasal vowels in their phonological inventory, and 60% of them prohibit nasal and oral vowels, according to Maddieson (1984, p. 60):

(...) nasals in the dental/alveolar area are the most common, but bilabials are also very frequent. Although velar nasals are far from rare, they are much less frequent than velar stops (...) On the other hand, palatal nasals are more common than might be expected from comparison with stop frequencies (...) No example of a phonemic uvular nasal is reported from the UPSID (...) Pharyngeal and glottal places of articulation are ruled out for nasals because of articulatory constraints.

Most languages do have one or more nasals. Almost 97% of the languages in the UPSID<sup>16</sup> database sample have nasals. None of the UPSID languages have more than 6 primary nasal consonants, which is the same maximum number of place contrasts for stops Maddieson (1984).

Based on Maddieson (1984), the UPSID database shows that if a segment is nasal, it is voiced in 93.1% of cases. Additionally, palatalized nasals are likely to be bilabial in 50.0% of cases, while labialized nasals tend to be velar in 63.6% of cases.

---

<sup>14</sup> The binary distinction between [+ nasal] and [- nasal] is assumed based on the fundamental articulatory difference between the open and close soft palate (CHOMSKY & HALLE, 1968). Such a description is assumed for all nasal elements, such as vowels and consonants.

<sup>15</sup> Nasal glides constitute a brief nasal realization linked to the other segments. There are three segments that form a boundary with nasal glides, namely prenasalized occlusive, vowels and nasal glides, and consonants with the oral attack.

<sup>16</sup> The Phonological Segment Inventory Database (UPSID) is a comprehensive database of phonological inventories of languages from around the world. It was compiled by Ian Maddieson and Kristin Precoda at the University of California, Berkeley. The UPSID contains information on the phoneme inventories of over 400 languages, representing a wide range of linguistic families and geographical regions.

Voiceless nasals are more likely to have a bilabial place of articulation than any other place, with a conformity rate of 30.6%.

Additionally, languages with nasals almost always possess /n/ (99.0% conformity), and the presence of /m/ implies the presence of /n/ in 99.3% of cases. Moreover, the presence of either /ɲ/ or /ŋ/ implies the presence of both /m/ and /n/ in 98.5% of cases.

The relationship between nasals and other consonants is consistent across different places of articulation, suggesting that the presence of a nasal at a specific place implies the presence of an obstruent in a similar location. Finally, the occurrence of prenasalized obstruent typically indicates the presence of a simple obstruent of the same class and place of articulation in 96.1% of cases.

Ferguson proposed universals about nasals, stating that almost every language has at least one primary nasal consonant (PNC) as a phoneme, characterized by a voiced nasal continuant allophone. In Maddieson's (1984) book, he argues that there are a few languages without PNCs, their number is small, and exceptions have been noted in the literature. For example, in the UPSID database, there are 10 languages without PNCs, with four of them lacking any nasal or nasalized segments.

These exceptions include Rotokas, Quileute, Puget Sound, and Mura. The remaining six languages without nasals have either prenasalized stops or nasalized vowels. Some languages in this group also exhibit nasal consonant phones.

The number of PNCs in languages typically ranges from one to six, with no UPSID languages having more than six PNCs. This number aligns with the maximum number of place contrasts for stops (MADDIESON, 1984, p. 61).

According to Ferguson (1963), if a language has only one primary nasal consonant (PNC), it is likely to be an apical nasal with a dental or alveolar place of articulation, represented by /\*n/. Among the languages listed in the UPSID database with only one nasal, the majority indeed have /\*n/.

However, there are exceptions such as Taoripi with /m/ and Mixtec with /ŋ/. Having only one nasal is considered relatively rare, with American languages showing

a higher prevalence of this pattern. The claim that every language with any nasal has /*n*/ is generally true, with very few exceptions.

Ferguson suggests that if a language has two PNCs, one of them will be /*m*/. This holds for most languages, with the bilabial nasal /*m*/ being present in most languages with two PNCs. There are also languages with additional nasal segments that exhibit differences in phonation type, length, or secondary articulation.

The generalization is extended up to the third PNC, where velar nasals are the most common, although there is a minor pattern with palatal or palato-alveolar nasals. Languages with four PNCs typically have a palatal nasal, while languages with five or six PNCs exhibit more variability in the inventory of nasals.

Ferguson also notes a connection between primary nasal consonants and obstruents, stating that the number of PNCs in a language is never greater than the number of obstruents. The presence of a nasal at a particular place of articulation usually implies the presence of an obstruent at the same place, with some exceptions involving palatal nasals.

Ohala (1975) developed explanations for nasal sound patterns, particularly those with a phonetic basis. The research suggests that the occurrence of nasals in languages can be attributed to their high distinctiveness.

Nasals are rarely confused with other consonant types and are reliably identified as nasals, making them valuable for effective communication. Numerous studies have demonstrated the distinctiveness of nasals as a class (MILLER & NICELY 1955; SINGH & BLACK 1966; SHEPARD 1972; GOLDSTEIN 1977).

Historical phonology also provides evidence that nasals are among the most stable sounds over time (MADDIESON, 1984). However, the presence of multiple nasals in a language is less easily explained based on distinctiveness alone, as nasals within the same class can be prone to confusion. While clear vowel transitions may alleviate this issue, the contrastiveness of nasals in consonant clusters is often limited due to both lack of distinctiveness and articulatory convenience.

The preference for certain places of articulation for nasals, such as dental/alveolar, bilabial, velar, and palatal, may have a perceptual basis. The palatal nasals were excluded from many phonetic and perceptual studies makes it challenging to draw definitive conclusions about their identification. However, some evidence suggests differences in the identification of /m, n, p/:

“Zee (1981) found that /n/ is most likely to be correctly identified, while House (1957) observed that /p/ is more prone to being misidentified as one of the other nasals. Specifically, Zee's research indicates that /p/ is particularly likely to be identified as /n/ following /j/, while Malecot's findings align with House's observation that /m/ is most likely to be identified as /n/ before /i/. Zee also notes that after /l/ or /e/, /m/ tends to be perceived as /n/” (MADDIESON, 1984, p.71).

Perceptual tendencies could be influenced by coarticulation effects and individual language habits. In languages that have voiceless nasals, there is usually a voiced portion before the vowel, which aids in their recognition Ohala (1975). Voiceless nasals are either disposed to loss due to indistinctness or regain their voicing over time. The rarity of voiceless nasals can be attributed to the difficulty in distinguishing them based on their noise spectrum Ohala (1975).

Patterns involving secondary articulations and the relationships between nasals and obstruents are likely connected to broader issues concerning the utilization of phonetic contrasts beyond nasals in isolation.

The nasalization of vowels is a prevalent phenomenon in the UPSID languages, with approximately 22.4% of the sample exhibiting a contrast between oral and nasalized vowels. The occurrence of nasalization is somewhat associated with the number of vowel qualities in a language, although the trend is less prominent compared to vowel length.

For instance, 21.2% of languages with 4-6 vowel qualities have contrastive nasalization, while 53.8% of languages with 10 or more vowel qualities exhibit this feature. However, it should be noted that the presence of nasalized vowels often involves distinct qualities compared to their corresponding oral counterparts. Only a

small percentage of languages with 10 or more vowel qualities include nasalized vowels that have unique qualities.

Linguistic systems with nasal vowels are unstable due to the complexity entailed in articulating and perceiving the nasality acoustically. Such characteristics exceed the existing differences in the oral vowel system.

Vowel nasality plays a crucial role in various languages. While some languages, such as French, Hindi, Lakhota, Navajo, and Brazilian Portuguese, exhibit phonemic nasality (DELVAUX, AMELOT, & ROSSATO, 2012), where nasality distinguishes word meanings, other languages utilize nasality differently. In these languages, nasality coarticulates with vowels, providing additional cues for listeners, even when nasality is not phonemically contrastive.

Vowel nasality is characterized by the lowering of the velum, which opens the velopharyngeal port, allowing air to flow through the nose and nostrils during vowel production. Unlike nasal stops, oral airflow is maintained concurrently with nasal airflow, resulting in additional nasal resonances and zeroes in the acoustic realization of the vowel. Listeners can perceive the difference in sound and infer the degree of velopharyngeal port opening, thus adjusting their perceptions accordingly.

Nasal vowels convey valuable information for contrasting nasality between vowels, reveal the timing and contour of nasality within the vowel, and allow for cross-speaker, cross-condition, and cross-linguistic comparisons.

Beddor's work (2009) highlights the absence of a single spectral criterion for labeling vowel nasalization onset in words. She suggests that the temporal alignment of the nasal and oral gestures plays a crucial role in determining the degree of vowel nasalization (BEDDOR, 2007).

In English, for example, when the nasal consonant was pre-voiced, it was longer, and the vowel nasalization was less extensive compared to when it was pre-voiceless. Similar patterns were observed in Ikalanga and other languages. However, Ikalanga showed a relatively constant-sized nasal gesture across different contexts (BEDDOR, 2007).



She also investigated the influence of vowel length on the relationship between vowel nasalization and nasal consonant duration in Thai and English. It was found that longer vowels were associated with shorter nasal consonant duration and less extensive vowel nasalization (BEDDOR, 2007).

French nasal vowels are phonologically specified to include nasal airflow, and they act as single segments in the language. French listeners interpret the meaning of a word or sentence correctly. Distinguishing between oral vowels and nasal vowels is crucial for understanding the intended message (COHN, 1990). In contrast, in English, nasality is generally not as crucial for word or sentence interpretation.

Demolin et al. (2003) examined the velar gesture and nasal airflow during the production of French nasal vowels. Magnetic Resonance Imaging (MRI) has provided information about the aperture of the velopharyngeal port during vowel production, showing variation in the shape of the port aperture across different nasal vowels. Aerodynamic studies have also revealed variability in the degree and timing of nasality during phonemically nasalized vowels (DELVAUX et al., 2008).

Furthermore, French nasal vowels can cause contextual nasalization of surrounding consonants, particularly in a carryover setting. However, this nasalization is unlikely to serve as a reliable cue for perceiving vowel nasality.

Differences have been observed between the articulation of oral and nasal vowels in terms of velar configuration and oral articulation. Nasal vowels exhibit stronger nasality, often with a plateau or peak in nasality toward the center of the vowel (Delvaux et al. 2002). The exact nature of the time course of nasality varies depending on context, token, and speaker.

Trubetzkoy (1949) introduced the notion that nasality represents a private opposition, while the concept of value was introduced by Jakobson, Fant, & Halle (1952). Nasality is considered a binary acoustic articulatory feature, encompassing various attributes.

Chomsky & Halle (1968) described nasality as a binary feature represented by the pattern [+nasal], which serves a distinctive function in natural languages. Oral sounds are contrasted with nasal sounds through the negative binary value [-nasal].

The nasalization is more than a binary value and encompasses a range of degrees and qualities. It is not simply a matter of a sound being either nasal or non-nasal, but rather a continuum of nasality that can vary in intensity and quality.

Nasalization can exhibit various degrees or levels of nasal airflow and resonance. This variation in nasalization can be observed across different languages, dialects, and even within individual speakers (LADEFOGED, 1975).

Several distinct variable voice qualities (LAVER, 1980, p.70) represent the articulatory geometry in the pharynx, the velum, the tongue movement, and the lips, which is why nasalization is a phenomenon resulting from a gestural constellation.

Nasality is a dynamic and complex phenomenon that involves various interactions among muscles, individual variations, and speech rate to achieve accurate nasalized production. It goes beyond the simple function of the velum and encompasses a range of synchronic processes.

In terms of degrees of nasalization, some vowels or consonants may exhibit a higher degree of nasality than others. For example, certain nasal vowels may have a stronger nasal airflow and resonance compared to nasalized vowels that arise through coarticulation with nasal consonants (CAGLIARI, 1977).

In this regard, nasality can be considered an unmarked resource in certain positions within the phonological system. It is viewed as a gesture or a dependent root node linked to the features of the velum (CLEMENTS, 1985; MCCARTHY, 1988; SAGEY, 1990; HALLE, 1995). Nasality serves as a phonological feature that interacts with other elements in the production of speech, allowing for the assimilation and modification of oral and nasal sounds.

However, nasality has variable characteristics in some languages. Some sounds may be partially nasalized or nasalized to differentiate them from oral sounds. The

gradient of the nasalization highlights the differences in the temporal pattern of nasal gestures (LADEFOGED et al., 1999).

These differences can be observed in the spectral changes. Acoustic aspects are generally found, and it is possible to find some regularities in various natural languages, even if the production's articulatory, aerodynamic, acoustic, and perceptual relationships are complex and variable.

The oral cavity and the nasal cavity must be coupled to produce nasal sounds. The cavities are different filters (resonators) with different dimensions, and the velopharyngeal opening performs the coupling of these two cavities.

The coupling results in a small cross-section, regulated by the movement of the soft palate (velum), which allows the airflow to escape through this bifurcation in the supraglottal cavity.

The nasal tube is smaller than the oral tube but is also a resonating cavity. The opening of the velopharyngeal sphincter is approximately 3.0 cm and is in the middle of the oropharyngeal cavity.

This opening is a sufficient air inlet that allows airflow to resonate to produce nasal resonances in the spectrum of a nasal vowel. Thus, the volume of airflow leaving the source is modified when entering the nasal cavity through the velopharyngeal opening. Figure 6 below shows the velum sphincter in red.

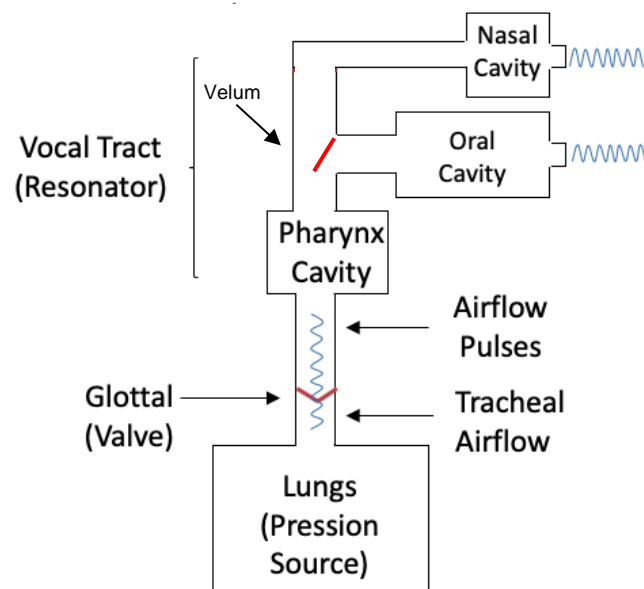


Figure 2: Model of speech production of the human vocal tract with the two tubes in parallel (nasal and oral cavities).

This figure represents the modeling of the vocal tract coupling to produce nasal sounds. The model of speech production with two tubes in parallel, representing the nasal and oral cavities, is a simplified representation of the human vocal tract. In this model, the vocal tract is conceptualized as a system of interconnected tubes through which airflow passes during speech production.

The nasal cavity, located behind the nose, and the oral cavity, including the mouth and throat, are the main components of this model. The two tubes run parallel to each other, allowing for independent control of airflow and sound production in each cavity.

During speech, air from the lungs is expelled and directed into either the nasal or oral cavity or a combination of both, depending on the specific speech sounds being produced. The movements and positioning of various articulatory structures, such as the tongue, lips, and velum, shape the airflow and create specific resonances within each cavity.

The nasal cavity contributes to the production of nasal sounds, where the air is directed through the nose, causing vibration of the nasal tissues and resulting in nasal resonance. The oral cavity, on the other hand, plays a crucial role in producing a wide range of speech sounds, including vowels and consonants, by manipulating the shape and size of the oral cavity and controlling the airflow through the mouth.

During the production of the nasal sound, the soft palate is lowered, which allows the airflow to spread in the nasal cavity, and the resonances can occur in the two cavities. The interaction between them can occur with a minimally significant opening.

An acoustic output is generated when the coupling between the two cavities occurs as a result of the total system configuration and resonances. The different degrees of coupling result in distinct degrees of nasalization; the value depends on the volume of the velum opening's cross-section (CURTIS, 1970).

By modeling the vocal tract with two parallel tubes, this simplified representation helps us understand how different articulatory configurations and airflow control

mechanisms contribute to the production of speech sounds. It provides a framework for studying the interactions between the nasal and oral cavities and their influence on speech production and sound quality. This type of articulation increases the acoustic resonance in the production of nasal sounds and allows both consonants and nasal vowels to be sustained during phonation for an extended period. This increase in resonance is primarily associated with the interaction between the oral and nasal cavities. When producing nasal sounds, both the oral and nasal cavities contribute to the resonance of the sound waves. Velum lowering is a key articulatory characteristic of nasal consonants, which is absent in oral stops.

Two cavities and the nasal sinus reverberate together in one system to produce nasal sounds. In mid-central vowel schwa, in the oral cavity and the pharynx, the oral frequencies resonate together at around F1 500, F2 1500 Hz, and F3 2500 Hz.

The frequencies of [ã] can vary depending on factors such as individual speech patterns and dialectal variations. However, on average, the formants of [ã] tend to fall within the following frequency ranges: Fn1 500 Hz to 900 Hz, Fn2 1200 Hz to 1700 Hz, and Fn3 2400 Hz to 2800 Hz. The anti-resonance locus is about 680 Hz and 2040 Hz.

In the production of nasals, opening the nasal cavity has several acoustic properties and consequences. Before the release of the closure, a nasal murmur occurs, characterized by the nasal radiation of sound energy (FRY, 1958, p. 118).

This nasal murmur exhibits a spectrum dominated by low-frequency prominence, typically around 250 Hz (FRY, 1958). Additionally, the opening of the nasal cavity introduces resonance not only in the pharyngeal and closed oral cavities but also in the nasal cavities themselves. This results in the nasal murmur, which is characterized by longer cavity lengths and lower-frequency resonance (FRY, 1958).

The nasal murmur in the nasals also exhibits antiresonances, which contribute to the damping of broad-band resonances and the absorption of acoustic energy. Constrictions in the closed oral cavity can decrease energy, while the soft walls of the nasal cavities absorb energy, resulting in lower energy levels (FRY, 1958).

The resonances of the nasal cavities typically fall between 800 and 2000 Hz, with the first formant (F1) of nasals being at lower frequencies and the second formant (F2) often aligning with the third formant (F3) of vowels (FRY, 1958, p. 118). The presence of antiresonance affects the peak of F1 and the energy of F2 in nasals preceding or following vowels (FRY, 1958).

This variability occurs because the nasal cavity is a tube that is parallel to the oral cavity. The nasal cavity has a more negligible effect due to the atmosphere because the mouth has a wider opening than the velum. The velum has a small area until the nostrils (around 13 cm).

It is possible to differentiate between oral and nasal sounds via a comparison of the acoustical spectrum. Nasal coupling affects the spectral characteristics of segments, as can be noted by moving the formants, increasing the bandwidth, and smoothing the harmonic energy peaks.

The emergence of a nasal murmur (nasal appendix) that can increase the length of nasal vowels (KRAKOW & HUFFMAN, 1993) also occurs in nasal sounds. The volume of the nasal airflow changes due to the articulators' movement (tongue and lips) and the extent of the velopharyngeal opening.

In the nasal vowel spectrum, the sound waves overlapping between the oral and nasal airflows change the resonances. The poles and zeros in the spectrum sound segment can constitute constructive or destructive interferences. Each additional zero can be paired with an extra pole.

Thus, the combination of the poles and zeros can eliminate or amplify the resonance. If an oral or nasal formant has the same bandwidth and fundamental frequency, this will result in a partial or total cancellation of resonance in some frequency ranges.

However, the spectral effects may vary if the formants have different bandwidths and frequencies, thus making acoustic analysis difficult. The increase in the energy intensity regions may decrease the intensity of the high-frequency harmonics caused by anti-resonances resulting in zeros in the spectrum (DICKSON, 1962).

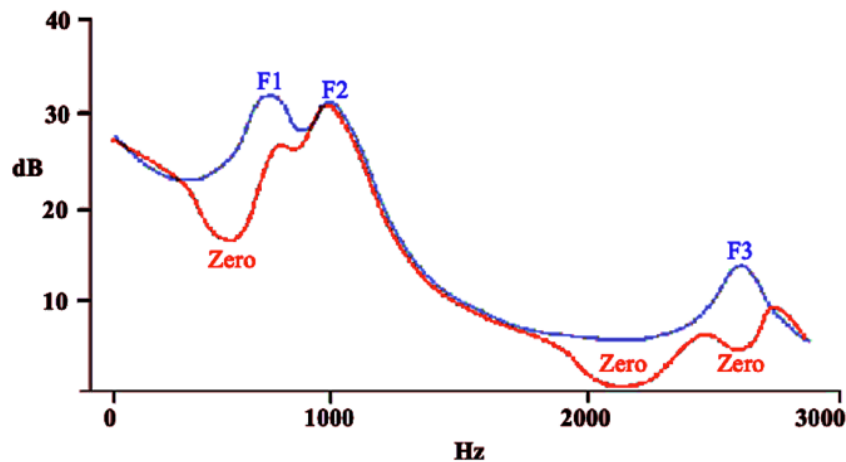


Figure 3: An example of overlapping graphics of poles and zeros from an oral vowel spectrum [a] in the blue line and the nasalized spectrum of [ã] in the red line. Adapted from Pickett (1980) and Styger & Keller (1994, p. 117).

These frequency interferences are generated in the nasal cavity in nasal segment formants, particularly in the first and second formant regions, as shown in figure 7. The nasal formants (anti-formants) are resonances that cause interferences in the oral sound.

Delattre (1954, 1955) observed that the amplitude reduction and widening of the first oral formant (F1) significantly nasalized the oral vowel. He also identified fixed peaks at 250 Hz and 2000 Hz and a variable peak at 900 Hz as the main formants of nasal sounds.

House & Stevens (1956) confirmed Delattre's observations regarding the amplitude reduction of F1 and the presence of a low-frequency peak (around 250 Hz) when the nasal and oral tracts are coupled. They also identified the appearance of a zero between 700 Hz and 1800 Hz due to the coupling.

Maeda (1982a, 1984) introduced the concept of a global flattening of spectra in the first and second formant region (200 to 2000 Hz) as an independent cue for nasalization. He also highlighted the significance of the two peaks (250 Hz and 1000 Hz) in nasal spectra as correlates of nasality.

The distance between these peaks was found to correlate with the magnitude of nasal coupling. Stevens et al. (1987) emphasized the perturbation of the low-frequency spectrum and the replacement of the first formant by a pole-zero-pole

combination as the main effects of nasalization. Hawkins and Stevens (1985) highlighted the importance of a very low-frequency peak, widening bandwidth, and the presence of a pole-zero pair resulting from acoustic coupling in the perception of nasality.

According to Pickett (1991), there is a low (formant) pole frequency that intensifies F0. In the region above F1, there is a zero (anti-resonance or low-frequency formant) above the pole, which decreases the spectrum's energy above the F1 area. If the vowel F1 is too high, it reduces the zero spectral power but, if F1 has a lower frequency, zero can be canceled, thus generating a level spectrum.

Nasal resonances alter the spectrum of the higher frequencies in the region of F2 and F3. There are changes in the spectrum range; additional spectra and zeros are introduced in the region (PICKETT, 1991).

According to Kent & Read (1992), the formants and anti-formants usually occur in pairs, as explained in the previous paragraph. Therefore, nasalization in the vowel segments introduces resonances and anti-extra resonances.

According to Fujimura & Erickson (1997), the nasal segments in the nasal vowel offset position behave similarly to glides and, in many languages, do not have specific features that differentiate them. The nasal appendix is tiny or does not exist in English words. Fujimura & Erickson found evidence of this in English data.

In the nasal appendix spectrum, F1 will have very low frequencies of around 200 Hz. Just above that, there will be a concentration of energy of about 800 Hz, which is a formant damped by a zero (FUJIMURA & ERICKSON, 1997). The amplitude decreases, the peaks are entirely cushioned and flat, and the formant loses energy due to nasal resonance. This resonance range at the end of nasal vowels has a variable acoustic duration depending on the speech context and individual characteristics (SOUSA, 1994).

According to Hawkins & Stevens (1985), the spacing between the pole and the zero expands as the pharyngeal opening increases. The inclusion of additional



resonance and the highest spectral prominence is proportional to the degree of the velum opening. The velopharyngeal opening area enlarges gradually.

There is a change in F1 frequency and an increase in its bandwidth beyond the pole-zero pair, which may add resonance to the vowel spectrum near F1. Such resonance would have a frequency of less than 1500 Hz. Thus, the additional zero pole pairs can fill the above spectral valleys below the F1 frequency.

Lindqvist & Sundberg (1972) indicated that the highest pole was about 450 to 650 Hz. Dang et al. (1994) noted the influence of paranasal sinuses<sup>17</sup> on the insertion of new zero-pole poles in the frequency range of around 400 Hz and 1300 Hz.

The previous discussion regarding the change in F1 frequency and the insertion of additional zero-pole pairs in nasals shows that the frequency and spacing of the zero-pole may vary from one individual to another:

Factors that can influence the pole-zero frequencies in nasals and, consequently, the resonant frequencies (formants) include the individual differences in the shape and size of the vocal tract and nasal cavity can lead to variations in the pole-zero frequencies. These differences can be influenced by factors such as anatomical variations, speaking style, and accent.

Also, the opening and closing of the velopharyngeal port play a crucial role in nasal resonance, creating the connection between the nasal and oral cavities. The position of the velum (soft palate) affects the coupling between the nasal and oral cavities, resulting in different resonant characteristics.

The volume of air inhaled and exhaled during speech can also influence the pole-zero frequencies. The amount of air in the vocal tract affects its shape and size, which, in turn, influences the resonant frequencies.

---

<sup>17</sup> Paranasal sinuses are a group of air-filled spaces located within the bones of the skull surrounding the nasal cavity. These sinuses are named based on their location and include the frontal, ethmoid, sphenoid, and maxillary sinuses. They are lined with a mucous membrane and communicate with the nasal cavity through small openings.

The context in which a nasal occurs within a word, or a sentence can also affect its acoustic properties. Coarticulation with neighboring sounds can lead to subtle changes in the vocal tract shape and subsequently impact the pole-zero frequencies.

The volume of air changes the size of the chest and extends the asymmetry in the nasal pits<sup>18</sup> that divided the nasal septum (BAVEGARD et al., 1993). The breath no longer plays a significant role in acoustic behavior in the high-frequency region at a frequency above 1300 Hz.

Feng & Castelli (1996) introduce a new interpretation for the gesture involved in producing nasal vowels, suggesting that the lowering of the velum creates a nasal target, similar to a nasal consonant, which serves as a reference for nasalization. This target is not necessarily achieved in the case of true nasal vowels but represents a dynamic trend from the oral configuration toward the nasal target.

The results of Feng & Castelli (1996) on pharyngonasal transfer functions exhibit complex spectra with numerous peaks and zeros. However, certain spectral components maintain relatively constant frequencies across measurements.

The idea of relating nasal vowels to consonants is not new and can be observed in phonological processes where nasal vowel systems often originate from the assimilation of oral vowels by adjacent nasal consonants. The introduction of the nasal target simplifies the concept of nasal vowels by considering the pharyngonasal configuration as a single tract with defined resonance frequencies.

### **2.1.1 BP: Phonetics and Phonology**

This section summarizes the phonetical and phonological inventories of BP. Our purpose is to classify articulatory features or characteristics involved in the production of vowels and consonants in this language. Even though BP and European Portuguese are from the same family, some phonemes are articulated differently.

---

<sup>18</sup> Nasal pits are the initial depressions or invaginations on the frontonasal. They are formed as small depressions or grooves in the frontonasal prominence. Subsequently, these nasal pits deepen and contribute to the development of the nasal structures, including the nasal cavities and external nose.

BP is a natural language with an egressive pulmonic airflow; the BP inventory has 33 phonemes: 15 vowels, 16 consonants, and two glides, in addition to the allophonic variants.

In BP, there is a tendency to produce open syllables CV stressed or unstressed, which is a suprasegmental parameter of this language (STREVENS, 1954; CÂMARA JR., 1972; ALBANO, 2001; PONTES, 1972). The IPA chart presents the consonantal phonemes.

Table 1: IPA Chart of BP Consonant Phonemes<sup>19</sup>.

	<i>Bilabial</i>	<i>Labio dental</i>	<i>Dental</i>	<i>Alveolar</i>	<i>Postal veolar</i>	<i>Retroflex</i>	<i>Palatal</i>	<i>Velar</i>	<i>Uvular</i>	<i>Pharyn geal</i>	<i>Glottal</i>
<i>Plosive</i>	p b		t̪ d̪	t d				k g			
<i>Nasal</i>	m			n			ɲ				
<i>Trill</i>				r							
<i>Tap</i>				ɾ							
<i>Fricative</i>		f v	s		ʃ ʒ			x ɣ	χ ʁ	ħ ʕ	h ɦ
<i>Affricate</i>					tʃ dʒ						
<i>Lateral</i>			l								
<i>Fricative</i>											
<i>Approximant</i>						ɻ	j	w			
<i>Lateral</i>				l			ʎ				
<i>Approximant</i>											

We now provide a summary of some of the phonological phenomena that occur in BP. Most archiphoneme realization variability occurs in BP because of the neutralization in the coda position (CÂMARA JR., 1982; CALLOU, LEITE & MORAES, 1998; CRISTÓFARO, 2002; D'ANGELIS, 2002; ALBANO, 2005; HORA, 2006). This phenomenon occurs in the production of fricatives, rhotics, laterals, and nasals, such as /S/, /R/, /L/, and /N/.

These phonemes lose the phonemic contrast when they occur in the final position of a syllable (coda). For examples of phonemic neutralization in the coda position: the phoneme /S/ is often realized as [s], resulting in a neutralization of the contrast between /S/ and /s/. For example, the word “**casas**” (houses) may be pronounced as [ˈkazas] instead of [ˈkazas].

The archiphoneme /R/ is often realized as a tap [ɾ], like the intervocalic /R/ in many dialects. This leads to a neutralization of the contrast between /R/ and other

<sup>19</sup> The symbols used in the transcriptions throughout this thesis are those used in the IPA alphabet. However, when citing the existing literature, the authors' original terminology and phonetic transcriptions have been retained.

possible coda consonants. For example, the word “**amor**” (love) may be pronounced as [‘amor] instead of [‘amoɹ].

The archiphoneme /L/ is often realized as a vocalic or syllabic [w] or [ɹ], resulting in a neutralization of the contrast between /L/ and other possible coda consonants. For example, the word “**sal**” (salt) may be pronounced as [‘saw] instead of [‘sal].

The archiphoneme /N/ is often realized as a nasalization of the preceding vowel without a distinct nasal release. This leads to a neutralization of the contrast between /N/ and other possible coda consonants. For example, the word “**bem**” (well) may be pronounced as [‘bẽ] or diphthongized [‘bẽjm].

The variability and neutralization of these phonemic contrasts in the coda position can occur in different phonetic contexts and inter-speaker variation. This variation may also be influenced by stylistic factors or individual speech patterns.

Some sibilants, such as /s/, /z/, /ʃ/, and /ʒ/, are distinct phonemes, but they can occur as allophones or variant realizations of the same phonemes in complementary distribution, in different phonetic contexts without causing change in meaning. This variation occurs in the Carioca dialect (Rio de Janeiro city) and in some northern states of Brazil, such as “**este**” [‘eʃtʃɪ] (this) or “**isqueiro**” [‘iʃkejɾʊ] (lighter). In other places in Brazil, the same words are pronounced as [‘estʃɪ] or [‘iskejɾʊ]. Thus, the sibilants in voiceless contexts or in the coda position became the postalveolar /ʃ/ and /ʒ/, as in “**turista**” (tourist) [tu‘riʒtɐ] or “**mesmo**” [‘mezɹmʊ] (same) (CRISTÓFARO, 2002).

There are two speech possibilities in BP for the pronunciation of /t/ and /d/; the plosives [t̪] and [d̪] can be pronounced as the allophones [tʃ] and [dʒ]. These affricate phonemes emerge in the context of /i/, /ĩ/, or /e/ in a weak position, resulting in consonant palatalization. Hence, the consonants became palatalized in these contexts.

This palatalization process occurs in various dialects, including those of Rio de Janeiro, São Paulo, Minas Gerais, and Bahia. Some examples of the most common palatalization processes are “**tia**” [‘tʃiɐ] (aunt) and “**dia**” [‘dʒiɐ] (day). However, in the

dialect from the state of Rio Grande do Sul, this palatalization process does not occur, and the pronunciation is as [ˈtiɛ] and [ˈdiɛ] (CRISTÓFARO, 2002).

The rhotic archiphoneme /R/ varies depending on the dialect and on whether the phoneme is velar, uvular, alveolar, pharyngeal, or glottal. The most frequent pronunciations are:

- Voiceless velar fricative [x].
- Voiceless glottal fricative [h].
- Voiceless uvular fricative [χ].
- Vibrant alveolar multiple [r].

The phonemes [r] and /R/ contrast in the intervocalic position as “**caro**” [ˈkaru] (expensive) versus “**carro**” [ˈkaβu], [ˈkahu] or [ˈkaɦu] (car). In the post-nuclear position, the rhotic allophones vary, as in “**carta**” [ˈkahtɐ], [ˈkaχtɐ] or [ˈkaxtɐ] (letter) versus “**carga**” [ˈkaɦgɐ], [ˈkaβgɐ] or [ˈkaɣgɐ] (load) (CRISTÓFARO, 2002).

In BP, the epenthetic vowel emerges in words such as “**fixo**” [ˈfikisu] (fixed), “**ficção**” [fikɪˈsẽw̃] (fiction), and **pneu** [piˈnew] (wheel) and “**látex**” (latex) [ˈlatekis]. The clusters occur in /b, k, d, f, g, p, t, v/, followed by /l/ or /r/, as in the word **flor** [ˈflor] (flower), **fraco** [ˈfraku] (weak) and **livre** [ˈlivrɪ] (free). In most dialects, the consonant /l/ is vocalized as [ʊ, w] at the end of syllables such as **sol** [ˈsɔw] (sun) (HORA, 2006); the only exception is in the southern region (Rio Grande do Sul state), where the sound is produced with velarization. The consonant /ɲ/ can emerge in [j] contexts.

The BP phonetic inventory includes nasal consonants, nasal vowels, nasalized vowels, diphthongs, and nasal triphthongs (CAGLIARI, 1977). There are nasal consonants, /m, n, ɲ/, which are labial, alveolar-dental, and palatal. These consonants in the onset position are fully specified. They have a point and a mode of articulation that is fully articulated. By contrast, the nasal segments will be produced without specification as n-like, or as a homorganic consonant in an offset.

Nasal consonants are a distinct category classified alongside other consonant types such as liquids, sonorants, and obstruents. The nasal feature [+nasal] sets nasal

consonants apart from oral consonants [-nasal]. Acoustically, nasal consonants exhibit formants with relatively constant frequencies. Producing a nasal consonant involves simultaneous articulatory gestures. The figures below show the point of articulation for nasal consonants.

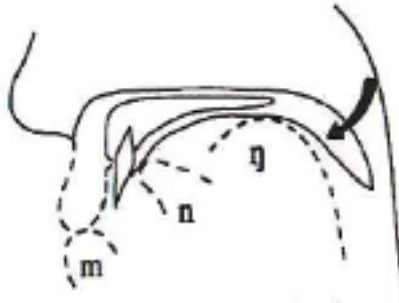


Figure 4: Adaptation of the vocal tract from Kent (1997, p. 64). Articulatory points of nasal consonants.

During the production of nasal consonants, there is a complete closure of the oral tract, which allows for the coupling of the oral and nasal cavities. This mode of articulation creates resonances within the nasal cavity due to the absence of airflow through the mouth. The soft palate is lowered, preventing a plosive release commonly found in oral stops, and instead, the airflow is expelled through the nostrils.

The dynamics of the nasal tract play a significant role in the resonance and radiation of airflow while the larynx produces sound. Nasal consonants expel air through the nostrils due to their mode of articulation. The velum and glottis gestures contribute to the production of the nasal murmur, which is an acoustic characteristic of nasal consonants (SOUSA, 1994).

Repp & Svastikula (1987) discussed the influence of syllabic context on vowel transitions and the articulatory point of nasal consonants. This information is crucial as it helps signal the place of articulation, affecting the perception of nasal consonants through coarticulation effects.

The spectrum of the nasal murmur does not vary significantly across nasal consonants, but it is characterized by low-frequency resonances caused by airflow from the nasal cavity (PICKETT, 1991). Fujimura (1962) studied English nasal consonants /m, n, ŋ/ and found their sound energy distribution within the average

frequency range of 800 to 2,300 Hz. The position of the anti-formant<sup>20</sup> differentiates them: [m] between 750 Hz and 1,250 Hz, [n] at 1,450 and 2,200 Hz, and [ŋ] above 3,000 Hz.

For [m], four formants are present in the main frequency range, while [n] has two formants and an anti-resonance in the F3 frequency, and [ŋ] has two formants and an anti-formant in the F2 frequency (FUJIMURA, 1962).

Nasal consonants tend to nasalize adjacent vowels through coarticulation in each vocal environment, with the oral vowel preceding a nasal consonant being nasalized for approximately 100 ms (PICKETT, 1991).

The nasalization is caused by the slow movement of the palatine velum during articulation, where the velum initiates the closing movement of the nasal cavity after the closure of the oral tract, but with a significant opening allowing the passage of air (Pickett, 1991).

Nasal consonants have distinctive features, including being typically voiced and exhibiting a sudden change in formant structure at the transition points between the nasal and adjacent vowel, known as a “fault transition.”

Additionally, they have varying nasal formants across different speakers, low F1 (N1) frequencies ranging from 250 to 300 Hz, weaker formants compared to vowel sounds, multiple antiformants, limited energy above 3500 Hz, and an antiformant in the 800 to 2000 Hz range corresponding to F2 in most vowels. The locations of formants and antiformants vary depending on the place of articulation and individual speaker variation.

The recognition of nasal consonants relies on a combination of formant, antiformant, and transition information, with transition information playing a significant role. For example, [m] exhibits an antiformant between 750 Hz and 1250 Hz, while N2 is typically found between 1000 Hz and 1300 Hz, resulting in a rising pattern of F2 towards the following vowels. [n] has N2 around 1000 Hz and an antiformant between

---

<sup>20</sup> Nasal zero or anti-formant is a frequency region where the energy of the sound is attenuated or canceled out due to the acoustic properties of the nasal cavity. It appears as a dip or notch in the spectrum of the nasal murmur.

1450 Hz and 2200 Hz, resembling alveolar stops. [ŋ] shows a primary antiresonance above 3000 Hz, with N2 around 1100 Hz and N3 at 1900 Hz, although there is variation among speakers (FUJIMURA, 1962).

The distinct acoustics of nasal consonants are characterized by their nasal murmur, formant structure, antiformants, and transitions, which contribute to their perception and recognition.

### **2.1.2 Oral and Nasal Vowels**

Vowels can be classified as oral and nasal, as follows: close, close-mid, and open-mid, and open; front, central, and back; rounded and unrounded (CÂMARA JR., 1970). The vocalic phonemes can be 'monophthongs, diphthongs, or triphthongs' (CAGLIARI, 1977, p. 5), which will be specified further.

In BP, the vocalic inventory in the tonic position has seven oral vowels, which are /a, e, ε, i, o, ɔ, u/. There are five vowels, /a, e, i, o, u/, in the pre-tonic syllable, and three weak vowels /e, ɪ, ʊ/ in the unstressed final position, resulting in neutralization between the opposition of /i/ and /e/ and /o/ and /u/.

In BP, vowel height alone does not directly differentiate stressed and unstressed syllables. Instead, the distinction between stressed and unstressed syllables is primarily marked by the presence or absence of vowel reduction, vowel duration, and changes in vowel quality.

In stressed syllables, the vowels tend to maintain their full quality and duration. On the other hand, in unstressed syllables, vowel reduction occurs, leading to a reduction in vowel quality and often a shorter duration. This reduction process typically involves the centralized pronunciation of vowels, resulting in a more centralized vowel sound compared to its full-quality counterpart.

While vowel height itself does not serve as the sole indicator of stress or unstress, it can play a role in the phonetic realization of unstressed vowels. In certain contexts, high vowels (/i, u/) may undergo more noticeable reduction and centralization compared to mid (/e, o/) or low (/a/) vowels. However, it's important to



note that other factors, such as syllable structure and surrounding phonetic context, also contribute to the realization of stressed and unstressed syllables. In the vocalic inventory of Brazilian Portuguese, the vowel /a/ is analyzed and produced as a back vowel. It is characterized as the lowest and widest vowel in terms of its tongue position and mouth opening. The vowel /a/ is pronounced with the tongue positioned towards the back of the mouth, resulting in a relatively open and wide vocal tract configuration.

In contrast, the vowel /i/ is considered the narrowest in terms of its tongue position and mouth opening. It is a front vowel produced with the tongue positioned towards the front of the mouth, creating a narrow and relatively closed vocal tract configuration.

The vowel /e/ is described as a front mid-high vowel. It is produced with the tongue positioned slightly lower and further back compared to /i/, resulting in a mid-high tongue position and a relatively moderate mouth opening.

Similarly, the vowel /o/ is produced as a back mid-high vowel. The tongue is positioned further back in the mouth compared to /e/, creating a mid-high tongue position, and the mouth opening is moderately wide. By analyzing and describing these vowel characteristics, we can understand the distinctions in tongue position and mouth opening that contribute to the different vowel qualities in Brazilian Portuguese. The vocalic diagrams of the phones and phonemes distribution are presented below:

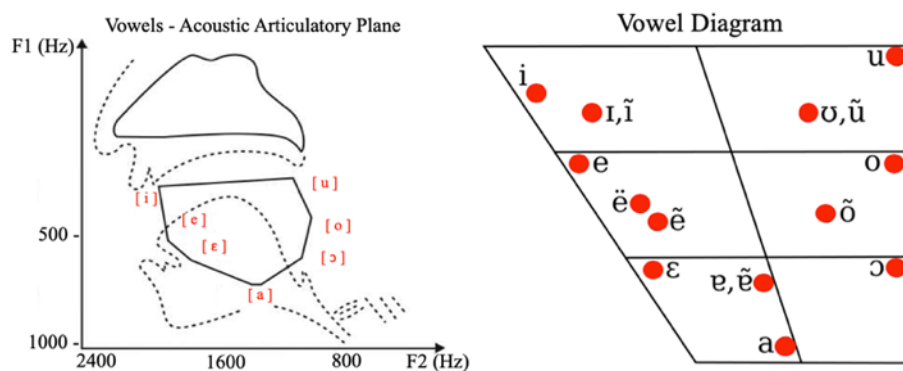


Figure 5: Acoustic articulatory adaptation from BP vowels (RUSSO & BEHLAU, 1993, p. 36) and its vowel diagram (BARBOSA & ALBANO, 2004, p. 229).

BP has five tonic nasal vowels, which are /ĩ/, /ẽ/, /ẽ̃/, /ũ/, /õ/ (CAGLIARI, 2007). All the oral tonic vowels can be nasalized through a coarticulation process with a nasal consonant.

The five nasal vowels are considered distinctive because they form minimal pairs, highlighting their phonological importance. They have distinct phonemic statuses and can create meaningful differences in words. This position was supported by Pontes (1972), Lacerda & Head (1966), and Teixeira, Moutinho, & Coimbra (2001), amongst other authors.

It has been proposed that nasality in a vowel is a binary feature that can result in the nasalized vowel, that shares the [+nasal] feature with the following nasal consonant, as in the example ['kẽ.me] "**cama**" (bed). This coarticulation process differs from phonemic nasal vowels, where the vowel itself undergoes nasalization as part of a nasalization process.

In this case, the nasal consonant can be deleted during derivation, resulting in a fully nasal vowel, as in ['kẽ̃.n.tu] "**canto**" (singing) (CAMARA JR, 1977). Regardless of their phonological status, the feature [+nasal] is absolute, and the vowel can exhibit it or not (Bisol, 1998).

As in the example ['kẽ.me] "**cama**" (bed), productions in BP exhibit nasalization but are not considered distinctive nasal vowels. These nasalized vowel productions are treated as allophones, which means they are variants or alternate forms of the phonemes and do not create minimal pairs.

The specific productions referred to as allophones are the nasalized versions of the oral tonic vowels in Brazilian Portuguese. For example, when an oral vowel such as /a/, /e/, /i/, /o/, or /u/ is followed by a nasal consonant, it undergoes nasalization as a result of coarticulation. These nasalized vowel productions are considered allophonic variations<sup>21</sup> of the respective oral vowels.

---

<sup>21</sup> Allophonic variation in the context of nasalization refers to the phonetic realization of nasalization as a gradient, phonetic coarticulation process. This means that the degree of nasalization can vary.

The nasalization of oral tonic vowels in Brazilian Portuguese is generally considered a phonetic coarticulation process rather than a categorical, phonological process. When an oral vowel is followed by a nasal consonant, there is a natural tendency for the vowel to become nasalized due to the influence and coarticulatory effects of the nasal consonant.

From a production perspective, the degree of nasality can vary depending on the degree of velopharyngeal port opening (CAGLIARI, 1977). Medeiros et al (2011) aerodynamic study indicates that nasal vowels have a higher nasal airflow compared to nasalized vowels. Therefore, nasality should not be treated as a binary phenomenon, as Chomsky's phonological approaches might suggest.

Nasal and nasalized vowels should be treated as a coarticulatory phenomenon because during coarticulation, the articulatory gestures of the vowel and the nasal consonant overlap and influence each other. As a result, the oral vowel can acquire nasal qualities, such as the resonance characteristics and airflow patterns associated with nasal sounds.

This nasalization process is a gradient in nature, meaning that the degree of nasalization can vary along a continuum. It is influenced by various factors such as the phonetic context, speaking rate, and individual speaker variations.

In the context of nasalization in Brazilian Portuguese, it is commonly understood as a phonetic coarticulation process rather than a categorical assimilation process. The nasalization is conditioned by the context, specifically the presence of a homorganic nasal consonant (CAGLIARI, 1970, p. 5).

Therefore, the nasalization of oral tonic vowels in Brazilian Portuguese is characterized by the influence of nasal consonants on the articulation and acoustic properties of the vowels. It is a gradient phenomenon that varies in degree depending on various factors.

The phonological representation of nasal vowels in BP has been a subject of debate among linguists. While nasal vowels in Portuguese, similar to French, exhibit

a distinctive effect, differing from languages like English and Spanish, the nature of this distinctiveness remains elusive.

Two main conceptions regarding the phonological status of nasal vowels have emerged: (i) the biphonemic hypothesis, suggesting that nasal vowels are composed of an oral vowel followed by a nasal consonant (/VN/); and (ii) the monophonemic hypothesis, proposing the existence of nasal vowels as independent phonological units (/Ṽ/).

Additionally, a less common hypothesis, put forth by Parkinson (1983) and Pimenta (2019), proposes a biphonemic nasal vowel composed of an oral vowel and a nasal vowel (/ṼṼ/), resembling a diphthong.

The monophonemic hypothesis suggests that nasal vowels are single phonemes, independent from their oral counterparts. According to this hypothesis, nasality is an inherent feature of these sounds and is not a result of assimilation or coarticulation with neighboring segments.

If nasal sounds are considered monophonemic units, it affects the analysis of phonological processes, such as assimilation and sandhi phenomena, where nasalization plays a crucial role. However, it is important to note that the monophonemic hypothesis is not without controversies.

Barbosa (2009) explores progressive assimilation, where a nasal segment influences a following sound, and regressive assimilation, where a nasal segment is affected by a preceding sound.

Câmara Jr. (1953), Cagliari (1977), Duarte & Teixeira (1979), Moraes & Wetzels (1992), Bisol (2002), Mateus (1975), and Mateus & Andrade (2000), defend the biphonemic hypothesis for the nasality of Portuguese vowels.

According to the biphonemic perspective, nasal vowels do not exist as independent phonological units; rather, they are viewed as oral vowels followed by a nasal consonant within the same syllable. The nasal resonance observed in nasal vowels is attributed to the presence of the following nasal consonant within the same syllable.

Câmara Jr. presents several arguments in favor of the biphonemic hypothesis, mostly based on the structural characteristics of Brazilian Portuguese and comparisons with other Romance languages. One of the primary arguments is related to the syllabic position of nasal vowels.

It is observed that nasal vowels function as closed syllables, similar to nasalized /l/, /z/, and /r/ sounds in Portuguese. Unlike what occurs with nasal vowels, elision often takes place in the presence of a vowel followed by another vowel.

However, this argument is met with criticism, as the absence of elision in certain sequences of nasal vowels can be attributed to factors such as the distinct nature of vocalic segments and stress patterns, rather than the presence of a nasal consonant in the phonological representation.

Another argument put forward by Câmara Jr. concerns the phonetic realization of the rhotic sound (/r/) in Portuguese. He claims that when nasal vowels are adjacent to the rhotic sound, the strong variant of /r/ is realized, suggesting the presence of a consonant closing the syllable instead of a nasal vowel. However, this argument is refuted by the existence of strong /r/ between vowels in various Portuguese words, indicating its productive distribution and compatibility with the monophonemic nasal vowels.

While Câmara Jr. presents persuasive arguments in support of the biphonemic hypothesis, some of these arguments face scrutiny and counterarguments from other linguists. The claim that elision patterns and the realization of the strong /r/ sound provide evidence for the biphonemic status of nasal vowels is not universally accepted. Alternative explanations, such as distinct stress patterns and the nature of vocalic segments, suggest that these phenomena may not necessarily stem from the presence of a nasal consonant in the phonological representation.

Moreover, the biphonemic hypothesis fails to account for several empirical phenomena that support the monophonemic hypothesis, such as the existence of minimal pairs differentiated only by nasal vowels, restrictions on vowel sequences, and the phonetic realization of nasal vowels in specific contexts.

The monophonemic hypothesis suggests that nasal vowels are independent phonological units, and nasalization is considered an inherent property of the vowel itself and is not a result of assimilation or coarticulation with neighboring segments.

If nasal sounds are considered monophonemic units, it affects the analysis of phonological processes, such as assimilation and sandhi phenomena, where nasalization plays a crucial role. However, it is important to note that the monophonemic hypothesis is not without controversies.

The biphonemic hypothesis of Câmara Jr. proposes that nasal vowels in Portuguese are composed of an oral vowel followed by a nasal consonant within the same syllable. While this hypothesis has been influential in the field, it faces criticism and challenges from proponents of the monophonemic hypothesis.

The hypothesis of the biphonemic above is the representation of the underlying structure of the combination / V + CN /, the nasal vowels are composed of an oral vowel followed by a nasal consonant within the same syllable, resulting in / VN/, as "**mundo**" [mũN'du] (world), "**cantar**" [kẽN'tar] (to sing), "**sangue**" ['sẽN.gu] (blood) or "**pingo**" ['pĩN.gu] (drop).

These examples reflect the biphonemic hypothesis and the perspective that nasal vowels are analyzed as a sequence of an oral vowel and a nasal consonant within the same syllable.

The VN (vowel-nasal) analysis considers nasal vowels as vowel-consonant sequences. This analysis was also proposed in generative terms by Lightner (1970) and Mateus (1975) and has been reinterpreted in autosegmental terms of CV phonology. According to this analysis, nasal vowels become associated with a second-time unit in the CV tier, representing compensatory lengthening in CV phonology. The VN analysis can account for the phonetic observation that nasal vowels are longer than oral vowels.

The nasalization of vowels is described as suprasegmental and fluctuating (BISOL, 1998). Nasalization involves the nasal resonance of a vowel sound, typically caused by the influence of a nasal consonant. The nasal feature extends beyond the

individual vowel segment and is influenced by the following nasal consonant and this influence can vary in degree or intensity depending on factors such as speech rate, emphasis, or context.

However, the nasal vowels are contrastive because both have two root nodes in the form of an oral vowel followed by a nasal consonant. That is, they carry distinct meanings and can differentiate words. In this view, both oral vowels (vowels produced without nasalization) and nasal vowels have underlying representations with two root nodes. One node represents the oral vowel itself, and the other node represents the nasal consonant that follows the vowel.

According to this perspective, nasalization occurs through assimilation, where the oral vowel assimilates or takes on the nasal quality of the following nasal consonant. This results in an underlying representation of a nasal consonant in the offset or final position of the vowel. This process is proposed by Cagliari (1977) and implies that nasalization involves the incorporation of nasal consonantal features into the vowel, creating a nasalized vowel. According to Câmara Jr. (1970), nasalization can also be a phonetic process.

The proposal of a double representation hypothesis by Câmara Jr. (1953; 1970), shows that nasality spreads leftwards to an underlying oral vowel from a neighboring nasal consonant. The nasal consonant is then deleted, except when followed by a plosive, in which case an audible homorganic nasal consonant might be inserted. This hypothesis suggests that Portuguese nasal vowels occur in heavy syllables and is supported by diachronic studies on the evolution of contrastive nasal vowels in Romance languages.

As we mentioned, the nasalization of vowels in BP has been a topic of discussion among phonologists, with different viewpoints proposed by various authors (LIPSKI, 1975; MATTE, 1975; CAGLIARI, 1977; PARKINSON, 1983; BISOL, 1998; SEARA, 2000; SOUSA, 1994; JESUS, 2002; MEDEIROS & DEMOLIN, 2006; MEDEIROS, 2007). One widely accepted structuralist perspective classifies the nasal vowel as a

nasal archiphoneme (CÂMARA JR., 1970), providing a basis for understanding the phonological status of nasal vowels in BP.

According to this viewpoint, the consonant assimilates the point of articulation of the vowel, as in “*tinta*” /tiNta/ [ˈtĩtɨ] (paint). The oral vowel becomes nasalized through assimilation, resulting in an underlying nasal consonant in the offset position (CAGLIARI, 1977).

Traditionally, it was believed that the nasal consonant disappeared during the historical sound change process in BP due to assimilation. However, an alternative perspective suggests that the assimilation process leads to the emergence of a velar nasal consonant without oral occlusion (LIPSKI, 1975). Shosted (2006) also identifies this occlusive consonant as a velar nasal /ŋ/ in the offset position. Examples such as “manga” [ˈmãŋga] (mango) and “branco” [ˈbrãŋku] (white) illustrate this assimilation process.

In the context where the nasal vowel is followed by a velar stop /k/ or /g/, velarization becomes more prominent. This description suggests that the emerging consonant shares the same place of articulation as the adjacent vowel.

The nasal consonant and the vowel are produced with a continuous or connected articulatory movement. It indicates that the articulatory gestures for the vowel and the nasal consonant are coordinated, with the consonant emerging seamlessly from the movement of the tongue during the production of the vowel. Thus, front vowels have constriction in the palatal region while back vowels exhibit velar constriction, according to this point of view the nasal feature should be in this articulation place.

Moraes & Wetzels (1992) regarding the duration of nasal vowels in Brazilian Portuguese, that nasal vowels were longer than their oral counterparts, with greater duration in stressed and pre-stress positions. The crucial distinction was observed in pre-stressed syllables before a plosive, where contrastive nasal vowels were significantly longer than oral vowels and allophonic nasal vowels.

In BP, the unstressed monophthong /a/ has the potential to undergo nasalization when followed by a nasal consonant in the syllable onset. This nasalization process



can be observed in words such as "**cama**" ['kã.ma] (bed), "**Camila**" [kã'mila ~ ka'mila] (noun), and "**Tamara**" [tã'mara ~ ta'mara] (noun).

Among the vowel sounds, the back-middle vowel [o] and the central vowel [a] are more susceptible to regressive nasalization in this process. The coarticulation between the movement of the tongue and the displacement of the velum contributes to this nasalization. When producing the unstressed vowel in conjunction with a nasal consonant, the tongue tends to retract further back and come closer to the soft palate region (MENDONÇA, 2019).

It is important to note that the occurrence of nasalization in this context is variable, meaning that both oral and nasal production can be observed in different contexts. Tonic vowels, regardless of stress, may exhibit nasalization, while unstressed vowels may or may not undergo nasalization.

Back vowels are more prone to nasalization, whereas front unstressed vowels may or may not assimilate. In certain cases, the presence of an accent can impede the nasal assimilation process. For example, in the word "**ruim**" [hũĩ ~ hu'ĩ] (bad), the nasal assimilation is halted due to the accent, whereas in "**ruína**" [hu'ine] (doom), nasalization occurs (MENDONÇA, 2019).

According to Abaurre & Pagotto (1996), the assimilation of nasalization<sup>22</sup> is influenced by socio-linguistic factors. There is little research on the vocalic nasality phenomenon in BP in phonetics, as most studies are phonological and consider the Structuralism approach or focus on acoustic analysis and descriptions.

Nasalization is seen as a complex production. The process of establishing consistent acoustic patterns or measurements for nasal assimilation phenomena is challenging. It implies the need for defining and establishing criteria or norms that can be used to analyze and compare nasalization effects in speech, particularly concerning nasal vowels and diphthongs. Additionally, accessing experimental

---

<sup>22</sup> The assimilation of nasalization signifies the incorporation or integration of nasal features into the pronunciation of certain vowels or consonants, resulting in the nasalization of the sound.

equipment to gather various types of data such as articulatory, aerodynamic, and imaging information during speech production poses a barrier (CAGLIARI, 2007).

According to Maia (1983), who described the characteristics of nasal vowels, vocalic nasality is a suprasegmental characteristic that modifies the quality of the oral vowel. These characteristics are the result of the coupling of oral and nasal cavities and changes in the geometry of the vocal tract.

The coupling of the oral and nasal cavities enables the production of nasal vowels without constriction of the oral tract. When airflow passes through the glottis, the air particles undergo modification in the filter, resulting in a specific output. This articulatory production creates a type of bifurcation in the tube and establishes a parallel resonant cavity through which the sound propagates.

Regarding the context of nasalization, a vowel is only considered nasal when it appears in a syllable context that does not assimilate a nasal consonant at the syllable boundary position. In BP, the vowel exhibits a more oral portion (GREGIO, 2006), indicating that it retains its oral characteristics despite the presence of nasality. For example, a nasal vowel maintains an oral component until the production of the nasal appendix, which refers to the transition between the vowel and the following nasal consonant.

Acoustically, the resonance of the nasal cavity extends to all productions of the nasal vowels. In BP, the vowel presents an oral vowel portion, a nasal vowel portion and a nasal appendix<sup>23</sup>. For example, in the word "**mãe**" [mãj] (mother), the nasal vowel /ã/ is produced with the oral cavity relatively open and the tongue in a back position. When the velum is opening allows the airflow to pass through the nose, resulting in nasalization.

In the beginning of the production of [ã] the velum is not opened, so this portion is oral, but nasalization modifies the global quality of the vowel and gives it a nasal

---

<sup>23</sup> Nasal appendix refers to a phenomenon observed in certain languages where nasal vowels occur with specific acoustic characteristics. It is characterized by the sudden loss of formant structure above 1000 hertz and the emergence of a very low frequency voiced nasal formant (Ohala & Ohala, 1991).

character. The coupling of oral and nasal cavities is gradual, that's results in this gradual nasalization of the vowel to finishing the realization with the nasal murmur.

The nasal murmur can be isolated from the vowel preceding a voiceless stop consonant. Three continuous phases are involved in producing the nasal vowel, as the movement of formants is gradual. Between the initial portion in the most oral part and the nasal appendix, there is a nasal vowel which is a nasal segment that has vocalic characteristics (SOUSA, 1994).

The three continuous phases involved in producing a nasal vowel are as follows: (a) oral phase: the vocal tract is primarily configured for producing an oral vowel sound. The tongue position, lip rounding, and other articulatory features are set for the specific oral vowel being produced; (b) nasal transition phase: the vocal tract begins to transition from the oral configuration to the nasal configuration.

When the velum (soft palate) is lowered, it allows air to flow into the nasal cavity while maintaining the oral articulatory position. This transition is gradual and involves the coordinated movement of articulatory structures; and (c) nasal phase: at the end of the articulation, the vocal tract assumes a nasal configuration.

The velum is fully lowered, creating a coupling between the oral and nasal cavities. The airflow passes through the nasal cavity, resulting in the nasalization of the vowel sound. The formant structure is modified due to the resonant properties of the nasal cavity, resulting in the characteristic acoustic properties of nasal vowels and the nasal appendix.

In our data is possible to analyze the gradual transition from oral to nasal configurations during the production of a nasal vowel. In aerodynamic, articulatory, acoustic, and imageries data, we can visualize the airflow variation, the articulatory movement, the movement of formants, and the velum lowering during the nasal vowel and nasal diphthong production.

### **2.1.3 Oral and Nasal Diphthongs**

Diphthongs emerged during the evolution of Latin, which had four oral diphthongs in its inventory, [ae], [aw], [ew], and [oe] which had evolved from oral monophthongs and oral diphthongs. Diphthongs are the combination of two vocal segments in the same syllable; the two possibilities are vowel + semivowel/semiconsonant or semivowel/semiconsonant + vowel.

Diphthongs are sounds that are described and analyzed in the field as a separated sound element connected by a transition between two or three elements as a vowel to glide<sup>24</sup> or vice versa. However, diphthongs are considered distinct units or entities within a syllable. They are not perceived as two individual vowel sounds, but rather as a single unit with a transitional movement between the two vowel elements.

Glide production occurs when there is a constriction in the vocal tract in the palatal [j] or velar [w] region, with lip rounding in the production of the latter. These sounds are represented by the phonetic symbols [j] and [w]<sup>25</sup>, respectively, and the nature of the diphthongs changes according to the tongue movement. The tongue position changes during the diphthong production, that is, the tongue develops a transition from one part of the diphthong (one vowel sound) to another, which affects the articulatory properties and the resulting formant movement

A glide is a sonorant segment that coarticulates with an adjacent vowel in the same syllable. Due to the mode of articulation and the constriction, this sound is a semi-consonant (a semivowel or an approximant consonant) and differs from vowels.

---

<sup>24</sup> When these sounds are in a syllable's peripheral positions, they are called glides or semivowels/ semi-consonant because they have less prominence compared to vowels. A glide is usually a non-syllabic segment.

<sup>25</sup> These sounds are glides or semivowels/ semi-consonant, a phonetic unit *like-vowel* with a shorter duration than the vowels. They are in a non-stressed position, and they are not permitted to be in the syllable's core (MATEUS, 2002, p. 172; MATEUS, 2003, p. 993).

Falling diphthongs are compounded by the glides /j/ or /w/ in the syllable; the same also occurs with rising diphthongs, and both can occur in unstressed or stressed syllables.

The phonological inventory of BP has oral and nasal diphthongs, as well as triphthongs. There are 28 oral diphthongs:

(a) Oral falling diphthongs: [aj, ej, εj, oj, ɔj, uj, aw, ew, εw, iw, ow, uw, ɔw];

(b) oral rising diphthongs: [ja, je, jo, jo, we, we, wε, wi];

(c) Nasal falling diphthongs: [ãw, ãj, ěj, õj, ũj]; and

(d) nasal rising diphthongs: [wã wĩ].

In falling diphthongs<sup>26</sup>, sounds start in a more sonorant sound position and pass to a less prominent sound position<sup>27</sup>. This sound change results in a drop in tension in the movement. That is, the syllable increases in acoustic volume at the beginning and decreases until the end. At the phonetic level of speech production, vowels and glides are at the syllable's core (MATEUS & ANDRADE, 2002, p. 285), and the rhyme is produced by a nuclear vowel followed by the glide<sup>28</sup>.

The cluster can occur in stressed, pre-stressed, and post-stressed syllables. This sound sequence is inseparable. The rising diphthong<sup>29</sup> is a sound sequence in which the most prominent segment is in the sound sequence passing from the glide to the vowel. Hence, the glide is at the core with the vowel in a complex syllabic core.

---

<sup>26</sup> In falling diphthongs, the beginning of the segment is more prominent than is the last part. In rising diphthongs, the last part of the segment is more prominent than is the beginning part; these kinds of diphthongs are preceded by a velar consonant (CAGLIARI, 1977, p. 6).

<sup>27</sup> The transition from vowel to glide follows the sonority principle and the decreasing sound scale.

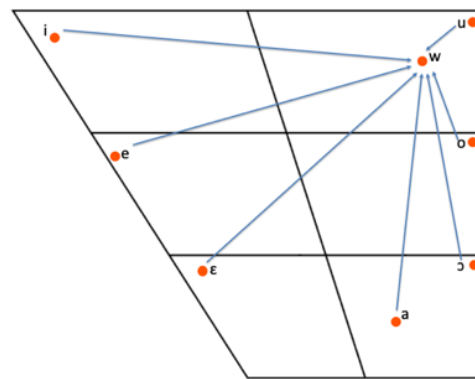
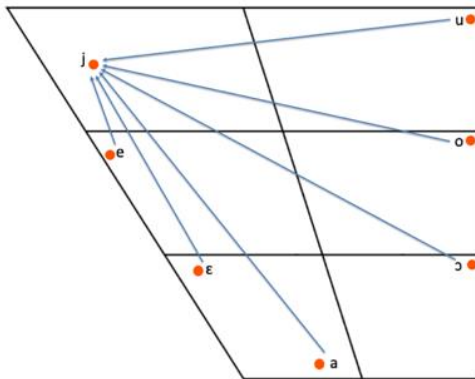
<sup>28</sup> As the glide is phonologically underspecified, both elements in the syllable are affected by the nasality assimilation process (MATEUS & ANDRADE, 2002, p. 285).

<sup>29</sup> The rising diphthong is described as if it did not exist at the phonological level because it assimilates the status of the preceding vowel (MATA DA SILVA, 1987, p. 388; BISOL, 1999, p. 725; ANDRADE & VIANA, 1994, p. 31; MATEUS & ANDRADE, 2000, p. 49; FREITAS, 2001, p. 215; OLIVEIRA, MOUTINHO & TEIXEIRA, 2005, p. 2933).

The distribution of these diphthongs is not similar; some of them occur in an initial, medial, or final position, and glides have a shorter duration and less intensity (BISOL, 1989, p. 215; OLIVEIRA, 2009, p. 72).

Furthermore, the perception of diphthongs often benefits from a noticeable pause or break between the two vowels (ANDRADE, 1998, p. 91-92; CABELEIRA & CORREIA, 2003, p. 357), which can be facilitated by factors such as accentuation or other linguistic cues. This pause allows for the separation and differentiation of the individual vowels in a sequence. Cagliari (1977, p. 13-14) presented the different diphthong qualities resulting from the sound movement that can be observed in the following chart:

(a) Oral falling diphthongs



(b) Oral rising diphthongs

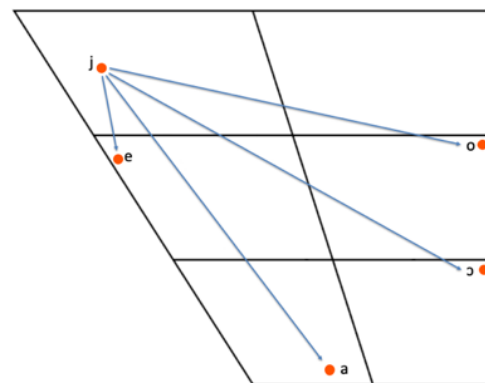
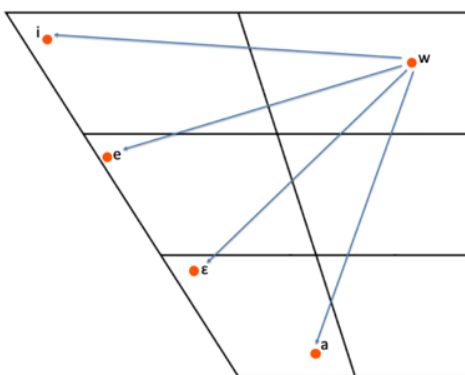


Figure 6: Diphthong chart diagram a) oral falling diphthongs and (b) oral rising diphthongs. Acoustic articulatory adaptation of nasal diphthongs diagram (CAGLIARI, 1970, p. 13-14). The arrows indicate the direction of tongue motion during speech production.

Diphthongization is a phonetic process that involves the transformation of a vowel into a diphthong. A diphthong is a sequence of two vowel sounds pronounced within a single syllable. During diphthongization, a vowel changes its articulation, typically from a monophthong (a single vowel sound) to a diphthong (a combination of two vowel sounds).

The process of diphthongization often occurs when a stressed vowel experiences a relaxation or reduction in its tense articulation (CHOMSKY & HALLE, 1968). As a result, the vowel unfolds or elongates, creating an opportunity for a semivowel (an approximant type of vowel-like sound) to be pronounced alongside it. This combination of the vowel and semivowel within the same syllable gives rise to a diphthong.

The specific nature of diphthongization varies depending on the language and dialect. The relaxation of the stressed vowel can lead to different patterns of vowel unfolding and the introduction of specific glides, resulting in various types of diphthongs (CHOMSKY & HALLE, 1968). These diphthongs can be characterized by the movement or glide from one vowel quality to another within a single syllable.

When the vowel breaks, a new sound is created. It is commonly observed that diphthongs, which involve a transition from one vowel to another, may subjectively feel longer than monophthongs, which are single vowel sounds.

However, the actual duration of vowel sounds can vary based on various factors, including language, individual speech patterns, and contextual influences. This diphthongization process results in longer vocalic segments, that operate within the

phonetic and phonological system of a language, adhering to patterns and principles of phonetic harmony.

Diphthongization is a phonetic process where a vowel transforms into a diphthong, typically triggered and influenced by the surrounding sounds or the boundary context. During diphthongization, the original vowel is perceived as two distinct segments: the first segment representing the vowel itself, and the second segment having a vocalic nature.

Diphthongization is vowel breaking<sup>30</sup> via the process of assimilating to a following vowel or consonant. The vowel breaking refers to the process of a vowel assimilating or adjusting its pronunciation to match a following vowel or consonant. This assimilation can result in the creation of a diphthong.

In many cases, diphthongization occurs when a stressed vowel is accompanied by an approximant, such as [j] or [w], added to its front or back position. This process can affect the speed of speech, and the efficiency of communication, and contribute to speech variation among individuals.

Diphthongization is a diachronic evolution. This process affected several words in the evolution of Latin to Portuguese: *ce(n)a* > *cea* > *ceia* (dinner); *lacte* > *leite* (milk); *me-u* > *meu* (my) and *credo* > *creo* > *credo* (creed).

Diphthongization emerges due to some boundary articulatory processes. As mentioned previously, // in the coda is vocalized in some BP dialects. In the same position, the allophone [j] can give rise to iodization<sup>31</sup> (diphthongization) in the preceding vowel, as in “*feroz*” [fe'roj̥s ~ fe'ro(j)ɛ] (fierce) or “*arroz*” [a'ʁoj̥s ~ ɐ'ʁo(j)ɛ] (rice) (BARBOSA & ALBANO, 2004, p. 230).

---

<sup>30</sup> The phenomenon of vowel breaking is a well-established linguistic term that refers to the process of breaking a long vowel sound into two shorter vowel sounds, typically caused by the presence of certain sounds such as consonants or glides (LADEFOGED & JOHNSON, 2014).

<sup>31</sup> Yodisation is the emergence of the sound /j/ because of phonological process caused by a sound changing in the palatal region (MEILLET, 1896).



Iodization also occurs by inserting a front high glide into the offset stressed syllable, as in “**mês**” [ˈmejs] (mouth) “**vocês**” [vo.ˈsejs] (you – plural), or in sandhi process, as in “**paz mundial**” [ˈpajz mūdʒiˈaw ~ ˈpa(j)ʒ mūdʒiˈaw] (world peace) and **dez anos** [ˌdɛjˈz\_ẽnu(j)s ~ ˌdɛjˈz\_ẽnuɐ] (ten years) (BARBOSA & ALBANO, 2004).

In rising diphthongs, the semivowel that is less prominent appears before the vowel because of higher prominence, as in /k<sup>w</sup>/ and /g<sup>w</sup>/. Cavalieri (2005) and Cristófaró (2002) described these sounds as complex consonants with a velarization or labialization feature.

Palatal glides can also occur in free variations, as in “**quiabo**” [kiˈabʊ ~ ˈkjabʊ] (okra) (BISOL, 2005, p. 112), while in the falling (descending) diphthong, the glide occurs in the less prominent offset position, as in “**sei**” [ˈsej] (know). Therefore, oral diphthongization is a phonetic sound change in which a monophthong becomes a diphthong. This phenomenon results in a vowel breaking with a glide insertion in the syllable (CÂMARA JR., 1986).

Some contexts are favorable due to coarticulatory gradience; for example, the oral diphthongization occurs by inserting a back high glide into the stressed syllable, as in “**boa**” [ˈbowa] (good) and “**soa**” [ˈsowa].

From the point of view of speech production, the oral diphthong is a sound continuum that is performed during a single respiratory impulse. Thus, as mentioned previously, the tongue moves from a more prominent position to a less prominent one.

Glides are speech sounds that fall between vowels and consonants and are characterized by their smooth transitions between adjacent sound segments. A glide is vowel-like or a consonant based on its degree of constriction (STEVENS, 1998). This constriction is narrow and therefore causes a significant drop in oral pressure,

but they are often produced without a complete vocal tract obstruction and do not exhibit abrupt discontinuities in the acoustic signal.

Thus, oral diphthongs are continuous sound that is more dynamic than vowels due to tongue transitions. Glides, such as /j/ and /w/ are characterized by relatively narrow constrictions in the oral cavity. These constrictions differ from those of vowels in terms of their degree and location. Glides have a constriction that is narrower than high vowels, resulting in a smaller cross-sectional area. This narrower constriction leads to a lower first formant frequency (F1) compared to high vowels.

The glide articulation mode is different from vowels because glide production requires a critical constriction of 0.17 cm in the oral cavity (STEVENS, 1980). Stevens compared the nature of the glide to that of nasal consonants<sup>32</sup> because, during the production of a glide, the vocal folds vibrate continuously. These sounds are produced with little or no airflow or pressure dropping above the glottis.

In glide articulation, the elevation of the tongue produces a slight constriction in the oral tract. The first formant frequency in glides and high vowels is influenced by the size of the cavity behind the constriction and the length of the constriction. By creating a larger cavity behind a long constriction, the first formant frequency is lowered. In the case of glides, both the primary and secondary constrictions contribute to the lowering of F1. However, vocal tract wall effects and limitations prevent F1 from reaching zero, even with a complete closure (HUNT, 2005).

The increased formant bandwidth in glides can be attributed to the acoustic conductance of the vocal tract walls and the kinetic pressure drop across the narrow constriction. These factors contribute to the overall acoustic distinction between glides and vowels (HUNT, 2005).

---

<sup>32</sup> The oral cavity is closed when producing nasal consonants and the airflow is directed to the nasal cavities.

The nasal diphthongs within the scope of this research follow diachronic changes, as in a) *veranu* > *verão* (summer); b) *pane* > *pão* (bread); c) *cane* > *cão* (dog), d) *manu* > *mão* (hand) and e) *bene* > *bem* (good) (BISOL, 2002). Nasal diphthongs occur frequently in BP, and there are four phonemes in the phonetic inventory: [ãw̃, ãĩ, õĩ, õj].

The diphthong [ãw̃] is very common and contrasts phonologically with [aw]. The first phoneme is produced in word-final syllables and can be stressed or post-stressed as a verbal suffix of the third person plural masculine superlatives, and in some nouns; for example, in the stressed syllable: “*não*” (no – adverb of negation) versus “*nau*” [naw] (synonymous of the boat – noun). Different nasal diphthong sound movements are presented in the following chart:

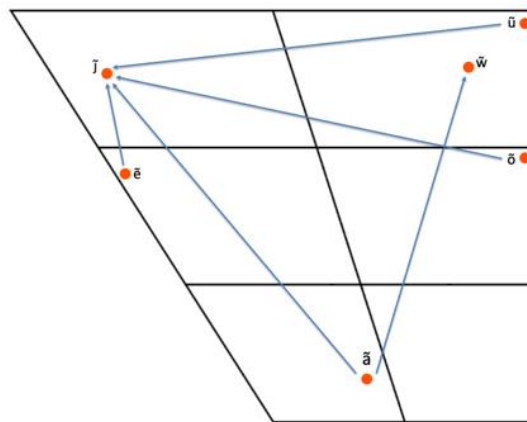


Figure 7: Diphthong chart diagram. Acoustic articulatory adaptation of nasal diphthongs diagram (CAGLIARI, 1970, p. 13-14). The arrows indicate the direction of tongue motion during speech production.

The diphthong [ãĩ] occurs in words such as “*mãe*” [mãĩ] (mother) and the post-stressed position, as in “*cantam*” [ˈkãtãw̃] (to sing – the third person plural). The segment [õĩ] occurs in the term “*muito*” [ˈmũĩtu] (much). The nasal diphthong [õj] occurs in stressed words such as “*leões*” [leˈõj] (lions).

The front nasal diphthong [ẽ̃] is an allophonic representation of the front nasal vowel [ẽ] and occurs in stressed words such as “*sem*” or “*tem*” and “*perdem*” [ˈpɛrtẽ̃] (to lose – third person plural) and “*sentem*” [ˈsẽ̃jtẽ̃] (to feel – third-person plural).

We listed some natural languages that contain nasal diphthongs, such as Koromfe<sup>33</sup>, Taiwanese<sup>34</sup>, Kuikuro<sup>35</sup>, Denesuline<sup>36</sup>, Birman<sup>37</sup>, Chinese Ningbo<sup>38</sup>, Kakka dialect<sup>39</sup>, Breton<sup>40</sup>, Polish<sup>41</sup>, Pumi and Bai<sup>42</sup>.

However, it is important to note that our knowledge about nasal diphthongs in natural languages is restricted by the availability of data, and there may be languages that have not been thoroughly documented or studied, which could potentially include nasal diphthongs.

Nasal diphthongs and nasal vowel diphthongization are complex and quite frequent segments in BP. However, the mode of production of these sounds can vary depending on the context, the vowel height, the degree of the soft palate opening, and the speech speed.

---

<sup>33</sup> Koromfe is a Gur language of the Niger-Congo family. It contains nasal diphthongs and triphthongs: [ãĩ], [wãĩ], [wẽĩ], and [wõĩ] (RENNISON, 1997).

<sup>34</sup> Thai has three contrasting nasal diphthongs: [ĩã], [ũã], [ũãỹ] (CHUNG, 1996).

<sup>35</sup> Kuikuro is a member of the Carib language family spoken in the Alto Xingu region in northern Mato Grosso. There is a nasal diphthong and a nasal triphthong in this language: [ũã] and [ũãĩ] (ESPÍRITO SANTO, 2010).

<sup>36</sup> Dënesųłin is a language spoken by the Chipewyan people in the northwest of Canada. This language is part of the Northern Athabaskan family, and contains four nasal diphthongs: [ẽĩ], [ãĩ], [ãũ], [õũ] (CHANG, 2008).

<sup>37</sup> Burmese is a language from the Sino-Tibetan family that is spoken by the Bamar ethnic group. There are four nasal diphthongs in its phonological inventory: [ẽĩ], [ãĩ], [ãũ], [õũ] (CHANG, 2008).

<sup>38</sup> Ningbo is a dialect in the Taihu family and is a subgroup of Mingzhou branch. This language is spoken on the south coast in the Chinese section. There are four nasal diphthongs: [ĩã, ũã, ũõ, ỹõ] (ZANG & HU, 2005).

<sup>39</sup> Tieshiko Hakka is a dialect in the Mandarin family, spoken in Xiaofeng province in China. The diphthong [ãũ] appears in the context of a velar nasal (HASHIMOTO & MANTARO, 2010)

<sup>40</sup> Breton is a language in the Celtic family and a minority language that is spoken in lower Brittany. It contains a nasal diphthong, [ãw̃], which also has the allophone [ãõ] (REENEN, 1982).

<sup>41</sup> Polish has three nasal diphthongs: [aũ̃], [uũ̃] and [iũ̃] (GUSSMANN, 1975).

<sup>42</sup> Pumi and Bai are Tibetan-Burma languages that are spoken in Yunnan province and contain the following nasal diphthongs: [jẽ̃], [jõ̃], [ɥẽ̃] and [wã̃] (NAMKUNG, 1996).

The production of nasal diphthongization in the Paulistano dialect of BP is a dialectal variation. Moreover, there are indications that this dialectal variation is a prevailing pattern or inclination within the broader context of the language.

It implies that this variation is not limited to isolated instances but rather occurs with a certain frequency in BP<sup>43</sup>. Understanding the variability in nasalization and the allophonic production in the coda is complex because it depends on physiological, dialectal, and perceptual factors.

Both triphthongs and diphthongs are complex vowels because the quality change occurs from a vowel element preceding or following a consonantal element such as /w/ or /j/. This perspective allows ‘(...) different rules for the distribution of the consonants and the vowels (...)’ (CAGLIARI, 1977, p. 8). Nevertheless, the perspective that considers different rules for the distribution of consonants and vowels in complex vowels (diphthongs and triphthongs) in BP will not be adopted in this thesis.

The production of nasal diphthongs and triphthongs is a distinguishing factor of Portuguese, and differentiates it from other Romance languages (TEIXEIRA, 2008). Regarding the peculiarities of these nasal segments, we consider the present research is necessary in the field because it analyzes new types of data pertaining to a rare phenomenon.

The dynamic modeling of the velopharyngeal trajectory in Brazilian Portuguese nasal diphthongs can provide valuable insights and contribute to our understanding of the velopharyngeal movements during the production of nasal diphthongs.

This analysis can provide a more comprehensive understanding of the specific articulatory gestures involved, their timing, coordination, and variability. Our

---

<sup>43</sup> The European Portuguese production of the nasal diphthongs is different from that of BP. We will not compare their production because it is outside of the scope of this work.

quantitative data on the velopharyngeal movements can help identify patterns and trends that may not be readily apparent through qualitative observations alone. This can help identify universal tendencies in the coordination and movement of the velum and pharyngeal structures during the production of nasal diphthongs.

By studying the velopharyngeal trajectory, dynamic modeling can provide insights into the mechanisms and control processes involved in producing nasal diphthongs. It can help us understand how the velum and pharyngeal structures dynamically interact and adjust during speech, shedding light on the coordination of articulatory gestures.

By examining the velopharyngeal trajectory in nasal diphthongs, we can gain insights into how coarticulatory processes operate in the context of nasalization. This can help us understand the principles of coarticulation. It is necessary to perform three concomitant gestures in the same time interval to produce nasal diphthongs; that is, the articulators of the front nasal glide are the closed glottis, the tongue motion, and the velum movement. In the nasal back diphthong, there is additional lip rounding.

So, can shed light on phonological constraints and concrete production which govern nasal diphthongs in Brazilian Portuguese and potentially in other languages as well. For example, if certain phonetic or articulatory patterns consistently align with specific phonological distinctions, it suggests that there may be universal constraints on the phonetic realization of nasal diphthongs.

Brazilian Portuguese nasal diphthongs represent a unique case within the broader context of nasal diphthongs. Analyzing their velopharyngeal trajectory can uncover language-specific characteristics, such as specific patterns of velar movement or pharyngeal shaping. This knowledge contributes to our understanding

of the phonetic and phonological properties of Brazilian Portuguese and allows for cross-linguistic comparisons.

The analysis of the velopharyngeal trajectory can provide evidence regarding the interplay between phonetics and phonology. It can help to investigate how phonetic and articulatory properties of nasal diphthongs align with phonological contrasts and rules, exploring the interaction between phonetics and phonology.

These insights can contribute to our understanding of the relationship between the physical properties of speech sounds and the abstract representations in the linguistic system.

#### **2.1.4 Anatomy and Physiology**

Speech is a complex act that depends on the coordination of the neural system. Some structures of the human body must work together for phonation to occur. The respiratory and digestive systems create a subsystem that is responsible for the phonatory and articulatory processes.

The breathing mechanism is orchestrated by the respiratory system in the human body. The primary function of this system is the absorption of oxygen and the elimination of carbon dioxide from the body. The inhalation and exhalation processes occur due to the variation between the pressure and volume.

The airways enter the nostrils and go through the larynx, following the trachea until the airflow reaches the lungs' alveoli, where gas exchange with the blood occurs. Figure 8 shows the diaphragm and the intercostal muscles being contracted, which increases the intrathoracic volume.

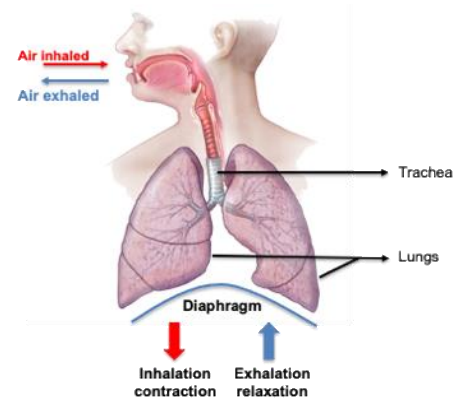


Figure 8: Illustration of inhalation and exhalation process. Adapted from Mayo Clinic website: <https://www.mayoclinic.org/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873>. © Mayo Foundation for Medical Education and Research. All rights reserved.

The alveolar pressure drops when the volume increases. The negative pressure helps the lungs to open to receive the airflow during the inhalation process and gas exchange.

The figure below shows that the thoracic cavity has less atmospheric pressure than the environment, following Boyle's law concerning the relationship between pressure and volume. The intrapleural pressure increases, the lung volume decreases during the expiration process, and the air flows out of the lungs.

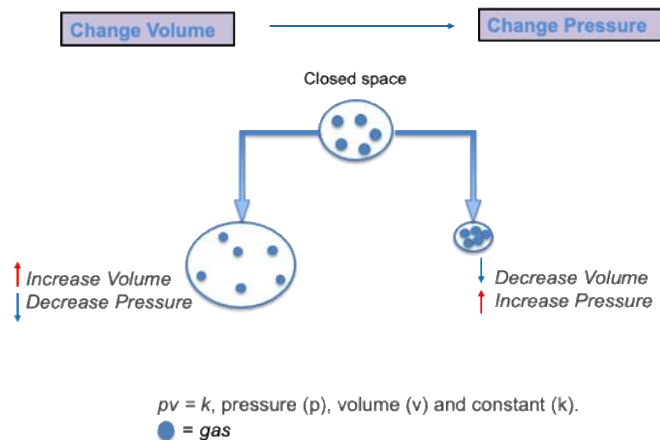


Figure 9: Illustration of Boyle's law. Time-temperature-transformation (TTT) is the relationship between pressure and volume when temperature and mass are constant.

The nasopharynx and the oropharynx are called the upper airway. This set of structures is responsible for swallowing, breathing, and speech production. The tongue motion changes the shape of the vocal tract. Figure 10 below shows the



respiratory system, the cavities of the vocal tract, some places of articulation in the vocal tract, and the glottis region.

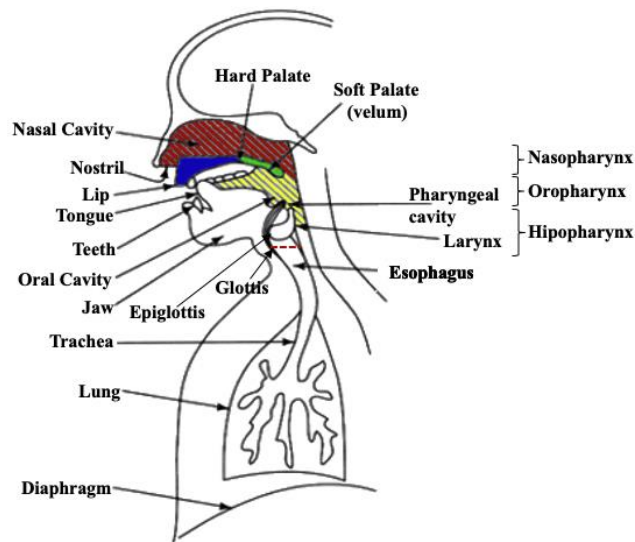


Figure 10: Illustration of the respiratory system adaptation (STEVENS, 1998). Midsagittal section of the vocal tract, vocal folds, and main cavities in the respiratory system.

The mucosa above the trachea comprises two horizontal tissue folds extending across the trachea wall from the thyroid cartilage to the arytenoid cartilage. Figure 11 shows the vocal folds (true vocal folds).

The vocal folds are tissues that are composed of cartilage, muscles, ligaments, arteries, veins, and nerves. Their functions are to produce sound and to prevent foreign bodies from entering the lungs through the airways (HUCHE & ALLALI, 1999).

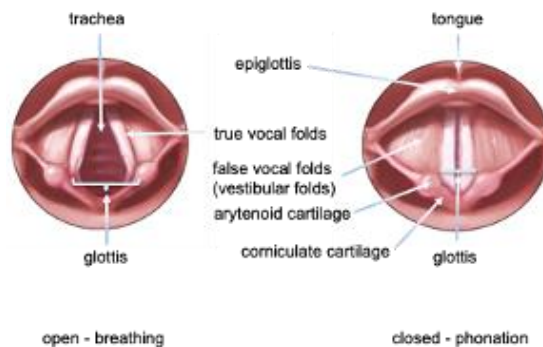


Figure 11: Illustration of open and closed vocal folds. Adapted from Mayo Clinic website: <https://www.mayoclinic.org/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873>. © Mayo Foundation for Medical Education and Research. All rights reserved.

The vocal folds are located inside the laryngeal-pharyngeal cavity in the larynx. The false vocal folds are also called vestibular folds and are located just above the vocal folds. The glottis is also located in this region.

The opening and closing of the glottis create a difference in pressure in the subglottal cavity. The vocal folds must be closed when the airflow passes through the glottis for vocalization to occur.

When the glottis is closed, the pressure inside the trachea increases. Thus, the trapped air causes the folds to separate, letting out air jets, and the airflow passes through the glottis, which is a narrow cavity.

The folds then close again, which generates an intermittent opening and closing movement, that emits a frictional signal. This type of movement causes the tension in vocal folds to vary and, when the airstream is driven into the larynx (DANGELO, 1997), the vibration of the vocal folds produces a resonant sound.

This process is crucial for vocalization. For example, the fundamental frequency is produced via the biomechanical action and individual physiological characteristics. The vocal folds must be abducted (open) when the airflow passes through the trachea to allow breathing to occur. The speech flow must be modulated by passing through the vocal folds and the articulators.

#### **2.1.4.1 Nasal Cavity**

The vocal tract has two main cavities that work together in nasal production; they are the oral and the nasal cavity, as shown in Figure 12 below. These cavities are resonators that modulate the acoustic signal during speech production.

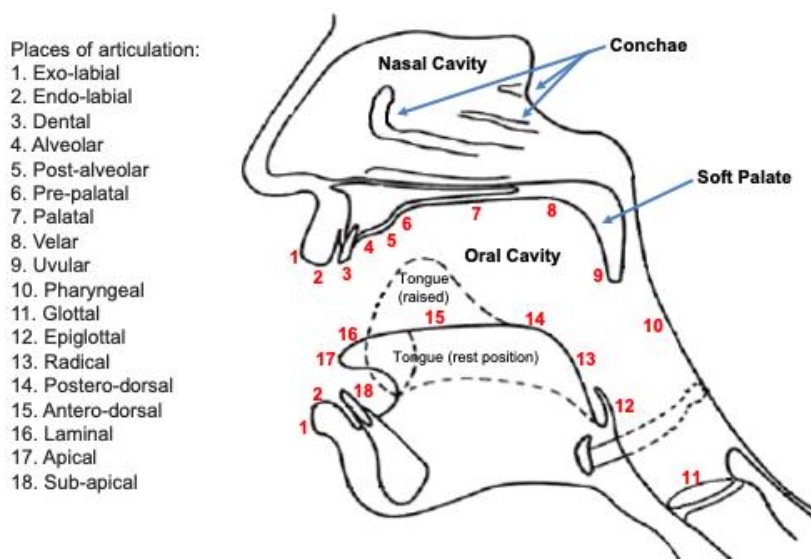


Figure 12: Sagittal mid-section of the oral and nasal cavities. Illustration of the places of articulation and active and passive articulators. Adaptation of the picture of the vocal tract in STEVENS (1998).

In Figure 12, the place of articulation refers to the region in which the airflow is modified by the articulation, for example, the lips or the tongue. The sound quality changes depending on the articulatory point and the vocal tract's geometry.

From the anatomical point of view, the oral and nasal cavities are connected by the soft palate<sup>44</sup>. The nasal cavity is divided into the vestibule, and the respiratory and olfactory segments, as shown in figure 13. The nasal cavity is mainly involved in the breathing process. In phonation, the nasal cavity acts as a static resonator that filters the oral resonances by filtering the speech.

The nasal cavity is a large cylindrical space in the region of the nostrils, the hard and the soft palate, and the brain, extending from the nasal alar anterior to the posterior choana, as shown in figures 13 and 14. This cavity is a cephalic part of the respiratory system.

The primary functions of the nose and paranasal sinuses are breathing, filtering the air, and expelling secretions from the paranasal sinuses. Additionally, the nose is involved in olfaction (sense of smell), while taste perception occurs primarily through

<sup>44</sup> This mucosa regulates its size via an opening and closing movement, creating a small gap called a velopharyngeal section. When it is open, it allows the airflow to pass through the nasal cavity.

taste buds on the tongue. The nasal passages play a role in warming and humidifying the inhaled air before it reaches the lungs (GREGIO, 2006).

The human skull possesses remarkable adaptations that make it both lightweight and highly resistant to impact, ensuring the protection of the brain and vital structures within. One crucial feature contributing to its lightweight nature is the presence of paranasal sinuses, air-filled spaces within the skull bones. These sinuses not only reduce the overall weight of the skull but also play a role in voice resonance and facilitate air circulation.

The anatomical structure of the nasal skeleton consists of an internal septum and an external nose (septal cartilage indicated in light blue in Figure 13). This region extends from the nostrils to the nasal choana posteriorly to the horizontal plate from the palatine bone (hard palate inferiorly). Figure 13 shows the nasal bones.

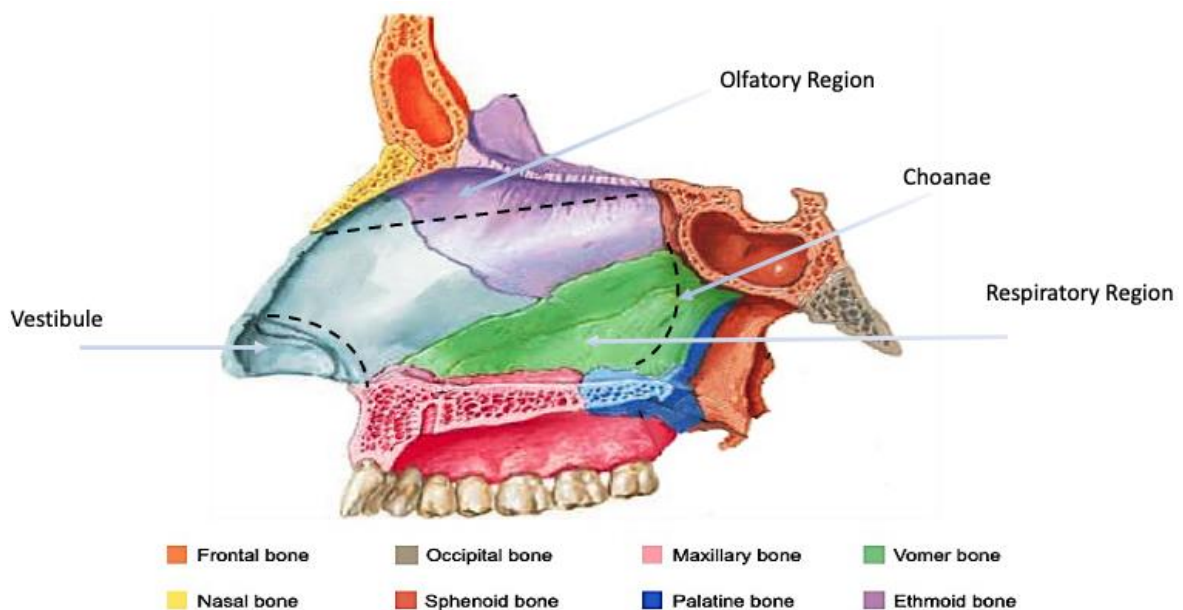


Figure 13: Lateral walls of the nasal cavity (75%). Medial view of the right side of the nasal septum (paramedian section) of the dry skull. The individual parts of bones are represented by different colors. Adapted from WOLF-HEIDDEGGER (2006, p. 70), NETTER (2000), and ARTMED (2000).

The lateral walls of the nasal cavity are irregular. The nasal septum forms the medial wall and is usually smooth. There is a thin bony structure on the lateral walls covered by mucosa called nasal conchae or turbinate.

There are three nasal conchae, namely the superior nasal concha, the middle nasal concha, and the inferior nasal concha. The space inside the nasal concha through which the airflow passes is divided into the upper, the middle, and the lower nasal meatus.

The nasal conchae are limited by the following structures: the blade, the perpendicular of the ethmoid bone, and the vomer. A thin flat bone forms the bottom of the cavity and the inferior part of the nasal septum. The upper and middle conchae are part of the ethmoid bone, whereas the inferior concha is the septal cartilage. The nasal conchae are shown in Figure 14:

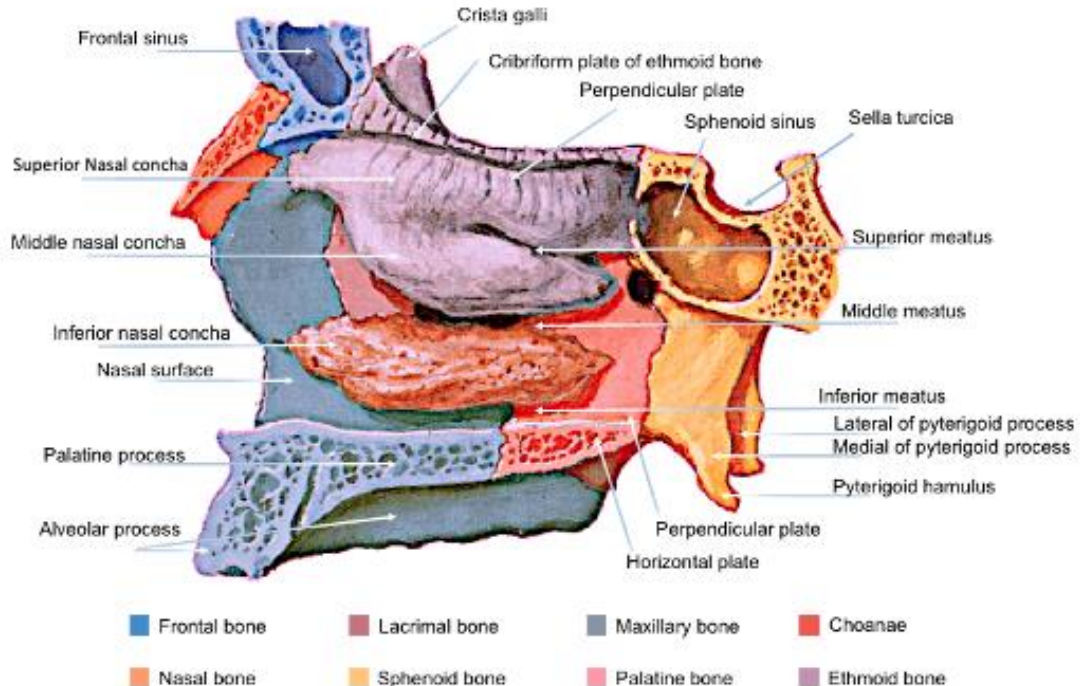


Figure 14: Lateral walls of the nasal cavity (75%). Medial view of bones of the nasal cavity, meatus, and nasal concha. Posterior view. The individual parts of bones are represented in different colors. Adaptation (WOLF-HEIDDEGGER, 2006, p. 70), NETTER (2000), and ARTMED (2000).

In the figure above, we can see that some of the bones that form the nasal cavity are in the posterosuperior space and extend to the superior nasal concha where the sphenoid sinus opens into the sphenoethmoidal recess.

The ethmoid bone forms the roof of the nasal cavity and separates the cavity from the brain and the eye sockets. The olfactory bulbs and nerves are on the side of the crista Galli between the frontal sinus bone, above the cribriform plate, and the upper part of the perpendicular plate of the ethmoid bone.

The palatine bones<sup>45</sup> form the floor of the nasal cavity: the palatine process in the maxillary bone (maxilla anteriorly), the horizontal plate and the perpendicular plate<sup>46</sup>. The connection of these bones at the midline forms an L-shape between the hard palate and the maxillary bones. The hard palate is responsible for dividing the nasal cavity from the oral cavity.

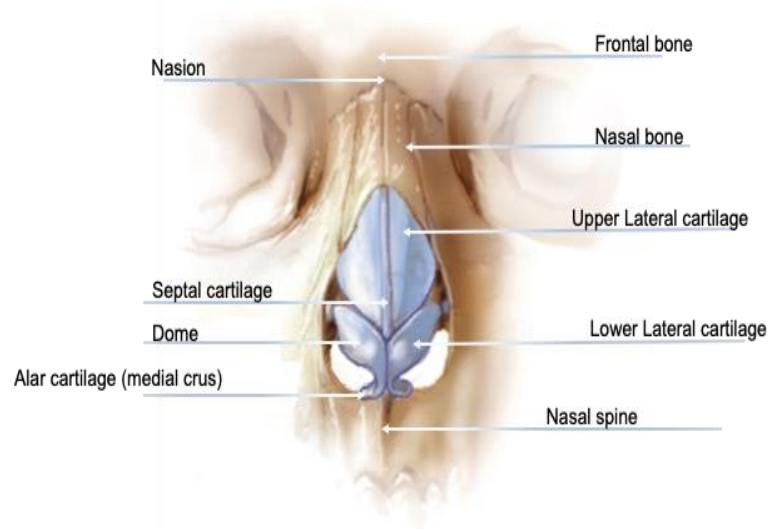


Figure 15: Frontal view of the nasal septum (60%). Bony components are shown in tan and cartilage in blue. Adapted from (WOLF-HEIDDEGGER, 2006, p. 68), NETTER (2000), and ARTMED (2000).

The nose is the outer part of the nasal cavity. The nose is covered with skin and stiff hair called vibrissae (HOLLINSHEAD et al., 1991; DANGELO, 1997). The septum

<sup>45</sup> The hard palate separates the oral and nasal cavities.

<sup>46</sup> The soft palate is composed of muscle fibers, commonly called the roof of the mouth, and its extension separates the nasal cavity from the oral cavity.

structure is composed of a perpendicular plate of the ethmoid bone, the vomer bone, the septal nasal hyaline cartilage and the maxillary crest. The root is the lowest part of the nose, and the upper part is called the summit or apex (SPENCE, 1991).

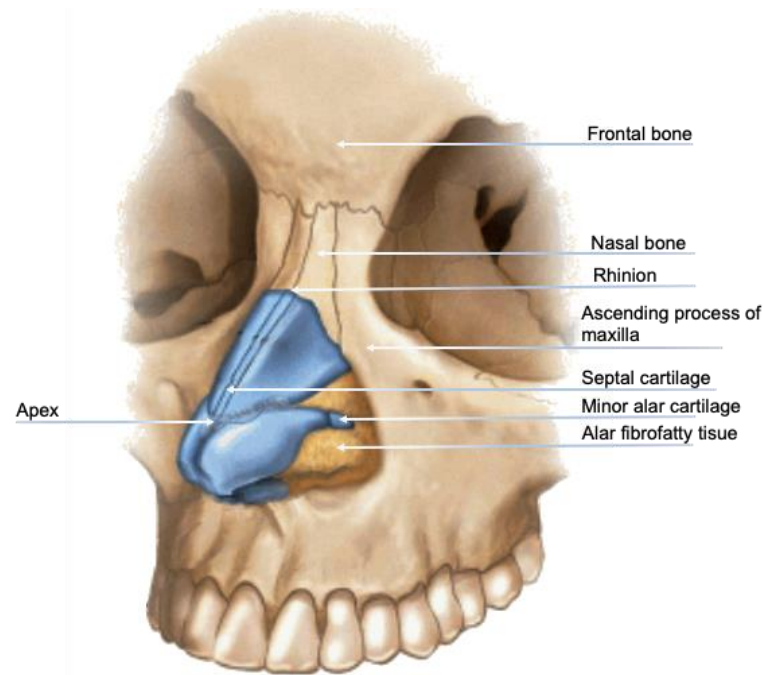


Figure 16: Anterolateral view of the nasal septum and the cartilage in the nasal septum. Bony components are shown in tan and cartilage in blue. Adapted from WOLF-HEIDDEGGER (2006, p. 28), NETTER (2000), and ARTMED (2000).

The nasal septum's cartilaginous bony structure is a protrusion in the center of the face that divides the nose into two small openings called the left and right nostrils. This region extends from the root (base) to the apex (the tip of the nose) and may deviate to the left or right in a curvilinear, concave, or convex shape.

The nose shape may affect its permeability, potentially influencing the ease of breathing through the nasal passages. Structural abnormalities or congestion in the nose could lead to partial nasal obstruction, making it more challenging to breathe through the nose. In such cases, individuals might compensate by relying more on mouth breathing. The total cross-section area of the nasal passage typically ranges from 1cm to 2cm (TEIXEIRA et al., 2001).

The nasal septum separates the nasal passages. These are asymmetrical, and the right side is usually smaller than the left side. The nasal cavities connect to the outside of the body through the nostrils and via the nasopharynx.

The nasal apertures, or the pyriform apertures, are usually called the nostrils. They are paired as two convex apertures that divide the nose into two cavities between the nasal vestibule and the posterior nasal cavity.

The lower ala cartilage partially surrounds the nostril, which is composed of the middle, lateral and medial crus. This flexible cartilage shapes the nasal tip. The lowest cartilage portion of the nose is formed by the footplate, the columella, and the nasal septal cartilage, as shown in figure 17 below.

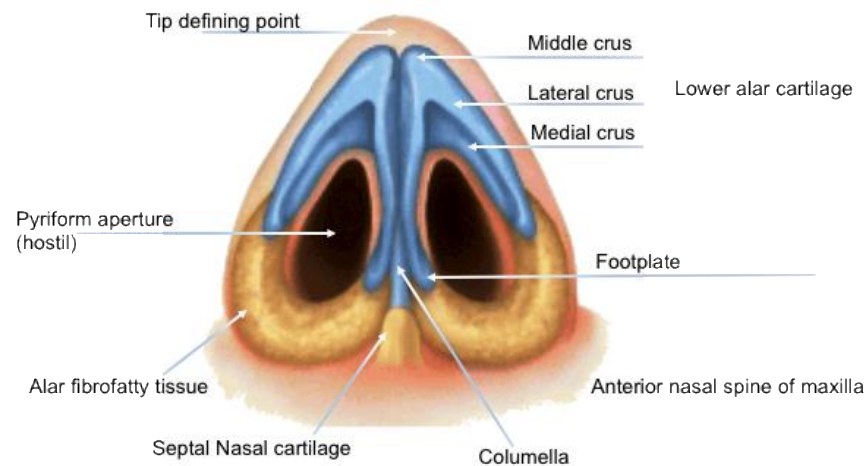


Figure 17: Inferior view. The base of the nose represents the paired alar cartilages and the tipping point of the nose. Adapted from WOLF-HEIDDEGGER (2006, p. 68), NETTER (2000) and ARTMED (2000).

The sinuses anatomy is subdivided into air passages found inside the bone's maxilla, frontal, sphenoid and ethmoid bones. These small cavities are divided by the jaws, and the frontal, sphenoidal and ethmoidal cells. They are lined with mucosa and are full of air.

The Sino nasal anatomy varies from one person to another (JACOB et al., 1990, p. 378) because the size varies and is slightly irregular. The sinuses are involved in mucus production in the nasal cavity (TEIXEIRA et al., 2001).



The ostia are openings that connect the nasal passages to the paranasal sinuses, and their length may vary between 0.4 to 0.6 cm. They are crucial for maintaining proper drainage and ventilation of the sinuses.

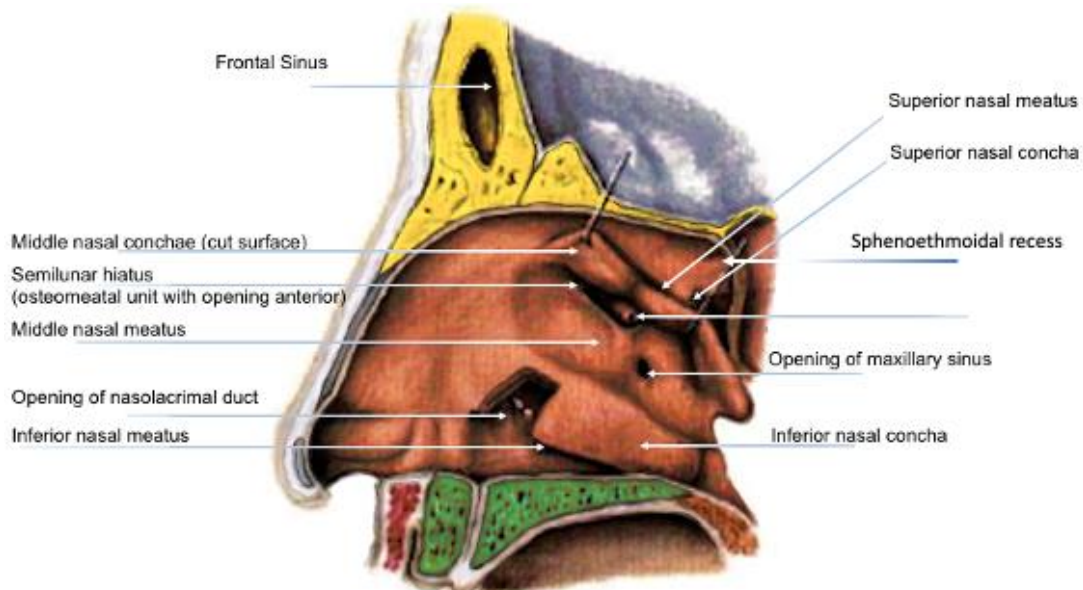


Figure 18: Lateral wall of the nasal cavity and the paranasal sinuses (75%). The paramedian cut the nasal septum. Paranasal sinus projection on the face. Adapted from WOLF-HEIDDEGGER (2006, p. 70).

The figure above shows the meatus of the nasal cavity. The space inside the nasal conchae is called the nasal meatus, and is divided into the upper, middle, and lower sections.

The lateral wall is covered by richly vascularized mucosa that allows a warm airstream to pass through them. The sphenoethmoidal recess is located in a posterosuperior space between the superior concha and the sphenoid sinus, in figure 18.

There are four paranasal sinuses<sup>47</sup>, of which the maxillary sinuses are the most important; they are located in the maxillary bone under the eyes and below the cheeks.

<sup>47</sup> Most paranasal sinuses increase in size from permanent tooth growth and continue to grow in accordance with the size and shape of the face (SPENCE, 1991).

The infraorbital nerve is innervated in this sinus, and it is the most prominent sinus with a volume of 15 ml. The ethmoid infundibulum helps to drain the maxillary sinus.

The frontal sinus is within the frontal bone and above the eyes and empties into the middle meatus. This sinus drains into the frontal recess via the middle meatus; it has the second-largest feature capacity volume on the face, and its volume in an adult can range from 4 ml to 8 ml.

The sphenoid sinus<sup>48</sup> is located posterior to the eyes, behind the nasal cavity's upper portion, and is almost in the central position of the skull. Its volume capacity may range from 4 ml to 5 ml, and it drains the superior meatus (TEIXEIRA et al., 2001) into the sphenoethmoidal recess located within the superior meatus.

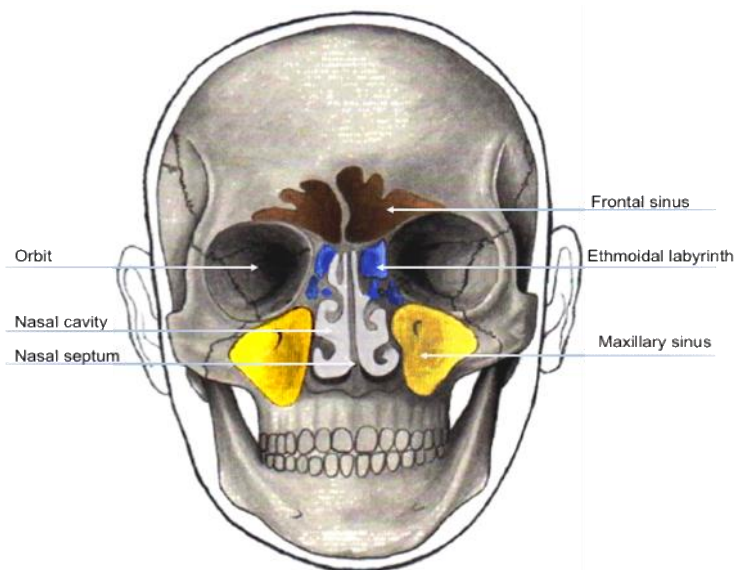


Figure 19: Nasal cavity and paranasal sinuses. The Paranasal sinus projection is in the anterior cranial fossa. Adapted from WOLF-HEIDDEGGER (2006, p. 74).

Finally, the ethmoid sinuses are located between the nose and the eyes, and the sphenoid sinuses are positioned within the sphenoid bone. They are small cavities called ethmoidal cells.

The ethmoidal cells are numerous, irregular and connect with the superior and middle meatus that composes complex labyrinths. They vary in quantity, as there may

<sup>48</sup> The sphenoid sinus may be divided or straight depending on individual characteristics.

be three to eighteen on each side; their total volume ranges from 2 ml to 3 ml (HOLLINSHEAD & ROSSE, 1991).

#### **2.1.4.2 Soft Palate**

The soft palate consists of a thin layer of fibers, muscle, tissue, blood vessels, nerves and glands that are in the back portion of the mouth between the oropharynx and the nasopharynx.

The velum is in the back portion of the mouth between the oropharynx and the nasopharynx, above the skull, lower tongue and pharynx. It is a flexible continuation of the hard palate that extends towards its back region.

The soft palate is the continuation of the pharynx and extends to the opening of the velum. The primary function of the velum is to connect the oral cavity, the pharynx and the nasal cavity.

Unlike the hard palate, the soft palate is mobile and has no bone. The back part of the velum has no connection with any structure other than the uvula, which is in the central process position along the midline.

The velum is 4 cm long, 2 cm wide, and 0.5 cm thick<sup>49</sup>. The velopharyngeal sphincter structure is connected to the lateral walls, posterior to the pharynx, and the soft palate (SOUZA, 2003). The figure below illustrates the velopharyngeal sphincter position in the supraglottic cavity.

(a)

(b)

---

<sup>49</sup> These dimensions are only approximate and vary from one speaker to another (STEVENS, 1998).

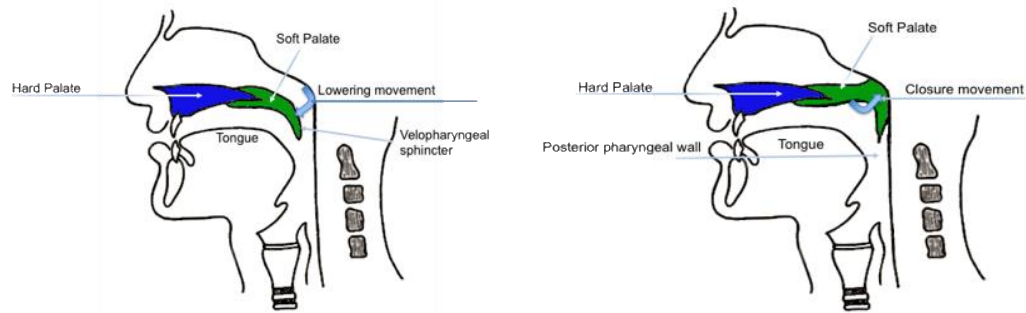


Figure 20: Midsagittal section of the vocal tract. Vocal tract and velopharyngeal sphincter movement representation. Adapted from Kent (1997, p. 189) and Shprintzen & Bardach (1995). In (a), the velum is opened, allowing the airstream to pass through the nasal cavity. In (b), the velum is closed in the default position.

When the velopharyngeal sphincter has lowered, it allows the passage of air through the nasal cavities. The default position is when the velum is elevated and the nasal cavity is closed or almost closed, which reduces the passage of the airstream (LEESON, 1970).

The velopharyngeal cross-section is in the posterior region of the nasopharynx cavity (GREGIO, 2006). This cavity ends at the choana, where it is bisected by the nasal septum.

The movement of the soft palate occurs via the articulation of five principal muscles and the uvula. The uvula is located at the tip of the velum, in the back region of the oral cavity.

The velum and the uvula play articulatory roles in producing nasal sounds and uvular and post-velar consonants (HARDCASTLE, 1976, p. 121).

The pharyngeal muscles are innervated by the pharyngeal branch of the vagus nerve, which is one of the branches of the vagus nerve responsible for supplying motor fibers to the muscles in the pharynx. The group of muscles includes the inferior constrictor muscle, the middle constrictor muscle, and the superior constrictor muscles that cover the pharynx, as shown in figure 21.

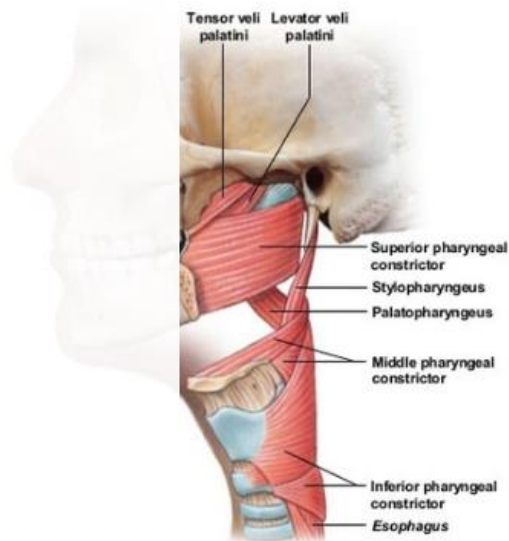


Figure 21: Adapted image of the lateral view of the muscles of the pharynx. All rights reserved to the website Teach Anatomy. Adapted from <https://teachmeanatomy.info/neck/viscera/pharynx/>.

The tensor veli palatini is responsible for keeping the tension of the soft palate to contract and raising the velum and for swallowing (HARDCASTLE 1976, p. 123). The levator veli palatini is responsible for elevating the soft palate. The palatoglossus is accountable for pushing the soft palate towards the tongue.

The palatopharyngeus maintains the tension of the soft palate and assists the pharynx when swallowing. The musculus uvulae are responsible for reducing the size of the uvula. The figure below illustrates the regions and the actions of these muscles.

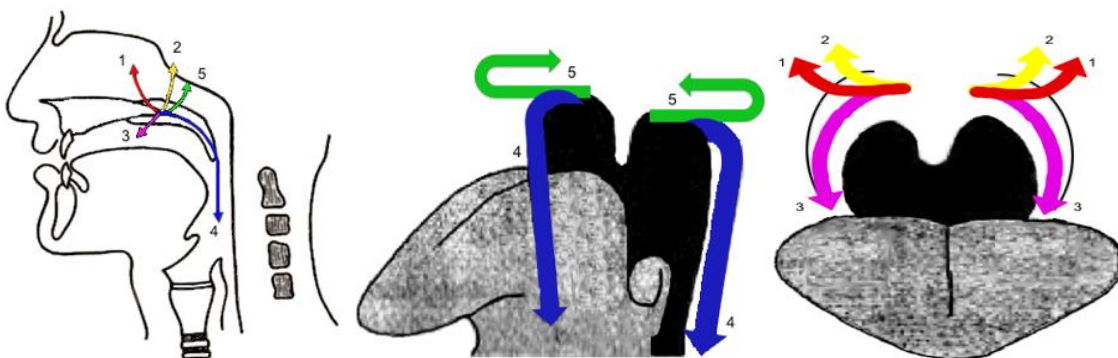


Figure 22: Soft Palate muscles: (1) Tensor Veli Palatini. (2) Levator Veli Palatini. (3) Palatoglossus. (4) Musculus Uvulae. (5) Palatopharyngeus. Adapted from Kent's illustration (1997, p. 190).

When the tensor palatini muscle stretches, it helps elevate the soft palate, opening the space between the oral cavity and the nasopharynx. The superior pharyngeal constrictor muscle and the palatal tensor relax, and the force of gravity aids the opening movement. (HARDCASTLE, 1976, p. 123).

The palatoglossus muscle connects the soft palate (also known as the velum) to the back part of the tongue of the thyroid cartilage and the lateral pharyngeal wall (HARDCASTLE, 1976, p. 124). Specifically, it originates from the soft palate and extends downward to attach to the side of the tongue near its base.

When the palatoglossus muscle contracts, it helps to lower the soft palate, allowing the nasopharynx to close off from the oral cavity, preventing food and liquids from entering the nasal passages during swallowing. This action directs the food and liquids into the esophagus and prevents aspiration.

During speech production, the palatoglossus muscle also plays a role in certain sounds. It assists in the movement of the soft palate, helping to produce specific speech sounds that require coordination between the soft palate and the tongue.

Although all individuals have the same set of muscles, they do not operate in the same way to produce achieve a specific set of gestures during voice production. The contraction of the palatoglossus muscle can be seen in Figure 23.

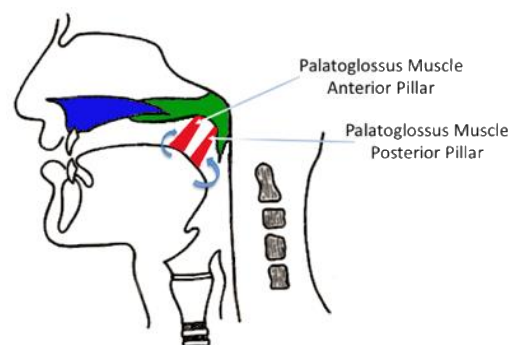


Figure 23: Lateral view of the palatoglossal arches, anterior and posterior pillars. Arrows indicate muscle movement. The picture was adapted from Kent's illustration (1997, p.191).

The palatoglossus is a muscle with fibers that are attached to the pharynx and connect the inferior surface of the velum to the edges of the tongue. It stretches down from the lower surface of the soft palate, divides into two parts, and connects to the tongue (HARDCASTLE, 1976, p. 124). This muscle, as mentioned previously, can lower the velum or raise the back of the tongue, thus helping to produce the velar consonants.

## CHAPTER 2: AERODYNAMIC ANALYSIS

---

### Résumé (Summation)

This chapter aims to demonstrate the aerodynamic aspects of the oral/nasal front/back diphthongization production in the *Paulistano* dialect. Thus, we analyzed and quantified the variation in airflow volume output resulting between the oral-nasal cavity coupling and the tongue movement. We analyzed the acoustic and aerodynamic consequences of velum lowering during the production of the front nasal vowel, the front oral diphthong, the oral back diphthong, and the back-nasal diphthongs. We demonstrate the acoustic and aerodynamic features, as well as the differences that make these elements distinct.

### 3. Introduction: Aerodynamic Analysis

In phonetics and phonology, aerodynamics is a method of analysis that assists in understanding the airflow during speech production. Aerodynamics provides valuable data that complements acoustic evidence in speech analysis.

Acoustic analysis alone can provide information about the sound properties and frequencies of speech sounds, but it may not directly reveal the articulatory actions involved in producing those sounds. On the other hand, aerodynamic analysis helps bridge this gap by shedding light on the airflow patterns and air pressures that occur during speech. This type of data makes it possible to infer which articulatory parameters are active during speech signals.

Speech analysis based on aerodynamic factors complements the acoustic evidence. The acoustic study becomes more complex when it comes to nasality because of the relationship between the oral and nasal tracts coupling and therefore the movement of the tongue.



The resonances transfer from the oral cavity to the nasal cavity and make it hard to see the oral formants in the spectrogram because of the nasal anti-resonances. Understanding the characteristics of nasalization and its gradual process of spreading in the acoustic properties is made easier by analyzing the airflow variation.

Through its volume and shape variation, we can use the aerodynamic data have provide valuable information about the volume and shape variations in the oral and nasal cavities during speech production. By examining these variations, we can gain insights into how the airflow behaves and influences the formants (resonant frequencies) of speech sounds. This type of analysis allows us to determine when there is a transfer from oral resonance to the nasal cavity.

So, using aerodynamic and acoustic data, it is possible to infer when and where the segment becomes more nasal. In our data, we found a correlation between the nasal murmur at the end of the articulation and the high peak of air in the nasal airflow.

In aerodynamics, fluid mechanics studies the behavior of fluids at rest, in motion, and the other forces acting on these systems. Flow can be understood as a vector field that crosses a surface. Speed variation is demonstrated mathematically via the continuity equation.

The continuity equation is a fundamental principle in fluid dynamics that expresses the conservation of mass for an incompressible fluid flowing in a closed system. It states that the rate of mass entering a given volume of fluid must be equal to the rate of mass leaving that volume. In mathematical terms, the continuity equation is represented as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

where:

$\frac{\partial \rho}{\partial t}$  is the rate of change of fluid density concerning time,

$\rho$  is the fluid density,

$\nabla$  is the del operator (vector differential operator),

$\cdot$  represents the dot product,

$\rho\mathbf{v}$  is the mass flux or mass flow rate, which is the product of fluid density ( $\rho$ ) and its velocity vector ( $\mathbf{v}$ ). The volume of a flow is measured in  $\text{m}^3/\text{s}$  (cubic meters per second).

In simpler terms, the continuity equation ensures that the amount of fluid entering a particular region must be equal to the amount of fluid leaving that region, considering there are no sources or sinks of mass within the system. This equation is used to study and analyze fluid flow behavior in various applications, including aerodynamics, hydrodynamics, and other branches of fluid mechanics.

If the flow passes through an area that is partially closed, this may increase its velocity. If the exit area for the flow is small, the rate will be higher and vice versa. Flow velocity can increase due to a reduction in section area and not because pressure increases (RESNICK, HALLIDAY & WALKER, 2009).

The subsystems involved produce the voice and airflow: The lungs (respiratory system) are the source of pressure and cause slight variations in the air pressure, while the glottis (glottal region) behaves like a valve by changing the resistance attributed to the airflow.

Impedance acts by creating semi-periodic waves, also known as air pulses. The vocal tract serves as a resonant filter for the sound generated by the vocal cords. As the vocal cords vibrate, they produce a complex sound waveform with various harmonics. While traveling through the vocal tract, certain frequencies are amplified (resonated), while others are dampened. The vocal tract, shaped and lengthened by

articulator movements such as the tongue and lips, determines specific resonant frequencies known as formants in speech (FANT, 1960; FLANAGAN, 1972).

The pneumotachograph is used to measure the airflow and pressure in speech production. The name for the speech airflow meter comes from the Greek: *pneumo* means 'air' and *tacho* means 'speed' (GICK, 2014). This tool measures the airflow in speech. It should be considered that production depends on the subsystems described in the introduction:

a) The respiratory system represents the lungs; it is the source of pressure and produces a slight variation in the tracheal air pressure.

b) The glottal region is where the glottis behaves similarly to a valve by changing the airflow resistance.

Many phonetic and phonology studies are using aerodynamic data to understand the variation in sound quality, as the equipment is affordable and enables the collection of data concerning variations in the airflow volume.

It is also possible to collect an extensive dataset, as the method is non-invasive. The results have a clear resolution that helps to extract the spatiotemporal parameters from oral and nasal productions.

However, there are some difficulties with the calibration process because this process takes some time; the calibration voltage is individual and should be done in each data recording. Another issue is that mask resonance can alter frequencies above 1000 Hz (BADIN et al., 1990).

Analyzing nasalization via aerodynamic volume variation is interesting because it allows us to indirectly infer certain coarticulation aspects in speech production. Coarticulation refers to the phenomenon where the articulation of one speech sound is influenced by the preceding or following speech sounds. In the context of

nasalization, the degree of velopharyngeal movement, which controls the opening and closing of the velopharyngeal port (the passage between the oral and nasal cavities), can indirectly indicate how speech sounds are influenced by the nasal environment.

For example, during speech production, if there is a gradual increase in nasal airflow (nasalization) while articulating certain vowels or consonants, it suggests that the velopharyngeal port is gradually opening, allowing air to flow into the nasal cavity. This coarticulation effect can modify the resonance properties of neighboring speech sounds, impacting their acoustic characteristics.

However, it's important to note that nasal airflow variation is not solely dependent on the velopharyngeal movement. Other factors, such as changes in the geometry of the oral tract (e.g., the shape and position of the tongue and lips), can also influence nasal airflow (KRAKOW & HUFFMAN, 1993). Therefore, analyzing aerodynamic volume variation helps in understanding how different articulatory adjustments work together to produce the unique features of nasalized speech sounds and how coarticulation plays a role in shaping speech production.

Aerodynamics studies the evolution of body movement, air integration with force, and solid objects. This term comes from the Greek *aer* (air) and *dunamikós* (mighty, strong, potent). In general, aerodynamics supports variations in sound intensity, volume, airflow (gas fluid), turbulence, and intraoral pressure studies in speech analysis.

The fundamental concepts of aerodynamics have been used for thousands of years, dating back to Ancient Greece. The formalization of this area began in the seventeenth century. This field was focused on developing aircraft and motion objects heavier than air (LILIENTHAL, 1891).

In 1721, Isaac Newton developed the air resistance theory. In 1738, Daniel Bernoulli created the Bernoulli principle, which describes the relationship between pressure, density and flow velocity. In 1757, Leonhard Euler described Euler's equations, which are applied to inviscid flow.

In the nineteenth century, Navier-Stokes described the most general equations governing fluid flow. In 1799, George Cayley identified the four aerodynamic forces of flight, namely weight, lift, drag vector, and thrust.

In addition, the first wind tunnel was developed by Francis Herbert Wenham in 1871. This development helped to validate more precise measurements of aerodynamic forces and supported the development of aviation and automobiles. Over time, computational fluid dynamics, supersonic and hypersonic flow were also developed.

### **3.1 Objectives**

This chapter aims to analyze the general characteristics of nasal diphthongization and their phonetic processes via the oral/nasal airflow and duration measurements. Accordingly, we aim to answer some crucial questions about the nature of nasal diphthongization and, if possible, to infer the mechanisms that contribute to the nasal signature during speech production.

### **3.2 Research Questions**

The research questions presented in this section 3.2 aim to analyze the general characteristics of nasal diphthongization and its phonetic processes. These questions explore how nasalization affects diphthong sounds and how aerodynamic features can be used to infer the articulatory targets of nasalization.

They also inquire about the differences between front and back-nasal diphthongs, the relationship between oral and nasal production characteristics, and whether nasalization leads to a change in sound quality.

On the other hand, the research questions presented in item 1.2 focus on aspects related to the role of tongue adjustment, tongue movement, nasal airflow variation, motor synchronization, velum displacement, and velarization or palatalization during speech production.

In summary, while both sets of questions are related to speech production and phonetic processes, the questions on item 1.2 focus on broader aspects of speech production, including nasality and tongue adjustment, whereas the research questions in this section are specifically centered on nasal diphthongization and its aerodynamics characteristics.

These questions explore the articulatory and aerodynamic aspects of speech production, with a specific emphasis on the mechanisms and processes involved in producing nasality and nasal sounds.

- Which aerodynamic features help to infer the articulatory targets of nasalization?
- Is there evidence that aerodynamic targets play a significant role in phonology?
- How are the productions of front and back-nasal diphthongs different?
- How are the features of oral and nasal diphthongs related?
- Is nasality the result of a change in sound quality?
- Is tongue displacement a result of nasality?
- What are the effects of the synchronization between velum and tongue movement?

- Will all the tokens always have a nasal airflow, a high peak, and a nasal appendix?
- Is nasal diphthongization a pattern?

Our aerodynamic data plays a crucial role in understanding various aspects of nasalization and nasal diphthongization. It provides valuable information about airflow patterns and pressure changes in the nasal cavity during nasalization, enabling the inference of specific articulatory gestures involved in producing nasal sounds.

Additionally, aerodynamic data unveils the influence of articulatory configurations, such as velum movement, on the acoustic properties of nasal sounds, thus contributing to the phonological patterns in a language.

In the case of nasal diphthongs, aerodynamic measurements help uncover distinctions in airflow and resonance between front and back-nasal diphthongs, shedding light on their unique articulatory and acoustic characteristics.

By comparing aerodynamic data from oral and nasal diphthongs, we gain insights into how the coupling of oral and nasal cavities affects the acoustic properties and perceived quality of diphthong sounds.

Moreover, aerodynamic data aids in correlating changes in nasality with modifications in resonance characteristics, providing a deeper understanding of how nasality impacts overall sound quality.

Additionally, it reveals possible relationships between nasality and specific tongue movements, contributing to the knowledge of articulatory aspects related to nasality.

The timing and coordination of velum and tongue movements elucidated through aerodynamic analysis, are crucial factors influencing the production of nasal sounds and their acoustic properties.

Furthermore, aerodynamic measurements allow us to determine the consistency of nasal airflow patterns, peak amplitudes, and other relevant characteristics across different tokens, providing valuable quantitative data for analysis.

Finally, aerodynamic data assists in identifying consistent patterns in nasal diphthongization, offering insights into typical articulatory configurations across speakers or linguistic contexts.

### **3.3 Theoretical Framework: Overview of Aerodynamics**

Nasality is a complex phenomenon, even for a non-pathogenic speaker, to produce because oral-nasal coupling changes the sound quality and sound perception. The airflow passes through two resonant cavities, the mouth, and the nose, simultaneously altering the resonance and propagation of sound waves. The aerodynamic device records air volume, which indicates pressure variations in the oral and nasal cavities, the glottis, or the subglottal cavity.

We can infer the soft palate mechanism of controlling (lowering and rising movements) and variations in the nasal airflow volume via aerodynamic analysis. The investigation of the airflow can indicate how the articulators behave during phonation.

The aerodynamic analysis provides valuable information about airflow patterns and volume variations in the nasal cavity during speech production. By examining the nasal airflow data, we can indirectly infer the mechanisms involved in controlling the soft palate's movements, including lowering and rising actions.

The investigation of airflow patterns can offer insights into how the articulators behave during phonation. For example, by analyzing the airflow data in coordination with other measurements, we can deduce the timing and coordination between the velum (soft palate) and the tongue movements. This information allows them to



understand how the articulators work together to produce specific speech sounds and how they influence the acoustic characteristics of these sounds.

Additionally, aerodynamic analysis helps identify variations in nasal airflow volume, which can provide clues about the degree of nasality in speech. Changes in nasal airflow volume can indicate differences in the size and shape of the nasal cavity during speech production, reflecting specific articulatory configurations. We believe that also can indicate oral constriction.

Nasal airflow data compare normal and pathological speech production conditions, and assist in the diagnosis of hyper-nasality, hypo-nasality, and nasal resonance insufficiency (WARREN et al., 1993; KRAKOW & HUFFMAN, 1993, p. 4).

The relationship of velum movement, airflow, and nasality described in the literature was based more on assumptions than on objective experimental evidence, mainly because the investigation of velic action during speech presents considerable difficulties (KRAKOW & HUFFMAN, 1993).

Despite all the technological advances, it is still difficult to collect data about the displacement of the velopharyngeal movement because many methods (EMMA, Nasal Endoscopy, and Magnetic Resonance Imaging) are invasive (author personal position) and require a specific place in which to record the data session, as well as professionals such as doctors or engineers. Some experimental methods, such as ultrasound and acoustic data, are not invasive and are accessible, but they cannot generate visual data showing velum movement or airflow variation.

A pneumotachograph is a specific device that is used most frequently in the field to measure air pressure and airflow; it involves a mask containing some sensors being attached to the speaker's face to register the airflow data.

Oral and nasal airflow data make it possible to draw inferences about tongue movement, lip protrusion, and velopharyngeal opening, as mentioned previously. This instrument provides an accessible and non-invasive method for large-scale data collection. The methodology requires strict control and individual calibration.

Oral and nasal airflow data allow us to draw inferences about mechanical aspects of articulatory movement. It is possible to infer from the volume variation in the oral and nasal airflow.

These variations reflect changes in the velopharyngeal opening and the oronasal coupling. The soft palate plays a crucial role as a lowering and rising control, affecting nasal airflow.

Positive or negative nasal airflow indicates that nasality is not a mandatory factor in nasal production or for the perception of this sound quality. The presence or absence of nasality in nasal sounds does not solely depend on the soft palate's position. This finding suggests that nasal sounds can exhibit a degree of variation in nasality, known as nasality gradience and the listeners may accept it (LAVÉRE, 1975).

Some physiological factors may influence the recording of nasal airflows, and the results of nasalization studies in aerodynamic experiments may have contradictory findings (CAGLIARI, 1970). According to Cagliari, three issues are controversial:

(a) The velum-pharyngealization of the velum that may or may not be completely closed. Thus, there may be a small amount of nasal airflow during all oral sounds.

(b) The nasalization of sounds can be perceived as nasalized even in the absence of measured or observable nasal airflow, or an oral vowel with low nasal airflow and a nasal sound that may also occur without velic lowering or nasal airflow escaping (VAN RIPER 1958; CAGLIARI, 1977).

(c) The amount of airflow can generate differences in the degrees of nasality (CAGLIARI, 1970).

Therefore, some external factors may cause changes in the measurement of the airflow, which would affect the interpretation of nasal airflow as changes in airflow resistance, and global changes in the airflow signal could change the characteristics of the data.

Krakow & Huffman (1993) propose that the movement of the velum (soft palate) influences the aerodynamic conditions of nasal airflow. However, it is important to note that this does not imply a straightforward relationship between the escape of nasal airflow and the opening of the velopharyngeal sphincter.

They suggest that while the movement of the velum plays a role in controlling nasal airflow, it does not mean that every instance of nasal airflow corresponds directly to the opening of the velopharyngeal sphincter. The relationship between these two factors is more complex and can be influenced by additional coarticulatory elements. The absence of a direct relationship between nasal airflow and the opening of the velopharyngeal sphincter means that other factors may also come into play. These can include the specific speech sounds being produced, individual variations in articulatory behavior, and other aspects of the speech production process.

We hypothesize that, if there is an escape of air through the nose, this implies that there is some opening of the velopharynx; if there is no airflow output, this suggests that the velum is closed or partially closed. If the nasal airflow remains constant, this may reflect that the velum and the oral articulators are in the default position.

According to Krakow & Huffman (1993), vibrations in the nose are usually more robust when the airflow is passing through the nose. These vibrations in the velum

tissue are related to the resonance that stems from the oral cavity. There are indications that vibration transfer occurs in sounds produced with more significant oral constriction, such as high vowels.

The velum experiences some vibration during all speech production, even if the velopharyngeal sphincter is closed. The nasal cavity retains the airstream during velopharyngeal closure or via the internal displacement of the lateral pharyngeal walls, even when the auditory tube is opened. In both cases, the amount of air released through the nostrils will be minimal.

According to Lubker & Moll (1965), there is a slight change in the size of the nasal cavity when speech begins, and the total volume of air exhaled through the nostrils is approximately 0.0042 liters.

Consequently, when the palatine velum is closed, it can retain airflow within the nasal cavity, allowing for both ingressive (inward) and egressive (outward) airflow. Additionally, the movement of the velum may generate a subtle noise, as observed in their study (LUBKER & MOLL, 1965).

Smith (1972) also demonstrated a small amount of incoming and egressive nasal airflow during occlusives produced with the soft palate closed. According to Cagliari (1977), the volume of nasal air was related to the threshold of perception.

Speakers will consider the threshold of nasalization<sup>50</sup> to determine whether a sound qualifies as being nasal sound (VAN RIPER & IRWIN, 1958; CAGLIARI, 1977) if the output of the nasal airflow is partial.

The elevation of the palate is also vital for the production of speech. Nevertheless, a small nasal airflow output during speech production does not

---

<sup>50</sup> The threshold of nasalization refers to a point or boundary at which the perception of nasalization changes. It represents the level of nasal airflow or nasal resonance at which a sound is perceived as having nasal qualities or being nasalized.

necessarily indicate nasalization. Instead, specific compression of the airflow inside the nasal cavity is required for nasalization to occur.

Physiological factors also can contribute to the production of nasality. According to Rousselot (1924, p. 526; CAGLIARI, 1977, p. 246), the vibration in the nose tissues may modify oral resonance during the production of consonants and vowels.

Concerning gender aspects and nasalization, it is observed that there may be differences between men and women in terms of the degree of nasality in their speech. Some studies have found that men tend to maintain a slightly open velum even during oral sounds, resulting in a greater degree of nasality in their speech compared to women. This difference in velar behavior contributes to variations in the size of the palatopharyngeal opening, which influences the degree of nasality perceived in the speech of individuals.

Warren (1967, p. 185) identified that sounds with a velic opening of 10 mm<sup>2</sup> could be perceived with an oral occlusion, and Greene (1964, p. 58) showed the same with a space of 8 mm<sup>2</sup>. However, some speakers judged the sound to be oral production only after the complete closure of the velopharyngeal sphincter.

The velopharyngeal cross-section also contributes to the proportional changes in nasal airflow. Total or partial constriction in the oral cavity increases or decreases the amount of airflow resistance.

### **3.3.1 Methodology**

The main concern of this section is the design of the aerodynamic analysis to describe nasal diphthongization based on variations in the nasal and oral airflows. Thus, we will describe how the design was conceived and implemented in this section.

In this section, we describe the method for comparing the speech production of the nasal diphthong /ãĩ/ with /aw/, and the nasal vowel /ẽ/ (the nasal allophone phonetically diphthongized as [ẽj]) with the oral diphthong /ej/.

We present data obtained from six native speakers of BP from the city of São Paulo. The parameters extracted for analysis were the volumetric variations of oral and nasal airflows and the acoustic duration of the target segment.

In the item 3.3.7, we determined the segmentation of the target segment using acoustic criteria. We looked for acoustics cues in changes in formant frequencies, analyzing the spectrogram.

### **3.3.2 Material**

Some devices are available to collect aerodynamic data in the field, but to develop this study of aerodynamics, we chose the portable EVA2 workstation to conduct our study. EVA is a pneumotachograph for assisted vocal evaluation (Evaluation Vocal Assistée) (TESTON, 1995). This device allows for the study of oral and nasal airflows during speech production.

The portable EVA2 is specially designed by SqLab ([www.sqlab.fr](http://www.sqlab.fr)) to study speech production disorders. Some of its main functions are to investigate and to assist in the diagnosis of velum disorders.

EVA2 allows for the recording of aerodynamic data, acoustic data, and electroglottograph (EGG)<sup>51</sup> data simultaneously. It is possible to analyze parameters

---

<sup>51</sup> This device records information about the movement of the vocal folds opening and closing and the periodicity of vocal folds vibration.

of speech production such as airflow volume, airflow rates, and intraoral pressure; the acoustics parameters are  $f_0$ , pitch and voice intensity.

The equipment is portable and is compatible with Windows® microcomputer devices; it records the acoustic signal and the aerodynamic data simultaneously. It is possible to analyze the correlation between the speech signal and aerodynamic features (DEMOLIN et al, 2006).

The input is recorded in three different channels—one channel for acoustic data, and another two channels for nasal airflow and oral airflow.

### **3.3.3 Equipment**

We chose to collect the data using the portable EVA2 device because the resulting data enables the inference of the articulators' movements during speech production. This device makes it possible to integrate the captured acoustic, aerodynamic, and EGG signals.

An integrated transducer in different channels records the oral airflow (AFo) and nasal airflow (AFn). A flexible silicone mouth mask records the AFo. The mask is attached to the speaker's face. The one that we used covers the entire mouth region because it is placed against the speaker's face to avoid air leaks. Thus, there is no interference in the articulatory movements.

As the mask does not cover the nostrils, to collect the AFn, two silicone pipes were placed inside the nostrils, and allowed for the free outward nasal airflow. Each silicone tube was 10 cm long and 0.5 cm in diameter. In front of the mask, a microphone was placed 8 cm away from the mouth to allow for the recording of the acoustic signal. The nasal sensor is located under the oral sensor in a vertical position, following the natural flow of air from the nose, as shown in figure 24.

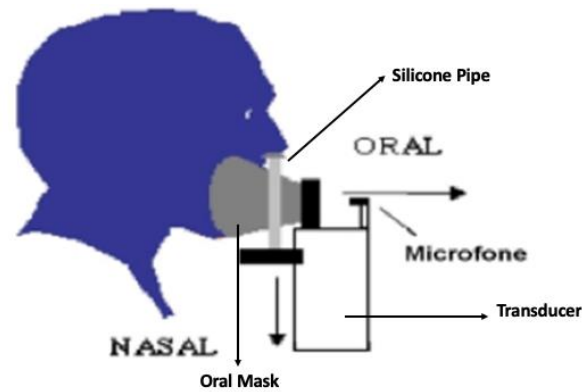


Figure 24: Sensor position for the EVA data recording. Image is taken from Ghio & Teston (2004).

The silicone pipe and the mask were adapted to the speaker's face and nostrils; they are connected to the transducer, which was connected to a microcomputer to record the data digitally. The computer controlled the instrumentation and the recording chain, while the transducer had an electronic manometer<sup>52</sup> to allow for calibration.

### 3.3.4 Corpus

The corpus contained words that undergo nasal diphthongization in an offset position, that is, the word-final position where nasal diphthongization occurs. We listed a group of words from Portuguese that had oral and nasal contrasts. Both the back-nasal diphthong and the front nasal vowel are phonetically diphthongized and are produced with a nasal appendix: [ãw̃] and [ẽj̃] (SOUSA, 1994; SEARA, 2000; JESUS, 2002; SHOSTED, 2006; MEDEIROS & DEMOLIN., 2006).

---

<sup>52</sup> A manometer is a device used for measuring the pressure of a gas or liquid. It typically consists of a U-shaped tube filled with a liquid, such as mercury or water, and the pressure difference between the two ends of the tube causes the liquid to rise or fall. The height of the liquid column indicates the pressure being measured. Manometers are commonly used in various fields, including physics, engineering, and fluid mechanics, to measure and monitor pressure levels.



The presence of the nasal appendix suggests that nasalization accompanies the diphthongs, possibly as a phonetic feature related to the nasal quality of the sounds, as a nasal murmur spreading during the production of the diphthongs [ãw̃] and [ẽj̃].

We also followed some observations by Krakow & Huffman (1993), who emphasized the control of the sound sequences and the consonantal boundary context. According to the cited authors, ‘velic movements in sequences containing a nasal consonant (e.g., bean) must be compared with movements in minimally conservative sequences containing only oral segments (e.g., bead), to separate the contextual effects of the nasal consonants from the intrinsic velic positions for the adjacent string’ (KRAKOW & HUFFMAN, 1993, p. 3).

We considered compiling a *corpus* that only contained oral and nasal minimal pairs<sup>53</sup>. However, this was deemed impractical because there are insufficient lexical items in the language to compile a workable *corpus*, comparing front nasal vowels and front oral diphthongs. There are only minimal pairs in the back oral and nasal diphthongs in Portuguese because these features are this phonological diphthong.

In total, we recorded twenty words containing the target segment within the tonic syllable. For the back diphthongs, we chose five oral and nasal monosyllabic words. As mentioned, all were minimal pairs in this case. Below is a list of words along with their corresponding phonetic transcriptions, as shown in Table 2.

We did not find minimal pairs to the oral counterparts for the front nasal vowel as we expected at the beginning of this experiment. We selected a group of five oral monosyllables, three nasal monosyllables, and two nasal disyllables.

---

<sup>53</sup> According to Krakow & Huffman (1993, p. 34) ‘it is appropriate to compare absolute nasal airflow levels on one segment in a minimally contrastive oral and nasal context. For example, airflow during the vowel of bid could be compared to that during mid and bin in English’.

The number of syllables in the word does not change the nasal diphthongization. According to Krakow & Huffman (1993), the consonantal variations in the onset do not cause any complications in the airflow analysis, for example, “*tem*” [tẽ] and “*sem*” [sẽ]. However, significant aerodynamic differences are related to the openness of the vowel sounds.

The differences in the absolute nasal airflow were expected to be found when comparing front and back vocalic production because the degree of oral constriction changes the airflow. The consonantal boundary can also influence the spread of the nasal appendix.

Certain boundaries or transitions between sounds can indeed have an impact on the articulatory or acoustic properties of the subsequent segments. To minimize these boundary influences, we carefully select words that have specific syllabic contexts. In this phonetic phenomenon, we focus on words that follow the syllabic context of CVG\_\_\_# (where V represents a vowel and G represents a glide) and C $\tilde{V}$  $\tilde{G}$ \_\_\_#<sup>54</sup> (where  $\tilde{V}$  represents a nasal vowel and  $\tilde{G}$  represents a nasal glide). By selecting words with these specific syllabic contexts, we aim to isolate and study the effects of nasalization on the target segments.

There was a variation in the occlusive consonants [p, b, t, d, k], the fricatives [s, f, h], the nasal consonant [m] and the approximate [l] at the onset of the target word. According to Krakow & Huffman (1993), the only lexical item that differs based on the context is the cluster [fr], but it does not cause a problem in the analysis.

The words that constituted the corpus are presented in Tables 2 and 3 and cover the front nasal vowel, the front vowel, the back oral and nasal diphthongs. The stimuli

---

<sup>54</sup> We used a very similar *corpus* in all the different experimental designs to maintain the cohesion in the thesis.

were presented in BP, and English glosses have been provided below the Portuguese words:

Table 2: Set of words with front oral diphthongs and front nasal vowels.

Oral diphthongs		Nasal diphthongs	
Stimulus	Gloss	Stimulus	Gloss
[paw]	wood	[pãw̃]	bread
[maw]	bad	[mãw̃]	hand
[taw]	these	[tãw̃]	so
[saw]	salt	[sãw̃]	sane
[kaw]	lime	[kãw̃]	dog

Table 3: Set of words with front oral diphthongs and front nasal vowels.

Oral diphthongs		Nasal Vowels	
Stimulus	Gloss	Stimulus	Gloss
[dej]	<i>give (1<sup>st</sup> SG)</i>	[tẽj̃]	<i>have (3<sup>rd</sup> SG)</i>
[sej]	<i>know (1<sup>st</sup> SG)</i>	[sẽj̃]	<i>without</i>
[nej]	<i>Ney (nickname)</i>	[nẽj̃]	<i>neither</i>
[frej]	<i>friar</i>	[hẽj̃]	<i>REM</i>
[lej]	<i>law</i>	[bẽj̃]	<i>good</i>

We inserted each word into the carrier sentences ‘**Digo** \_\_\_\_ **todo dia**’ (I say \_\_\_\_ every day) [dʒigu \_\_\_\_ kade dʒiɛ] and ‘**Digo** \_\_\_\_ **cada dia**’ (I say \_\_\_\_ each day) [dʒigu \_\_\_\_ todʊ dʒiɛ]. We chose the syntactic structure of the carrier sentence to avoid prosodic changes during the recording sessions.

We asked the speakers to use a declarative intonation and a standard speech rate when reading the sentences. We aimed to place the phrasal accent on the keyword.

The sentence ‘Digo \_\_\_\_ todo dia’ (I say \_\_\_\_ every day) with the stop /t/ in the syllable boundary position is a control sentence, to avoid the assimilation of the articulation point between the nasal glide and the voiceless consonant.

In the context of nasal place assimilation, the velarization of the nasal appendix can be considered a secondary assimilation process. Specifically, when the consonant /k/ is involved.

We believe that the velarization or palatalization of the nasal sound is not directly influenced by the adjacent sound, but rather results from other articulatory processes, such as the assimilation of the articulatory place of the preceding vowel. It is important to investigate and demonstrate that the velarization or palatalization is a consequence of these additional articulatory factors.

The carrier sentences with the phrases for the data collection were written on cards and were presented randomized to each participant. We presented the sentences in an orthographic form for ease of reading.

All the participants repeated the sentences three times, resulting in 360 tokens (20 diphthongs x 6 speakers x 3 repetitions) for the /k/ boundary context and 120 tokens for the /t/ boundary context, and used these data in the descriptive and quantitative analyses. Although some speakers produced a short pause after the emission of the keyword, we did not control for intensity because it was not the aim of our analysis.

### **3.3.5 Speakers**

As mentioned previously, we recorded six speakers<sup>55</sup> for this study. All the informants were BP speakers who were born in São Paulo. To avoid dialectal or gender variation, we selected two groups of native speakers of the *Paulistano* dialect.

The male group consisted of three speakers with an average age of 45, while the female group consisted of three speakers with an average age of 35; all the

---

<sup>55</sup> There are more participants in this experiment than in the other experiments. It was easier to collect aerodynamic data than it was EMMA data. We used a very similar corpus for consistency.

speakers were healthy, and they did not report any speech pathologies. All the speakers also had university degrees.

The following table shows the acronyms used to refer to the participants. The initials were used to identify each speaker when necessary.

Table 4: Speakers identified via their initials. The table is divided according to gender.

Speaker (Age)	
Male	Female
BH (41 years)	PC (41 years)
LZ (44 years)	LC (40 years)
OD (50 years)	RD (25 years)

### 3.3.6 Data Processing

In order to extract coherent data from the oral and nasal airflow rates, we needed to consider various aspects at the time of the data collection. For example, we needed to control the environment to avoid the acoustic room being subject to fluctuations in temperature or humidity. We also needed to ensure that the device and the use thereof respected ethical rules and local regulations.

It was essential to ensure that the airstream (respiration) was bidirectional to avoid distortion or a rapid variation in the volume of the airflow. The volume sensor must be small; otherwise, the channel for the oral airflow could be contaminated by resistance from the channel with nasal airflow or vice versa.

It was also important to ensure that the sensors did not obstruct the articulatory movements. The masks and the silicone tubes were adapted to the speakers' facial structures to preserve all the actions. Some articulatory aspects also needed to be controlled, such as lip turbulence that could interfere with the aerodynamic measurements.

EVA2 is a pneumotachograph system with a stable sound pressure transducer that improves the acquisition of articulatory data. The device is sufficiently small to register airflow variation in the order of  $1.0 \text{ cm}^3/\text{s}$ .

The resistance was equivalent to 1% of the intraoral pressure of a non-pathologic speaker. The resistance level was set to  $10 \text{ dm}^3/\text{s}$  for breathing, corresponding to 60 dB, to decrease the non-linearity of the aerodynamic turbulence during speech production.

The speakers were silent and held their breath while the device was being calibrated. The EVA2 was calibrated in the minimum position (zero  $\text{dm}^3/\text{s}$ ) on each repetition. For the aerodynamic analysis, the representation of the air volume resulted in a graph that ranged from  $-100 \text{ dm}^3/\text{s}$  to  $300 \text{ dm}^3/\text{s}$ .

For acoustic analysis, the frequency range was 8 kHz, but we preferred to keep it at 5 kHz with a 5 ms hamming window, which generated a wideband spectrogram and enabled the visualization of the formants.

We used two software programmed to analyze and extract the data. The Signal Explorer<sup>56</sup> allows for the visualization of a graph ranging from  $-60 \text{ dB}$  to  $60 \text{ dB}$ . We used a 25 ms window. Linear Predictive Coding (LPC) has 48 analysis points, and the Fast Fourier Transform (FFT) has 512 points.

These values allowed the generation of spectrograms and graphs that discriminated between the analyses of the airflow and the acoustic signal because they filtered the non-significant interferences from the signal. Phonedit<sup>57</sup> is software that allows for data recording, processing, and analysis. It was possible to filter the

---

<sup>56</sup> Signal Explorer was developed by A. Soquet & D. Demolin.

<sup>57</sup> Phonedit MFC Application Company: LPL (Aix-provence, France) - [www.lpl-aix.fr](http://www.lpl-aix.fr)

airflow and pressure signals. Each data recording resulted in five file extensions: *f0*, *int*, *inf*, *oaf*, and *wa1*<sup>58</sup>.

### 3.3.7 Segmentation

The visual inspection<sup>59</sup> of the vocalic segments allowed for the segmentation of data. This approach facilitated the aerodynamic and acoustic descriptions of the target segments.

By analyzing the spectral energy distribution and waveform characteristics, we observed changes in the energy distribution and amplitude throughout the diphthong [aw] and [ej].

To extract the measurements of each phoneme, we needed to divide the vowel segment into two parts, the oral vowel and the oral glide, and its nasal correspondence into three parts, the nasal vowel, the nasal glide, and the nasal appendix.

By deconstructing phonetic phenomena into smaller, constituent parts, we can examine the interactions and relationships between these units. This approach allows us to study the intricate patterns and processes that occur during speech production and perception. By viewing complex phonetic events as sequences of elementary units, we gain a deeper understanding of the underlying mechanisms and structures involved in speech sounds. This analysis enables us to explore the complexities of phonetics in a systematic and detailed manner.

---

<sup>58</sup> Each extension refers to the parameters collected by EVA2. The fundamental frequency was stored in the *f0* file extension, the intensity in *.int*, nasal airflow in *.naf*, the oral airflow in *.oaf* and the waveform in *.wa1*.

<sup>59</sup> Visual inspection is an initial qualitative analysis method in which the main aim is the visualization and annotation of the characteristics of the object. This observation is the basis of the data analyses.

The acoustic segmentation was performed using the oscillogram, the spectrogram (Hz) by time (*ms*), and the spectrogram showing the frequency bands in which the energy concentration was present.

Diphthongs often exhibit formant transitions, where the frequency bands associated with the vowel gradually shift as the articulation transitions from one vowel quality to the glide. So, we used these formant transitions as criteria for segmenting the diphthong. Through figure 25, is possible to analyze the formant transition:

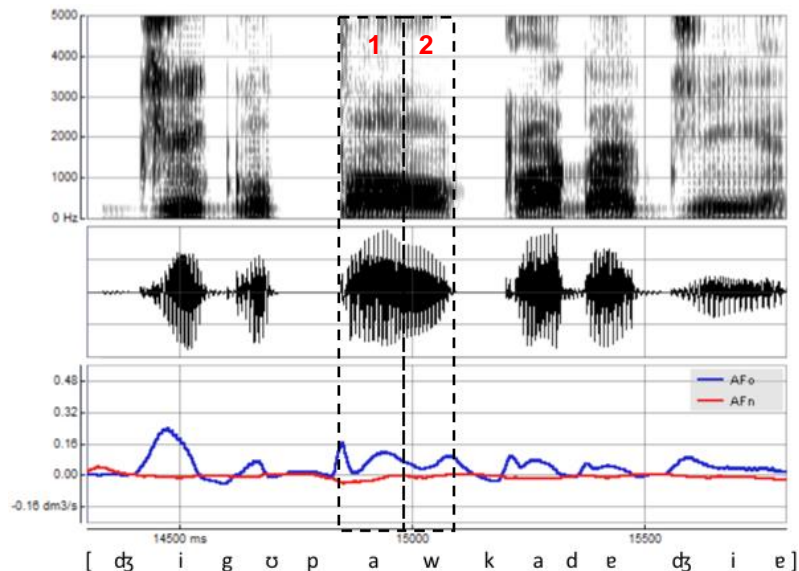


Figure 25: This figure displays a spectrogram (top), oscillogram (middle), and graphs depicting oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'pau'. The speaker associated with the data is identified as BH. The number one represents the oral vowel, while number two represents the oral glide.

The formant transition from [a] to [w] (1 to 2) is approximate and can vary depending on individual and phonetic factors. Initial [a]: F1 = 700 Hz, F2 = 1200 Hz; mid-transition: F1 gradually decreases to around 500 Hz, while F2 starts to rise to around 1500 Hz; and final [w]: F1 = 300 Hz, F2 = 1800 Hz.

In the example above (figure 25), we can note that the waveform for [aw] was laringalized. During the articulation of the diphthong, changes in the spectral energy distribution and waveform characteristics can be observed.



The initial [a] is characterized by a strong energy concentration in the lower frequency range (0-1000 Hz) and a relatively high amplitude, indicating a stronger intensity. As the articulation transitions from [a] to [w], there is a gradual shift in energy distribution towards higher frequencies, with increased energy in the higher frequency range (1000-4000 Hz).

The waveform exhibits gradual changes, reflecting the changing intensity of the diphthong during this transition. Finally, the final [w] is characterized by a strong energy concentration in the higher frequency range (2000-4000 Hz) and a relatively high amplitude, indicating a stronger intensity.

These observations of spectral energy distribution, transition, and waveform characteristics provide valuable cues for segmenting the different stages of the [aw] diphthong and understanding the articulatory transitions within it.

During the transition from [a] to [w], there is a decrease in amplitude in the waveform pattern, suggesting a shift in vocal tract articulation. This change corresponds to the moment when the tongue moves from the open back position for [a] to the rounded position for [w].

Figures 25 and 26 present examples of the acoustic segmentation for the acoustical duration extraction of the segments to the oral and nasal diphthongs, respectively.

In the spectrogram below, we present the back nasal diphthong. The acoustic segmentation related to formant transitions involved identifying the regions in the spectrogram where these shifts in formant frequencies occur. In the case of [ãw̃], the transition from the nasal vowel [ã] to the glide [w̃] can be visually observed as a gradual change in the formant structure in figure 26:

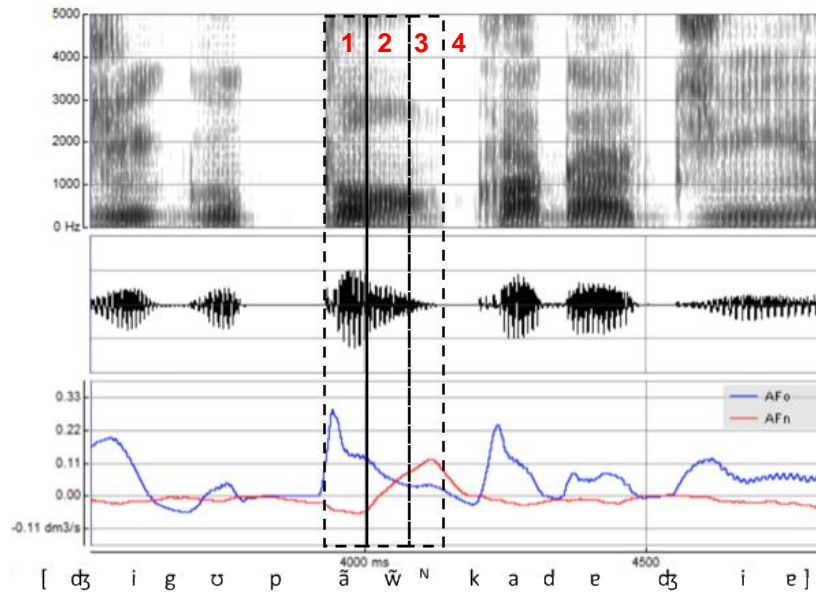


Figure 26: This figure illustrates a spectrogram (top), oscillogram (middle), and graphs representing oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'pão'. The data corresponds to the speaker identified as BH. In the figure, the number one represents the oral vowel, and the number two represents the oral glide.

During the initial [ã] (in 1), the spectral energy is predominantly concentrated in the lower frequency range (0-1000 Hz) with relatively less energy in higher frequencies. This is reflected in the spectrogram of the reinforcement of oral resonances because of nasal formants. The waveform exhibits a higher amplitude, indicating a stronger intensity.

As the articulation transitions from [ã] to [w̃] (1 to 2), there is a notable shift in the spectral energy distribution. The energy gradually shifts towards higher frequencies (1000-4000 Hz) while decreasing in the lower frequency range. This reflects in the spectrogram the dampening of oral resonances. The waveform displays a loss of intensity of the diphthong.

In the middle stage of the diphthong [w̃] (2 to 3), the spectral energy distribution shows a strong concentration in the higher frequency range (2000-4000 Hz) with relatively less energy in the lower frequencies until the nasal appendix (3). The waveform maintains a relatively low amplitude, indicating a sustained intensity of the nasal murmur.

The amplitude of the waveform shows a decrease at the transition point between [ã] and [ũ]. A sudden increase in amplitude suggests a shift in vocal tract configuration, potentially indicating the onset of the glide [ũ].

The waveform exhibits a clear shift in the shape, as the articulation transitions from the nasal vowel [ã] to the nasal glide [ũ]. This change in waveform shape reflects the alteration in vocal tract configuration during the diphthong.

The waveform exhibits a relative stability phase, that indicates that the vocal tract has reached a consistent configuration for [ã], such as the stable configuration of the nasal glide [ũ].

The first part of the diphthong [ã] had more energy than the nasal glide. The amplitude of the waves was more significant in the nasalized vowel than in the nasal glide. In the second portion, the sound wave corresponds to the nasal glide [ũ], which gradually lost power until this energy dissipated completely in the spectrogram. In the third portion, the nasal appendix is present at the end of the articulation.

During the articulation of the diphthong [ej], changes in the spectral energy, formant transition, distribution, and waveform characteristics provide insights into the transition stages within the diphthong.

In figure 27, is possible to note the formant transition from [e] to [j]: initial [e]: F1 = 400 Hz, F2 = 2000 Hz; mid-transition: F1 gradually rises to around 600 Hz, while F2 starts to decrease to around 1500 Hz; and final [j]: F1 = 800 Hz, F2 = 1000 Hz.

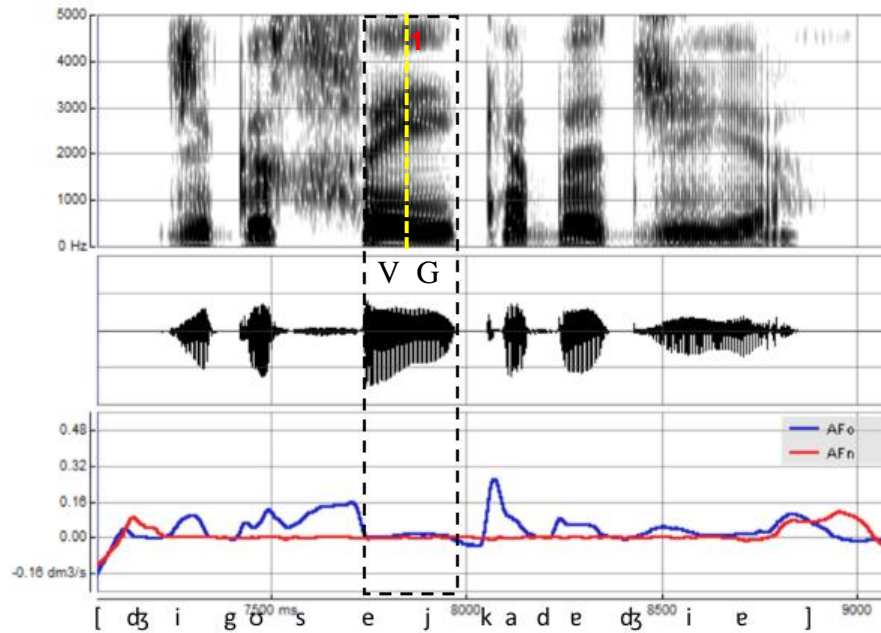


Figure 27: This figure displays a spectrogram (top), oscillogram (middle), and graphs depicting oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'sei'. The data corresponds to the speaker identified as LC. In the figure, the number one represents the oral vowel, and the number two represents the oral glide. The letters represent the vowel portion (V) and the glide (G) portion.

In the transition from [e] to [j], there is a decrease in amplitude and a distinct change in waveform pattern. This signifies the movement of the tongue from the mid-high front position for [e] to the high front position for [j], as well as the transition from a relatively steady-state vowel to the glide-like consonant. Additionally, the waveform exhibits a gradual decrease in amplitude or a smoother transition from the [j] to the following segment, indicating the glide-like articulation.

The initial [e] is characterized by a strong energy concentration in the higher frequency range (2000-4000 Hz) and a relatively high amplitude, indicating a stronger intensity. As the articulation progresses from [e] to [j], there is a gradual shift in energy distribution towards lower frequencies.

The energy in the higher frequency range decreases, while the energy in the lower frequency range (0-2000 Hz) increases. Finally, the final [j] is marked by a strong

energy concentration in the lower frequency range (0-2000 Hz) and a relatively high amplitude, indicating a stronger intensity.

In the spectrogram below, in figure 28, we present the front nasal diphthong [ẽ̃]. The acoustic segmentation related to formant transitions involved the transition from the nasal vowel [ẽ̃] to the glide [j] (1 to 2), which can be visually observed as a gradual change in the formant structure.

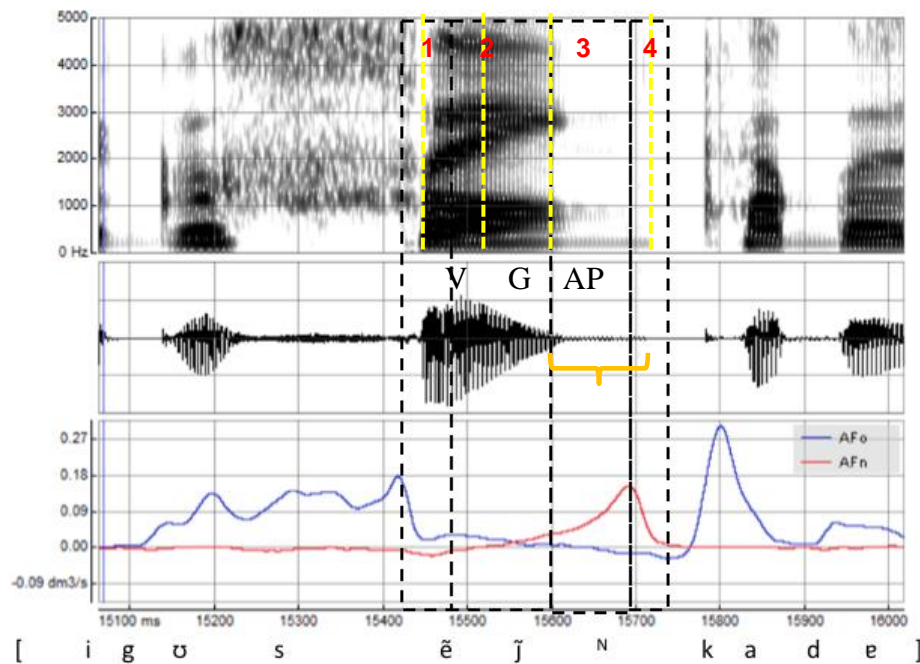


Figure 28: This figure presents a spectrogram (top), oscillogram (middle), and graphs illustrating oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'sem'. The data corresponds to the speaker identified as LC. In the figure, the number one represents the oral vowel, and the number two represents the oral glide. The letters represent the vowel portion (V), the glide (G) portion and the nasal appendix (AP)

During the initial [ẽ̃] (in 1), the spectral energy is predominantly concentrated in the higher frequency range (2000-4000 Hz) with relatively less energy in lower frequencies and the absence of resonances in the medium frequency zone. This is reflected in the spectrogram by the reinforcement of nasal resonances in low and high frequencies. The waveform exhibits a higher amplitude, indicating a stronger intensity.

As the articulation transitions from [ẽ̃] to [j] (1 to 2), there is a shift in the spectral energy distribution. The energy gradually shifts towards lower frequencies (0-1000

Hz) while decreasing in the higher frequency range, in a union of formants. The waveform displays a gradual loss of intensity in the diphthong.

In the middle stage of the diphthong [ɪ] (2 to 3), the spectral energy distribution shows a strong concentration in the lower frequency range (0-1000 Hz) with relatively less energy around 300 Hz.

In the nasal appendix (3 and 4), the spectrogram as the energy gradually dissipates all frequencies. The waveform maintains a relatively low amplitude, indicating a sustained intensity of the nasal murmur resulting from the vocal cords' vibration.

The segmentation of the nasal appendix duration involved identifying the boundaries of the nasal appendix segment within the nasal glide and the silence of the following consonant, in this case, /k/. The yellow key represents the acoustic waveform and spectrogram signal of the nasal appendix. To identify the nasal appendix boundaries, we analyzed acoustic characteristics based on visual observations. We noted the damping in frequencies and intensity and the presence of nasal murmur.

The aerodynamic data were extracted from variations in the air volume's behavior over time and the high volume of the nasal airflow in the segments. The data were used to generate a graph representing the change in air volume ( $\text{dm}^3/\text{s}$ ) and time (ms). The air volume was represented in a chart ranging from  $-100 \text{ dm}^3/\text{s}$  to  $300 \text{ dm}^3/\text{s}$ .

The segmentation and data extraction of the oral airflow (AFo) and nasal airflow (AFn) parameters were segmented into three parts. We infer that the maximum nasalization volume refers to when the velum is completely open and the tongue is in contact with the velar/palatal region, thus obstructing the oral cavity and directing air to the nasal cavity.

Figures 26 and 28 present the aerodynamic segmentation for velum displacement and duration extraction in oral and nasal diphthongs, respectively. In (1), there is a representation of the initial velum movement, and the nasal airflow is under zero. In (2), the velum is lowering, and the nasal airflow is positive. In (3), nasalization has a high airflow rate. In (4), the nasalization airflow rate drops, and the velum is rising.

We infer that the maximum opening of the velum can be identified through a combination of visual inspection and analysis of the nasal airflow trajectory. By examining the nasal airflow data, we looked for the boundary between the glide and the nasal appendix.

In this point, between 2 and 3 (in figure 26 and 28), the velum reached its widest position, indicating maximum opening. This was determined by observing the trajectory of the nasal airflow and identifying the moment when the nasal airflow started to be elevated quickly and reached a high-volume level, suggesting the velum was fully open.

We believe that the high peak of nasalization between 3 and 4 (in figure 26 and 28), is related to tongue position in the glide target. The identification of the tongue position in the glide target is considered the constriction of the oral track in the nasal glide phase.

The onset of the velic rising duration interval was determined by identifying the moment when the nasal airflow started to decrease from its maximum value. This was observed as a downward trend from the nasal peak level of maximum nasal airflow. The target point was marked as the time when the nasal airflow began to decline.

The offset of the velic trajectory interval was identified as the point when the nasal airflow reached zero  $\text{dcm}^3/\text{s}$  or approached a negligible level (in 4). We

established the moment when the nasal airflow trajectory showed a complete cessation or near cessation. This indicated that the velum was closed and the oral airflow started to be positive again.

### 3.3.8 Parameters

The analyzed aerodynamic parameters were extracted from the behavior of air volume concerning time, which generated a graphic representing the change in air volume ( $\text{dm}^3/\text{s}$ ) and time (ms). The airflow variation was related to the movement of the articulators in the oral cavity and in the nasal cavity.

To determine the parameters, we employed a combination of qualitative analysis and careful examination of the acoustic, articulatory, and aerodynamic features in our data. This analysis involved a close observation of the patterns and changes in the nasal airflow data to identify the relevant time points during the closure movement of the velum.

By closely examining the nasal airflow waveform and tracking its trajectory over time, we were able to determine the specific parameters of interest. Based on the data, we established standardized abbreviations to represent the articulatory movements for further analysis and data quantification.

- Negative duration of AFn (**NegAFn**) (ms): This is the measurement of the trajectory of the nasal airflow shape when the aerodynamics become negative (under  $0 \text{ dm}^3/\text{s}$ ).
- Velic opening duration (**Vod**) (ms): This measures the duration of the positive AFn trajectory. We infer that it is the beginning or somewhere very close to the moment of velic lowering. At this moment, the nasal airflow becomes positive. We considered this to be the first phase of the nasal airflow trajectory.



- Duration of glide constriction (**Dgc**) (ms): This is the measurement when the velum is at its maximum opening and the tongue is in the glide target position, creating a constriction in the oral cavity. We considered this to be the second phase in the nasal airflow trajectory.
- Velic rising duration (**Vrd**) (ms): This is the measurement of the duration of the nasal airflow shape when the velum begins its closure movement, leaves the target position, and completes the movement when the airflow reaches 0 dm<sup>3</sup>/s. We consider this to be the third phase of AFn.
- Nasalization Peak (**NP**) (dm<sup>3</sup>/s): This measures the maximum value of the AFn concentration during the production of the target segment.
- Total duration of the velic trajectory (**Tdvt**) (ms): This is the total measurement of the duration of the AFn shape. All these velum trajectories integrate the three phases of the nasal airflow contour (the first phase, the second and the third phase). This measurement is the inference based on the sweeping gesture of the soft palate.

For the acoustic data extraction, the oral diphthongs were divided into two portions. The initial portion refers to the oral vowel and the second to the oral glide. The nasal diphthong and the diphthongized nasal vowel were divided into three parts, the nasalized vowel, the nasal glide, and the nasal appendix, as shown in Section 3.3.7.

The durational values that were extracted are presented in milliseconds and follow the acoustic segmentation of the spectrogram, as shown in Section 2.3.9 of this chapter. The initials of the acoustic durations are listed below:

- **AdVn** – Acoustic temporal duration of the nasalized vowel (ms).
- **AdGn** – Acoustic temporal duration of the nasal glide (ms).

- **AdAPn** – Acoustic temporal duration of the nasal appendix (ms).
- **AdVs** – Acoustic temporal duration of the oral segment (ms).

### **3.4 Results**

This section describes the results of the qualitative and quantitative aerodynamic data analyses of the target segments. The results are presented in tables, graphs, spectrograms, oscillograms, and airflow graphs. In this section, we intend to provide answers to our questions about nasal diphthongization.

Based on the visual inspection of the spectrogram, we noted that the oral diphthongs were characterized by two phases and that the nasal diphthongs and the diphthong vowel had three stages.

There is a fully specified vowel in the first phase of oral diphthongs, and there is a glide in the second phase. There is a formant movement and harmonics that reflect the gradual movement of the tongue between the vocalic element target to the glide. The places in the spectrogram with the greatest concentration of energy acoustically reflect the articulatory effects of the vocal tract configuration.

A vowel occupies the central position of the syllable, and the coda position is occupied by a glide and not by a reduced vowel. The spectrum of the glide decreases in energy due to the partial constriction in the oral tract.

The first four formants generated from the speech signal are the main parameters for performing an acoustic analysis of speech that is focused on the vowel segments.

The formants are correlated in the following way:

(a) The articulatory movement in the vertical plane is linked to F1; for example, the height of the tongue and the opening of the mandible.

(b) The articulatory movement in t in the horizontal plane is linked to F2. The movement of anteriorization or posteriorizing of the tongue.

(c) F3 is related to oral constriction.

(d) F4 correlates with the shape of the larynx and pharynx (KENT & READ, 1992).

Between the vowel and the glide target, there is a movement of formants that reflects the gradual movement of the tongue. The vowel has a specific and distinct quality in terms of its articulation and acoustic properties, while the glide has the vowel quality of vowels but with a slower articulation than a consonant, as evidenced by the spectrum's decreased energy. The partial constriction generates this lower energy in the oral tract, which is the defining condition of a glide.

We cannot confirm that nasal diphthongs are articulated in the same way as their oral counterparts using the aerodynamic data. The nasality feature stems from the articulatory gesture of the velum coupling between the oral cavity and the nasal cavity.

We note that the articulatory process of the movement and positioning of the velum, specifically in relation to the coupling between the oral cavity and the nasal cavity, has acoustic and aerodynamic consequences and produces distinct oral and nasal phonemes. However, the aerodynamic and acoustic data cannot show the tongue movement or the oral geometry.

The spectrum of a nasal vowel element is more complex than that of an oral one. The coupling generates poles and zeros, which correlate with the volume of the acoustic mass that passes through the two cavities.

Increased coupling between the oral and nasal cavities results in a more pronounced presence of pole zeros in the spectrum of vowel segments. The

movement of the soft palate during articulation makes it challenging to accurately measure the acoustic frequencies, which vary depending on the extent to which the velum is lowered.

The degree of coupling affects the dominance of resonances during different phases of the glottal cycle, with stronger coupling leading to greater decay of energy in oral resonances and a more prominent nasal resonance in the vowel spectra (PRASAD & YEGNANARAYANA, 2020).

The larger size of the nasal tract contributes to the widening of F1 and a decrease in its amplitude for nasalized vowels. On average, nasalized vowels demonstrate lower levels of F1 broadening and amplitude compared to their oral counterparts (PRASAD & YEGNANARAYANA, 2020).

In F1, the sound quality changes, leaving it medium/low, and consequently decreasing the energy of the spectrum and the amplitude of the sound waves. Therefore, it is challenging to standardize acoustic frequencies when they are influenced by nasal resonance. Thus, we note that nasal resonances have a dampening effect on low-frequency formants.

The increase in vowel duration<sup>60</sup> in the presence of nasalized resonances can be attributed to the spreading of energy and acoustic characteristics associated with nasality. When a vowel is nasalized, the coupling between the oral and nasal cavities leads to a broader spectral distribution and a longer duration of the nasalized resonances. This spread of energy extends the temporal extent of the vowel, resulting in an overall increase in vowel duration.

---

<sup>60</sup> Vowel duration refers to the length of time that a vowel sound is produced in speech. In the context of nasalization, the presence of nasal resonances can affect the temporal characteristics of the vowel, leading to an extended duration compared to non-nasalized vowels.

We cannot confirm that there is always an overlap between resonance and anti-resonance. The overlap depends on each case. Nevertheless, there is an increase or weakening of the formant bandwidth if the pole-zero is close to an oral frequency.

In our production data, the specific phonetic environment where a velar consonant (/k/) precedes the nasal diphthong and the presence of the velar onset time of /k/ could influence the movement of the formants, to a velar formant transition. Therefore, we also used the alveolar-dental stop /t/ to control the boundary influence resulting from the onset time of the following consonants.

In a spectrogram, the upward movement of F3 and F4 indicates a change in the acoustic characteristics of the sound, specifically related to the tongue position and the shape of the vocal tract during the articulation of the alveolar consonant.

It is visible as a diagonal trajectory that starts at 2000 Hz gradually moves upwards. It can be observed as a change in the position of the dark bands or energy concentration in the spectrogram over time.

Therefore, there is a short upward movement of F1 and F2. This formant movement can be seen in figure 29 in the dotted circle. There are cases in which there was a slight elevation of F3. These characteristics were more evident (STEVENS, 1998, p. 357).

Our data did not reveal any relevant velarization or palatalization influenced by the alveolo-dental consonant on the configuration of the oral formant. Figure 29 shows a vocalic murmur at the end of the oral glide articulation, as indicated by the key in the spectrogram.

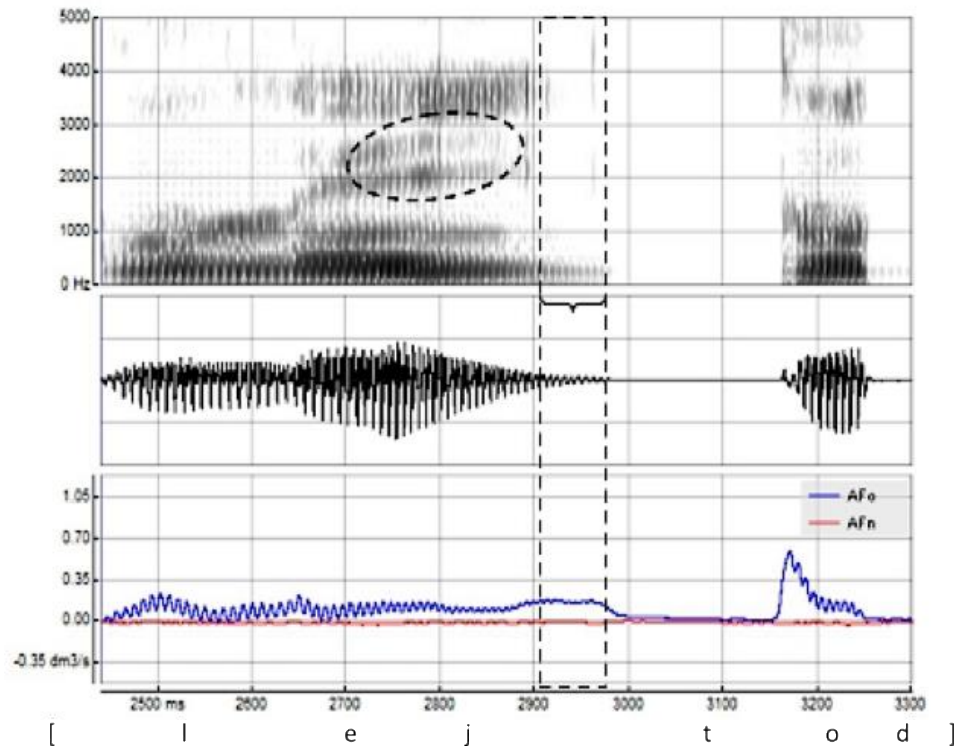


Figure 29: The figure consists of three panels. The top panel displays a spectrogram, the middle panel shows an oscillogram, and the bottom panel presents two-line graphs. The blue line represents the oral airflow, and the red line represents the nasal airflow. The keyword analyzed in this figure is 'lei'. The speaker for the data is identified as BH.

The murmur represents the final portion of the glottal airflow before the closure of the vocal folds during the production of nasal glide. It represents a period of reduced or weakened voicing, often characterized by a lower intensity or amplitude of the vocal fold vibration.

The final few pulses of glottis vibration spread thoroughly the following unvoiced consonant. There is a transitional phase between the voiced and unvoiced segments, where the vocal folds gradually cease vibrating before the onset of the unvoiced consonant. There is the stop silence after the dashed square. Thus, the consonant onset matched the absence of airflow caused by the closing of the oral cavity.

We observed the formant movement approximation of F2 and F3 in the glide spectrum in Figure 29. In general, when analyzing formant movement in glides, it is expected to observe a convergence of formants as the tongue moves towards a

constriction position. However, in some cases of our data, there was no convergence between the formants.

The movement of the formants in high-frequency regions indicates that the tongue is rising due to the glide constriction. This formant convergence can also demonstrate the posteriorization of the tongue body, even in alveolar stop contexts.

### **3.4.1 Mask Resonance**

According to Ghio & Teston (2004), the captured mask resonance is a response to the rapid variation of airflow. There were two types of problems in the data, namely the measures of delay and disturbances.

Depending on the positioning of the mouth mask, the time of the airflow meter's response may cause phenomena that delay measurements if there is a rapid variation in the flow rate. This delay refers to the sensor's response time or frequency bandwidth. Thus, connections such as pressure sensors or filtering were applied to the signal.

Pressure sensors are fast, and sensitive and behave in the same way as microphones by recording variations in acoustic pressure in the spectrum. The airflow sensor extends the vocal tract and may distort speech sounds depending on their transfer functions.

The mask has a specific volume, space, or air volume that exists between the mask itself and the speaker's face when the mask is worn. The mask is designed to fit snugly against the face, and this creates an enclosed space between the mask and the face, which can be considered as an additional volume in the system.

Accordingly, additional cavities within the sensor measuring tube can also produce certain acoustic effects, such as adding additional resonances in the acoustical signal. The presence of acoustical resonances or vibrations in the system,

such as those caused by the mask or the additional cavities, can lead to unwanted acoustic distortions in the recorded signals.

To address this issue, calibration processes are necessary to properly adjust and configure the system in order to minimize or eliminate these acoustic distortions and ensure accurate measurement and recording of the desired signals. (GHIO & TESTON 2004, p. 57).

We chose not to extract the formants for this thesis because we noticed a resonant range of close to 1000 Hz in all the acoustic data regarding the mouth mask's vibration.

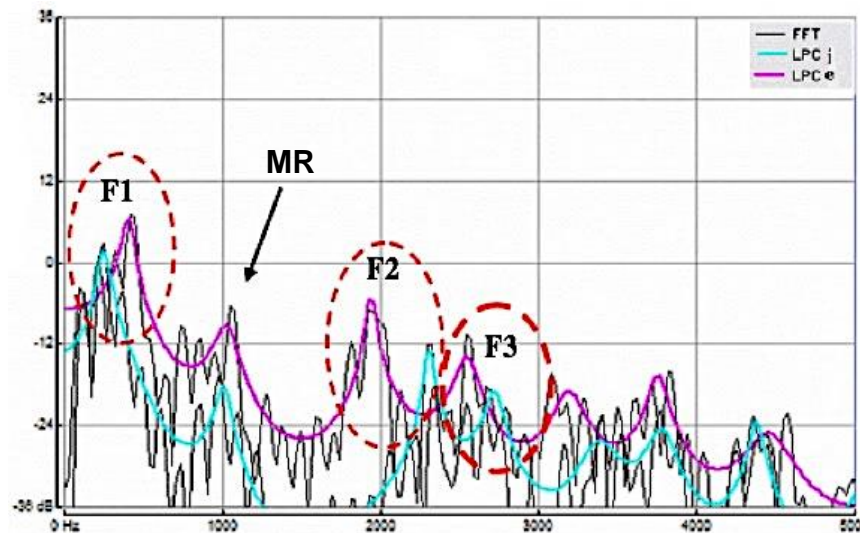


Figure 30: The figure displays the overlap of FFT (black line) and LPC spectra with the oral front vowel [e] (green line) and the glide [j] (pink line) within a 25 ms window. The analyzed word in this figure is 'frei'. The speaker for the data is identified as OD. Dashed circles are used to mark the positions of F1, F2, and F3 in the spectra. Additionally, an arrow represents the mask resonance (MR).



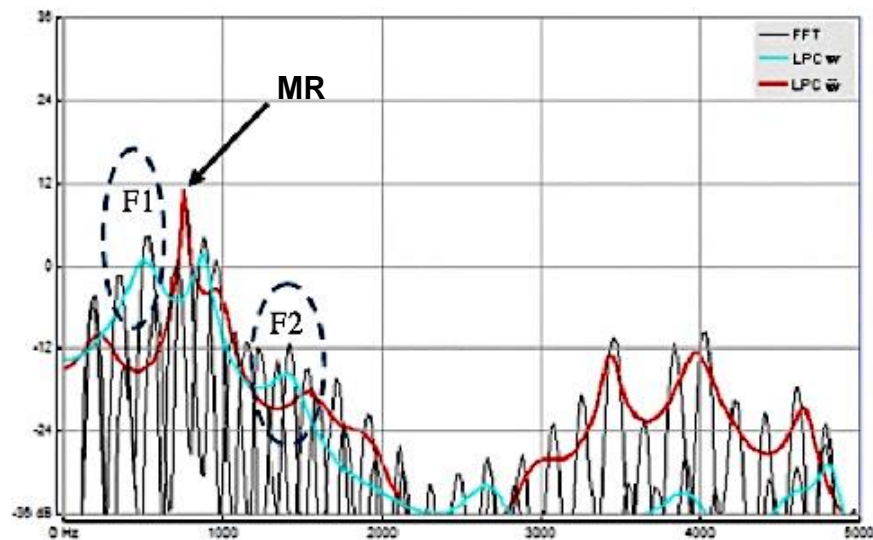


Figure 31: The figure illustrates the overlap of FFT (black line) and LPC spectra of the oral glide [w] (red line) and the nasal glide [w̃] (green line) within a 25 ms window. The analyzed words in this figure are 'sal' and 'são'. The speaker for the data is identified as LC. Dashed circles are used to mark the positions of F1 and F2 in the spectra. Additionally, an arrow represents the mask resonance (MR).

The frequency shown in figures 31 and 32 is a formant resulting from the vibration of the mask on the speaker's face. Ghio & Teston (2004) found that this extra energy concentration added to the spectrum during the recording of acoustic data and aerodynamics.

The presence of this frequency invalidates the acoustic analysis of the formants, as this resonance is constant and is within a frequency range that undergoes the addition of poles and zeros in nasal segments.

The mask resonance frequency was extracted and grouped in a table. We are convinced that this constant frequency was an alteration that was external to the oral and nasal cavities in this procedure. Mask resonance frequency values are presented in Table 5. The empty element ( $\emptyset$ ) represents the impossibility of determining the resonance of a flow frequency.

Table 5: Total values of the average (A) and the standard deviation (SD) in the resonance generated by the vibration of the oral mask during the production of the oral and nasal segments. N = 36.

Phoneme	MR Frequency (Hz)	
	Average (SD)	Speaker

	Male						Female					
	OD		BH		LZ		PL		LC		RD	
[e]	1067	(25)	991	(31)	1046	(28)	1065	(73)	997	(72)	1030	(34)
[j]	1046	(36)	1002	(42)	1042	(19)	1090	(78)	945	(57)	965	(33)
[a]	955	(34)	900	(56)	957	(37)	1045	(37)	1059	(34)	1082	(44)
[w]	813	(53)	739	(41)	768	(36)	847	(40)	881	(44)	847	(40)
[ē]	1666	(47)	1030	(48)	1057	(59)	1048	(89)	1064	(68)	1057	(79)
[ĵ]	1048	(54)	926	(95)	1028	(34)	973	(60)	1025	(47)	944	(59)
[ã]	1051	(34)	915	(127)	909	(99)	857	(69)	977	(37)	1052	(105)
[w̃]	781	(35)	∅	∅	857	(82)	761	(73)	778	(51)	892	(26)
<b>Total</b>	1053	(40)	929	(63)	958	(49)	961	(65)	966	(51)	984	(53)

The presence of the Mask Resonance frequency around 1000 Hz invalidates the acoustic formants analysis. This resonance is constant and within a frequency range that suffers from the addition of poles and zeros in nasal segments.

### 3.4.2 The Negative Nasal Airflow

Based on our data, we identified a negative nasal airflow at the beginning of the velum displacement. According to Cagliari (1977), negative airflow in nasalization aerodynamics changes the configuration of the formants. The author pointed out that the initial movement of the soft palate produced this negative airflow.

In our data, a negative airflow preceding nasal consonants and nasalized vocal segments was noted. We hypothesize that the negative airflow was related to the initial movement of the velum.

The small amount of initial displacement energy applied to the articulator created a slight change in the system, and the airstream inside the nostrils was a slight breath, which created a negative displacement of air particles.

The velum is closed in the default position and initiates the lowering. Furthermore, this first movement causes negative airflow. At the beginning of the nasal vowel articulation, we realized that the negative nasal airflow did not create an

acoustic effect in the spectral resonance. This spectral portion has oral formant energy and appears to be less nasalized.

When the opening velum movement begins and increases the velopharyngeal cross-section degree, the spectrogram gradually becomes more nasalized. The small initial displacement of energy applied to the articulator creates a minor change in the system and causes a slight intake of breath via the nostrils. Thus, there is an ingressive displacement of the air particles.

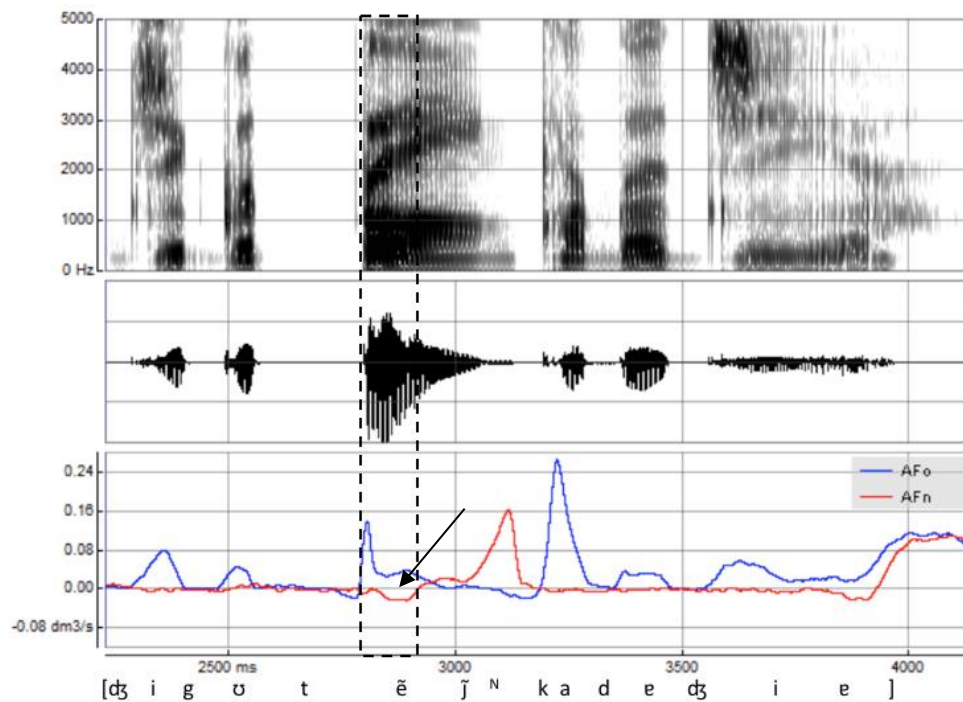


Figure 32: This figure includes a spectrogram (top), oscillogram (middle), and graphs showing oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'tem'. The data corresponds to the speaker identified as LC. The arrow in the figure indicates the negative airflow.

Figure 32 shows that the nasal airflow is slightly negative at the start of the nasal diphthong articulation. We infer that the negative airflow is due to the gradual lowering and opening movement of the soft palate.

The negative AFn was not used as an analysis parameter in this study because the negative nasal airflow volume and durational measures varied significantly. As discussed in the introduction in this thesis, the negative nasal airflow variability may

be related to the presence of positive or negative nasal airflow in some cases; thus, it may not be an obligatory factor for the production and perception of nasality (LAVIER, 1975). Hence, we did not analyze the acoustic effect of the negative AFn.

### 3.4.3 Back Diphthong

In the oral back diphthong [aw], the partial obstruction of the airflow in the oral cavity generated acoustic effects in the spectrogram due to the gradual articulatory movement between two vocalic elements with different types of constrictions.

In the figure below, each element's target position is presented to illustrate the position of the tongue during the production of the diphthong [aw]. This illustration makes it easy to see the points of articulation.

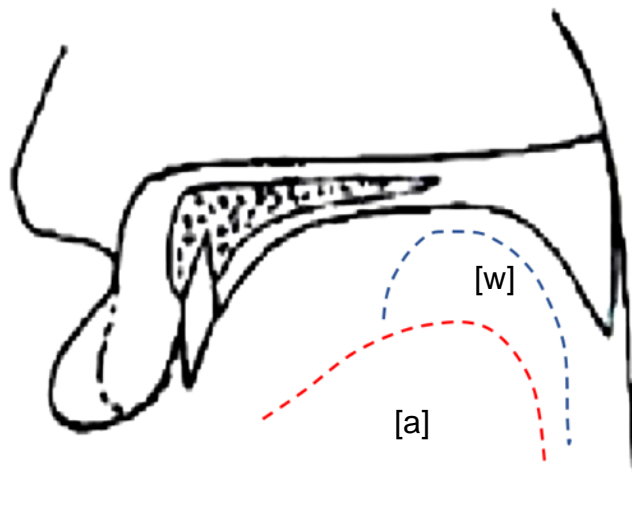


Figure 33: Illustration of mid-sagittal section of the oral cavity. Illustration of target positions for each phoneme [a] and [w]. Adapted from Kent (1997, p. 66).

From the figure above, it is possible to infer the trajectory of the tongue inside the oral cavity towards the posterior part from the fixed sequence of the articulation points of the two target sounds. Thus, it is possible to observe the volume increase in the front part of the oral cavity. This volume is the result of the glide target position.

Thus, in the spoken the word [saw], there is an F2 upward movement to around 2000Hz; this initial transition is due to the stop occlusion release of the fricative [s], and the formants undergo the onset of movement from the tongue posteriorizing to reach the oral back-vowel target.

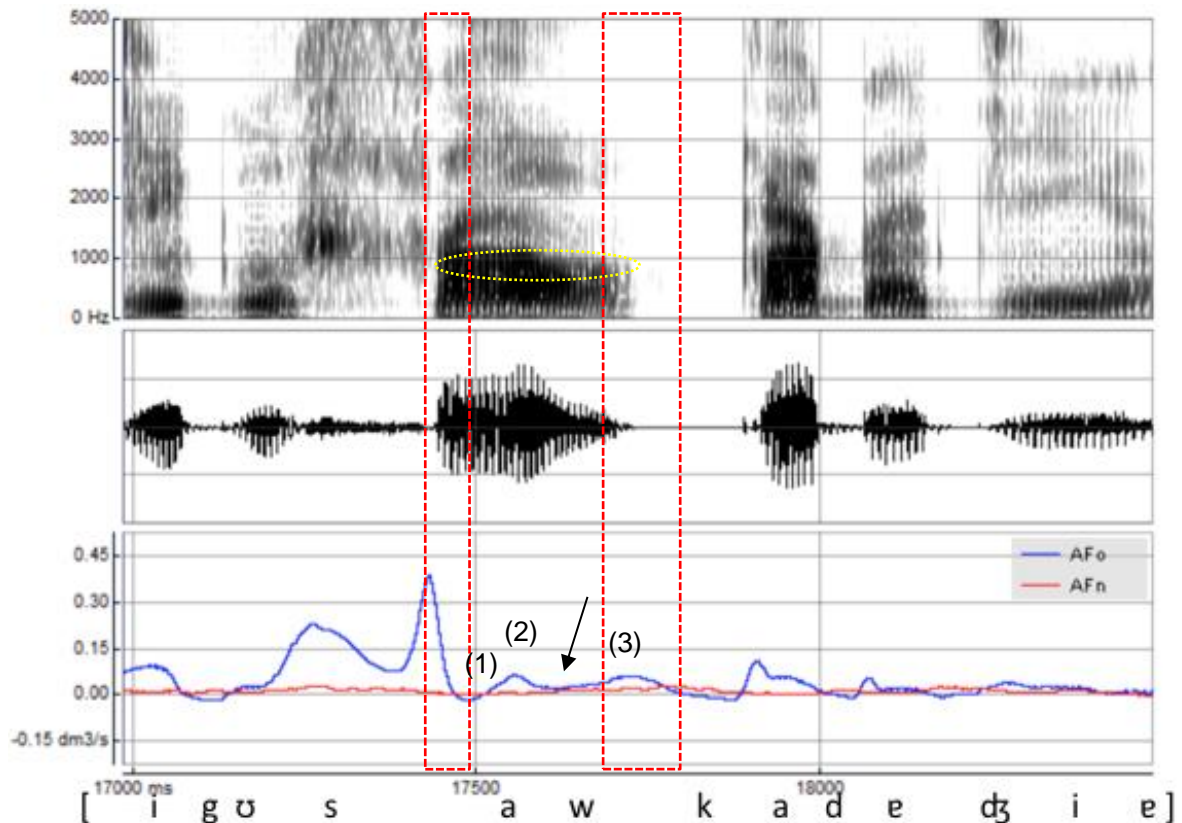


Figure 34: This figure displays a spectrogram (top), oscillogram (middle), and graphs showing oral airflow (blue) and nasal airflow (red line). The keyword examined in this figure is 'sal'. The data pertains to the speaker identified as BH.

In Figure 34, the dashed rectangle in red indicates the vowel onset time and the formant's transition due to the relationship between the release of the fricative consonant and the onset of vocal fold vibration. The dashed circle in yellow indicates the energy concentration of the mask resonance frequency formants. It is possible to note that the MR is very close to F1.

Accordingly, we can say that, in back diphthongs, there is a greater concentration of energy in this region of lower frequencies. However, the oral back glide has less

energy than does the antecedent oral vowel. There is a greater intensity concentration in the F1 and F2 regions due to the increased back articulation of the diphthong because of the lip rounding.

When the back articulation of the diphthong is combined with lip rounding, it further enhances the intensity concentration in these formant regions. Lip rounding can affect the acoustic properties of the vocal tract, altering the resonant frequencies and overall intensity distribution.

Therefore, lip rounding should be considered as an additional factor contributing to the intensity concentration in the F1 and F2 regions in the context of back diphthongs. The glide oscillogram of [w] has less frequency than does the oral central vowel [a].

As our speakers did not have any speech pathologies, the nasal airflow was close to or below zero  $\text{dm}^3/\text{s}$  in the oral segment. It is possible to infer the trajectory of the articulators inside the oral cavity via the variation in the oral airflow.

By analyzing the characteristics of the oral airflow, such as its rate, direction, and turbulence, we can infer the trajectory and relative positions of the articulators during different speech sounds.

These inferences are based on the knowledge of how specific articulatory movements correspond to specific airflow patterns and how they shape the acoustic output of speech.

It is important to note that while oral airflow provides insights into articulatory movements, it does not directly provide a precise spatial representation of the articulators inside the oral cavity.

In figure 34, (1) represents the vowel onset gesture [a]. (2) represents the oral vowel's target, and the vowel's offset then coincides with the beginning of the glide

articulation [w]. At the beginning of the second articulation, indicated by the second dashed rectangle (3), the gesture will end when the lips close entirely, and the AFo passes the zero line.

The arrow shown in the aerodynamic data marks the airflow trajectory and the point at which there is a velar constriction in the oral tract, which is the glide target. The glide target results from the constriction of the oral cavity and the posteriorization of the tongue. In general, the AFo did not approach or become negative. There is only a partial blockage of the oral tract occurred.

Aerodynamically, in glide [w], the volume of oral airflow that passes through the cavity is reduced, and the air is expelled in a jet form with greater volume. The glide release begins in this articulatory moment and reaches the target constriction, and these gestures conclude with the lip rounding in the offset position.

In the glide offset, the energy concentration of the beginning of the vowel/glide dissipates over time because of the tongue movement and the secondary labial articulation, which is indicated by the solid black arrow.

There was no oral-nasal coupling in the oral segments but, since there we observed the movement of formants, we can infer that this was due to changes in the volume in the oral tract, which altered the oral resonance.

#### **3.4.4 Back-Nasal Diphthong**

The difference between the oral and nasal segments may be attributed to more than the soft palate activation concomitant with the tongue movement and the lip rounding. The velic lowering movement results in the airflow being expelled through the nostrils.

We cannot infer that the same articulatory process occurs in the back-nasal diphthong [ãũ] as in its oral counterpart, nor can we claim that the articulatory process occurs in the same way in the oral cavity.

Acoustically, this coupling changes the frequency values and increases the bandwidth, inserting poles and zeros into the spectrum. Anti-resonances generated by the nasal airflow caused all these effects, and such acoustic effects were not found in the oral counterparts.

When there is oronasal coupling, we infer that the tongue movement will adapt to the new oral tract geometry to maintain the harmonic pattern to differentiate between the oral and the nasal sound quality.

In the figure below, the aerodynamic changes reflect the shape of the formants. The initial movement of displacement of the soft palate in [m] produces a negative airflow. The number (1) indicates the negative nasal airflow. The phoneme [m] the shape of the nasal airflow can be described as a plateau.

The arrow in Figure 35 shows the NP, representing the nasal appendix in the final portion of the glide; (2) indicates the nasal appendix.

To identify the nasal appendix, we observed the dampening of the frequency range of 1000 Hz, the decrease in energy around the frequency range of 3000-4000 Hz; and the presence of nasal formants, such as the nasal antiformant. The nasal antiformant manifests as a dip or decrease in energy around the frequency range of 2000-3000 Hz, providing evidence of nasalization and indicating the presence of the nasal appendix.

The nasal appendix is generally of relatively short duration compared to the preceding nasalized segment. Its identification can be inferred by considering the preceding nasal segment and observing the subsequent closure of the velum. The



AFn shape has a high ascending movement related to the oral constriction that moves the airflow from the oral cavity to the nasal cavity.

We noticed a concurrence between the AFo and AFn at the beginning of the nasal diphthong articulation. The AFo was positive and became negative due to the closure of the oral cavity.

At this point, no air was passing through the oral cavity. Nevertheless, a considerable amount of nasal airflow passed through the nasal cavities as a result of the constriction of the oral tract, which allowed us to infer that the soft palate also remains open during the entire process of diphthong production.

In our data, we noticed that the high peak of nasalization at the glide articulation offset was in the nasal appendix region. We infer that this appendix emerged due to an increase in the volume of nasal air when the mouth is closed, and the vocal folds vibrate.

The beginning of the AFn's ascending shape was related to the glide constriction, but the high NP was related to the closure of the oral cavity and not to the narrowing caused by the velopharyngeal section closing.

In the figure below, (1) indicates the negative airflow and the elevation of oral airflow in [ʊ], (2) indicates the high peak of the nasal airflow, and the circle indicates the convergence of F3 and F4.

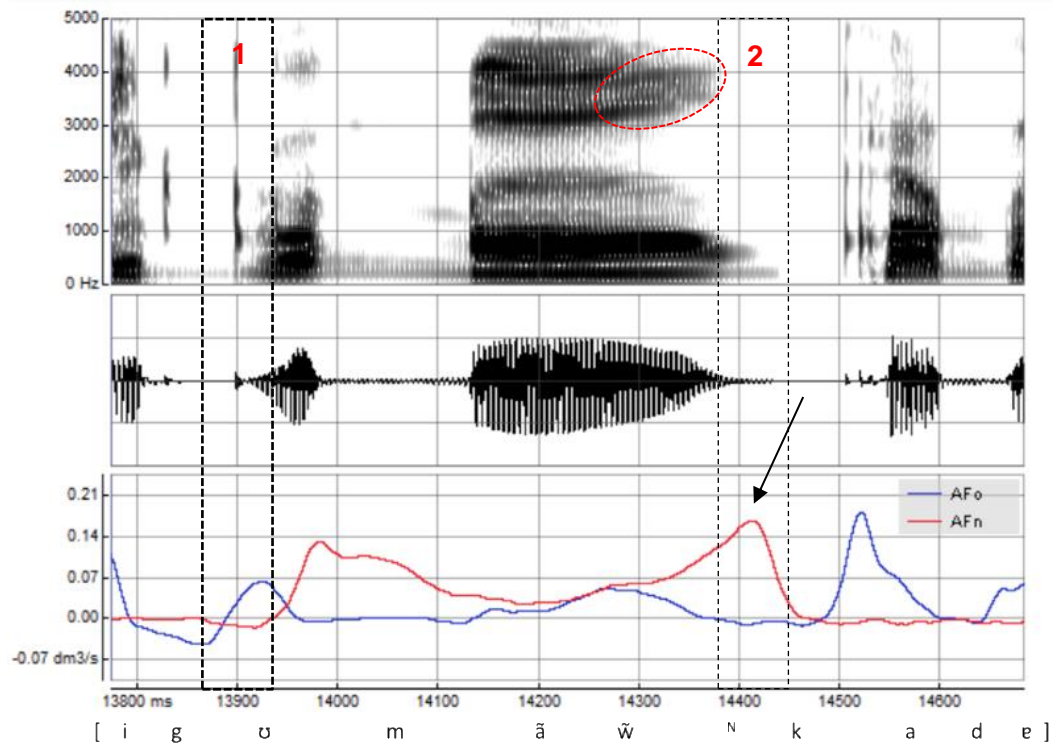


Figure 35: This figure presents a spectrogram (top), oscillogram (middle), and graphical representations of oral airflow (blue) and nasal airflow (red line). The analyzed keyword in this figure is 'mão'. The data corresponds to the speaker identified as LC.

The nasal consonant [m] is characterized by a distinct low-frequency energy concentration, seen as a dark band or in the lower frequency range. The closure of the lips creates a complete obstruction in the oral cavity. The velum is lowered, allowing air to pass through the nasal cavity. Acoustically, [m] is characterized by a low-frequency energy concentration, often seen as a dark band or in the spectrogram.

The closure of the lips for the [m] is released, allowing the airflow to pass through the oral cavity to produce the back nasal vowel. The nasal consonant [m] transitions into the following vowel segment consonant attenuates the movement of the formants.

The velum keeps lowered, and the vowel [ã] is characterized by an open vocal tract and a lowered velum, allowing nasal airflow. The beginning of the vowel [ã] of the spectrogram of the word “*mão*” is nasal due to the nasality assimilation. Acoustically, the vowel [ã] exhibits a relatively low-frequency energy concentration, often associated with the nasal formants.

F1 is close to the 700 Hz region, and there is a reinforcement of resonance caused by the RM and nasal resonances associated with the articulation of the velic gesture around FN1 250-500 Hz and FN2 between 500-1000 Hz.

We noted that the nature of nasal production always had a convergence of high-frequency formants in the final portion of the glide, as indicated by the dashed circle. The coarticulatory effects of oral constriction and velum lowering can contribute to the convergence of high-frequency formants. The third and fourth formants in the acoustic spectrum of speech become close to each other, it is an overlap.

The convergence of F1 and F2 are probably Helmholtz resonators because of lip rounding for [w]. The convergence of F3 and F4 formant transition is a common attribute of constriction on the velar region, as there is the spectral prominence of the F3 peak because F3 and F4 could be modeled as closed tube resonators. The movement of the formants reflects the transition between the vowel and the glide during articulation, which increases the airflow passing through the nasal cavity.

If we compare the FFT and the LPC of [a] with [ã], the differences between the nasal and oral segments mainly occurred at the energy amplitude, resulting in the bandwidth and the distance between the formants increasing.

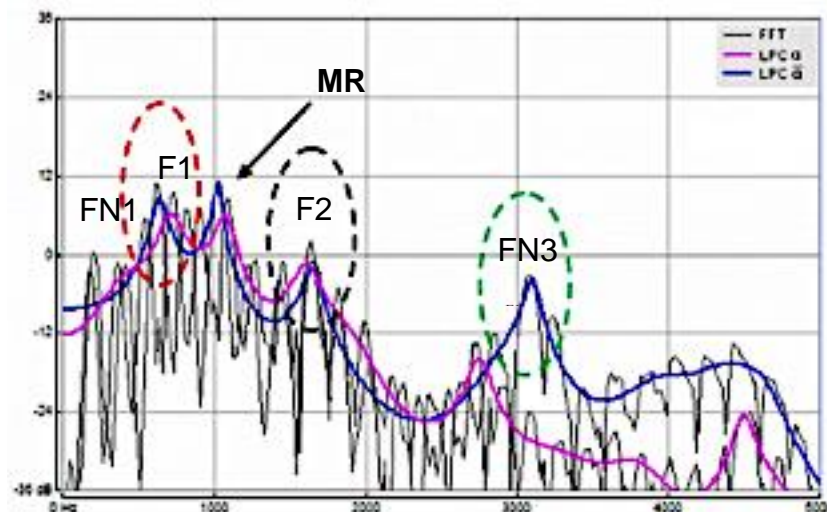


Figure 36: The FFT (black line) and LPC spectrum overlap of the oral central vowel [a] (pink line) and the nasal vowel [ã] (blue line) in the 25 ms window. Words 'sal' a 'são'. Speaker PL. Dashed circles mark F1, F2, and FN3. The arrow represents MR.

We noticed that the FN1 of [ã] had a slightly lower frequency difference than did [a] due to the initial opening of the nasal cavity and the low frequency of the nasal resonances that overlapped with the oral resonances. Thus, for [ã,], the energy concentration of the FN1 was around 500 Hz.

While there was an increase of energy in the frequencies corresponding to the FN1 compared to their oral formants, as well as a widening bandwidth, FN1 and FN2 remained relatively stable when comparing the oral and nasal formants. In this example, the FN3 was reinforced and moved towards high frequencies. Thus, FN3 had a narrowing bandwidth, and the amplitude increase was generated by a pole.

The nasal glide is articulated with the tongue the backward movement of the tongue, a labial constriction, and velic lowering, thus reducing resonance in the oral cavity. This alteration results in the introduction of anti-resonances in the spectrum. These antiformants occur at specific frequencies and cause a decrease in energy in those frequency ranges. In the case of velarization, one of the effects is the attenuation or absence of energy in the region between 2000 and 3000 Hz.

The reduction in resonance in the oral cavity during the articulation of a nasal glide is primarily due to the presence of zeros in the acoustic spectrum. Zeros occur when there is a cancellation or attenuation of certain frequencies in the vocal tract transfer function. In the case of nasalization, the coupling between the oral and nasal cavities leads to the introduction of anti-resonances

Figure 37 shows the comparison of the FFT and the LPC of [w] with [w̃]. The differences between the nasal and oral segments mainly occurred at the strength of the peak in the spectrum indicated by the reinforcement of the MR and the widening bandwidth.

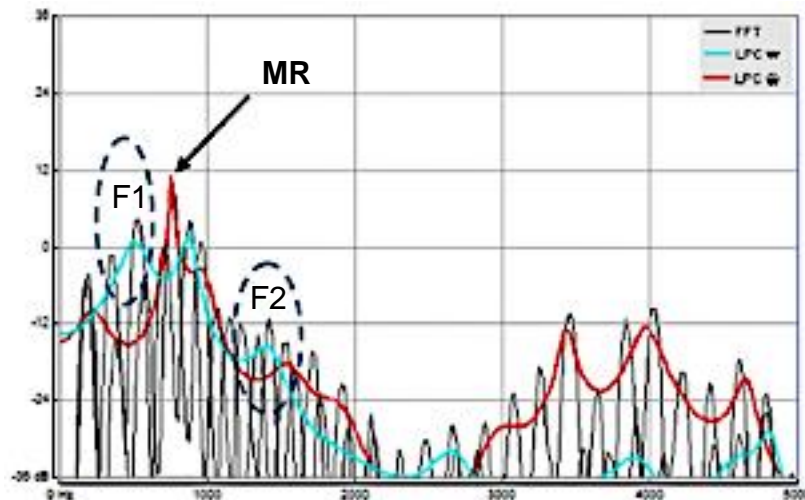


Figure 37: The FFT (black line) and LPC spectrum overlap of the oral central vowel [w] (green line) and the nasal vowel [w̃] (red line) in the 25 ms window. Words 'sal' and 'são'. Speaker LC. Dashed circles mark F1 and F2. The arrow represents MR.

There was a relatively low-frequency energy level compared to other frequency regions in the FFT and the LPC for the oral [w]. By examining the amplitude or intensity values at different frequency points and comparing their relative magnitudes, there is a damping in the region between 2000 Hz and 3000 Hz, resulting in the acoustic effect of glide constriction.

The black arrow shows where there was MR in the nasal glide. The MR had a nasal pole with a considerable amplitude and a very narrow bandwidth in this example. The dashed circles indicate the presence of F1 and F2 in the oral glide. Nasal formants appeared to be extremely damped in the spectrum.

From this comparison between the nasal and oral spectrum, we understand that the nasal vowel and nasal glide are both segments influenced by nasal resonances. Based on Figure 36, the central nasalized vowel exhibits a prominent nasal peak at FN3, indicating the influence of nasal resonances during this segment.

If we compare the vowel to its oral counterpart, the vowel was nasalized because there was a variation in the formant pattern. Aerodynamically, the velum movement had already begun, and the nasal airflow was increasing gradually.

We observed that the nasal gesture, or the nasalization of vowels, is typically a result of the assimilation of adjacent elements in Brazilian Portuguese (BP). By assimilation, we mean the influence and modification of the vowel by neighboring nasal sounds or nasal articulatory gestures.

### 3.4.5 Front Diphthong

During the production of the oral front diphthong [ej], the articulatory change between two vowel elements has different degrees of constriction. In the figure below, each element's target position to produce the front diphthong is indicated to illustrate the movement of the tongue inside the nasal cavity. This illustration makes it easy to see the points of articulation.

Figure 38 represents the point and mode of articulation of the front oral diphthong consisting of [e], represented by the dotted line, and the continuous line, which indicates the glide [j], respectively.

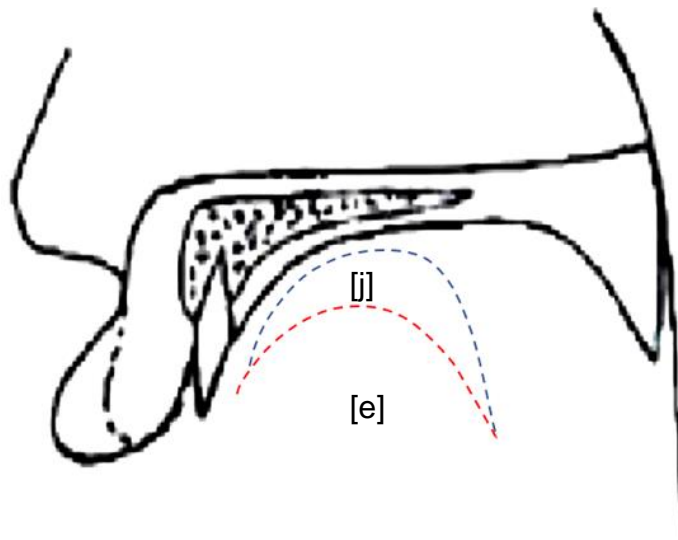


Figure 38: Illustration of the mid-sagittal section of the oral cavity. Illustration of target positions for each phoneme [e] and [j]. Adapted from Kent (1997, p. 66).

Figure 38 illustrates that during the articulation of the palatal glide, the tongue volume increases, resulting in a narrowing of the palatal region. Specifically, the back

region of the oral cavity exhibits a larger volume compared to the front part of the tract. This indicates that when the constriction for /j/ is made, the back region of the oral cavity becomes larger than the front region.

The articulatory changes in the oral tract's geometry result in variations in the acoustics and aerodynamics. However, it cannot be inferred that the oral gestures have the same articulatory characteristics when the soft palate produces the nasal diphthong.

We illustrate some acoustic characteristics of front diphthongs in the spectrogram for [dej] in Figure 39 below. [e] is a mid-front vowel, and its articulation involves a relatively open oral cavity, resulting in a more uniform distribution of energy across the spectrum. In contrast, [j] is a palatal glide with a constriction in the palatal region, leading to a concentration of energy in the higher frequencies, particularly in the region associated with F2 and F3.

As the palatal gesture characterizes the vowel [e] and the glide [j], the slight decrease in the frequency of F1 can be attributed to both the movement of the tongue body towards the anterior part of the oral cavity and the raising of the tongue. These articulatory movements result in a narrowing of the pharyngeal-oral constriction, which leads to a decrease in the resonant frequency of the F1 (STEVENS, 1989).

At the beginning of the vowel, a formant transition gradually elevates F2, which approaches F3. According to Stevens' tube models, the rising of F2 is primarily influenced by the relative lengths of the front and back cavities rather than the raising of the tongue. The tube models explain that as the tongue moves towards the palatal region, the length of the front cavity increases, leading to a higher resonant frequency and thus an elevation of F2.

In the central portion of the vowel [e], we noticed that the formants of F2 were already showing a gradual ascendancy. This formant elevation is a specific characteristic of the diphthongization in front segments.

The upward movement of F2 is related to the elevation of the tongue in the palatal region, which resulted in a lack of energy in the region of medium frequencies between 500 Hz and 2500 Hz. In the spectrogram, the MR was around 1000 Hz.

The ascending formant movement is necessary for the perception of the quality of the glide. We infer that the glide has an articulatory point that is further back than the vowel [e] because a higher formant frequency of F2 suggests a more anterior articulatory position.

In the case of [e] and [j], the vertical movement of the tongue towards the palate primarily affects the degree of constriction and influences the resonances of the vocal tract, including F1. The change in tongue height impacts the cavity length, which in turn affects the frequency of the lowest resonance, F1, according to the principles of tube modeling proposed by Stevens.

As the tongue is raised towards the palate, the oral cavity lengthens, resulting in a longer vocal tract and a decrease in F1 frequency. The relationship between tongue height and F1 frequency is generally inverse, meaning that higher tongue position corresponds to lower F1 frequency.

The differentiation between [e] and [j] is mainly attributed to the articulatory point and manner of the constriction, which influence the higher formants such as F2 and F3. F2-F3 transition frequencies are typically higher in [j] compared to [e] due to the palatal constriction and the positioning of the tongue closer to the hard palate.

The dissipation of energy in the diphthong over time is primarily due to the tongue movement towards the palatal constriction, which affects the acoustic properties of



the vocal tract. As the tongue moves towards the palatal region, it changes the shape and size of the oral cavity, altering the resonances and acoustic characteristics of the sound. This alteration in constriction redistributes the energy across different frequency ranges, resulting in a decrease in acoustic amplitude as the diphthong progresses.

This transition results in a reduction in format energy and a reduction in the oral cavity due to palatal constriction. According to Stevens (1989, p. 330), a glide has a slower release with a more significant constriction and no contact between the tongue and the palate during articulation. The yellow arrow shows energy losses in the spectrogram, which may have been the result of glide constriction.

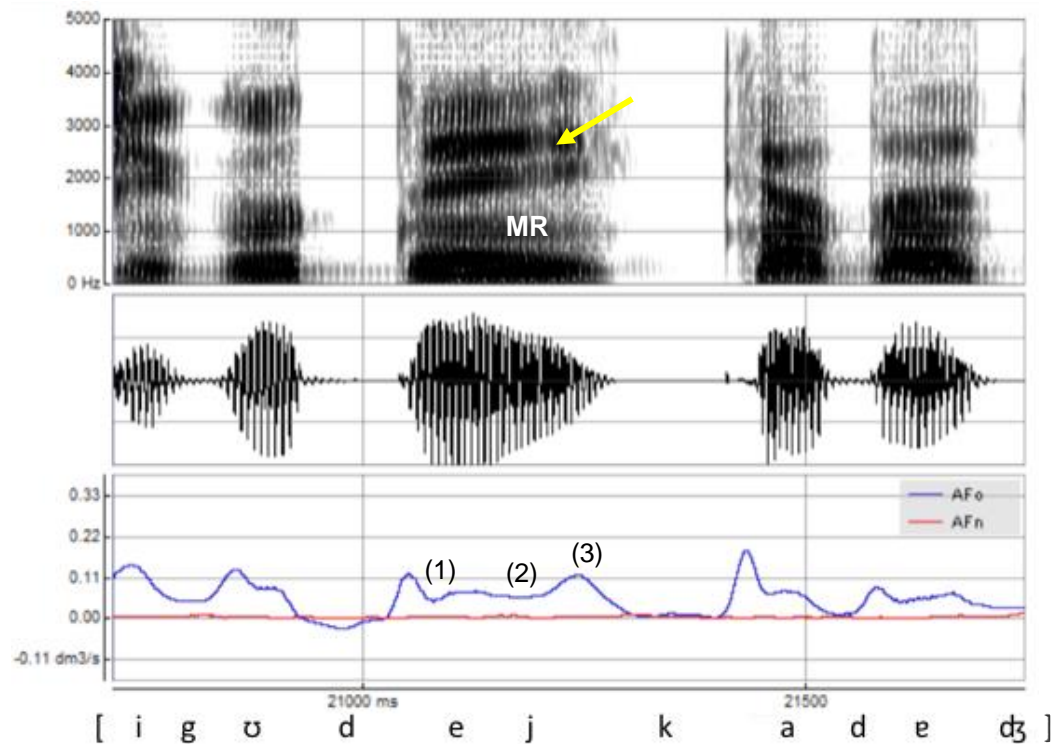


Figure 39: This figure presents a spectrogram (top), oscillogram (middle), and graphics depicting the oral airflow (blue) and nasal airflow (red line). The keyword used for analysis is 'dei', and the speaker for this data is LC.

The AFo shape was similar in the production of the front and back diphthongs. We infer that this occurred because the articulatory manner was the same and transitioned between a more prominent segment and a less prominent one.

In the production of the [ej] diphthong, there is an open and unobstructed airflow as the articulatory position transitions from the initial vowel [e] to the glide [j]. This is reflected in the oral airflow shape, where the airflow maintains a relatively stable and continuous pattern.

As the tongue moves towards the palatal region to produce the [j] glide, a partial constriction occurs in the oral cavity, causing a reduction in the oral airflow. The inference of (1) being the onset of the vowel gesture can be supported by observing changes in the oral airflow pattern.

Prior to (1), there is a reduction in the oral airflow associated with the closure or constriction of the preceding consonant. The subsequent initiation of the vowel gesture is indicated by the resumption or increase in the oral airflow.

The inference of glide constriction in (2) can be made by examining the changes in the oral airflow during the production of the [j] glide. During the constriction phase, there is a decrease in the volume of the oral airflow compared to the preceding vowel. This reduction in oral airflow corresponds to the narrowing or constriction of the oral cavity as the tongue moves towards the palatal region.

The identification of the offset of the glide articulation in (3) may involve a release of the constriction, resulting in changes in the volume of flow rate, so it is possible to infer the offset of the glide articulation.

The airstream was expelled as a jet during the glide constriction phase due to the narrowing or constriction of the oral cavity. The airflow velocity and amplitude during this phase are higher compared to the preceding as the vowel.

In oral production, the nasal airflow became zero  $\text{dm}^3/\text{s}$  position. Therefore, no air was passing through the nasal cavity. It indicates that during that production, there is a complete closure or blockage of the nasal passage. This occurs during articulatory

configurations when the velum is raised to close off the nasal cavity, preventing air from passing through it.

In Figure 40, we compare the FFT and the LPC of [e] with [j]. The general differences between the vowel and glide segments occurred at the formant position and the peak energy amplitude. In the figure below, the MR is around 1000 Hz.

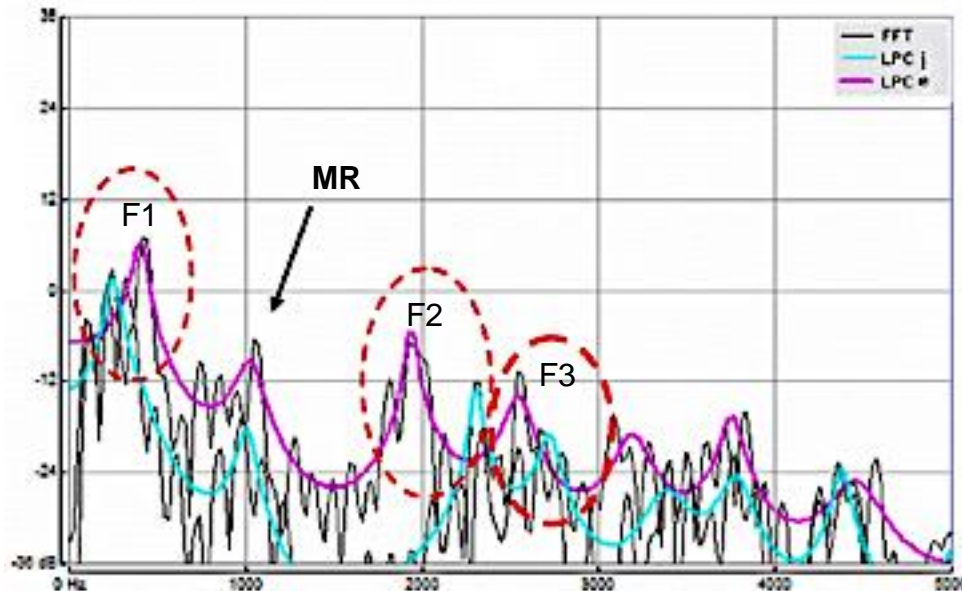


Figure 40: The FFT (black line) and LPC spectrum overlap of the oral central vowel [e] (pink line) and the nasal vowel [j] (green line) in the 25 ms window. Words 'frei' and 'sem'. Speaker: OD. Dashed circles mark F1, F2, and F3. The arrow represents MR.

Hence, we can observe that the merging or convergence of F2 and F3 in the glide portion [j] was primarily influenced by the specific place of articulation. The constriction created in the palatal region resulted in increased airflow resistance during the production of the glide, a characteristic that was not observed in the vowel segments. According to Stevens (1998), the variation in F3 does not influence glide perception since there is F2 ascendancy. Therefore, as the high formants have a higher frequency but lose intensity, there is less energy in the frequency region.

### 3.4.6 Front Nasal Vowel

In this section, we analyze the production of the nasal vowel [ẽ], which is produced via diphthongs with [ẽj]. As there is a glide insertion during the production of the nasal vowel, we expect that the comparison of the articulatory process of the diphthong nasal vowel and the oral diphthong will be similar and that we will be able to illustrate the palatalization process during speech production.

The entire spectrogram signal had an approximation of between F2 and F3. F1 lost energy, and F4 had rising formant movement. The low frequency of F1 indicated that the tongue was elevated and, when F2 was high, this indicated that the tongue was moving horizontally. This type of transition also suggests a velar articulation (STEVENS, 1998). However, our data were not able to demonstrate any velarization. The arrows show the increase in nasal airflow and the shape culminating in a peak in the airflow.

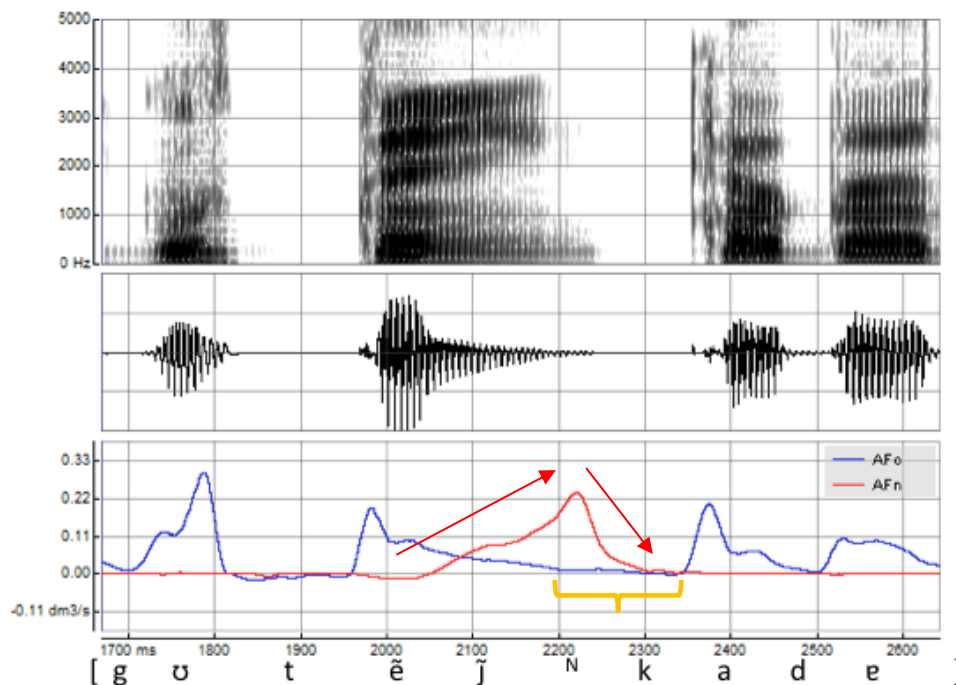


Figure 41: This figure displays a spectrogram (top), oscillogram (middle), and graphics representing the oral airflow (blue) and nasal airflow (red line). The arrow indicates the direction of positive nasal airflow. The keyword analyzed in this context is 'tem', and the speaker for this data is OD.

In this section, we analyze the production of the nasal vowel [ẽ̃], which is accomplished diphthongs with [ẽ̃j̃]. As there is a glide insertion during the production of a nasal vowel, we expect that comparing the articulatory process of the diphthong nasal vowel and the oral diphthong will be similar and that we will be able to illustrate the palatalization process during speech production.

The entire spectrogram signal had an approximation of between F2 and F3. F1 lost energy, and F4 had a rising formant movement. The low frequency of F1 indicates that the tongue is elevated and when F2 is high shows that the tongue is moving horizontally. This type of transition also suggests a velar articulation (STEVENS, 1998). However, our data were not able to indicate any velarization.

The F2 elevation observed in the spectrogram was the result of the tongue moving towards the palatal region from its initial position during the production of the glide. As the tongue approached the palatal region, it reached the target configuration for the glide articulation, leading to the elevation of F2. Thus, nasal diphthongization results in a glide with the same region of articulation as the vowel.

The synchronization of gestures between the lowering of the soft palate and the movement of the tongue body results in the emergence of the nasal appendix. This can be observed in the waveform as a feature associated with nasalization.

The airflow graphic shows similar characteristics between the back-nasal diphthongs and the nasalized front vowel. The AFn begins its trajectory with the negative airflow, and, in the spectrogram, we can see that there was more acoustic energy in this portion, which indicates a more oral portion of the vowel.

The trajectory of AFn was ascending, and the AFo was descending and gradually decreased as the tongue moved to reach the glide target. The AFn gradually became

positive, making the entire segment nasalized. The beginning of lowering of the soft palate began together with the vowel onset, but the AFn is negative.

We believe that negative nasal airflow occurs when there is a complete closure of the velopharyngeal port, preventing air from passing through the nasal cavity. However, when the velum gesture is active, it can cause an upward movement of the soft palate, redirecting the airflow through the nasal cavity.

We infer that there is a plateau in the form of the AFn. A plateau in the form of the nasal airflow (AFn) indicates that there is a stable and sustained airflow through the nasal cavity, suggesting that the soft palate is open and allowing air to pass through.

This coordination with the constriction of the oral cavity, particularly in the context of a glide, suggests that there is a significant degree of soft palate opening to facilitate the desired airflow pattern. The plateau in AFn indicates a balance between oral and nasal airflow, indicating the coordination between the two gestures.

The highest point in the nasal airflow shape, where the two arrows meet in Figure 41 shows the location of the peak of the nasal airflow. At this point, the AFo was close to 0 dm<sup>3</sup>/s, which indicates the oral cavity constriction, as shown in the yellow key. Thus, the airflow in the oral tract was directed into the nasal cavity until the soft palate closed.

This airflow shape was characteristic of the front nasal vowel and the back-nasal diphthong segments analyzed thus far. However, the nasal air volume was more prominent in the front nasal vowels than it was in the back-nasal diphthongs, as we will demonstrate in the following items in this thesis.

With regard to the front and back nasal vowels, the flat low frequencies were added without a spectral prominence of up to 1220 Hz. This spectrum dampening was

the result of the bandwidth increasing in FN1 to the front nasal vowel and the bandwidth increasing in FN2 to the back-nasal diphthong.

The addition of an FN dampened the low-frequency resonances. The addition of around 250 Hz resonance was the result of oronasal coupling, which blocked the presence of a spectral peak. Figure 42 represents the LPC and the FFT, illustrating the discussion above.

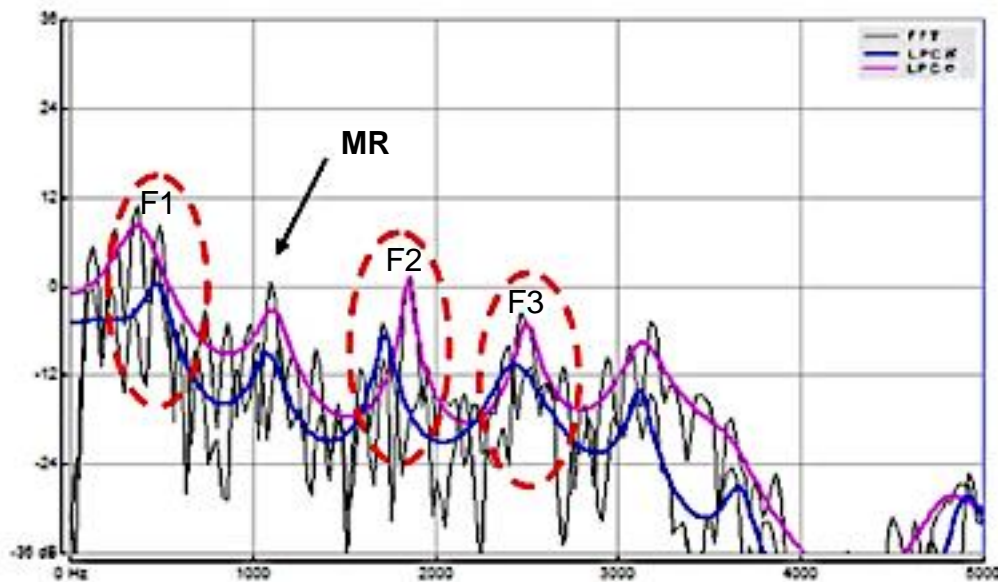


Figure 42: This figure presents the overlap of the FFT (black line) and LPC spectrum of the oral vowel [e] (pink line) and the nasal vowel [ẽ] (blue line) in a 25 ms window. The analyzed words are 'sei' and 'sem', spoken by the speaker OD. Dashed circles indicate the positions of F1, F2, and F3. The arrow represents the mask resonance (MR).

In the context of nasalization, the formants, including F1, are typically lowered in frequency due to the nasal resonance. The lowering in the resonances is a result of the interaction between the oral and nasal cavities. F1 was close to the 250 Hz nasal resonance; for this reason, it was more inclined to form a pole (STEVENS, 1998).

The low-frequency resonances associated with nasality contribute to the attenuation or dampening of the formants, rather than reinforcement. Therefore, the decrease in the amplitude of the formants, as indicated by the dotted circles in red in Figure 41, is a consequence of the nasalization process. The influences of these low-

frequency resonances characterize the nasalization and result in a shift and damping of the formant frequencies.

In the comparison between the nasal vowel [ẽ] and the oral vowel [e], what we observe is a shift in the frequency of the first formant (FN1) towards higher frequencies in the nasal vowel compared to the oral vowel. This shift indicates that the nasal resonance has an influence on the resonance characteristics of the vowel, causing a change in the frequency placement of the first formant.

Additionally, when comparing the spectral segments of the nasal vowel [ẽ] and the oral vowel [e], we notice an increase in the bandwidth of the formants. This means that the range of frequencies around the formants becomes broader in the nasal vowel. On the other hand, the amplitude of the sound waves decreases in the nasal vowel compared to the oral vowel. This attenuation of the amplitude is a characteristic effect of nasalization.

A characteristic of front vowels is that they exhibit harmonics in the higher frequency regions due to the elevation of the tongue body. This means that front vowels tend to have more energy in the higher frequency harmonics. However, it's important to note that all vowel sounds have harmonics in the higher regions due to the glottal source, which generates energy at the fundamental frequency and its integer multiples (harmonics).

The elevation of the tongue body in front vowels affects the resonances and formant frequencies of the vocal tract. As the tongue moves towards the front region, it alters the shape and size of the vocal tract, leading to shifts in the resonant frequencies. These shifts can result in formants like F2 and FN2 being located in higher frequency regions compared to back vowels.



We observed that high-frequency formants are less susceptible to damping effects compared to low-frequency regions. This means that high-frequency formants tend to maintain their energy and spectral characteristics more effectively. The influence of nasal cavity resonances, characterized by pole zeros, is more pronounced in the medium and low-frequency formants.

Figure 43 displays the comparison of the spectrograms, LPC spectra, and FFT spectra between the oral glide and the nasal glide. In the oral glide, F1 exhibited a higher amplitude for its frequency peak compared to FN1 in the nasal glide. Additionally, the oral glide had a narrower bandwidth for F1 compared to FN1.

The influence of the nasal resonances increased the bandwidth. Thus, the FN1 peak was dampened, became flat, and generated energy concentrations in the low frequencies. The nasal production sound became lower than the oral sounds (DELVAUX, 2003). There was a wide bandwidth in the high-frequency nasal formants, and FN2 and FN3 were damped as shown in the red line in the figure below.

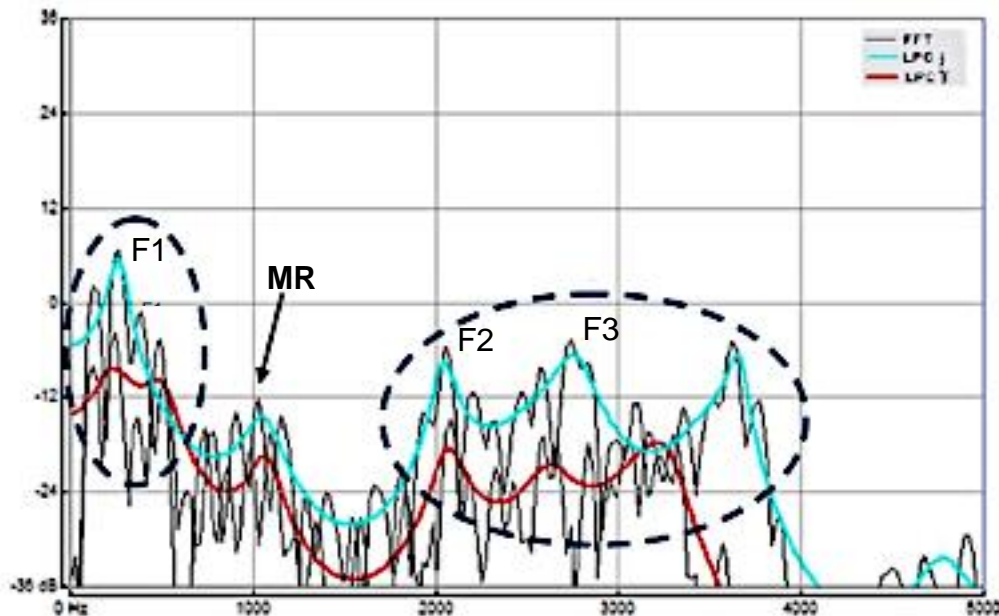


Figure 43: This figure shows the overlap of the FFT (black line) and LPC spectrum of the oral vowel [j] (blue line) and the nasal vowel [ɲ] (red line) in a 25 ms window. The analyzed words are 'sei' and 'sem', spoken by the speaker OD. Dashed circles indicate the positions of F1, F2, and F3. The arrow represents the mask resonance (MR).

The front nasal glide is the result of nasal diphthongization. The emergence of the glide is a concrete phenomenon that occurs due to the rising of the tongue. This allophonic process may be caused by the change in geometry, which forces the tongue to adapt to maintain the quality of the vowel and to reach the target of the sliding constriction in the palatal region.

### **3.4.7 Nasal Appendix**

In this section, we present the data from the aerodynamic analysis and the acoustic configuration of the nasal appendix in nasal diphthongization. The nasal appendix is an acoustic characteristic that is associated with nasalization in Brazilian Portuguese (BP), as documented by Sousa (1994). It represents a specific acoustic pattern observed in our data, which indicates the presence of nasalization.

Nasal spreading refers to the articulatory movement of the tongue and soft palate during nasalization. It involves the spreading of nasal resonance within the vocal tract. In the analyzed context, nasal spreading may play a perceptual role and contribute to the nasalization process.

However, it is important to note that Sousa (1994) reported in her thesis that there were cases where a particular speaker did not produce a nasal appendix. This finding suggests that the presence or absence of the nasal appendix may vary among individuals, highlighting the individual differences and complexities involved in nasal articulation in BP.

The acoustic portion of the nasal appendix indicates a narrow constriction in the oral tract due to the synchronization of gestures between the tongue elevation and the velum lowering.

Acoustically, there were nasal appendix resonances after the nasal glide. The most salient resonances in the spectrum were in the low-frequency regions near F0 and the high-frequency regions above 3000 Hz. The arrows show the increase and the decrease in the nasal airflow. There is a high peak of nasal airflow in (1).

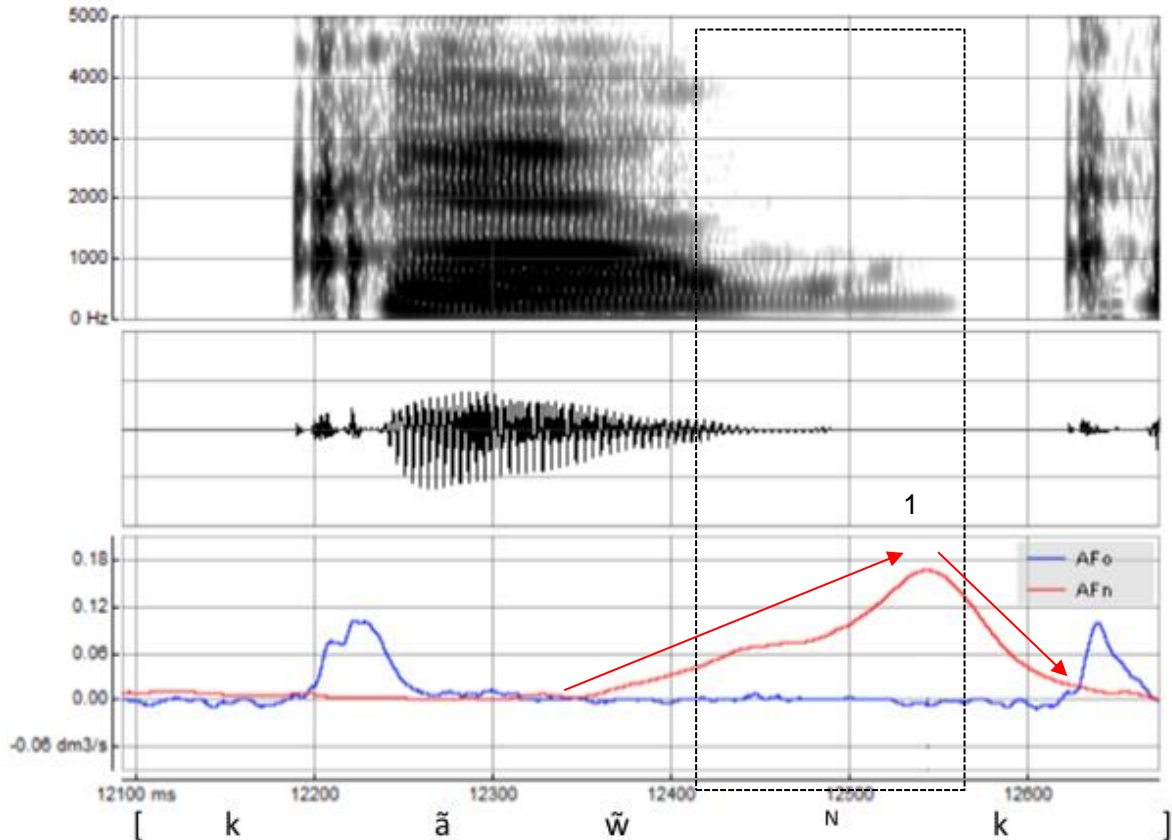


Figure 44: This figure illustrates the negative airflow indicated by the arrow. It includes a spectrogram (top), oscillogram (middle), and graphs showing the oral airflow (blue) and nasal airflow (red line). The keyword analyzed in this figure is 'cão', spoken by the speaker OD.

According to Demasi & Demolin (2009), there was a peak in nasal airflow that coincided with the target of the glide articulation in the nasal airflow configuration. In the spectrogram, we infer that the absence of the nasal appendix murmur became apparent when the glottis opened, resulting in the lack of resonances in the spectrum.

Based on this observation, we infer that the closure of the velum occurred simultaneously with the opening of the glottis, followed by the release of the subsequent consonant. As the nasal cavity has different characteristics in each

speaker, it may influence different formant configurations. In addition, the frequencies may vary according to the degree of velum lowering and the size of the oral cavity.

Although we did not find a formant pattern in the nasal appendix, we noticed that both the front nasal vowel and the back nasal diphthong had the same nasal airflow shape that was produced in both stop contexts next to /t/ and /k/.

The nasal appendix is a segment that has more resonances in the low-frequency region between 250 Hz and 500 Hz. There is an increased presence or emphasis of resonant frequencies within the low-frequency region of 250 Hz to 500 Hz. These resonances can be observed as peaks or formants in the spectral analysis of the nasal appendix. These resonances modify the F0 values and reinforce their energy, which indicates a higher concentration of energy or amplitude within that frequency range, suggesting a stronger acoustic presence of nasalization in that segment.

Thus, we noticed different regions of energy concentration in the nasal appendix amongst the speakers. Consequently, we did not extract the frequencies because it was difficult to measure the damped nasal frequencies by adding the mask resonance.

According to Teixeira (2007), in European Portuguese, a full nasal consonant can occur in the same context as the voiceless stop consonant /p/. An example of such a word is "campo" /'kãpu/ which is transcribed as ['kã̃pu]. In this word, the nasal consonant /m/ appears before the voiceless stop /p/.

In Brazilian Portuguese (BP), the nasal murmur assimilates the articulatory point of the glide, in European Portuguese, there may be degrees of nasalization depending on the context. This comparison is being made between the two varieties of Portuguese to highlight the differences in the degree and assimilation of nasalization in specific phonetic contexts.

The nasal appendix is the reflex of the vocalic nasalization that generates a sound element with a nasal murmur, which is a characteristic typically associated with nasal consonants. This production is the result of the vibration of the vocal folds when there is oral constriction and velum displacement.

The constriction blocks the oral cavity and drives the airflow to the nostrils. Therefore, the nasal diphthongization produces the nasal glide /j/ by producing the oral constriction in front nasal vowels.

The velum articulation has its dynamics (open and closing movement) which produces changes in the oral geometry, associated with tongue elevation resulting in nasal diphthongization. The nasal appendix results from a nasal coarticulation that assimilates the nasal vowel and glide region.

The acoustic nature of the nasal glide can be nasal because there is an increase in the AFn in its acoustic portion, which refers to a more significant velic opening, while there is an influence of the negative AFn and the rising AFn in the vowel, which can make the sound oralized in the begging, or nasalized by assimilation of the nasal feature. This depending if there is an anticipatory<sup>61</sup> or carryover<sup>62</sup> movement of the velic opening.

### **3.4.8 Duration and Nasality Rate**

---

<sup>61</sup> Anticipatory Nasalization: In this context, the nasalization of a vowel occurs in anticipation of a following nasal consonant. The vowel preceding the nasal consonant becomes nasalized to match the upcoming nasal quality. For example, in the word "camp" /kæmp/, the vowel /æ/ is nasalized in anticipation of the nasal /m/ sound.

<sup>62</sup> Carryover Nasalization: In this context, the nasalization of a vowel persists or carries over from a preceding nasal consonant. The vowel following the nasal consonant retains the nasal quality. For example, in the word "run" /rʌn/, the vowel /ʌ/ is nasalized due to the influence of the preceding nasal /n/ sound.

In this section, we measure the duration parameter to compare how nasality influences the acoustic temporal behavior of the target segments analyzed in this thesis. From a temporal perspective, we confirm how the aerodynamic data behave through the temporal trajectory of the nasal airflow.

We emphasize that the acoustic durational representation of speech is not equivalent to the durational representation of airflow. We observed a variation between the vocalic segments in both groups of nasal diphthongs, the vowel and nasal glide. Table 6 shows the duration values<sup>63</sup> for each segment.

Table 6: Absolute acoustic duration values for each segment. All gender groups.

Absolute Acoustic Duration ( <i>ms</i> ). N =36.									
		Average (SD)							
		Segment							
	Speaker	[aw]		[ãw̃]		[ej]		[ẽj]	
Male	BH	254	(17)	244	(10)	231	(14)	250	(17)
	LZ	286	(34)	334	(26)	245	(18)	327	(34)
	OD	253	(13)	246	(20)	207	(18)	256	(24)
	<b>Total</b>	<b>264</b>	<b>(28)</b>	<b>295</b>	<b>(31)</b>	<b>228</b>	<b>(23)</b>	<b>278</b>	<b>(43)</b>
Female	RD	279	(15)	324	(26)	290	(16)	322	(23)
	PL	234	(20)	275	(25)	206	(10)	232	(18)
	LC	276	(13)	285	(19)	238	(19)	274	(30)
	<b>Total</b>	<b>263</b>	<b>(26)</b>	<b>295</b>	<b>(31)</b>	<b>245</b>	<b>(37)</b>	<b>276</b>	<b>(44)</b>

The average acoustic duration<sup>64</sup> of the total production of the target segments, to back nasal diphthongs was 295 ms; for the front vowel, the total duration was around 276 ms. Thus, the front nasal vowel reduced the acoustic duration compared to the back-nasal diphthong. The graphs below compare the absolute duration values (ms) to compare the oral and nasal outputs of the target segments.

<sup>63</sup> The acoustic analysis was made in Signal Explorer, which allow us the examination of spectrograms and waveforms. We manually segmented the speech signal based on changes in formants and energy, as explained in item Segmentation.

<sup>64</sup> The duration was measured by calculating the time difference between the start and end points of the segment.

Chart 1: Absolute acoustic duration (ms). Comparison between the target segment's production and the male group (N =36).

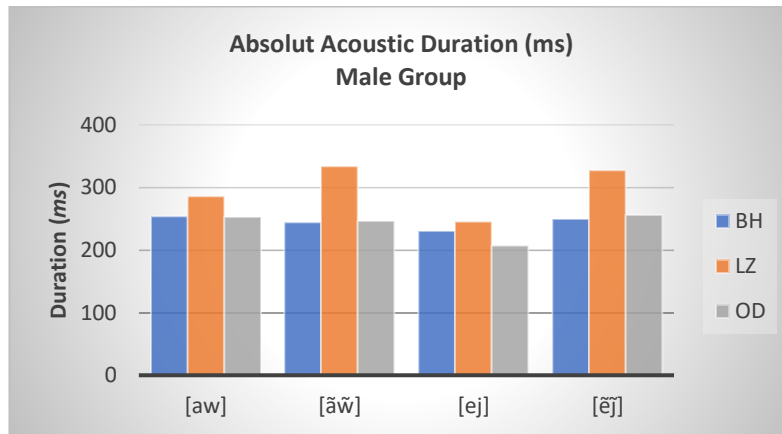
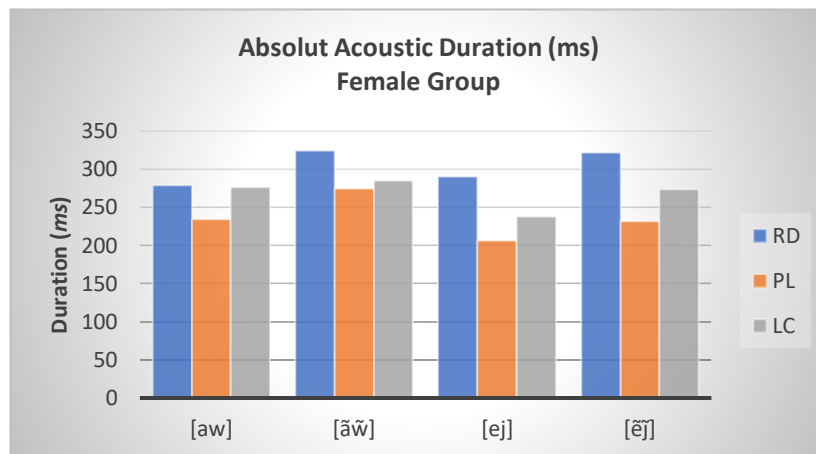


Chart 2: Absolute acoustic duration (ms). Comparison between the target segment's production and the female group (N =36).



Moraes & Wetzels (1992) pointed out that sequences with nasal elements were acoustically longer; phonologically, they would be a bimoraic syllable. Thus, the second mora is a floating segment with place traces.

The concept of mora is a unit of timing or weight in phonology, and it can have implications for syllable structure and rhythm in languages. The claim that sequences with nasal elements are acoustically longer and phonologically represent a bimoraic syllable suggests a correlation between nasal segments and increased duration.

In our data, the nasal segments were longer than the oral segments, the magnitude of this difference may not necessarily provide strong evidence for the theoretical proposal of an additional mora.

To support the theoretical proposal of an additional mora, it would be necessary to consider other phonological and prosodic factors in addition to duration, such as stress patterns, syllable structure, and rhythmic patterns within the language.

The oral front diphthong had an average of 228 ms and 245 ms for males and females, respectively. The nasal front diphthong had an average of 278 ms and 276 ms for males and females, respectively.

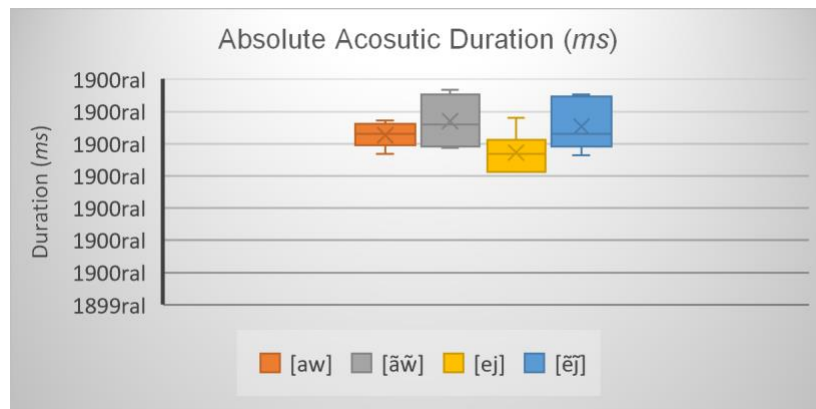
The oral back diphthong had an average of around 264 ms for both gender groups. The nasal back diphthong had an average of around 295 ms for both gender groups.

Our results showed interesting patterns, in terms of average durations for specific gender groups. The nasal back diphthong demonstrated a longer duration for both gender groups.

To visually compare the duration differences, Chart 3 was provided, illustrating the absolute acoustic durations (in milliseconds) for the target segments' production. This chart allowed for a clear comparison between the durations of the target segments in both their oral and nasal forms.

Chart 3: Absolute acoustic duration (ms). Comparison between the target segment's production and both target's productions (N =36).





In the total average values of all the target segments, the nasal segments were longer than the oral segments. The nasal segments may have been longer than the oral ones because of the position of the unpacking<sup>65</sup> nasal vowel (PARADIS & PRUNET, 2000) resulting in the glide in the coda (semi-consonant).

We divide the acoustic duration of the nasal diphthongs into three parts based on changes in the spectrogram for analyzing the temporal characteristics of different components within the nasal diphthongs.

(a) The acoustic temporal duration of the nasalized vowel (AdVn) represents the duration of the vowel segment during the nasal diphthong, where nasalization is present.

(b) The acoustic temporal duration of the nasal glide (AdGn) represents the duration of the glide segment during the nasal diphthong.

(c) The acoustic temporal duration of the nasal appendix (AdAPn) represents the duration of the nasal appendix, which is the segment following the nasal glide in the nasal diphthong.

---

<sup>65</sup> Unpacking is breaking down the properties of a segment into individual segments. A widely known example of unpacking is separating nasal vowels into vowels and nasal consonants. So, this phenomenon is represented as  $\tilde{V} > VN$ . This is a way to keep the [+nasal] feature in the segment (PARADIS & PRUNET, 2000). In our case, nasal vowels can be represented as a sequence of a nasal vowel followed by a nasal glide, maintaining the underlying [+nasal] feature of the segment.

The points (a), (b), and (c) represent acoustic temporal durations. They are intended to capture the duration of specific acoustic components within the nasal diphthongs, following the segmentation provided previously.

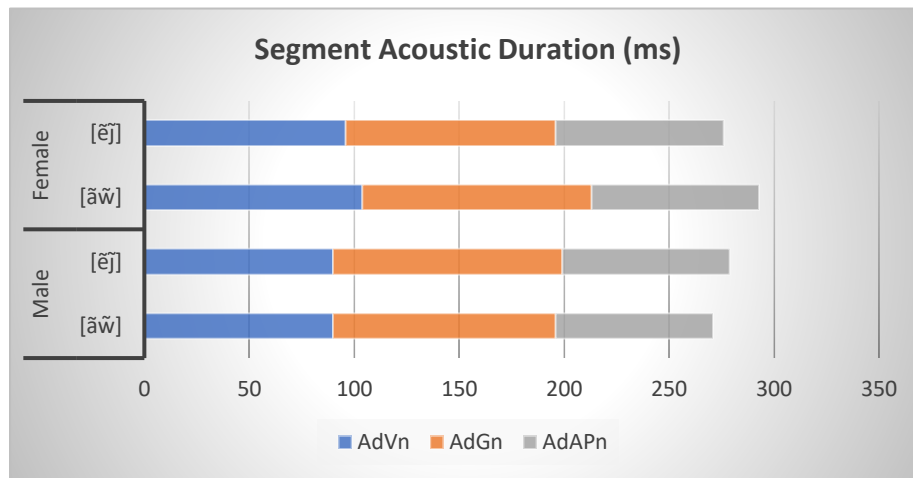
Table 7: Segment acoustic duration values for the nasal segment: AdVn, AdGn, and AdAPn. All gender groups.

Segment Acoustic Duration (ms). N =36.							
		Average (SD)					
		AdVn		AdGn		AdAPn	
Segment	Speaker						
[ãw̃]	BH	86	(12)	84	(29)	74	(47)
	LZ	110	(15)	127	(32)	88	(18)
	OD	75	(30)	108	(24)	63	(22)
	<b>Total</b>	<b>90</b>	<b>(19)</b>	<b>106</b>	<b>(28)</b>	<b>75</b>	<b>(29)</b>
[ẽj̃]	BH	86	(15)	76	(20)	88	(23)
	LZ	117	(54)	125	(25)	85	(21)
	OD	76	(16)	110	(28)	70	(20)
	<b>Total</b>	<b>93</b>	<b>(25)</b>	<b>109</b>	<b>(20)</b>	<b>80</b>	<b>(20)</b>
[ãw̃]	RD	127	(18)	110	(18)	87	(18)
	PL	86	(20)	104	(17)	83	(25)
	LC	100	(38)	114	(25)	71	(17)
	<b>Total</b>	<b>104</b>	<b>(25)</b>	<b>109</b>	<b>(20)</b>	<b>80</b>	<b>(20)</b>
[ẽj̃]	RD	128	(15)	107	(15)	87	(19)
	PL	72	(23)	86	(17)	74	(20)
	LC	87	(16)	108	(34)	79	(17)
	<b>Total</b>	<b>96</b>	<b>(18)</b>	<b>100</b>	<b>(22)</b>	<b>80</b>	<b>(19)</b>

The back-nasal vowel accounted for 34% and 36% of the total segment, while the back-nasal glide was 38% and 36% for male and female speakers, respectively. The nasalized front vowel accounted for 34% and 36% of the total segment, while the glide was 38% and 36% for male and female speakers, respectively.

Nevertheless, there was not the same variation in duration in the nasal appendix. The nasal appendix accounted for 28% of the acoustic production of the nasal diphthong for both groups, as shown in the following chart.

Chart 4: Average of the segment acoustic duration (ms). Comparison between the target segment's production in the male and female groups (N =36). Values for the nasal segment: AdVn, AdGn, and AdApn.



From the total average of the nasal diphthong's values for the two groups of subjects, the nasalized vowel [ã] had an average duration of 90 ms and 104 ms for men and women, respectively. The nasal glide [j] had an average duration of 109 ms and 100 ms, for men and women, respectively. The nasal appendix had an average duration of 75 ms and 80 ms for men and women, respectively.

From the total average of the values for the two groups of subjects for the diphthongized front vowel, the nasalized vowel [ẽ] had an average duration of 93 ms and 96 ms for men and women, respectively. The nasal glide [ũ] had an average duration of 106 ms and 109 ms for men and women, respectively. The nasal appendix had an average duration of 80 ms for men and women, respectively.

The acoustic duration of the nasal vowel was AdVn 33% and 36% in the segment to [ãũ] for males and females, respectively. The AdGn was 39% and 37% for the respective gender groups, and the AdApn was 28% and 27%, respectively.

The acoustic duration of the nasal vowel was AdVn 32% and 35% of the segment to [ẽĩ] for males and females, respectively. The AdGn was 39% and 36% for the

respective gender groups, and the AdApn was 29% and 29%, respectively. Thus, the proportions were extremely similar.

The nasal appendix is a phonetic phenomenon characterized by the presence of a nasal consonant-like element at the end of a nasal vowel. It occurs when there is a release of nasal airflow following the closure of the oral articulators.

In our case, the cause of the nasal appendix can be attributed to the coarticulatory effects between the nasal glide and the following consonant. During the production of a nasal glide, the velum is lowered, allowing airflow to pass through the nasal cavity.

When transitioning from the nasal glide to a consonant, the velum begins to close, blocking the nasal airflow. However, due to the inherent temporal overlap of speech gestures, there can be a brief period where the velum is in the process of closing, leading to the release of nasal airflow and the production of the nasal appendix.

Sousa (1994) found values ranging from 40 to 100 ms, while Jesus (2002) obtained 10 to 65 ms for the nasal murmur<sup>66</sup>. Thus, we infer that our data showed the nasal spreading too. After glide constriction in the oral cavity, the folds continued to vibrate at an average of 80 ms, and we considered the duration of the production of the nasal appendix.

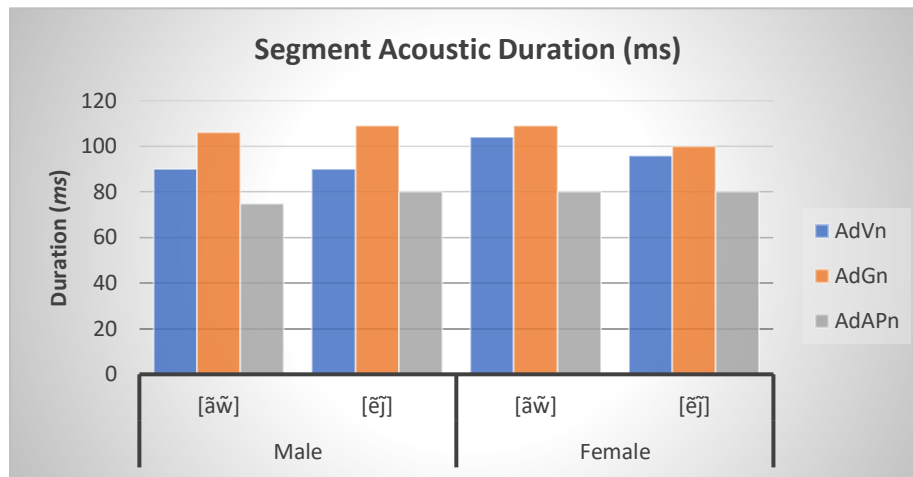
This value may vary depending on factors such as speech rate, phonetic context, and individual articulatory patterns that can influence the occurrence and characteristics of the nasal appendix; for more details, see Medeiros (2007). Chart 5

---

<sup>66</sup> The nasal murmur, also known as the consonant feature, consonant segment, or nasal segment, is characterized by a low-frequency formant that occurs at the end of a nasal vowel, following the articulatory closure. The nasal murmur exhibited a low-frequency characteristic and did not display spectral information indicative of a transition to the subsequent consonant (JESUS, 2002).

is the same as Chart 4 but emphasizes each element's values for the target nasal segment.

Chart 5: Average of segment acoustic duration (ms). Comparison between the target segment's production in the male and female groups (N =36). Values for the nasal segment: AdVn, AdGn, and AdAPn.



Thus, we can see that, in the average values for each target segment shown, the nasal glide had the most prolonged acoustic duration in all occurrences. In diphthongs, the first component was usually longer and more robust than the second.

However, as this was a nasal diphthongization phenomenon, the glide probably had the expansion of nasality and nasal resonances, and increased presence of nasal acoustic characteristics during the production of the nasal glide, such as nasal resonance and airflow through the nasal cavity.

The acoustic duration measure refers to the temporal displacement of the oral articulatory gesture. However, we can infer the opening and closing of the velum movement based on the aerodynamic data.

The airflow trajectory refers to the variation in the output volume in speech signals. Therefore, we also analyzed the behavior of the nasal airflow during production to predict the soft palate's gesture dynamics.

The velum activation began when the AFn was negative<sup>67</sup>. Despite this, we do not know how this influences the acoustic signal. The significant opening of the velum began when the AFn became positive.

There was a gradual increase of the AFn. When the AFn and the AFO intersected, the tongue had probably reached the glide target. Following this, the AFn had a high elevation in the nasal appendix region.

The nasal peak is a characteristic of production. The last phase of the velum gesture is the closure. At this point, the Afn became negative. The table below shows the time values extracted from the nasal airflow (ms).

Table 8: Duration of the nasal airflow shape. Values for the nasal segment: Total duration of the velic trajectory (Tdvt) (ms). Velic opening duration (Vod) (ms). Duration of glide constriction (Dgc) (ms). Velic rising duration (Vrd) (ms) in the male and female groups.

Duration of Nasal Airflow Shape (ms). N =36.									
Average (SD)									
Segment	Speaker	Tdvt		Vod		Dgc		Vrd	
[ãw̃]	BH	291	(86)	122	(48)	88	(58)	81	(14)
	LZ	299	(70)	124	(42)	85	(64)	90	(38)
	OD	247	(43)	123	(31)	50	(18)	72	(13)
	<b>Total</b>	<b>279</b>	<b>(66)</b>	<b>123</b>	<b>(40)</b>	<b>74</b>	<b>(47)</b>	<b>81</b>	<b>(22)</b>
[ẽj̃]	BH	360	(66)	182	(43)	97	(36)	81	(15)
	LZ	291	(32)	107	(53)	95	(38)	89	(31)
	OD	276	(48)	121	(17)	69	(21)	86	(24)
	<b>Total</b>	<b>309</b>	<b>(49)</b>	<b>137</b>	<b>(38)</b>	<b>87</b>	<b>(32)</b>	<b>85</b>	<b>(23)</b>
[ãw̃]	RD	201	(25)	100	(17)	46	(16)	55	(12)
	PL	272	(37)	139	(29)	65	(18)	68	(14)
	LC	261	(50)	129	(38)	75	(24)	57	(12)
	<b>Total</b>	<b>245</b>	<b>(37)</b>	<b>123</b>	<b>(28)</b>	<b>62</b>	<b>(19)</b>	<b>60</b>	<b>(13)</b>
	RD	311	(28)	141	(51)	108	(41)	62	(10)

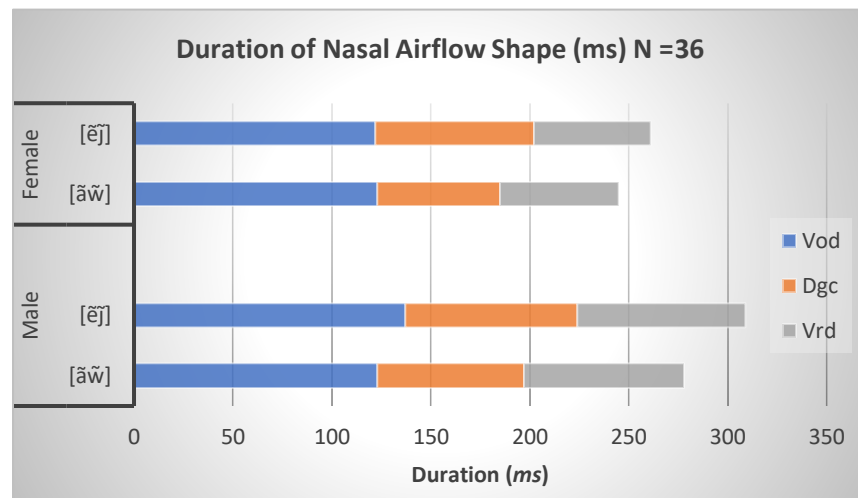
<sup>67</sup> The negative airflow (AFn) represents the absence or minimal nasal airflow, indicating that there is no nasal airflow passing through the nose. On the other hand, positive nasal airflow indicates the presence of nasal airflow passing through the nose. When velum activation starts, there is a transition from the default position of the velum (closed or raised) to a lowering movement, allowing for nasal airflow. This results in an ingressive nasal airflow, where air enters through the nose.

[ẽ]	PL	226	(27)	103	(15)	55	(20)	68	(20)
	LC	245	(31)	121	(41)	78	(18)	46	(9)
	<b>Total</b>	<b>261</b>	<b>(29)</b>	<b>122</b>	<b>(36)</b>	<b>80</b>	<b>(26)</b>	<b>59</b>	<b>(13)</b>

The average of the absolute duration of the Afn trajectory (positive values) to [ãw̃] was 279 ms and 245 ms for males and females, respectively. The average of the absolute duration of the Afn trajectory to [ẽj̃] was 309 ms and 261 ms for males and females, respectively. These values represent the linear trajectory of the duration of the nasal airflow.

The entire trajectory of the velum can also be divided into three phases, which are the opening phase, the maximum aperture, and the closing phase. The velum begins the lowering movement in the initial phase; in the medial phase, the velum remains open (in the glide constriction) and the velum begins the closing movement in the final phase. The chart below presents the temporal values for the three phases of the velum displacement (ms).

Chart 6: Average duration of the nasal airflow shape. Values for the nasal segment: Velic opening duration (Vod) (ms). Duration of glide constriction (Dgc) (ms). Velic rising duration (Vrd) (ms) for the male and female groups. N= 36.



The temporal nasal airflow shape was **Vod** 44% of [ãw̃]; **Dgc** accounted for 27% and Vrd for 29% in the male group. To [ẽj̃], **Vod** accounted for 44%, **Dgc** for 28% and

Vrd for 28%. The temporal nasal airflow shape was **Vod** 50% of [ãw̃], **Dgc** was 25% and **Vrd** was 25% in the female group. To [ẽj̃], **Vod** was 47%, **Dgc** was 31% and **Vrd** was 22%.

It is interesting to note that, although the duration was different, the temporal proportion of the articulatory gesture of the velum was quite similar in the two groups of speakers. Both groups' opening gestures were slower than were the closing gestures (DELVAUX, 2003; AMELOT, 2004; DEMASI & DEMOLIN, 2009). Therefore, the opening velum trajectory accounted for more than 40% of the total velic displacement.

The table below shows the average values for the maximum rate of the nasal airflow during the target production.

Table 9: Volume variation in nasal airflow shape. Values for the nasal segment in high NP (dm<sup>3</sup>/s).

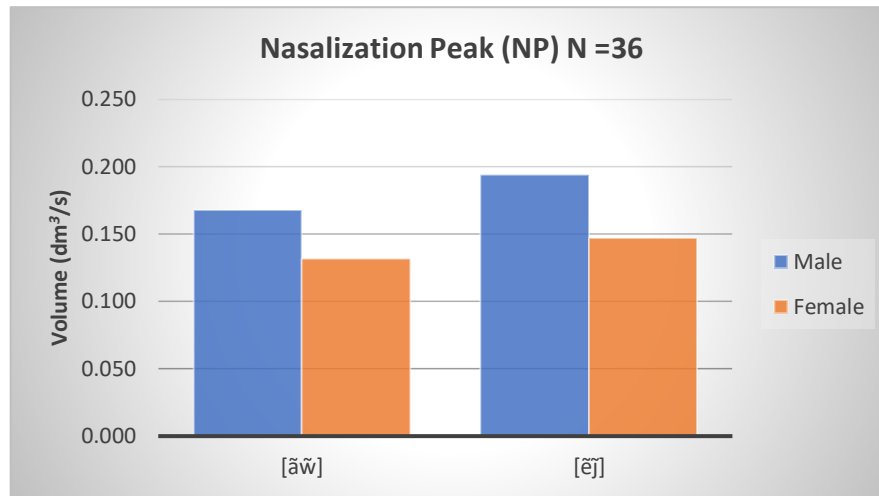
NP (dm <sup>3</sup> /s). N =36.				
Average (SD)				
Speaker	Segment			
	[ãw̃]	[ẽj̃]		
<b>BH</b>	0.109 (0.33)	0.133 (0.32)		
<b>LZ</b>	0.185 (0.67)	0.209 (0.68)		
<b>OD</b>	0.215 (0.67)	0.252 (0.68)		
<b>Total</b>	<b>0.168 (0.56)</b>	<b>0.194 (0.56)</b>		
<b>RD</b>	110 (0.25)	0.139 (0.61)		
<b>PL</b>	144 (0.68)	0.143 (0.49)		
<b>LC</b>	141 (0.32)	0.162 (0.30)		
<b>Total</b>	<b>132 (0.42)</b>	<b>0.148 (0.47)</b>		

Initially, we noticed that the male group had a higher nasal airflow rate than the female group; Delvaux (2003) also found this variation in French nasal vowels. To [ãw̃], the NP had an average of 0.168 dm<sup>3</sup>/s and 0.132 dm<sup>3</sup>/s in the male and female



groups, respectively. To [ẽĩ], the average was 0.194 dm<sup>3</sup>/s and 0.148 dm<sup>3</sup>/s, respectively.

Chart 7: Average duration of the nasal airflow shape. Values for nasal segment NP (dm<sup>3</sup>/s) for the male and female groups.



We compared the rate of volume variation in the nasal airflow; in the male group, the [ãũ] was 8% higher than [ẽĩ]. In the female group, it was 6%. The general NP had more volume in the males' production than it did in the females' production. Comparing the production of the NP by females and males revealed that the NP was 12% higher for [ãũ] and 14% higher for [ẽĩ].

The NP in our data had a high-volume peak. Medeiros, D'Imperio & Espesser (2008, p. 131) found an average to high nasal airflow peak in a nasal vowel; the value of the nasal appendix nasal rate was much larger than it was in a nasal consonant, which was 0.060 dm<sup>3</sup>/s. The authors' results revealed that the nasal appendix had an average of 0.090 dm<sup>3</sup>/s.

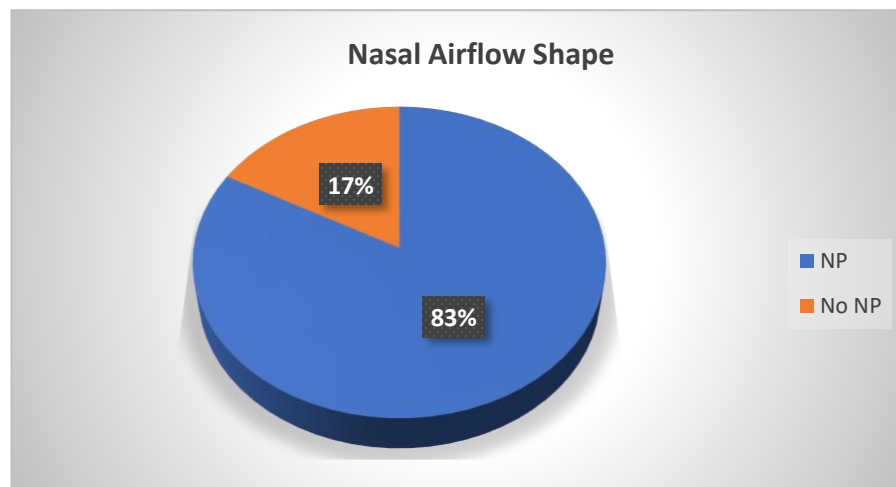
We note that characterizing the duration of the oral and nasal airflows is essential for understanding the velum dynamics during diphthongization because it provides insights into the coordination and timing of articulatory movements.

We inferred three distinct phases of the velum movement from the aerodynamic data, as reflected by the variations in the positive nasal airflow volumes. The temporal duration of the oral and nasal airflows provides insights into the coordination and timing of articulatory movement of the velum dynamics during the process of diphthongization.

The velum opening movement was longer than the closing movement in the initial phase. The high peak of nasal air occurred during the production of the nasal appendix.

In our data, 83% of the occurrences had the Afn pattern of high nasal airflow (DEMASI & DEMOLIN, 2009). Thus, most of the data had a high peak at the end of the articulation, while the Afn suffered some system disturbance in 17% of the cases, with the result being no high jet in the nasal airflow.

Chart 8: Occurrences of NP. The 3-phase standard for Afn (N=120).



The three phases of the velum trajectory are associated with articulatory movements in the oral tract and the coupling of the oral and nasal cavities, resulting in the high nasal airflow. The synchronization of the velum and the tongue movement

causes a modification of the volume in the oral cavity and changes the quality of the sound output.

### **3.5 Partial Conclusion**

Our data demonstrated that acoustic durations were distinct from aerodynamic durations, as airflow is modified due to coarticulation processes. In our data, the  $A_{fn}$  became positive when the degree of the velopharyngeal opening was significant at the end of the acoustic production of the vowel.

Nevertheless, the influence of nasal resonance on the spectrogram was noted from the middle portion of the vowel. In this process the vowel sound becomes more similar to a neighboring sound affecting the middle portion of a vowel, causing it to become nasalized.

The nasal quality of the neighboring sound influences the vowel, resulting in nasalization of the vowel's acoustic characteristics. The assimilation can be observed in the spectrogram by changes in the formant values, reduction in wave amplitude, and widening of the bandwidth (spread of energy across frequency). In oral diphthongs, the formants are more uniform than in nasal segments, nasal resonances aren't visible in the spectra and amplitude is more uniform.

Changes in the trajectory of the soft palate have a non-linear effect on the amount of nasal airflow. Due to the assimilatory phonetic process, in which sounds become more similar to neighboring sounds in terms of their articulatory or acoustic properties. We cannot say that there is always an overlap between resonances and anti-resonance, as the overlap depends on each case (KRAKOW & HUFFMAN, 1993).

So, assimilation changes the quality of a neighboring sound affecting the surrounding vowel, causing it to become nasalized. The assimilation of nasality is gradual and results in changes to the formant values, amplitude, and bandwidth of the vowel sound.

In our nasal production data, the vowel is indeed nasalized, due to the influence of nasal resonance, characterized by formant structures associated with the nasal cavity. This influence is observed in the spectrogram through changes in formant values, amplitude, and bandwidth. Nasal resonances influence on all nasal diphthongs, the vowel is nasalized, the glide is nasal, and also produces the nasal appendix.

The vowel was nasalized because it was influenced by nasal resonance, with a decrease in the amplitude and an increase in the bandwidth, which consumed sound energy, and a slight dampening of the peaks of the wave (FUJIMURA, 1962).

The nasal quality of the glide is primarily determined by its articulatory configuration and the lowered velum. We infer that the velum is in its lower position, which allows more airflow volume to pass through the nasal cavity during its production. The effects of anti-resonances further contribute to the nasal quality by shaping the spectral characteristics of the sound.

The relationship between the geometry of the oral tract, the degree of constriction in the oral cavity, and velum lowering are associated with the production of nasalization. When there is a slight constriction in the oral cavity and an opening of the velum, it allows for a greater coupling between the oral and nasal cavities. This coupling enhances the perception of nasalization by influencing the resonance properties of the sound, resulting in a greater presence of nasal resonances and characteristics in the acoustic signal.

The high nasal airflow peak in the nasal appendix, as mentioned by Cagliari (1977) and corroborated with our data, indicates that there is a concentration of energy in the low-frequency ranges associated with this nasal segment. This can be attributed to the coupling between the oral and nasal cavities during nasalization, leading to the specific acoustic characteristics observed in the nasal appendix.

The volume of air passing through the nasal cavity during nasalization is influenced by both the size of the velopharyngeal opening and the geometry of the oral tract (KRAKOW & HUFFMAN, 1993); hence, it is possible to draw inferences about the duration of the velum opening and closing.

The size of the velopharyngeal opening directly affects the amount of air that can flow through the nasal cavity. A larger opening allows for a greater volume of nasal airflow. However, the geometry of the oral tract, including the degree of constriction in the oral cavity, also plays a role in determining the volume of nasal airflow.

The shape and size of the oral cavity, as well as the position and movement of the articulators, contribute to the overall airflow dynamics and can impact the amount of air that enters the nasal cavity.

The nasal airflow can provide valuable information about the duration of the velopharyngeal opening and closing during speech production. By analyzing the nasal airflow pattern, we inferred the timing and duration of the velopharyngeal opening and closing gestures.

When the velopharyngeal opening begins, the nasal airflow increases, indicating the initiation of the opening movement. The duration of the positive nasal airflow corresponds to the period when the velopharyngeal port remains open, allowing air to pass through the nasal cavity. As the velopharyngeal port starts closing, the nasal

airflow decreases, indicating the closure of the port. When nasal airflow is negative (under the zero), it represents the period of velopharyngeal closure.

By examining the temporal characteristics of the nasal airflow, such as the duration of positive and negative phases, we gained insights into the timing and coordination of the velopharyngeal opening and closing gestures. This information helped us in understanding the velopharyngeal function during speech production of our data and its relationship to nasality.

Therefore, it is important to consider both the size of the velopharyngeal opening and the geometry of the oral tract when discussing the volume of air passing through the nasal cavity during nasalization. These factors work together to influence the nasal airflow and contribute to the overall perception of nasality in speech.

When analyzing the  $A_{fn}$ , we found a dynamic model for the trajectory of the aerodynamic trace for the BP diphthongs for [ãw̃] and [ẽj̃]. In this model, we related the amount of air volume expelled by the nostrils to the articulation mode of the active articulators (such as the tongue and velum) during the coordination and interaction of articulatory movements that occur during the production of speech sounds, which in turn affects the resulting acoustic properties of the sounds. Figure 45 shows the  $A_{fn}$  pattern for both segments.

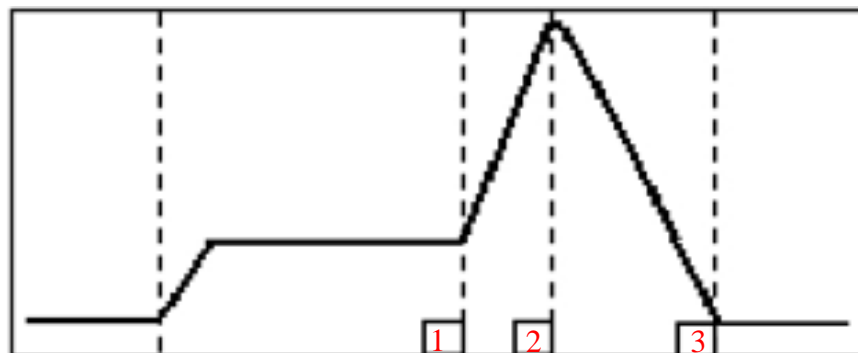


Figure 45: Nasal Airflow Shape. Representation of the  $A_{fn}$  pattern for our nasal segment data for both groups of speakers.

The nasal airflow shape represents the three phases of velum movement. The velum begins lowering in (1); (2) indicates the constriction of the oral tract and (3) indicates the velum closing.

The maximum opening amplitude can be inferred from the temporal duration of phase 1 to phase 2. We infer that phase 2 is related to the high velum opening gesture when the velum reaches its target position. When the velum closes, the articulator returns to its default position, as indicated in phase 3.

The maximum opening of the velum can be inferred when the nasal airflow reaches a stationary and uniform movement. The duration of the nasal airflow, in this context, refers to the temporal span of the velum movement rather than the maximum opening amplitude. The duration provides information about the time it takes for the velum to complete its movement from a closed position to an open position.

It is important to note that the maximum velum opening amplitude can vary depending on factors such as the specific segment being produced or the phase of the breath cycle. The duration of the nasal airflow, however, can provide insights into the overall temporal characteristics of the velum movement during the production of nasal sounds.

In our data, we found a global organization between the tongue and the velum movement, despite the temporal asynchrony observed in reaching their respective targets. When the velum reaches its target, the velum is already at its maximum opening position, but the tongue has not yet reached its target to glide production.

It suggests that the timing of onset and target gestural movements of the velum and tongue are not synchronized temporally. But there is a synchronization between the tongue and velum movement, because when there is a glide constriction in the oral tract, the articulatory movement of velum closing starts. At this moment, when the

tongue reaches its degree of constriction results in an increase in the nasal airflow volume.

Thus, as the constriction caused by the tongue is the smallest, the greater the volume of airflow that must pass through the nasal cavity, and the opposite is also true. Therefore, the jet of airflow is the result of the constriction in the oral cavity and not of the size of the opening in the velopharyngeal section (DEMASI & DEMOLIN, 2009).

According to your findings, the volume of airflow passing through the nasal cavity is influenced by the degree of oral constriction rather than the size of the opening in the velopharyngeal section.

These findings align with the notion that the overall coordination between the tongue and velum movements plays a role in regulating nasal airflow during speech production.

The specific timing and coordination of these movements can vary among individuals and contexts. So, our study contributes to the understanding of the complex interplay between articulatory gestures and airflow dynamics during the production of nasal sounds.

As the vowel becomes more closed (higher in the vowel space), the rate of nasalization tends to increase, and vice versa. This relationship suggests a proportional relationship between vowel height and the amount of nasal airflow. This observation is based on the work of Cagliari (1977).

In the two groups of segments chosen for this analysis, nasal diphthongization in the São Paulo dialect was characterized by a formation of a constriction in the glide portion. This constriction was mainly evident in the acoustic data via the movement of formants.



In the spectrum of front nasal vowels, as well as in the point of articulation of the glide in the palatal region and the back nasal diphthong in the velar region, there is an accentuated upward shift or increase in the formant values in the higher frequency ranges. This implies that the formants associated with those higher frequencies exhibit acoustic energy more prominent in these segments. The increase in formant values in the higher frequency ranges suggests specific acoustic characteristics of nasal diphthongization.

The process of nasal diphthongization involves the transformation of an oral vowel into a nasalized vowel and nasal glide segments. This occurs when oral vowels are influenced by adjacent nasal segments but remain oral due to phonological distinctiveness, although they have the option to become nasal (MEDEIROS, 2007). The nasalization of vowels preceded by nasal consonants is explained by the gradual and controlled movements of the velum (CLUMECK, 1976).

Aerodynamic data from Amelot & Michaud (2006) and articulatory data from Oliveira & Teixeira (2007) support the notion that the closing of the velum overlaps with the following articulatory gesture, resulting in partial nasalization of the latter gesture. These findings shed light on the specific articulatory movements involved in this process.

In our data, we observed that the elevation of the tongue creates a constriction in the vocal tract, allowing for increased nasal airflow into the nasal cavity. This constriction contributes to the production of the nasal glide segment.

Additionally, the closing movement of the velum plays a role in the production of the nasal appendix. These coordinated articulatory actions are responsible for the formation of nasal diphthongs. It is important to clarify that although the tongue

movement is controlled with a degree of constriction, it is not directly determined by the phonological structure of the language.

## CHAPTER 3: FIBERSCOPE ANALYSIS

---

### Résumé (Summation)

This chapter analyzes the velum movement in nasalized diphthongs in BP. The study was conducted using a nasofiberscope to study the size and the movements of the velopharyngeal opening during the production of monosyllabic words containing the oral and nasal diphthongs [ej], [ẽj], [aw] and [ãũ]. The data showed that the velum lowering was gradual, and the velum section movements could be observed in the velopharyngeal opening thought (PNG). These data confirm previous observations showing that there are several phases in nasalized BP diphthongs. The velum lowering has three different phases: onset, target and offset. The findings showed that the anticipatory velum lowering was variable, while the closing movement was stable, and the nasal appendix appeared in all

### 4. Introduction to Fiberscope Analyses

This chapter presents a physiological study of velopharyngeal articulation in BP speech production. Thus, this research analyzed nasofiberscope<sup>68</sup> photonasograph images. This data pertains to typologically rare and understudied nasal diphthongs in BP.

This type of data is complex and relevant to the field, due to its uniqueness it contributes to advancing knowledge and understanding about nasalization and velopharyngeal port movement. However, the complexity of collecting such data limits the amount of data available for studying nasality and nasal diphthongization because the methodology is invasive.

The photonasograph can monitor the opening of the velopharyngeal port, which is a passageway between the oral and nasal cavities using light variation Ohala

---

<sup>68</sup> Fibro nasopharyngoscopy is a medical procedure that allows for the visualization of bodily regions to diagnose diseases in the nasal cavities, rhino pharynx, oropharynx, hypopharynx and supraglottic, glottic and subglottic larynx. It is an invasive examination, and some people may experience discomfort.

(1971). The light is utilized to gather information about the status or position of the velopharyngeal port. This light is then converted into an electrical signal, specifically voltage, for further analysis or processing.

This chapter provides videos and pictures of the velopharyngeal shape, as well as the general characteristics of the velum's dynamics and its trajectory during the articulatory movement in the speech signal.

We followed the structure and methodology of Amelot (2004), who studied the same phenomenon in French nasal vowels. Amelot (2004) conducted a study in her thesis on the movement of velum and pharyngeal walls during speech production.

Amelot (2004) conducted a study on the acoustic segmentation of nasal vowels and examined the timing of velum lowering and wall opening concerning nasal vowels. The study measured the time delay between the onset of velum movement and the acoustic onset of the vowel and the time delay between the onset of wall movement and the acoustic onset of the vowel.

The study confirmed that the velum begins to open before the start of the nasal vowel and takes approximately 50 ms to reach the maximum velar opening. The movements were less progressive than regressive, with the velum closing before the end of the nasal vowel in some cases. The degree of velar opening varied among nasal vowels and was influenced by the surrounding consonantal environment. The findings supported the hypothesis that articulatory constraints play a significant role in velar movements during nasalization.

She also investigated the movements of the velum and the hypothesis that the velum takes longer to open than to close. The opening movement involved measuring the distance between the beginning of the phoneme and the maximum velar opening, while the closing movement involved measuring the distance between the end of the

phoneme and the maximum velar opening. The results showed that the velum takes less time to open than to close.

The study provided insights into the timing and coordination of velum and wall movements during nasal vowels, as well as the duration patterns of nasal phonemes in different linguistic contexts.

The results showed that velum movements start before the target vowel and end after, similar to wall movements. The average measurements indicated that wall movements end at the same time as velum movements.

The study revealed a relatively good synchronization between velum movements and wall movements. The percentage of cases where velum movement synchronized with wall movement was examined.

The study discussed the implications of velum opening and wall opening on the closure of the velopharyngeal port. The findings confirmed that the velum is lower for nasal vowels compared to oral vowels, and it is lower for open vowels compared to closed vowels, which aligns with the literature. Regarding intra-speaker variability, the study agreed with McKern et al.'s (1970) findings that the variability in velopharyngeal port closure force was influenced by the speaker's sex.

Niimi et al. (1982) found that wall movements were parallel to velum movements, indicating synchronous motion. The study confirmed that the walls move symmetrically with the velum. Comparisons between isolated and contextual data showed discrepancies with Moll & Daniloff (1971) but supported the results of Moon et al. (1994), where the data for each speaker were relatively similar in both isolated and contextual contexts. The study confirmed the presence of anticipation and propagation of velum and wall movements, with differences in opening and closing movements depending on the consonantal environment

Amelot also conducted a preliminary study on the movements of the velopharyngeal walls and their contribution to phoneme production. When the velum is open, there is a significant distance between the walls, as seen in nasal vowels, while the walls are close together when the velum is closed, and it was observed that their velopharyngeal wall movements varied in intensity.

Based on her data, the study concluded that wall movements are not essential for describing the opening and closing movements of the velopharyngeal port. The walls were considered complementary to velum movements and only amplified them. However, the study acknowledged the importance of the walls in pathological conditions such as hypernasality, where the deficiency of the velum could be compensated by pharyngeal movements.

The opening of the velum is influenced by tongue position, and her study confirms that the velum is lower for the vowel /a/ and the consonant // in spontaneous speech. Additionally, the maximum velum opening seems to play a significant role in velum movement.

Amelot's work provides insights into the temporal and spatial coordination of velum and pharyngeal wall movements during nasal vowel production. The findings highlight the importance of anticipatory movement over perseveration in articulatory coordination.

The opening of the velum is influenced by tongue position, and the study confirms that the velum is lower for the vowel /a/ and the consonant // in spontaneous speech. Additionally, the maximum velum opening seems to play a significant role in velum movement.

In Amelot's (2004) study, various findings related to velum movements were observed at the beginning of a sentence, there are multiple possible openings for the velum, whether the phoneme is oral or nasal.

The total duration of velum movement is longer than the acoustic duration of nasal vowels. There is no significant difference in the total duration of velum movement based on the consonantal context. More relaxed speech leads to larger velum movements.

Velum movement is more prominent for nasal vowels at the end of a word. In most cases, the velum starts opening before the acoustic onset of the nasal phoneme. Anticipation of velum movement is greater in relaxed speech.

The duration of velum movement closure is less important than the duration of anticipatory movement. It is also shorter than the duration of nasal airflow persistence. Closure movement duration is longer for phrases compared to non-words, but shorter in spontaneous speech compared to read phrases. Some instances were found in spontaneous speech where the velum closed before the acoustic completion of the nasal phoneme.

The timing of maximum velum opening and the onset of the nasal phoneme differs between isolated vowels and vowels in sequence. Maximum velum opening occurs before the nasal vowel when pronounced in isolation, but approximately halfway or in the first half of the nasal phoneme when in a sequence.

The duration of velum opening movement is roughly the same as the duration of velum closure movement, regardless of the nasal vowel, consonantal context, or speech style. The velum is more open when the nasal phoneme is in the final position of a word. In spontaneous speech, there is a compensatory effect where the velum is less open, but the articulatory movement is longer.

Amelot's (2004) analysis of velum movement traces showed that even for oral vowels, the velum undergoes slight variations in movement, with differences observed between low and high vowels. Nasal vowels exhibited clear opening movements, characterized by three phases: an opening phase, a maximum opening phase, and a closing phase.

Different patterns were observed for the beginning and end of the opening movement, which was simplified for analysis. The opening movement could start abruptly after a plateau, oscillate with small opening and closing movements before a significant closing movement, or oscillate with small movements and then open abruptly.

The opening movement had a regular pattern and could be segmented into three phases. The maximum opening could either transition immediately to the closing phase or maintain the maximum opening for a short duration before closing.

The velum was maximally open during breathing phases, and there were variations in opening during pauses, which were more significant than for nasal phonemes.

At the beginning of a phrase, the velum could be high before phonation, but typically, it opened before the first phoneme and continued closing, reaching maximum closure in the middle or end of the phoneme. At the end of a phrase, the velum started lowering before the acoustic offset of the last phoneme, returning to its original position to allow the speaker to resume breathing.

Amelot's (2004) study observed that the velum could be open for the vowel /a/ depending on the context. However, it found only one instance of a nasal vowel without a velum opening, specifically for the word "ben". The velum movement did not propagate as much as nasal airflow. Opening of the velum during hesitations, marked



by vowel elongation, confirmed the findings from nasal airflow data, as the open velum allowed air to pass through the nasal passage.

Temporal measurements were conducted to analyze the duration of nasal phonemes. The overall duration measurements were similar to the aerodynamic findings, with nasal vowels being longer than oral vowels.

The duration of nasal vowels was influenced by the surrounding consonantal and vocalic environment, with shorter durations observed when surrounded by voiceless consonants and longer durations when surrounded by /l/.

The presence of different vowels influenced the duration of velar movement, but the effect was not consistent. Velar movement anticipation was greater when the nasal vowel appeared at the end of a word.

We also followed the work of Lovatto et al. (2007), who analyzed the velopharyngeal aperture using a fiberscope during the production of the nasal vowels. The authors investigated the velar movements involved in the production of nasal vowels in BP spoken in the region of Rio Grande do Sul.

In the authors' corpus, the nasal vowels were positioned at the beginning, in the middle, and at the end of some pseudo words. The study examines the nasal vowels /ẽ, ĩ, ũ/ in initial, medial, and final positions within nonwords. The authors measure velum movements using fiberoptic video-recording synchronized with acoustic recording.

The acoustic duration of the vocalic nasal portion (V) and the nasal appendix (N) of the nasal vowels was analyzed. The relative timing of velum movements and the acoustic onset and offset of V and N was also investigated.

The results indicate that the nasal appendix is observed in 85% of the nasal vowel productions. The duration of the nasal appendix had a mean of about 88 ms to

/ẽ/, 86 ms to /ĩ/, and 100 ms to /ũ/ (n=6). According to the authors, the nasal appendix was not always present in all contexts in BP, suggesting that the nasal appendix is not a necessary condition for producing nasal vowels.

The authors concluded that, when the nasal vowel was in the medial or final position, the maximum lowering of the velum may occur before or during the production of the nasal appendix, resulting in variability. Lovatto et al. (2007) also did not observe a difference in the velum lowering movement during the production of nasal vowels, and the segments that preceded and followed the target nasal vowel were plosives and fricatives.

Lovatto et al. (2007) demonstrated the anticipation of the velar movement before the nasal vowel, and that the maximum extent of velum lowering could occur before the nasal appendix or during its production. The initial anticipation of velum opening to the nasal vowels is about 262 ms to /ẽ/, 300 ms to /ĩ/, and 229 ms to /ũ/ (n=6).

The authors further suggested that nasalization was not a binary feature [+ - nasal] because the opening of the soft palate was necessary, but was not the only, condition for nasal vowel production.

The study confirms that V plus N (Vn) is acoustically longer than the corresponding oral vowels. The duration of Vn suggests that it consists of two phonemes rather than a single vowel. However, the presence of N is not obligatory, and there is variability in the lengthening of Vn. The authors note that the phonologization of nasality in stressed and unstressed syllables in BP may differ.

The findings of the study indicate that the anticipation of velar movement before the nasal vowel in BP is consistent, although it is influenced by the position of the vowel. These observations align with similar data observed in French (AMELOT, 2004) using the same segmentation criteria. However, in BP, the timing of the

maximum lowering of the velum is more variable compared to French. In BP, the maximum lowering can occur either before the nasal tail or during it.

As the cited authors, in our study, we used a fiberscope and photo-nasography (PNG) to provide images of the nasal cavity through the production of the phonological opposition: /aw/ versus /aũ/ and /ej/ versus /ẽ/ and its allophonic realization: [ãũ] and [ẽj]. Thus, we aimed to investigate the velopharyngeal motor actions during non-pathological speech production.

PNG (photography nasoendoscopy) is an instrument used for visualizing the movement of the velum inside the nasal cavity through static photographs and videos. It also allows for a dynamic evaluation of the nasopharyngeal cavities, the velopharyngeal sphincter (velum), and the nasopharyngeal walls.

The main objective of PNG is to measure the velopharyngeal aperture by assessing the intensity of light that passes through the velopharyngeal port (velar transillumination). PNG provides a straightforward interpretation in terms of the opening and closing of the velopharyngeal port during speech (VAISSIÈRE, HONDA, AMELOT, MAEDA & CREVIER-BUCHMAN, 2010).

By employing nasoendoscopy and PNG, we can gather detailed data and observations about the mechanisms and functions of the velopharyngeal port during speech production. This provides valuable insights into the movement of the velum and the coordination of the pharyngeal structures during vocal emission. Specifically, our study aims to contribute to our understanding of the velopharyngeal port's function, investigating the variability of nasal diphthongization phonemes and the role of the velopharyngeal port in the production of nasal sounds. We seek to explore how the movement of the velum affects the acoustic characteristics of nasalization.

The velopharyngeal port plays a crucial role in speech production by regulating the airflow and sound resonance between the oral and nasal cavities. When the velum is raised and contacts the posterior pharyngeal wall, the port is closed, preventing nasal airflow. Conversely, when the velum is lowered, the port opens, allowing air to pass through the nasal cavity.

Closure of the velopharyngeal port is essential for producing non-nasal sounds such as stops, fricatives, and vowels. It ensures that the airflow is directed solely through the oral cavity, leading to oral resonance and articulation.

Imaging techniques, including nasoendoscopy, enable real-time visualization and evaluation of the velopharyngeal port function during speech. They provide valuable insights into the movements and coordination of the velum, pharyngeal walls, and surrounding structures.

Our study aims to expand our understanding of the velopharyngeal port's function by employing nasoendoscopy and PNG, allowing for dynamic image analysis. By visualizing and characterizing the movements and coordination of the velopharyngeal sphincter and pharyngeal walls during speech production, particularly focusing on nasal phonemes in Brazilian Portuguese, we aim to contribute to the scientific knowledge of the palate region and enhance our understanding of velopharyngeal sphincter movements, such as lowering and closing.

Previous research has provided some understanding of the velopharyngeal sphincter movements during speech production. It is known that the velopharyngeal port plays a critical role in controlling the airflow and sound resonance between the oral and nasal cavities.

When the velum is raised and contacts the posterior pharyngeal wall, the port is closed, preventing nasal airflow. Conversely, when the velum is lowered, the port

opens, allowing air to pass through the nasal cavity. However, the precise dynamics and variability of these movements in different phonetic contexts, particularly in BP, require further investigation.

Our study aims to contribute to this knowledge by utilizing nasoendoscopy and PNG techniques to capture dynamic images and videos of the velopharyngeal sphincter movements. By analyzing these visualizations, we can gain more detailed insights into the specific behaviors and coordination of the velum, lateral pharyngeal wall, and posterior pharyngeal wall during different speech sounds and phonetic contexts in BP.

Through our investigation, we seek to expand our understanding of the velopharyngeal sphincter movements in BP and explore potential variations and patterns that may exist. This can lead to a deeper comprehension of the mechanisms involved in producing nasal and oral sounds, as well as shed light on the variability of velopharyngeal port function in different individuals and phonetic contexts.

Ultimately, our study aims to contribute to the existing knowledge by providing more detailed and specific information about the velopharyngeal sphincter movements in BP and their impact on speech production.

#### **4.1 Objectives**

The objective of this study is to evaluate various aspects related to velopharyngeal movements and their influence on the differentiation between oral and nasal sounds in Brazilian Portuguese (BP) during nasal diphthongization. The study aims to address research questions about the timing and degree of velum lowering and wall opening during nasal vowels, as well as the duration of velar opening compared to velar closing. Additionally, the study will examine velum movements

preceding and following the target front nasal vowel and back nasal diphthong, exploring the parallel movements between the velum and pharyngeal walls.

The presence of anticipation and propagation of velum and wall movements will be analyzed, with a focus on variations in opening and closing movements based on different nasal front vowels and nasal back diphthongs. The relationship between velum opening and the distance between pharyngeal walls during nasal vowels and their closure during non-nasal sounds will also be investigated. The study will further explore anticipatory and postponed movement in articulatory coordination.

Furthermore, the study aims to observe multiple possible openings for the velum, considering both oral and nasal phonemes, and analyze the three-phase pattern (opening, maximum opening, closing) observed in nasal vowels (AMELOT, 2004). Different patterns of opening movements, including abrupt openings, oscillations, and abrupt closures (AMELOT, 2004), will be examined for their occurrence and characteristics. Variability in the maximum lowering of the velum, occurring either before or during the nasal appendix, will also be assessed.

The study seeks to provide insights into the control and coordination of velopharyngeal movements during speech production in BP. The findings will contribute to a better understanding of the specific mechanisms involved in nasal diphthongization and enhance our knowledge of the phonetic processes associated with the velopharyngeal region.

## **4.2 Research Questions**

The research questions outlined in the study are important as they aim to address various aspects related to nasal diphthongization and the velopharyngeal port's role in speech production. By answering these questions, we can deepen our

understanding of the mechanisms involved in nasality and contribute to the existing body of knowledge in the field.

- To evaluate the opening and the closing of the velopharyngeal port trajectories during nasal diphthongization.
- To describe global velar articulation.
- To understand the velum adjustment and the role of the velopharyngeal port movement (PNG data) during the production of nasality.
- To describe the velum movement in the correlation of the acoustical signal with the front nasal vowel and the back nasal diphthongs.
- To determine the role of the influence of neighboring sounds in the production of nasality.
- To compare if the velum gestures for the front and back nasal segments are the same.

Evaluating the opening and closing trajectories of the velopharyngeal port during nasal diphthongization is crucial for understanding the specific articulatory movements involved in producing nasal sounds. This investigation builds on previous studies by providing detailed insights into the temporal and spatial characteristics of velar adjustments during nasalization.

Describing global velar articulation helps in comprehending the overall coordination and movement patterns of the velum during speech production. This knowledge contributes to our understanding of velopharyngeal functioning and its role in shaping oral and nasal acoustic properties.

Understanding the velum adjustment and the role of lateral wall movement, as observed through PNG data, sheds light on the interplay between different articulatory components involved in the production of nasality. This investigation expands upon

existing research by providing a comprehensive analysis of velopharyngeal mechanisms and their contributions to nasal sound production.

Describing the velum movement in correlation with the acoustic signal of front nasal vowels and back nasal diphthongs enables us to explore the relationship between articulatory gestures and resulting acoustic characteristics. This analysis helps elucidate the specific acoustic cues associated with nasal phonemes and their variations, contributing to the broader understanding of speech perception and production.

Determining the role of the influence of neighboring sounds in the production of nasality investigates how adjacent sounds influence the articulatory behavior of nasal segments. This analysis builds upon previous work by examining the impact of neighboring phonetic context on velopharyngeal coordination and nasalization patterns.

Comparing the nature of velum gestures for front and back nasal segments expands our knowledge of the articulatory variation in nasal diphthongization. By examining differences in velopharyngeal movements between these segments, we can better understand the specific articulatory strategies employed in producing different nasal sounds.

By evaluating the opening and closing trajectories of the velopharyngeal port during nasal diphthongization, this study aims to gain insights into the precise movements and coordination involved. Describing velar articulation contributes to understanding the overall coordination and patterns of velum movement. By examining the role of lateral wall movement and its correlation with nasality production, the study aims to understand the contribution of these factors to nasal sound production.



Investigating the correlation between velum movement and acoustical signals for front nasal vowels and back nasal diphthongs provides valuable information on the acoustic cues associated with nasalization. Obtaining information about soft tissue, particularly the velum, enhances our understanding of its anatomical properties and their impact on speech production.

Determining the role of assimilatory location in nasal sound production explores how adjacent sounds influence velopharyngeal coordination and nasalization patterns. Comparing the nature of velum gestures between the front and back nasal segments helps to uncover articulatory variations in nasal diphthongization.

This study aims to contribute to the broader aims of the thesis by deepening our understanding of the velopharyngeal port's role in nasal diphthongization, elucidating the articulatory mechanisms involved, and providing insights into the production and perception of nasal sounds in Brazilian Portuguese. The findings will contribute to the field of speech production research and enhance our knowledge of the phonetic processes associated with nasal diphthongization.

### **4.3 Endoscopy**

Endoscopy is a technique that is used to access the interior of the human body; it involves the uses of a flexible, small-diameter instrument that adapts to the body cavity to minimize the risk of perforation (NUNES, 2018).

The word endoscopy originated from the Greek *endon* (inside) and *skopeo* (to look) and means 'to look inside' (SCHWAB & SINGH, 2010). Primitive tools and techniques related to modern endoscopy were already described in ancient Egypt (1700-1600 BC).

The endoscope is a medical device that is used to diagnose, prevent, control, treat or alleviate various pathologies (SONGARA, SHARMA, GUPTA V. K & GUPTA,

P., 2010) and allows for the visual inspection of the internal cavities of the human body. Organs can also be manipulated without requiring an incision, or only a tiny incision (GROSS & KOLLENBRANDT, 2009).

In 1713, René de Reaumur presented the first glass fibers. In the 1790s, Claude Chappé presented the 'optical telegraph', while Graham Bell invented the optical telephone system in 1880. In 1930, Heinrich Lamm presented the first study using fiber optics to transmit images and, in 1952, Harold Hopkin used glass fibers as an endoscope for medical applications (NUNES, 2018).

In order for the endoscope image to be of good quality, a neutral energy source of light is essential to avoid burns or electric shocks. In addition, qualified professionals are required to operate the device.

The German physician Philipp Bozzini invented the first rigid endoscope; it was a simple, vase-shaped, tubular aluminium apparatus that was 35 cm high and covered in leather (ENGEL, 2003). The instrument was called *Der Lichtleiter* or 'The Light Conductor' (MORGENSTERN, 2005).

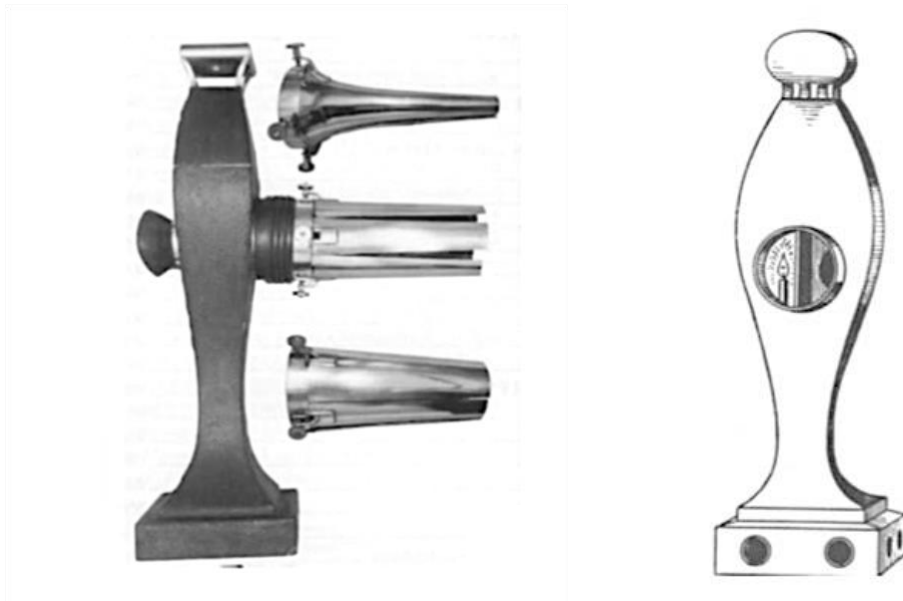


Figure 46: *Der Lichtleiter* or 'The Light Conductor' (MORGENSTERN, 2005).

This endoscope had an optical part with two lenses. A wax candle was used as a light source and the other lens was used for observation. It was a primitive instrument that allowed for the examination of the urethra, vaginal canal and pharynx.

In 1853, the French surgeon Antoine Jean Desormeaux developed the first cystoscope. The light was reflected inside the optical tube via mirrors, thus increasing the light source intensity (BERCI & FORDE, 2000). This device was used to examine the urinary tract, bladder and female gynecological tract (NEZHAT, 2005).

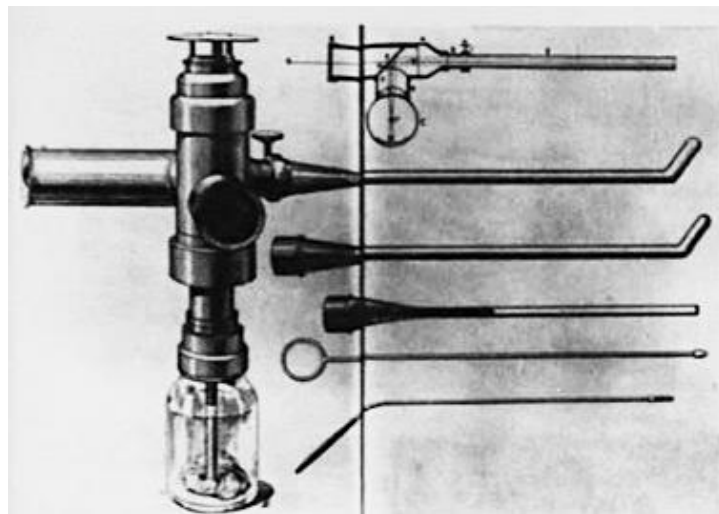


Figure 47: Antoine Jean Desormeaux's endoscope (VERGER-KUHNKE, REUTER & BECCARIA, 2007).

In 1873, Trouvé created an endoscope with a prism system, which increased the field of view by about 90 degrees. The light source did not overheat and did not need cooling; hence, it was safer and provided good visibility (NEZHAT, 2005).

In 1879, the German Maximilian Nitze created a cystoscope that had three thin aligned lenses that increased the field of view; in 1888, the light source of cystoscope was replaced by an incandescent lamp.



Figure 48: Illustration of Nitze's cystoscope (GRUNERT & OERTEL, 2011).

Lange and Meltzing created photo endoscopy in 1898. This system had a small chamber added to the endoscope tube, thus making it possible to obtain images of cavities. Nevertheless, the images had to be printed (GROSS & KOLLENBRANDT, 2009).

The first fiberscopes were created in the twentieth century. These devices were limited and rigid. In 1932, Georg Wolf and Rudolph Schindler created the first semi-flexible gastroscope, which had a rigid part and then became flexible tube (NUNES, 2018).

In 1954, Hopkins created a fiberscope consisted of fiber optics, image transmission and cold light. A light point entered the tube at one end and exited at the other. In 1956, Basil Hirshowitz developed a device made of a more flexible material.

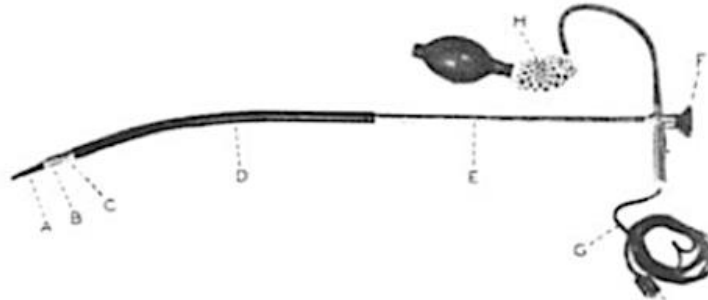


Figure 49: Image of Rudolph Schindler's endoscope (DE GROEN, 2017).

In 1983, the Welch Allyn company in the United States converted the images into electrical signals; thus, it was possible to transmit them to a monitor (MARSH, 1996). In 1964, the Olympus Optical Company developed cameras with fiberscopes, which allowed it to take photographs and observe the camera image in real time (HARA, TOBITA & TSUNODA, 1967).



Figure 50: Image of the modern video endoscope (GAAB, 2013).

This type of endoscope has replaced the fiberscope. The video endoscope consists of an endoscope, a video processor, a light source, water, air and monitors for viewing the video (GAAB, 2013).



Figure 51: Image of the flexible endoscope (DI IEVA, TAM, TSCHABITSCHER and CUSIMANO, 2014).

The flexible endoscope essentially consists of a long, thin, flexible tube, containing a light source and a micro-video camera at its end, and is sufficiently flexibility to travel through irregular structures. The device is subdivided into two types, which are optical fiber and video (GAAB, 2013).

#### **4.3.1 PNG**

PNG is an instrument that allows for the visualization and evaluation of oropharyngeal cavities. In this study, this system was used to investigate the velopharyngeal port's lowering and closing movement, as well as the movement of the velopharyngeal port during speech production.

PNG employs a fiberscope to illuminate the nasal cavity and a phototransistor to capture the images. The video system captures the velum movements, with assists in the investigation of the shape of the velum articulation to evaluate the nasal cavity's surface and the velum height.

The PNG data make it possible to observe the velum's distance from the posterior wall. A light source and a photosensor are placed in one of the speaker's nostrils (OHALA, 1971). The light source is placed above the velopharyngeal port (inside the nasal cavity) and a photosensor is located below it (or vice versa).

The change in the light intensity that passes through the velopharyngeal port allows for the evaluation of the velopharyngeal aperture. This data collection mechanism is called velar transillumination.

The use of PNG in the field of phonetics and phonology can help explain various issues related to speech production and the functioning of the vocal tract. Specifically, PNG allows for the visualization and analysis of the velopharyngeal port, providing valuable insights into the movements and coordination of the velum and pharyngeal walls during speech.

PNG enables the assessment of the velopharyngeal aperture, which is crucial for understanding the mechanisms involved in velopharyngeal closure and opening. It can help explain issues related to velopharyngeal insufficiency, such as velopharyngeal incompetence or velopharyngeal dysfunction, which may lead to speech disorders like hypernasality or nasal air escape.

PNG allows for the visualization of the velum and the nasal cavity, providing insights into the movements and adjustments involved in nasalization. It can help explain issues related to nasal coarticulation, nasal assimilation, and the production of nasal sounds in different phonetic contexts.

By observing the movements of the velum and pharyngeal walls, PNG can contribute to the understanding of articulatory gestures involved in speech production. It can help explain issues related to the coordination and timing of articulatory movements, especially in the context of nasal diphthongization or other complex phonetic processes.

PNG can provide evidence for phonological processes involving nasalization, such as nasal harmony or nasal spreading. It can help explain issues related to the interaction between phonetics and phonology, shedding light on how phonological patterns manifest in articulatory and acoustic properties.

This is a tool valuable for investigating and explaining various issues in phonetics and phonology, specifically those related to velopharyngeal function, nasality, articulatory phonetics, and phonological processes. By directly observing and analyzing the movements and adjustments of the velopharyngeal structures, PNG provides insights that contribute to a deeper understanding of speech production and the phonetic properties of languages.

The imaging data collected through PNG are obtained using small and semi-portable hospital equipment. While the technique is safe, it is considered invasive and requires data collection to be conducted under local anesthesia and supervised by a doctor. Consequently, this type of data is less commonly obtained in fieldwork settings.

The collected data allows for the accurate visualization of the soft palate in two dimensions, providing high-quality images. PNG enables a careful analysis of the opening and closing of the velopharyngeal port, offering valuable information about the dynamics and characteristics of the velopharyngeal mechanism.

#### **4.4 Theoretical Framework**

A fiberscope is a light source and is an optical method that can transmit information from one point to another. The system uses light/laser and converts electric signal variations into light signals. Accordingly, many studies of the vocal tract have used a flexible fiberoptics endoscope.

Several studies of speech production have analyzed velum movements using fiberscope data because it is possible to observe the lateral and posterior pharyngeal walls, the velum, and the larynx.

The fiberscope allows for the study of the inside of the vocal tract without the use of radiation. The lens is not positioned in the same place but depends upon the target articulatory location. It is possible to observe the movements of the velum and the pharyngeal walls inside the nasal cavity, as well as to view the velum in a rest position or falling further onto the soft palate.

We used the methodology described by Ushijima and Sawashima (1972), Benguerell et al. (1977 a), Niimi et al. (1982), Amelot (2004), and Lovatto et al. (2007) to measure the velar height from the video images. The velum can be higher, lower, or at a distance from the posterior wall.

Experiments using a fiberscope can produce an extensive database because there is no time limit during the recording session. The results provide reliable evidence of coarticulation in real-time. The device can be used in conjunction with other kinds of instruments, such as acoustic and aerodynamic recording devices.

To develop this research using nasofiberscope data, we based our study on the works of Amelot (2004) and Lovatto et al. (2007); these authors had the same experimental approach, and their contributions in terms of methods and data analyses were essential in our research.



Lovatto et al. (2007) investigated anticipatory velum lowering in their study, specifically focusing on nasal vowels. The participants in their study were asked to produce non-words with a disyllabic structure (/CoVnCoVo/, /CoVoCoVn/, and /VnCoVo/) where Vn represented nasal vowels (/ẽ, ĩ, ũ/), Co represented oral consonants (/p, b, f/), and Vo represented oral vowels (/a, i, u/). It is important to note that Vn was stressed in these non-words. The non-words were placed in the carrier sentence "Diz \_\_\_ pra ele" ([dʒis\_\_\_'pra'eɪ]) which translates to "Say \_\_\_ for him". Each sentence was repeated twice, resulting in a total of 54 utterances.

Lovatto et al. (2007) examined anticipatory velum lowering specifically in the context of contrastively nasal vowels within the given carrier sentence and non-word structures. This finding suggests that when there is a greater spatiotemporal articulation between the vowel and nasal consonant, it can lead to anticipatory adjustments, resulting in the anticipation of velum lowering before the nasal vowel. This is supported by previous studies (SOLÉ, 1995; COHN, 1990; BEDDOR, 2009; MALECOT, 1960; FOWLER & BROWN, 2000) that highlight the role of spatiotemporal coordination between vowels and nasals, leading to anticipatory effects.

Beddor et al. (2013) study focused on vowel nasalization and its role in distinguishing oral vowels from nasal vowels. The degree and duration of nasalization were found to be important factors in perceiving nasalization and differentiating nasal vowels from oral vowels. The perception of the nasalization varies depending on the degree of nasalization present in the vowel. The nasal characteristics are present in the vowel sound, such as the amount of nasal airflow or resonance.

The coarticulation of adjacent speech sounds influences one another in terms of their articulatory gestures. The acoustic information suggests that as the duration of

coarticulation between a vowel and a following nasal consonant increases, indicating nasalization may become longer and more prominent.

These earlier studies shed light on the velopharyngeal investigation of the nasal cavity and velum articulation using an experimental approach and proposed models for velopharyngeal movements during speech production in different languages. From the previous studies, we made predictions and formulated our hypotheses for this research, as mentioned at the beginning of this chapter.

Another study involved fiberoptic videoscopic from Elgendy (2001) focused on investigating the dynamics of the back cavity structure during pharyngeal articulation. His study recorded 9 speakers of Egyptian Arabic, which revealed that the velopharyngeal port is open during the production of pharyngeal and laryngeal consonants in non-nasal contexts.

One of the objectives of his study was to verify the existence of a connection between nasal and pharyngeal articulation. To examine these phenomena, Elgendy employed video registrations using the fiberoptic technique to capture the motion pictures of the naso-, oro-, and laryngeopharynx.

The articulatory dynamics associated with the production of consonants in the back cavity of the vocal tract were monitored using a nasal fiberscope. This approach aimed to provide a thorough understanding of the interaction between nasal, pharyngeal, and laryngeal articulation.

Elgendy's thesis (2001) examined various aspects of articulatory dynamics, including vowel height and its relation to velar elevation, different types of coarticulation, the distinction between true pharyngeal consonants and pharyngealized consonants, as well as the potential connection between nasal and pharyngeal articulation.

In his thesis, the author presented experimental results concerning the velum's dynamics and the relationship of the lateral and posterior pharyngeal walls to words that exhibit pharyngealization and compared this to words with nasal articulation in different phonetic contexts.

The author presented frames with the phoneme representing the shape of the velopharyngeal port. The results indicated that, even in a non-nasal context, the velum was partially lowered, and the degree varied according to the phonetic context and the word structure. Thus, in the nine speakers analyzed, the velopharyngeal wall was partially open. The same author still found some inter-speaker variation.

Elgendy observed that the degree of port opening varied based on the location of constriction in the pharynx and the height of the vowel following the pharyngeal consonant. The author aimed to explore the underlying mechanism of producing pharyngeal consonants and to provide a more comprehensive understanding of their complexity. He also investigated the interaction between nasal, pharyngeal, and laryngeal types of articulation, examining the influence of both aerodynamic and mechanical factors.

Additionally, Elgendy discussed various phonetic and phonological observations related to pharyngeal articulation, nasalization, and laryngealization. He noted perceptual judgments indicating a significant degree of nasality associated with pharyngeal consonant production.

Elgendy also explored the impact of velum and larynx movements on the size of the pharyngeal cavity, highlighting how the direction of these movements can lead to expansion or reduction of the cavity. By investigating the anatomical relationships between various muscles and cartilages involved in pharyngeal articulation, Elgendy

aimed to shed light on the synergistic and antagonistic actions that contribute to the overall mechanism.

Elgendy's thesis focused on the anatomy and dynamics of the vocal tract, particularly the pharyngeal region, and its relationship to articulatory movements. He discussed the conventional understanding of the vocal tract as a three-branch tube, consisting of the front (oral), back (pharyngeal and laryngeal), and nasal branches. The pharynx, a longitudinal cylindrical tube connecting the nasal and laryngeal cavities, was described as partially open towards the oral cavity.

The interplay between the muscles: constrictor and levator muscles, plays a crucial role in shaping the pharyngeal cavity. The contraction of these muscles leads to a reduction in the cavity size. On the other hand, the levator muscles primarily contract in the lower pharynx (SONESSON, 1970), while the upper pharynx undergoes passive compression.

Elgendy also discussed the impact of the palatopharyngeus muscle on the size of the velopharyngeal port. This muscle, arising from the soft palate, inserts into the lateral pharyngeal wall and the thyroid cartilage. Contraction of the palatopharyngeus muscle causes the soft palate to lower when the thyroid and pharyngeal wall are fixed.

Conversely, if the soft palate is fixed due to the contraction of an antagonist muscle, such as during swallowing, the raising of the thyroid can passively pull down the soft palate.

This action leads to an opened velopharyngeal port, facilitating nasal-oral coupling to varying degrees depending on the force of contraction applied to the palatopharyngeus muscle. The specific activation of either the extrinsic laryngeal muscles or the palatopharyngeus muscle during swallowing remains unclear (ZEMLIN, 1968).

Elgendy's thesis also considered the relationship between vowel height, velar elevation, and pharyngeal size. It was reported that the degree of velar elevation varies for different vowels and is often related to vowel height or openness. For instance, the velum tends to have a higher position for high vowels compared to low vowels (MOLL, 1962; LUBKER, 1968; SEAVER & KUEHN, 1980).

The height of the vowel is determined by the position of the tongue in the vocal tract, which, in turn, affects the size of the pharynx. This relationship between vowel height and pharyngeal size has implications for the acoustic output signal, particularly in terms of formant frequencies.

Coarticulation, which refers to the influence of neighboring sounds on each other, can arise from factors other than timing variations due to mechanical-inertial factors. One source of coarticulation is the pre-programmed neural commands that coordinate articulatory movements.

In anticipatory coarticulation, articulator movements are initiated in advance, causing smearing of features from preceding sounds onto subsequent ones. For example, the velum can start lowering well before a nasal consonant occurs in an utterance, as observed in American English (MOLL & DAMLOFF, 1971).

On the other hand, carryover coarticulation involves the transfer of features from an initial sound to following sounds if they do not conflict with the features of those sounds. The extent of coarticulation is influenced by factors such as the range of coarticulation exhibited by different articulators and the phonetic context. Speaking rate can also play a role (STEVENS & HOUSE, 1963; VAN SON, 1991).

Elgendy's thesis also addressed a study of perception, when a speaker listens to words containing true pharyngeal segments, trained ears may perceive some degree of nasal coupling. This can be detected through the intonation pattern, where

a drop in the F0 contour occurs when the pharyngeal segment is produced. Acoustic signals associated with true pharyngeal consonants, such as /ʔ/, compared to the preceding vowel, may indicate the presence of constriction in the larynx itself (Elgendy, 1982).

Furthermore, observations and speculations are pointing to a possible connection between nasal and pharyngeal articulation. Some regional dialects, like contemporary Iraqi Arabic, exhibit a substitution of voiced pharyngeal /i/ with a nasal consonant (e.g. "he gave" pronounced as /ʔant a/). The pharynx, being anatomically linked to the nasal, oral, and laryngeal cavities, suggests a potential interaction between nasal, pharyngeal, and laryngeal articulation. His study aimed to explore this connection, providing evidence for a relationship between nasal and pharyngeal articulation. The nasalization and pharyngealization of certain sounds demonstrate the interplay between these articulatory features (HETZRON, 1969; KOTBY & BARAKAH, 1979).

Elgendy (2001) aimed to show the degree of displacement and the target position for the target articulation. This variation was more significant in intervocalically pronounced words, such as /sʕɑʕʕar/ 'priced'. The velopharyngeal port must be sufficiently wide to indicate an evident nasal-oral coupling, but the degree of the velum lowering was affected by the vicinity of the following consonants /ʕ, s, m and n/. However, for words such as /maenh/ 'giving', the velum was closed during the final /h/ production. For uvular pharyngeal consonants such as /ʕ, ɣ/, the velum opening was higher in all the speakers.

Elgendy's thesis focuses on investigating the possible connections between the velopharyngeal orifice area and the point of constriction along the pharynx. The study observations help us to raise questions about nasal articulation.

The thesis assumes that the constriction in the pharyngeal cavity involves an interaction between velar and pharyngeal activities, as incomplete closure of the velopharyngeal port during the production of non-nasal consonants and vowels may result in coupling between the nasal cavity and the rest of the vocal tract.

Amelot's thesis revealed that the relationships between articulatory movements, nasal airflow, and perception of nasality are not linear. The opening movement of the velopharyngeal port is a necessary but not sufficient condition for nasal airflow production. Similarly, nasal airflow is a necessary but not sufficient condition for the perception of vowel nasality.

These conclusions are based on observations that demonstrate a lack of temporal relationship between the onset of velum opening and the appearance of nasal airflow. There is also no correlation between the size of the velopharyngeal opening and the amount of air exiting. Furthermore, there is no relationship between the anticipation of nasal airflow and the time required to perceive a nasal vowel.

In most productions, the velum begins its movement before the acoustic signal where evidence of oral closure during the nasal phoneme becomes perceptible or detectable. It signifies the onset of nasal airflow and the transition from the preceding non-nasal sound to the nasal sound, the slower the rate of speech, the more magnitude of anticipatory degree of coarticulation.

The articulatory movement of the soft palate during the production of nasal vowels is typically more gradual or slower when the nasal vowel occurs at the end of a word. This phenomenon can be attributed to phrase-final lengthening, which is a tendency for the final syllables or sounds in a phrase or utterance to be elongated or prolonged. This lengthening effect may result in a more gradual closure of the soft palate for nasal vowels in word-final position.

In Amelot's study (2004), it was found that nasal vowels have a longer duration of nasal airflow compared to non-words containing nasal vowels. This indicates a delay or slower onset of nasal airflow in these pseudo words, meaning that the nasal airflow starts later in time compared to other types of words.

The duration of the nasal appendix is longer for nasal vowels than it is for non-words with nasal vowels. Amelot (2004) also showed that the nasal airflow depends on the surrounding segmental context of the nasal vowel. It includes the adjacent phonemes or sounds that occur before and after the nasal vowel. The specific context surrounding the nasal vowel can influence the nasal airflow. For example, certain phonetic factors such as occlusives (consonants with complete closure) may result in minimal nasal airflow, particularly in non-words.

Amelot (2004) concluded that the degree of soft palate opening at the beginning of a sentence varied depending on whether the phoneme following it was oral or nasal. In this context, the reference point is the phoneme that immediately follows the soft palate opening. The distinction between oral and nasal phonemes affects the extent or magnitude of the soft palate opening at the beginning of the sentence.

The nasal airflow is minimal in occlusives, particularly in non-words. The voicing also changes the volume of the nasal airflow, and the regressive nasality is present to a moderate degree (AMELOT, 2004).

Amelot (2004) concluded that there were several possible degrees for the soft palate opening at the beginning of the sentence, which depended on whether the phoneme was oral or nasal. The velar movement does not change in the function of the surrounding consonants. In spontaneous speech, the velum is closed before the acoustic end of the nasal phoneme.



This implies that there is a temporal mismatch between the articulatory gesture of closing the velum and the acoustic manifestation of the nasal sound. This may result in a brief period of overlap between the nasal and the following sound, potentially influencing the perception and acoustic characteristics of both sounds.

The closure of the velum before the acoustic end of a nasal vowel can affect the nasal resonance and duration of the vowel. It may result in reduced nasalization or altered nasal resonance characteristics. This temporal dissociation between articulation and acoustics can impact the perception and intelligibility of nasal vowels, as the expected acoustic cues associated with nasality may be altered.

In Amelot's thesis (2004), it was concluded that the soft palate opening at the beginning of the sentence varied depending on whether the phoneme being produced was oral or nasal. It suggests that the degree of soft palate opening (velar movement) remains constant regardless of the surrounding consonants.

This finding is important as it indicates that the articulatory movement of the soft palate is not influenced by the adjacent consonants. It suggests that the soft palate movement is primarily driven by the phonemic distinction between oral and nasal sounds rather than the immediate phonetic context.

The implications of this finding could relate to the consistency and stability of the articulatory gestures associated with oral and nasal phonemes, contributing to the overall intelligibility and clarity of speech production.

#### **4.4.1 Velopharyngeal Sphincter**

Many studies have described the physiology of the soft palate, the velopharyngeal sphincter (henceforth VPhS), and all the pathologies that may affect it. In 1868, a German anatomist, H. von Luschka, described the anatomy of the soft palate and the pharynx in humans without pathological conditions (HOLLENDER,

1959; LOFIEGO, 1992).

The soft palate and the pharynx are two anatomical structures that are arranged in a three-dimensional configuration between the oral and nasal cavities (NOBREGA, 2002; CAMARGO et al., 2001), and are responsible for separating these two cavities (GEORGIEVSKA-JANCHESKA et al., 2016).

Moreover, the soft palate and the pharynx compose the velopharyngeal sphincter. This isthmus plays an essential role in the physiological maintenance of velopharyngeal movement and speech production.

The velopharyngeal mechanism consists of a synergistic action involving the movements of the palate, the lateral walls, and the posterior pharyngeal wall. The mechanism is bounded by the transition zone between the nasopharynx and the oropharynx (SHPRINTZEN & BARDACH, 1995).

The velopharyngeal region is composed of the palatine muscle, the palatine vein tensor, the uvula muscle, the superior pharyngeal constrictor, And the palatopharyngeal, the palatoglossal and the salpingopharyngeal arches (DICKSON, 1972; LOFIEGO, 1992; CASSEL & ELKADI, 1995; ALTMANN, 1997).

The most important muscles for maintaining the physiology of this region and the functions thereof are the soft palate and tensor veli palatini. The velopharyngeal sphincter corresponds to the area composed of the muscles that lift the velum and the lateral and posterior walls of the pharynx.

The essential function of the velopharyngeal muscles is to close the nasopharyngeal isthmus. This closing mechanism involves an anatomical concept of a circular muscle band in a hollow viscera and is responsible for maintaining the alternation between the respiratory and digestive tracts of the pharynx when swallowing (SILVA, 2008).

Regarding the structures that constitute the VPhS and its motor actions, the velopharyngeal closure represents a valvular sphincter operation that is essential for motor functions such as breathing, swallowing, vomiting, sucking, middle ear ventilation, speech production, and vocal resonance, and allows the nasal airflow escape during the production of all nasal sounds (BRUNNER, 2005).

The correct velopharyngeal closure is also essential for maintaining the oral pressure during a speech to evaluate perceptual aspects of speech production, which can provide some information about gestural changes over time (MARSH, 2003, 2004).

There is a valve mechanism that has the function of contracting, closing, and opening the sphincter to perform activities. The closing movements are triggered by pneumatic (airflow) and non-pneumatic (no airflow) activities as a result of the muscle group interaction of the elevation, sphincter and depressor movements (SHPRINTZEN, 1974; MATSUYA, 1974).

The muscular mechanism is coordinated and synergistically activates the velopharynx structures simultaneously; thus, it can link or separate the oropharynx and nasopharynx cavities (CASSEL & ELKADI, 1995; TABITH JUNIOR, 1997).

Thus, the velum can change its rest (default) position to possible positions such as stretching, closing, and opening (coupling). The soft palate accomplishes its normal physiological activities via these shape changes. The default position indicates that it can be characterized as a 'velar readiness position' (LADEFOGED, 1981, p. 99).

Through the correct respiratory functions and the precision of the sphincter movements, the velopharyngeal sphincter allows the air to escape into the nasal cavity, thus maintaining the appropriate resonance and articulation of nasal sounds in speech production. The opening (coupling) movement can be observed by comparing

oral and nasal sounds.

Some dysfunctions of the velopharyngeal mechanism structures are associated with hypernasality or low intraoral pressure (WITT, 2007; SMITH & KUEHN, 2007). When someone presents with a velopharyngeal insufficiency, the reason could be physiological, such as the lack of palate tissue or a short palate extension. Velopharyngeal incompetence is a neuromotor deficiency that can be corrected via compensatory movement articulations during oral consonant emission, for example (TRINDADE et al., 1996).

The closing movement of the velum is characterized by contact between the palate area and the posterior pharyngeal wall. The soft palate has already achieved its trajectory in the direction of touching the posterior pharyngeal wall before the speaker emits an oral or nasal sound (BZOCH et al., 1959; ALTMANN, 1997). Stretching occurs when there is a soft palate extension. This increased muscle length can occur during the production of oral consonants or the vowel /u/ (ALTMANN, 1997).

A concentric closing movement of the VPhS occurs when the palatal vein elevator muscle exerts upward and backward force, and the palatopharyngeal and palatoglossal muscles provide antagonistic force.

In motor actions that require VPhS closure, these three muscles form two anatomical and functional straps and have a central meeting point in the third middle region of the soft palate, which is called the functional center of the velum (SHPRITZEN et al., 1974; MOON & KUHEN, 2004).

Two velopharyngeal closure patterns have been described, namely pneumatic and non-pneumatic patterns. Motor actions related to blowing, whistling, and speech are characterized by pneumatic closure, in which an air jet is directed at different flow velocities in the oral cavity.

Motor actions, such as vomiting and swallowing, are characterized by non-pneumatic closure, and are triggered to prevent fluids, mucus, and even solids from passing into the nasopharynx (SHPRINTZEN et al., 1974, 1981; MATSUYA et al., 1974).

There are four distinct types of VPhS closure patterns, namely coronal, sagittal, circular, and circular with a Passavant fold, which can vary individually in their configurations (SHPRINTZEN, 1995). According to Lofiego (1992), the velopharyngeal has a three-dimensional closure mechanism in which the sphincter operates as a valve mechanism.

The four distinct types of velopharyngeal sphincter closure patterns are relevant for speech because they play a crucial role in regulating the airflow during speech production. The VPhS acts as a valve mechanism, controlling the opening and closing of the passage between the oral and nasal cavities. The specific configuration of the closure pattern determines the extent to which air is allowed to pass through the nasal cavity during speech.

During oral speech sounds (consonants and non-nasal vowels), the VPhS should be closed to prevent any nasal airflow, directing all the exhaled air through the oral cavity. However, during the production of nasal sounds (nasal vowels and nasal consonants), the VPhS must be open to allow airflow through the nasal cavity. The different closure patterns influence how effectively the velopharyngeal port can close and open, impacting the resonance and quality of speech sounds.

Proper closure and control of the VPhS are essential for normal speech production. Any malfunction or inadequacy in the closure patterns can lead to conditions such as velopharyngeal insufficiency or velopharyngeal incompetence, which can result in hypernasal speech or nasal air escape, affecting speech

intelligibility and articulation. Therefore, understanding the different closure patterns and their configurations is crucial for speech therapists and clinicians working with individuals with speech disorders related to velopharyngeal function.

The figure below represents the measurements of velum movement patterns in tandem with velopharyngeal wall movements:

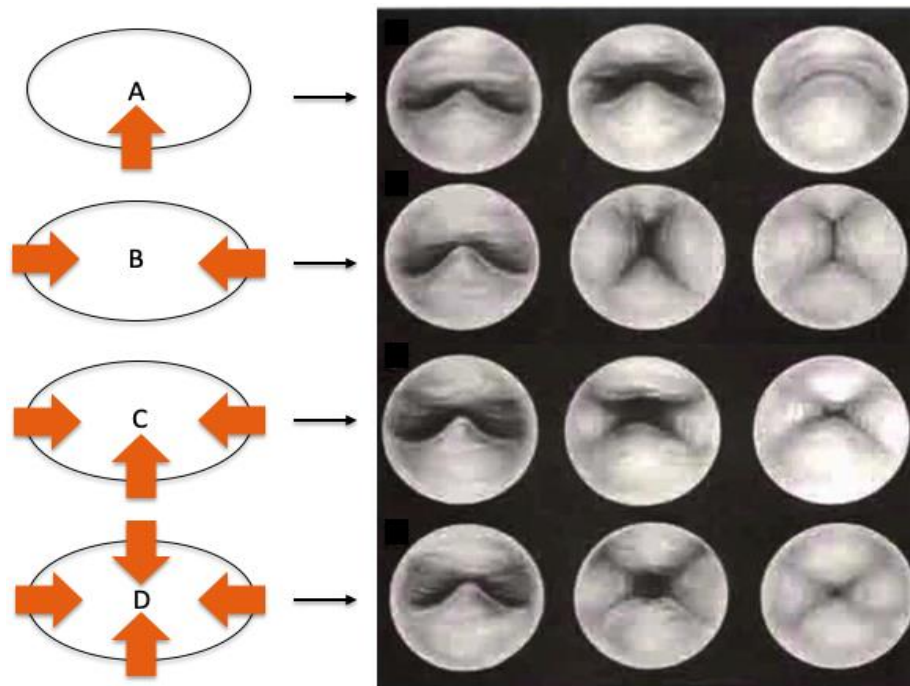


Figure 52: Measurements of opening and closing movements of the velum (SHPRINTZEN, 1995). Velopharyngeal patterns. (A) Coronal pattern. (B) Sagittal pattern. (C) Circular pattern. (D) Circular with Passavant's ridge pattern (SHPRINTZEN, 1995).

According to Shprintzen (1982), the VPhS closure patterns can be classified as:

- a) Coronal: Velum closure is mainly due to the anteroposterior movement of the soft palate, with relatively little movement of the lateral pharyngeal walls and no posterior wall movement. It is the most common movement of the soft palate towards the posterior pharyngeal wall. Velopharyngeal closure is mainly accomplished via a posteriorizing movement towards the back of the pharynx with relatively little lateral pharyngeal wall movement.

- b) Sagittal: The closing mechanism occurs by moving the lateral pharyngeal wall, with little involvement of the soft palate and almost zero involvement of the posterior wall—movement in the lateral walls refers to distance in the medial sense. Velopharyngeal closure is mainly accomplished via lateral pharyngeal wall, movement towards the midline with relatively little contribution from the movement of the velum.
- c) Circular: The movement of the soft palate and lateral walls is almost the same, but there is no balance motor action of the posterior wall movement way between the lateral walls and the soft palate. Velopharyngeal closure involves equal contributions by the posteriorizing movement of the velum and the movement of velopharyngeal port towards the midline.
- d) Circular Passavant's ridge: This produces motor activity that is almost identical to that described previously but differs in that it indicates the participation of the posterior pharyngeal wall and a circular closure that is associated with the formation of a muscular crease in the posterior pharyngeal wall. This pattern is essentially the same as the circular pattern, except that there is also movement or bulging of the posterior pharyngeal wall (Passavant's ridge).

According to Melo Vieira (2003), normal individuals without a cleft palate have a velar constriction ranging from 0 to 5.50 mm. The size of the velar constriction for nasality varies according to the speaker. For example, hypernasality, which refers to audible nasal air emission, can occur when the velopharyngeal opening is relatively large.

## **4.5 Methodology**

This section describes how we collected and analyzed the data to examine the vertical velum activity, the velum movement of lowering and closing, and the velopharyngeal port displacement during nasal diphthongization in speech production. The nasofiberscope is a technique that is used in medical and speech-language research due to its high-quality results.

### **4.5.1 Equipment**

We chose to collect the image data using a nasofiberscope because it is possible to evaluate the velum displacement during speech production through this device. This technique is minimally invasive and enables the recording of traditional endoscopy data. We used a cold light and a small camera in a flexible fiberscope.

The three main parts of the device are the objective lens with the sensor, the light source channels, and the instrument channel. The connection section connects the nasofiberscope to the video processor.

The fiberscope method allows for the video recording of velum height through the detection of the highest point from the nasal side of the velum. A phototransistor with a mini cold light illuminates the nasal cavity and the velopharyngeal port.

A fiberscope was used to illuminate the nasal cavity, and a separate tube with a photosensor was placed in the pharynx to detect velar transillumination. A specially designed 3 mm flexible silicone tube with a micro-camera (Olympus OTV-SF) connected to the fiberscope was inserted into the middle turbinate in the speaker's nostrils to visualize the nasal cavity.

A miniature tungsten light illuminated the cavity above the velum, and a small photosensor captured transillumination via the open velopharyngeal port while it was



inserted in the nasal cavity. A phototransistor with a mini cold light illuminated the nasal cavity and its walls, as shown in figures 53 and 54:

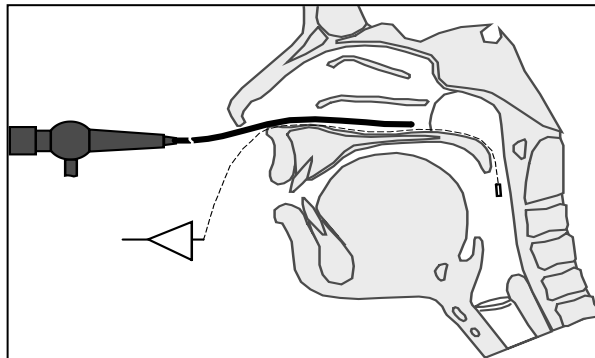


Figure 53: Nasofiberscope video-endoscopy position to collect data (VAISSIÈRE, HONDA, AMELOT, MAEDA & CREVIER-BUCHMAN, 2010).

The fiberoptics consisted of two fine glass fibers, which were flexible because they were as thin as a human hair. A protective plastic coating covered the entire fiber, thus increasing its overall size and protecting the surface. One of the tubes was a guiding light, and the other had a lens that transmitted the image onto a screen. There were two concentric lasers, an inner core and an outer core with a refractive index.

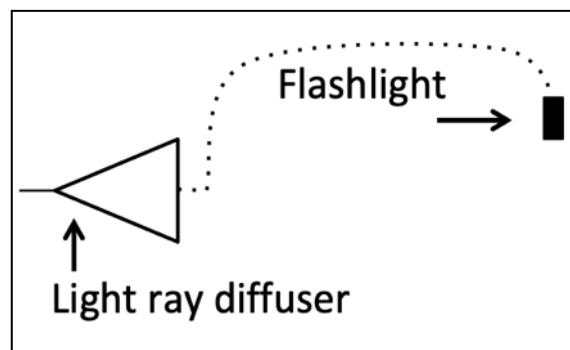


Figure 54: Illustration of the light diffusion inside the cavity (VAISSIÈRE, HONDA, AMELOT, MAEDA & CREVIER-BUCHMAN, 2010).

The guiding light illuminates the region and consists of a flexible tube made of glass and plastic. The fiber optic output is the light itself, which can modulate the intensity and move the lens that transmits the image onto a screen (HECHT, 1999).

The fiber optics provide the light source, modulate its intensity, leave the source, and enter an optical fiber, which carries it to a receiver. The receiver converts the light back into electronic form to be transmitted to devices on the other end.

The fiberscope is inserted either into the oral canal or into the nasal canal to analyze the velum's behavior. The position of the fiberscope does not interfere with the movements of the velum. The flexible tube is inserted into the nasal duct, is placed on the hard palate at the edge of the soft palate, and passes through the middle and lower horns.

The fiberscope cable is held inside the nasal cavity at the nostril level to ensure that the fiberscope does not move. Thus, it is possible to observe the nasal surface, the velopharyngeal port, the degree of velum movement, the velum's default position (reference point), and respiration. The flashlight was placed 2 cm away from the image sensor.

A transducer allowed for the recording of the videos; it captured 25 images per second, resulting in one picture every 10 ms. A microphone was attached to the endoscope to record the acoustic data simultaneously.

The audio channel and the video were synchronized, and the velopharyngeal cavity was monitored online via an endoscope screen. The data were digitized on a computer.

The acoustic signal was sampled at 48 kHz with a resolution of 16 bits. The unit of measurement for the video was pixels. The data were digitized and processed via MATLAB using homemade software and Praat..

### 4.5.2 Corpus

As in the previous experiment, the corpus included words that undergo nasal diphthongization. Nasal diphthongization refers to a phonological process where a vowel sound in a word becomes nasalized and starts to resemble a diphthong (a gliding sound that changes in quality during its pronunciation).

The words of our corpus are monosyllables that change their vowel sound, resulting in the nasal diphthongization. This phonetic process is an allophonic variation, that occurs in specific contexts, without changing the word's meaning. The nasal diphthongization occurs as an allophonic variation, in certain conditions as the final position of a word or a syllable. Our corpus included a group of words from BP that had oral and nasal contrasts.

We recorded twenty monosyllabic words to study the velum displacements, the velum lowering and rising, and the velopharyngeal port dynamics through images.

Both the back nasal diphthong /ãĩ/ and the front nasal vowel /ẽ/ are phonetically diphthongized and are produced with a nasal appendix: [ãĩŋ] and [ẽjŋ]. The word set covered the opposing oral and nasal sequences of diphthong production, including minimal and near pairs in BP.

These words produced examples of the gestural articulation dynamics undergoing a diphthongization process lies in the phonetic realization and the presence of a nasal appendix in the produced forms. This nasal vowel and nasal diphthongs have a gliding quality during their production, which is characteristic of diphthongs. In contrast, non-diphthongized nasal vowels would have a steady-state, non-glide quality. The presence of the nasal appendix, represented by [ŋ] and [j], indicates that the nasalization extends beyond the vowel and includes a subsequent

nasal glide sound, suggesting a transition from the nasal vowel to the nasal (semi) consonant.

The process of diphthongization during vowel nasalization indicates that the nasalization affects the phonetic realization of the vowel, causing it to glide or transition between two articulatory positions. This is different from underlying nasal diphthongs, which are already specified in the phonemic representation of a language, but it can undergo additional processes of closing oral capacity during speech production.

All the chosen words have a CVG or a C $\tilde{V}$  $\tilde{G}$  sequence during speech production, which includes the nasal vowels and the oral and nasal diphthongs.

This experiment aimed to facilitate the comparison and analysis of the velum's displacement during the production of oral and nasal vowel diphthongization.

The words were recorded only in the carrier sentence 'Digo\_\_\_\_\_ todo dia' ([dʒigʊ\_\_\_\_\_ ptodʊ dʒiɐ] – 'I say \_\_\_\_\_ every day'). We placed a voiceless dental consonant /t/ after the target word to minimize potential influences that occur between two segments, specifically at the end of one segment (the target word) and the beginning of the subsequent segment (the following voiceless dental consonant /t/). These word boundary effects can sometimes impact the articulation of the coda (the final part) of the target word.

By adding the voiceless dental consonant /t/ after the target word, we aim to minimize any potential influences or coarticulatory effects in the velar region, that might occur due to the transition from the final segment of the target word to the initial segment of the following word. This can help isolate the nasal articulatory characteristics of the nasal diphthongization without interference from the following segment.

The corpus was almost the same as the one described in Chapter 2. The tables below show the stimuli in BP with glosses in English:

Table 10: Set of Words with Back Oral and Nasal Diphthongs.

Oral Diphthongs		Nasal Diphthongs	
Stimulus	Gloss	Stimulus	Gloss
[paw]	wood	[pãw̃]	bread
[maw]	bad	[mãw̃]	hand
[taw]	these	[tãw̃]	so
[saw]	salt	[sãw̃]	sane
[kaw]	lime	[kãw̃]	dog

Table 11: Set of Words with Front Oral Diphthongs and Front Nasal Vowels.

Oral Diphthongs		Nasal Vowels	
Stimulus	Gloss	Stimulus	Gloss
[dej]	give (1 <sup>st</sup> SG)	[těj]	have (3 <sup>rd</sup> SG)
[sej]	know (1 <sup>st</sup> SG)	[sěj]	without
[nej]	Ney (nickname)	[něj]	neither
[fej]	friar	[hěj]	REM
[lej]	law	[běj]	good

The speaker repeated each carry-sentence three times, resulting in 60 tokens. The data were collected in the European Hospital Georges Pompidou in Paris. As this was an invasive experiment, a doctor, Lise Crevier-Buchman, conducted the recording section. A topical anesthetic is spread on the surface of the speaker's mucous nasal membrane using a cotton swab.

The speaker was calm and comfortable following the fiberscope insertion. We asked her to repeat the stimuli twice using a standard speech rate. The phrases had been written on paper. The speaker remained seated in a chair during the reading of the carrier sentence. The recording session lasted about two hours..

### 4.5.3 Speaker

The participant in this experiment was a 30-year-old Brazilian woman who was a native speaker of the BP Paulistano variant from São Paulo City. She did not have any speech pathology and had been living in Paris for less than a year at the time of the experiment. The fiberscope, a device used to visualize the vocal tract during speech production, was inserted 14 cm inside her nasal cavity to capture the articulatory movements and velopharyngeal activity.

It's important to note that the participant in this experiment was also the author of the thesis, which means I was actively involved in conducting all the experiments within the thesis, including the aerodynamic study and the articulatory study (referred to as RD in both cases). As the subject and the researcher, I hope to provide valuable insights and data for the investigation of velopharyngeal functions and nasal vowel articulation in Brazilian Portuguese.

#### **4.5.4 Data Processing**

We used the methodology described by Ushijima and Sawashima (1972), Benguerell et al. (1977), Niimi et al. (1982), and Amelot (2004) to process the velar height and pharyngeal wall displacement based on the video images

The fiberscope recorded the opening and closing of the velopharyngeal port and not the velum lowering and rising movements; hence, our data are an estimated value derived from the frontal view of half of the velum. The image output considered each speaker's cavity size to calculate the values. The part of the fiberscope tube was attached to the speaker's nostril to reduce the optical distortion.

The audio signals were digitized, converted, and stored in several separate files on a computer hard disk. The simultaneous audio and image recordings decreased the number of frames per second. The video was divided into frames according to the

horizontal and vertical axes. The velum movement during the target word had an average of 10 frames per second.

The engineer, Angelique Amelot, created scripts in MATLAB to extract the pixel variation in each frame obtained using the fiberscope. The acoustic data were segmented and tagged in Praat. Each image frame of the soft palate movement had its position variation aligned along the x-y axis. This alignment aided in comparing and measuring the movements of the soft palate across different frames. The homemade software used the values to plot the graphics' spectrogram, velar height, and pharyngeal wall displacement.

The back pharyngeal wall displacement and the upward or downward movements of the velum were calibrated along the y-axis. The measure was extracted between two points along the y-axis, and the back edge of the velopharyngeal port displacement was measured along the x-axis.

The nasopharyngeal lateral wall displacement in the midsagittal plane, at the point below the velopharyngeal region, was measured on an x-axis perpendicular to the y-axis.

The velar height and pharyngeal wall displacement were plotted on a relative scale between 0 and 15 pixels in a time function for each utterance. The highest position of the velum was the maximum value on the y-axis. The lowest position of the velum occurs during breathing.

#### **4.5.5 Segmentation**

The segmentation of the acoustic duration was performed using the spectrogram (Hz). Segmentation of the acoustic duration involves dividing the speech signal into smaller intervals based on formant transition between vowel and glide segments.

The segmentation of the image data facilitated the extraction of pixel values for the displacement of the velum and the velopharyngeal port. We measured the size of the horizontal section, the vertical section, and the maximum velum opening, as shown in the figure below (raw data frames).

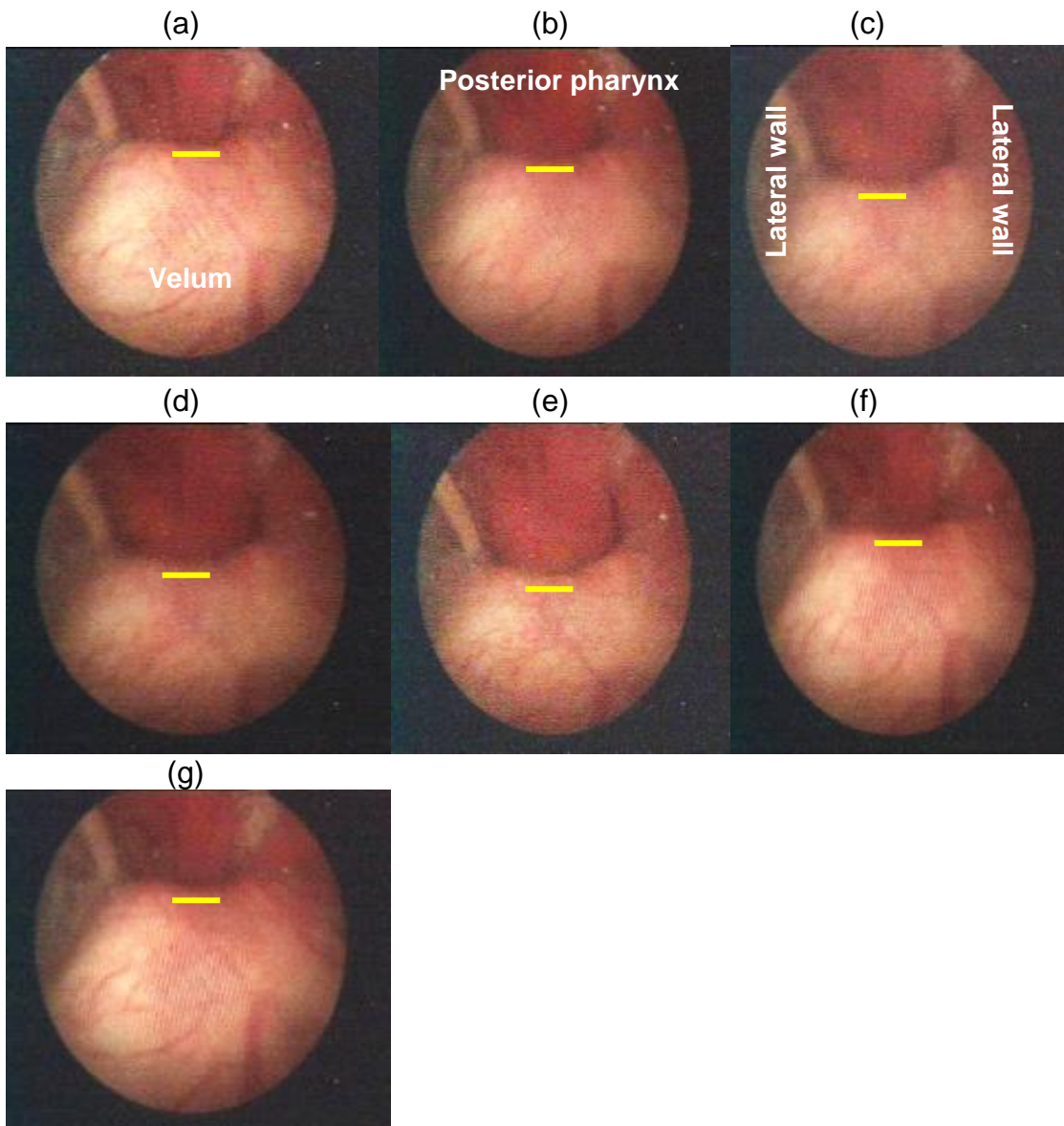


Figure 55: Nasal cavity anterior. Images of the velum's movement during the production of the word 'cão'. The velum's opening movement to the velum's lowering movement are shown from left to right. Speaker RD.

The visualization of the nasal cavity was possible due to the variation between the intensity and the transillumination passing through the cavity. It is possible to observe the spatial variation in velum displacement in each frame. The velum begins



the lowering movement from (a) to (e). In (a) there is the completely closed velum and in (e) a completely open velum during speech production.

We created a default measurement to represent the reference point based on the light diffusion inside the cavity. This reference point was determined during the breathing phase when the soft palate was in the lowest aperture position, as shown in Figure 56.

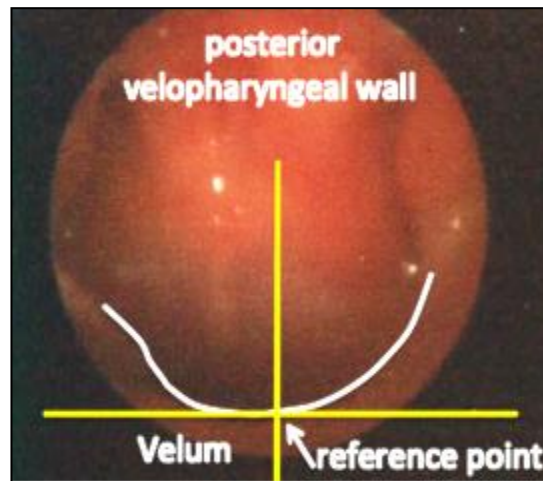


Figure 56: Nasal cavity anterior. A frame showing the reference point (RF). Adapted from Amelot's (2004) methodology. Speaker RD

We measured the delta between the lowest and highest points in the picture to evaluate the soft palate movement and the PNG data. In Figure 57, we can see the difference between the velum's position and the RF point; (a) the lowering and (b) the rising movement.

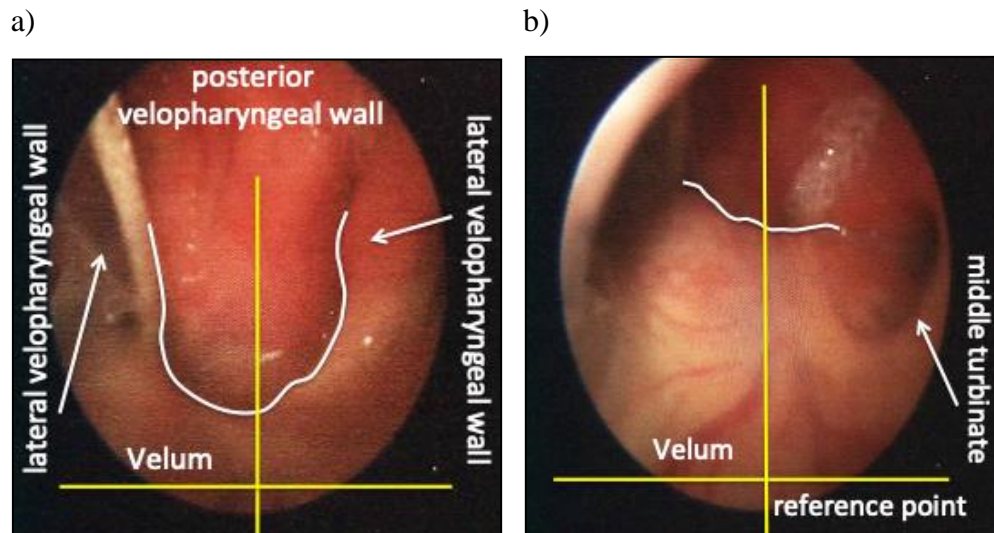


Figure 57: Nasal cavity anterior. Example of velar measurement. The frames show the velum's lower and higher points. Adaptation of Amelot's (2004) methodology. Speaker RD.

It is possible to indicate the opening and closing of the velum along the waveform based on the correlation between the fiberscope and the acoustic signal, as shown in Figure 58:

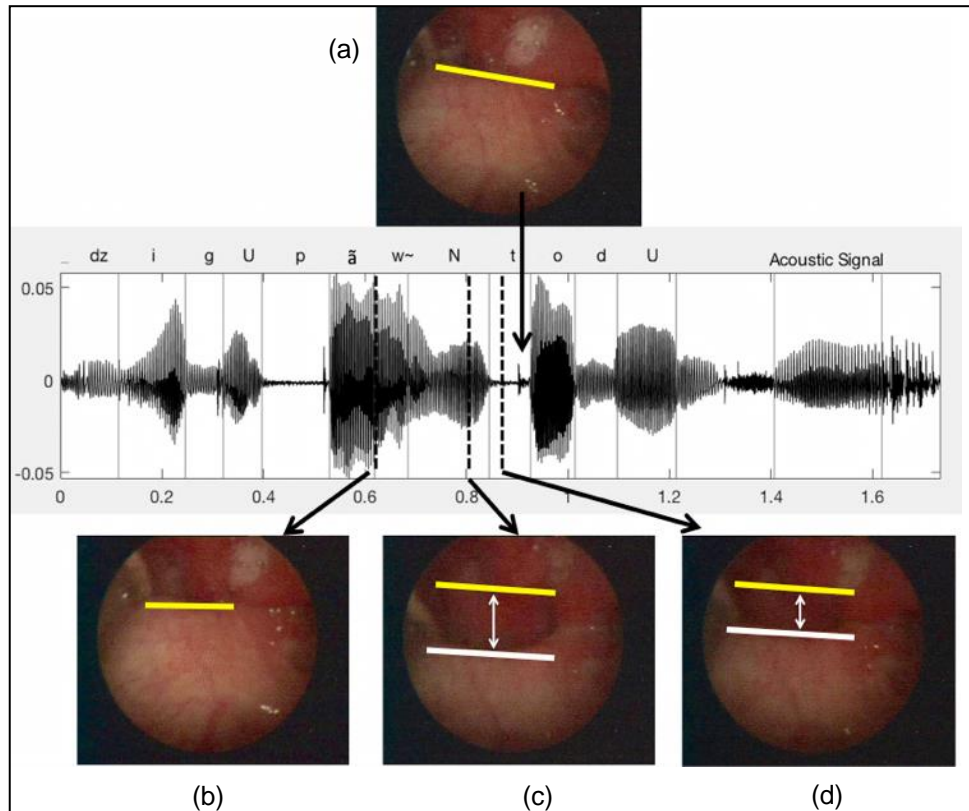


Figure 58: Anterior view of the nasal cavity. Image frames and the oscillogram. The alignment of the nasofiberscope frames and the acoustic output data. Target word: 'pão'. Adapted from Amelot's (2004) methodology. Speaker RD.

To evaluate the soft palate's movement, we measured the delta between the lowest and highest points (a-c) in line with previous work by Amelot (2004). In (a), the image represents the position of the closed soft palate. In (a) and (b), the velum is closed, and in (c), the velum is in the target position.

Diphthongs are considered in this thesis as one continuous segment, but for acoustic segmentation, we divided the diphthongs into different segments based on their oral and nasal production. This division is not based on separating them into distinct segments but rather on differentiating between the oral and nasal phases within the same continuous diphthongal trajectory.

For oral production, the diphthong was segmented into two segments, while for nasal production, it was segmented into three segments. The nasal production was segmented into the nasal vowel, the nasal glide, and the appendix.

The criteria for segmentation of nasal production into the nasal vowel, the nasal glide, and the appendix were likely based on acoustic and articulatory characteristics observed in our data. The segmentation criteria follow:

- Nasal Vowel: The beginning phase of the diphthong where the main vowel is nasalized. This phase typically exhibits characteristics of a nasalized vowel sound, such as lowered formants and the presence of nasal airflow.

- Nasal Glide: This phase occurs after the nasal vowel and is characterized by a transitional movement from the nasal vowel to the subsequent consonantal-like part. It might involve a gradual transition from the nasalized vowel to the articulatory position for the following glide.

- Appendix: The appendix corresponds to the more consonantal part of the nasal diphthong, where there is complete closure in the oral cavity, and the nasal

airflow is redirected through the nasal cavity. This phase represents the nasal murmur, after the glide at the end of the diphthong.

This approach allows for a more streamlined analysis of the dynamic changes in the acoustic signal over the course of the diphthong's production, considering both the oral and nasal phases.

By combining the segments, we believe that we can gain a better understanding of the overall temporal characteristics and acoustic properties of the diphthong during both oral and nasal contexts. The goal was to facilitate data extraction and analysis by focusing on one dynamic segment over time, rather than analyzing each segment separately.

Each picture in the figure above represents an instant during the speech production of each phoneme.

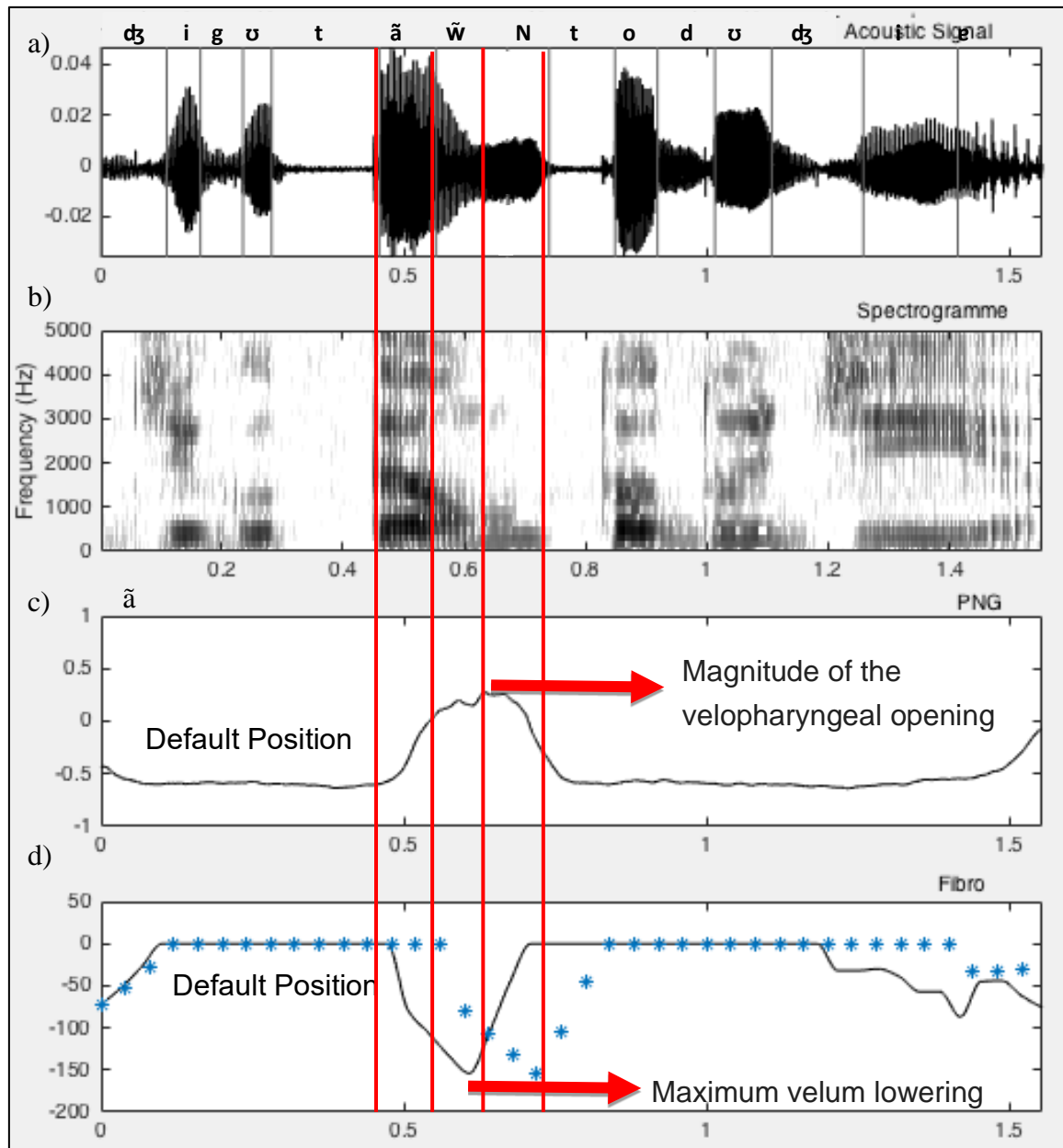


Figure 59: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'tãõ'. Speaker RD. [N] is the nasal appendix notation. The stars in blue are the fiberscope measurements for each frame. The red arrows indicate the acoustical segmentation.

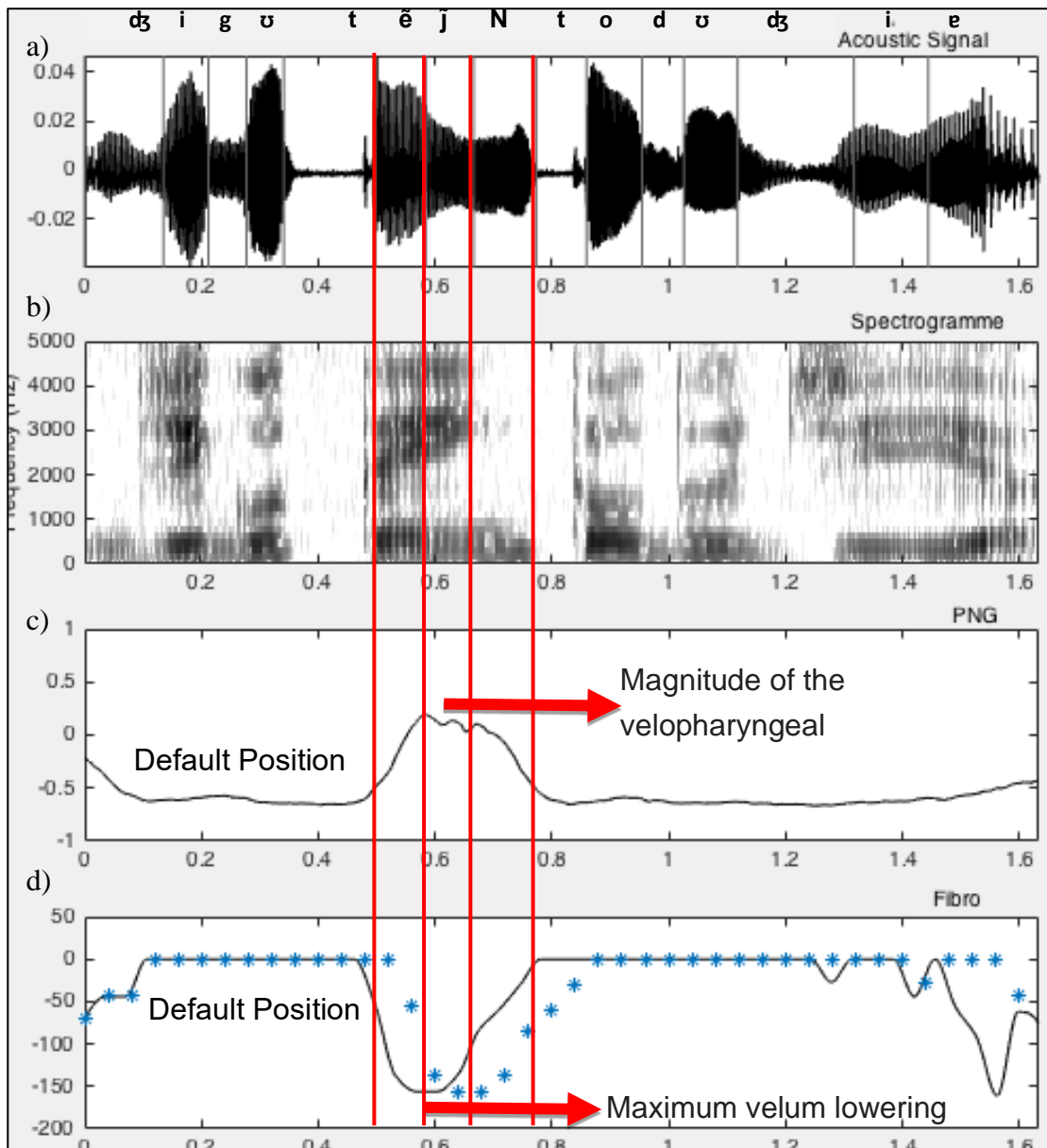


Figure 60: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'tem'. Speaker RD. [N] is the nasal appendix notation. The stars in blue are the fiberscope measurements for each frame. The red arrows indicate the acoustical segmentation.

The oral production was segmented into two portions regarding the format trajectory of the sound production. Thus, we segmented the vowel and the glide. The onset of a glide is characterized by begging of movement from high-frequency formant, which reflects the changing vocal tract configuration during the sound production. The figure below represents the duration of each production.

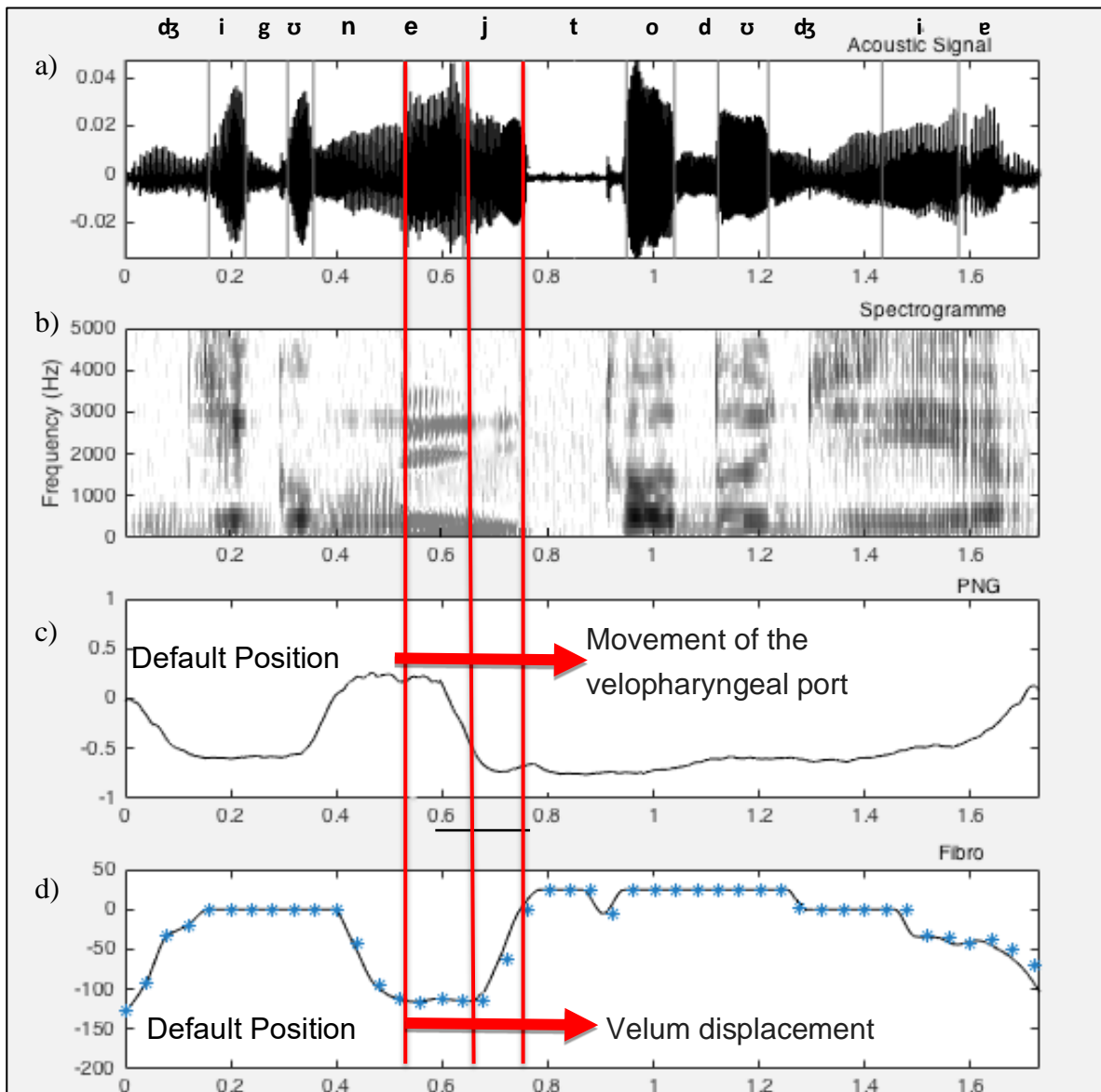


Figure 61: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'nei'. Speaker RD. The stars in blue are the fibroscope measurements for each frame. The red arrows indicate the acoustical segmentation.

We can see the velopharyngeal port movement (PNG data) and the velum movement in these figures. The peak in the PNG data indicates the velopharyngeal aperture in the nasal segments.

The PNG graphic represents the measure of the velopharyngeal port movement. The fibroscope graphic in (d) represents the soft palate movement. This movement refers to the movement of lowering. The maximum opening was when

the velum was in the lower position, which we inferred was the target position for producing nasalization.

#### **4.5.6 Parameters**

The velum frame variation was extracted from the image parameters to characterize the temporal duration of the velum movement (ms) and the spatial velum movement (px). Using this information, a graphic representation of the velum's spatial change over time was generated. This graphic illustrates the articulatory movement of the velum during speech production.

The synchronization of the fiberscope images with the acoustic signal was likely performed to precisely timestamp the video frames corresponding to specific points in the acoustic signal. This synchronization allows us to analyze and interpret the soft palate movement image-by-image in the videos concerning the corresponding acoustic signal.

By aligning the acoustic signal with the video frames, we were able to accurately link specific moments of the velum's movement captured in the fiberscope images with the corresponding acoustic signal, such as the onset and offset of the soft palate and the vocal tract configuration during speech production.

The velum's movement parameters are:

- Velum oscillation: Velum oscillation refers to three different behaviors observed during the beginning stage of the velum lowering. These behaviors are (a) plateau, where the velum remains stable at a certain position, (b) peak before opening, where the velum reaches its highest point before starting to lower, and (c) instability or vibratory movement<sup>69</sup>,

---

<sup>69</sup> The term "vibratory movement" was mentioned for the instability or fluctuations observed in velum.



where the velum shows slight fluctuations or tremors during its lowering process. These behaviors can be analyzed to understand the dynamics of velum movement during speech.

- The lowest point of the soft palate represents the velum lowering fiberscope data (px).
- The movement of the nasopharyngeal walls represents the movement of the velopharyngeal port (PNG data) (px).
- Velum movement (anticipation) or after (postpone) the vowel onset time.
- Acoustic duration (ms) of oral and nasal sounds. To measure and compare the durations of oral and nasal sounds during speech production.

## **4.6 Results**

This section describes the results of the qualitative and quantitative image data analyses of the soft palate movement and the velopharyngeal port (PNG data) movement during the speech production of nasalization target segments. We intend to provide answers about the velum displacement during nasal diphthongization using these data.

The investigation delves into the intricate movement of the velum during the articulation of nasal diphthongs, aiming to unravel the dynamics of this complex speech process. Through detailed analyses of image data, including PNG and fibro displacement graphics, we sought answers to a series of pertinent questions.

We explored whether the velum exhibits distinct patterns of movement or behavior at different stages of nasal diphthongs, encompassing their onset, duration, and offset. Additionally, this study investigated potential variations in velum

displacement between nasal segments, such as back nasal diphthong and front nasal vowel.

By comparing the velum's movement during nasal diphthongization to its behavior in the production of oral diphthongs and other vowel sequences, we aimed to gain insights into the differential articulatory processes involved. Furthermore, this investigation sought correlations between velum movement and the acoustic characteristics of nasal diphthongs, like nasalization degree and duration.

This study aspired to shed light on the critical role of the velum in the production of nasal diphthongs and enhance our comprehension of this intricate aspect of speech production. The results are presented in tables, graphs, spectrograms, oscillograms, PNG, and fibro (fiber) displacement graphics.

In the figure below, it can be noted that the PNG movement begins gradually. There is an elevated movement of the lateral walls at the location of the nasal appendix. At this moment, the velum begins the lowering movement per se.

The red lines depicted in the figure represent the identification of the velopharyngeal aperture, which was evaluated by analyzing the intensity of light passing through the velopharyngeal port using PNG with fiberscopy.

The data displayed in the figure provides evidence of the velopharyngeal port opening during the production of nasal diphthongs. The transillumination signal initiates a gradual rise right at the onset of the nasal vowel. Following the vowel onset, the port undergoes a slight opening after a few milliseconds, reaching its maximum opening during the latter part of the nasal vowel.

The velopharyngeal port remains open even after the onset of the subsequent consonant /t/. During this phase, there is a postponement, which leads to some degree

of nasalization of the initial portion of the voiced consonant /t/, although this nasalization is typically not discernible to the listener.

The blue lines represent the soft palate movement. To determine the magnitude of the velopharyngeal opening we have looked for the point where displacement deviates significantly from the baseline.

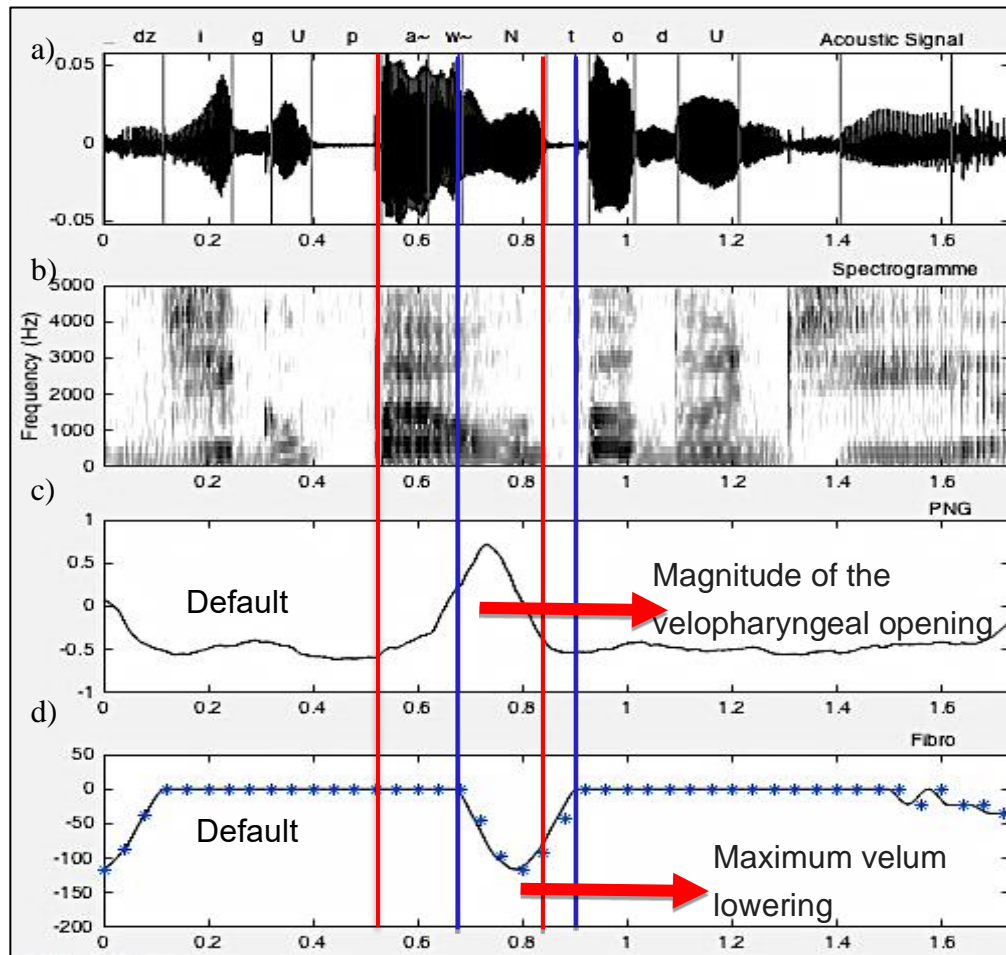


Figure 62: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'pão'. Speaker: RD. The stars in blue are the fiberscope measurements for each frame.

Regarding the nasal vowel onset time, the formant frequencies of the nasal glide (F2) move towards the formant frequencies of the preceding nasal vowel (F1). This movement of the higher frequencies towards the lower frequencies reinforces the overall nasal quality of the sound, the higher frequencies are weakened during the nasal vowel release, contributing to the perception of a nasal sound, as in (a) and (b).

The timing when the velum lowering begins from the nasal glide production, and the maximum soft palate opening coincides with the production of the nasal appendix. The velum lowering has an impact on the velopharyngeal port. The velum lowering, which plays a crucial role in creating the velopharyngeal opening, begins during the production of the nasal glide.

This means that as the speaker transitions from the nasal vowel to the nasal glide, the velum starts to lower and open the velopharyngeal port. The maximum opening of the soft palate, indicating the largest velopharyngeal port aperture, occurs during the production of the nasal appendix. The velopharyngeal port initially moves horizontally, then moves together with the lowering of the soft palate in order to create a circumference of the velopharyngeal section.

The soft palate and the velopharyngeal port magnitude do not match. The soft palate movement starts at any time before the vowel is released. The velopharyngeal port can begin to move before or after the soft palate lowers. The movement of the velopharyngeal port is more unstable than the soft palate movement; thus, this does not appear to be a parameter that characterizes nasalization.

During the activation of the velopharyngeal sphincter, the movement of the initial velopharyngeal port may change the direction of the airflow inside of the nasal cavity and result in the negative nasal airflow that we demonstrated in Chapter 2.

In the front nasal vowel, the F2 moves towards F3. The nasal appendix is the last part of the nasal diphthong and is represented by the nasal murmur (N). The velum and the velopharyngeal port have the same behavior in contrast to the back nasal segments.

In the velopharyngeal port, the peak on the graphic indicates the greatest opening of the velopharyngeal port between. The coordinated movements of the soft

palate and the velopharyngeal port movement during speech production are described as a trajectory that includes a "drawdown" phase, which likely corresponds to the initial lowering or movement of the soft palate, followed by a "valley" phase where the soft palate reaches its lowest point during speech production.

The coordinated actions of the soft palate in conjunction with the velopharyngeal port movement create and control the velopharyngeal port's opening and closing. The soft palate moves in conjunction with the velopharyngeal port.

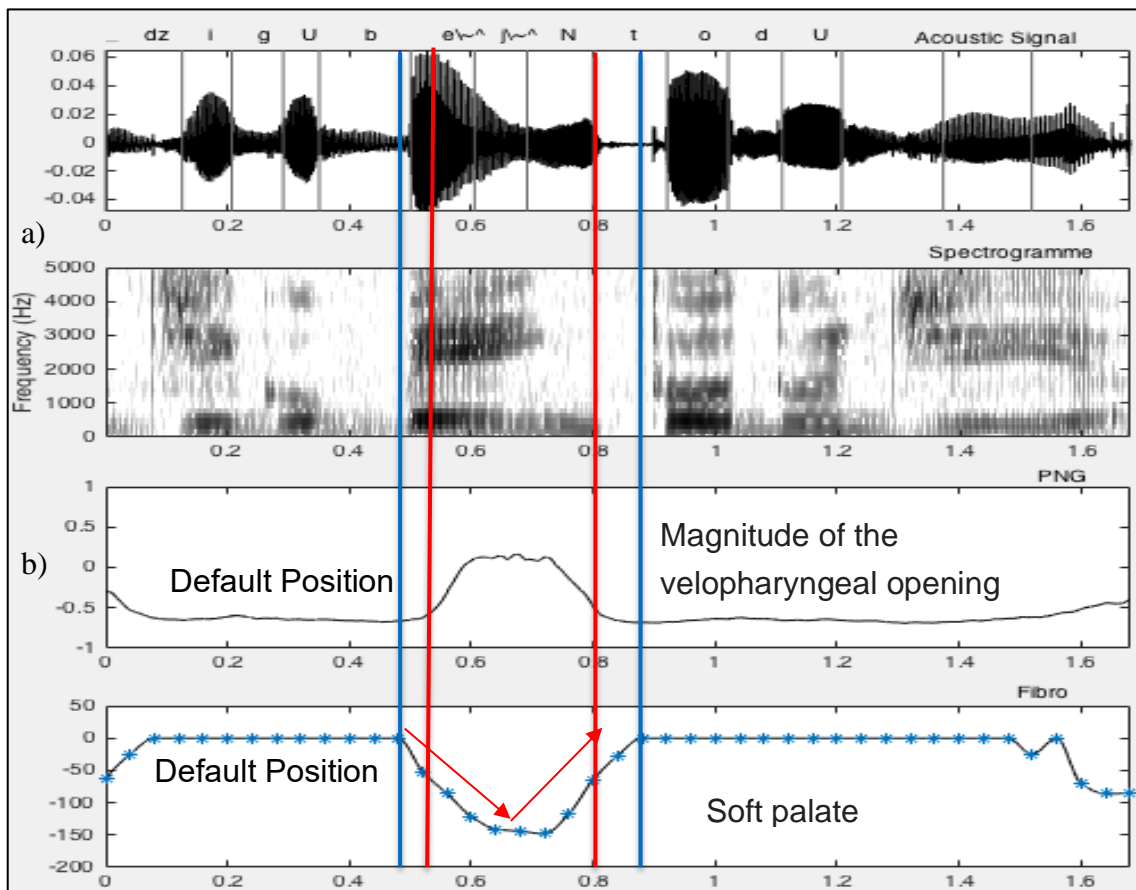


Figure 63: The alignment of the data. (a) Waveform, (b) spectrogram, (c) PNG transillumination and (d) lowering and closing velum movements. Target word: 'bem'. The stars in blue are the fiberscope measurements for each frame.

Figure 63 shows that the soft palate gesture and the velopharyngeal port are not aligned but coordinated. When the soft palate has a wider aperture, the velopharyngeal port has a greater distance between the walls. The values extracted

from the maximum velum aperture and the velopharyngeal port movement are presented in the following tables.

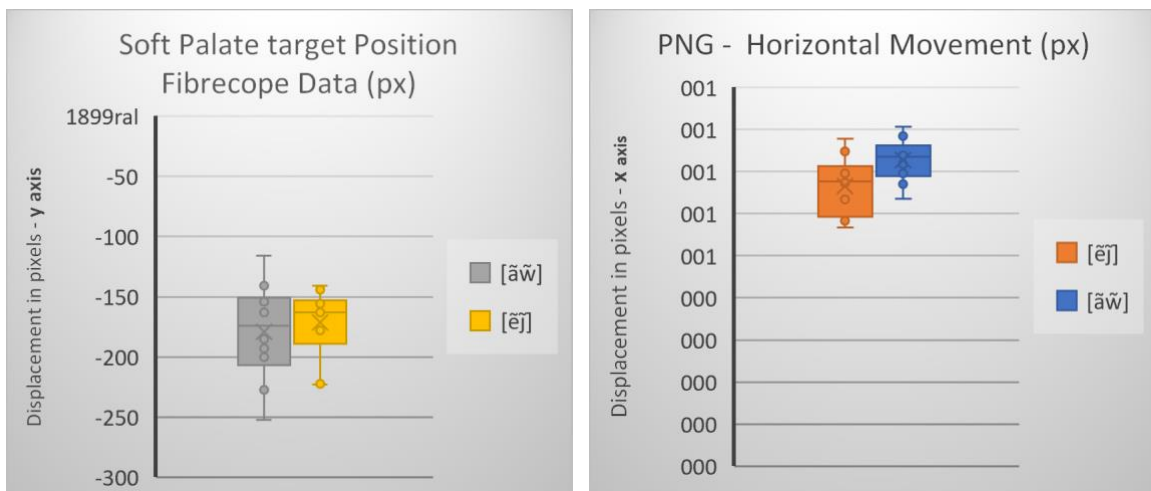
Table 12: Soft palate target values extracted from the fiberscope data. The velopharyngeal port values extracted from the PNG data. Back and front nasal segments. Average (Standard Deviation). The unit of measurement is Pixel (px) (N = 10).

Soft Palate Target Position Fiberscope Data (px) Vertical Movement		Velopharyngeal Port PNG Data (px) Horizontal Movement		
Segments				
[ãw̃]	[ěj]	[ãw̃]	[ěj]	
-185	-144	0,74	0,67	
-163	-163	0,78	0,57	
-252	-223	0,72	0,75	
-161	-222	0,75	0,63	
-116	-163	0,73	0,70	
-141	-163	0,81	0,60	
-200	-159	0,74	0,70	
-193	-178	0,67	0,78	
-227	-141	0,70	0,68	
-154	-156	0,64	0,58	
<b>A(SD)</b>	<b>-179 (41)</b>	<b>-171(29)</b>	<b>0,72 (0.05)</b>	<b>0,665 (0.07)</b>

The average for the soft palate moving in a downwards direction was  $-179$  px. The value was negative because the soft plate default position was closed, or zero on the x–y axis. Therefore, when the soft palate begins its lowering movement, the movement occurs from the pharyngeal back wall towards a central region in the vocal tract. The velum has a diagonal movement that will be described in the next chapter.

The charts below present the spatial variation in the maximum velum aperture and the maximal velopharyngeal port displacement.

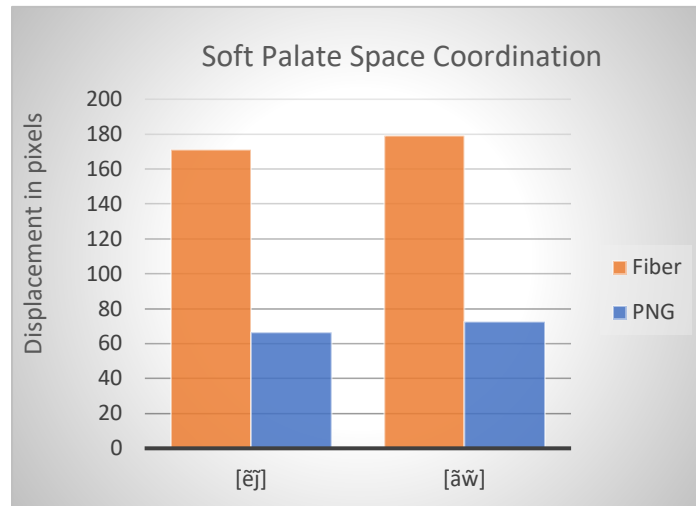
Chart 9: Spatial velum movement in the fiberscope data and the PNG data for soft palate displacement. Front and nasal diphthongs: Average (SD) px (N =10).



The soft palate and the velopharyngeal port movement have individual, homogeneous values for both nasal segments because the average of the soft palate opening is similar and the values for the maximum velopharyngeal port distance are similar. These values were expected because the movements are not overly complex, and the nature of the segments in this analysis was similar.

The velum trajectory has a homogenous displacement for the velum and the velopharyngeal port displacement. As for the mean displacements in pixels, they are 171 pixels for the Fiber data in both nasal target segments, and 67 pixels and 73 pixels for the front and back nasal segments, respectively. However, it is possible to note the difference in the values between the measurements of the soft palate's lowest point (fiberscope data) and movement of the velopharyngeal port (PNG data), as shown in the following table.

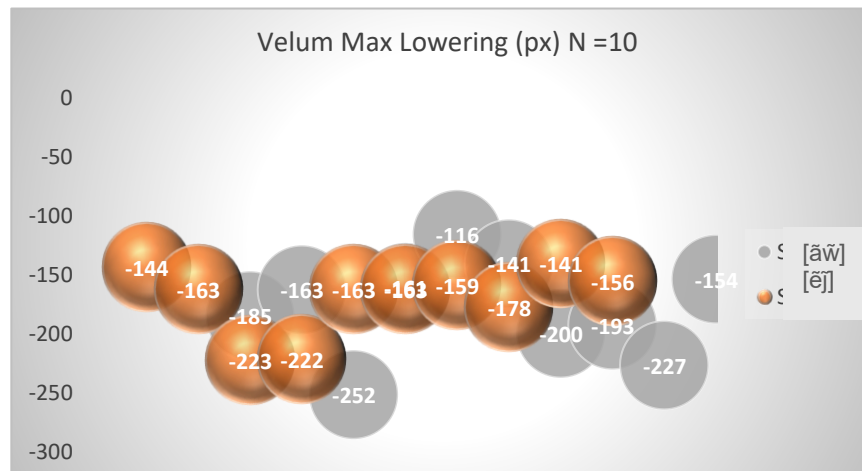
Chart 10: Spatial velum movement in the fiberscope data and the PNG data for soft palate displacement. Front and nasal diphthongs: Average (SD) px (N =10).



There were a few differences between the values for the front and back nasal diphthongs. These differences were more substantial when comparing the data for the degree of spatial variation between the fiberscope and the velopharyngeal port.

The discrepancy between the measurements of the soft palate's lowest point (fiberscope data) and the movement of the velopharyngeal port (PNG data) could be attributed the difference in the magnitude of the velopharyngeal port movement and the soft palate lowering. The graphic below shows the spatial variation in the velum lowering.

Chart 11: Maximum velum opening. Fiberscope data for the soft palate displacement. Front and nasal diphthongs: Average (SD) px (N =10).





Our data confirmed that the soft palate had an aperture (PNG data) and a lowering movement (fiber data). Unlike in the aerodynamic data, it was possible to see the velum aperture and the lowering in the images of the nasal cavity and the velum displacement. We will also see this kind of movement in the articulatory data but with greater trajectory detail.

RD exhibited a coordinated movement pattern involving the soft palate and velopharyngeal port during speech production. This coordination can be described as "circular" due to the synchronized activity of both the soft palate and lateral walls. However, it's essential to clarify that the study did not confirm the specific motor action of the posterior wall movement.

Consequently, the velopharyngeal closure process was likely facilitated by both the posteriorizing movement of the velum and the inward movement of the velopharyngeal port walls toward the midline axis.

The starting action related to the timing and position for the velum movement was variable for RD. One of the reasons for this is that the soft palate moves cyclically (MATSUO & PALMER, 2009) during speech production; the soft palate may have muscle spasms, brief involuntary contractions, shocks, spasms or tremors (ZADIKOFF, LANG & KLEIN 2006).

Thus, we realized that, at the beginning of the velum lowering and at the end of the closing movement, the velum has three different movement behaviors: (a) a plateau, (b) a peak before opening, or (c) instability (over the vibratory movement). This gestural behavior followed the criteria proposed by Amelot (2004), as shown in Figure 64:

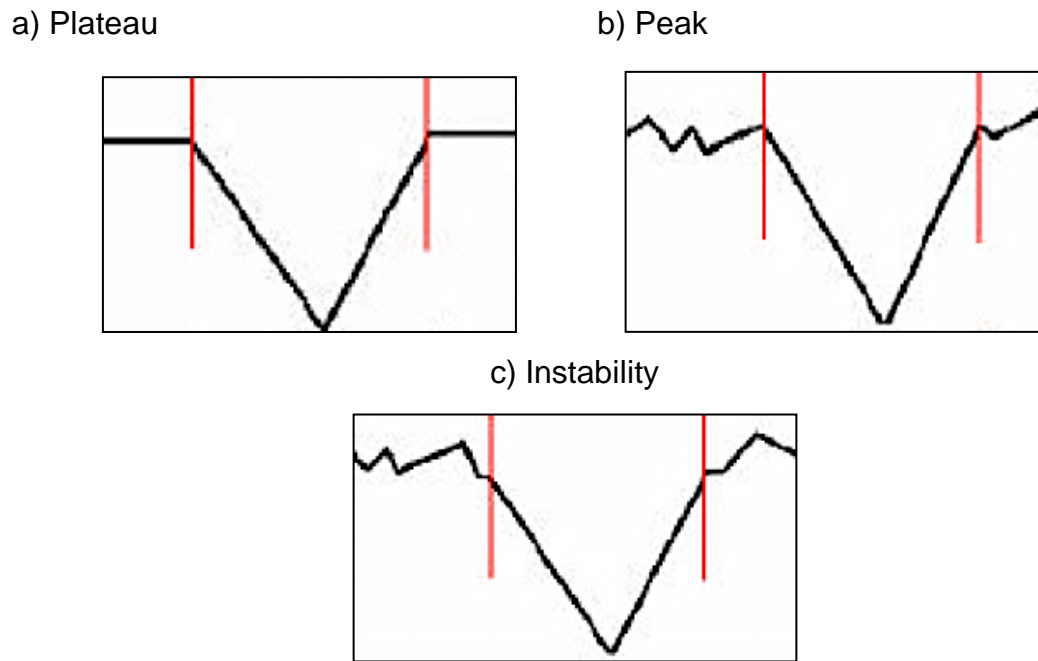
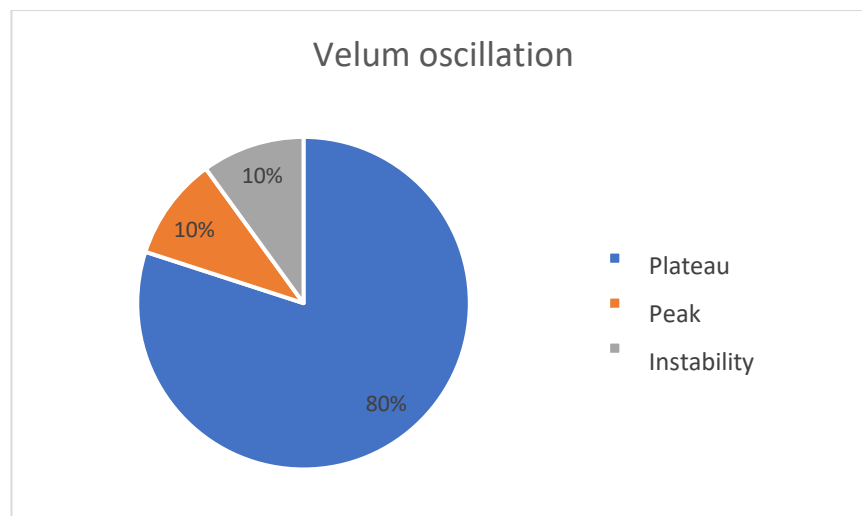


Figure 64: Illustration of the velum's behavior. The red lines represent the beginning and end of the velum displacement. Adapted from Amelot's (2004) work.

The velum oscillation to (a) occurred in 80% of the cases. In these cases, the velum did not undergo any oscillation. In (b), the peak was a type of shock, resulting in an elevation appearing in 10% of the cases and, in (c), the instability represented a tremor with small opening and closing movements, accounting for 10% of the cases. The chart below represents the proportion of the three kinds of velum behavior.

Chart 12: Spatial velum movement in the fiberscope data and PNG data for soft palate displacement. Front and nasal diphthongs: Average (SD) px (N =20).



Determining the movement onset and offset of the soft palate involves analyzing the image data obtained through fiberscopy. So, we observed each frame from the fiberscope data and extracted the pixel value variations corresponding to the soft palate movement.

By aligning these values and examining their variations, we traced the trajectory of the velum throughout the movement. To identify the movement onset, we searched for significant changes or deviations in the soft palate's position from its resting state.

The soft palate starts to move away from its highest point when the fiberscope detects a noticeable change in its position. Similarly, to determine the movement offset, we looked for the point where the soft palate returns to its resting position after the movement. This is when the soft palate reaches its lowest point or when the fiberscope detects a change indicating the end of the movement.

We conducted the same type of analysis for the velopharyngeal port distancing movement. The table below shows the raw data. The values for '*mau*' and '*nem*' were not included, as the soft palate was already open during the production of the vocalic nasal segments.

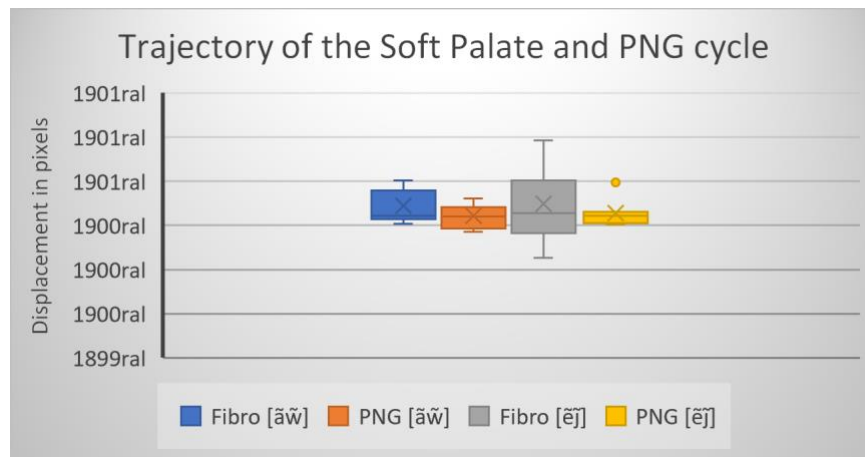
Table 13: The soft palate and the velopharyngeal port movement in a spatial path over time. Fiberscope data and PNG data for the front and back nasal segments. Average (SD) ms. (N =8).

Soft Palate Trajectory		Velopharyngeal Port Cycle	
Fiberscope Data (px)		PNG Data (px)	
Vertical Movement		Horizontal Movement	
Segments			
[ãw̃]	[ẽj̃]	[ãw̃]	[ẽj̃]
318	402	341	305
304	395	285	321
314	492	361	396
322	320	317	331
367	327	320	302

	402	226	331	323
	379	283	292	311
	317	394	290	320
<b>A (SD)</b>	<b>328 (57)</b>	<b>356 (83)</b>	<b>324 (26)</b>	<b>342 (50)</b>

The values show a temporal difference between the velum movement and the velopharyngeal port. The movements were linked but were not mechanically dependent because the velum can move horizontally without a vertical movement, as we will see in the next chapter.

Chart 13: Trajectory of the soft palate and the velopharyngeal port cycle. Fiberscope data and PNG data for the front and back nasal segments. Average (SD). Values in pixels. (N =8).



Despite the general differences in the displacement of the soft palate and the velopharyngeal port, a correlation coefficient of 0.75 indicates a relatively strong positive linear relationship. We calculated the correlation coefficient ( $r$ ) to measure the strength and direction of the linear relationship between the displacement of the soft palate and the velopharyngeal port for both front and back diphthongs ([ãw̃] and [ẽj]). A correlation coefficient of 0.75 indicates a relatively strong positive linear relationship.

The analysis aimed to examine how the displacement of the soft palate (measured using fiberscopy) and the movement of the velopharyngeal port (measured using PNG) were related during the production of nasal diphthongs.

The strong linear correlation ( $r = 0.75$ ) suggests that there is a consistent relationship between these two variables, meaning that as one increases, the other tends to increase as well. However, it's important to note that while there is a correlation between the soft palate and velopharyngeal port movement, they may not be equivalent due to different activation times.

The correlation between the onset of the nasal vowel and the velum onset (lowering movement) can be divided into two behaviors, which are when the velum movement begins before (anticipated) or after (postponed) the nasal vowel onset time. The delta between the acoustic durations of the oral and nasal gestures revealed that the beginning of the soft palate lowering/opening was variable. The results of this variation are presented in the following tables. In the delta variation for the words 'mãõ' and 'nem'<sup>70</sup>, the velum was already lowered during production. For the item 'mãõ', the mean was 232/131 ms in the fiberscope data and 307/204 ms in the velopharyngeal port for each repetition. For 'nem,' the average for the two occurrences was 219/249 ms in the fiberscope data and 214/234 (18) ms in the velopharyngeal port in each repetition.

Table 14<sup>71</sup>: Delta of the temporal coordination. Opening: nasal vowel onset and velum lowering (ms). Average (SD). Speaker RD (n=20),

Velum Lowering – Temporal Opening Variation (ms)							
[ãw̃]				[ẽj]			
Anticipation		Postpone		Anticipation		Postpone	
Fiber	PNG	Fiber	PNG	Fiber	PNG	Fiber	PNG
232*	307*	146	25	22	36	65	26
21	13	49	46	219*	214*	57	
131*	204*	41		22	22	41	
	13	20		249*	234*		

<sup>70</sup> In Table 14, the values for the velum and velopharyngeal port displacement during the production of the nasal consonants are marked with \*.

<sup>71</sup> In Table 14, the values for the velum and the velopharyngeal port displacement after the production of the voiced glottal fricative consonants are marked with \*\*.

3	152	33	52
44	24	117**	152**
37	49	137**	184**
29			41
			29

In Table 14, each row represents a different instance of the nasal diphthong production, specifically for the nasal vowels [ãĩ] and [ẽĩ]. The table is organized into four columns, representing different temporal coordination scenarios between the onset of the nasal vowel and the velum lowering (velopharyngeal port opening) during the speech production of these nasal diphthongs. The columns are labeled as follows:

- Anticipation - Fiber: The time (in milliseconds) representing the onset of the velum lowering as measured by fiberscope, when it occurs before the onset of the nasal vowel.

- Anticipation - PNG: The time (in milliseconds) representing the onset of the velopharyngeal port opening as measured by PNG, when it occurs before the onset of the nasal vowel.

- Postpone - Fiber: The time (in milliseconds) representing the onset of the velum lowering as measured by fiberscope, when it occurs after the onset of the nasal vowel.

- Postpone - PNG: The time (in milliseconds) representing the onset of the velopharyngeal port opening as measured by PNG, when it occurs after the onset of the nasal vowel.

Each entry in the table represents the duration of the velum lowering or velopharyngeal port opening relative to the nasal vowel onset time for a specific occurrence of the nasal diphthongs [ãĩ] and [ẽĩ].

The "Anticipation" rows indicate that the velum lowering or velopharyngeal port opening begins before the nasal vowel, while the "Postpone" rows indicate that it begins after the nasal vowel. The values in each cell represent the time difference in

milliseconds between the onset of the velum lowering or velopharyngeal port opening and the onset of the nasal vowel for the corresponding condition.

When producing back nasal diphthongs, the soft palate only begins the opening movement before the vowel during the production of the nasal consonants, as was expected. However, the beginning of the velum activity (velopharyngeal port opening) occurred 20 ms before the release of the vowel, specifically during the production of the consonant /s/. This suggests that the velum started to lower before the vowel was released or during the /s/ consonant.

The anticipatory soft palate movement may be related to the trigger word context. The nasalization spread appeared to follow the sonority principle, which decreases the degree of nasalization according to the scale's sonorous scale. However, for the front nasal vowel, the soft palate and the velopharyngeal port began the opening movement before the nasal vowel in most cases.

The duration of anticipatory velum movement can be demonstrated during the production of the token [hẽ̃]. For one-word repetition, the average to Fiber and PNG data is 117 and 137 ms, and for second-word repetition, the average to Fiber and PNG data is 152 and 184 ms. The influence of the fricative sounds in the onset of a syllable, the influence of nasalization gesture. This means that the velopharyngeal port was already beginning to open in anticipation of the nasal vowel, affected by the consonants that preceded the nasal vowel.

The velum and the velopharyngeal port were active during the production of the nasal appendix and moved to open the velopharyngeal section. The maximum opening of the soft palate coincided with the central portion of the nasal appendix. The closure began at the moment at which the soft palate reached its target and reversed

the movement. The following table shows the delta duration for the soft palate and the velopharyngeal port closure.

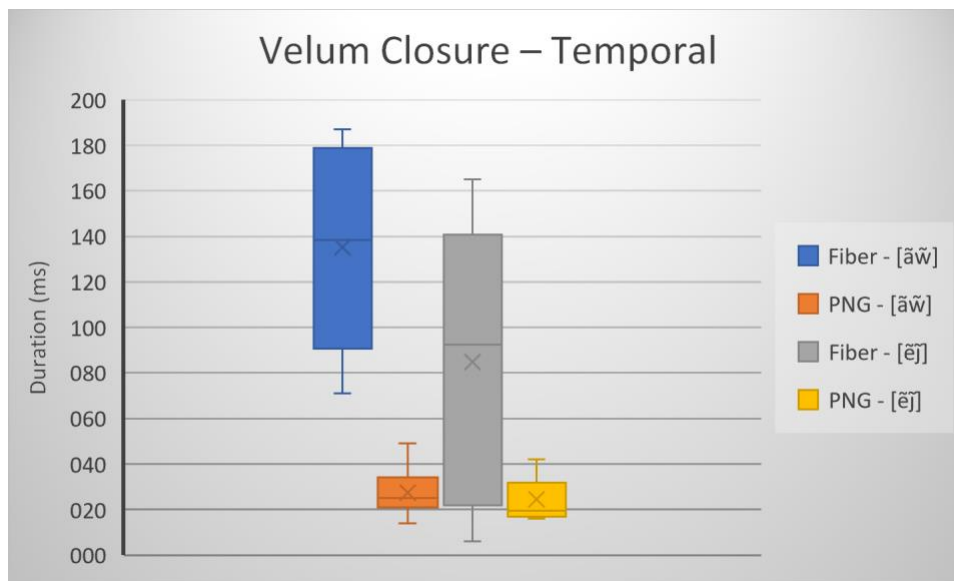
Table 15: Delta of the temporal coordination (ms). Postpone to closing movement of the velum. Average (SD). Speaker RD (n =20).

<b>Velum Closure – Temporal Closing Variation (ms)</b>				
<b>[ãw̃]</b>		<b>[ẽj]</b>		
<b>Fiber</b>	<b>PNG</b>	<b>Fiber</b>	<b>PNG</b>	
103	21	129	17	
93	35	165	40	
175	26	95	26	
111	49	34	17	
71	24	6	18	
177	21	161	20	
84	34	27	42	
166	23	134	29	
184	27	90	16	
187	14	7	19	
<b>A(SD)</b>	138	28	80	25

The closing movement of the velum and the velopharyngeal port was in the same phase because they coincided. The PNG variation shows that the movement of the velopharyngeal port was similar for both target segments, which determined the correct global velum movement. The velopharyngeal port movement was not variable in these data but was variable in nasal vowels. The chart below shows the similar behavior of the samples.

Chart 14: Soft palate and the velopharyngeal port cycle. Fiberscope data and PNG data for the front and back nasal segments. Average (SD) ms. (N =20).





The velum closing and the velopharyngeal port movement are in the same phase as the consonant onset in the following context. The phase relationship can be inferred from the timing and coordination of the movements observed in the fiberscope (velum closing) and PNG (velopharyngeal port movement) data concerning the onset of the consonant.

The mean for the movement of the velum closure in the fiber data from [ãw̃] and [ẽj] was 28 ms and 25ms, respectively. The Pearson correlation is  $r = 0.29$ ; in the PNG data, the velopharyngeal port closing movement is  $r = 0.21$ . The weak correlation between the movement of the velum closure and the velopharyngeal port closing movement during the production of [ãw̃] and [ẽj]. This weak correlation suggests that the two movements do not closely align or occur in the same period.

The weak correlation probably occurred because the correct velopharyngeal closure is essential for maintaining the intraoral pressure. Thus, there was a synchronic movement between the soft palate and the velopharyngeal port.

The moderate linear relationship between the two variables represents the similar behavior of the closing movements. This result corroborates the pattern of

nasal airflow, which undergoes a steep drop after the peak of nasal air. The image data also corroborated the articulatory data that demonstrated the shortest and fastest trajectory for the velum closure.

#### 4.6.1 Duration

The acoustic duration data that were collected simultaneously with the imaging data corroborated our aerodynamic data results and the analyses of nasal vowels in the literature. Both nasal vowel and nasal diphthong were longer than were the oral segments. The average duration for [aw] was 279 ms and 252 ms for [ej]; it was 279 ms for [ãw̃] and for [ẽj̃].

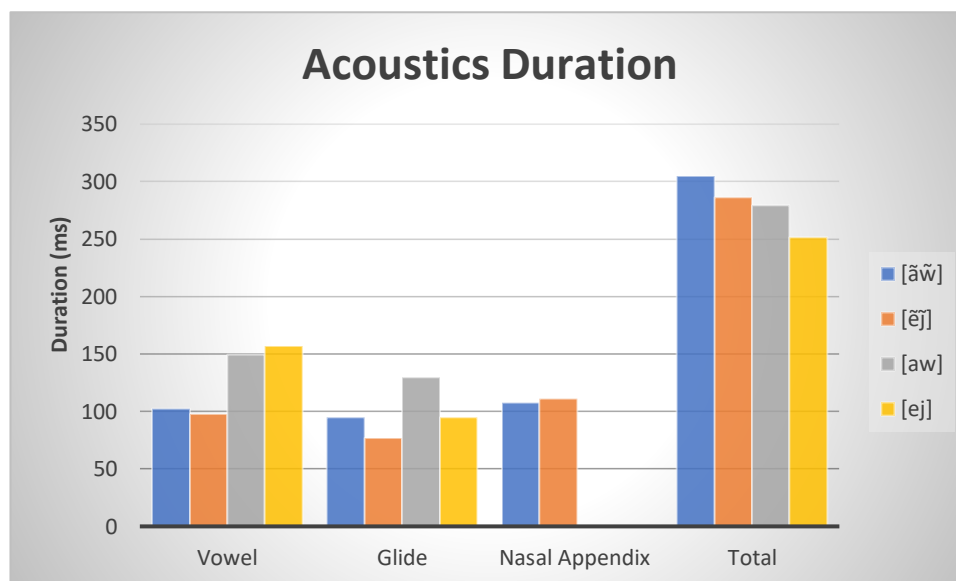
Table 16: Acoustic duration (ms) of front and back oral diphthongs. Average (SD). Speaker: RD. (N =10).

Acoustic Duration (ms). Front and Back orals segments. Speaker RD. (N =10)						
Total Duration		Vowel		Glide		
Segment						
	[aw]	[ej]	[aw]	[ej]	[aw]	[ej]
	295	273	137	147	158	126
	293	254	120	168	173	86
	255	252	173	140	81	112
	287	227	136	156	151	72
	281	231	150	173	132	58
	281	325	162	208	119	117
	296	198	170	108	127	89
	307	215	156	128	152	87
	222	212	134	138	88	73
	274	331	155	203	119	128
A (SD)	279 (25)	252 (46)	149 (17)	157 (32)	130 (30)	95 (24)

Table 17: Acoustic duration (ms) of front and back oral and nasal diphthongs: Average (SD). Speaker: RD. (N =10).

Acoustic Duration (ms). Front and Back nasal segments. Speaker RD. (N =10).								
Total Duration		Nasal Vowel		Nasal Glide		Nasal Appendix		
[ēj]	[āw̃]	[ēj]	[āw̃]	[ēj]	[āw̃]	[ēj]	[āw̃]	
300	287	105	127	86	95	109	65	
302	315	97	100	72	88	133	128	
296	315	89	89	53	66	154	160	
267	295	95	92	64	107	108	96	
317	320	95	83	108	92	114	145	
292	274	109	114	83	103	100	57	
265	322	97	112	65	114	103	96	
255	376	115	140	61	111	79	124	
291	266	79	78	99	101	113	87	
276	279	97	89	78	71	101	119	
<b>A</b>	<b>286</b>	<b>305</b>	<b>98</b>	<b>102</b>	<b>77</b>	<b>95</b>	<b>111</b>	<b>108</b>
<b>(SD)</b>	<b>(20)</b>	<b>(32)</b>	<b>(10)</b>	<b>(20)</b>	<b>(18)</b>	<b>(16)</b>	<b>(20)</b>	<b>(33)</b>

Chart 15: Average of the Duration of Acoustic Nasal Segments: Average (SD) px (N = 10).



We compared the total duration averages for [aw] and [āw̃]; the Pearson correlation coefficient was  $r = 0.58$ ; this value was  $r = 0.23$  for [ej] and [ēj]. Thus, the linear relationship between the duration of the back nasal diphthong and the back oral diphthong had a moderate linear relationship compared to the relationship between the front nasal vowel and the front oral diphthong.

The nasal appendix was present in all the cases; its total duration had an average of 109 ms and its duration was similar in the context of an occlusive coda boundary. The relationship between [ãw̃] and [ẽj̃] was weak because the Pearson correlation was equal to  $r = -0.31$ . This negative value was because the first segment was a true nasal diphthong, and the second segment was a nasal vowel with a diphthongized nature.

The boundary context with the consonant /t/ probably influenced the linear correlation between the two variables. A wider spread of nasal resonances that occurs due to the silence of the stop can be predicted.

The influence of boundary edges and context gradience can trigger varying degrees of nasality in the speech sounds Turnbull, Seyfarth, Hume & Florian Jaeger (2018), leading to a greater or lesser nasality gradience. This process can result in the lengthening or reduction of the nasal appendix, which is the nasal portion of a diphthong. In the context of the study, this could potentially explain some of the variations observed in the velum closure and velopharyngeal port closing movement during the production of nasal diphthongs [ãw̃] and [ẽj̃].

#### **4.7 Partial Conclusion**

We conclude that, in BP, nasal diphthongization is the result of the timing variation in the velum mechanism of opening and closing movements associated with the velopharyngeal port movement. The coordination between each mechanism affects the production of nasal vowels and diphthongs.

The soft palate moves cyclically during speech production and is linked temporally to oral production, even the soft palate movement and the oral gestures (e.g., tongue movements for vowel articulation) do not have identical timing or are not

perfectly synchronized, because there is some temporal difference between the two movements. The evidence for the temporal linkage between the soft palate and oral production could be observed in the timing of peak velum opening, which occurs at the end of the acoustic vowel interval.

The velum maintains a consistent behavior observed in the velum's movement during the onset of nasal diphthongs. In 80% of cases, the soft palate opening begins with a plateau, which means that during the velum opening movement, there might be some minor oscillatory motion or small vibrational patterns observed in the velum tissue, but after the onset the opening movement is stable.

In most cases, the velum begins its movement after the vowel onset; thus, there is a temporal delay or difference in the timing between two events, in this case, between the onset of the nasal vowel and the beginning of the velum movement. It means that the velum opening does not start exactly at the same time as the nasal vowel onset, but slightly later.

The velopharyngeal port begins its movement before or after the nasal onset. The velopharyngeal port movement is not stable and is not synchronized with the velum displacement.

The velopharyngeal port displacement had a smaller amplitude than did the soft palate displacement. The duration of the velum trajectory refers to the time it takes for the velum to complete its opening and closing movements during the production of nasal diphthongs. It appears that this duration is similar or uniform between the velum displacement (soft palate movement) and the lateral pharyngeal walls (velopharyngeal port movement).

Our data showed that the velum trajectory and the narrowing of the velopharyngeal port were these parameters play a crucial role or have a significant

impact on velopharyngeal control, which refers to the coordination and control of movements of the velum and the velopharyngeal port during speech production.

In comparison, the anticipatory velum lowering varied depending on the lowering of the soft palate and the velopharyngeal port movement. Thus, the soft palate offset was more coordinated than the onset movement because the closing movement was coordinated with the onset time of the boundary consonant. Thus, the fiberscope results support or confirm the findings obtained from the aerodynamic. The coordination of the soft palate offset in relation to the end of the vowel has a shorter duration in the aerodynamic and articulatory data as well.

Even if the nasal vowel is a monophthong, both segments have a nasal appendix as the result of the tongue rising and a velum carryover articulation because both the velum and the velopharyngeal port complete their movements after the acoustic production of the nasal segment. Figure 8 illustrates the temporal gestural coordination of nasal coarticulation. According to Lovatto (2007), our results could be schematized as follows:

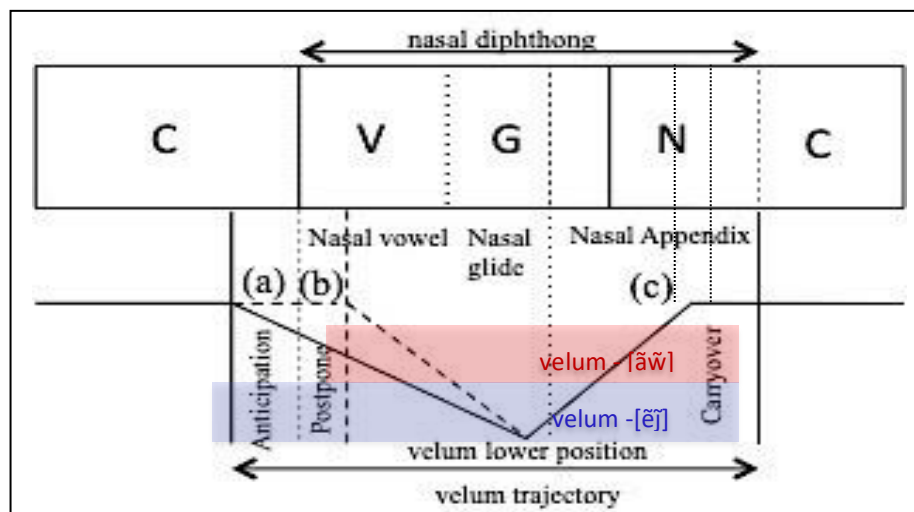


Figure 65: Illustration of the temporal velum trajectory. Adapted from Lovatto (2007).

This illustration shows the gestural synchrony between oral and nasal articulation to mark the end of the nasal diphthongization in the coda. The figure

presents the coordination between the movements of oral articulators (e.g., tongue) and the velum during the production of nasal diphthongs. The top panel represents the segment, while the bottom panels represent the velum activity.

The solid vertical lines represent the specific time points in the gestural coordination, while the dashed vertical lines might indicate the unstable transitions between different phases of segment production.

There is no coordination between the velum lowering and the nasal vowel onset (a, b). In the offset articulation (c), the relationship of the acoustic nasal appendix, the velum closing, and the velopharyngeal port distance were more stable and coordinated in our data due to the left boundary /t/. In all our data, we find a nasalization carryover. The velopharyngeal port close movement sometimes ends after the onset of the following consonant /t/, resulting in a carry-over, the first part of the voiced consonant /t/ is nasalized with nasal spreading resonances.

The lack of coordination suggests that the velum lowering movement does not align precisely with the onset of the nasal vowel. We have measured the timing variability of both the velum lowering onset relative to the nasal vowel onset and the velum target achievement relative to the nasal vowel offset to support this partial conclusion.

There is a temporal variation in velum lowering during the production of the nasal diphthongs [ãĩ] and [ẽĩ]. The variations are observed in both the fiberscope (Fiber) and the velopharyngeal port movement (PNG) data.

For [ãĩ], the velum lowering onset shows anticipation in several instances, with durations ranging from 21 ms to 232 ms (Fiber) and 13 ms to 307 ms (PNG). Additionally, some cases exhibit postponement with durations from 3 ms to 44 ms (Fiber) and 13 ms to 152 ms (PNG).

Similarly, for [ẽĩ], the velum lowering onset displays anticipation with durations from 22 ms to 146 ms (Fiber) and 22 ms to 25 ms (PNG). postpone is also observed with durations from 20 ms to 49 ms (Fiber) and 24 ms to 49 ms (PNG).

These variations in velum lowering durations suggest different degrees of coordination and timing in the opening movement of the velum during the production of the nasal diphthongs [ãĩ] and [ẽĩ].

The temporal variation in velum closure during the production of the nasal diphthongs [ãĩ] and [ẽĩ] showed more stability. The mean for the movement of the velum closure in the fiber data from [ãĩ] and [ẽĩ] was 28 ms and 25ms, respectively. We observed in both the fiberscope (Fiber) and the velopharyngeal port movement (PNG) data, for [ãĩ], the velum closure durations range from 71 ms to 187 ms (Fiber) and 14 ms to 42 ms (PNG); for [ẽĩ].

We observed the velum closure durations range from 6 ms to 165 ms (Fiber) and 16 ms to 40 ms (PNG), this depends on the anticipatory or postponement of the velum movement during the vowel onset time. But the average durations to close movement for velum closure are 138 ms (Fiber) and 80 ms (PNG) for [ãĩ], and 80 ms (Fiber) and 25 ms (PNG) for [ẽĩ].

These variations in velum closure durations indicate differences in the timing and coordination of the closing movement of the velum during the production of the nasal diphthongs [ãĩ] and [ẽĩ].

So, this study can help us understand the movement of the velum and the velopharyngeal port during the nasal diphthongization occurring in the coda position of the syllable, which is typically word-final position.



We observed differences in the timing of velum lowering and the nasal vowel, the speech production system may make compensatory adjustments to achieve a proper degree of nasalization and the presence of a nasal appendix.

In nasal front vowels, nasalization starts at the same time as the nasal vowel is released, whereas in back diphthongs, the release of the vowel is more oral, suggesting a different pattern of nasalization timing. Thus, this different type of phasing between the vowel and the opening of the velum must be compensated for by the presence of the nasal appendix to maximize the perception of the nasalization (BEDDOR et al., 2013).

## CHAPTER 4: ARTICULATORY ANALYSIS

---

### Résumé (Summation)

This chapter aims to analyze the articulation of the oral and nasal diphthongs of BP through articulatory data collected by the Electromagnetic Midsagittal Articulography 2D device. We will provide a brief overview of the instrument used. We will then describe the methodology and how our data were acquired through the EMMA 2D device. Finally, we will explain the details of the displacements of the electromagnetic pellets analyzing the following parameters from our data: (a) Velum temporal-displacement trajectory; (b) Trajectory of temporal displacement of the body and back of the tongue; (c) Spatial variation and (d) durational rates.

### 5. Introduction: Articulatory Analysis

Speech production is an intricate process that involves precise coordination between various articulatory organs. Understanding how sounds are formed and manipulated within the vocal tract is essential for gaining insights into the phonetic and phonological properties of languages. Brazilian Portuguese (BP), like many other languages, exhibits a remarkable phonetic feature known as nasalization, where certain nasal vowels and nasal diphthongs present a nasal appendix after the maximal lowering of the velum, associated with an oral constriction that leads to a great volume of airflow through the nasal cavity. This phonetic phenomenon gives rise to the characteristic nasal resonance found in nasalized speech in the Paulistano dialect.

In this study, we focus on the articulation of front nasal vowel and back nasal diphthongs in BP. Specifically, we analyze articulatory data collected using the Electromagnetic Midsagittal Articulography 2D (EMMA 2D) device to explore the

movements of the tongue and soft palate during the production of /ẽ/ and /ãw̃/. These nasal sounds are particularly diphthongized, as they represent distinct phonetic elements nasal vowels, and nasal diphthongs, this contributes significantly to understanding the phonetic and phonological system of the language.

The motivation behind this chapter stems from the intricate nature of articulatory processes involved in producing nasal diphthongs in BP. While nasalization has been studied in various languages, understanding the precise articulatory strategies underlying the production of nasalized diphthongs remains a complex task. To address this, we explore the role of lingual adjustment in nasalization, the trajectory of the velopharyngeal opening, and the interaction of both gestures.

The analysis of lingual adjustment helps us to comprehend how the tongue dynamically adapts its position to achieve accurate oral targets to maximize the volume of nasal airflow while maintaining the integrity of the target diphthongs. Additionally, by comparing the nature of gestures in oral and velum movement, we aim to gain a deeper understanding of how nasalization is impacted by the tongue and the soft palate movements, comparing it to purely oral articulations.

The primary objective of this study is to describe the timing of the articulation onset, target e offset of the tongue, and velum gestures of front nasal vowel and back nasal diphthongs in the Paulistano dialect. Utilizing data collected with the EMMA 2D device, we seek to comprehend the specific articulatory gestures involved in producing the nasal diphthongization /ẽ/ and /ãw̃/. Furthermore, we aim to compare these gestures with those observed in oral diphthongs (/ej/ and /aw/) to identify any distinct or similar articulatory patterns between oral and nasal segments.

The findings from this research hold broader implications for the fields of phonetics, phonology, and articulatory studies. Understanding the articulatory and

nasal diphthongs contributes to a more comprehensive understanding of phonetic variations and phonological patterns in BP.

The use of EMMA 2D in articulatory analysis offers a valuable and non-invasive tool to investigate speech production processes. EMMA provides researchers with real-time data on the movements and positions of articulators, enabling a detailed examination of the tongue, soft palate, and other relevant articulatory structures during speech production. This technology has been widely employed in phonetics and speech research to study various speech sounds, including diphthongs, vowels, consonants, and coarticulatory phenomena.

One of the primary advantages of using EMMA 2D to analyze articulatory is its ability to provide precise articulatory data. The system utilizes electromagnetic sensors to track the movement of tiny electromagnetic pellets placed on articulatory structures such as the tongue and soft palate. This allowed us to obtain detailed information about the spatial and temporal positions of the tongue and velum articulators during the target segment production.

EMMA 2D facilitates the study of coarticulation effects, which refer to the influence of neighboring sounds on the production of nasal diphthongization. In the context of nasal diphthongs, neighboring oral and nasal gestures impact each other, leading to coordinated movements of the tongue and soft palate.

By analyzing the articulatory data captured by EMMA 2D, we explored how coarticulatory effects contribute to the production of nasal diphthongization, enhancing our understanding of the intricate processes involved in nasalization.

While EMMA 2D is a powerful tool, it also has some limitations. The placement of electromagnetic sensors on articulators might cause slight discomfort, potentially affecting speech production. Additionally, the small sample size of participants and

the need for specialized equipment and expertise can limit the accessibility of EMMA 2D.

One observation in our data is the presence of partial constriction of the oral cavity during nasal diphthongization. In some cases, total constriction of the oral cavity is also possible, particularly due to glide constriction in the palatal or velar region.

Our analysis highlights differences in tongue movement and velum lowering between different informants (RD and GV). For example, RD demonstrated a more stable tongue-displacement trajectory than GV. Additionally, the nasal diphthong [ãũ] had a higher and more back target position than the corresponding oral diphthong [aw] for both informants. These findings suggest that nasalization involves specific adjustments in tongue and velum movements to achieve the anticipated nasal resonance.

Our study also observed vowel breaking in the front nasal vowel [ẽ], resulting in the formation of two elements: the middle vowel [ẽ] and a palatal glide [j]. This vowel breaking mirrors the articulatory production of the front oral diphthong, confirming the production of oral cavity constriction in the palatal region during nasal diphthongization.

Regarding the velopharyngeal movement, it was observed that the beginning of back nasal diphthongs is initially oral before becoming nasalized. The velum trajectory followed a diagonal displacement, starting from a back region and progressing through a central region. The coordination between the soft palate gesture and the tongue gesture, with different activation times, contributed to the nasalization process.

The EMMA 2D data provided detailed information on the timing, duration, and coordination of articulatory gestures involved in nasal diphthongization. The analysis of this data allowed for a deeper understanding of the complex processes underlying

nasalization and its impact on speech production in Brazilian Portuguese. Moreover, the comparison between oral and nasal segments, as well as between different informants, contributed to the refinement of theoretical models of speech production, particularly in the context of gestural phonology.

Our articulatory analysis using EMMA 2D has provided valuable insights into the production of nasal diphthongs in Brazilian Portuguese. The findings shed light on the specific articulatory adjustments and movements involved in nasalization, offering a more comprehensive understanding of the phonetic and phonological properties of nasal diphthongization.

We hope that this research can contribute not only to the field of phonetics, phonology, and speech production but also has potential applications in language acquisition, speech pathology, and cross-linguistic studies of nasalization in natural languages. Studying the interaction of velarization or palatalization with velum lowering can reveal how these articulations impact the nasality of diphthongs and how they influence the articulatory adjustments required for nasalization.

In the following sections of this chapter, we present a detailed overview of the methodology, data collection process using the EMMA 2D device, and the analysis of qualitative and quantitative data. Subsequently, we discuss the results in the context of gestural phonology and control theory, offering valuable insights into the intricate relationship between articulatory gestures and nasalized speech. Finally, we conclude by summarizing the critical points addressed and the outcomes obtained in this study, thereby contributing to a broader understanding of nasal diphthongization in Brazilian Portuguese and its relevance beyond linguistic boundaries.

## **5.1 Objectives**

Nasalization is a distinctive and complex phonetic feature found in many languages, including Brazilian Portuguese. Nasal diphthongs involve intricate articulatory adjustments, as they combine both oral and nasal elements. Our motivation for describing the articulation of front nasal vowels and back nasal diphthongs in the Paulistano dialect is to compare the nature of gestures in oral and nasal segments.

Nasalization can introduce variations in speech sounds, affecting their acoustic properties and phonetic realization. By studying front nasal vowel and back nasal diphthongs, we can investigate how different nasalized speech sounds are articulated and how their acoustic properties differ from their purely oral counterparts. This exploration of phonetic variation is essential for characterizing the specific articulatory patterns unique to nasalization in BP.

Nasalization involves coarticulation effects, where neighboring articulatory gestures influence each other during speech production. By comparing the nature of gestures in oral and nasal segments, we can gain insights into how nasalization impacts the coordination of tongue and soft palate movements.

Investigating the articulation contributes to the broader phonetic typology of nasalization in natural languages. This knowledge can aid in comparing nasalization patterns across different languages and identifying universal or language-specific articulatory strategies related to nasalized speech sounds.

Our data allows other linguists to explore cross-linguistic similarities and differences in articulatory patterns. This analysis can reveal how nasalization is realized in the target segments produced in the Paulistano dialect and can provide insights into the phonetic and phonological universals and language-specific variations related to nasal diphthongs.

Overall, the motivation for these specific objectives is to advance our understanding of nasalization in Brazilian Portuguese and its articulatory complexities. By focusing on front and back nasal diphthongs and comparing them with oral diphthongs. Our findings can have implications not only for the study of phonetics and speech production but also for language-specific phonological analyses and cross-linguistic investigations of nasalization in natural languages.

## **5.2 Research Questions**

The motivation for studying the role of lingual adjustment in nasalization stems from the complexity of nasal diphthongization. Nasal diphthongs involve the coordination of both oral and nasal articulatory gestures. The position and shape of the tongue play a crucial role in achieving the desired oral-nasal balance during the production of nasalized speech sounds.

Lingual adjustment suggests the dynamic and precise movement of the tongue to achieve the appropriate oral-nasal balance during the production of nasal diphthongs. By investigating the role of lingual adjustment, we aim to understand how the tongue adapts its position and shape to produce nasal diphthongs accurately.

The velopharyngeal trajectory is another crucial aspect to study in the context of nasal diphthongization. The velopharyngeal port plays a role in controlling the airflow into the nasal cavity during the production of nasalized speech sounds. Understanding the trajectory of the velopharyngeal port provides valuable information on the movements of the soft palate (velum), onset, target, and offset.



This investigation of the velopharyngeal trajectory can shed light on how the velum opens and closes during nasal diphthongization and how it interacts with the oral articulators.

Velarization and palatalization is a hypothesis that refers to the result of tongue articulation during the production of the back nasal and front nasal vowel target segments. The interaction of these articulations with velum lowering influences the overall articulatory patterns during nasal diphthongization. This interaction can reveal how these articulations impact the nasality of diphthongs and how they influence the articulatory adjustments required for nasalization. Our research question aims to explore:

- The role of lingual adjustment in the nasalization.
- The velopharyngeal trajectory.
- The interaction of velarization or palatalization with velum lowering.

### **5.3 Theoretical Framework: Velum Activity**

Laboratory phonology is a hybrid experimental methodology that investigates the relationship between phonetic and phonological components through experimental approaches and aims to address all speech components, such as the biodynamic and cognitive aspects of human speech. This methodology is looking for phonological patterns through interdisciplinary scientific experiments to understand human speech in its organizational structure.

We aim to analyze the nasal diphthongization from a biomechanical perspective. We aim to examine how the articulators are physically arranged, shaped, and configured to improve articulatory efficiency and the dynamics of the physical movement of the tongue and velum during speech. Examining how the articulators

help us to understand the organization of the articulatory gestures involved in speech production, specifically the tongue and velum.

The Electromagnetic Midsagittal Articulography 2D is used to collect articulatory data during speech production. EMMA 2D tracks the movements in the spatiotemporal displacement of small electromagnetic pellets attached to specific points on the tongue and velum. The EMMA 2D system records these pellets, providing detailed information on the positions and movements of the tongue and velum during speech production.

EMMA 2D allowed us to examine the articulators in real-time data on the spatial arrangements and movements of the tongue and velum. The system captures the 2D displacement of the electromagnetic pellets, allowing us to visualize and analyze the precise movements of these articuli in the vocal tract.

EMMA 2D provides information on the exact positions of the electromagnetic pellets attached to the tongue and velum, allowing to map out the spatial configurations of these gestures during speech production. The system tracks the temporal displacement of the pellets, revealing how the tongue and velum move during the production of nasal diphthongs. This information helped us to understand the dynamic articulatory patterns involved in nasalization, and even confirm our initial hypothesis of velarization and palatalization of the target segments. EMMA 2D also allowed us to examine how the tongue and velum coordinate their movements during speech production, providing insights into the coordination of articulatory gestures involved in nasal diphthongization, by comparing the articulatory data from oral and nasal segments. Overall, our data provided valuable information to understand the biomechanics of nasal diphthongization and the coordination of articulatory gestures involved in producing nasalized speech sounds.

We aim to understand how speakers use the articulators to evaluate the configuration of the tongue, the velum, and the vocal tract geometry during speech production: The main objective of the study is to investigate how speakers utilize their articulators (tongue and velum) to understand the specific configurations and movements of these articulatory organs during speech production. This includes exploring how the tongue and velum coordinate with each other and the overall vocal tract geometry to achieve nasalization of specific speech sounds.

Thus, through static and dynamic EMMA data on the trajectory of the tongue and velum, it is possible to understand how the speaker controls temporal gestural orchestration, observing articulator behavior during speech production.

This articulatory study with images helps verify articulatory compartment, giving a concrete analysis of articulatory variations. Using the EMMA device, it is possible to visualize the displacement of the tongue and velum on the midsagittal plane during articulation. This plane is typically used for tongue and velum configuration analysis (BERTI et al., 2016).

The articulatory data is a kind of approach that helps to understand the spatiotemporal articulation, as the manner, the articulatory point, and the secondary articulations in sounds produced in the back region of the oral cavity.

The acoustic signal results from the vibration of the vocal folds and the configuration of the oral-track geometry; the variations in the source and the filter result in changes in the quality of the acoustic output.

Because the oral cavity has active and passive articulators, it is not easy to determine how they behave. The acoustic spectrogram provides clues about the articulators' performance but not the exact picture of the oral geometry.

Tongue and velum articulation display highly complex neuromuscular activity (INOUE, ONO, HONDA & KURABAYASHID, 2007; LADEFOGED & MADDIESON, 1996). Motor orchestration is complex, but it is essential for nasalization.

Vocal nasalization is generated through the modulation of the tongue, which partially obstructs the airflow in the oral cavity associated with soft palate movement. A series of adaptations and specific muscle activations are necessary (LADEFOGED, 2001; TITZE, 2008) to perform the speech.

According to Song & Schwarz (2009), speech combines oral motor fluency and phonetic symbolism through the muscular dynamics of swallowing and expectoration. The speech-motor system results from the evolution of the respiratory and digestive systems and shares the same anatomy, which later was adapted to language development (ROZIN, 1999).

The pharyngeal cavity is used for breathing, feeding, and vocalization. Motor control of this cavity is essential for nasalization, as it is composed of the tongue, the soft palate, the pharyngeal walls (posterior and lateral), and the epiglottis. The pharynx controls the breathing airway and, when contracted, pushes the bolus into the esophagus for swallowing.

The soft palate moves cyclically during eating, respiration, and speech. In speech, soft-palate movement can be associated with tongue and jaw movement.

The soft palate muscles, tongue, and pharyngeal walls control the airway, controlling ventilation and gas exchange (KOKAWA, SAIGUSA, AINO, MATSUOKA, NAKAMURA, TANUMA, et al., 2006) during breathing. However, their motor controls also produce nasal sounds in speech.

The soft palate is lowered during nasal breathing, opening the velopharyngeal section. During oral breathing, the velum is closed, separating the nasal cavity from

the pharyngeal airway. The levator veli palatini and palatoglossus muscles are activated to determine the position of the palate and the type of breathing. During pharyngeal swallowing, the soft palate is closed.

During swallowing and chewing, the soft palate occupies different temporal relationships with jaw and tongue movements. The soft palate moves upward when the jaw opens; when the jaw closes, the soft palate moves downward. Palatal movement varies individually, but the timing of palatal movement and jaw movement is highly consistent (KOKAWA, SAIGUSA, AINO, MATSUOKA, NAKAMURA, TANUMA, et al., 2006).

Thus, the soft palate plays an essential role in breathing, chewing, and speaking. The movement of the velum is articulated and controlled together with the other articulators, such as the tongue and jaw. The muscles of the velic action are spatiotemporally controlled. Thus, the soft palate and nasal pharyngeal walls, which are also used during speech production, are controlled and coordinated to produce nasalization.

The velum during speech production is cyclical too, just as it is during breathing and chewing. The velum undergoes specific patterns of movement to achieve different speech sounds, including nasal consonants, nasal vowels, and even during normal breathing.

During speech production, the velum has specific onset, target, and offset timing, which means it opens, reaches the target, and closes at precise moments in coordination with other articulators like the tongue and jaw. These coordinated movements allow to production of nasal sounds (nasal consonants, nasal vowels, or nasalized vowels) when the velum is lowered, allowing air to flow through the nasal cavity.

Different degrees of lowering can produce nasal consonants, nasal vowels, or breathing. The degree of velum lowering plays a crucial role in determining the specific speech sound being produced. Thus, the coordinated velum movements in speech production play a vital role in achieving the different speech sounds and this coordination allows for the cyclical and precise control of airflow through the oral and nasal cavities during speech.

## **5.4 Methodology**

This section will explain how we collected and analyzed the data to examine the vertical velum activity, the velum's lowering and closing movements, and tongue displacement during the production of nasal diphthongization. The EMMA is a high-quality device used in speech-language research.

We used the EMMA 2D device to record articulatory data during the production of nasal diphthongs in Brazilian Portuguese. This allowed them to examine the vertical velum activity, velum's lowering and closing movements, and tongue displacement. The data collected from two female speakers were processed and analyzed to gain insights into the mechanics of nasal diphthongization in speech production.

### **5.4.1 Material**

In this study, we propose to analyze nasal diphthongization through the dynamics of the articulators during speech production. Thus, we set a comparative corpus to be collected in an *Electromagnetic Midsagittal Articulography* 2D device.

The data were recorded in an acoustic room in Gipsa-Lab<sup>72</sup>. Carsten's EMMA system (AG100) was used.

To use the 2D EMMA system, reference pellets must be placed on the upper incisor and in the nose region to control head movement. This mapping is essential to the conversion of the signal data into numerical data.

The calibration of the 2D system is achieved using the variation between the nose sensor and the upper incisor. Some studies show that the 2D EMMA system may produce inaccurate or erroneous data if there is a misalignment between the position of the lateral pellets (those placed on the sides of the articulators) during data collection. This misalignment can lead to discrepancies in the measurements of the articulatory movements, potentially affecting the accuracy of the recorded data. This possibility does not imply that the system is less effective than the 3D system (BRYAN, 2002).

Comparisons between EMMA 2D and 3D EMMA data have found that they are similar along the x-y axis (SERRURIER & BADIN, 2005). The recording made via the alignment of the sensors in both systems relies on the articulators' trajectories in time.

Both the systems showed variation in signal amplitude from the articulatory trajectories of the tongue-tip sensor (BRYAN, 2002). 2D and 3D EMMA systems displayed differences in the amplitude of the recorded signals when tracking the movements of the tongue-tip sensor. This variation in signal amplitude implies that there might be slight differences in the measurements of tongue movements between the two systems. These differences could potentially influence the precision and

---

<sup>72</sup> The GIPSA-lab, Grenoble Images Speech Signal Automation, is a research unit of CNRS, Grenoble-INP, and the University of Grenoble-Alpes. It conducts theoretical and applied research on signals and systems. The Emma 2D data for our study was collected in "Systèmes Linguistiques, Dialectologie & Oralité" (SYLDO) department, that is focused on linguistic systems, dialectology, and orality.

accuracy of the data obtained from each system, especially when studying fine articulatory details. However, it is not possible to say if these variations represent different strategies employed by the speaker, because both devices can suffer magnetic variations if the sensors are positioned in the soft palate region.

In addition, the images and the overall shapes of the alignment generated by the two systems are similar. Thus, we collected our data with both systems. However, the 3D data displayed some signal interference, and we discarded it.

It is necessary to calibrate the EMMA 2D device before the collection session to minimize the loss of information and to obtain an extensive and comprehensive characterization of the results.

#### **5.4.2 Equipment**

As mentioned, we used EMMA, developed by Carstens Medizinelektronik, to register the articulatory data. The EMMA AG100 (EMMA 2D) makes possible the visualization of the articulator position and the movement inside the mouth in the midsagittal plane.

This technique allows large amounts of data on the articulator behavior in a 2D trajectory through the x-y coordinates to be obtained.

The EMMA is a kinematic tracking system that quantifies articulatory movement through small electromagnetic variations in the field. This creates a magnetic field which induces alternating currents of different frequencies in the sensors attached to the articulators. The electrical signal resulting from the magnetic variation is transmitted to the transducer and digitized.

The signal captures the articulators' natural movements in real time (KRAKOW & HUFFMAN, 1993). The device has 12 pellets. Three of these are placed in the head,



one is placed in the nose region, and the others are placed on the articulators.

To enable the data collection process, the speaker wears a heavy helmet, which plays a vital role in generating the alternation in the magnetic field necessary for the system's functioning and helps to calibrate the head movement and rotation, the pellet's spatial displacement during the speech data collection. The photos of the speakers wearing the helmets during the recording sessions are below in Figures 66 and 67.

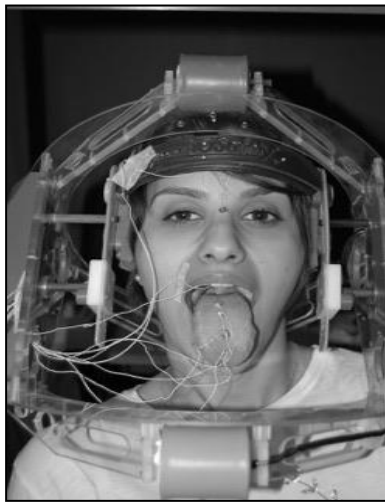


Figure 66: Speaker RD wearing the helmet used to calibrate the 2D EMMA system.



Figure 67: Speaker GV wearing the helmet before the recording session.

Head movement and rotation were determined and calibrated by the position of the helmet that the speaker wore during data collection as we mentioned, and it can be seen in Figures 66 and 67.

These images give an idea of the helmet size and the position of the sensors placed on the speaker. The speakers quickly adapted to the sensors. No one complained about discomfort or had a gag reflex.

Engineer Christophe Savariaux dried the informant's tongue with gauze and applied special glue to the articulation points we wanted to analyze. At no time did the speakers feel any discomfort.

We did not ask the informant to produce or sustain a nasal vowel to attach the sensor to the soft palate. The only thing we requested was that the informant not retract their tongue towards the soft palate to avoid the sensors touching and potentially causing a gag reflex, which could happen if the sensor glued on the tongue dorsum touched the uvula or the soft palate.

The EMMA 2D system is generally stable, but interference occurred between the pellets in the back region (specifically, the pellet on the tongue back and the velum) due to their proximity. This interference was observed during the third repetition of the data collection session for speaker GV, and these particular data were discarded. Interference was identified through visual inspection of the recorded data..

### **5.4.3 Corpus**

We set a corpus to compare the articulatory movements of front oral vowels, nasal diphthongs, and nasal vowels in the syllable nucleus position. We used the same group of words from the fiberscope study. Thus, we used the same words to make our generalizations more homogenous. The corpus was composed of monosyllable minimal pairs and similar pairs, all real words, as can be seen below..

Table 18: Set of words for back oral and nasal diphthongs.

Oral Diphthongs		Nasal Diphthongs	
Stimulus	Gloss	Stimulus	Gloss
[paw]	wood	[pãw̃]	bread
[maw]	bad	[mãw̃]	hand
[taw]	these	[tãw̃]	so
[saw]	salt	[sãw̃]	sane
[kaw]	lime	[kãw̃]	dog

Table 19: Set of words for front oral diphthong and front nasal vowels.

Oral Diphthongs		Nasal Vowels	
Stimulus	Gloss	Stimulus	Gloss
[dej]	give (1 <sup>st</sup> SG)	[těj]	have (3 <sup>rd</sup> SG)
[sej]	know (1 <sup>st</sup> SG)	[sěj]	without
[nej]	Ney (nickname)	[něj]	neither
[frej]	friar	[hěj]	REM
[lej]	law	[běj]	good

Each target word was inserted into the control carry sentence: “Digo \_\_\_\_\_ todo dia”: [dʒigʊ\_\_\_\_tod dʒie] (“I say \_\_\_\_\_ every day”) to register all data. The use of the dental stop /t/ following the target word is done to neutralize articulatory differences that may occur due to the context of nasal diphthongization. The dental stop /t/ is a non-nasal sound that requires the velum to be closed, preventing any nasal airflow during its production.

By inserting the /t/ after the target word, we also prevent the velar assimilation that can occur, if the context was /k/, avoiding the assimilation processes that could happen in aerodynamic study. By carefully controlling the context and articulatory environment, we can ensure more precise and accurate data collection during the study of nasal diphthongization.

During data collection, the sentences were presented randomly on a microcomputer screen. In this study, there were more repetitions because of the complexity of the recording type.

Each sentence was read and repeated five times. We also requested the speakers to repeat the sentences at a standard speech rate. Each set resulted in 50 tokens for analysis.

#### **5.4.4 Speaker**

In this study, two female speakers, both approximately 30 years old, took part as participants. They were Ph.D. students and native speakers of Brazilian Portuguese (BP), specifically the Paulistano dialect spoken in the city of São Paulo. The recordings were conducted during separate data collection sessions. Both speakers had been living in France for less than two years at the time of the study.

To quantify and analyze the data, we assigned the notation "RD" to the speaker from the first recording session and "GV" to the speaker from the second session. RD is the same speaker who participated in the aerodynamic and fiberscope studies mentioned earlier. Additionally, RD is also the author of this thesis and took part in all the experiments conducted throughout the research.

The EMMA method is considered invasive, as it requires placing small electromagnetic sensors on various articulators, such as the tongue, lips, jaw, and velum. However, some speakers may have a more tolerant gag reflex than others, making it feasible to collect data and quantify the displacement of these articulators.

#### **5.4.5 Data Processing**

Acoustic and articulatory data were collected synchronously in the same recording and on the same time scale. The speaker stayed in a room with acoustic treatment. Data collection was performed and supervised by the engineer Christophe Savariaux at Gipsa-Lab at Université Stendhal in Grenoble, France.

EMMA data were processed by a Chebyshev low-pass filter in MATLAB. The Chebyshev filter is widely utilized in audio, communication systems, and medical equipment (e.g., EEG machines) due to its steep roll-off and adjustable cutoff frequency. The cutoff frequency, which specifies when the filter begins attenuating the signal, is a critical parameter typically provided as a range (e.g., 10 Hz to 100 kHz).

We used Chebyshev low-pass filter filters to sharp transition to remove high-frequency noise and artifacts from the data, resulting in a smoother representation of the articulatory movements.

Chebyshev filters are known for their sharp transition between the passband and stopband, which allows them to achieve a specified level of variation in the passband with a lower-order filter compared to other filter types.

The qualitative and quantitative analyses of the analysis parameters were developed through TRAP homemade software (SAVARIAUX, 2004), a MATLAB interface software developed by the engineer responsible for the recording session.

Trap (Traitement des Signaux de Parole) is a custom-made software developed by SAVARIAUX in 2004. This software is designed to process and interpret the signals obtained from EMMA 2D, EPG, ISIS, WAVE, and Matlab Binary data. The signals acquired through electromagnetoencephalography (EMA) using the Carstens EMA system is processed and analyzed using the TRAP software, which comes in two versions: v3 with five coils (including two reference coils) and v4 with ten coils (including two reference coils). The positional data (x, y) and tilt of each coil are stored in .0xx (v3) or .mat (v4) files, while the corresponding acoustic signal is saved in .mxx or .wav files.

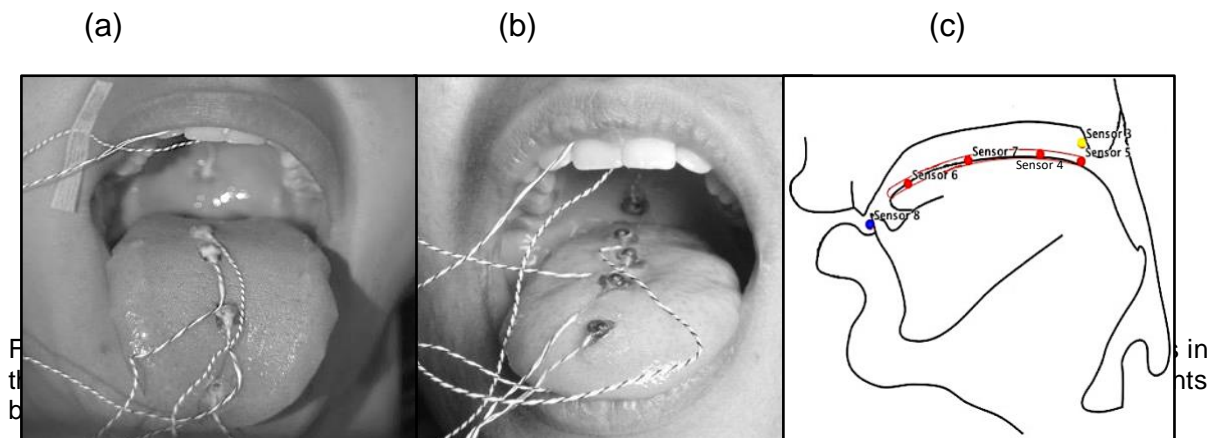
The TRAP software, developed using Matlab, operates in interactive or command-line mode and consists of a main window and a synchronous window synchronized by a temporal cursor.

The software processes trajectories by rotating them to align the hard palate horizontally. Additionally, it performs normalization concerning the reference pellets and applies low-pass filtering.

The software conducts analyses on each pellet's trajectory, including position, velocity, and acceleration, while identifying minima and maxima of position, velocity, or acceleration within the entire window or between two selected time points. Zero crossings are detected, and average values are calculated. TRAP also generates sagittal cuts and animations of the pellets' movements.

The versatility of TRAP allows for efficient and comprehensive signal processing, making it a valuable tool for researchers to analyze articulatory movements during speech production.

Data were recorded using midsagittal EMMA (Carstens AG100) to map the intraoral articulatory movements. The articulatory movements were recorded using pellets attached to the following parts of the tongue: tongue tip (TT; 6), tongue blade (TBL; 7), tongue post-dorsum (TPD; 2), tongue back (TBCK; 5) and the velum (VEL; 3), in addition to the nose (NS), upper incisors (UI), and lower incisors (JAW; 8).



The pellets were placed along the midsagittal plane and were spaced at approximately 1.5 cm. The pellet on the velum was approximately 1.2 cm from the pallet glued in the tongue dorsum.

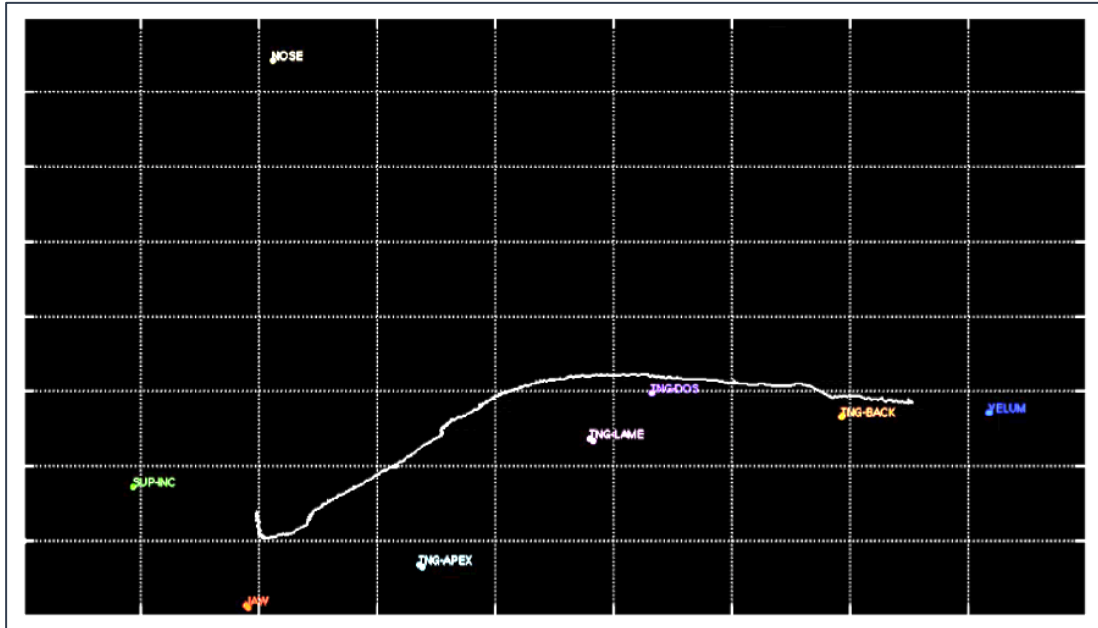


Figure 69: Pellets in rest position in the oral cavity. RD. Midsagittal view of the oral tract.

The pellets were placed along the midsagittal plane and were spaced at approximately 1.5 cm. The pellet on the velum was approximately 1.2 cm from the pallet glued in the tongue dorsum.

The reference pellets aimed to calibrate and control the movement of the head. The audio was recorded simultaneously. Each pellet was, on average, less than 0.5 cm in size. To collect the “bite plane,” the speaker, still with the pellets glued to the articulators, bit a small acrylic plate and remained with the articulators at rest. This measure is the reference point for all pellets. It is the 0 position (zero to the x-axis and y-axis). The bite plane is shown in Figure 70..

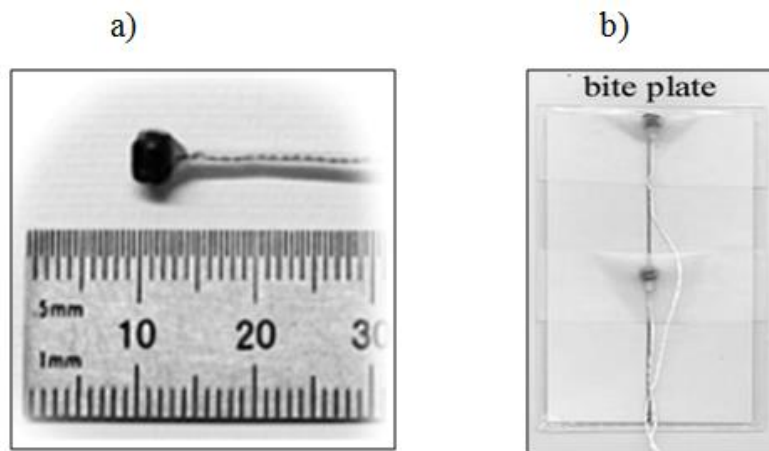


Figure 70: a) Image of the electromagnetic pellets used for the collection of articulatory data, manufactured by the company Carstens. b) Illustration of acrylic plate bitten by the participant at the end of data collection.

The reference point for the EMMA data collection was established at the bite plane. During this process, the speaker, with the pellets attached to the articulators, bit a small acrylic plate and remained with the articulators at rest. This bite plane position was considered the default position or "zero" for all the electromagnetic pellets' movements. The reference point served as an intersection between the x- and y-axes in the Cartesian coordinate system.

From the reference point measure, the researchers were able to extract the spatial size of the oral tract. To calculate the distance between the reference point (bite plane) and the highest point of the hard palate, engineer Christophe Savariaux measured all the extension of the hard palate by passing a pellet through it. This allowed them to determine the spatial size and positions of the various articulatory movements during speech production.

The results of the rest position of the articulators, the bite plane, and the hard palate extension measurements can be seen in figure 69 and illustrated in the x-y coordinates in figure 71, in the next item.

#### **5.4.6 Segmentation**



The images and the general alignment generated in the system are high quality. Thus, we were able to elucidate the intricate articulatory features of the trajectory of the soft palate, which is the main objective of this research.

The values obtained result from the articulators' spatial variation in relation to the bite plane and hard palate measurements.

The size of the hard palate was determined by manipulating a pellet projected across the entire length of the hard palate. In this way, the positive values are projected between the palate and the reference point, while the negative values are below the zero points.

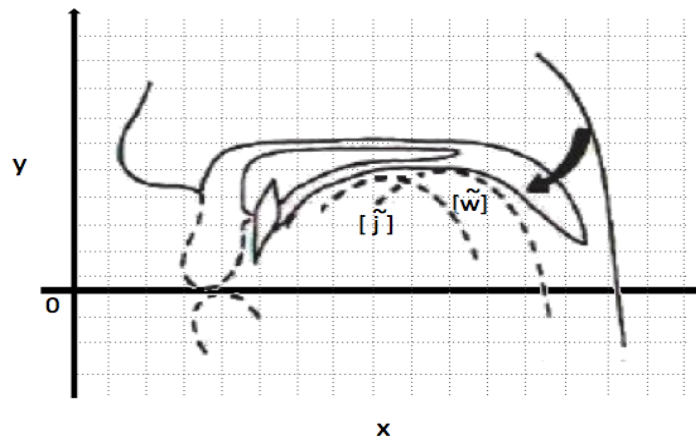


Figure 71: Illustrative image of the phonatory apparatus and the medium sagittal cut placed on the axis of the x-y coordinates. Illustration of the positioning of the target nasal glides.

When using the x-y coordinate system to quantify the movement data of articulators within the oral cavity, it is necessary to define a reference point, which is considered the zero value (x-y). When the articulators move during speech production, their positions are recorded as x-y values relative to this reference point.

By assigning the zero value (x-y) as the reference point, any movement of the articulators that occurs above this point is considered a positive value, and any movement below it is considered a negative value. This quantification was adopted to facilitate data analysis, but it does not mean that the articulator moves along a

trajectory with negative values.

The use of positive and negative values relative to the reference point (zero value) serves only as a convention to represent the movement of articulators in the oral cavity and simplify the analysis of data collected using the x-y coordinate system.

The entire length of the hard palate measure begins at the soft palate surface boundary, passing through the hard part of the palate to the teeth. This line was drawn in our figures to represent the values of the displacement of a pellet on the hard palate surface.

There is tongue displacement to the velar region in back segments, where the back pellet may play a more active role. Thus, we will analyze the pellets placed on the tongue dorsum.

In front segments, the tongue moves in the palatal region, where the post-dorsal pellet may play a more active role. Thus, we will analyze the pellets placed in the tongue body region and on the tongue dorsum. For all tokens, we will analyze the velum trajectory.

We extracted the values from graphics generated by the TRAP program. Each graphic displays the temporal trajectory of pellets on the x- and y-axes. To extract the quantitative data, we selected a specific point on the chart, which automatically generated a table containing all the values on the x-y axis from that location in space for all pellets selected in the window of trajectories.

These values were exported to a table in Excel, where we selected the parameters to analyze. The figure below shows examples of these raw graphics executed by the TRAP program.

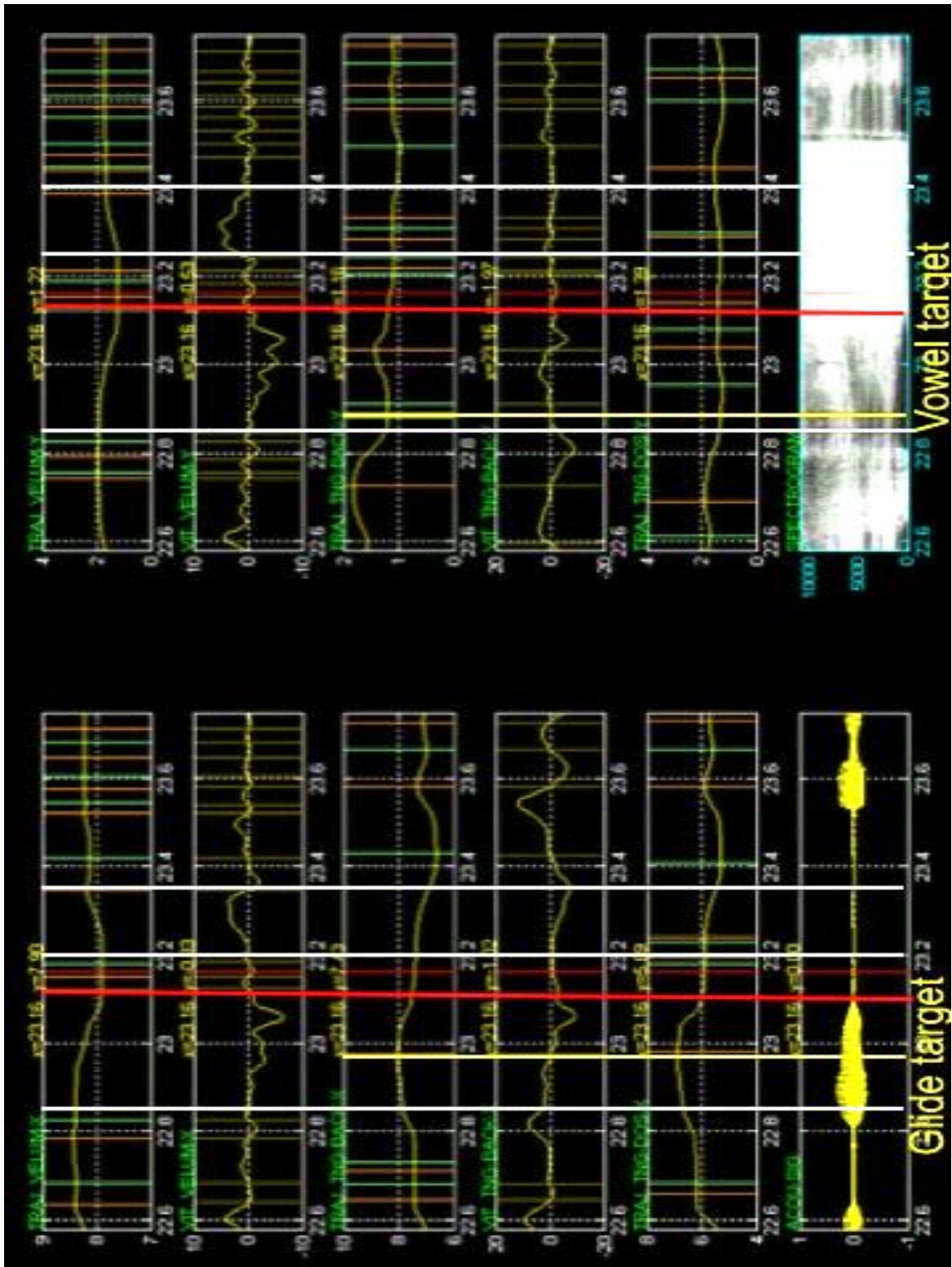


Figure 72: Raw EMMA data. Pellet displacement using the x-y coordinate plane (cm). Targets and articulatory movements [kãw̃]. RD.

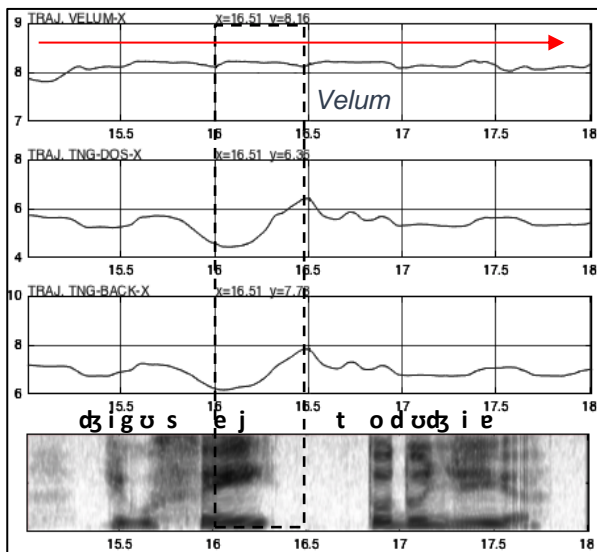
Figure 72 shows the raw graphics of the pellets' displacement trajectories. Each image can combine 10 pellet trajectories, the spectrogram, and the oscillogram. Thus, it is necessary to check the point where the pellet tracing changes direction to extract the vertical/horizontal pellet-movement values.

Each window represents a point on the x- and y-axes. In the graphs, the x-axis represents time and spatial variation horizontally. The y-axis represents the temporal variation in the vertical. White lines represent the onset and offset of soft palate movement. The red line represents the velum target, and the yellow line represents the glide or vowel target.

The displacement of the articulators during speech production can be related to reaching their targets and changing the direction of movement. The displacement of the articulators involves the difference between the final position and the initial position of the articulators. The displacement represents the distance and direction of the movement carried out by the articulators during speech production. We believe that the change of changing the direction of movement is related to gestural reaching the target. Upon reaching the articulatory target, a change in the direction of the articulator's movement may occur, representing a change in the articulation of speech sounds.

The following windows display the pellets' spatiotemporal displacement on the x-y axis. The variables are TRAJ VELUM, VIT VELUM, TRAJ TNG-DOS, and TRAJ TNG-BACK. These initials refer to the name given to the pellet glued in the articulator. The quantitative results are in cm. The movement of the soft palate (y-axis) at the beginning and end of speech production represents breathing. The line trajectories in the graphics represent the pellet's movement..

Velum and tongue **horizontal movements**



Velum and tongue **vertical movements**

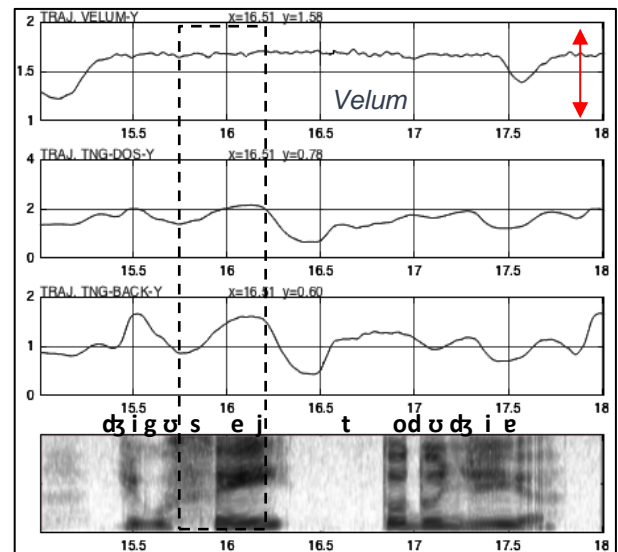
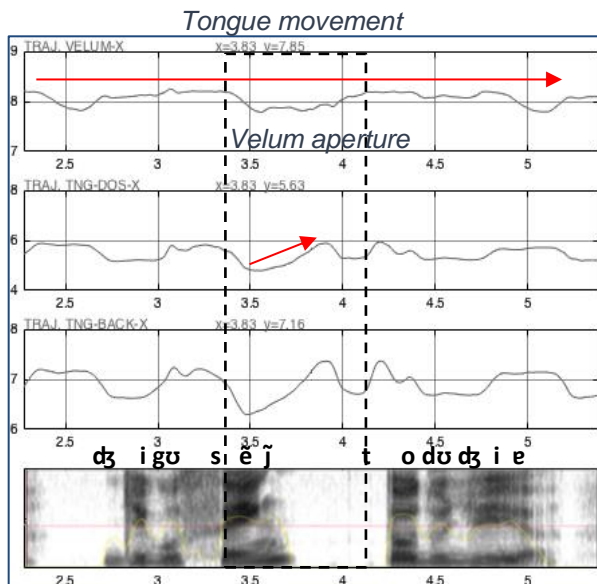


Figure 73: EMMA data. Pellets displacement using the x-y coordinate plane (cm). Targets and articulatory movements [e]. Speaker: RD.

Velum and tongue **horizontal movements**



Velum and tongue **vertical movements**

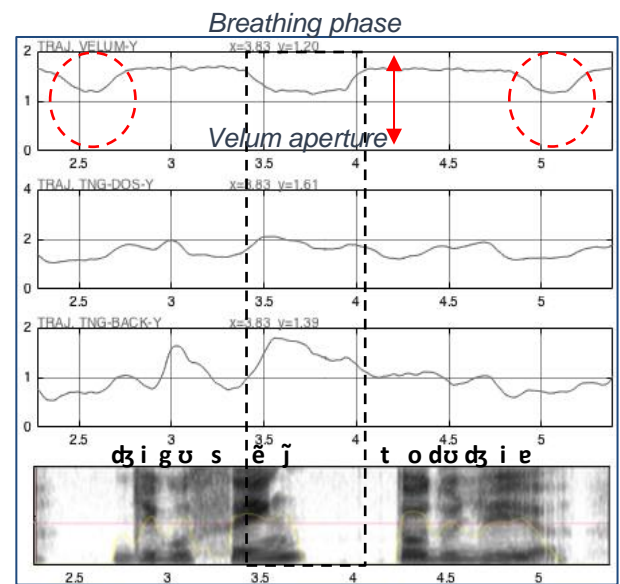


Figure 74: EMMA data. Pellet displacement using the x-y coordinate plane (cm) Targets and articulatory movements [ē]. Speaker: RD.

The red arrows represent the direction of movement of the pellets in horizontal and diagonal displacement. These arrows illustrate the direction of displacement of the articulators during speech production. They indicate the path followed by the pellets as they move within the articulatory system. The horizontal arrows in all

graphics from figure 73 and 74 represents the horizontal velum displacement and the vertical arrows represent the horizontal movement of the velum. The diagonal arrow in Figure 74 indicates the tongue elevation.

The circles indicate the breathing phase at the beginning and end of each sentence production. They mark specific points in the spatiotemporal displacement trajectory of the pellets where the speaker takes a breath before starting to speak (beginning) and after finishing a sentence (end).

The vertical dotted lines represent the temporal duration of the velum during the target segments. In the context of this study, these lines likely demarcate the period during which the velum is actively or not involved in the production of specific speech segments.

These visual representations aid in understanding the articulatory movements and the role of the velum during speech production. They provide valuable insights into the coordination and timing of the articulators involved in the production of various speech sounds.

For the segmentation of temporal duration, we used the same method as in the other experiments. We segmented the oral target production in two segments and the nasal target production in three segments.

To extract the pellet displacement values, we focused on the linear spatial values related to changes in the displacement trajectory of the target segment. The soft palate, when not active, remains mostly in the same position, which is considered the initial position (default).

When there is a change in the displacement trajectory, it indicates that the soft palate is active. The default position (closed) corresponds to the onset of the soft palate, where it begins its downward movement. The maximum position reached

during this movement is considered the target, and when the soft palate returns completely to the initial position, it is considered the offset. These temporal and spatial values are what we extracted to analyze the movements of the soft palate during the production of different speech units.

To extract the target tongue position, we used the same strategy mentioned above. During the production of the target word, when the tongue trajectory changed direction, it indicated that the tongue had reached its target at that moment and initiated a new trajectory to reach another target.

Our data analysis provided valuable insights into the articulatory features of the trajectory of the soft palate and tongue during speech production. It contributed to our research objective and offered valuable information for studying speech motor control and the production of speech sounds.

Understanding the temporal duration of the velum during the target segments allowed us to assess its active involvement in specific target segments. Analyzing the spatial variation and displacement trajectory of the articulators, such as the soft palate and tongue, allowed us to comprehend the patterns of movement involved in producing these articulators.

By studying the tongue trajectory, we gained insights into the precise coordination and movement patterns of the tongue during speech production.

#### **5.4.7 Parameters**

The extraction of our measurements provided insights into the nature of articulatory movements during nasal diphthongization. By examining the velum and tongue trajectories, spatial variation, and the duration of gestural displacement, we

gained a deeper understanding of the coordination and timing involved in target segment production. We extracted the following parameters from our data:

- **Velum Temporal-Displacement Trajectory:** This measurement provides insights into the movement patterns of the velum during nasal diphthongization. We extracted measurements of velum opening closing and target.
- **Trajectory of Temporal Displacement of the Body and Back of the Tongue:** This measurement provides insights about the tongue movement patterns during the production of nasal diphthongs.
- **Spatial Variation:** This measurement provides insights about the spatial variation in the movements of the soft palate and tongue.
- **Duration of Gestural Displacement and Acoustic Duration:** This measurement provides insights about the timing and coordination of articulatory gestures during the production of these diphthongs. The acoustic duration provides a measurement of temporal aspects of speech production.

## **5.5 Results**

In this section, we will analyze the results obtained from EMMA's data. To do this, tables and charts will be presented with averages and standard deviations of all the repetitions performed by the speaker.

As in other experiments, we made a qualitative visualization of the pellets, allowing the extraction of quantitative data and an effective and cohesive descriptive analysis.



The quantitative data compare displacement and timing for both tongue and soft palate movement to oral diphthongs, nasal vowels, and nasal diphthongs. Thus, we analyzed the synchrony between the soft palate's spatiotemporal trajectory and the tongue's spatiotemporal trajectory.

The tongue articulatory region analyzed was in the dorsum and the post-dorsal regions. The reference point for all measures was the 0 cm Cartesian coordinate plane.

The horizontal and vertical orientation used was to determine the oral cavity's size from the calculation of delta spatial variation. To determine the oral cavity length, we calculated the difference between the position of the JAW pellet in the rest position (reference point) and the final portion of the palate size on the x-axis.

To determine the oral cavity height, researchers identified the high point on the hard palate and used it as a reference point on the y-axis. By measuring the distance from this reference point to the soft palate or tongue pellet. This measurement indicates the anterior-posterior dimension of the oral cavity during speech production.

The values for RD are 6.5 cm on the x-axis and 2.10 cm on the y-axis. For GV, the x-axis value is 7.0 cm, and the y-axis value is 2.40 cm. By comparing the values between RD and GV, it was observed that GV had a larger oral cavity than RD.

This comparison provides insights into the articulatory differences between different speakers and how their oral cavities contribute to speech production. The horizontal orientation and spatial variation measurements offer valuable information for understanding the role of the oral cavity in shaping speech sounds and articulatory movements.

The figures below display a visual representation of the raw data generated by the TRAP program. The arrows show the velum pellet, and the trace is the register of the gesture trajectory.

The images presented here are static representations, but in the TRAP program, the visualization is dynamic and continuous over time. This dynamic visualization enables us to observe the articulators' movements in space-time, showing the continuous variations as speech is produced.

To illustrate how the pellets' trajectory occurred during speech production, we carefully selected specific frames from the participant's speech in the articulatory animation provided by TRAP. These selected frames serve as snapshots, allowing us to analyze and explain the movement patterns of the pellets and understand how the articulators coordinate and change positions during the production of speech sounds.

To interpret the type of graphic below, (e.g., in Figure 75) we need to understand that there are electromagnetic pellets glued on the articulators and that they move during speech according to the articulator movement.

Once we understand that there are electromagnetic pellets glued on the articulators and that they move during speech according to the articulator movements, we can observe the displacement trajectory of the pellets in space-time. Thus, the pellets move synchronously with the articulators, resulting in the trajectory traced on the chart.

We observe the displacement trajectory of the pellets in space-time through graphics from the TRAP program. The electromagnetic pellets that were attached to the articulators were tracked as they moved along with the articulatory movements, and the changes in trajectory direction gave us information about the gestural movement.

In the figure below, we can see the spatial path left by the pellet for a certain period – that is, the time the informant was speaking. When this path results in a continuous line, the pellet (articulator) moves slowly. When the line is dashed, it means that the pellet is moving quickly. Figure 75 represents the production of an oral diphthong; we can notice the lack of soft palate vertical movement. The arrow points out the displacement of the horizontal velum trajectory.

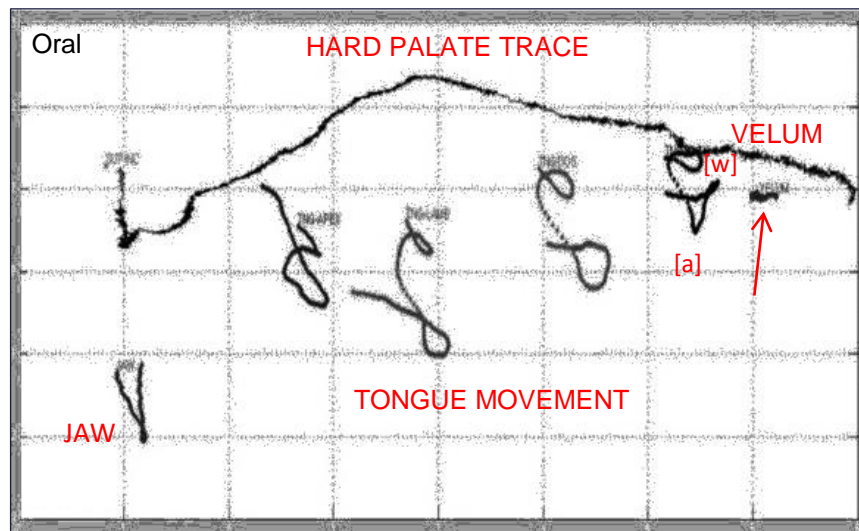


Figure 75: Graphic of x-y coordinates in cm. Spatial displacement of the pellets. Midsagittal view of the oral tract. Stimulus [aw]. Speaker: GV.

During the observation, researchers noticed that the pellet attached to the soft palate vibrated because it is in a region of fibrous muscle tissue that tends to vibrate during speech production. This vibration is a natural characteristic of the soft palate's movement during speech.

When velum (soft palate) is an active gestural in speech production, its displacement can be described as both horizontal and vertical, this kind of movement creates a specific pattern of displacement in the diagonal. Figure 76 shows a nasal diphthong; we can notice the horizontal and vertical soft palate movement. The arrows point out velum displacement. Arrows indicate the velum trajectory.



is vowel breaking, resulting in the presence of the palatal glide. Arrows indicate tongue trajectory from the vowel target to the glide target and the velum trajectory.

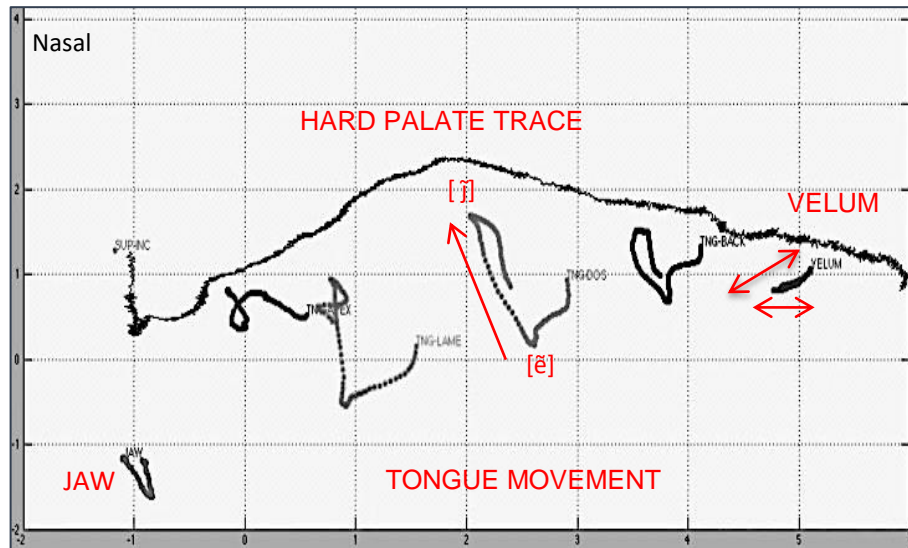


Figure 78: Graphic of x-y coordinates in cm. Spatial displacement of the pellets. Midsagittal view of the oral tract. Stimulus [ɛ̃j̃]. Speaker: GV.

The process of transitioning from the nasal front vowel /ɛ̃/ to the production of the nasal palatal glide [j̃] is the result of the vowel breaking. The raising of the tongue is one of the central characteristics of the nasal diphthongization.

Comparing front nasal and oral segments, the tongue movement, trajectory, and dynamics are similar. Comparing back nasal segments, tongue movement, trajectory, and dynamics are similar. In back segments, the tongue has a different dynamic, the gestures seem to be more accurate.

The gesture amplitude was small for both nasal segments, and the tongue's target position was reached more dynamically in the nasal vocalic segment; the tongue's target was more elevated and was in some cases in a back position.

### 5.5.1 Velum Movement

In this section, we will analyze velum displacement during the nasal segment's production. We noticed that the velopharyngeal sphincter is involved in both the

vertical (opening) and horizontal (closing) movement. Thus, the velum displays a double movement: lowering/closing. The opening movement generates oronasal coupling.

Velum's movement during this experiment performed an oblique trajectory, starting in a back position and moving to a more frontal position when the velum reached its target and it's returned to the original position. The velum's default position is closed.

In (a), below, the velum's temporal displacement on the x-axis shows that the velum moved from a more back region to a more front position. Thus, the velum was moving away from the pharynx. Arrows indicate the fixed point of the glide target, and the key indicates the velum trajectory.

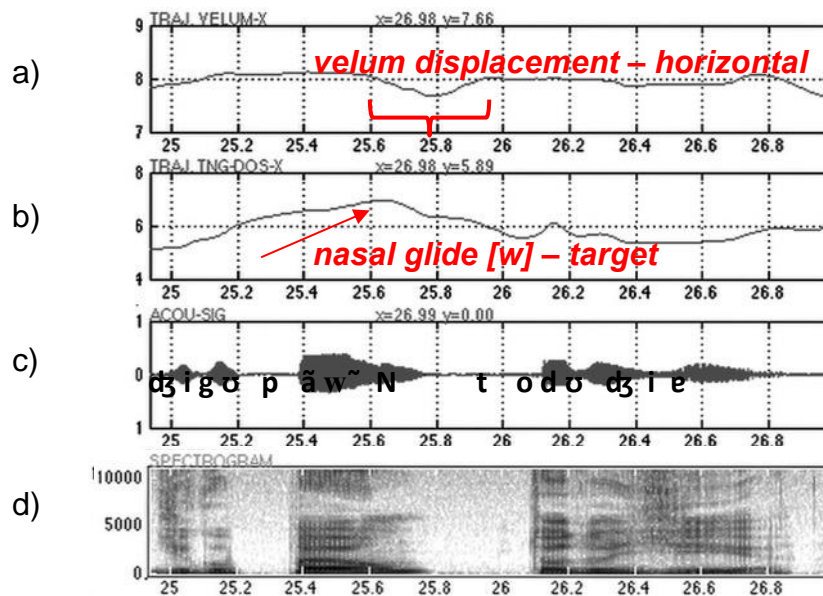


Figure 79: Articulatory trajectory of the velum and the tongue back on the x-axis in cm. Speaker: RD.

The red key in graphic (a) shows on the x-axis the horizontal temporal displacement of the pellet glued to the velum. From this information, it is possible to extract the horizontal aperture parameter from the velum section. In the figure (b), the red arrow represents tongue-back displacement in the horizontal.

Glide production constricts the oral cavity, conducting airflow to the nasal cavity, which dampens the oral resonances, creating a velar nasal appendix. Thus, the resonances continue to pass through the nasal cavity until the production of consecutive consonants begins.

Another factor that can influence the duration of the nasal appendix is a pause or the speed rate. Windows (c) in figure 80 refers to the acoustic signal. This figure shows on the x-y axis the space-time trajectory of the pellet glued to the velum. Arrows indicate the inflection point in the velum trajectory.

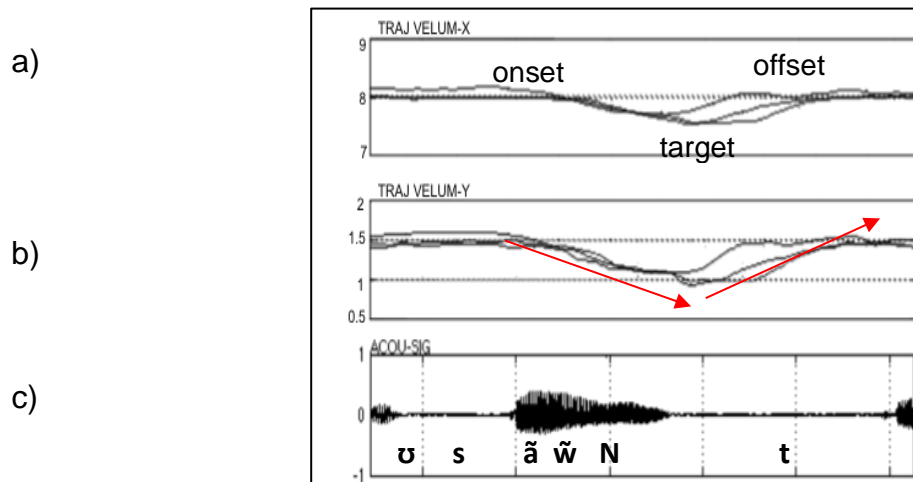


Figure 80: Articulatory trajectory of the velum and the tongue back on the x-y-axis in cm, and acoustic signal. Speaker: RD.

In (a) and (b), the vertical spatial displacement of the soft palate on the y-axis shows the displacement of the velum pellet during the production of the nasal diphthong.

A sharp drop represents velum articulation moving downwards to generate the soft palate opening. The trajectory of the velum can be divided into three phases: opening, maximum opening, and closing of the soft palate.

The three phases of the velum trajectory were likely identified based on the visual analysis of the data obtained from the electromagnetic pellets attached to the soft

palate. These patterns could indicate distinct phases of the velum's movement during speech production:

- **Opening:** This phase likely corresponds to the initial downward movement of the velum until the velum target (maximum opening).

- **Maximum Opening:** In this phase, the velum reaches its maximum lowering position, creating the maximum opening of the soft palate during speech production. This allows for maximum airflow through the nasal passage, which is relevant for certain speech sounds. The velum can rest in this position waiting for the tongue to reach the glide constriction.

- **Closing:** This phase involves the raising movement of the velum, resulting in the closure of the soft palate.

In the example given above, velum developed both a vertical and horizontal trajectory. There was a variation of position in both movements from a higher spatial point to a lower spatial point, representing the velum opening.

When the velum reaches its target position, it signifies the opening of the nasal passage, allowing air to flow through the nasal cavity. This opening is necessary for producing nasal sounds. However, to produce the nasal glide, the oral cavity must narrow. This creates a constriction that leads airflow through the nasal cavity. When the tongue reaches the glide target, the velum starts its closure movement, because this movement is directly related to the onset time of the following consonant.

### **5.5.2 Soft Palate Spatiotemporal Trajectory**

In this section, we will analyze the spatial path and duration of the movement of the soft palate during nasalization. Our purpose is to compare spatiotemporal variations in the initial velum position, the maximum opening position, and the final position.



Our purpose is to observe the development of the velum trajectory and compare the velum trajectories of the two speakers to determine if they were similar or distinct.

Thus, we divided the displacement of the velum into three moments:

- The initial position of the velum (TRAJ VEL I).
- The maximum opening position of the velum (TRAJ VEL M).
- The closing position of the velum – that is, the moment when it moves from the maximum opening position to the default position (TRAJ VEL F).

### **5.5.2.1 Back Segments**

Table 20 presents the soft palate spatial displacement data for the nasal back diphthong [ãw̃], as measured in two speakers: RD and GV (N = 25). The table provides average and standard deviation values for the soft palate's movement in both the x and y coordinates (in centimeters). The data include measurements for the onset, target, and offset positions of the soft palate during the production of the nasal back diphthong.

From this table is possible to understand the average spatial displacement of the soft palate during the production of the [ãw̃] diphthong for both speakers. The x-coordinate represents the horizontal movement, while the y-coordinate represents the vertical movement of the soft palate. The average values give an indication of the typical soft palate movement for each position (onset, target, and offset) in both speakers, RD and GV. The standard deviation values provide an estimate of the variability or dispersion of the data points around the average, indicating the degree of consistency or variation in the soft palate's movement.

In Table 20, below, the raw and average values refer to the velum movement. Thus, the average velum opening was 8.19 cm and 5.08 cm on the x-axis and 1.62 cm and 1.02 cm on the y-axis for RD and GV, respectively.

The average maximum velum opening was 7.92 cm and 4.89 cm on the x-axis and 1.31 cm and 0.76 cm on the y-axis for RD and GV, respectively. The average velum default position was 8.20 cm and 5.06 cm on the x-axis and 1.62 cm and 1.00 cm on the y-axis for RD and GV, respectively.

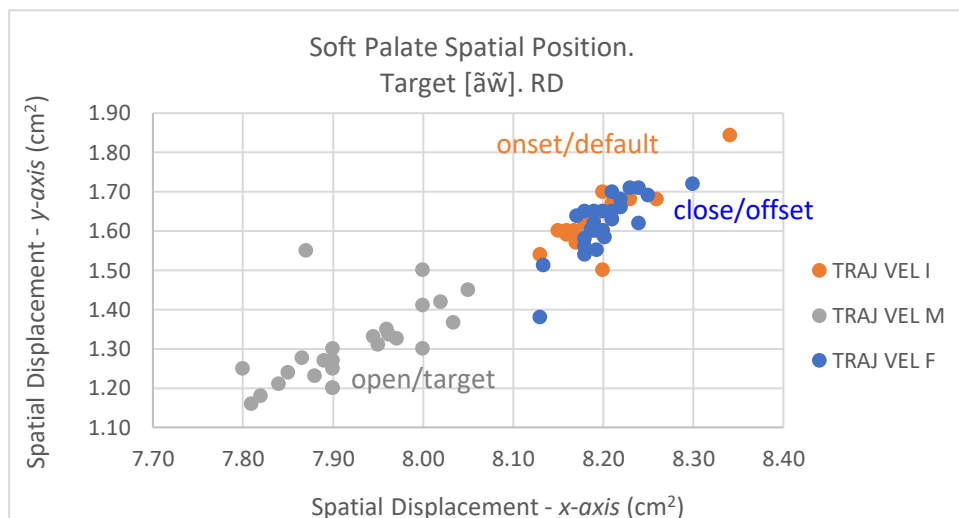
Table 20: Soft palate spatial displacement. Average and standard deviation. Production [ãw̃]. Coordinates x-y (cm). Speakers: RD and GV (N =25).

**Soft Palate Spatial Displacement (cm). Nasal Back Diphthong [ãw̃]. Speakers: RD and GV (N =25).**

	RD						GV					
	Onset		Target		Offset position		Onset		Target		Offset position	
	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)
<b>A</b>	8.19	1.62	7.92	1.31	8.20	1.62	5.08	1.02	4.89	0.76	5.06	1.00
<b>SD</b>	0.04	0.06	0.07	0.10	0.03	0.07	0.04	0.07	0.03	0.03	0.03	0.06

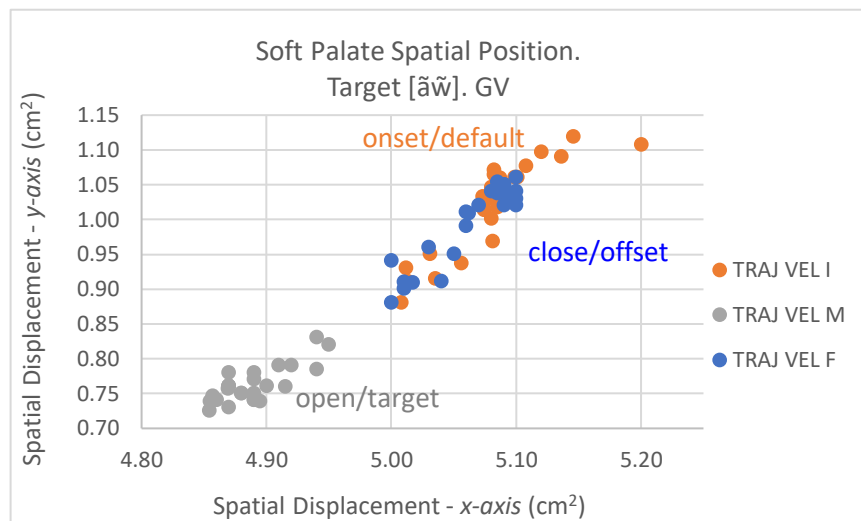
Charts 16 and 17 represent the soft palate in onset (starting position), target (opening movement endpoint), and offset (closing endpoint) (TRAJ VEL M).

Chart 16: Position of soft palate spatial displacement. Production [ãw̃]. Coordinates x-y (cm). Speaker: RD. (N =25).



The position of the pellet of the velum articulation is represented in Charts 1 and 2. From the data, we can see that the soft palate's movement can be divided into three gestures: a) onset (opening), b) target (maximal opening), and c) offset (closure). These three gestures appeared in RD and GV.

Chart 17: Position of soft palate spatial displacement. Production [ãĩ]. Coordinates x-y (cm). Speaker: GV. (N =25).



We calculated the variation between the initial position of the pellet and the default position, to measure the spatial variation of velum displacement and determine whether there was and different parameter for lowering the movement and nasal signature. However, the velum displacement had the same cycle trajectory for all data.

Table 21 presents the average values and standard deviations of the soft palate's delta spatial variation during the production of the back nasal diphthong [ãĩ]. The delta spatial variation represents the difference between the onset and the target of the soft palate.

For the back nasal diphthong [ãĩ], the average delta spatial variation for speaker RD is 0.38 cm on the x-axis and 0.47 cm on the y-axis. Speaker GV, on the other hand, has an average delta spatial variation of 0.19 cm on the x-axis and 0.26 cm on

the y-axis. These values indicate the magnitude of spatial displacement that occurs between the onset and target of the soft palate during the production of the back nasal diphthong [ãw̃].

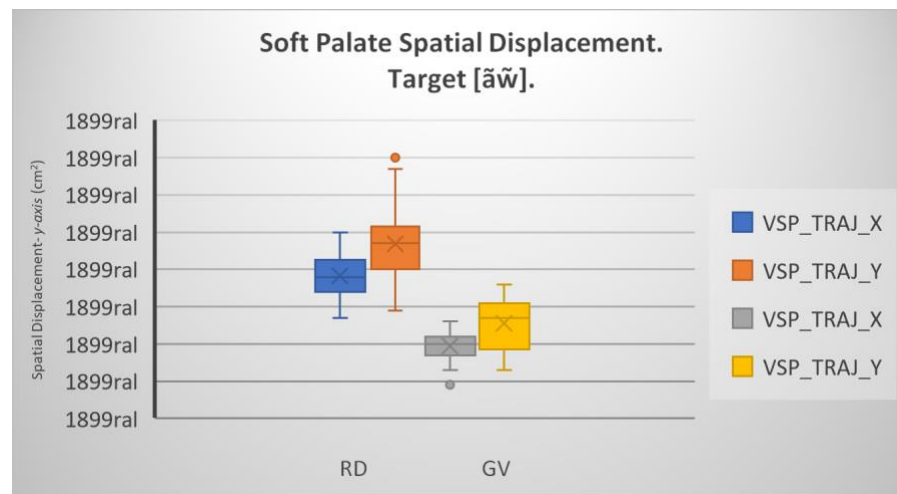
Table 21: Soft palate delta's spatial variation. Averages and standard deviations. Production [ãw̃]. Coordinates x-y (cm). Speakers: RD and GV (N =25).

<b>Soft Palate Delta's Spatial Variation (cm). Back Nasal Diphthong [ãw̃]. RD and GV (N =25).</b>				
	<b>RD</b>		<b>GV</b>	
	(x)	(y)	(x)	(y)
<b>A</b>	<b>0.38</b>	<b>0.47</b>	<b>0.19</b>	<b>0.26</b>
<b>SD</b>	<b>0.06</b>	<b>0.10</b>	<b>0.04</b>	<b>0.07</b>

Spatial variation is visualized on the chart below. In this chart, the overall average of the data and its standard deviation is shown. RD had a general spatial variation more significant than GV. This variation occurred because of the biological differences between the subjects, such as the size of the oral cavity and the shape of the hard palate.

The chart below displays the results representing vertical (x-axis) and horizontal (y-axis) velum motion (VSP\_TRAJ). In this case, both speakers displayed a more significant displacement along the vertical axis than along the horizontal axis. These values reflect the velum opening. This articulatory result complements the results of the fiberscope experiment.

Chart 18: Soft palate spatial displacement. Delta between onset and velum target. Production [ãw̃]. Coordinates x-y (cm). Speaker: RD and GV. (N =25).



We verified that the length between the opening position and the maximum velum opening (y-axis). This lowering movement had more magnitude amplitude than the horizontal (x-axis) movement of the velum.

We calculate the correlation between the x-y coordinates of the velum trajectory for each group of nasal segments, and the obtained Pearson coefficient value was 0.89, indicating a strong positive correlation. This result implies that there is a consistent and close relationship between the velum's horizontal and vertical movements for each group of nasal segments, suggesting a coordinated pattern of movement.

However, when comparing the velum trajectory between speakers RD and GV, the correlation coefficient was negative ( $r = -0.17$ ), indicating a weak negative correlation. This means that there is little consistency in the relationship between the velum's horizontal and vertical movements between these two speakers. The negative correlation suggests that their velum movements on the x-y axis diverge and do not follow the same spatial values. This may be related to physiological factors.

In the table below, we analyze the duration (ms) of the spatial movement of the soft palate during the production of back nasal diphthongs. From this data, we determined the average and standard deviation of the values extracted from the

temporal duration of the opening and closing movements of the velum in addition to the total duration of the movement.

The opening movement refers to the beginning of soft palate movement when it moves from the default position to the “maximum opening” of the velum. This is the point where the velum reaches its articulatory target and is ready to return to the default position. This reverse path will be called the “closing” trajectory.

Table 22: Soft palate spatial displacement. Absolute duration (opening and closing movements [ms]). Production [ãõ]. Speakers: RD and GV. (N =25).

Total Duration of Soft Palate Displacement Back Nasal Diphthong						
	RD			GV		
	Lowering	Closing	Total	Lowering	Closing	Total
A	262	158	420	206	192	398
SD	60	52	80	58	59	79

Table 22 presents the data on the absolute duration of soft palate spatial displacement during the production of the back nasal diphthong [ãõ] for speakers RD and GV. The durations are measured in milliseconds (ms).

For the opening movement of the soft palate, RD had an average duration of 262 ms, while GV had an average duration of 206 ms. For the closing movement, RD had an average duration of 158 ms, and GV had an average duration of 192 ms.

The total duration (VEL\_TD) of the complete soft palate displacement is the sum of the opening movement (VEL\_OP\_D) and the closing movement (VEL\_CL\_D). For speaker RD, the total duration was 420 ms (262 ms + 158 ms), and for speaker GV, it was 398 ms (206 ms + 192 ms).

The total duration was, on average, similar for both speakers, but we cannot say that said duration is a standard feature of velum articulation because only two speakers were analyzed.

Chart 19: Soft palate trajectory. Average duration to opening trajectory and the closure of the soft palate. Total duration (ms). Production [ãw̃]. RD and GV (N =25).

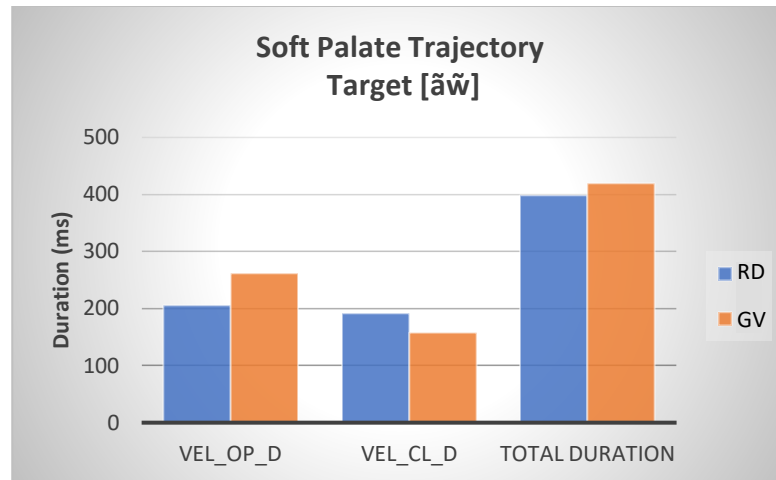


Chart 19 shows the average total duration of velum lowering and rising displacement. The total duration represents the velum's complete movement from beginning to end.

The opening movement took longer than the closing movement. This difference may be correlated to tongue movement. Additional visual representations will be analyzed in the next items of this description showing plots comparing the trajectories of the velum and tongue.

When the glide reaches the target position, the velum is also in the target position and starts the rising movement before the release of the subsequent consonant. This suggests that the velum begins its rising movement before the subsequent consonant is released, which might indicate a coordinated action between the velum and the tongue.

This process increases the duration of the lowering movement because the velum remains open, "expecting" the tongue to reach its target. This discussion about the coordination of the velum with the tongue gives us and global comprehension of the gestural coordination, but specific quantification of this coordination will be provided in the next items.

### 5.5.2.2 Front Segments

In this section, we will examine the velum trajectory and process to produce the front nasal vowel. We will also present the same type of analysis as in the above section, albeit with the articulatory data of the back nasal diphthongs.

We analyzed only the first three repetitions performed by GV. The last two repetitions suffered an electromagnetic interference in the soft palate pellet, making it impossible to study the production data of these nasal diphthongs.

The table below displays averages, and standard deviations to produce the front nasal vowel. The average values for velum opening displacement were 8.24 cm and 5.09 cm on the x-axis and 1.70 cm and 1.05 cm on the y-axis for RD and GV, respectively.

Table 23: Soft palate spatial displacement. Averages and standard deviations. Production [ẽ̃]. Coordinates x-y (cm). Speakers: RD (N =25) and GV (N =15).

Soft Palate Spatial Displacement (cm). Front Nasal Vowel [ẽ̃]. Speakers: RD (N =25) and GV (N =15).												
	RD						GV					
	Onset		Target		Offset position		Onset		Target		Offset position	
	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)
A	8.24	1.70	7.86	1.23	8.15	1.67	5.09	1.05	4.85	0.82	5.10	1.05
SD	0.06	0.08	0.06	0.05	0.06	0.04	0.01	0.03	0.04	0.03	0.01	0.02

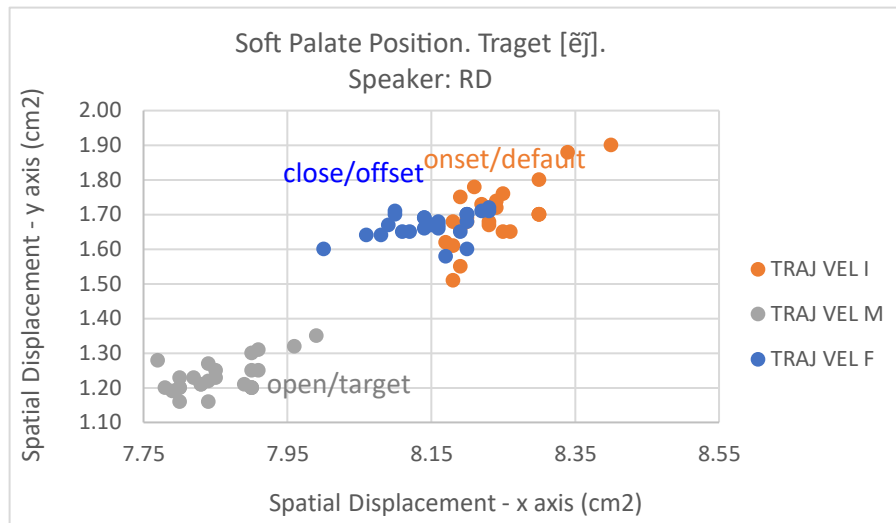
The velum has the same pattern as the soft palate in onset (starting position), target (opening movement endpoint), and offset (closing endpoint) to the back nasal diphthongs and the nasal front vowel. The values in the table and the charts represent the static position of velum.

From these data, we further verified that the soft palate displayed the same three movements: onset (opening movement), target (the endpoint of lowering), and offset



(the endpoint of closing), as explained in the previous section, and that these three gestures movement occurred similarly for RD and GV.

Chart 20: Position of soft palate spatial displacement. Production [[ $\tilde{e}$ ]]. Coordinates x-y (cm). Speaker: RD. (N =25).



The target position is the moment that the velum is in [+ open] and [+ front] (TRAJ VEL M). Thus, as with the back nasal diphthongs, this medial position is the target for the vocalic segments analyzed in this thesis. In our data, these gestures of onset, target, and offset appear to be velum's patterns because they are present in the production of both speakers. The nasalization also increased gradually during speech production of the nasal target segments.

Chart 21: Position of soft palate spatial displacement. Production [[ $\tilde{e}$ ]]. Coordinates x-y (cm). Speaker: GV. (N =15).

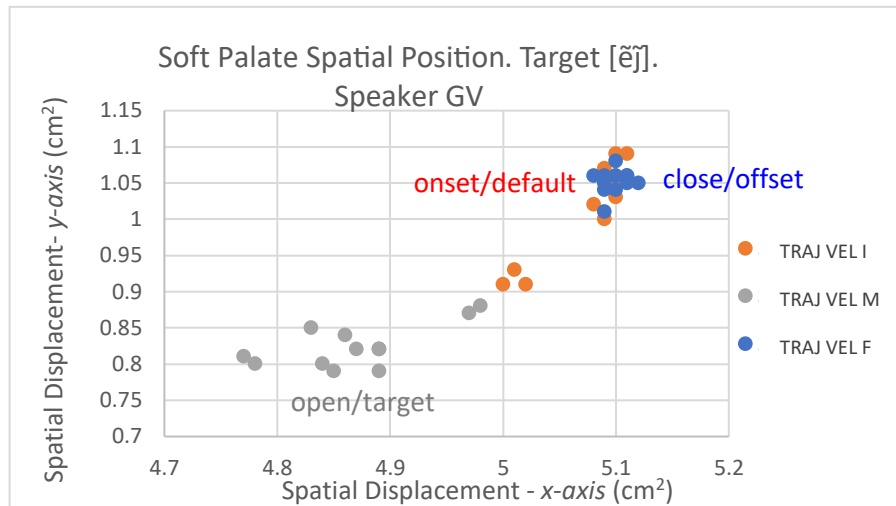


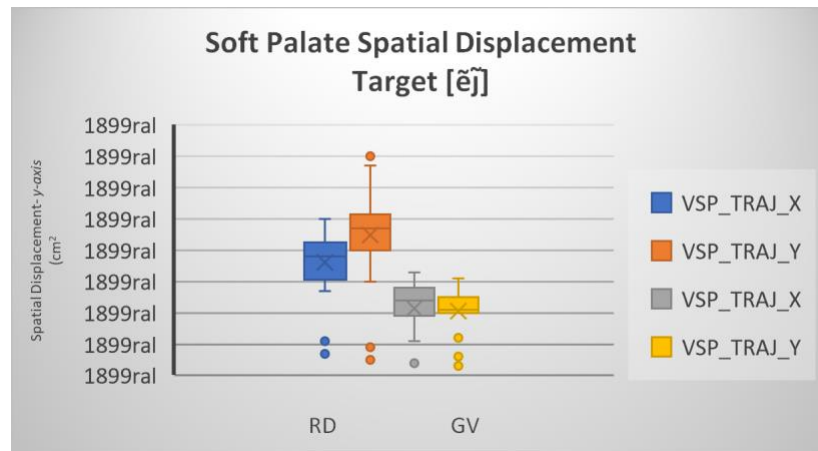
Table 24 shows the averages and standard deviations of soft palate displacement during front nasal oral vowel production. As with the back nasal diphthong analysis, this value represents the delta between the onset position minus the target position.

Table 24: Averages and standard deviations of the spatial delta of soft palate displacement. Front nasal vowel [ẽĩ]. RD (N =25) and GV (N =15) on the x-y axis (cm).

Space Delta – Velum’s Displacement. Front Nasal Vowel [ẽĩ]. Speakers: RD (N =25) and GV (N =15) on the X-Y Axis (cm).				
	RD		GV	
	(x)	(y)	(x)	(y)
A	0.36	0.45	0.21	0.21
SD	0.10	0.14	0.09	0.08

In this way, we verified that the spatial sweep between the opening position and the maximum opening of the velum on the x-y axis had a horizontal spatial average variation of 0.36 cm and 0.21 cm for RD and GV, respectively, and spatial variation on the vertical plane was 0.45 cm and 0.21 cm for RD and GV, respectively. Spatial variation is visualized in the chart below. In this chart, the visualization of the overall average of the data and its standard deviation is shown.

Chart 22: Soft palate spatial displacement. Lowering trajectory. Production [ẽĩ]. Coordinates x-y (cm). Speaker: GV. (N =15).



The chart displays the results representing the vertical (y-axis) and horizontal (x-axis) motion of the velum (VSP\_TRAJ). In this case, RD had a more significant displacement on the horizontal axis than on the vertical axis. GV had a more significant displacement on the vertical axis than the horizontal axis. These values reflect the velum's opening movement.

Velum movement is the result of the orchestration of pharyngeal muscles. If the velum does not have sufficient lowering movement, the velopharyngeal port must compensate with a larger opening. There must be a sufficient section for the passage of air into the nasal cavity.

The values of duration of velum displacement are displayed below. We extracted these values from the spatial movement of the soft palate during speech production. These values were extracted from the spatial movement of the soft palate during speech production.

The duration of lowering is calculated as the difference between the start and endpoint of the velum opening movement. The duration of closing is calculated as the difference between the target endpoint and the endpoint of the velum closing movement. The total duration represents the time from the start of the velum opening until the endpoint of the closing velum movement.

The values for the duration of velum displacement are displayed in table 23

below. This table presents averages and standard deviations of the values representing the temporal duration of the velum movement.

Table 25: Soft Palate Duration (opening and closing movements [ms]). Production [ẽ̃]. Speakers: RD (N =25) and GV (N =15).

Total Duration of Soft Palate Displacement (ms) Front Nasal Vowel Diphthongized [ẽ̃]. Speakers: RD (N =25) and GV (N =15).						
	RD			GV		
	Lowering	Closing	Total	Lowering	Closing	Total
A	207	198	405	213	209	420
SD	56	52	66	52	25	62

The opening movement of RD had an average duration of 207 ms and that of GV had an average duration of 213 ms. The average duration of the closing movement was 198 ms and 209 ms for RD and GV, respectively.

The total duration (VEL\_TD) of the complete displacement of the soft palate is found by adding the duration of the opening (VEL\_OP\_D) and closing (VEL\_CL\_D) movements, resulting in the following total values: 405 ms and 420 ms for RD and GV, respectively.

Chart 23: Soft palate trajectory. Average duration of opening trajectory and the closing of the soft palate. Total duration (ms). Production [ẽ̃]. Speakers: RD (N =25) and GV (N =15).

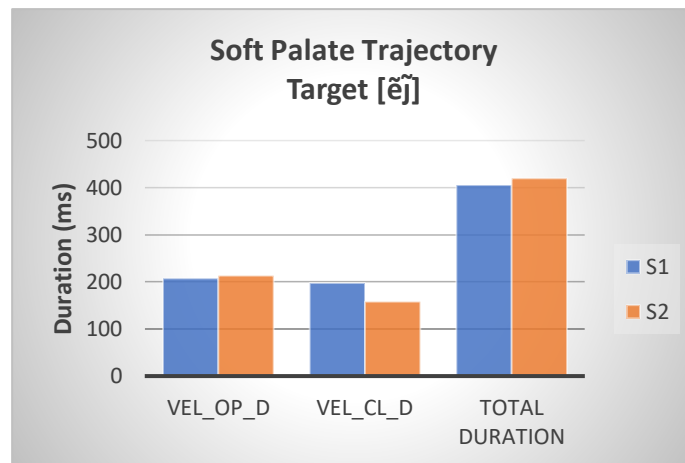


Chart 20 shows the values corresponding to velum lowering and rising displacement. The duration is like the back nasal diphthong. The opening duration was longer than that of the closing movement. Nevertheless, we cannot conclude that this difference is significant. The average of the total duration was similar for RD and GV.

### **5.5.3 Tongue Movement**

This section will analyze the trajectory and duration of tongue movement during the production of oral diphthongs, the front nasal vowel, and back nasal diphthongs. We aim to examine the dynamic trajectory of the tongue movement reaching the target for the vowel and the glide.

We will examine the tongue's spatial variation at two fixed points in its trajectory. This description aims to describe tongue motion and to determine whether nasal gesture activation modifies its trajectory. We will compare the tongue target position during the production of vowels and glides for both oral and nasal segments.

The values are extracted from the movement of the pellets TRAJ\_TNG\_DOS (dorsum/body) and TRAJ\_TNG\_PDOS (post-dorsal/back). It is important to remember that the tongue body, dorsal, and post-dorsal pellets move together.

#### **5.5.3.1 Back Diphthongs**

This section will analyze tongue trajectory during oral and back nasal diphthong production. The words are minimum pairs. For back segments, we will investigate the tongue back (post-dorsal) pellet (TRAJ\_TNG\_PDOS) because of the glide articulatory position.

The graphics below represent the movement of the pellets on the tongue back (TRAJ\_TNG\_PDOS) on the vertical and horizontal plane. Arrow indicates the trajectory between the vowel target and the glide target.

To identify the targets of the vowel and glide, we analyzed the inflection point. This point refers to a change in curvature in a trajectory. The curve changes from concave upward to concave downward, or vice versa. The inflection point marks the beginning of a new movement. Thus, when the curve of the gesture trajectory departs from an inflection point, it means that the gesture has already reached its target and changed its trajectory to reach another target.

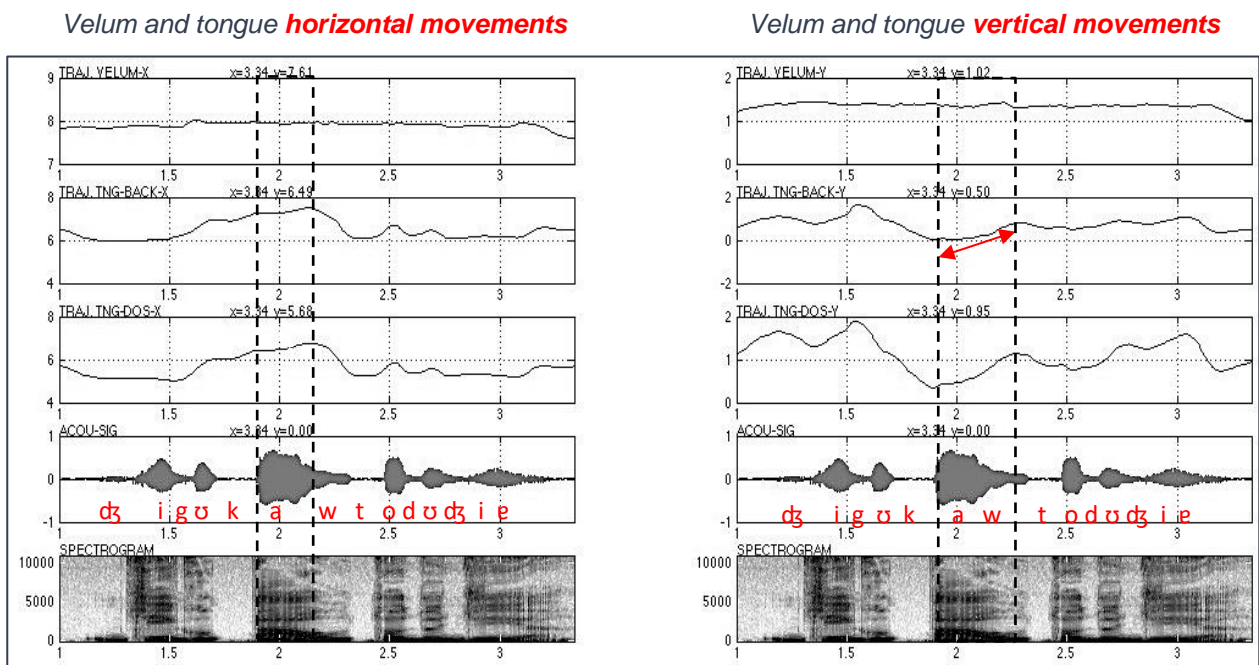


Figure 81: EMMA data. Pellets displacement on the x-y coordinate plane (cm). Targets and articulatory movements [aw]. Speaker: RD.

The arrows indicate the tongue movement during the production of the back oral diphthong. The upper window represents soft palate movement. The movement of the velum was relatively stable on both axes despite slight vibrational variations. The

vibration of the velum did not have any effect during speech increase. Thus, the velum remained closed during the production of back oral diphthongs as expected.

In the example above, the tongue left the vowel target and moved towards a more dorsal and higher region in the oral cavity. When the articulator reached the constriction of the glide, a falling trajectory occurred.

During the production of the back diphthong, the back part of the tongue underwent an elevation that created constriction in the oral cavity. The velarization of tongue movement in the horizontal plane was expected.

The phoneme /a/ moved from a more central region and reached a higher region to produce the glide /w/. Oral resonances decrease when we compare vowel and glide production. To produce the back oral diphthong, the tongue, and lips must be synchronized. The lips' rounding movement contributes to constricting the oral cavity and reducing resonances from the mouth.

Articulatory gestures are dynamic and can overlap, which makes it more challenging to define the onset and offset of the tongue gesture. This is because these events occur while the articulation of the subsequent phoneme is already taking place, leading to overlapping movements.

Figure 82 shows the displacement of the pellets during the production of the word [maw], an oral back diphthong. It is possible to visualize the tongue articulator's trajectory clearly because the gestures have less articulated each other. The velum displayed displacement because it produces the [m] consonant.

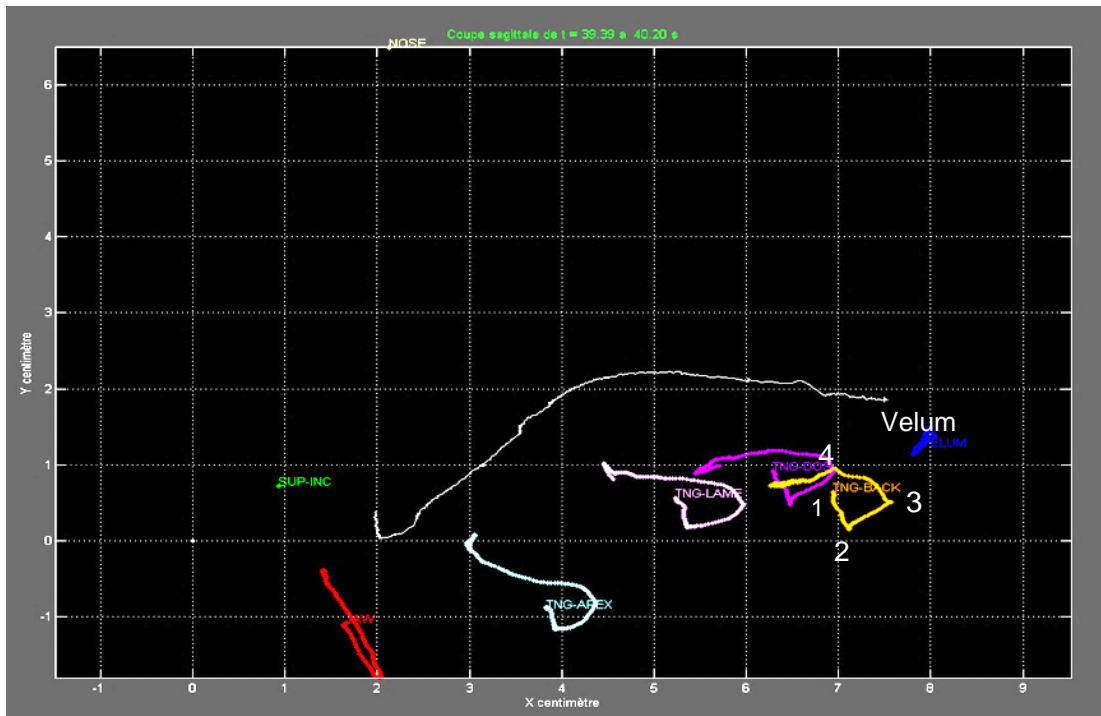


Figure 82: EMMA data. Pellets displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [maw]. Midsagittal view of oral cavity. Speaker: RD. The explanation for the numbers 1-4 will be given in the text.

In the back oral diphthong, in (1), the tongue is moving from the target [ʊ]. In (2), the tongue reaches the vowel target [a]. In (3), the tongue is back and rises to reach the glide target [w], and in (4), the tongue starts the anteriorization movement to produce the consonant /t/.

The magnitude of the tongue movement was small, less than 0.5 cm backward and 1 cm upwards. Thus, the tongue began the movement in the back position to produce the diphthong and finished in an advanced position towards the palatal region, due to the production of the dental stop /t/.



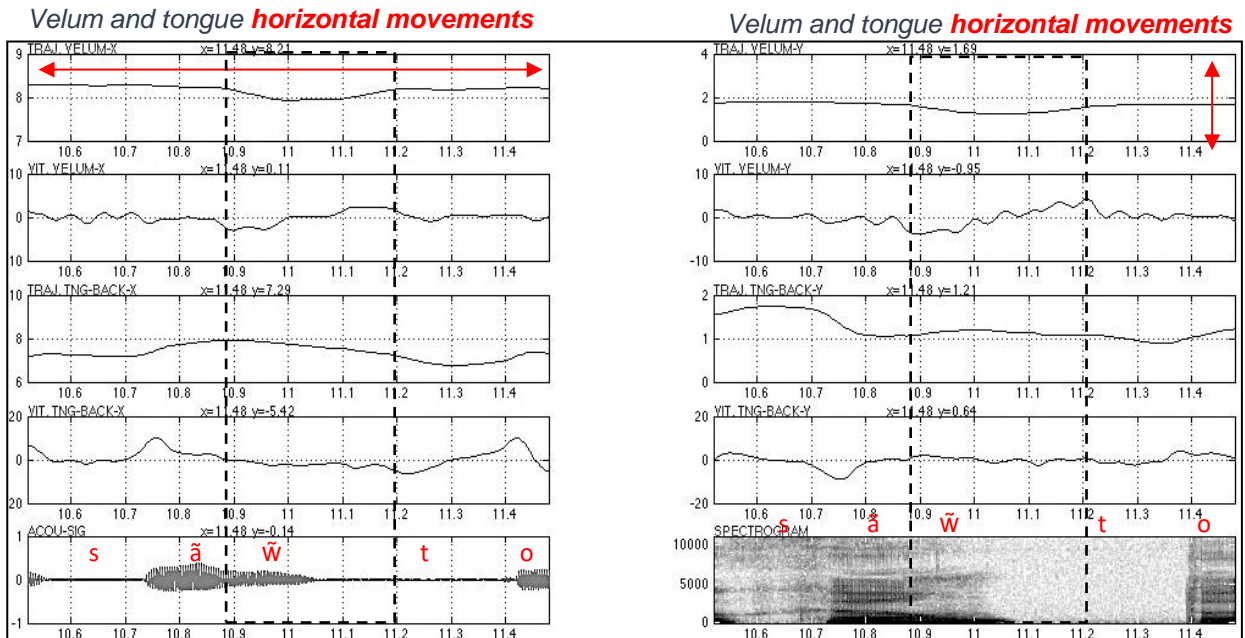


Figure 83: EMMA data. Pellet displacement using the x-y coordinate plane (cm). Target segment [ãw̃]. Speaker: RD.

The back nasal diphthong is a phoneme composed of a nasalized vowel, a glide, and the nasal appendix. In the nasal production of [ãw̃], the tongue played a different role than in producing the oral diphthong. The nasal vowel [ã] was produced in the narrower and back positions in these cases and moved to a velar position to reach the target position of the glide [w̃]. In some cases, the movement associated with velum lowering closed the oral cavity.

Figure 83 shows the pellets glued in the tongue. The figure represents tongue action during the word [pãw̃]. The velum was active because it is involved in the production of back nasal diphthongs.

As in oral production, the tongue was moving from the target [ʊ] to reach the vowel target [a]. After this, the tongue reached the glide target [w] and started the anteriorization movement to produce the consonant /t/; the velum was active during all tongue trajectories.

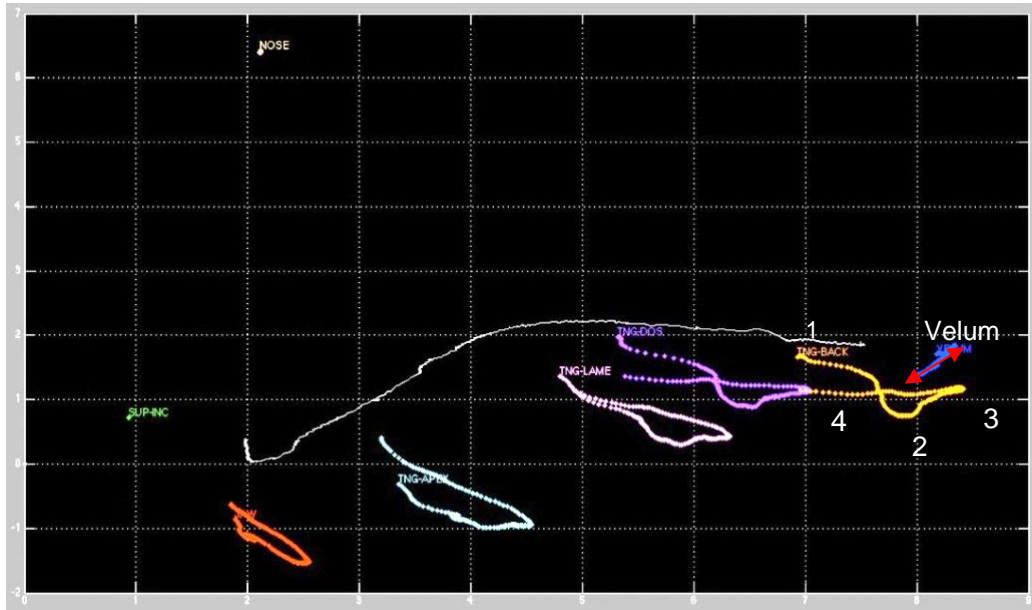


Figure 84: EMMA data. Pellet displacement on the the x-y coordinate plane (cm). Tongue targets and articulatory movements of [pãw̃]. Midsagittal view of oral cavity. Speaker: RD. The explanation for the numbers 1-4 will be given in the text.

In this drawing, a displacement from a central-back region to a back region is visible. The tongue moved from a central-back region in (1) to reach the target position of the nasal vowel [ã] in (2) and moved to a back position in (3) to reach the glide target of [w̃]. In (4), the tongue starts the anteriorization movement to produce the consonant /t/. The tongue trajectory displayed a spatial variation of about 1 cm and was lifted to the post-dorsal region.

In Figure 85 below, to produce the oral diphthong, GV controlled the gestures using the same strategy as RD. In the graph, one can see the velum oscillation in the vertical axis. Thus, GV maintained the same strategy as RD to produce the oral diphthongs, keeping the specific parameters of action, time, and symbolic information for the tongue back trajectory.

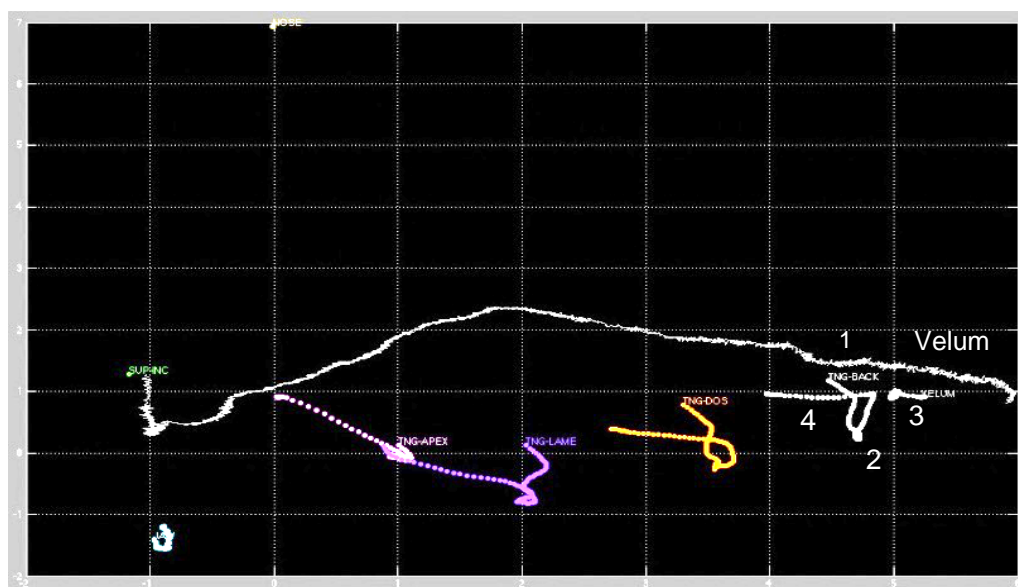


Figure 85: EMMA data. Pellet displacement on the the x-y coordinate plane (cm). Tongue targets and articulatory movements of [paw]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-4 will be given in the text.

In the back oral diphthong, in (1), the tongue is moving from the target [ʊ]. In (2), the tongue reaches the vowel target [a]. In (3), the tongue is back and rises to reach the glide target [w], and in (4), the tongue starts the anteriorization movement to produce the consonant /t/.

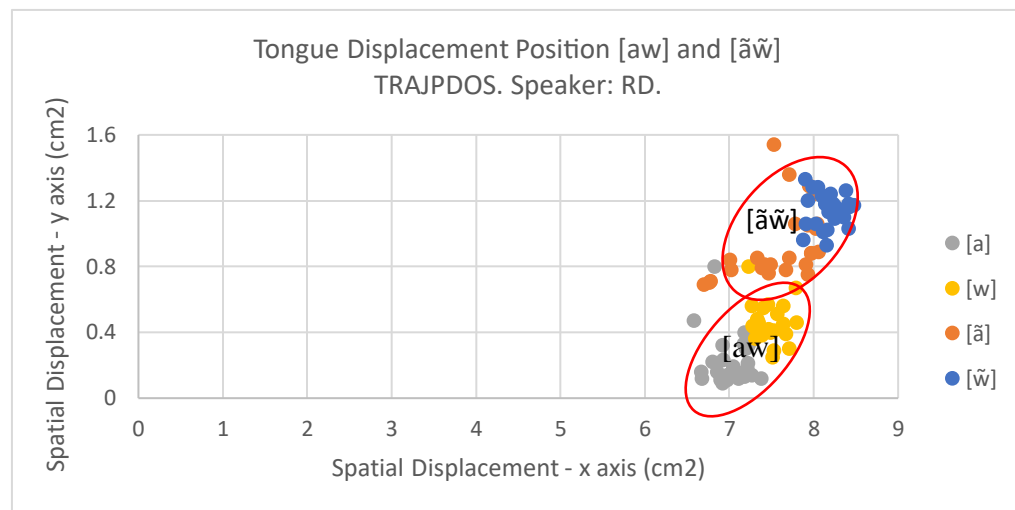
In table 26 below, we compare the averages, and standard deviations of the trajectories of the diphthongs [aw] and [ãw̃]. For RD, the average values for the target position of [a] are 7.02 cm and [w] 7.47 cm on the x-axis and 0.23 cm and 0.46 cm on the y-axis. The average values for the target position of [ã] are 7.58 cm and [w̃] 8.17 cm on the x-axis and 0.93 cm and 1.14 cm on the y-axis.

Table 26: Averages and standard deviations of tongue-displacement position: post-dorsal. Oral and back nasal diphthong: [aw] and [ãw̃]. RD (N =25) (cm).

Tongue-Displacement Posiiton: Post-Dorsal. Diphthong [aw] and [ãw̃]. Speaker: RD (N =25) (cm).								
	[a]		[w]		[ã]		[w̃]	
	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)
A	7.02	0.23	7.47	0.46	7.58	0.93	8.17	1.14
SD	0.21	0.15	0.16	0.12	0.43	0.22	0.17	0.10

The data for GV shows the target of the oral back vowel as more anterior than the nasalized vowel. As the tongue was already in a higher position, fewer horizontal movements produced the back nasal diphthong. Depending on tongue dynamics, the glide target position was more constricted more posterior, or both.

Chart 24: Tongue position: post-dorsal. Back oral diphthong [aw] and nasal [ãw̃]. RD (N =25) (cm).



The chart 24 displays the global position, comparing the oral and nasal segments. Thus, nasal diphthongization made the tongue rise with a back movement. The synchronism between the velum and the tongue produced a constriction in a velar region. In some cases, this constriction may become an occlusion of the oral cavity.

In the figure below, which represents the back nasal diphthong produced by GV, it was observed that GV's production of the nasal back diphthong differed from RD's. In GV's case, the tongue did not reach each target position gradually, as was observed in RD's production.

Instead, the gestures corresponding to (2) and (3) overlapped, meaning that the movements of the tongue for these two targets occurred simultaneously or closely together in time. This overlapping of gestures indicates a different articulatory pattern

compared to RD's sequential and gradual movement between targets. In (4), the tongue starts the anteriorization movement to produce the consonant /t/.

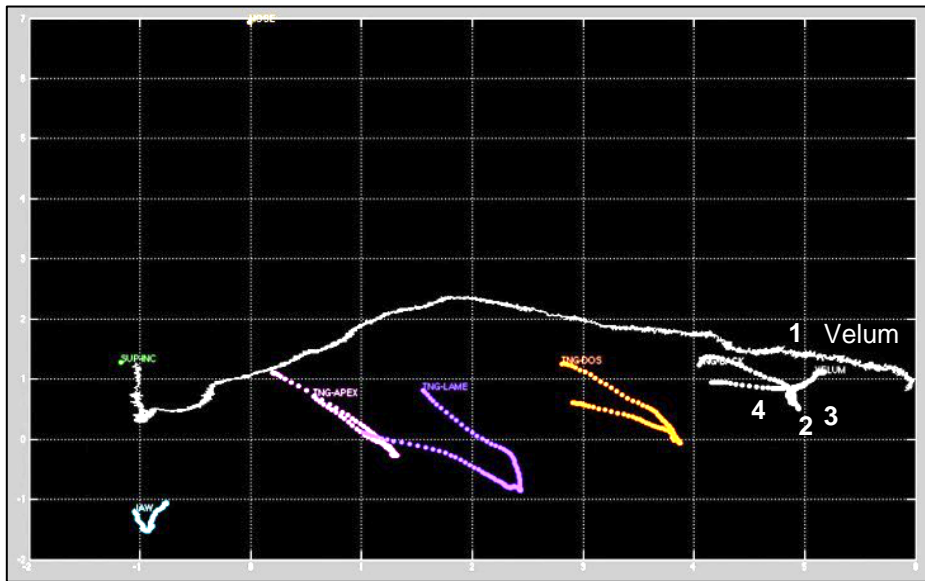


Figure 86: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [pãw̃]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-4 will be given in the text.

According to Pouplier & Goldstein (2007), adjustments in timing gestures and activation magnitude result in overlapping in articulatory gestures and smaller gesture magnitudes.

By combining findings from the studies of Albano (2001), Rodrigues (2007), Berti (2006), and Berti & Ferreira-Gonçalves (2012), it is suggested that the symbolic and dynamic nature of speech production, along with the intricate coordination of articulatory movements, contributes to the generation of these gradient processes. These processes allow for a range of variations in the production of nasal diphthongs, which enables speakers to produce distinct phonetic contrasts with precision and intermediate articulatory settings.

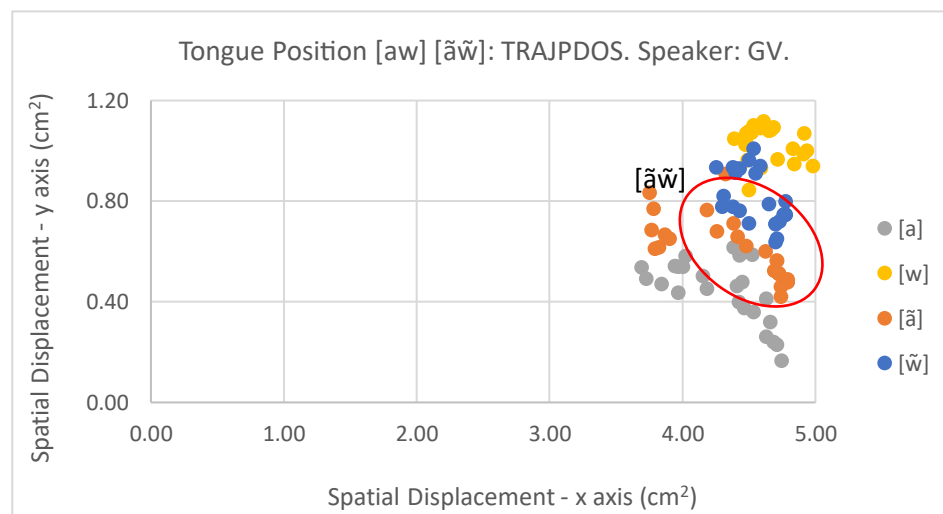
Table 27: Averages and standard deviations of tongue-displacement position: post-dorsal. Oral and back nasal diphthong: [aw] and [ãw̃]. Speaker: GV (N =5) (cm).

**Tongue-Displacement Position: Post-Dorsal. Later  
Diphthong [aw] and [ãw̃]. Speaker: GV (N =25) (cm)**

	[a]		[w]		[ã]		[w̃]	
	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)
A	4.30	0.45	4.64	1.03	4.35	0.61	4.56	0.81
SD	0.32	0.13	0.17	0.07	0.40	0.12	0.17	0.10

The chart below displays the distribution of average values for the target position of [a], that is 4.30 cm and [w] 4.64 cm on the x-axis and 0.45 cm and 1.03 cm on the y-axis. The average values for the target position of [ã] are 4.35 cm and [w̃] 4.56 cm on the x-axis and 0.61 cm and 0.81 cm on the y-axis.

Chart 25: Tongue Position: post-dorsal. Oral and nasal back diphthong: [aw] and [ãw̃]. Speaker: GV. (N =25) (cm).



The oral back vowel target was 0.05 cm lower than the back vowel target for the nasalized vowel. The production average indicates that the vowel underwent a slight change in its height. Therefore, we infer that the tongue is in a high position. The tongue raising seems necessary for nasalization because the nasal vowel is higher than the oral vowel.

The table below shows the values of the distance between the tongue back and the soft palate during the production of the back nasal-glide target. This measure

represents when the soft palate is in its lowest position. However, this does not necessarily mean that the tongue back is in the most back position of its trajectory.

These gesture trajectories are out of phase because the gestures of the tongue and velum have different timing activation, even if they are part of the same gestural constellation. We measured the spatial delta between the tongue dorsum position during the velum maximal opening position. Table 28 shows these values.

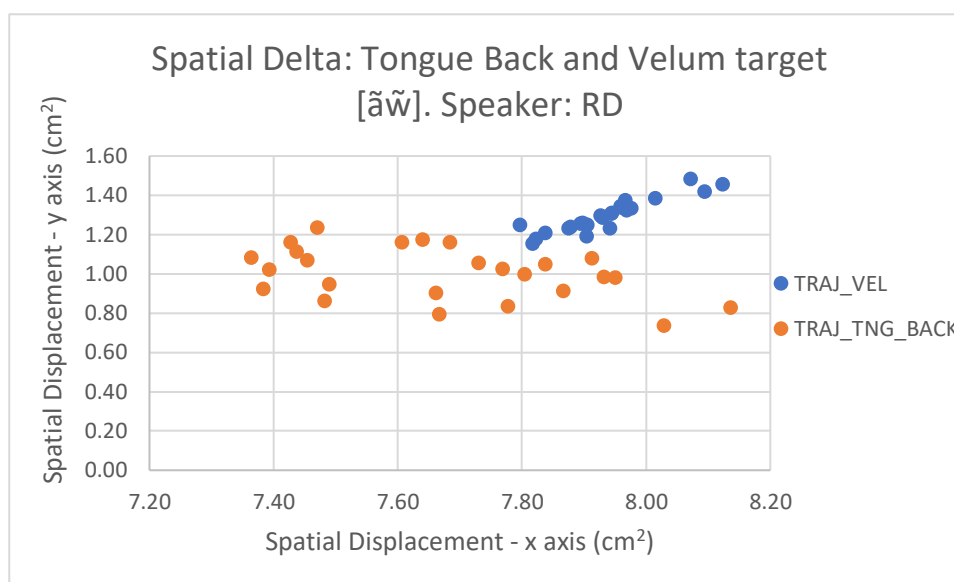
Table 28: Averages for delta from tongue-displacement trajectory and velum trajectory. Back nasal diphthong [ãĩ]. Speaker: RD. (N =25) (cm).

Spatial Delta – Tongue Back and Velum Target. Back Nasal Diphthong [ãĩ]. Speaker: RD.					
TRAJ_VEL		TRAJ_TNG_BACK		DELTA	
x	y	x	y	x	y
7.97	1.32	7.91	1.08	0.06	0.24
7.84	1.21	7.87	0.91	-0.03	0.3
7.8	1.25	7.8	1	0	0.25
7.82	1.16	7.47	1.24	0.35	-0.08
7.94	1.23	7.64	1.17	0.3	0.06
7.94	1.31	7.45	1.07	0.49	0.24
7.88	1.23	7.39	1.02	0.49	0.21
7.82	1.18	7.38	0.92	0.44	0.26
7.9	1.26	7.43	1.16	0.47	0.1
7.9	1.19	7.36	1.08	0.54	0.11
7.89	1.26	7.66	0.9	0.23	0.36
7.96	1.35	7.84	1.05	0.12	0.3
7.95	1.31	7.67	0.79	0.28	0.52
7.88	1.24	7.49	0.95	0.39	0.29
7.93	1.3	8.14	0.83	-0.21	0.47
7.96	1.34	8.03	0.73	-0.07	0.61
8.07	1.48	7.77	1.02	0.3	0.46
7.93	1.28	7.78	0.83	0.15	0.45
7.9	1.25	7.48	0.86	0.42	0.39
7.98	1.33	7.73	1.05	0.25	0.28
7.97	1.37	7.44	1.11	0.53	0.26
8.01	1.39	7.68	1.16	0.33	0.23
7.9	1.25	7.95	0.98	-0.05	0.27
8.09	1.42	7.93	0.98	0.16	0.44
8.12	1.46	7.61	1.16	0.51	0.3

In the table above, the values from the delta of the variables of the tongue back and velum demonstrate that the back nasal glide produced a constriction in the oral tract.

The values highlighted in dark grey suggest a closure in the vocal tract. When the delta results in a negative number, this means that the tongue and velum trajectory crossed in a posterior position. The constriction in the velar region occurred in 25% of the back nasal pronunciations performed by RD. The closure is the result of the velum position. The velum movement reached an advanced position compared to the tongue back. The chart below shows this spatial distribution:

Chart 26: Tongue Position trajectory: post-dorsal. Oral and back nasal diphthong: [ãw̃]. Static value of the pellet. RD. (N =25) (cm).



During the production of the back nasal diphthong, RD had five instances of vocal-tract closure, as shown in Table 28. The distance between the tongue back and the velum target was, on average,  $x = 0.25$  cm and  $y = 0.29$  cm. The table and the chart below show the values and the distribution of this data for GV.

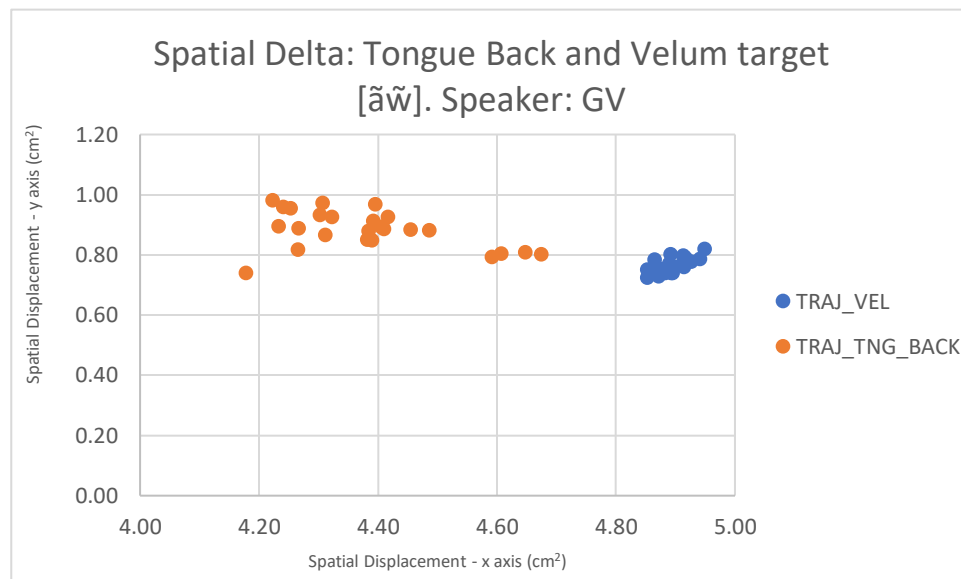
Table 29: Values for delta from tongue-displacement trajectory and velum trajectory. Back nasal diphthong [ãw̃]. Speaker: GV. (N =20) (cm),

Spatial Delta – Tongue Back and Velum target. Back Nasal Diphthong [ãw̃]. Speaker: GV.					
TRAJ_VEL		TRAJ_TNG_BACK		DELTA	
x	y	x	y	x	y
4.92	0.79	4.40	0.97	0.52	-0.18
4.94	0.78	4.46	0.88	0.48	-0.1
4.88	0.74	4.32	0.93	0.56	-0.19



4.88	0.75	4.25	0.95	0.63	-0.2
4.86	0.75	4.31	0.97	0.55	-0.22
4.91	0.80	4.61	0.80	0.3	0
4.95	0.82	4.49	0.88	0.46	-0.06
4.88	0.74	4.18	0.74	0.7	0
4.89	0.75	4.23	0.89	0.66	-0.14
4.87	0.76	4.24	0.96	0.63	-0.2
4.89	0.80	4.67	0.80	0.22	0
4.85	0.72	4.38	0.85	0.47	-0.13
4.87	0.73	4.39	0.85	0.48	-0.12
4.89	0.77	4.30	0.93	0.59	-0.16
4.90	0.74	4.22	0.98	0.68	-0.24
4.87	0.78	4.65	0.81	0.22	-0.03
4.91	0.77	4.41	0.89	0.5	-0.12
4.90	0.76	4.31	0.87	0.59	-0.11
4.89	0.74	4.38	0.88	0.51	-0.14
4.91	0.76	4.27	0.89	0.64	-0.13
4.85	0.75	4.41	0.89	0.44	-0.14
4.87	0.75	4.59	0.79	0.28	-0.04
4.86	0.74	4.27	0.82	0.59	-0.08
4.93	0.78	4.42	0.93	0.51	-0.15
4.86	0.74	4.39	0.91	0.47	-0.17

Chart 27: Tongue-displacement trajectory: post-dorsal. Oral and back nasal diphthong: [ãũ]. Static value of the pellet. Speaker: GV. (N =25) (cm).



During the production of the back nasal diphthong, GV had no instances of vocal-tract closure. The distance between the tongue back was, on average,  $x = 0.50$  cm forwards to the velum, and the velum target was  $y = 0.12$  cm downwards from the tongue.

### 5.5.3.2 Front Segments

This section will examine tongue articulatory behavior during the production of the front nasal vowel [ẽ] and the front oral diphthong. The parameters will be extracted from the pellets in the TRAJ\_TNG\_DOS (tongue blade) and TRAJ\_TNG\_PDOS (tongue back) positions.

The front oral diphthong is characterized by a middle vowel /e/ in the syllable core and by a glide in the nucleus position. Thus, the tongue moves from a lower position to a higher position. The tongue elevation is a parameter of nasal diphthongization.

Phonologically, the nasal vowel [ẽ] should not activate the tongue elevation gesture in the symbolic category. However, the constriction of the gesture is activated due to nasal unpacking, resulting in nasal diphthongization and the emergence of the palatal glide.

Recasens (1994) showed that palatal-[l] productions can be velarized in Catalan, depending on the lateral position in syllables and the adjacent vowel quality. The F1 frequencies are related to the jaw opening movement and adjacent vocalic environment.

According to Recasens et al. (1993), the palatal consonants [ʎ, ɲ, j] create a constriction in the palate region. Nevertheless, only imaging data can reveal if the place of articulation is on the entire palate or restricted to a specific area, the alveolar palatal [ʎ, ɲ], and the palatal [j]. Thus, Recasens et al. (1993) the palatal consonants are divided into two classes through EMMA's data because the lateral consonants activate different articulatory regions when produced.

The graphics below represent the tongue back pellets (TRAJ\_TNG\_DOS). Arrows indicate velum and tongue moment and the fixed point of the glide target.

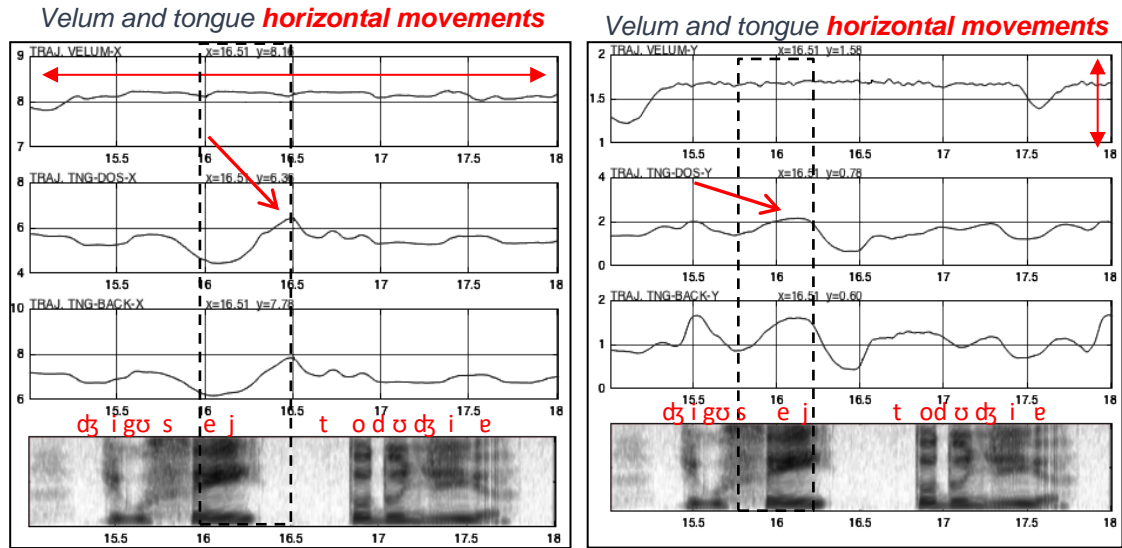


Figure 87: EMMA data. Pellets displacement using the x-y coordinate plane (cm). Target segment [ej]. Speaker: RD.

In Figure 87, we see the two images representing front oral-diphthong production. The pellet movement from the soft palate displayed a slight oscillatory movement although the soft palate was closed. The soft palate lowering at the beginning and at the end of speech production is related to the breathing.

The oral diphthong is composed of a vowel and a syllable glide core. In the front diphthong, the tongue starts from a mid-close position [e] and moves forward closer position [j]. The displacement of the pellet glued to the tongue body displayed a lifting movement on the y-axis. The soft palate oscillated during the production of the oral segment.

The tongue gradually transitioned from one position to another, creating glide constriction. The tongue's rising movement was a default parameter for the correct articulation of the front diphthong and its perception. The figure below displays the pellet's trajectory during the pronunciation of the word "sei".



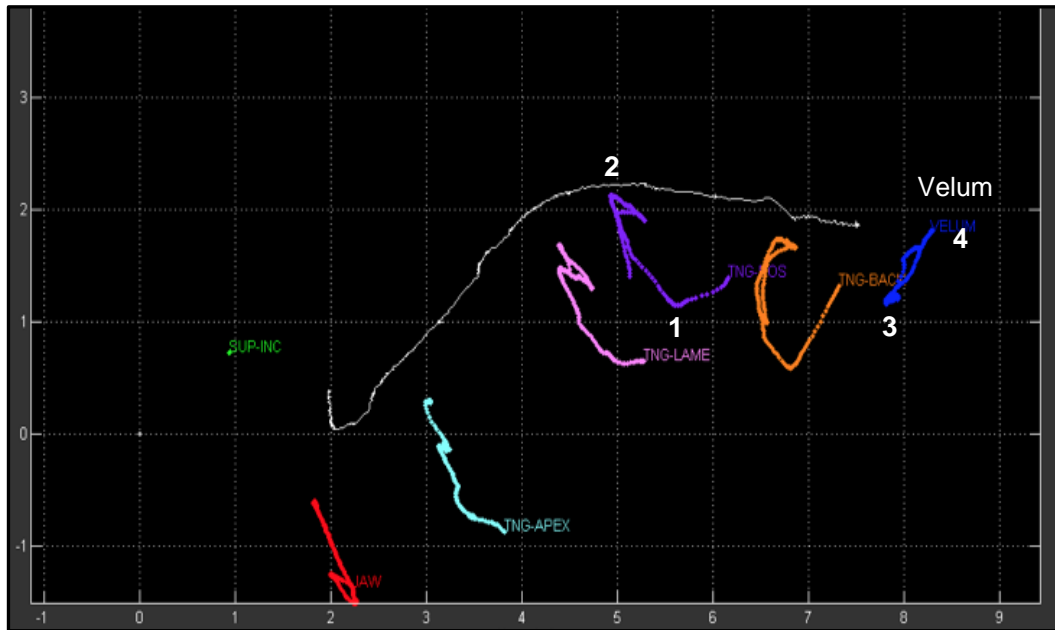


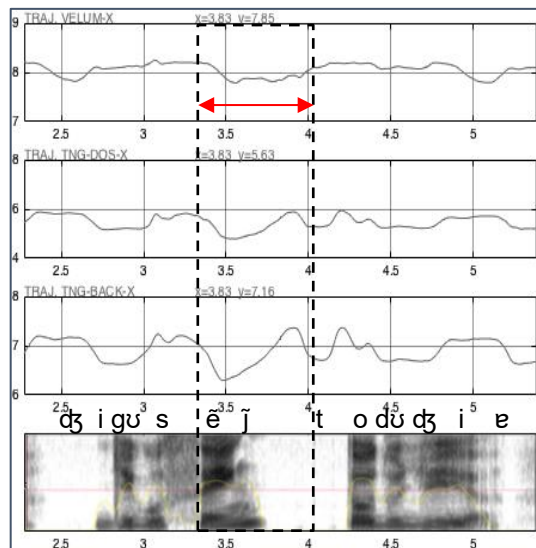
Figure 89: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [sɛj]. Midsagittal view of oral cavity. Speaker: RD. The explanation for the numbers 1-4 will be given in the text.

In the context of a boundary with the stop /t/, the front segments are produced with less tongue displacement because the articulation of the phoneme occurs in the front region of the oral cavity. To produce back diphthongs, the tongue goes towards a back region and then forward to a front region to articulate the stop /t/.

The graphics below represent the pellets on the tongue back (TRAJ\_TNG\_DOS) in the vertical and horizontal plane. Arrows indicate velum displacement in vertical and horizontal movements.

The graphics below represent the tongue back pellets (TRAJ\_TNG\_BACK). Arrows indicate velum and tongue moment and the fixed point of the glide target.

Velum and tongue **horizontal movements**



Velum and tongue **horizontal movements**

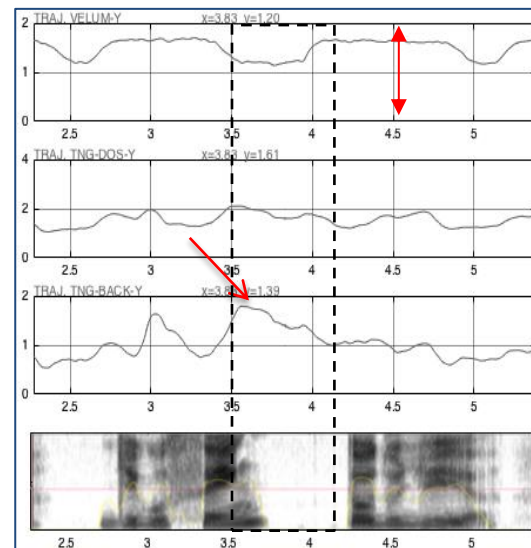


Figure 90: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Target segment [sē]. Speaker: RD.

In Figure 89, the images represent the soft palate movement in synchrony with tongue-body movement. The elevation of the tongue body's rising movement towards the palatal region indicates the presence of glide from the phenomenon of nasal diphthongization. In nasal vowels, production should not involve constriction in the oral cavity. For resonance to occur in the nasal cavity, both the oral cavity constriction and velum lowering must be involved in the production of nasal sounds.

It is essential to note that in nasal diphthongization, the tongue may experience some constriction as it transitions from the nasalized vowel to the glide. However, this constriction alone is not sufficient to produce resonance in the nasal cavity. To achieve resonance in the nasal cavity, the velum must also be lowered, allowing air to flow through the nasal passages.

The vowel's onset is oral, without nasalization, because soft palate displacement started before the vowel. The onset and duration of the articulatory gestures involved in producing the oral portion of the diphthong (vowel and glide) were different from

those involved in producing the nasal portion. The timing of the soft palate displacement, indicating the switch to nasal articulation, started after the oral vowel's onset, showing that the nasal gesture was activated later than the oral gesture.

The duration of the movements of the tongue and the velum were different from each other (temporal duration), and the timing of when each movement started or became active was also different (activation time). This could result in the velum mechanism remaining active even after the palatal glide reached its target, contributing to a high volume of nasal airflow.

The table below represents the articulatory target values for each phone during the articulation of the target word. For RD, the average values for the target position of [e] are 5.37 cm and [j] 4.85 cm on the x-axis and 1.42 cm and 2.09 cm on the y-axis. The average values for the target position of [ẽ] are 5.42 cm and [j̃] 5.08 cm on the x-axis and 1.32 cm and 2.07 cm on the y-axis.

Table 30: Averages and standard deviations of tongue-displacement position: tongue body. Oral and front nasal segments: [ej] and [ẽj̃]. Speaker: RD. (N =25) (cm).

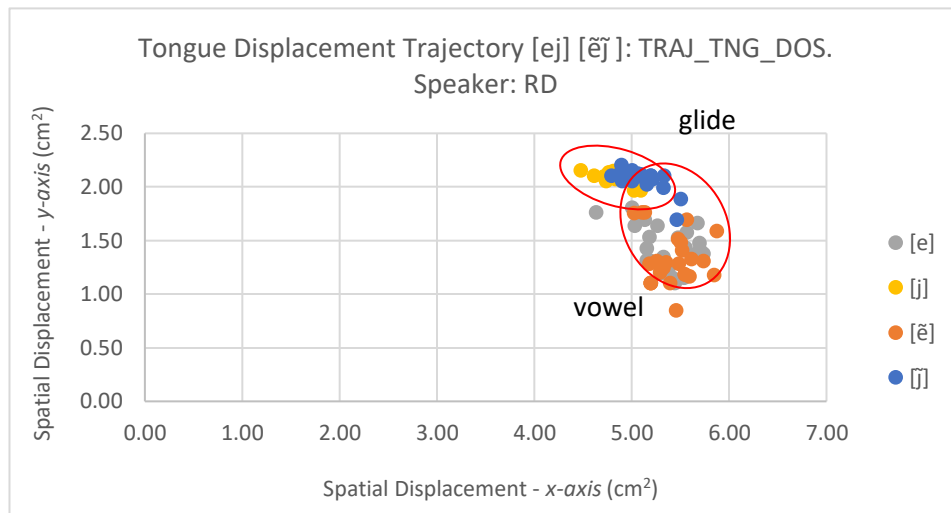
Tongue Displacement Position: Tongue Body. Front Segments [ej] and [ẽj̃]. RD (N =25) (cm)								
	[e]		[j]		[ẽ]		[j̃]	
	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)
<b>A</b>	<b>5.37</b>	<b>1.42</b>	4.85	2.09	<b>5.42</b>	<b>1.32</b>	5.08	2.07
<b>SD</b>	<b>0.26</b>	<b>0.20</b>	0.14	0.04	<b>0.22</b>	<b>0.24</b>	0.18	0.10

From the values of RD, we can observe that the target of the front oral vowel was produced around the same region of the front nasal vowel in the oral cavity. The average of the target achievement indicates that the vowels /e/ and /ẽ/ had the same spatiotemporal characteristics.

Consequently, the nasal parameter is a fundamental difference between these segments. A vertical elevation of the tongue body went towards the palatal region to

produce the oral and nasal palatal glide. In Chart 28, it is possible to observe the presence of the vowel and the glide for oral and nasal production. We would like to highlight the similarities between the front nasal vowel and the oral diphthong, and the presence of the nasal glide in the production of the front nasal vowel.

Chart 28: Tongue-displacement position: tongue body. Oral and front nasal diphthong: [ej] and [ẽj]. Static value of the pellet. RD (N =25) (cm).



The front nasal vowel and the oral diphthong share the same information regarding constriction location and degree. Thus, the associated gesture variables were structured and coordinated to implement the articulators' actions.

When GV produced the front oral diphthong, the tongue body had the same kind of movement as in RD. The tongue gradually went from a lower position to a higher position to create glide constriction. This movement can be observed in (1) and (2). The velum had a vibratory movement, but there was no lowering movement. The figure below shows the pellet's trajectory during the pronunciation of "seɪ" by GV. The tongue-rising movement seems to be a parameter for nasal diphthongization. The tongue had the same target positions and location in oral and nasal segments.



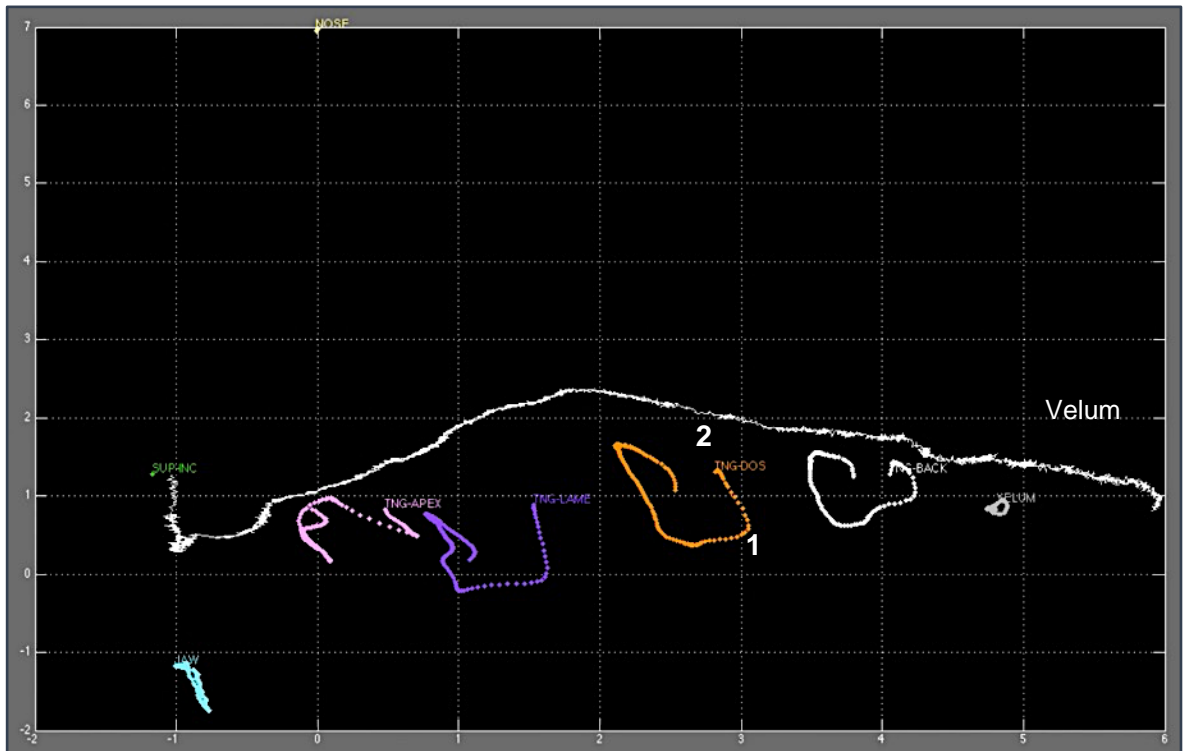


Figure 91: EMMA data. Pellet displacement on the x-y coordinate plane (cm). Tongue targets and articulatory movements of [sej]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-2 will be given in the text.

Figure 92 shows the pellet trajectories for the tongue and velum displacement for GV.

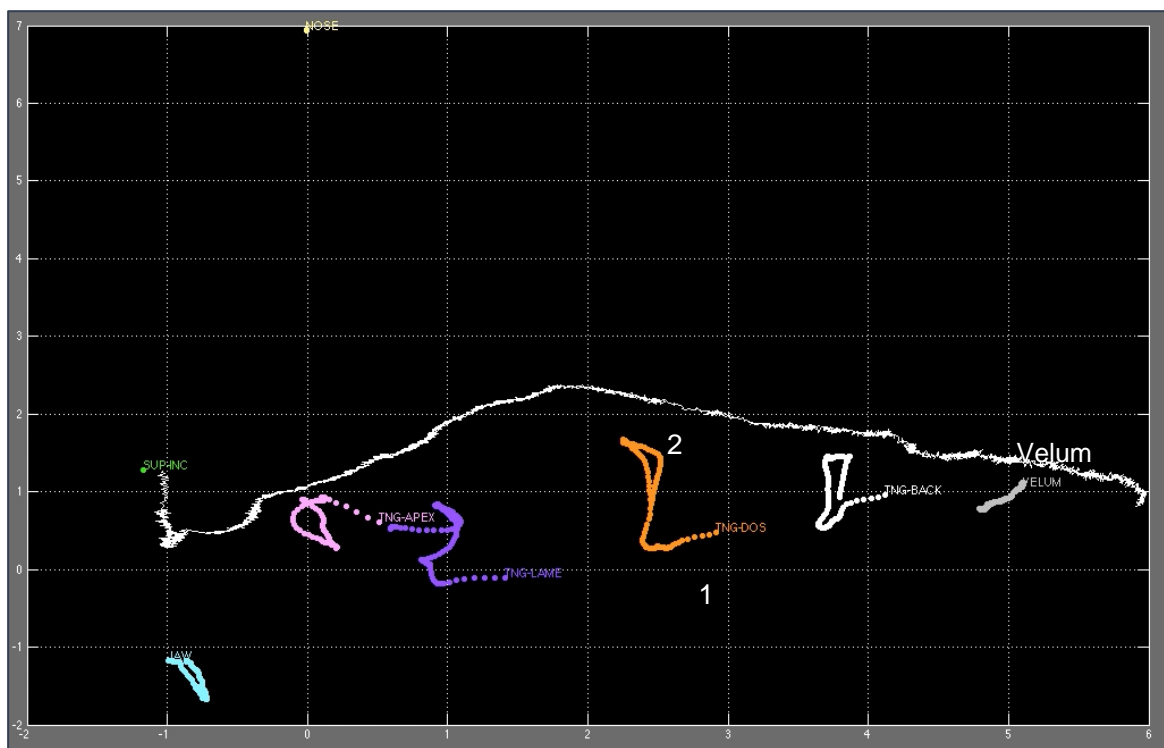


Figure 92: Emma data. Pellets displacement in x-y coordinate plane (cm). Tongue targets and articulatory movements of [sẽ]. Midsagittal view of oral cavity. Speaker: GV. The explanation for the numbers 1-2 will be given in the text.

The numbers (1) and (2) show the tongue body displacement during speech production. In table 31 below, we compared the average, and standard deviation values of the trajectory of the tongue body during the production of the diphthongs [ej] and [ẽj]. The values from Table 31 represent the articulatory target of each phone articulated by GV.

The average values for the target position of [e] are 2.70 cm and [j] is 2.09 cm on the x-axis and 0.41 cm and 1.62 cm on the y-axis. The average value for the target position of [ẽ] is 2.88 cm and [j] is 2.17 cm on the x-axis and is 0.52 cm and 1.63 cm on the y-axis.

Table 31: Average and Standard Deviation of Tongue target position: Tongue Body. Oral and Nasal Front Segments: [ej] and [ẽj]. GV (N =25) (cm).

Tongue-Displacement Position: Tongue Body. Front Diphthong [ej] and [ẽj]. Speaker: GV. (N =25) (cm)								
	[e]		[j]		[ẽ]		[j]	
	(x)	(y)	(x)	(y)	(x)	(y)	(x)	(y)
A	2.70	0.41	2.09	1.62	2.88	0.52	2.17	1.63
SD	0.22	0.20	0.12	0.08	0.51	0.24	0.21	0.08

The oral glide [j] was in a front palatal region if compared to the vowel [e]. The average values show a spatial variation around 0.60 cm between the vowel and the oral glide in the horizontal. Vertically, the difference was around 1.20 cm.

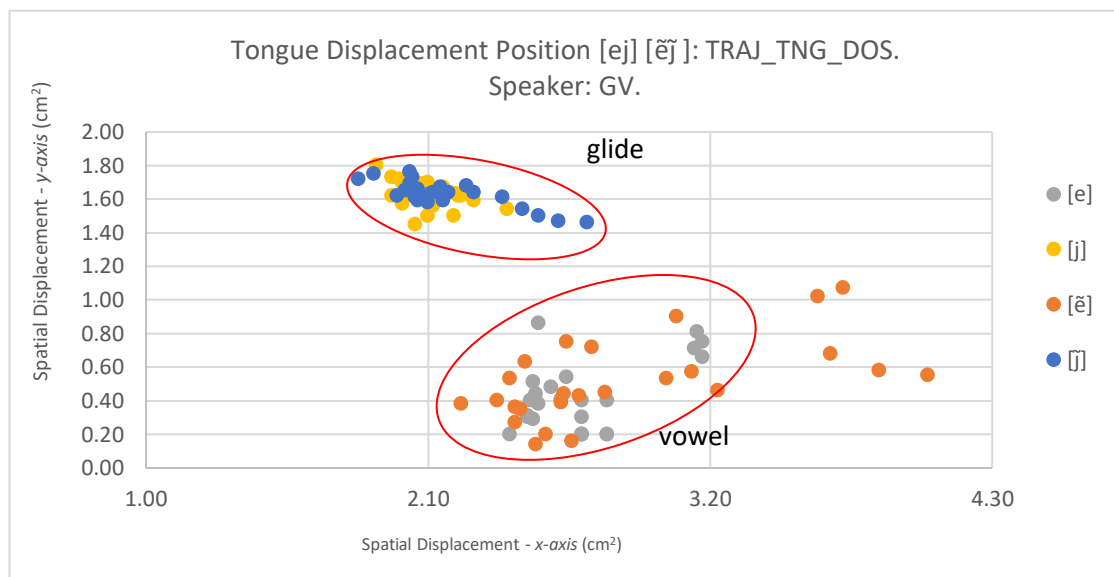
The nasal glide [j] was also in a front palatal region if compared to the nasal vowel [ẽ]. The mean values show a spatial variation in the horizontal, around 0.71 cm, between the vowel and the nasal glide. In the nasal segments, this difference was around 0.70 cm. Vertically, the difference was around 1.10 cm.

The nasal vowel target was around 0.10 cm higher than the oral vowel,

comparing the horizontal position of the oral vowel [e] and the nasal vowel [ẽ]. We can say if this difference is significant or not. We didn't apply any statistical tests, because of the small sample.

The oral glide [j], comparing the production with the nasal glide [j̃], exhibited no relative differences in the horizontal position. The constriction degree remained the same. Thus, there was no oral cavity occlusion because GV's hard palate is at 2.3 cm on the y-axis. The difference between the tongue glide constriction and the hard palate position had an average of 0.67 cm. The chart below compares the spatial variation for front production.

Chart 29: Tongue-displacement position: tongue body. Front oral diphthong [ej] and nasal [ẽj]. Static value of the pellet. Speaker: GV. (N =25) (cm).



Few differences were comparing the oral and nasal production in front segments. The presence of the nasal gesture resulted from nasal diphthongization that raised the tongue to a higher position and created a constriction, resulting in the nasal glide during the production of the nasal vowel /ẽ/. Thus, the similar spatiotemporal characteristics lead us to infer that tongue distinctive features were equivalent.

When a back nasal diphthong is produced, it involves the coordinated movement of the tongue and the soft palate. In an anti-phase relationship, the movements of the tongue and the soft palate are coordinated in such a way that the soft palate starts its lowering movement (allowing air to pass through the nasal cavity) after the tongue has completed the production of the nasal-vowel segment.

The anti-phase relationships between gestures involve more than just a sequential execution of movements; they involve precise coordination and timing to achieve the desired speech sound (NAM, GOLDSTEIN & SALTZMAN, 2010).

The temporal timing activation of the velum onset and vowel onset is different. The difference between the two gestures was 63 ms for RD and 63 ms for GV. Thus, the nasal vowel has its beginning with less nasal resonance influences.

The closing of the soft palate, the offset of voicing, and the onset time of the following consonant displayed a relationship in phase. In all analyzed cases, the soft palate ended the closing movement before the onset of the stop /t/. The temporal relation between velum closure and the onset time of the stop /t/ was 82 ms for RD and 58 ms for GV. The table below shows the delta between the tongue and velum.

Table 32: Averages of temporal acoustic duration the delta variation between nasal-vowel onset and velum lowering. Delta variation between velum closure and stop consonant onset. Speakers: RD (N =25) and GV. (N =15) (ms).

**Temporal Delta Variation – Nasal-Vowel Onset and Velum Onset. Velum Offset – Stop VOT. (ms)**  
Speakers: RD. (N=25) and GV (N =15).

RD				GV			
[ãw̃]		[ẽj̃]		[ãw̃]		[ẽj̃]	
VowOnst/ VelOnst	VelOff/ SVOT	VowOnst/ VelOnst	VelOff/ SVOT	VowOnst/ VelOnst	VelOff/ SVOT	VowOnst/ VelOnst	VelOff/ SVOT
58	0	0	27	48	0	52	65
-178	58	10	80	-164	52	31	88
37	75	-180	0	63	50	-132	68
56	135	0	39	40	73	30	107
66	145	-57	46	46	23	10	80
55	115	5	0	48	25	48	79

-170	140	5	48	-162	53	14	109	
60	117	-140	78	79	55	-126	92	
72	122	0	30	86	73	31	117	
60	23	-61	107	23	79	36	102	
49	67	5	162	77	19	25	99	
-186	98	0	37	-175	51	30	108	
71	112	-141	70	71	82	-142	90	
87	62	36	90	70	39	27	105	
55	53	0	0	65	90	34	80	
58	24	0	54	59	21	∅	∅	
-149	150	9	80	-159	108	∅	∅	
66	56	-160	70	74	79	∅	∅	
82	137	0	83	67	32	∅	∅	
69	79	32	102	66	68	∅	∅	
64	37	8	92	84	47	∅	∅	
-132	0	0	17	-122	58	∅	∅	
86	53	-172	74	67	78	∅	∅	
62	131	0	50	69	86	∅	∅	
37	70	-52	61	61	109	∅	∅	
<b>A</b>	<b>63</b>	<b>82</b>	<b>10</b>	<b>60</b>	<b>62</b>	<b>58</b>	<b>27</b>	<b>93</b>

For RD, the average duration is 63 ms, indicating that the velum lowering starts around 63 ms after the onset of the nasal vowel [ãĩ]. For GV, the average duration is 10 ms, suggesting that the velum lowering begins around 10 ms after the onset of the nasal vowel [ãĩ]. For RD, the average duration is 82 ms, which means the velum closure ends about 82 ms before the onset of the stop consonant following the nasal vowel [ãĩ]. For Speaker GV, the average duration is 60 ms, indicating that the velum closure ends approximately 60 ms before the onset of the stop consonant following the nasal vowel [ãĩ].

These results show that there is a difference in the temporal coordination of gestures between the two speakers. RD exhibits longer durations in both cases compared to GV, which may imply differences in the coordination of velum lowering and closure during the production of the back nasal diphthong [ãĩ].

The timing of velum closure and stop /t/ onset during the production of the back nasal diphthong [ãĩ] differs between the two speakers (RD and GV), and there is also a difference in the timing of velum lowering after the onset of [ãĩ].

For RD, the average duration is 62 ms, indicating that the velum lowering starts around 62 ms after the onset of the [ẽ̃]. For GV, the average duration is 58 ms, suggesting that the velum lowering begins around 58 ms after the onset of the [ẽ̃]. For RD, the average duration is 27 ms, which means the velum closure ends about 27 ms before the onset of the stop consonant following the [ẽ̃]. For GV, the average duration is 93 ms, indicating that the velum closure ends approximately 93 ms before the onset of the stop consonant following the [ẽ̃].

RD exhibits a slightly longer duration of velum lowering after the onset of [ẽ̃] compared to Speaker GV, suggesting a delayed velum lowering in this context. However, for the velum closure and stop /t/, RD shows a much shorter duration compared to Speaker GV.

These findings suggest that the coordination of velum and tongue movements during the production of [ẽ̃] is different for the two speakers. The differences in the timing of velum lowering and closure may contribute to the distinct acoustic characteristics of the nasal diphthong [ẽ̃] produced by each speaker. As with the previous analysis, it's important to consider that these values represent averages based on the analyzed cases,

These results suggest that in both speakers RD and GV, the closing of the soft palate (velum lowering) and the offset of voicing are coordinated in phase. Specifically, the soft palate closing before the onset of the following stop consonant /t/. This phase relationship ensures that the nasal resonance (velum closure) is effectively separated from the oral portion of the stop consonant onset.

By closing the soft palate before the release of the stop consonant, the nasal passage is effectively closed off, allowing air to be channeled exclusively through the oral cavity during the stop consonant's release. This prevents any unwanted nasal

airflow during the production of oral stops like /t/, resulting in a crisp and distinct oral release of the consonant.

This phase relationship observed in both speakers demonstrates the precise coordination and timing required for the smooth transition from a nasalized vowel to a voiceless oral stop, contributing to the overall clarity and intelligibility of speech production.

The differences in the temporal timing between the two speakers indicate individual variation in the coordination of velum movements with other speech events. However, the general pattern of the soft palate completing its lowering movement before the onset of the stop consonant is consistent across both speakers.

The synchrony of gestures took place when the soft palate had a notable opening. This opening allowed the internal pressure in the oral cavity to equalize with atmospheric pressure, thus limiting the release of the consonant. In the context of our study, if the consonant following the nasal vowel were a fricative or a sonorant consonant, the soft palate might have extended its activity beyond that consonant.

The figures below illustrate the synchronized temporal trajectory of the velum's opening and closing movements, derived from the acoustic signals recorded for speakers RD and GV.

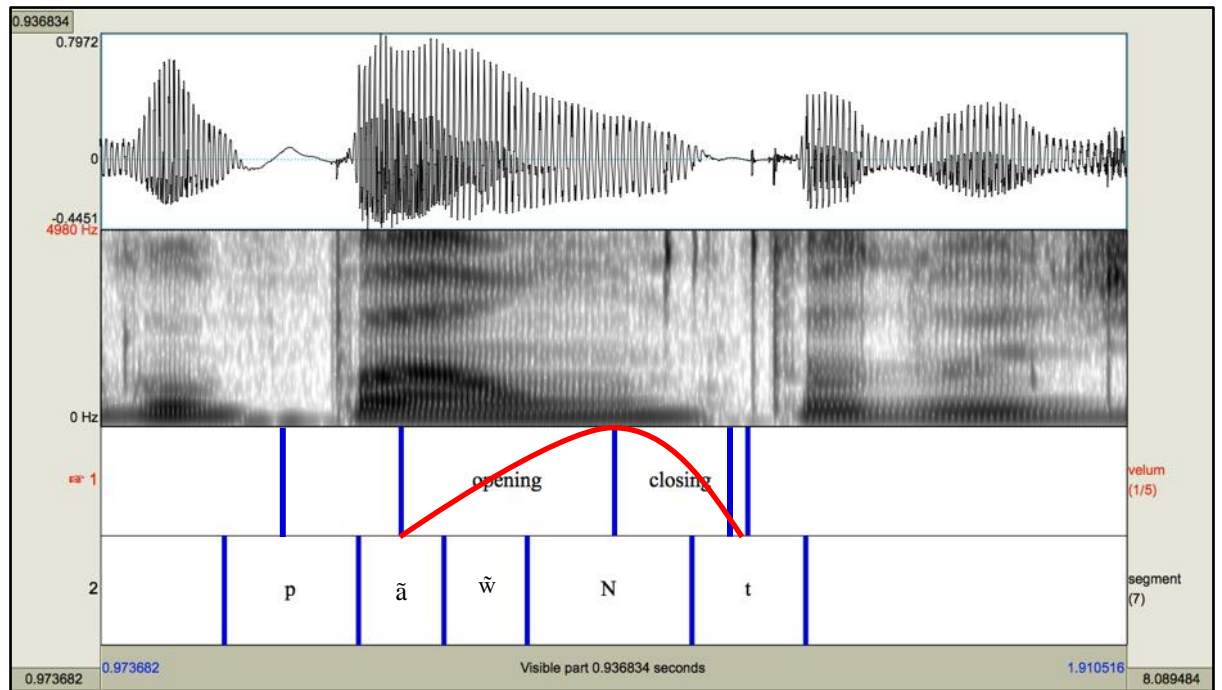


Figure 93: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: “*pãõ*”. Speaker: RD.

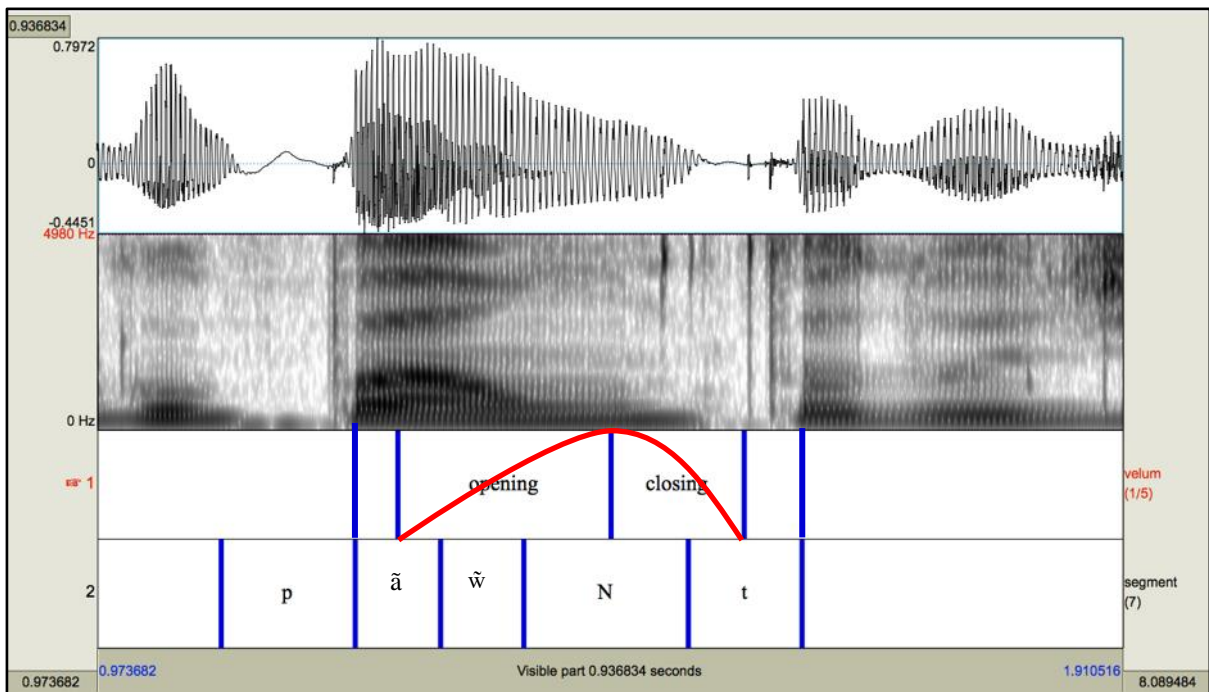


Figure 94: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: “*pãõ*”. Speaker: GV.



In the front nasal vowel, the timing of the soft palate-opening gesture differed from that observed in the back nasal diphthongs. Specifically, the soft palate-opening gesture exhibited greater synchrony with the front nasal-vowel release.

The soft palate started opening after the vowel was released 10 ms for RD and 27 ms for GV. This synchrony was probably caused by the tongue position being similar in the front oral diphthong and the front nasal vowel.

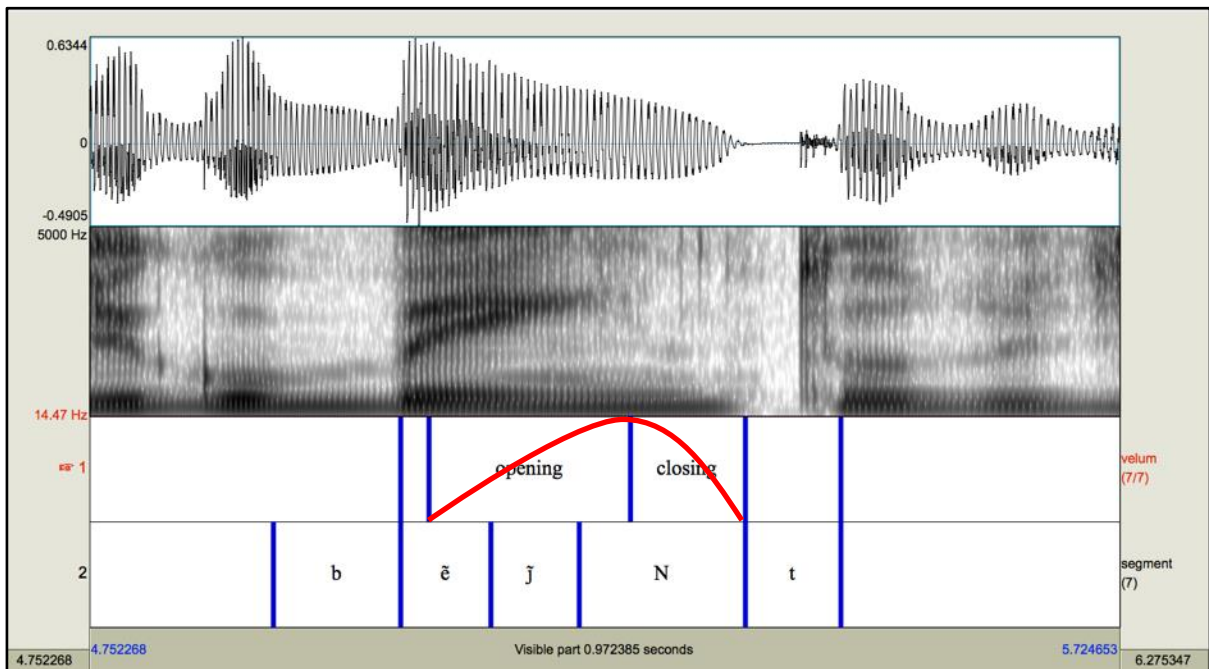


Figure 95: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: "bem". Speaker: RD.

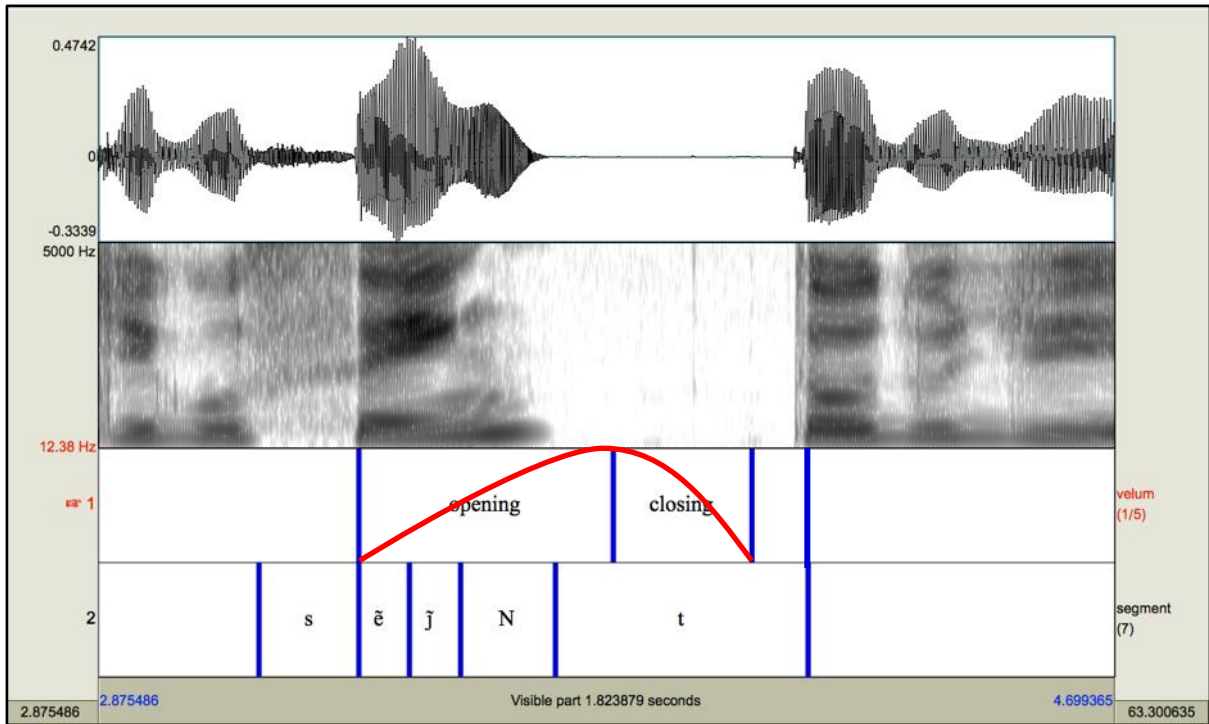


Figure 96: Oscillogram (top) and spectrogram (middle). (1) Textgrid with the temporal velum activation. (2) Textgrid with the phonetic transcription. Keyword: “bem”. Speaker: GV.

In some cases, velum activity began together with the vowel onset time for both speakers while producing a front nasal vowel. As we noted in our image study, the opening of the soft palate started before the nasal vowel. Velum anticipation occurred while producing the nasal consonants /m/ and /n/, which are highlighted in clear blue in table 32, and in producing the voiced glottal fricative /h/, which is highlighted in grey. This kind of behavior was noticed in the fiberscope data.

In cases of the back nasal diphthong, the velum began its activity after the vowel was released. For the front nasal vowels, the velum’s lowering movement began almost simultaneously with the release of the nasal vowel. From this, we can infer that the greater volume of nasal airflow in the front nasal vowel may be related to the greater synchrony between the soft palate’s lowering and the vowel’s release.

Despite the velum being considered a passive articulator, it moved and acted together with the language trajectory in a gestural constellation to maximize the effect of nasality in speech.

## **5.6 Partial Conclusion**

Based on the data and analysis presented, it was observed that nasal diphthongization involves a partial constriction of the oral cavity. This suggests that during the production of nasal diphthongs, the oral cavity experiences some degree of narrowing. The oral cavity is not completely closed but rather exhibits a partial obstruction or constriction. In some cases, a total constriction of the oral cavity might also be possible. Additionally, the constriction of the oral cavity could be in the palatal or velar region, and it is attributed to glide constriction.

In 25% of the cases produced by RD, an oral-block condition was observed due to the coordination of the tongue's back movement and velum lowering. Specifically, after the tongue reached its target position, the velum initiated the closing movement. The targets for the glide and the velum were in phase, indicating a synchronization between these gestures. The closure of the velum was dependent on the constriction caused by the glide movement.

The front nasal vowel can be analyzed as consisting of three distinct parts: an oral vowel, a nasal glide, and a nasal appendix. Similarly, both the back nasal diphthong and front nasal vowel involve tongue elevation, with the back nasal diphthong directed towards the velar region and the front nasal vowel towards the palatal region. The magnitude of the tongue gesture was more prominent, and the articulatory targets were extended, resulting in clearer and more defined movements. We noticed the gesture's configuration during the visual analysis of the pellets' motion.

The back oral diphthong exhibited more hyperarticulated target positions and tongue trajectories compared to the back nasal diphthong. As a result, the point of articulation for the back oral diphthong was further back than in the case of the back nasal diphthong.

Oral diphthong production demonstrated higher stability and more consistent tongue movement, leading to consistent and accuracy in reaching its intended articulatory target. On the other hand, the back nasal diphthong showed less regular tongue movement due to the unstable environment, resulting in variability in the constriction region and articulatory manner of the tongue gesture. These variations were observed to vary according to the individual speaker.

In the figure below, we illustrate how the tongue reached its target. In our results, RD maintained a more stable tongue-displacement trajectory than GV, but in both informants, the nasal vowel  $[\tilde{a}]$  had a target higher than the oral vowel  $[a]$ .

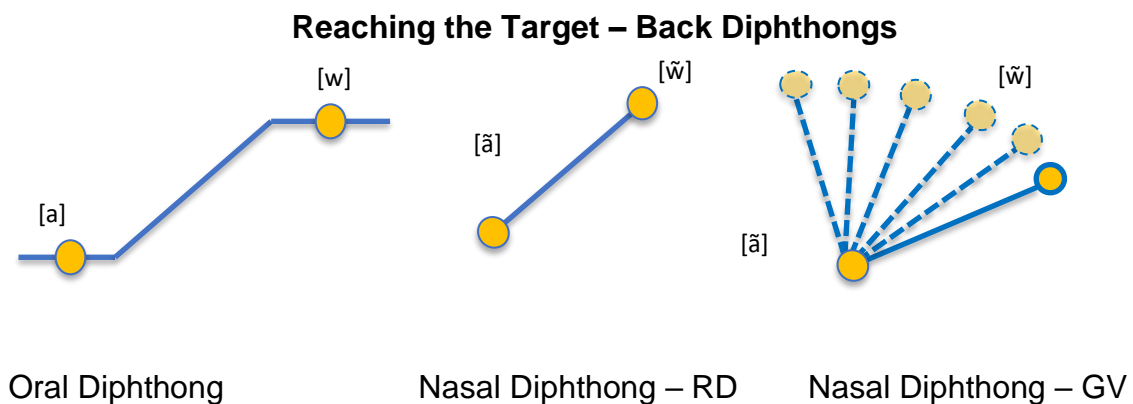


Figure 97: Illustration of the tongue-displacement model.

For RD, the nasal vowel target  $[\tilde{a}]$  was positioned about 0.56 cm (x-axis) further back and in a higher position, and about 0.70 cm (y-axis) further back than the  $[a]$  target. The  $[\tilde{w}]$  target was about 0.70 cm higher than the position of the  $[w]$  target.

RD performed the total occlusion of the oral cavity at times, which demonstrates that the nasal velar consonant can appear in some cases. However, this process depends on extralinguistic and physiological factors. For GV, the nasal vowel target [ã] was positioned about 0.5 cm above the target of [a]. At the same time, the [w̃] target was closer to 0.08 cm, which is below the position of the [w] target.

The precision of the articulatory movement changed according to the shape of the palate, generating slight individual variations. Probably, the variation in GV's speech production occurred because of instability and physiological factors.

During the production of the front nasal vowel, vowel breaking occurred, and this vowel was split into two elements: the middle vowel [ẽ] and a palatal glide [j̃]. The nasal diphthongization of the vowel [ẽ] was a mirror of the articulatory production of the front oral diphthong for both RD and GV. Thus, this similarity confirms the production of oral-cavity constriction in the palatal region.

### Reaching the Target – Front Segments

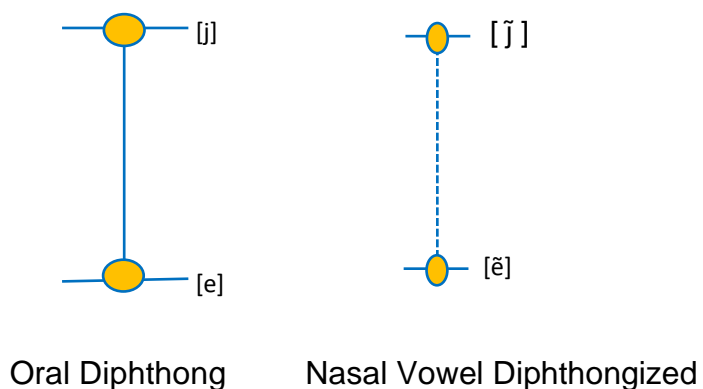


Figure 98: Illustration of the tongue-displacement model for RD and GV.

The velum-opening trajectory began when the vowel was released. Through the temporal comparison of the acoustic and articulatory data, we noticed that in back diphthongs, the vowel was oralized for more or less 63 ms for RD and 62 ms for GV

and the production was gradually produced, becoming nasalized. For the front vowel, the values were lower: 10 ms and 27 ms for RD and GV, respectively.

The velopharyngeal movement created a section around  $x = 0.27$  cm in width and  $y = 0.31$  cm in height for RD. For GV, the values are  $x = 0.19$  cm in width and 0.26 cm in height for the back nasal diphthong. For the diphthongized nasal vowel [ẽ], these values are  $x = 0.38$  cm in width and  $y = 0.46$  cm in height for RD, and for GV the values are  $x = 0.24$  cm in width and  $y = 0.23$  cm in height. In the figure below, we illustrate a diagonal trajectory performed by the gesture of the soft palate.

### Reaching the Target – Nasal Vocalic Segments

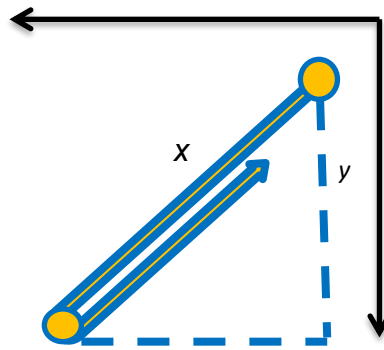


Figure 99: Illustration of velum-lowering model for RD and GV.

The velum trajectory exhibited a diagonal displacement, beginning in a back region and transitioning through a central region. Despite having different activation times, there was a correlation between the gestures of the soft palate and the tongue. The velum movement was divided into sequential gestures like opening, target, and closing.

The initial portion of the vowel was oral because, in most cases, there was no anticipatory soft palate movement in the back nasal diphthongs, while in the front nasal vowels, there was an anticipatory process. To back segments, a small portion is oral, and to front segments, all the portion was nasalized, at which point there was a gradual

soft palate opening and the displacement of the tongue to a higher position (constriction of the glide). The spatiotemporal movement and coordination between the velum and tongue displacement resulted in the nasal appendix and nasal diphthongization.

## CHAPTER 5: CONCLUSION

---

The goal of this thesis was to present a hypothesis about the process of palatalization and velarization of nasal vowel segments in the Paulistano dialect, from a Laboratory Phonology perspective, based on the analysis of the temporal and gradual characteristics of the articulatory gestures involved in nasal diphthongization.

Therefore, we utilized four experimental perspectives: aerodynamic, acoustic, imagery, and articulatory. The acoustic data assisted our investigation by providing qualitative data to support the quantitative analysis of the three experiments described in this thesis.

The bi-phonemic hypothesis challenges the conventional view of nasal diphthongs as sequences of two separate phonemes, suggesting instead that some nasal diphthongs are formed through the assimilation of a vowel and a nasal consonant into a single phonemic unit.

The traditional view of nasal diphthongs in BP considers them as sequences of a vowel and a nasal consonant in rapid succession, where the nasal consonant is often represented as a coda after the vowel. For example, the word "bem" [bẽj] is conventionally analyzed as consisting of the vowel /e/ followed by the nasal consonant.

In contrast, our hypothesis proposes that nasal diphthongs like "bem" are a sequence of two distinct segments (/ẽ/ and /j/). Instead, it suggests that the nasal glide /j/ is the result of nasal vowel /ẽ/ diphthongization, resulting in diphthong as a single, unified phonemic unit.



According to our hypothesis, the nasal glide becomes an integral part of the vowel's articulation and is not analyzed as a separate consonant, even if this is an allophonic phenomenon.

The term "bi-phonemic" is appropriate to our hypothesis because it refers to the perceived dual nature of these nasal diphthongs (vowel and glide), where they combine elements of both vowels and nasal glide. The nasalization of the vowel is not considered a separate segment, but rather an inherent part of the vowel itself.

In our thesis, we also used the backgrounds of Speech Control Theory (KINGSTON & DIEHL, 1994) and Gestural/Articulatory Phonology (BROWMAN & GOLDSTEIN, 1992) to analyze our hypothesis. Speech Control Theory provided insights into how to understand the planning and execution of articulatory gestures during the nasal diphthongizing production. It also helped us understand how speakers coordinate and time their articulatory movements.

Gestural/Articulatory Phonology, on the other hand, allowed us to understand the dynamic nature of speech gestures and their interplay during the process of nasal diphthongization. This theory emphasizes that speech is produced through coordinated gestures that target specific articulatory goals. It provided a framework to analyze how the tongue and soft palate gestures interacted during the assimilation processes between vowel and nasal.

By utilizing these theoretical frameworks, we were able to show that the assimilation processes between vowel and nasal led to the emergence of a nasal glide with the vowel articulatory point. We observed how the gestures of the tongue and soft palate dynamically adjusted and coordinated to create a nasal glide along with the vowel articulation.

The temporal and gradual characteristics of these gestures were analyzed to understand their precise timing and coordination. We found evidence supporting the hypothesis that the assimilation processes between vowel and nasal result in the formation of a nasal glide.

Timing Theory suggests that oral and nasal gestures occur simultaneously, leading to similar durations and temporal overlap between oral and nasal airflows. Timing Theory also proposed sequential nasalization either before or after the oral articulation, resulting in different durations of nasal and oral airflows.

Coarticulation Theory emphasizes mutual influence and temporal overlap between nasalization and oral articulation, predicting slightly different durations depending on the specific coarticulatory patterns. The Gestural Coordination Theory predicted tightly synchronized durations of nasal and oral airflows, reflecting the coordination of velum and oral gestures.

Examining the velopharyngeal port's opening and closing trajectories during nasal diphthongization provided crucial insights into specific articulatory movements involved in nasal diphthong production. Describing global velar articulation and velopharyngeal port contributed to understanding velopharyngeal functioning and its role in shaping oral and nasal acoustic properties.

Analyzing velum movement in correlation with acoustic signals of front nasal vowels and back nasal diphthongs helped us to explore the relationship between articulatory gestures to understanding the role of assimilatory adjacent sounds influencing velopharyngeal coordination and nasalization patterns.

Comparing velum gestures for front and back nasal segments expanded knowledge of articulatory variation in nasal diphthongization and the specific strategies employed in producing different nasal sounds.

In this conclusion, we examined the coordination of tongue movement and soft palate using symbolic representations, phonetic planning, and speech implementation. The abstract notations that represent the linguistic provided a way to analyze and understand the coordination and timing of articulatory gestures in nasal diphthongization.

We examined the instances of gestural overlaps, where the tongue and nasal gestures occur simultaneously or sequentially, paying attention to the timing and duration of each gesture concerning the others. We investigated the coarticulation patterns from the oral and nasal segments, looking for patterns of anticipatory or carryover nasalization.

We used our findings to conclude the phonetic and phonological properties of nasal diphthongization. Consider how the symbolic representations capture the articulatory coordination and temporal relationships involved in nasal diphthongization.

So, our goal was to understand the effects of space and time on articulatory, aerodynamic, and acoustic output from a phonetic and phonological experimental perspective to identify the nature of nasal diphthongization (vowel, glide, and nasal appendix), that one element can combine elements of both vowel and nasal glide.

Our study aimed to understand how the temporal synchronization and timing precision of the oral and nasal articulatory movements. Our findings suggest that the temporal precision of the velopharyngeal port is synchronized with the tongue movement in the symbolic phonological representations.

Our goal was to develop a speech production system to find the patterns of nasal diphthongization production, that combines symbolic phonological representations and phonetic planning to explain the motor implementation of the speech components,

specifically the airflow variations and formant movements, the velopharyngeal port movement, and the articulatory trajectories.

Thus, we analyzed phonological and phonetic planning through variations in the concrete implementation of speech (TURK & SHATTUCK-HUFNAGEL, 2020). In addition, we aimed to understand the gestural coordination pattern of the VN during the production of the front nasal vowel and the back nasal diphthong, from the perspective of gesture phonology and speech control theory. These theories supported our goal of describing and analyzing nasal diphthongization and the dynamic movement of the soft palate during speech production.

We described the contrasts between oral and nasal vowel segments (a nasal vowel and a nasal diphthong) to explain the gestural overlap between articulatory gestures. We also investigated the different degrees of gestural activation (GOLDSTEIN et al., 2007).

As nasal diphthongization is a complex phenomenon to study. Thus, we developed three experimental analyses to measure gesture coordination, time, and the gesture magnitude of the tongue and the soft palate trajectory to conclude that Camara Jr., Linpski, and Shosted were correct in their hypotheses of the biphonemic nature of the nasal vowels and the emergence of a velar constriction.

Through our data, we noticed that there is an assimilation process of the articulation point, reiterating the biphonemic thesis of Camara Jr. (1964). However, the assimilated articulation point is from the vowel. Linpski (1975) is also correct in concluding that there is a constriction of the oral cavity in the velar region. Our findings corroborate with these authors' hypothesis, but, and contribute to showing the production of velar and palatal constriction. In our data, the blockage of oral track occurred in a few productions from the same speaker, during the velar nasal glide.

In the case of the production of the velar nasal glide, there is both total blockage of the oral cavity (referring to the production of the emerging velar nasal consonant proposed by Shosted, 2006) and partial blockage. So, we came to this conclusion from the analysis of our three experimental approaches.

Chapter 2 was based on the results of the EVA2 device, which was used to collect variations in oral and nasal airflows during speech production. Chapter 3 described the use of a nasofiberscope to collect image data and discussed the measurement of the opening and closing of the soft palate and the velopharyngeal port.

In Chapter 4, we discussed the EMMA 2D articulatory data. These data helped us to understand the articulators' trajectories and their relationship in space and time. This device provided a more detailed analysis of the production of nasal diphthongs. Thus, we generalized about the velum displacement in vocalic segments based on the phonetic details.

We can confirm that diphthongs are complex syllable nuclei via the dynamic approach to gestural coordination and the phasing relationships between oral and nasal vowel/glide segments. The nasal vowel undergoes segmentation due to nasality, which adds one more gesture to the gestural constellation, thus making the syllable complex.

Our data analysis suggested that oral diphthongs were produced in two stages and nasal diphthongs in three stages. The tongue plays an important role in the production, and velum also plays an essential and dynamic role.

Oral diphthong gestures are characterized by having two targets with different heights and constrictions. Back nasal diphthongs and front nasal vowel gestures are

characterized by having three targets. The tongue has two targets, and the velum has one target. As shown by Cagliari (1977), all nasal vowels in BP are diphthongized.

The timing duration of the tongue and velum gestures play a crucial role in achieving the nasal quality of diphthongs beyond simple oronasal coupling. This is because the velum movement directly affects the configuration of the vocal tract, acting as a significant parameter in shaping the acoustic properties of the sound.

In the case of nasal diphthongs, the timing of both tongue and velum movements is critical because the velum movement directly affects the vocal tract's geometry. In nasal diphthongs, the timing of the velum closure and opening is synchronized with the tongue movement. The timing and coordination of these gestures are crucial to achieving the appropriate nasality of the diphthong. This highlights the intricate coordination and control involved in speech production and the role of the velum as a significant parameter in shaping the vocal tract's acoustic properties.

Consequently, we believe that the movement of the velum is not the only gesture that is responsible for producing the oral-nasal distinction, but that the highest position of the vowels (tongue) is also important.

When the velopharyngeal opening is active, oral constriction appears to contribute to the high volume of nasal airflow, which alters the sound's spectral properties via nasal resonance.

Negative nasal airflow before the beginning of nasal vowel production coincided with velopharyngeal port movement in some cases. The effect that we observed in the aerodynamic data was probably related to this movement of the side walls.

In terms of the acoustic-articulation relationship, the tongue gesture in the back nasal diphthong had a different gestural organization from the back oral diphthong. By

contrast, the front nasal vowel gestural organization was similar to the front oral diphthong.

However, in both cases, there was a change in the tongue trajectory concerning the location and degree of constriction and the soft palate gesture activation, which characterizes nasal diphthongization.

The coordination between the tongue and the soft palate during nasal diphthongization is achieved through an organized and interconnected system known as a gestural constellation.

Importantly, this coordination is not in a linear phase, meaning that the gestures do not follow a simple sequential pattern. Instead, the tongue and soft palate gestures move simultaneously but at different times. Specifically, the tongue gesture is activated before the soft palate gesture, as depicted in the figure below.

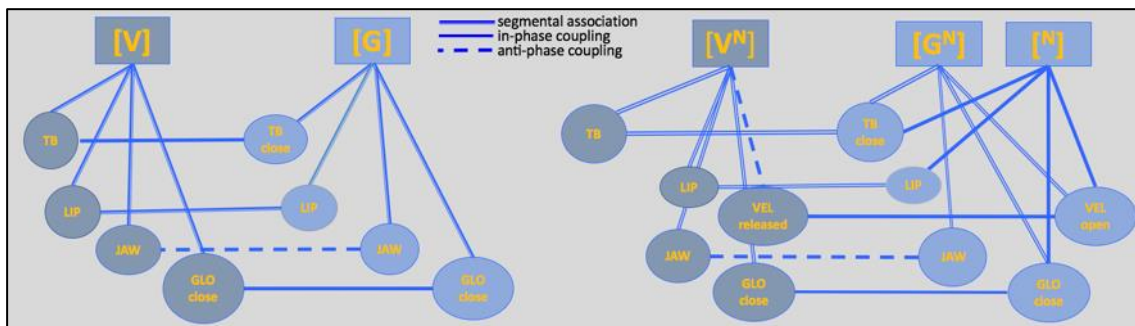


Figure 100: Gesture guideline comparing the phases of gestures in the production of oral diphthongs and nasal diphthongs. The solid lines show the motions that are phase-coupled, and the dashed lines represent the gestures' phase.

During nasal diphthongization, the coordination between the tongue and the soft palate is orchestrated by nonlinear planning oscillators. Each speech gesture involved in this process is associated with its oscillator, and its activation is triggered at a specific phase.

When the tongue and soft palate gestures require synchronization, their respective oscillators become coupled, enabling them to establish a stable pattern of

relative phasing during the planning stage. Once this pattern stabilizes, the activation of each gesture is initiated by its corresponding oscillator, leading to a consistent and accurate timing relationship between the two gestures.

The timing and the phasing relationships between the articulatory gestures were different in the syllables shown above. The oral and nasal gestures established a phase relationship. The velum gesture activation time was variable and had an unstable relationship with the oral gesture. The phase seems to be a characteristic of nasalization in our data. The activation of the nasal gesture begins after the vowel gesture and characterizes the phase relationship, represented by the dotted line.

The figure below shows a dynamic model of gestural stages comparing the oral diphthong to the nasal diphthong/vowel to illustrate how the gestures are organized and coordinated within the syllable from the coupled oscillator model (NAM and SALTZMAN, 2003; GOLDSTEIN et al., 2007a). The gestural notation was adapted from the model by Tilsen (2014).

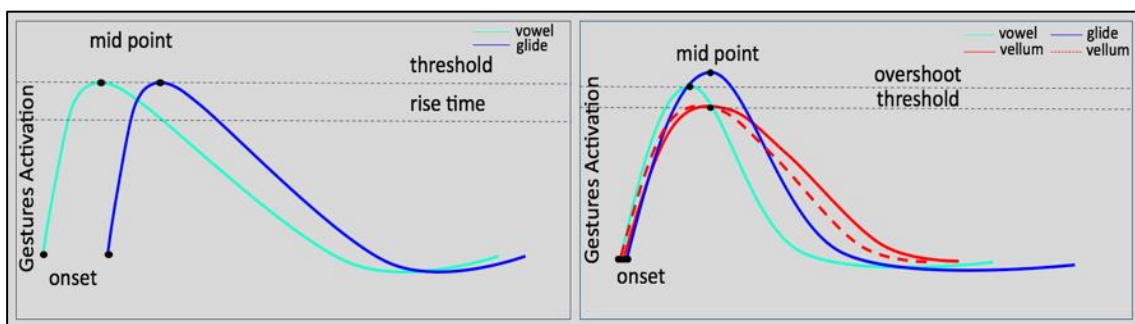


Figure 101: Dynamic oscillatory modeling (Tilsen, 2014) illustrates the coordination planning from the activation of tongue height and soft palate trajectories. The tongue gestures are co-selected and coordinated, while the nasality gesture is competitively selected in relation to the oral gesture.

We suggest that nasalization generates changes in the tract's geometry, causes the tongue gesture to exceed the intrinsic time limit, and creates a constriction in the tract by expanding its trajectory.

The relationship between the temporal coordination of the tongue gesture and the soft palate is less stable. The tongue trajectory generates oral cavity constriction.



There is a competition between the oral and nasal gestures, which is caused by the strong coupling in VN, making coordination between these gestures difficult (NAM, GOLDSTEIN & SALTZMANN, 2009) due to the realization of the glide.

This assumption about inter-gesture coordination within the syllable corroborated the hypothesis of stability in the gestures involved in producing the oral syllable VG. The coupling between gestures in the VG syllabic pattern was the strongest, making it more stable and easier to parameterize than a complex syllabic pattern such as VN > VGN.

The nasal glide modified the syllable structure to make it more stable and harmonic. An assimilatory process used vowel constricts location to orient the glide constriction. The emerging glide constriction increases the nasal airflow and amplifies the nasal resonances in the acoustic signal.

In the case of the back nasal diphthong, the tongue gesture varied between complete or reduced articulation. In the diphthongized nasal vowel, the tongue gesture attained its full magnitude, and the soft palate gesture maintained its complete magnitude. However, the airflow volume was directed to the nasal cavity in both cases because of glide constriction.

Nasal diphthongization is a vowel shift resulting from gesture coordination in the phase and variable magnitude. The articulatory gesture is characterized by intrinsic time and overlapping gestures during speech production (GOLDSTEIN, 2008; FOWLER & BROWN, 2000) to overshoot the perceptive thresholds.

The nasal gesture overlaps with the oral gesture due to the coordination of articulatory gestures resulting from mismatches between the temporal aspects and the magnitude activation of these gestures. The emergence of palatalization and velarization serves to maintain phonemic contrast, considering the inter-gestural

coordination and the occurrence of varying degrees of tongue and velum gesture overlap.

The development of the symbolic language aspect is attributed to the stable dynamics (BERTI, 2006) of both the tongue and the velum's gesture trajectories, as described by Berti (2006). The articulatory gesture of the soft palate shows a consistent trajectory in its three phases: onset (opening), target, and offset (closing). However, the velum's temporal oscillator exhibits an initial phase delay at the beginning of the vowel.

The gesture overlaps or the excessive activation of the magnitude depends on the speech rate, the speech style (VAN LIESHOUT & GOLDSTEIN, 2008), social stratification, and the educational level, but we didn't manipulate these variables in our study. For this reason, nasal diphthongization is an ongoing phenomenon in the São Paulo dialect, but not in all dialects of BP.

When examining nasalization data, we analyze their particularities and how they fit into the laboratory and gestural phonology models. However, we did not take into consideration all aspects of nasalization, which is widely produced in Brazilian Portuguese, and we also did not consider its perception, for example.

This research was done to find out how complex syllables and phonological primitives with a motor, symbolic, or dynamic nature are made. This makes it possible to look at the steps in the process of speech production and how they change from one step to the next. We hope that this work will provide a comprehensive description of the data to the field and speech sound changing and speech modeling.

## REFERENCES

---

ALBANO, E. C. (2001). *O Gesto e Suas Bordas: Esboço de Fonologia Acústico-Articulatória do Português Brasileiro*. Campinas: Mercado de Letras.

ALBANO, E. C. (2005). Sobre o Abrimento 3 de Mattoso Câmara: Pistas Fonotáticas para a Classe das Líquidas (On Mattoso Câmara's Aperture 3: Phonotactic Clues To The Liquid Class The Joaquim Mattoso Câmara Junior's Phonological Ideas). *Estudos da Língua(gem)*, 2(1), 45-66.

ALBANO, E. C. (2012). Uma introdução à dinâmica em fonologia, com foco nos trabalhos desta coletânea. *Revista da ABRALIN*, 2, 1-30.

ALTMANN, E. B. C. (1997). Anatomia e fisiologia do esfíncter velofaríngeo. In: ALTMANN E. B. C. *Fissuras labiopalatinas* (4th ed., pp. 133-156). Carapicuíba: Pró-fono Departamento Editorial.

AMELOT, A. (2004). *Etude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français*. PhD Dissertation, Université Paris III.

BAKEN, J., & ORLIKOFF, J. (1993). *Clinical measurement of speech and voice*. San Diego: Singular.

BARBOSA, P. A. ALBANO, E. C. (2004). Brazilian Portuguese. *Journal of the International Phonetic Association*, 34(2), 227-232.

BAVEGARD, M., FANT, G., GAUFFIN, J., & LILJENCANTS, J. (1993). Vocal tract sweep tone data and model simulations of vowels, laterals, and nasals. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, 4, 43-75.

BEDDOR, P. S. (2009). A coarticulatory path to sound change. *Language*, 85, 785–821.

BEDDOR, P. S. (2009). A coarticulatory path to sound change. *Language*, 85(4), 785-821.

BEDDOR, P. S., MCGOWAN, K. B., BOLAND, J. E., COETZEE, A. W., &

BRASHER, A. (2013). Perceptual time course of coarticulation. *Journal of the Acoustical Society of America*, 133(4), April.

BERCI, G., & Forde, K. A. (2000). History of endoscopy: What lessons have we learned from the past? *Surgical Endoscopy*, 14, 5–15.

BERNSTEIN, N. (1967). *The coordination and regulation of movements*. London: Pergamon Press.

- BERTI, L. C. (2006). Aquisição incompleta do contraste entre /s/ e /ʃ/ em crianças falantes do português brasileiro. (Doctoral dissertation). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.
- BERTI, L. C., & FERREIRA-GONÇALVES, G. (2012). A aquisição do contraste entre /t/ e /k/ sob a ótica dinâmica. *Revista da ABRALIN / Associação Brasileira de Linguística*, Número especial 2. Organização de Eleonora Albano. Natal, RN: UFRN.
- BISOL, L. (1991). O ditongo em português. *Boletim da ABRALIN*, 11, 51-58.
- BISOL, L. (1994). Ditongos Derivados. *DELTA*, 10, 123-140.
- BISOL, L. (1996). *Introdução a Estudos de Fonologia do Português Brasileiro*. Porto Alegre: EDIPUCRS.
- BISOL, L. (1998). A nasalidade, um velho tema. *DELTA*, 14, 27-46.
- BISOL, Leda. (1989). O ditongo na perspectiva da fonologia atual. *DELTA*, 185-224.
- BJÖRK, L. (1961). Velopharyngeal function in connected speech - studies using tomography and cineradiography synchronized with speech spectrography. *Acta Radiologica, Supplement 202*, 1-94.
- BJUGGREN, G., & FANT, G. (1964). The nasal cavity structures. *Speech Transmission Laboratory - Quarterly Progress and Status Report*, 4, 5-7.
- BROWMAN, C., & GOLDSTEIN, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155-180.
- BUDER, E. H. (2005). Determination of Nasality with Acoustic and Aerodynamic Parameters. *The Cleft Palate-Craniofacial Journal*, 42(4), 379-387.
- BZOCH, K. R., GRABER, T. M., & AOBA, T. (1959). A study of normal velopharyngeal valvins for speech. *Cleft Palate Bulletin*, 9, 3-10.
- CAGLIARI, L. C. (1977). An experimental study of nasality with particular reference to Brazilian Portuguese. Ph.D. Dissertation, University of Edinburgh.
- CAGLIARI, L. C. (1977). *Nasalização em português. Contexto*.
- CAGLIARI, L. C. (1999). *Fonologia do Português. Análise pela Geometria de Traços e pela Fonologia Lexical*. Author Edition. Campinas.
- CAGLIARI, L. C. (2007). *Elementos de fonética do Português Brasileiro*. São Paulo: Paulistana.
- CALLOU, D., & LEITE, Y. (1990). *Iniciação à fonética e à fonologia*. 4th Edition. Rio de Janeiro: Zahar.

- CÂMARA JR., J. M. (1970). *Estrutura da língua portuguesa*. Petrópolis: Vozes.
- CÂMARA JR., J. M. (1979). *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão.
- CÂMARA JR., J. M. (1984). *Problemas de linguística descritiva*. Petrópolis: Vozes.
- CAMRAN, N. (2005). *History of Endoscopy*. Society of Laparoendoscopic Surgeons. Retrieved from <http://laparoscopy.blogs.com/endoscopyhistory/>
- CASSEL, D., & ELKADI, H. (1995). Anatomy and physiology of the palate and velopharyngeal structures. In R. J. Shprintzen & J. Bardach (Eds.), *Cleft Palate Speech Management: A Multidisciplinary Approach* (pp. 45-58). New York: Mosby.
- CASTRO, S. V. (199?). *Anatomia fundamental*. 3rd Edition. McGraw-Hill do Brasil.
- CHILDERS, D., & DING, C. (1991). Articulatory synthesis: nasal sounds and male-female voices. *Journal of Phonetics*, 19, 453-464.
- CHOMSKY, N., & HALLE, M. (1968). *The sound pattern of English*. New York: Harper and Row.
- CLEMENTS, G. N. (1985). The Geometry of Phonological Features. *Phonology Yearbook*, 2, 225-252.
- CLEMENTS, G. N., & HUME, E. V. (1995). The internal organization of speech sounds. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 225-252). London: Blackwell.
- CLOSURE, P. (1982). An electromyographic study of velopharyngeal-nasal closure. *Folia Phoniatrica*, 34, 246-257.
- CLUMECK, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, 4, 337-351.
- COHN, A. C. (1990). Phonetic and phonological rules of nasalization. *UCLA Working Pap. Phonetics*, 76, 1-224.
- CONTERNO, G., CIELO, C. A., & ELIAS, V. S. (2010). Fissura palatina reparada: fechamento velofaríngeo antes e durante o som basal. *Braz. j. otorhinolaryngol.*, 76(2), 185-192.
- CROFT, C. B., SHPRINTZEN, R. J., & RAKOFF, S. J. (1981). Patterns of velopharyngeal valving in normal and cleft palate subjects: a multi-view videofluoroscopic and nasendoscopic study. *Laryngoscope*, 91(2), 265-271.
- CURTIS, J. (1968). Acoustic of speech production and nasalization. *Cleft Palate and Communication*, pp. 27-60.

CURTIS, J. (1970). The acoustics of nasalized speech. *Cleft Palate Journal*, 7, 380–396.

DANG, J., HONDA, K., & SUZUKI, H. (1994). Morphological and acoustical analysis of the nasal and the paranasal cavities. *The Journal of the Acoustical Society of America*, 96(4), 2088-2100.

DANGELO, J. C. (1997). *Anatomia humana sistêmica e segmentar*. São Paulo: Atheneu Publishing Company.

DE GROEN, P. R. (2017). History of Endoscope [Scanning Our Past]. *Proceedings of the IEEE*, 105(10), 1987-1995.

DELATTRE, P. (1954). Les attributs acoustiques de la nasalité vocalique et consonantique. *Stud. Ling.*, 8(2), 103-109.

DELATTRE, P., & Freeman, D. C. (1968). A dialect study of American r's by x-ray motion picture. *Linguistics*, 6(44), 29-68.

DELVAUX, V. (2003). *Contrôle et connaissance phonétique: les voyelles nasales Du français*. PhD Dissertation, Université Libre de Bruxelles.

DELVAUX, V., AMELOT, A., & ROSSato, S. (2012). Invariant nasality in the labiodental flaps of Belgian French. In *Proceedings of the 13th Annual Conference of the International Speech Communication Association (Interspeech 2012)*, 2846-2849.

DELVAUX, V., METENS, T., & SOQUET. (2002). French nasal vowels: acoustic and articulatory properties. In *7th International Conference on Spoken Language Processing (ICSLP 2002)*. Published 16 September.

DEMASI, R., & DEMOLIN, D. (2009). A dinâmica dos movimentos do véu em ditongos nasais do português brasileiro: Aspectos Aerodinâmicos. *X Congresso Nacional De Fonética e Fonologia. IV Congresso Internacional de Fonética e de Fonologia. Comunicação Oral*. Niterói.

DEMOLIN, D. (2012). The experimental phonology. *Revista da ABRALIN*, 11(1). Retrieved from <https://revista.abralin.org/index.php/abralin/article/view/1136>.

DI IEVA, A., TAM, M., Tschabitscher, M., & Cusimano, M. D. (2014). A journey into the technical evolution of neuroendoscopy. *World Neurosurgery*, 82(6), E777–E789.

DICKSON, D. R. (1962). An acoustic study of nasality. *Journal of Speech and Hearing Research*, 5, 103-111. *Disorders Related To Cleft Lip And Palate* (2004). 5th ed. Boston: Little and Brown.

DURAND, T. M., & LIDDELL, S. K. (2009). An articulatory investigation of word-final English /t/ using electromagnetic articulography. *Journal of the Acoustical Society of America*, 125(1), 396-410.

ELGENDY, A. (2001). Aspects of Pharyngeal Coarticulation. [Thesis, fully internal, Universiteit van Amsterdam]. LOT.

ENGEL, R. M. E. (2003). Epochs in Endourology: Philipp Bozzini — The Father of Endoscopy. *Journal of Endourology*, 17(10), 859–863.

ENTENMANN, G. L. (1976). The development of nasal vowels. Ph.D. Dissertation: University Texas.

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

FANT, G. (1970). *Acoustic theory of speech production*. Paris: Mouton.

FENG G., & KOTENKOFF, C. (2004). Vers un nouveau modèle acoustique des nasales basées sur l'enregistrement bouche-nez séparé. In *Actes des XXVème Journées d'Etude sur la Parole*, Fès, Maroc, pages 213-216.

FENG, G. & CASTELLI, E. (1996). Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. *The Journal of the Acoustical Society of America*, July 1996. DOI: 10.1121/1.414967.

FENG, G., & CASTELLI. (1996). Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. *Journal of the Acoustical Society of America*, 99, 3694-3706.

FERGUSON, C. A. (1975). Universal tendencies and normal nasalization. In *Nasal Fest: Papers from a Symposium on Nasal and Nasalization*. California: Department of Linguistics, Stanford University, pp. 175-196.

FOWLER, C. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8, 113-133.

FOWLER, C. A., & BROWN, J. M. (2000). Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Percept. Psychophys.*, 62, 21–32.

FOWLER, C., RUBIN, P., REMEZ, R., & TURVEY, M. T. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production* (pp. 373-420). New York: Academic Press.

FUJIMURA, O. (1962). Analysis of Nasal Consonants. *Journal of the Acoustical Society of America*, 34(12), 1865-1875.

FUJIMURA, O., & ERICKSON, D. (1997). Acoustic phonetics. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 65-115). Oxford: Blackwell Publishers.

GAAB, M.R. (2013). INSTRUMENTATION: Endoscopes and equipment. *World Neurosurgery*, 79, RD4.e11-RD4.e21.

GAFOS, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory*, 20, 269-337.

GHIO, A., & TESTON, B. (2004). Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *Proceedings of the International Conference on Voice Physiology and Biomechanics*, 55–58.

GOLDSMITH, J. (1976). *Autosegmental Phonology*. Master Dissertation: MIT.

GOLDSTEIN, L. (2008). Emergence of discrete gestures. In *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, Espanha. Universitat Autònoma de Barcelona.

GOLDSTEIN, L., POUPLIER, M., CHEN, L., SALTZMAN, E., & BYRD, D. (2007). Dynamic action units slip in speech production “errors”. *Cognition*, 103, 386–412.

GRACCO, V. L., & NYE, P. W. (1993). Magnetometry in Speech Articulation Research: Some misadventures on the road to enlightenment. *Proceedings of the ACCOR Workshop on Electromagnetic Articulography in Phonetic Research*, Munchen, Germany, 91-104.

GREENE, M. C. L. (1964). *The Voice and its Disorders*. 2nd Edition. London: Pitman.

GREGIO, F. N. (2006). Configuração do trato vocal supraglótico na produção das vogais do português brasileiro: dados de imagens de ressonância magnética. Master Dissertation: Pontifícia Universidade Católica de São Paulo.



GROSS, S., & KOLLENBRANDT, M. (2009). Technical Evolution of Medical Endoscopy. *Acta Polytechnical: Journal of Advanced Engineering*, 49(2), 15–19.

GRUNERT, P., & OERTEL, J. (2011). Technical and Clinical Evolution of Modern Neuroendoscopy. In *Advances in Endoscopic Surgery*, Cornel Iancu (Ed.), InTech, 175-190.

GUYTON, A. C. (1988). *Fisiologia humana*. 6th Edition. Rio de Janeiro: Guanabara Koogan.

HAJEK, J. (1991). The hardening of nasalized glides in Bolognese. In P. M. Bertinetto, M. Kenstowicz, & R. Krakow (Eds.), *Nasals, Nasalization and the Velum* (pp. 225-249). San Diego: Academic Press.

HALLE, M. (1995). Feature Geometry and Feature Spreading. *Linguistic Inquiry*, 26.

HARA, Y., Tobita, Y., & Tsunoda, H. (1967). Intra-gastric Photography: Gastrocamera with Fiberscope. *Gastroenterology*, 94(3), 337-343.

HARDCASTLE, W. J. (1976). *Physiology of Speech Production - An Introduction for Speech Scientists*. London: Academic Press.

HARRELL, A. G., & Heniford, B. T. (2005). Minimally invasive abdominal surgery: Lux et veritas past, present, and future. *American Journal of Surgery*, 190, 239–243.

HAWKINS, S. (1992). An introduction to task dynamics. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II. Gesture, segment, prosody* (pp. 9-25). Cambridge: Cambridge University Press.

HAWKINS, S., & Stevens, K. N. (1985). Acoustic and Perceptual Correlates of the Non-nasal-nasal Distinction for Vowels. *Journal of the Acoustical Society of America*, 77, 1560-1575.

HUBER, J., Stathopoulos, E., & Sussman, J. (2004). The control of aerodynamics, acoustics, and perceptual characteristics during speech production. *Journal of the Acoustical Society of America*, 166.

HUCHE, F. L., & Allali, A. (1999). *A voz: anatomia e fisiologia dos órgãos da voz e da fala* (2nd ed.). Porto Alegre: Artmed.

HUNT, E. H. (2005). *Acoustic Characterization of the Glides /j/ and /w/ in American English*. Unpublished master's thesis, Massachusetts Institute of Technology.

JACOB, S. W., Francone, C., & Lossow, W. J. (1990). *Anatomia e Fisiologia Humana* (5th ed.). Rio de Janeiro: Editora Guanabara.

JESUS, M. S. V. (2002). Estudo fonético da nasalidade vocálica. *Estudos Linguísticos: estudos em fonética e fonologia do português*, 205-224.

KEATING, P. (1985). Universal phonetics and the organization of grammars. In V. Fromkin (Ed.), *Phonetic Linguistics* (pp. 115-132). Academic Press.

KEATING, P. (1990). The window model of coarticulation: articulatory evidence. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I* (pp. 451-470). Cambridge University Press.

KEATING, P. A. (1990). The window model of coarticulation: articulatory evidence. In J. Kingston & M. Beckman (Eds.), *Laboratory Phonology I* (pp. 451-470). Cambridge University Press.

KELSEY, C. A., Woodhouse, R. J., & Minifie, F. D. (1969). Ultrasonic observations of coarticulation in the pharynx. *Journal of the Acoustical Society of America*, 46(4), 1016-1018.

KELSO, J., Saltzman, E., & Tuller, B. (1986). The dynamical perspective on speech production: data and theory. *Journal of Phonetics*, 14, 29-59.

KENT, R. D. (1997). *The Speech Sciences*. San Diego: Singular Publishing Group Inc.

KENT, R., & Read, C. (1992). *The acoustic analysis of speech*. San Diego: Singular Publishing.

KINGSTON, J., & Diehl, L. (1994). Phonetic knowledge. *Language*, 70, 419-454.

KOKAWA, T., SAIGUSA, H., AINO, I., MATSUOKA, C., NAKAMURA, T., TANUMA, K., YAMASHITA, K., & NIIMI, S. (2006). Physiological Studies of Retrusive Movements of the Human Tongue. *Journal of Voice*, 20(3), 414-422.

KRAKOW, R. A., BEDDOR, P. S., & Goldstein, L. M. (1988). Coarticulatory influences on the perceived height of nasal vowels. *Journal of the Acoustical Society of America*, 83(3), 1146-1158.

KRAKOW, R., & HUFFMAN, M. (1993). Instruments and techniques for investigating nasalization and velopharyngeal function in the laboratory: an introduction. In M. Huffman & R. Krakow (Eds.), *Nasals, nasalization, and the velum*. San Diego: Academic Press.

LACERDA, A., & HEAD, B. F. (1975). Análise de sons nasais e sons nasalizados do Português. *Revista do Laboratório de Fonética Experimental de Coimbra VI*, 5-71.

LADEFOGED, P. (1974). *An Elements of acoustic phonetics*. The University of Chicago Press.

LADEFOGED, P. (1975). *A course in phonetics*. Harcourt Brace Jovanovich.

LADEFOGED, P. (1982). *A Course in Phonetics*. San Diego: Harcourt.

LADEFOGED, P., & MADDISON, I. (1999). *The sounds of the world's languages*. Cambridge: Blackwell Publishers.

LEESON, T. S. (1970). Sistema respiratório. In T. S. Leeson (Ed.), *Atlas de Histologia* (pp. 337-360). São Paulo: Atheneu.

LIBERMAN, A. M., & MATTINGLY, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36.

LIBERMAN, A. M., & STUDDERT-KENNEDY, M. (1978). Phonetic perception. In R. Held, H. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology*, vol. VIII, 'Perception' (pp. 1-36). Heidelberg: Springer-Verlag.

LIEBERMAN, P., & BLUMSTEIN, S. E. (1988). *Speech Physiology, Speech Perception, and Acoustic Phonetics*. Cambridge: Cambridge University Press.

LINDBLOM, B. (1983). Economy of speech gestures. In M. MacNeilage (Ed.), *The Production of Speech* (pp. 217-246). New York: Springer-Verlag.

LINDBLOM, B. (1983). *The Production of Speech*. In M. MacNeilage (Ed.), *Economy of speech gestures* (pp. 217-246). New York: Springer-Verlag.

LINDBLOM, B. (1990). On the communicative process: speaker-listener interaction and the development of speech. *AAC Augmentative and Alternative Communication*, 6, 220-230.

LINDQVIST-GAUFFIN, J., & SUNDBERG, J. (1976). Acoustic properties of the nasal tract. *Phonetica*, 33, 161-168.

LIPSKI, J. (1975). Brazilian Portuguese vowel nasalization: secondary aspects. *Canadian Journal of Linguistics*, 20, 59–77.

LOFIEGO, J. L. (1992). *Fissura lábio-palatino*. Rio de Janeiro: Revinter.

LOSKEN, A., WILLIAMS, J. K., BURSTEIN, F. D., MALICK, D. N., & RISKI, J. E. (2006). Surgical correction of velopharyngeal insufficiency in children with velocardiofacial syndrome. *Plastic and Reconstructive Surgery*, 117(5), 1493-1498.

LOVATTO, L., AMELOT, A., CREVIER-BUCHMAN, L., BASSET, P., & VAISSIÈRE, J. (2007). A FIBERSCOPIC ANALYSIS OF NASAL VOWELS IN BRAZILIAN PORTUGUESE. *ICPhS XVI*. Saarbrücken, 6-10 August 2007.

LUBKER, J. F., & MOLL, K. L. (1965). Simultaneous oral-nasal airflow measurements and cinefluorographic observations during speech production. *Cleft Palate Journal*, 257–272.

MADDIESON, I. (1984). *Patterns of Sounds*. Cambridge University Press. Cambridge Studies in Speech Science and Communication.

MAEDA, S. (1982). The role of sinus cavities in the production of nasal vowels. *Proceedings IEEE Int. Conf. ASSP*, 2, 911-914.

MAIA, E. M. (1985). *No reino da fala - A linguagem e seus Sons* (3rd ed.). São Paulo: Ática.

MALECOT, A. (1960). "Vowel nasality as a distinctive feature in American English," *Language*, 36, 222–229.

MARCHESAN, I. Q., & JUNQUEIRA, P. (1998). Atipia ou adaptação: como considerar os problemas da deglutição. In P. Junqueira & B. Dauden (Eds.), *Aspectos atuais em terapia fonoaudiológica* (pp. 1-17). São Paulo,

MARSH, B. R. (1996). Historic Development of Bronchoesophagology. *Otolaryngology- Head and Neck Surgery*, 114, 689–716.

MARSH, J. L. (2004). The evaluation and management of velopharyngeal dysfunction. *Clin Plast Surg*, 31, 261-270.

MATEUS, M. H. M. (1975). *Aspectos da fonologia do português*. Lisboa: Instituto Nacional de Investigação Científica.

MATSUO, K., & PALMER, J. (2009). Coordination of mastication, swallowing and breathing. *Japanese Dental Science Review*, 45(1), 31-40.

MATTOSO CÂMARA JR., J. (1969). *Problemas de Lingüística Descritiva*. Petrópolis: Vozes.

MATTOSO CÂMARA JR., J. (1977). *Para o Estudo da Fonêmica Portuguesa* (2nd ed.). Rio de Janeiro: Padrão.

MAYO CLINIC. (n.d.). Vocal cord paralysis. *Mayo Clinic Family Health Book*, 5th Edition. Retrieved from <https://www.mayoclinic.org/diseases-conditions/vocal-cord-paralysis/symptoms-causes/syc-20378873>

MCCARTHY, J. (1988). Feature geometry and dependency: a review. *Phonetica*, 43, 84-108.

MCKERNS, D., & BZOCH, K. (1970). *Variations in velopharyngeal valving: The factor of sex*. Gainesville, Florida.

MEDEIROS, B. R. (2007). Vogais nasais do português brasileiro: reflexões preliminares de uma revista. *Revista Letras* 72 (UFPR), 165-188.

MEDEIROS, B. R. de, & DEMOLIN, D. (2006). Vogais nasais do português brasileiro: um estudo de IRM. *Revista da ABRALIN*, 5(1/2). Retrieved from <https://revista.abralin.org/index.php/abralin/article/view/943>

MEDEIROS, B., D'IMPERIO, and R. ESPESSER. (2008). La voyelle nasale en portugais brésilien et son appendice nasal: étude acoustique et aérodynamique. *Journées d'Etude sur la Parole*.

MEDEIROS, Beatriz Raposo de. (2007). Nasalidade em PB e a questão da coda: vogal ou consoante? *Anais. Belo Horizonte: Faculdade de Letras da UFMG*.

MEILLET, A. (1896). *Varia. Mémoires de la société de linguistique de Paris*, 9, 136–159.

MEYERHOFF, W. L., & SCHAEFER, S. D. (1991). Physiology of the nose and paranasal sinuses. In M. Paparella & D. A. Shumrick (Eds.), *Otolaryngology. Vol I - Basic Sciences and Related Principles* (pp. 315-332). Philadelphia: W. B. Saunders Company.

MOLL, K. L. (1962). Velopharyngeal closure on vowels. *Journal of Speech and Hearing Research*, 17, 30–77.

MOON, J., & FOLKINS, J. (1991). Coordination of velopharyngeal muscle activity during positioning of the soft palate. *Cleft Palate Craniofac Journal*, 101, 1332-1335.

MOON, J., KUEHN, D., & HUISMAN, J. (1994). Measurement of velopharyngeal closure...

MORAES, J. A., & WETZELS, W. L. (1992). Sobre a duração dos segmentos vocálicos nasais e nasalizados em Português. Um exercício de Fonologia Experimental. *Cadernos de Estudos Lingüísticos*, 23, 153-166.

MORAIS-BARBOSA, J. (1962). Les voyelles nasales portugaises: interprétation phonologique. In A. Sovijarvi & P. Aalto (Eds.), *Proceedings of the IVth International Congress of Phonetic Sciences* (pp. 1-28). The Hauge: Mountain.

MORGENSTERN, L. (2005). The 200th Anniversary of the First Endoscope: Philipp Bozzini (1773-1809). *Surgical Innovation*, 12(2), 105–106.

NAM, H., & GOLDSTEIN, L., & SALTZMANN, E. (2009). Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marisco, & I. Chitoran (Eds.), *Approaches to Phonological Complexity* (pp. 1-28). Mouton de Gruyter.

NAM, H., & SALTZMAN, E. (2003). A Competitive, Coupled Oscillator Model of Syllable Structure. Paper presented at the Proceedings of the XIIth International Congress of Phonetic Sciences, Barcelona, Spain.

NAM H., GOLDSTEIN L., SALTZMAN E. (2010). "Self-organization of syllable structure: a coupled oscillator model," in *Approaches to Phonological Complexity*, eds Pellegrino F., Marisco E., Chitoran I., (New York, NY: Mouton de Gruyter.

NETTER, F. H. (2000). *Atlas de anatomia humana* (3rd ed.). Porto: Artmed.

NIIMI, S., BELL-BERTI, F., & HARRIS, K. S. (1982). Dynamic aspect of velopharyngeal.

NUNES, D. X. D. (2018). *A História do Endoscópio*. Clínica Universitária de Otorrinolaringologia. Master Dissertation.

ODELL, J. S. (2002). *International Negotiation, Creating Data on International Negotiation Strategies*. Springer, 39.

OHALA, J. (1981). Speech timing as a tool in phonology. *Phonetica*, 38, 204-217.

OHALA, J. (1983). The origin of sound patterns in vocal tract constraints. In P. MacNeilage & D. A. Daniloff (Eds.), *The production of speech* (pp. 189-216). New York: Springer-Verlag.

OHALA, J. (1986). Against the direct realist view of speech perception. *Journal of Phonetics*, 14, 75-82.

PARADIS, C., & PRUNET, J-F. (2000). Nasal Vowels as Two Segments: Evidence from Borrowings. *Language*, 76(2), 324-357.

PARKINSON, S. (1983). Portuguese nasal vowels as phonological diphthongs. *Lingua*, 61, 157-177.

PEGORARO-KROOK, M. I., DUTKA-SOUZA, J. C. R., & MARINO, V. C. C. (2008). Nasoendoscopy of velopharynx before and during diagnostic therapy. *Journal of Applied Oral Science*, 16.

PICKETT, J. M. (1991). Consonants: nasal, stops and fricative manners of articulation. In R. J. Baken & R. G. Daniloff (Eds.), *Readings in clinical spectrography of speech* (pp. 194-202). San Diego: Singular Publishing Group.

PIERREHUMBERT, J. (1980). The phonology and phonetics of English intonation. Ph.D. Dissertation (Unpublished), MIT.

PIGGOTT, G. L. (1987). On the autonomy of the feature nasal. *CLS* 23, 223-238.

PLAST SURG. (2004). Vol 31(2), 261-269.

PORT, R. F., & VAN GELDER, T. (Eds.). (1995). *Mind as motion: explorations in the dynamics of cognition*. Cambridge, Massachusetts: The MIT Press.

PRASAD, R., & YEGNARAYANA, B. (2020). A study of vowel nasalization using instantaneous spectra. Preprint submitted to *Computer Speech and Language*.

QUILIS, A. (1981). *Fonética acústica de la lengua española*. Madrid: Gredos.

REED, D., & LEITE, Y. (1947). The segmental phonemes of Brazilian Portuguese: Standard Paulista dialect. In K. Pike (Ed.), *Phonemics* (pp. 194-202). Ann Arbor: University of Michigan.

REPP, B. H. (1982). Phonetic trading relations and context effects: new experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92, 81-110.

RESNICK, R., HALLIDAY, D., & WLAKER, J. (2009). *Fundamentals of Physics*. Vol. 2. 8th Ed. Rio de Janeiro: LTC.



ROMAN, J., FANT, G., & HALLE, M. (1952). Preliminaries to Speech Analysis: The distinctive features and their correlates. Technical Report 13, Massachusetts: Acoustics Laboratory: MIT Press.

ROSSATO, S., BADIN, P., & BOUAOUNI, F. (2003). Velar movements in French: an articulatory and acoustical analysis of coarticulation. In ICPHS-15, 3141-3144.

ROUSSELOT, L'Abbé. (1897-1901). Principes de phonétique.

SAVARIAUX, C. (2014). Trap: Logiciel de traitement des signaux de parole [Manual]. Version 6. Département Parole et Cognition, GIPSA-lab, Grenoble-Alpes University – France.

SCHWAB, K., & SINGH, S. (2010). An introduction to flexible endoscopy. *Surgery*, 29(2), 80–84.

SEARA, C. (2000). Estudo acústico-perceptual da nasalidade das vogais do português brasileiro. Ph.D. Dissertation. Florianópolis: Universidade Federal de Santa Catarina.

SOLÉ, M.-J. (1995). Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech*, 38, 1-23.

SONGARA, R. K., SHARMA, G. N., GUPTA, V. K., & GUPTA, P. (2010). Need for harmonization of labeling of medical devices: a review. *Journal of Advanced Pharmaceutical Technology and Research*, 1(2), 127-144.

SOUSA, E. M. *Para a caracterização fonético-acústica da nasalidade no Português do Brasil*. Master Dissertation: Universidade Estadual de Campinas, 1994.

SOUZA, M. C. Q. (2003). Características Espectrais da Nasalidade. Master Dissertation. Escola de Engenharia de São Carlos.

SPENCE, A. P. (1991). *Anatomia humana básica* (2nd ed.). São Paulo: Manole Ltda.

STEVENS, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3-46.

STEVENS, K. N. (1998). *Acoustic Phonetics*. Cambridge: MIT Press.

STEVENS, K. N. (1999). Articulatory-acoustic-auditory relationships. In W. Hardcastle & J. Laver (Eds.), *The Handbook of Phonetic Science* (pp. 462-506). Oxford: Blackwell.

STEVENS, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustic Society of America*, 111, 1872-1891.

STEVENS, K. N. (2005). Features in Speech Perception and Lexical Access. In D. E. Pisoni & R. E. Remez (Eds.), *Handbook of Speech Perception* (pp. 125-155). Cambridge, MA: Blackwell.

STORY, B. H. (1995). Physiologically based speech simulation using an enhanced wave-reflection model of the vocal tract. Ph.D. Dissertation, University of Iowa.

STYGER, T., & KELLER, E. (1994). Formant synthesis. In E. Keller (Ed.), *Fundamentals of Speech Synthesis and Speech Recognition: Basic Concepts, State of the Art, and Future Challenges* (pp. 109-128). Chichester: John Wiley.

Technical Report:

TEIXEIRA, A., MOUTINHO, L. C., & COIMBRA, R. (2001). Acerca das vogais nasais do Português Europeu. *Revista da Universidade de Aveiro, Letras*, 18, 241-274.

TROUBETZKOY, N. S. (1949). *Principes de phonologie* (Translated by J. Cantineau). Paris.

TURK, A., & SHATTUCK-HUFNAGEL, S. (2020). *Speech Timing* (Oxford Studies in Phonology and Phonetics). OUP Oxford. Kindle edition.

TURNBULL, R., et al. (2008). Nasal place assimilation trades off inferrability of both target and trigger words. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 9(1), 1-27.

VAN LIESHOUT, P. H. H. M., & GOLDSTEIN, L. M. (2008). Articulatory phonology and speech impairment. In M. J. Ball, M. R. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 467–479). Malden: Blackwell Publishing.

VERGER-KUHNKE, B., REUTER, A., & BECCARIA, L. (2007). La biografía de Philipp Bozzini (1773-1809) un idealista de la endoscopia. *Actas Urológicas Españolas*, 31(5), 437–444.

VIEIRA, J. M. (2003). Perfil espectrográfico da hipernasalidade de fala de mulheres portadoras de fissura palatina. Master Dissertation in Bioengineering, Universidade de São Paulo, São Carlos.

VIEIRA, J. M. (2003). Perfil espectrográfico da hipernasalidade de fala de mulheres portadoras de fissura palatina. São Carlos: Bioengenharia, Universidade de São Paulo.

VOGEL, A. P., et al. (2009). Comparing Speech Aerodynamics in Speakers with Cleft Palate and Healthy Controls: Nasality Index and Multidimensional Voice Program. *The Cleft Palate-Craniofacial Journal*, 46(3), 311-320.

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

WOLF-HEIDEGGER. (2006). Atlas de anatomia humana. Edited by Petra Kopf-Maier. Translated by Hécio Werneck. Vol. 2. Rio de Janeiro: Guanabara Koogan.

XAVIER, M. F., & MATEUS, M. H. M. (Cosmos, Ed.). (s.d.). Dicionário de termos linguísticos - V. Lisboa: Cosmos.

# Dynamic modeling of the velopharyngeal trajectory in Brazilian Portuguese nasal diphthongs

## **Abstract**

The objective of this Ph.D. research is to understand the characteristics of nasal diphthong and front nasal vowel production in the Paulistano dialect of Brazilian Portuguese through experimental phonetic and phonological analysis. In this particular case, the nasalization of speech sounds involves a phonetic process known as nasal diphthongization. This allophonic process alters the quality of vowels and nasal diphthongs. To illustrate this phenomenon, we compare the pronunciation of the diphthong 'ão' [ãõ] with its oral equivalent 'au' [aw]. Similarly, we examine the production of the diphthongized nasal vowel /ẽ/, represented as [ẽ̃], and the production of the oral diphthong [ej]. The study comprises an aerodynamic analysis (pneumotachograph), a 2D electro-articulography (electromagnetic midsagittal articulography) study, and a nasal fiberscope study. A corpus of 20 words repeated three times by native speakers of the São Paulo dialect was created. The general results suggest that nasal diphthongs are a co-articulatory phenomenon involving gestural assimilation, leading to the emergence of unspecified nasal consonants called nasal glides. This phenomenon is also associated with the formation of nasal appendixes. The articulatory process alters the quality of resonance influenced by tongue movements associated with the opening and closing of the velum. As a result, the synchronization of tongue movement with soft palate movement causes nasal gliding, which propagates its resonance and creates nasal appendixes. Depending on the degree of co-articulation of the nasal vowels, either velar or palatal constriction may occur. Consequently, we conclude that nasal diphthongization modifies the geometry of the vocal tract, not only changing the shape of the vocal tract but also the position and type of tongue articulation. This complex articulation leads to the emergence of narrowed vowel segments that sound more consonant-like. Based on these findings, we can consider nasal diphthongization as a transitional phenomenon that plays a phonetic role in enhancing perception.

**KEYWORDS:** *Brazilian Portuguese/BP, Nasal Diphthongization, Experimental Phonetics and Phonology and Speech Timing.*

UNIVERSITE SORBONNE NOUVELLE

ED 622 – LANGAGE SCIENCES

ILPGA – Institut de Linguistique et Phonétique Générales et Appliquées.

8 Avenue Saint Mandé 75012 PARIS

PARIS

# **La modélisation dynamique de la trajectoire vélo-pharyngée dans les diphtongues nasales du Portugais Brésilien**

## **Résumé**

L'objectif de cette recherche doctorale est de comprendre, d'un point de vue expérimental, les caractéristiques de la production de la diphtongue nasale postérieure et de la voyelle antérieure nasale en portugais brésilien à l'aide d'analyses phonétiques et phonologiques. La nasalisation de la voix est affectée par un processus phonétique appelé diphtongaison nasale. Ce processus allophonique est un phénomène qui modifie la qualité des voyelles et des diphtongues nasales. Pour démontrer ce phénomène, nous comparerons la prononciation de la diphtongue « ão » [ãõ] avec celle du phonème équivalent oral « au » [aw]. De même, nous comparerons la production de la voyelle nasale diphtonguée / ẽ /, réalisée comme [ẽ̃], avec celle de la diphtongue orale [ej]. Nous avons créé un corpus de 20 mots répétés trois fois par des locuteurs natifs du dialecte de São Paulo. Afin de présenter les caractéristiques générales de la production des voyelles nasales et des diphtongues nasales ainsi que leurs mécanismes articulatoires, nous avons réalisé des études aérodynamiques, d'imagerie et articulatoires. Cette étude comprend une étude aérodynamique (pneumotachographe), une étude d'électro-articulographe 2D (articulographe électromagnétique sagittal médian) et une étude de fibroscope nasal. Les résultats généraux suggèrent que la diphtongaison nasale est un phénomène consonantique coarticulatoire d'assimilation gestuelle. Ceci est interprété comme une occurrence de consonnes nasales non spécifiées, appelées semi-voyelles/semi-consonnes nasales, résultant du contexte de l'émergence d'appendices nasaux. Ce processus altère la qualité de la résonance en raison de l'influence des mouvements de la langue associés à la descente et à la fermeture du voile du palais. Ainsi, la synchronisation des mouvements de la langue avec ceux du voile du palais se traduit par des semi-voyelles/semi-consonnes nasales, dont les résonances se propagent, générant l'appendice nasal et le processus de nasalisation. Selon le degré d'articulation des voyelles nasales, il peut y avoir une constriction vélaire ou palatine. De cela, nous concluons que la diphtongaison nasale modifie la géométrie du tube vocal, la position et le lieu d'articulation de la langue. Cette articulation complexe conduit à l'apparition de segments vocaliques. L'élévation des voyelles produit une transition vers une semi-voyelle/semi-consonne, avec l'émergence d'appendices nasaux homorganiques, rendant le segment sonore plus consonantique. À partir de ces résultats, nous pouvons considérer que la diphtongaison nasale est un phénomène transitionnel qui joue un rôle phonétique pour renforcer la perception.

***MOTS CLÉS : Portugais Brésilien, Diphtongaison nasale, Phonétique et Phonologie Expérimentale et Synchronisation de la parole.***

UNIVERSITE SORBONNE NOUVELLE

ED 622 – LANGAGE SCIENCES

ILPGA – Institut de Linguistique et Phonétique Générales et Appliquées.

8 Avenue Saint Mandé 75012 PARIS

# **A modelização dinâmica da trajetória velofaríngea nos ditongos nasais do português brasileiro**

## **Resumo**

O objetivo desta pesquisa de doutorado é o uso da análise Fonética e Fonológica para compreender as características de produção dos ditongos nasais e da vogal anterior nasal no português brasileiro, a partir de um viés experimental. No dialeto Paulistano, a nasalização vocálica passa por um processo fonológico denominado ditongação nasal. Esse processo alofônico é um fenômeno que pode alterar a qualidade das vogais e dos ditongos nasais. Para demonstrar esse fenômeno, compararemos a produção do ditongo “ão” [ãw̃] e sua contraparte oral “au” [aw]. Da mesma forma, compararemos a produção da vogal nasal ditongada / ẽ /, produzida como [ẽj̃], com o ditongo oral [ej]. Montamos um corpus com vinte palavras que foram repetidas três vezes, por falantes nativos do dialeto Paulistano. Para apresentarmos as características gerais da produção da vogal nasal e do ditongo nasal, bem como seus mecanismos articulatórios, estabelecemos um estudo aerodinâmico, um estudo de imagem e um estudo articulatório. Esta pesquisa contém um estudo aerodinâmico (pneumatógrafo), um estudo eletro-articulógrafo 2D (articulógrafo eletromagnético mediano sagital) e um estudo com um naso-fibrocópio. Os resultados gerais sugerem que a ditongação nasal é um fenômeno co-articulatório de assimilação gestual resultante do surgimento de uma consoante nasal não especificada, chamada de glide nasal, associada ao contexto e ao surgimento do apêndice nasal. Esse processo altera a qualidade das ressonâncias, devido à influência do movimento da língua associado ao abaixamento e fechamento do velum. Consequentemente, o movimento da língua e sua sincronização com o movimento do véu palatino resultam no surgimento da glide nasal que espalha suas ressonâncias gerando o apêndice nasal. Dependendo do grau de coarticulação da vogal nasal, uma constrição velar ou palatina pode surgir. Assim, concluímos que a ditongação nasal altera a geometria do trato, o ponto e o modo de articulação da língua. Essa articulação complexa resulta no surgimento de um segmento vocálico constricto e o alçamento vocálico produz a propagação da glide, bem como o surgimento do apêndice nasal homorgânico, tornando o som mais consonantal. A partir desses resultados, podemos considerar que a ditongação nasal é um fenômeno de transição que desempenha um papel fonético para intensificar a percepção.

***PALAVRAS-CHAVE: Português Brasileiro, Ditongação Nasal, Fonética e Fonologia Experimental e Tempo de fala.***

UNIVERSITE SORBONNE NOUVELLE

ED 622 – LANGAGE SCIENCES

ILPGA – Institut de Linguistique et Phonétique Générales et Appliquées.

8 Avenue Saint Mandé 75012 PARIS