



Universidade de Brasília (UnB)

Instituto de Letras (IL)

Departamento de Linguística, Português e Línguas Clássicas (LIP)

Programa de Pós-Graduação em Linguística (PPGL)

**FONÉTICA E FONOLOGIA DA LÍNGUA ARUTANI:
UMA DESCRIÇÃO ARTICULATÓRIA, ACÚSTICA E FONOLÓGICA DO IDIOLETO
DA ÚLTIMA FALANTE NO BRASIL**

Brasília

2025

GABRIEL FERREIRA BARROS

**FONÉTICA E FONOLOGIA DA LÍNGUA ARUTANI:
UMA DESCRIÇÃO ARTICULATÓRIA, ACÚSTICA E FONOLÓGICA DO IDIOLETO
DA ÚLTIMA FALANTE NO BRASIL**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística do Departamento de Linguística, Português e Línguas Clássicas da Universidade de Brasília, como parte dos requisitos necessários para a obtenção do grau de Mestre em Linguística.

Área de concentração: Teoria e Análise Linguística.

Orientador: Professor Doutor Thiago Costa Chacon

Brasília

2025

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

Ff Ferreira Barros, Gabriel
Fonética e fonologia da língua Arutani: uma descrição articulatória, acústica e fonológica do idioleto da última falante no Brasil / Gabriel Ferreira Barros; orientador Thiago Costa Chacon. Brasília, 2025.
241 p.

Dissertação (Mestrado em Linguística) Universidade de Brasília, 2025.

1. Fonética. 2. Fonologia. 3. Fonética acústica. 4. Línguas indígenas. 5. Salvaguarda. I. Costa Chacon, Thiago, orient. II. Título.

GABRIEL FERREIRA BARROS

**FONÉTICA E FONOLOGIA DA LÍNGUA ARUTANI:
UMA DESCRIÇÃO ARTICULATÓRIA, ACÚSTICA E FONOLÓGICA DO IDIOLETO
DA ÚLTIMA FALANTE NO BRASIL**

Banca examinadora

Professor Dr. Thiago Costa Chacon (Presidente)

Universidade de Brasília

Professor Dr. Ronaldo Manguiera Lima Júnior (Membro interno)

Universidade de Brasília

Professor Dr. Magnun Rochel Madruga (Membro externo)

Universidade Federal de Minas Gerais

Professora Dra. Flávia de Castro Alves (Membra suplente)

Universidade de Brasília

AGRADECIMENTOS

A Deus, pois sem Ele nada disso seria possível.

Ao meu orientador, professor Thiago Costa Chacon, que me orientou desde a graduação. Obrigado por toda a parceria nas pesquisas e nos trabalhos que desenvolvemos juntos. Os conselhos e orientações que venho recebendo do senhor extrapolam a vida acadêmica.

À minha família, meus pais e meu irmão, muito obrigado por todo apoio, desde de sempre, e por sempre estarem presentes. Pai e mãe, obrigado por sempre terem sido um exemplo para mim e por me mostrarem o poder da educação na mudança de vida e da realidade do mundo. Mateus, meu irmão, obrigado por sempre me ajudar em tudo o que precisei.

Ao povo Arutani, em especial à senhora Juanita Martinez. Agradeço, também, ao povo Xiriana (Ninam), em especial Albino Xiriana. Obrigado por serem sempre a constante lembrança da importância da valorização e da preservação das línguas e culturas indígenas.

Aos povos indígenas do Brasil que são exemplo de resistência, luta e resiliência.

À Universidade de Brasília, por ser uma universidade pública e de qualidade, que foi e continua sendo a minha segunda casa, a qual me abriu o caminho para o mundo da pesquisa e da educação.

Aos meus professores, tanto os professores do PPGL quanto os da época de graduação, sou grato a cada um de vocês pelos ensinamentos, pela aprendizagem e pela troca de experiências. Agradeço em especial: Ronaldo Lima Jr., Dionei Gomes, Marina Magalhães, Marcus Lunguinho, Antônio Augusto e a todos os professores do NTL Indígenas.

Aos meus amigos da Universidade de Brasília, que fiz desde a graduação. Em especial: Débson Ramos, Marina Félix, Mateus Motta, Diogo Koga, Pedro Araújo, Isabela Reis, Rodrigo do Prado e todos os meus colegas do NTL Indígenas.

RESUMO

A língua Arutani é uma língua isolada com apenas uma falante viva no Brasil. Neste trabalho, descrevemos e analisamos a fonética e a fonologia do Arutani falado por Juanita Martinez, com o objetivo de documentar e apoiar a salvaguarda dessa variedade da língua. Além do caráter científico e linguístico desta pesquisa, ela possui algumas motivações políticas muito fortes, como trazer evidência para a diversidade linguística no Brasil e para o risco de silenciamento e de extinção de línguas como o Arutani. Para a descrição e a análise do Arutani, fizemos o uso dos dados presentes no Dicionário Arutani (disponível em Chacon *et al.*, 2020). A partir do dicionário, utilizamos tanto os áudios que o compõem quanto registros escritos da língua, bem como algumas transcrições fonéticas. Utilizamos muitos recursos da fonética acústica, tanto práticos quanto teóricos, com o intuito de fazermos uma descrição o mais precisa e embasada possível dos fones existentes na língua. Analisamos e descrevemos as vogais orais (tônicas e átonas), as vogais nasais e nasalizadas, *glides* (por serem foneticamente mais próximos às vogais) e hiatos existentes na língua. Além disso, foi possível apresentarmos como o acento se manifesta a partir de uma perspectiva acústica na fala de Juanita. Após isso, realizamos um estudo sobre os segmentos consonantais. Com isso, analisamos e descrevemos os segmentos oclusivos, oclusivos com articulação secundária, nasais, fricativos e um único caso de tepe. A partir do inventário fonético completo, partimos para a análise fonológica e fonêmica. Para isso, agrupamos os fones em vogais e consoantes; e fizemos o contraste entre sons foneticamente semelhantes (SFS) em ambientes idênticos e análogos para a identificação de fonemas e alofones. Em seguida, definimos, com base no sistema da língua, o *status* fonológico dos *glides*; apresentamos como se dá a nasalidade vocálica; e descrevemos o acento de uma perspectiva métrica. Por fim, analisamos as estruturas silábicas, a fonotática do Arutani e fenômenos que surgem nestes contextos, como consoantes heterossilábicas e a ambissilabidade de alguns fonemas.

Palavras-chave: Arutani; fonética; fonologia; Juanita Martinez; salvaguarda; diversidade linguística; fonética acústica; fonêmica.

ABSTRACT

Arutani is an isolate language with only one living speaker in Brazil. In this study, we describe and analyze the phonetics and phonology of the Arutani spoken by Juanita Martinez, with the aim of documenting and safeguarding this variety of the language. Beyond its scientific and linguistic contribution, this research is also motivated by strong political concerns, such as raising awareness about Brazil's linguistic diversity and the risk of silencing and extinction faced by languages like Arutani. For the description and analysis, we used data from the Arutani Dictionary (available in Chacon et al., 2020). From this dictionary, we relied on both its audio recordings and written records of the language, as well as some phonetic transcriptions. We employed a wide range of tools from acoustic phonetics, both practical and theoretical, in order to provide the most accurate and well-grounded description possible of the language's phones. We analyzed and described the oral vowels (both stressed and unstressed), nasal and nasalized vowels, glides (due to their phonetic similarity to vowels), and vowel hiatuses in the language. In addition, we were able to present how stress manifests acoustically in Juanita's speech. Following that, we conducted a study of the consonantal segments, analyzing and describing the plosives, plosives with secondary articulation, nasals, fricatives, and a single case of a tap. Based on the complete phonetic inventory, we then proceeded to the phonological and phonemic analysis. To do this, we grouped the phones into vowels and consonants and performed contrasts between phonetically similar sounds (PSS) in identical and analogous environments to identify phonemes and allophones. Subsequently, we defined the phonological status of glides within the system of the language, presented the characteristics of vowel nasality, and described stress from a metrical perspective. Finally, we analyzed syllable structure, Arutani's phonotactics, and phenomena that arise in these contexts, such as heterosyllabic consonants and the ambisyllabicity of certain phonemes.

Keywords: Arutani; phonetics; phonology; Juanita Martinez; safeguarding; linguistic diversity; acoustic phonetics; phonemics.

LISTA DE FIGURAS

Figura 1 - Mapa dos rios Alto Paragua e Alto Uraricaá e atuais comunidades indígenas locais (tradução nossa).....	16
Figura 2 – Na foto à esquerda temos Juanita Martinez e na foto à direita temos Albino Xiriana (de azul), Thiago Chacon (de amarelo) e Juanita.....	18
Figura 3 – Dicionário Arutani.....	18
Figura 4 – Um modelo simples de produção de vogais: tubo reto de seção transversal uniforme fechado em um lado (por uma membrana vibradora que simula as pregas vocais) e aberto no outro (correspondendo à abertura bucal).....	51
Figura 5 – Segmentação das sessões de gravação no ELAN.....	59
Figura 6.1 – TextGrid gerado pelo ELAN exportado no Praat.....	60
Figura 6.2 – TextGrid gerado pelo ELAN exportado no Praat ampliado.....	60
Figura 7 – Arquivo “.txt” gerado pelo ELAN aberto no Audacity.....	61
Figura 8 – Segmentação da sentença no Praat.....	62
Figura 9 – Segmentação da palavra koxipan ‘caranguejo’ no Praat.....	63
Figura 10 – Tabela extraída do Dicionário Arutani e com as transcrições feitas com base nas análises acústicas.....	64
Figura 11 – Quadro de dispersão inicial das vogais tônicas orais (Script 1).....	71
Figura 12 – Quadro de dispersão inicial das vogais tônicas orais (Script 2).....	71
Figura 13 – Quadro de dispersão inicial das vogais tônicas orais com marcações para reanálise (Script 1).....	72
Figura 14 – Quadro de dispersão inicial das vogais tônicas orais com marcações para reanálise (Script 2).....	73
Figura 15 – Histogramas de F1 e F2 de [a].....	76
Figura 16 – Histogramas de F1 e F2 de [i].....	76
Figura 17 – Gráfico de dispersão inicial das vogais orais tônicas.....	77
Figura 18 – Boxplot dos valores de F1 das vogais orais tônicas.....	78
Figura 19 – Boxplot dos valores de F2 das vogais orais tônicas.....	78
Figura 20 – Histograma de F1 e F2 de [ɐ].....	79
Figura 21 – Gráfico de dispersão final das vogais orais tônicas.....	81
Figura 22 – Quadro de dispersão final das vogais tônicas orais.....	82
Figura 23 – Imagem acústica de kirixa.....	83
Figura 24 – Imagem acústica de kayari.....	84
Figura 25 – Imagem acústica de karo.....	85
Figura 26 – Imagem acústica de kuyuru.....	86
Figura 27 – Quadro de dispersão das vogais átonas orais.....	88
Figura 28 – Imagem acústica de kurata.....	89

Figura 29 – Imagem acústica de sambrero.....	90
Figura 30 – Imagem acústica de arihan.....	91
Figura 31 – Imagem acústica de poromã.....	92
Figura 32 – Imagem acústica de pururu ata.....	93
Figura 33 – Gráfico comparativo de F1 e F2 das vogais tônicas orais e átonas orais.....	94
Figura 34 – Imagem acústica de wasi sã.....	100
Figura 35 – Imagem acústica de xipirin ti.....	103
Figura 36 – Imagem acústica de kari.....	103
Figura 37 – [ĩ] vs. [i].....	104
Figura 38 – Imagem acústica de ariran.....	105
Figura 39 – Imagem acústica de katara.....	105
Figura 40 – [ẽ] vs. [a].....	106
Figura 41 – Imagem acústica de mo.....	107
Figura 42 – Imagem acústica de mamo.....	107
Figura 43 – [õ] vs. [u].....	108
Figura 44 – Imagem acústica de apakono.....	109
Figura 45 – Imagem acústica de ako.....	109
Figura 46 – [õ] vs. [o].....	110
Figura 47 – Imagem acústica de xipuya.....	112
Figura 48 – Imagem acústica de xaray.....	113
Figura 49 – Imagem acústica de [mʊ. 'ja].....	114
Figura 50 – Imagem acústica de [moj. 'ja].....	115
Figura 51 – Tentativa de segmentação de [j].....	115
Figura 52 – Imagem acústica de oi.....	117
Figura 53 – Imagem acústica de sui.....	117
Figura 54 – Imagem acústica de aay.....	118
Figura 55 – Imagem acústica de toropia.....	118
Figura 56 – Imagem acústica de imaa.....	119
Figura 57 – Imagem acústica de oo.....	120
Figura 58 – Imagem acústica de kariya.....	120
Figura 59 – Imagem acústica de kipa.....	124
Figura 60 – Imagem acústica de pasa.....	125
Figura 61 – VOT de [p] em pasa.....	126
Figura 62 – Imagem acústica de kaptiaw.....	126
Figura 63 – Traços acústicos de [p] em kaptiaw.....	127
Figura 64 – Imagem acústica de kapti.....	128

Figura 65 – Segmentação entre [p̄] e [t].....	128
Figura 66 – Imagem acústica de [v. 'tap].....	129
Figura 67 – Imagem acústica de [v. 'tap̄].....	130
Figura 68 – Imagem acústica de [v. 'tab̄].....	131
Figura 69 – Imagem acústica de koa muya.....	132
Figura 70 – Análise acústica de [b].....	132
Figura 71 – Imagem acústica de koa.....	133
Figura 72 – Imagem acústica de muya.....	133
Figura 73 – Imagem acústica de mata.....	134
Figura 74 – Imagem acústica de toari.....	135
Figura 75 – Imagem acústica de matkia.....	136
Figura 76 – Imagem acústica de xiraka.....	137
Figura 77 – Medição do VOT de [k].....	137
Figura 78 – Imagem acústica de kokap.....	138
Figura 79 – Análise espectral do burst das consoantes oclusivas.....	142
Figura 80 – Imagem acústica de akwa.....	146
Figura 81 – Imagem acústica de okya.....	146
Figura 82 – Imagem acústica de arako.....	147
Figura 83 – Imagem acústica de iki.....	148
Figura 84 – Imagem acústica de kokama.....	150
Figura 85 – Espectro de [m].....	152
Figura 86 – [m] vs. [v].....	153
Figura 87 – Formato de onda de [m].....	154
Figura 88 – Imagem acústica de koratana.....	154
Figura 89 – Espectro de [n].....	155
Figura 90 – [n] vs. [v].....	156
Figura 91 – Formato de onda de [n].....	157
Figura 92 – [m] vs. [n].....	158
Figura 93 – Imagem acústica de tamkaña.....	159
Figura 94 – Formato de onda de [ɲ].....	159
Figura 95 – Espectro de [ɲ].....	160
Figura 96 – [ɲ] vs. [ẽ].....	161
Figura 97 – Imagem acústica de kosa.....	163
Figura 98 – Imagem acústica de oxa.....	164
Figura 99 – Espectro de [s].....	165
Figura 100 – Espectro de [ʃ].....	166

Figura 101 – [s] vs. [ʃ].....	167
Figura 102 – Imagem acústica de pasa.....	168
Figura 103 – Imagem acústica de pixa.....	168
Figura 104 – Imagem acústica de awxihi.....	169
Figura 105 – Imagem acústica de awhi.....	170
Figura 106 – Espectro de [h].....	171
Figura 107 – Imagem acústica de xapi amaxãn.....	172
Figura 108 – [ʃ] vs. *[ʂ].....	173
Figura 109 – Imagem acústica de yapururu.....	175
Figura 110 – Comparação entre duas ocorrências de [r].....	176
Figura 111 – Zoom na comparação entre duas ocorrências de [r].....	177
Figura 112 – Imagem acústica de xiraka.....	177
Figura 113 – Resumo da estrutura silábica do Arutani.....	212

LISTA DE TABELAS

Tabela 1 – Comparação entre a fala de Juanita e o “Uruak” registrado em Migliazza (1978).....	36
Tabela 2 – Quadro fonético das consoantes.....	67
Tabela 3 – Quadro fonético das vogais orais.....	68
Tabela 4 – Quadro fonético das vogais nasais e nasalizadas.....	68
Tabela 5 – Quadro fonético dos ditongos crescentes e decrescentes encontrados.....	68
Tabela 6 – Layout e trecho da tabela produto do Script 1.....	70
Tabela 7 – Comparação dos valores de formantes extraído pela análise LPC e pela análise FFT	74
Tabela 8 – Médias dos valores de duração, pitch, intensidade, F1, F2 e F3 das vogais tônicas orais.....	80
Tabela 9 – Médias dos valores de duração, pitch, intensidade, F1, F2 e F3 das vogais átonas orais	87
Tabela 10 – Contraste entre os valores de F1, F2 e F3 das vogais tônicas orais e das vogais átonas orais.....	94
Tabela 11 – Comparação duração, pitch e intensidade entre as vogais tônicas orais e suas contrapartes átonas orais.....	95
Tabela 12 – Normalização dos valores de duração, pitch e intensidade das vogais orais por meio de z-score.....	97
Tabela 13 – Análise comparativa das diferenças de duração, pitch e intensidade após normalização.....	97
Tabela 14 – Média das diferenças dos valores de duração, pitch e intensidade após normalização..	98
Tabela 15 – Dados acústicos das vogais tônicas nasalizadas.....	101
Tabela 16 – Dados acústicos das vogais átonas nasalizadas.....	101
Tabela 17 – Quantitativo das vogais orais e suas contrapartes nasais e nasalizadas.....	102
Tabela 18 – Comparação dos valores de F1, F2 e F3 entre vogais orais e nasais.....	102
Tabela 19 – Ditongos crescentes e decrescentes.....	111
Tabela 20 – Valores de F1 (inicial vs. final) e F2 (inicial vs. final) na transição formântica.....	114
Tabela 21 – Hiatos existentes na língua.....	116
Tabela 22 – Quantitativo de consoantes oclusivas.....	122
Tabela 23 – Valores de VOT das consoantes oclusivas.....	139
Tabela 24 – Valores de F1 e F2 de transição das consoantes oclusivas.....	140
Tabela 25 – Análise dos espectros dos bursts.....	143
Tabela 26 – Quantitativos das consoantes oclusivas velares com articulação secundária.....	144
Tabela 27 – Valores de F1 e F2 de transição em [k ^w] e [k ^j].....	146
Tabela 28 – Comparação da duração entre [w] vs. [u] e [j] vs. [i].....	148
Tabela 29 – Quantitativos das consoantes nasais.....	149

Tabela 30 – Comparação entre os valores dos formantes nasais de [m] e [n].....	157
Tabela 31 – Quantitativo de consoantes fricativas.....	163
Tabela 32 – Valores de F1 e F2 de transição de [s] e [ʃ].....	165
Tabela 33 – Comparação dos dados espectrais de [ʃ] vs. *[ʂ].....	173
Tabela 34 – Quando fonológico das consoantes.....	180
Tabela 35 – Quando fonológico das vogais.....	180
Tabela 36 – Comparação de [e] vs. /i/.....	197

SUMÁRIO

1. Introdução.....	16
1.1 Objetivos.....	19
1.2 Povo, cultura e língua Arutani.....	20
1.2.1 Biografia e perfil sociolinguístico da última falante de Arutani no Brasil.....	24
1.2.2 Bibliografia sobre a língua Arutani.....	29
1.2.3 A língua Arutani: uma língua isolada?.....	33
1.3 Contexto e realidade da pesquisa desenvolvida.....	40
1.4 Justificativa.....	42
2. Referencial teórico e metodológico.....	49
2.1 Contribuições da fonética articulatória, da fonética acústica e da fonologia.....	49
2.2 Origem dos dados e elaboração do corpus.....	56
2.3 Processamento e preparação dos dados.....	58
2.4 Descrição e análise dos dados.....	64
3. Análise Acústica.....	67
3.1 Segmentos Vocálicos.....	69
3.1.1 Tratamento dos Dados.....	69
3.1.2 Vogais Tônicas Orais.....	80
3.1.3 Vogais Átonas Orais.....	86
3.1.4 Diferenças Acústicas entre Vogais Tônicas e Átonas Orais.....	93
3.1.5 Vogais Nasais e Nasalizadas.....	98
3.2 Glides.....	110
3.3 Hiatos e estratégias para a demarcação de fronteira silábica.....	116
3.4 Segmentos Consonantais.....	121
3.4.1 Consoantes Oclusivas.....	121
3.4.2 Consoantes Oclusivas Velares com Articulação Secundária.....	143
3.4.3 Consoantes Nasais.....	149
3.4.4 Consoante Fricativas.....	162
3.4.5 Tepe.....	174
4. Análise Fonológica.....	179
4.1 Consoantes.....	181
4.1.1 Oclusivas não-vozeadas /p/, /t/, /k/.....	181
4.1.2 Oclusiva glotal [ʔ].....	182
4.1.3 Oclusivas velares com articulação secundária: /kʷ/ e /kʲ/.....	182
4.1.4 Nasais: /m/, /n/, /ɲ/.....	184

4.1.5 [b] como alofone de /m/ e /p/.....	185
4.1.6 Fricativas: /s/ vs. /ʃ/ vs. /h/.....	186
4.1.7 [s] e [ʃ] em variação livre:.....	187
4.1.8 Palatais: /j/ e /ɲ/.....	187
4.1.9 Glides: /w/ e /j/.....	188
4.1.10 Contraste entre /t/ vs. /s/ vs. /ʃ/.....	188
4.1.11 Outros contrastes:.....	189
4.2 Vogais.....	191
4.2.1 Contraste das vogais em ambiente tônico:.....	191
4.2.2 Contraste das vogais em ambiente átono:.....	195
4.2.2.1 Sobre o estatuto fonêmico de [e] e /i/:.....	197
4.2.3 Contraste das vogais em ambiente tônico vs. átono.....	198
4.3 Segmentos ambivalentes.....	199
4.4 Nasalidade Vocálica.....	201
4.5 Sílabas e Fonotática.....	205
4.5.1 Ataque e coda.....	206
4.5.2 Núcleo silábico.....	210
4.5.3 Sequências Consonantais Heterossilábicas.....	210
4.5.4 Ambissilabidade.....	211
4.5.5 Resumo da estrutura silábica.....	211
4.6 Acento.....	212
5. Considerações finais.....	215
6. Referências.....	217
Anexo I.....	222
Anexo II.....	227
Anexo III.....	228
Anexo IV.....	236
Anexo V.....	238
Anexo VI.....	241

1. Introdução

O Arutani, conhecido também como Uruak e Awaké¹, é uma língua isolada – ou seja, não faz parte de nenhuma família linguística conhecida (CAMPBELL, 2010, p. 16) – e foi até pouco tempo tida como uma língua extinta (cf. Labrada *et al.*, 2020; Coppens, 2008). Segundo Migliazza (1978, p. 135) e Chacon *et al.* (2020), os Arutani viveram, desde os primeiros relatos, em comunidades ao longo do Alto Paragua (Venezuela) e bacias do rio Uraricaá (Terra Indígena [TI] Yanomami, Brasil) na fronteira Venezuela-Brasil. No seguinte mapa, pode-se observar as comunidades Ninam, da etnia Xiriana, nas quais também vivem falantes de Arutani: No Brasil temos a comunidade Saúba, na Venezuela temos as comunidades Kosoiba, Kavaimakén e Colibri.

Figura 1 - Mapa dos rios Alto Paragua e Alto Uraricaá e atuais comunidades indígenas locais (tradução nossa)



Fonte: Labrada *et al.* (2020, p. 171)

¹ Em nossa pesquisa, desenvolvemos melhor a discussão sobre esses diferentes etnônimos e se eles se referem, realmente, a um mesmo povo e a uma mesma língua. Essa problemática será apresentada e analisada com base no recente trabalho de Chacon e Labrada (2023).

Com relação à vitalidade da língua e cultura desse povo, Coppens (2008) ressalta o fato de que, para uma avaliação mais substancial a respeito da real situação dos Arutani, seriam necessárias novas e exaustivas pesquisas. Tendo em mente essa necessidade, foi exatamente o que foi feito pelo Dr. Thiago Chacon, juntamente com outros pesquisadores e apoiadores, no *Projeto de Salvaguarda da Língua e Cultura Arutani*. Durante a realização de uma pesquisa em 2011 no rio Ericó (afluente do Uraricaá, TI Yanomami), junto aos Ninam, esse grupo de pesquisadores recebeu relatos de que haveria a possibilidade da existência de indivíduos Arutani vivendo juntamente a outro grupo Ninam, tendo, no caso, possíveis falantes de Arutani do lado brasileiro da fronteira. Enquanto no Brasil a etnia Arutani aparecia como extinta nos últimos censos, segundo o Censo de 2011 do *Instituto Nacional de Estadística* da Venezuela, existiam cerca de 20 Arutani vivendo em território venezuelano (INE, 2015). Logo, já se sabia da existência de pessoas que se identificavam como Arutani, mas não quantos ainda falavam a língua, e se falavam, quais seriam seus graus de fluência. Entretanto, sem nenhum relato oficial da existência de indivíduos Arutani vivendo no Brasil.

Após consultas, trocas de informações e muitas buscas, os pesquisadores encontraram Juanita Martinez, uma senhora que possui atualmente 81 anos, aproximadamente, e que é a última falante de Arutani no lado brasileiro da fronteira. Juanita nasceu na Venezuela na comunidade de Makô e é moradora atual da comunidade Saúba no Brasil², onde vive com os Ninam na TI Yanomami. Juanita não fala o português, mas é falante de Ninam e possui bastante proficiência em Arutani.

² Entendemos que essa divisão geográfica é uma mera convenção política de caráter colonizador, mas atualmente é assim que os trabalhos estão sendo divididos entre os pesquisadores envolvidos.

Figura 2 – Na foto à esquerda temos Juanita Martinez e na foto à direita temos Albino Xiriana (de azul), Thiago Chacon (de amarelo) e Juanita.



Juanita é a principal consultora desta pesquisa e é a principal colaboradora do projeto de salvaguarda do Arutani no que tange às pesquisas feitas no Brasil³. Para chegar aos dados atuais existentes sobre os Arutani, tanto no Brasil quanto na Venezuela, foram necessárias algumas pesquisas mais direcionadas a partir de 2017. Desde então estamos desenvolvendo um dicionário *on-line* do Arutani de domínio público.

Figura 3 – Dicionário Arutani



³ Quanto aos estudos sobre os (semi-)falantes na Venezuela, eles estão sendo desenvolvidos por Jorge Labrada e Francia Medina, pesquisadores já citados anteriormente e que fazem parte de pesquisas importantes como as em Chacon *et al.* (2020) e Labrada *et al.* (2020).

Esse dicionário pode ser encontrado em Chacon *et al.* (2020)⁴ e já dispõe de transcrições fonéticas simples e áudios de palavras e de frases. A partir disso, além do dicionário, algumas informações de caráter demográfico foram coletadas. Segundo Chacon *et al.* (2020), a etnia Arutani conta com uma população de 20 indivíduos, os quais 16 estão do lado venezuelano da fronteira e 4 no lado brasileiro, sendo que todos vivem entre pessoas do povo Ninam. Além de Juanita, existem outros cinco (semi-)falantes que vivem no Alto Paragua na Venezuela, sendo três na comunidade de Kosoiba e dois na comunidade Kavaimaken. Ademais, os autores ainda relatam a possibilidade da existência de outro falante vivendo em meio aos Ninam no Alto Paragua, na comunidade de Colibri.

O encontro com Juanita foi de grande relevância para a linguística, tendo em vista que existem poucos materiais sobre a língua (sejam escritos ou audiovisuais), assim como é listado em Labrada *et al.* (2020). Esse encontro contribuirá não apenas para esta pesquisa, mas também para a ciência em si, já que poderemos, graças aos avanços tecnológicos atuais, estudar e compreender melhor a língua, colaborando para uma maior diversidade de estudos sobre as línguas do mundo e tornando mais precisos estudos futuros de cunho comparativo. Levando todas essas questões em consideração e, principalmente, o fato da pouca descrição desenvolvida sobre a língua Arutani é que se propõe nesta pesquisa a descrição linguística do Arutani, tendo como foco a fonética e a fonologia da língua. Nossa análise e descrição será com base na falante do lado brasileiro, ou seja, no idioleto de Juanita.

1.1 Objetivos

O trabalho desenvolvido nesta dissertação tem como propósito base a descrição e a análise linguística do Arutani, a partir do idioleto de Juanita, tendo um enfoque sobre a fonética e a fonologia da língua, contribuindo para sua salvaguarda, e, também, para o aprimoramento do Dicionário Arutani com as nossas transcrições. Pretendemos, também, avançar nas pesquisas de descrição da língua do projeto de *Descrição, Tipologia e Teorias Linguísticas* do Programa de

⁴ Disponível em: <<https://japiim.museudoindio.gov.br/dic/arutani/>>.

Pós-graduação em Linguística (PPGL) da Universidade de Brasília (UnB) e do *Núcleo de Tipologia e Línguas Indígenas* (NTLIndígenas).

A análise fonética terá como objetivo principal a descrição a partir das características acústicas de cada um dos fones encontrados em nossas transcrições fonéticas, como fones consonantais e vocálicos. Além dos segmentos, analisaremos os suprasegmentos (elementos prosódicos) do Arutani, levando em consideração duração, *pitch* e intensidade. Isso possibilitará uma análise em que se terá: (1) a identificação de todos os fones da língua; (2) sua classificação articulatória; e (3) a análise de propriedades acústicas que evidenciam cada classe de som. Ademais, será possível avaliar, numa perspectiva acústica, quais desses elementos suprasegmentais (prosódicos) são mais relevantes para a marcação do acento na fala de Juanita.

Após a análise fonética completa, o foco será na parte fonológica da língua. Para isso, almeja-se desenvolver uma análise fonêmica para a identificação e descrição dos fonemas e dos processos fonológicos encontrados na língua. Além disso, pretendemos explicar o *status* fonêmico dos segmentos ambivalentes na língua; o processo de nasalização vocálica; os padrões silábicos possíveis e questões fonotáticas; e como se dá a estruturação do acento de uma perspectiva fonológica.

Em Chacon *et al.* (2020) propomos, preliminarmente, cinco vogais grafêmicas <i>, <e>, <a>, <o> e <u>; e 12 consoantes grafêmicas <p>, <t>, <k>, <s>, <x>, <h>, <r>, <m>, <n>, <w>, <y>, <ñ> baseadas na fala de Juanita. Estes grafemas foram sugeridos como uma forma prática de escrita para apoiar o trabalho de registo e preservação da língua, permitindo o registro de uma quantidade significativa de dados antes que fosse possível realizar análises fonético-fonológicas mais detalhadas. Sendo assim, a partir dos resultados encontrados, pretendemos fazer as alterações necessárias no dicionário Arutani, tendo como base as análises e descrições fonológicas desenvolvidas nesta pesquisa.

1.2 Povo, cultura e língua Arutani

Os trabalhos já realizados sobre os Arutani, ou que mencionam sua existência, evidenciam que ainda se sabe muito pouco sobre esse povo e sua língua, ressaltando ainda mais a

importância desta pesquisa na ampliação desses conhecimentos. Para os referenciais desta pesquisa, foi elaborada uma pesquisa bibliográfica, servindo como fonte de apoio de dados, tanto em questões linguísticas quanto etnográficas. Alguns trabalhos que tratam dessas questões sobre os Arutani, e que nos servem como base, são, por exemplo, Armellada e Matallana (1942), Migliazza (1967;1978), Loukotka (1968), Rodrigues (1992), Coppens (2008), Pisemskaya (2009, 2010), Jolkesky (2016), Labrada *et al.* (2018), Ferreira *et al.* (2019), Chacon *et al.* (2020), Labrada *et al.* (2020), Chacon (2019, 2020a, 2020b) e o mais recente trabalho de Chacon & Labrada (2023).

Ao falarmos de estudos sobre os Arutani, temos os clássicos trabalhos de Migliazza (1978) e de Coppens (2008)⁵. Neles são apresentados dados importantes sobre onde os Arutani vivem e informações populacionais. Segundo os autores, os Arutani viviam na região dos rios Paragua e Uraricaá na fronteira entre Brasil e Venezuela.

Quanto à questão populacional, Migliazza (1978) traz por volta de 15 pessoas, dos quais apenas 5 eram falantes, já Coppens (2008), no ano de 1983, relata uma população de 10 indivíduos vivendo na Venezuela. Entretanto, há uma nota de rodapé em que Coppens (2008, p. 752) cita um outro trabalho de Migliazza e afirma que os números populacionais da pesquisa de Migliazza coincidem com as que com as suas, pois Migliazza teria relatado a existência de 10 Arutani, – ou Uruak, como ele chama – vivendo na Venezuela no rio Paragua e 6 vivendo no Brasil no Alto Uraricaá. Além disso, são apresentados alguns etnônimos, que, segundo esses autores, referem-se à mesma etnia, são eles: Arutani, Urutani, Uruak e Awake, dependendo da região em que se localizavam.

Migliazza (1978) buscou fazer uma comparação linguística entre as línguas Makú, Sapé e Uruak, e trazer um breve panorama sobre onde os povos que falavam essas línguas viviam e suas relações étnicas. Ele apresenta uma lista de 100 palavras na língua Uruak, o que contribuiu para estudos comparativos de pesquisas que visam contrastes dialetais e até reconstruções. Já Coppens (2008) fez uma descrição mais etnográfica, apresentando questões mais culturais e de costumes do povo que ele chama de Uruak e das relações deles com outros povos.

⁵ Seu trabalho data de 1983, mas a edição a qual tivemos acesso é a segunda edição de 2008.

O povo Arutani possuía diversos costumes particulares, mas, também, eram influenciados por outros povos com quem eles tinham contato, assim como todos os povos já conhecidos que interagiam com outros povos. Apesar da escassa documentação sobre as relações culturais sobre eles, Coppens (2008) descreve que a organização social dos Arutani, com relação a parentescos e alianças matrimoniais, era bastante complexa e expõe que os Arutani realizavam matrimônios interétnicos. Para eles, os povos mais importantes, em grau de prioridade, para a realização de casamentos eram os Ninam (Yanam), em segundo grau os Sapé e depois os Pemón. Revelando-se um povo que também praticava a exogamia linguística na região da Amazônia.

O autor também nos informa sobre os ritos que marcavam tanto a puberdade feminina quanto a masculina. Para as mulheres o rito de iniciação ocorria quando elas menstruavam pela primeira vez e tinham que ficar isoladas entre três a quatro dias e tinham seu cabelo cortado. Os homens precisavam passar por testes de força, como serem picados por vespas sem que expressassem qualquer feição de dor (COPPENS, 2008, p. 766).

Coppens (2008) também relata a relação dos Arutani com o nascimento e com a morte das pessoas que compunham sua comunidade e que faziam parte de sua etnia. Quanto ao nascimento, o autor apresenta que após o parto a mãe e o pai da criança que nasceu deveriam passar por um período de descanso estrito de cinco dias, podendo se alimentar apenas de pequenos peixes. Já sobre a morte, esse autor retrata que a pessoa que falecia era enterrada horizontalmente, na maioria das vezes dentro da casa onde a pessoa falecida vivia, mas às vezes era enterrada do lado de fora também. O ato de chorar parecia ser algo importante para eles, já que, segundo Coppens (2008, p. 766, tradução nossa) “Os parentes próximos do falecido tinham que deixar correr livremente as lágrimas. Se as limpassem com as mãos, poderiam virar mar, correndo risco de não chegarem ao céu”⁶.

Com base em duas entrevistas, uma delas dividida em duas seções de gravação, feitas pelo Dr. Thiago Chacon em 2017 e em 2020 – intituladas “biografia_texto1_8-12-17”, “biografia_entrevista_20-12” e “biografia_entrevista2_20-12” –, Juanita relatou um processo de

⁶ Texto original: “*Los parientes cercanos del difunto tenían que dejar correr libremente las lágrimas. De limpiarlas con la mano, podrían convertirse en un mar, con peligro de no llegar el alma al cielo*”.

enterro diferente do descrito em Coppens (2008). Ao ser perguntada se o seu grupo antigo cremava as pessoas igualmente aos Xiriana (Ninam), Juanita e um de nossos colaboradores disseram que os Arutani não cremavam os mortos como outros povos, mas os enterravam e depois tiravam os ossos para cremá-los. Hoje em dia, na comunidade Saúba, eles contam que o processo é semelhante, as diferenças são que atualmente existem duas formas, as quais são escolhidas pela família da pessoa falecida: (1) a família pode optar por enterrar seu familiar falecido dentro de casa, ou (2) podem colocá-lo no jirau – que é uma espécie de armação/estrutura de madeira, algo parecido com uma bancada, ou cama, típico em algumas regiões do norte da América do Sul – onde o morto é cremado.

Os Arutani, ainda segundo Coppens (2008), acreditavam estar cercados por dois tipos de espíritos, os bons e os maus, sendo os bons chamados de *Mawari*, os quais viviam no céu. Dentre esse tipo bom existia o espírito *Uti*, conhecido como criador da Terra e dos homens, sendo ele superior aos outros espíritos. Já os espíritos maus moravam nas colinas e corredeiras.

Para eles, o ser humano era composto por três partes distintas: corpo, *katuma* (que seria a ‘sombra’, ou traduzido como ‘dobro’) e *tirakwate* ‘alma’. Eles acreditavam que a sombra desaparecia quando a pessoa morria e que a alma tinha dois destinos, ou o céu ou o *anekëri*, parecido com a ideia que se tem popularmente no ocidente de inferno, localizado onde o sol se põe. A alma iria para um desses dois lugares a depender do comportamento de cada um (COPPENS, 2008, p. 767). Por fim, Coppens (2008) explica que, para os Arutani, o único intermediário entre o mundo espiritual e o mundo físico era o *imu*, uma espécie de pajé. Existia também uma relação entre mestre e aprendiz para que fossem passados os ensinamentos e outro *imu* fosse formado – para isso, eram necessários meses de preparação.

A situação de perda da transmissão da língua Arutani que resultou no seu número reduzido de falantes atual (cf. Chacon *et al.*, 2020) começa a fazer sentido ao observarmos o que Coppens (2008, p. 768, tradução nossa) apresenta sobre os câmbios culturais e o desaparecimento da língua Arutani:

Os Uruak estão prestes a desaparecer como uma cultura diferenciada. Observamos anteriormente que o último núcleo tribal foi fisicamente e

culturalmente integrado aos Yanam. Esta assimilação é tão avançada que a maioria dos Uruak, em vez de falar entre si a sua própria língua, usa a língua Yanam como veículo normal das suas conversas diárias. Devido a esta circunstância, é provável que muitos deles tenham esquecido aspectos importantes de sua própria língua.

Labrada *et al.* (2020) apresentam três principais pontos pelos quais se deveu de maneira inicial essa diminuição populacional dos Arutani, que seriam: (1) os conflitos e os deslocamentos por conta de outros grupos indígenas maiores, como os Marakaná, tidos como “inimigos mortais” dos Arutani, e depois por grupos Caribe e Yanomami; (2) os ataques de grupos Caribe com o objetivo de capturar pessoas para serem escravas; e (3) as epidemias de doenças como sarampo.

Outro ponto que contribuiu para essa diminuição foram os casamentos e a escolha dos falantes de Arutani em falar as línguas desses outros povos (cf. Labrada *et al.*, 2020; Coppens, 2008). Portanto, o que podemos supor é que ocorreu uma relação de bilinguismo não igualitário entre os Arutani e os Ninam, nessa relação de exogamia linguística que ocorria/ocorre na região. Com isso, um povo que já era historicamente minoritário – dentro de uma outra minoria que são os povos indígenas –, acabou sendo assimilado linguística e culturalmente pelos Ninam.

Quanto à língua falada pelos Arutani, segundo Labrada *et al.* (2020), o *status* seria de uma língua isolada, adormecida ou gravemente ameaçada. Os autores ressaltam, ainda, que “Arutani” é como Juanita se identifica quanto à sua etnia e refere-se quanto à sua língua, por isso, iremos continuar nos referindo a ela assim até o final deste trabalho. Labrada *et al.* (2020, p. 172, tradução nossa) ainda com base em relatos de Juanita e membros de comunidades Ninam do Uraricaá, afirmam que “os Awaké e os Arutani eram dois grupos étnicos distintos, falando variedades linguísticas diferentes, mas relacionadas” – essa questão será melhor explicada e desenvolvida ao apresentarmos o trabalho de Chacon & Labrada (2023) na seção 1.2.3.

1.2.1 Biografia e perfil sociolinguístico da última falante de Arutani no Brasil

A seguir, apresentaremos a biografia de Juanita e, juntamente, tentaremos traçar seu perfil sociolinguístico, para que, assim, possamos entender melhor parte de sua história, sua trajetória de vida, seu contexto social e quais os reflexos que isso tudo tem, ou pode ter, em sua fala.

Analisar seu perfil sociolinguístico pode, além disso, evidenciar as diferenças entre seu idioleto, os doculetos de que temos conhecimento e os trabalhos futuros descrevendo a fala dos falantes em território venezuelano.

Para contarmos a biografia de Juanita e descrevermos seu perfil sociolinguístico, utilizamos como base em gravações de entrevistas feitas pelo Dr. Thiago Chacon com Juanita em 2017 e 2020. Essas seções de entrevistas foram transcritas e documentadas por meio do *software* ELAN, o qual possibilitou que essas informações fossem incorporadas ao *corpus* do projeto de salvaguarda do Arutani. Ru Anita Xiriana Yanomami (*Ru Anita* é uma corruptela do nome em espanhol, Juanita) é uma senhora de aproximadamente 82 anos de idade. Nascida, segundo sua identidade, em 1942 na Venezuela – esta data é aproximada, já que seu registro foi feito quando Juanita já era mais velha. Ela é moradora, atualmente, da comunidade Saúba, próxima ao rio Uraricaá, do lado brasileiro da fronteira entre o Brasil e a Venezuela, na TI Yanomami em meio aos Ninam, também conhecidos na região como Xiriana. Nos dias atuais, ela é a última/única falante de Arutani em território brasileiro segundo as informações mais recentes.

Ela é filha de Kapaorio⁷ (sua mãe) e Xitoraw (seu pai), sendo sua mãe Arutani e seu pai Xiriana. Kapaorio e Xitoraw tiveram dois filhos juntos: Juanita e Pedro Martins. Kapaorio já havia sido casada com outro homem (o qual não foi identificado por nome por Juanita), com quem teve três filhas: Ana, Kôxi e Kariô, que moram possivelmente na comunidade Colibri – na Venezuela – e que talvez também saibam falar Arutani, caso ainda estejam vivas.

Retomando algumas informações importantes sobre Juanita, ela nasceu, mais especificamente, na comunidade Makô, próxima ao igarapé Mamoram, que desemboca na cachoeira Mãihã. Essa comunidade está localizada no rio Alto Paragua, na região da Venezuela. A comunidade Makô era uma comunidade Arutani e ficava longe da beira do rio, em uma região mais plana, pois era um local mais apropriado para o cultivo de certos alimentos e frutas, como banana. Nessa comunidade, Juanita conta que cresceu com várias crianças que também falavam Arutani, mas que, infelizmente, todas já faleceram. Ela contou que era chamada de *Kari* por seus

⁷ Seu nome vem de um apelido dado pelas pessoas de sua comunidade, pois quando era criança, Kapaorio adoeceu após comer um peixe. *Kapa* significa ‘comer, mastigar’ em Arutani, mas não sabemos ao certo se existe uma relação entre o verbo e seu nome.

familiares, espécie de apelido, que significa ‘menina’. Algum tempo depois, um pouco mais crescida, Juanita, sua família e outras pessoas do seu grupo fundaram outra comunidade, agora já mais próxima ao rio e que, segundo ela, fica perto de duas serras grandes. Essa nova comunidade ficava acima da cachoeira Mãihã.

Após um tempo, enquanto eles viviam nessa comunidade, o pai de Juanita acabou falecendo. Por conta disso, seu irmão, Pedro Martins, acabou arranjando um casamento entre Juanita, sua irmã, e Ramón, seu amigo. Ramón (já falecido) era Xiriana e foi escolhido por Pedro para ser o esposo de sua irmã pelo fato de ter um motor de barco e saber pilotá-lo, algo muito importante e necessário para a locomoção na região. Juanita afirma que no início não queria casar-se com Ramón. Porém, já casados, Juanita e Ramón se mudaram para a comunidade Kosoiba – uma comunidade que ainda existe na região e que possui três (semi-)falantes de Arutani (cf. Labrada *et al.*, 2020). Juanita e Ramón tiveram cinco filhos, mas apenas dois sobreviveram, Trina e Jesus (Rezú), nascidos na comunidade Koakap – Juanita não nos informou onde e nem quando ela viveu nessa comunidade, mas sabemos, pelos seus relatos, que ficava na Venezuela e que foi depois de ter se mudado para Kosoiba. Uma outra informação importante é que Ramón, pelo o que nos foi informado, já havia tido outra esposa, tanto que, quando foi viver com Juanita, levou consigo um filho do outro casamento.

Após isso, Juanita e sua família cruzaram a serra Parima – que marca a divisa entre Brasil e Venezuela na região – e vieram para o lado brasileiro da fronteira. Por decorrência de uma doença comum daquele período, seu marido, Ramón, acabou falecendo quando já estavam vivendo no Brasil. Morando atualmente na comunidade Saúba, ela repassa seus conhecimentos para as novas gerações e contribui para a documentação e salvaguarda de sua língua materna. Apesar de tudo, e de não ser sua intenção, Juanita acabou tornando-se um verdadeiro e explícito exemplo da resistência e da luta indígena no Brasil, reforçando que o multilinguismo e a diversidade linguística nacional ainda são uma realidade. Mesmo com todo o genocídio que vem ocorrendo nas TI Yanomami – justamente onde Juanita vive atualmente – e com a drástica diminuição no atual número de falantes das diversas línguas dos povos tradicionais brasileiro (cf. Rodrigues, 1966; Garcia, 2014), a simples existência de pessoas como ela já nos faz ter uma

esperança de um futuro mais diverso, tanto cultural quanto linguisticamente. Espera-se que toda essa riqueza imaterial seja um dia, verdadeiramente, valorizada em âmbito nacional e internacional.

Traçar um perfil sociolinguístico é sempre uma tarefa complexa, pois é necessário que sejam levados em consideração diversos conceitos e definições que nem sempre abrangem a profundidade dessa relação entre linguagem, indivíduo e contextos sociais. Em uma breve tentativa de fazer isso, ressaltamos que: Juanita é uma falante bilíngue, falando fluentemente o Ninam e lembrando o Arutani com fluência. Ela aprendeu Arutani principalmente com a sua mãe (que era Arutani) e Ninam com o seu pai (que era Xiriana) e com pessoas com quem sempre teve contato. Juanita diz que aprendeu primeiramente o Arutani, porém o Ninam acabou tornando-se sua língua dominante, provavelmente por conta de sua maior frequência de uso e pelo prestígio regional.

Essa “troca” de L1⁸ é algo comum e que já acontece há décadas entre os Arutani e os povos da família Yanomami. Coppens (2008, p. 768) relata que algo parecido aconteceu em uma proporção comunitária com os indígenas Uruak (Arutani) vivendo entre os Yanam (Ninam). Em seu relatado, o autor apresenta que a cultura Uruak estava próxima da extinção, pois o último grupo tribal havia sido completamente assimilado pelos Yanam, sendo que esse nível de integração levou os Uruak a adotarem a língua Yanam no dia a dia, em vez de usarem o seu idioma nativo, o que provavelmente resultou no esquecimento e desuso da língua.

Entretanto, ressaltamos que não julgamos tais fatos como positivos ou negativos, estamos apenas relatando a realidade de um contexto multilíngue. Com relação a isso, Bortoni-Ricardo (2014, p. 26, com adaptações) destaca que: “É sempre bom reiterar que a situação do multilinguismo é mais comum no mundo do que a de monolinguismo” e ressalta que “a cultura brasileira é avessa a admitir a característica multilíngue de nosso país”. Esse primeiro fato nos mostra que, talvez por questões históricas e sociais de possíveis conflitos e dominações na região

⁸ Entendemos L1 como sua língua dominante, a qual possui mais fluência e uso social; e L2 como sua segunda língua de mais fluência e uso social. Ademais, definimos aqui língua materna como a língua herdada pela mãe (Arutani) e língua paterna como a língua herdada pelo pai (Xirian/Ninam). No caso de Juanita, sua L1 era o Arutani, a qual passou a ser sua L2 e o Ninam sua atual L1 (para um aprofundamento da discussão, ler Spinassé, 2008).

da TI Yanomami em que o Arutani era mais comumente falado, apesar de ter sido (e ainda ser) uma língua minoritariamente falada entre a população da região, ocorresse algum tipo de estigmatização da língua e um prestígio de outras (para um panorama sobre estigma e prestígio cf. Bagno, 2007).

Como já apresentamos, Juanita cresceu em uma comunidade em que outras crianças falavam Arutani, assim como ela, e, atualmente, na comunidade do Saúba ela não tem ninguém para conversar em sua língua materna, pois seus filhos, Trina e Jesus (Rezú), sabem apenas poucas palavras da língua da mãe. Ademais, seu irmão, Pedro Martins, também não fala Arutani, pois, apesar de serem filhos da mesma mãe, ele não teve tanto contato com a língua, pois foi levado preso ainda muito novo pelas autoridades policiais da Venezuela e voltou a viver novamente entre seu povo quando já estava mais velho, como conta Juanita. Porém, Juanita relatou quais foram as últimas pessoas com quem pôde conversar em Arutani. Foram duas mulheres, uma senhora chamada Apôto e outra chamada Xapa (uma senhora cega e já falecida).

A partir deste panorama, do perfil e do contexto sociolinguístico de Juanita, é importante o destaque de alguns conceitos da sociolinguística e como eles se aplicam ao Arutani falado por Juanita. Para isso, apresentaremos inicialmente os conceitos de **variedade linguística**, **dialeto** e **idioleto**⁹. Com foco na objetividade e aplicação conceitual, não nos aprofundaremos na problemática entre as definições dos termos “dialeto” e “idioleto”, pelo contrário, optamos pela definição pontual apresentada em Bagno (2007, p. 47), na qual ele explica que uma variedade linguística representa uma das diversas maneiras de falar uma língua, associada a fatores sociais como origem, idade, gênero, classe social e nível de instrução. A sociolinguística, fundamentada na ideia de heterogeneidade, considera que toda língua é composta por essas variedades. O mesmo autor ainda afirma que apesar de “dialeto” ser um termo antigo usado desde a antiga Grécia, atualmente, os linguistas optam por usar o termo “variedade”, ou seja, são sinônimos abarcando o mesmo conceito. Quanto ao conceito de “idioleto”, Bagno (2008, p. 48) diz que este termo “designa o modo de falar característico de um indivíduo, suas preferências vocabulares, seu modo próprio de pronunciar as palavras, de construir as sentenças etc.”.

⁹ Para esta discussão não é tão relevante a conceitualização e aplicação de termos como “socioleto” e “cronoleto”, por exemplo, já que não se aplicam ao contexto de análise e descrição do Arutani, a qual possui apenas uma falante.

A partir de uma perspectiva sociolinguística, podemos dizer que Juanita possui seu próprio idioleto que é produto de toda a sua história de vida e dos diversos contextos sociais e linguísticos em que ela esteve e está inserida. Além disso, podemos, também, observar que sua variedade é característica de alguém que viveu em diversas comunidades sendo influenciada por diferentes contextos sociolinguísticos. Sendo assim, sua variedade do Arutani pode ser “resumida” como a variedade de uma mulher de aproximadamente 82 anos; natural da comunidade Makô na Venezuela, que viveu em diferentes comunidades multilíngues; não-escolarizada no sentido tradicional; viúva; mãe de dois filhos; uma das matriarcas da comunidade Saúba onde vive atualmente; e que teve contato constante com a língua Ninam, a qual passou a ser sua L1, sendo substituída pelo Arutani, sua língua materna.

Um ponto importante, o qual vale destaque sobre o perfil sociolinguístico de Juanita, é a questão da obsolescência e do atrito que o Arutani teve para ela com relação ao Ninam. Essas duas línguas estavam disputando espaço no meio social em que Juanita estava inserida, sendo o Ninam uma língua mais usada que o Arutani e comunicativamente mais necessária/produtiva para ela, sem contar com o fato de Juanita ficou bastante tempo sem falar Arutani, voltando a utilizá-la apenas em 2017.

1.2.2 Bibliografia sobre a língua Arutani

Há trabalhos que apenas mencionam a existência do povo ou da língua Arutani, como em: Pisemskaya (2009, 2010) que também denomina o povo como Arutani; Rodrigues (1992) ao listar vários povos indígenas brasileiros, também os chamando de Arutani; Ferreira *et al.* (2019) também fazem o mesmo. Campbell (2010) menciona o Arutani, chamado por ele de Awaké, em uma lista de línguas isoladas da América do Sul, fazendo, assim, a classificação filogenética da língua. Algo parecido foi feito por Loukotka (1968), que, também, em uma lista de línguas, apresenta o Arutani como uma língua isolada, porém apresenta alguns dados da época sobre distribuição geográfica dos falantes. Jolkesky (2016) faz um panorama sobre os extratos

lexicais¹⁰ possivelmente compartilhados pelos Arutani com outros grupos etnolinguísticos, como o suposto grupo “Dohu”, Makú, Sapé, Tukáno e Warao.

Armellada e Matallana (1942) realizaram uma pesquisa durante uma exploração no rio Paragua entre 1933 e 1936, talvez a mais antiga de que temos registro. Lá eles encontram um povo, o qual os autores identificam como Arutani. Com relação à população étnica, eles relatam a existência de 11 Arutani. Nesse trabalho, também, são listadas cerca de 200 palavras em Arutani, que cobrem campos lexicais da (i) natureza; outro campo que os autores chamam de (ii) natureza humana e, juntamente, o léxico de família; (iii) etnografia; (iv) reino animal; (v) reino vegetal; outro grupo de palavras que eles chamaram de (vi) cores e qualidades; também, outros grupos de palavras que se referem a (vii) numerais e pronomes; e, por fim, algumas frases. Essa pesquisa tem um valor histórico muito importante, já que se trata da primeira vez que se tem algum registro escrito sobre os Arutani.

Em Migliazza (1967, p. 170-171) temos diversas informações importantes e que nos dão pistas sobre os Arutani. O autor aponta que a língua, a qual ele chama de Awaké, era falada há tempos atrás nos rios Paragua e Uraricaá e fala que os topônimos da região são todos de origem dessa língua. Migliazza (1967, p. 170-171, *grifos do autor*, com adaptações) aponta que:

Os Awaké autodenominam-se [orotani], e contam que suas terras foram invadidas primeiro pelos Karib que eles chamam de Porokotó e depois, quando estes retiraram-se para o médio Parágua, (onde são os atuais Arekuna), chegaram os Yanomami, que chamavam de Kasrapai ou seja *lábio comprido*. Os Awaké são chamados pelos Yanomami de [iriák], e atualmente casam-se com os Yanomami e são bilíngües, falando as línguas Ninam e Awaké. A nova geração fala somente a língua Ninam (Yanomami). Os adultos que ainda falam Awaké são 17, que moram no alto Uraricaá e alto Parágua.

Destacamos essa autodenominação que o autor transcreve como [orotani], muito próximo à autodenominação de Juanita, que é Arutani. Acreditamos que pode ter sido um nome que sofreu mudança e variou com o tempo até chegar ao nome atual ou apenas uma compreensão equivocada da pronúncia por parte do autor, é difícil escolher uma das hipóteses com poucos dados à disposição. Outro ponto importante é como os Yanomami chamavam esse povo, que é

¹⁰ Essa amostragem da língua é pouco relevante para trabalhos de caráter fonético-fonológico ou histórico-comparativos, já que são realizados a partir de uma ortografia e não nos fornece muitas informações no âmbito fonético da língua.

transcrito por Migliazza como [iriák], próximo ao etnônimo Uruak comum na literatura e utilizado por Coppens (2008) para falar sobre este mesmo povo. Por fim, temos o número reduzido de falantes, no caso, 17, os quais são mais velhos e são bilíngues nas duas línguas. O último ponto relevante é o de que os mais jovens não falam mais a língua, o mesmo fato relatado em Coppens (2008).

De uma perspectiva linguística, Migliazza (1967, p. 171) propõe a existência de 13 fonemas consonantais: p [p], t [t], k [k,g], ? [ʔ]¹¹, s [s], š [š]¹², h [h], r [r̥]¹³, m [m], n [n], ñ [ɲ], w [u], y [i]¹⁴; e de 5 vogais: i [i, ɪ], e [e, ɛ], a [a, ə], o [o, ɔ] e u [u]. Em um segundo trabalho publicado 11 anos após esse de 1967, Migliazza (1978, p. 135-136) apresenta a seguinte proposta do quadro fonético da língua que agora ele chama de Uruak: as oclusivas: p, t, k, ʔ; as fricativas: s, š, h; as nasais: m, n, ñ; o *flap*: r; as semivogais (chamadas assim pelo autor): w, y; e as vogais: i, e, a, ə, o, u.

Labrada *et al.* (2018) fazem um panorama e apresentam dados históricos e linguísticos sobre os Arutani e os Sapé em que os autores explicam a quase extinção desses dois grupos. Labrada *et al.* (2020) talvez seja um dos trabalhos mais completos sobre os Arutani. Apesar de não ser muito extenso, nele são apresentadas informações sobre a língua e o povo; mapas sobre distribuição de seus falantes; citam textos que trazem informações sobre a língua e a cultura Arutani; e o *status* atual da língua, do número de falantes e das pesquisas que estão sendo desenvolvidas (todas essas informações já foram citadas e apresentadas anteriormente).

Além disso, nos relatórios das pesquisas de campo feitos por Chacon (2019, 2020a e 2020b), explicita-se as metodologias de pesquisa de campo viagem à comunidade de Saúba. Esses relatórios são muito importantes para esta pesquisa, já que os áudios produtos das sessões de gravação realizadas nessas viagens de campo compõem o Dicionário Arutani (CHACON *et al.*, 2020), que é de onde estamos retirando a maioria dos dados para esta pesquisa.

¹¹ Acreditamos que ʔ [ʔ] seja a representação gráfica de uma oclusiva glotal [ʔ], isso fica evidente em Migliazza (1978).

¹² É possível que š [š] seja a representação gráfica da consonante fricativa pós-alveolar [ʃ].

¹³ O que o autor representa graficamente como r [r̥] é o *tepe* [ɾ], o qual o autor classifica com *flap*.

¹⁴ Migliazza classifica w [u], y [i] como semiconsoantes, que equivalem às semivogais [w] e [j], respectivamente.

Há, também, o recente trabalho de Chacon & Labrada (2023) que trata sobre o quão isolada é a língua Arutani, fazendo uma análise comparativa com empréstimos linguísticos e os diferentes doculetos encontrados sobre a língua. Esse trabalho é importante para a nossa pesquisa pelo fato de apresentar evidências sobre as variedades do Arutani ao propor qual seria o *status* ou categorização dessas variedades e qual seria a hipótese mais precisa sobre sua filiação genética (ver seção 1.2.3).

Por fim, existe, também, o dicionário multimídia Arutani (CHACON *et al.*, 2020), disponibilizado *on-line*, em que é apresentamos um grande *corpus* da língua, o que contribui bastante para a documentação da língua, sendo um produto muito importante para a valorização não só da língua, mas, também, da cultura Arutani. O dicionário traz visibilidade para a existência da língua e serve de recurso didático tanto para o ensino de linguística, por exemplo.

Quanto à questão da fonética e fonologia do Arutani falado por Juanita, por se tratar de um trabalho inicial de uma língua pouco descrita, iremos ter como referência os trabalhos anteriores, mas nos embasaremos no trabalho no de Chacon *et al.* (2020) em que há um breve resumo sobre a estrutura do idioleto, veja: “Na fala de Juanita Martinez, o Arutani possui cinco vogais fonêmicas (a, e, i, o, u) e 12 consoantes fonêmicas (p, t, k, ts, s, x [ʃ] ~ [ʂ], h, r, m, n, w, y [j] ~ [ɲ])”. Portanto, pretendemos averiguar estas hipóteses e propor, caso necessário, certas mudanças. Outro trabalho recente é o de Cabral (2022) na descrição da sintaxe da língua, explicando o Arutani como uma língua SOV, ou SV em casos de verbos intransitivos, apresentando alinhamento ergativo-absolutivo.

Um último trabalho sobre o Arutani é o de Zamponi (2023); nele o autor faz a análise de áudios gravados por Migliazza que nunca haviam sido analisados. A partir desses áudios Zamponi estabelece um inventário fonológico com cinco vogais /i, e, a, o, u/ e quatorze consoantes. As consoantes seriam: /p/ [p, ɸ], /t/ [k, g], /ʔ/, /ts/, /s/, /ʃ/, /h/, /m/, /n/, /ɲ/, /ɾ/, /w/ e /j/.

Em suma, a bibliografia sobre os Arutani e sua língua mostrou-se escassa, pouco detalhada num geral e de muito difícil acesso. Alguns textos antigos foram digitalizados, outros

encontram-se apenas em museus e bibliotecas fora do Brasil, tanto que Labrada *et al.* (2020) citam trabalhos que estão inacessíveis ou que podem ter sido perdidos sobre a língua. Essa realidade nos mostra ainda mais a importância do trabalho que estamos fazendo com o dicionário Arutani (cf. Chacon *et al.*, 2020), o qual é o principal catalisador dos trabalhos que estamos desenvolvendo atualmente, como este de descrição da fonética e fonologia do Aratani e outros de caráter sociolinguístico, tipológico e histórico-comparativo.

1.2.3 A língua Arutani: uma língua isolada?

O Arutani é atualmente classificado como uma língua isolada, levando em consideração os trabalhos supracitados. Entretanto, para chegarmos a uma classificação filogenética, precisamos levar muitas questões em consideração, a primeira delas é a compreensão clara do que é uma língua isolada. Campbell (2010, p. 16, tradução nossa) explica que

[...] uma língua isolada é uma língua que não tem parentes, isto é, que não tem nenhuma relação genética demonstrável com qualquer outra língua. É uma língua que não se demonstrou ser descendente de nenhuma língua ancestral que tenha outros descendentes (filhas). Assim, as línguas isoladas são, na verdade, famílias linguísticas com apenas um membro.

A partir disso, surge outro conceito necessário, que é o de “família linguística”. Novamente, Campbell (2010, p. 16-17) conceitua que, para se identificar uma família linguística, é necessário que existam evidências suficientes de que um conjunto de línguas descendam de uma mesma língua ancestral. O autor faz a ressalva de que uma família linguística pode ser composta por várias línguas ou por apenas uma língua, que é o caso das línguas isoladas.

Campbell (2010, p. 28) ainda mostra que as línguas isoladas não são tão raras como aparentam ser, existindo cerca de 136 línguas isoladas no mundo, o que daria um terço das 420 famílias linguísticas conhecidas. Campbell ainda destaca que “As línguas isoladas não são muito diferentes das línguas que têm parentes. Línguas isoladas poderiam facilmente ter parentes perdidos ou poderiam ter se diversificado em pequenas famílias de línguas relacionadas” (tradução nossa).

Além disso, existe uma diferença entre “línguas isoladas” e “línguas não classificadas”. Campbell (2010, p. 20, tradução nossa) explica que “As línguas isoladas não são agrupadas em

classificações genéticas maiores porque para elas existem dados e as comparações desses dados com outras línguas não revelam parentesco linguístico”. O ponto principal que difere esses dois tipos de línguas é a possibilidade de comparação com base em dados suficientemente necessários. Por isso, o autor ainda apresenta que uma língua não classificada é aquela sem documentação suficiente para determinar se possui parentes linguísticos. O Arutani é uma língua que está deixando de ser pouco documentada, graças ao projeto de salvaguarda, e veremos mais à frente o que se pode concluir até agora sobre a sua classificação filogenética.

Um ponto a ser levado em consideração dentro da discussão sobre a classificação de uma língua isolada ou uma família com duas ou mais línguas tem a ver com a definição do que são línguas diferentes ou diferentes variedades de uma mesma língua. Chacon *et al.* (2014, p. 69) explicam que a língua, da perspectiva da linguística, trata-se de um conceito baseado numa abstração nos níveis sistêmico e simbólico-político. O nível sistêmico, segundo eles, é “referente à língua enquanto um sistema baseado num conjunto de normas, signos e convenções sociais que substancia a comunicação entre indivíduos e grupos sociais”. Já o nível simbólico-político, é o nível de classificação que se refere “aos valores socioculturais, políticos e ideológicos construídos historicamente que definem uma língua *versus* outras línguas ou uma língua que abarca diversas variedades internas (dialetos, sotaques, etc.)”.

No trabalho de Chacon e Labrada (2023), vemos claramente a aplicação desses conceitos. Dessa forma, a proposta de classificação do Arutani feita por esses autores leva em consideração esses dois níveis de análise. Os autores partem dos relatos orais de Juanita sobre uma diversidade de grupos falantes de Arutani no passado. Segundo ela, **Arutani** seria a autodesignação de seu grupo, cujos habitantes seriam originários das cabeceiras do rio Uaricaá. **Awake** seria a designação de um grupo do baixo Uaricaá, já extinto. Os Awake falavam uma língua relacionada e, também, semelhante ao Arutani, porém eram inimigos dos Arutani e dos povos de língua Caribe que viviam no Uaricaá. Outro grupo seria conhecido como **Xinaroya**, composto por indivíduos Arutani, os quais viviam distantes e eram considerados canibais. Por último, temos o grupo conhecido como **Sa’si**, que possuíam uma língua próxima ao Arutani, porém

faziam o uso do xibolete (marca na fala para identificar um determinado grupo linguístico) [saʔsi] no final das frases, sendo que alguns deles vivam entre os Arutani.

A partir disso, precisamos levar alguns pontos importantes em consideração para a classificação de línguas ou variedades. A primeira é que, no nível simbólico-político, não é possível apresentar o Arutani como uma mesma língua ou mesma variedade do Awake, já que, apesar de serem línguas relacionadas no nível sistêmico, segundo Juanita, eram grupos inimigos (nível simbólico-político). Um segundo ponto é que, apesar de não termos registros e documentação da variedade Sa'si, só de haver o relato da existência de um xibolete já nos evidencia uma clara estratégia intencional de distinção tanto no nível sistêmico quanto no nível simbólico-político.

Além disso, os autores demonstram a diversidade interna entre os diferentes registros do Arutani – sejam os de variedades com falantes atuais como aquelas variedades documentadas historicamente. Os autores propõem o Arutani e o Awake como variedades irmãs, levando em consideração o nível sistêmico e o nível simbólico-político. Os autores trazem a hipótese de “diferentes socioletos contando histórias diferentes desde o início do período colonial” (CHACON & LABRADA, 2023). Com base nisso, podemos dizer que são possíveis variedades de uma mesma língua, ou podemos pensar numa pequena família de línguas bastante próximas entre si, mas apenas de uma perspectiva sistêmica/estrutural. Já de uma perspectiva política, social, cultural e identitária já seria algo mais complexo, visto que são línguas, ou variedades, com etnônimos diferentes e com grupos com diferentes histórias internas.

Apesar de não ser o principal foco desta pesquisa, de modo a colaborar com a hipótese de Chacon e Labrada (2023), elaboramos um estudo comparativo prévio entre 14 palavras das 100 que compõem a lista feita por Migliazza (1978) da variedade Uruak e das respectivas 14 palavras correlatas com o mesmo significado disponíveis no dicionário Arutani em Chacon *et al.* (2020).¹⁵

¹⁵ Para esta comparação não estamos nos baseando, ainda, na fonética ou na fonologia das variedades, mas, sim, nas convenções ortográficas seguidas pelos autores, as quais nos dão uma noção e uma ideia básica dos possíveis fonemas ali representados por grafemas.

Tabela 1 – Comparação entre a fala de Juanita e o “Uruak” registrado em Migliazza (1978)

Arutani		Português
Migliazza (1978)	Chacon <i>et al.</i> (2020)	Glosa
<i>kaykate/ka-</i>	<i>kaykate</i>	‘você’
<i>kari</i>	<i>kari</i>	‘mulher’, ‘irmã mais nova’
<i>kina</i>	<i>kina</i>	‘pessoa’
<i>yopsa</i>	<i>yopsa</i>	‘passarinho’
<i>toari</i>	<i>toari</i>	‘cachorro’
<i>maykate/ma-/tsa-</i>	<i>maikate</i>	‘eu’
<i>šawi</i> ¹⁶	<i>xawi</i>	‘alto’
<i>šapi</i>	<i>xapi</i>	‘árvore’
<i>ayta</i>	<i>ayata</i>	‘este’ (locativo), ‘ali’
<i>maya</i>	<i>maña</i>	‘o que’
<i>sikiپی</i>	<i>kipisi</i>	‘pequeno’
<i>kitate</i>	<i>kitati</i>	‘tudo’
<i>kaʔtyaw</i>	<i>kaptiaw</i>	‘muito’

Com base na tabela anterior, pode-se dividir os termos listados em três grupos de pares de palavras: (1) em que as palavras são idênticas e não sofreram nenhum tipo de mudança linguística ou ortográfica; (2) outro em que as palavras aparentam ser as mesmas, mas com diferenças ortográficas; e (3) pares de palavras em que ocorre algum tipo de mudança linguística (seja fonética ou qualquer outro tipo de mudança/variação). Neste último caso, elas podem ser sinônimas na língua, ou palavras que sofreram mudanças linguísticas drásticas ou, até mesmo, palavras que não são correspondentes apesar de possuírem a mesma semântica. Para que essas afirmações sejam confirmadas, seria necessária uma análise mais aprofundada – seja uma reconstrução com base no Método Histórico-Comparativo, por exemplo, ou uma comparação com outras línguas, supondo-se que tenha ocorrido algum dos diferentes tipos de empréstimo.

Para o primeiro grupo de palavras temos as seguintes correspondentes na Tabela 1:

- (1) a. *kaykate*

¹⁶ Entenda os grafemas <š> e <x> como o fone [ʃ].

- ‘você’
- b. *kari*
‘mulher’
- c. *kina*
‘pessoa’
- d. *yopsa*
‘pássaro’
- e. *toari*
‘cachorro’

Em (1) não há nenhum tipo de diferença entre os dois registros, seja por convenção ortográfica ou por algum tipo de mudança linguística. Já em (2) temos as seguintes correspondentes quando comparamos os registros em Migliazza (1978) com os de Chacon *et al.* (2020), respectivamente:

	Migliazza (1978)	Chacon <i>et al.</i> (2020)
(2) a.	<i>maykate</i> ‘eu’	<i>maikate</i> ‘eu’
b.	<i>šawi</i> ‘alto’	<i>xawi</i> ‘alto’
c.	<i>šapi</i> ‘árvore’	<i>xapi</i> ‘árvore’

O par em (2a) está no segundo grupo pelo simples fato de ser um possível caso que interpretação da língua ou uma preferência ortográfica, já que não podemos afirmar qual foi realmente o som em posição de coda silábica escutada por Migliazza, se seria o glide [j] ou a vogal alta anterior não-arredondada [i]. Dessa forma, não podemos afirmar se trata-se de um ditongo ou um hiato. Já em (2b) e em (2c) ocorreu apenas uma questão de preferência e convenção ortográfica, já que o que na escrita foram representados como <š> e <x> são na verdade a mesma consoante fricativa pós-alveolar não-vozeada [ʃ]. Com isso, partimos para o terceiro grupo de palavras, o qual já possui mudanças mais significativas.

Em (3) temos as seguintes correspondentes na Tabela 1 quando comparado os registros em Migliazza (1978) e Chacon *et al.* (2020), respectivamente:

	Migliazza (1978)	Chacon <i>et al.</i> (2020)
(3) a.	<i>ayta</i> 'este' (locativo), 'ali'	<i>ayata</i> 'este' (locativo), 'ali'
b.	<i>maya</i> 'o que'	<i>maña</i> 'o que'
c.	<i>sikiپی</i> 'pequeno'	<i>kipisi</i> 'pequeno'
d.	<i>kitate</i> 'tudo'	<i>kitati</i> 'tudo'
e.	<i>ka'tyaw</i> 'muito'	<i>kaptiaw</i> 'muito'

Para esta breve análise, estamos tomando uma perspectiva diacrônica, partindo do registro mais antigo, de 1978, para uma fala mais atual, sendo o idioleto de Juanita. Nessa perspectiva, em (3a) teria ocorrido uma epêntese (ou inserção) (cf. Campbell, 2013), na qual ocorre a inserção da vogal <a>. Já para a mudança que ocorre em (3b), há duas possibilidades de análise: (i) a primeira é a de que seja o mesmo fone interpretado de maneiras distintas pelos diferentes pesquisadores já que [j] e [ɲ] são sons foneticamente muito semelhantes, ou (ii) a segunda, em que ocorreu um fortalecimento do glide [j] <y> para a consoante nasal palatal vozeada [ɲ] <ñ> (ou seja, [j] > [ɲ]). Vale ressaltar que em Chacon *et al.* (2020) evidenciamos essa variação entre [j] e [ɲ] na fala de Juanita, o que sugere uma mudança em curso

Por outro lado, em (3c) ocorreu uma metátese, que seria, com base em Campbell (2013, p. 35), a transposição de sons dentro de uma palavra. Segundo o autor, “Muitas vezes, acredita-se que a metátese seja encontrada principalmente apenas em alterações esporádicas, mas a metátese também pode ser uma alteração regular” (tradução nossa). Isso pode ocorrer facilmente no português do Brasil como em ‘sílaba’ > *‘síbala’. Portanto, a mudança *sikiپی* > *kipisi*, é um ótimo exemplo de metátese, podendo ser, ou não, uma forma já consolidada na fala de Juanita. Se já

tiver sido consolidada, essa mudança pode ser uma variação própria de seu idioleto ou herdade de seu contato com outros falantes de Arutani.

No caso de (3d) temos uma variação entre *kitate*, em Migliazza (1978), e *katati*, em Chacon *et al.* (2020). Neste caso, há uma variação entre <e> e <i>, os quais podem corresponder, respectivamente, a [e] e [i]. Em *katati*, sabemos que se trata de [i]. Essa variação entre [e] e [i] entre os dialetos, como entre a fala de Juanita e outros doculetos, é muito comum nas variedades do Arutani de uma perspectiva diacrônica. Exploraremos mais esta questão na seção 4.2.2.1.

Em (3e), teríamos inicialmente a hipótese de uma mudança de uma pausa glotal [ʔ] para uma oclusiva bilabial não-vozeada [p]; porém, essa mudança não faz sentido, pois não se pode prever o desenvolvimento de um ponto de articulação bilabial no contexto da palavra. Uma melhor hipótese para a relação entre [ʔ] e [p] é a de que é comum no Arutani que oclusivas em posição de coda silábica variem ocorrendo com ou sem plosão (*burst*), ou seja, [p] ~ [p̚]. Assim sendo, o que pode ter ocorrido, na verdade, é que no registro de Migliazza (1978) tenha ocorrido a produção de um som próximo a um [p̚], que é acusticamente semelhante a uma pausa glotal. Uma outra hipótese seria que, pelo fato de a língua estar em processo de extinção, Juanita tenda a ter uma fala mais conservadora, uma forma de reforçar traços que possivelmente foram apagados ou sofreram algum tipo de lenição na fala registrada em Migliazza (1978). Essa possível fala mais conservadora pode ser uma estratégia para reforçar traços identitários. Sendo assim, estaríamos analisando um processo de debucalização, o qual deixou de ocorrer na fala de Juanita Juanita, portanto, encontramos a seguinte sequência de mudanças: [p] > [ʔ] > [p]. Dessa forma, fica clara a importância de estudos acústicos para sanar possíveis imprecisões e dúvidas como esta, assim como fizemos para este trabalho. Veremos mais sobre isso nas seções de análise e descrição.

Apenas uma análise mais aprofundada nos esclarecerá se há ou não uma mudança entre a variedade Uruak de Migliazza - mais antiga - e a variedade de Juanita - mais moderna. Uma outra possibilidade é que se trate de dialetos diferentes, sem relação de sucessão entre eles.

1.3 Contexto e realidade da pesquisa desenvolvida

O *status* atual de vitalidade da língua Arutani, com apenas uma falante no Brasil e mais dois na Venezuela, pode parecer algo raro, mas esse tipo de situação só vem crescendo no mundo. Devido ao seu número exíguo de falantes já em idade avançada, e devido ao fato de que não há pessoas aprendendo a língua, o Arutani foi documentado em um projeto de salvaguarda, em que o foco não foi, e não é, de revitalização, mas, sim, de documentação e de registro para o conhecimento de sua existência para gerações futuras e para a contribuição para com a linguística e outras ciências. Portanto, os dados coletados com base nas sessões de gravação não foram muito extensos.

As realidades de pesquisa de campo não são as mesmas de pesquisas laboratoriais, ainda mais quando falamos do contraste entre línguas quase extintas *versus* línguas amplamente faladas. Dessa forma, temos o desafio de descrever uma língua com apenas uma interlocutora, com 82 anos e que, segundo os relatórios de campo, não sustentava fazer e trabalhar em longas sessões de gravação, mesmo que se tratando de falas espontâneas e de contação de histórias e de compartilhamento de seus conhecimentos. Juanita rapidamente sentia-se cansada e isso refletiu na quantidade de palavras registradas em nosso *corpus*, sem contar com o fato de que, num primeiro contato com os pesquisadores, ela demorou a lembrar de certas palavras, já que fazia tempo que não falava Arutani, pelo fato de não ter ninguém com quem conversar, apenas falava Arutani em sonhos, como relatou. Porém, Juanita foi adquirindo cada vez mais fluência e segurança no uso da língua com o passar do tempo durante as sessões de documentação.

O trabalho em campo com povos de recém contato, como os Arutani e os Ninam, não é simples, tendo em vista o difícil acesso às comunidades e a dificuldade de se levar equipamentos até lá, visto que algumas viagens demoram dias subindo o rio até chegar ao destino final, como é o caso da comunidade Saúba, onde Juanita mora. Saindo de Brasília é possível pegar um voo até Boa Vista que dura um pouco mais de 3h e 30 min., em seguida é preciso fazer uma viagem de 2h e 30 min. até o município de Amajari e após isso uma outra viagem de quase três dias subindo o rio até a comunidade Saúba. Outras dificuldades são a falta de acesso à comunicação nessas comunidades, além do fato de Juanita não aguentar mais fazer essas viagens longas até a cidade.

Dessa forma, muitos pesquisadores de laboratório dirão que as condições em que os áudios foram gravados não são as mais adequadas, visto que não têm o isolamento acústico necessário e “apropriado” para uma análise acústica. Dirão, também, que o *corpus* deveria ser maior, mas isso pode ser explicado pelas dificuldades já listadas e pela urgência em se documentar a língua ao máximo, tendo em mente que, caso Juanita falecesse, não teríamos mais falantes vivos no Brasil. Veremos neste trabalho que é possível, sim, fazer análises acústicas detalhadas e cientificamente embasadas, mesmo sem as condições adequadas, de modo que este trabalho poderá servir como exemplo para pesquisas futuras que lidam com estudos em fonética e fonologia de línguas indígenas.

Atualmente, são poucas as pesquisas de análise acústica de línguas indígenas no Brasil. Devemos aproveitar e fazer o máximo de uso das ferramentas atualmente desenvolvidas, já que décadas atrás não era possível ser feito o trabalho com base nos recursos tecnológicos que foram desenvolvidos até os dias de hoje. Com a quantidade de avanços e *softwares* a que temos acesso, torna-se evidente que pesquisas de análise acústica com línguas indígenas devem ser feitas com o foco de tornar mais precisa certas análises e respaldar ou refutar trabalhos antigos que foram realizados antes da existência de tais recursos.

Trabalhos iniciais de documentação, como este, precisam de um elevado grau de detalhamento, para que tenhamos o mínimo de equívocos e inconsistências em análises futuras, pois se existem imprecisões no âmbito fonético-fonológico, isso influenciará com certeza nos âmbitos morfológicos, sintáticos e discursivos, por exemplo. Com isso, trazemos em voga o fato de que línguas à beira da extinção também merecem um elevado grau de detalhamento, visando, assim, contribuir, também, para estudos tipológicos e comparativos.

Foi necessário ter uma certa coragem para desenvolver esta pesquisa, tanto por estarmos trabalhando com uma língua isolada e quase extinta, quanto por mostrarmos que trabalhos que aparentam ser quase impossíveis ou inviáveis de serem feitos podem ocorrer desde que haja um estudo profundo; o uso/auxílio de novas tecnologia e metodologias cientificamente embasadas; a utilização de recurso estatísticos e qualitativos comunicando-se entre si; e uma predisposição para que o trabalho seja feito com rigor. Para isso, analisamos nosso *corpus* várias vezes, foram

feitas diversas correções ao longo do processo. Além disso, foi necessário ter até mesmo uma certa paixão e um propósito bem definidos nos âmbitos científicos, sociais e políticos que envolvem esta pesquisa. Portanto, não existem condições perfeitas para a análise acústica quando tratamos de documentação de línguas indígenas, pois as pesquisas são feitas em campo no meio das comunidades indígenas.

Por fim, um aprendizado que tivemos no desenvolvimento desta pesquisa foi o de que não existem dados ruins, mas, sim, mau uso de dados. Sendo assim, buscamos fazer uma análise e uma descrição da fala de Juanita utilizando o máximo de recursos disponíveis e extraíndo o máximo de informações que estavam ao nosso alcance, mas sempre respeitando os limites que os próprios dados nos imponham.

1.4 Justificativa

Sobre esta pesquisa, existem motivações tanto de caráter pessoal quanto de caráter científico e acadêmico. Após um contato mais profundo com a linguística funcional e a linguística descritiva nos primeiros anos da graduação, surgiu em mim o desejo, por inspiração em trabalhos de alguns professores, de um dia poder descrever e analisar uma língua ainda pouco descrita. Por falta de conhecimento, eu acreditava ser algo impossível, pelo fato de observar a “abundância” de linguistas já trabalhando nessa área e por ver a quantidade considerável de pesquisas sendo realizadas no Brasil e no exterior com foco na descrição de línguas indígenas. No entanto, essa vontade continuou existindo, ainda mais ao conhecer melhor o trabalho de alguns professores durante a graduação – como os trabalhos do professor Dr. Thiago Chacon, com os Kubeo e os Ninam, do professor Dr. Dionei Gomes, com os Mundurukú, e da professora Dra. Altaci Rubim, com seu povo, os Kokama – que elaboraram gramáticas, desenvolveram descrições e projetos de valorização linguística. Esses trabalhos precursores na descrição de algumas línguas sempre me motivaram como pesquisador/linguista. Acredito na importância dessas pesquisas para seus respectivos povos, tanto de uma perspectiva científica quanto cultural para a reafirmação de suas identidades neste atual contexto pós-moderno (cf. Bauman, 2005; Hall, 2019).

Este trabalho também se justifica pela escassez de estudos linguísticos sobre a língua Arutani e a necessidade de se ampliar as bases de dados descritivos na linguística para que se possa compreender melhor o fenômeno universal da linguagem e contribuir para a tipologia linguística. Tal afirmação é embasada em Rodrigues (1966, p. 4-5), em que são apresentadas algumas das principais tarefas e compromissos que a linguística deveria ter no Brasil. No caso da investigação das línguas indígenas, o autor, em seu trabalho pioneiro, expõe que o estudo de novas línguas é essencial para o avanço do conhecimento linguístico, pois cada língua investigada contribui com manifestações únicas da linguagem humana, ampliando a experiência do linguista, desafiando teorias estabelecidas e abrindo novos horizontes para a compreensão do fenômeno da linguagem. Além do mais, com a descrição de uma língua pouco conhecida, surge a possibilidade da descoberta de novos padrões linguísticos, ampliando os conhecimentos sobre o funcionamento da linguagem humana. Dessa forma, buscamos, primeiramente, uma descrição precisa dos fones e fonemas da língua, para que, com isso, possamos avançar com mais segurança na descrição e análise nos âmbitos morfológicos e sintáticos, por exemplo.

Nossa pesquisa visa abranger, de uma perspectiva fonética, a descrição e a análise acústica do Arutani. No Brasil, ainda há poucos estudos de fonética acústica sobre línguas indígenas. Isso leva a algumas problemáticas. A primeira seria a desvalorização das línguas indígenas por certas áreas da linguística. A outra seria o fato de que o não uso da fonética acústica abre a possibilidade para equívocos de análise, já que muitos pesquisadores utilizam apenas a análise perceptual (ou oitiva) como recurso para suas transcrições fonéticas. Apesar de ser um bom e tradicional recurso, uma análise oitiva pode enganar certos ouvintes, até os mais treinados, visto que nossos ouvidos estão mais condicionados a interpretar os diversos sons existentes em outras línguas com base em aproximações aos sons das línguas que já falamos. Sem contar que, como pesquisadores, devemos buscar a utilização do máximo de recursos que os avanços tecnológicos nos permitirem, tendo sempre em mente o que a tecnologia pode ou não contribuir para com as nossas pesquisas. Dessa forma, justifica-se o uso da fonética acústica neste trabalho como um recurso que colabora para que venhamos evitar o máximo de equívocos na análise e na descrição dos segmentos e suprasegmentos da língua. Com isso, evidencia-se o

fato de que os trabalhos de descrição, documentação e análise de línguas indígenas e os estudos em fonética acústica devem se complementar sempre que necessário.

Um propósito transversal, que faz parte da essência deste trabalho, é a valorização não só da língua, mas, também, do povo, da cultura e da identidade Arutani, não apenas para os que ainda permanecem e resistem, mas também por todos os ancestrais Arutani. Essa valorização irradia de maneira direta, ou indireta, para todos povos nativos deste continente, transcendendo o caráter científico deste trabalho e tornando-o, também, político, uma vez que buscamos por meio de nossa pesquisa evidenciar ainda mais a realidade multilingue do Brasil e do continente sul-americano. Uma dessas formas de valorização é o Dicionário Arutani (cf. Chacon *et al.*, 2020). Assim, a linguística torna-se mais uma vez transdisciplinar, provando ser um meio de transformação de diversas realidades. Com isso, estamos correndo contra o tempo para analisarmos e descrevermos o máximo que conseguirmos sobre a língua, levando em consideração a idade avançada de Juanita e que, com o passar dos anos, as línguas no mundo estão cada vez mais desaparecendo.

Para Juanita e os poucos Arutani que ainda existem, este trabalho servirá para eternizar sua língua na forma de uma monografia científica. Além disso, procuramos mostrar que apesar de todos os processos violentos e genocidas que ocorreram na América no período da colonização européia (cf. Quijano, 2005), ainda assim, o multilinguismo e a resistência indígena são uma realidade neste continente e, principalmente, no Brasil (cf. Garcia, 2014). Dessa forma, este trabalho oferece uma descrição da fonética e fonologia do Arutani, mas também se alinha tanto para as questões políticas de preservação das línguas, para as lutas por direitos indígenas e para a demarcação de terras, quanto para a linguística e para o aumento do número de línguas descritas no mundo.

Apesar de todas as políticas historicamente impostas e atualmente praticadas, a resistência indígenas ainda mostra seus frutos, e Juanita é um exemplo claro disso. Esta pesquisa tem como motivação trazer mais visibilidade para as reivindicações e as problemáticas enfrentadas pelos povos indígenas, reforçando que apesar de tentarem exterminar os povos que aqui já viviam, eles mostraram, novamente, que, apesar de alguns serem declarados como

praticamente extintos, ainda continuam **vivos**. O contato com a última falante de Arutani em território brasileiro reforça tudo isso, e nos mostra a força dessas pessoas e o poder que a língua tem em processos de resistência.

No Brasil, já existem algumas políticas públicas que tratam a respeito da diversidade linguística nacional. Mesmo que indiretamente, elas respaldam e evidenciam a importância de pesquisas como esta, que buscam a análise e a descrição de línguas indígenas faladas no Brasil, contribuindo, assim, para a preservação do multilinguismo nacional e evitando que casos de línguas extremamente ameaçadas como o Arutani venham continuar existindo. Uma dessas políticas é embasada no Decreto nº 7.387, de 09 de dezembro de 2010, em que foi instituído o Inventário Nacional da Diversidade Linguística (INDL). Observe o que falam os Artigos 1º e 2º deste decreto:

Art. 1º - Fica instituído o Inventário Nacional da Diversidade Linguística, sob gestão do Ministério da Cultura, como instrumento de identificação, documentação, reconhecimento e valorização das línguas portadoras de referência à identidade, à ação e à memória dos diferentes grupos formadores da sociedade brasileira. [...] Art. 2º - As línguas inventariadas deverão ter relevância para a memória, a história e a identidade dos grupos que compõem a sociedade brasileira (BRASIL, 2010).

A partir disso, observa-se que providências legais vêm sendo tomadas desde de 2010, mas desde então os povos indígenas brasileiros estão tendo que lutar cada vez mais em defesa de suas terras, de suas culturas e de suas línguas para que possam continuar existindo.

Outra política que vem sendo feita, é a de oficialização e cooficialização de línguas indígenas. Rubim (2016) explica que a cooficialização de uma língua eleva o *status* da língua indígena ao nível da língua dominante, enquanto a oficialização a reconhece formalmente a língua indígena como forma de garantir que seus falantes possam usá-la plenamente. No Brasil, a cooficialização tem fortalecido línguas indígenas e de imigração, promovendo uma política de valorização dos direitos linguísticos. Contudo, a implementação prática dessas línguas ainda enfrenta desafios: o reconhecimento jurídico muitas vezes não é seguido de ações concretas para que essas línguas tenham o mesmo destaque que o português. Em São Gabriel da Cachoeira, por exemplo, há dificuldades na aplicação das medidas legais, apesar do avanço que os povos

indígenas alcançaram em prol do respeito aos direitos das comunidades linguísticas (RUBIM, 2016, p. 86-87).

Com relação a outras leis brasileiras, pouco se avançou quando se trata dessa discussão sobre ensino de línguas indígenas como forma de contribuição para a preservação dessas línguas. Quanto à Constituição Federal de 1988, no Capítulo III que trata sobre educação, cultura e desporto em seu Art. 210 é apresentado no §2º que: “O ensino fundamental regular será ministrado em língua portuguesa, assegurada às comunidades indígenas também a utilização de suas línguas maternas e processos próprios de aprendizagem” (BRASIL, 1988). Ademais, em seu Art. 231, no Capítulo VIII destinado aos “índios” é exposto que “São reconhecidos aos índios sua organização social, costumes, línguas, crenças e tradições, e os direitos originários sobre as terras que tradicionalmente ocupam, competindo à União demarcá-las, proteger e fazer respeitar todos os seus bens” (BRASIL, 1988). Contudo, o Art. 13, desse mesmo documento, afirma que “A língua portuguesa é o idioma oficial da República Federativa do Brasil”, evidenciando a visão colonialista ainda presente nas bases legais do Brasil, o que reforça a necessidade da continuação da luta a favor políticas de oficialização e cooficialização de línguas indígenas.

Outra lei bastante importante para a educação brasileira é a Lei nº 9.394 de 1996, conhecida como Lei de Diretrizes e Bases da Educação Nacional (LDB). Esta lei apresenta que, no que tange ao ensino médio no Brasil, “O ensino da língua portuguesa e da matemática será obrigatório nos três anos do ensino médio, assegurada às comunidades indígenas, também, a utilização das respectivas línguas maternas” (BRASIL, 1996), alteração feita apenas em 2017. O Plano Nacional de Educação (PNE), elaborado decenalmente, em sua Meta 5, que é de “alfabetizar todas as crianças, no máximo, até o final do 3º (terceiro) ano do ensino fundamental” (BRASIL, 2014), apresenta como quinta estratégia:

5.5) apoiar a alfabetização de crianças do campo, indígenas, quilombolas e de populações itinerantes, com a produção de materiais didáticos específicos, e desenvolver instrumentos de acompanhamento que considerem o uso da língua materna pelas comunidades indígenas e a identidade cultural das comunidades quilombolas (BRASIL, 2014).

Tudo isso evidencia uma políticas extremamente simplistas, e muitas das vezes repetitivas, e pouco inovadoras, com relação aos direitos políticos na área da educação dos povos

indígenas do Brasil e no ensino de suas respectivas línguas. O que se espera, na verdade, são leis e políticas mais efetivas que valorizem ainda mais a diversidade linguística no Brasil e as culturas dos povos indígenas brasileiros. Políticas públicas mais assertivas no âmbito da valorização e ensino de línguas indígenas é importante para que evitemos que casos como o do Arutani se repitam e continuem acontecendo.

O desaparecimento de línguas no mundo sempre foi uma realidade, ainda mais se olharmos de uma perspectiva diacrônica. Além disso, a língua também foi constantemente utilizada, no decorrer dos séculos, como um instrumento de imposição da identidade e de opressão de outros povos, portanto, como uma forma de poder. Na história da humanidade, tem-se vários exemplos de situações em que a língua foi utilizada dessa forma, como no período da colonização europeia nas Américas. Segundo Quijano (2005, p. 117), o encontro violento entre diferentes povos na América trouxe a noção de “raça” como base para novas identidades sociais, como índios, negros e mestiços, e redefiniu outras já existentes. Termos como “espanhol” e “português”, que antes indicavam apenas origem geográfica, passaram a ter uma dimensão racial em relação às novas identidades. Com isso, essas identidades foram vinculadas a hierarquias sociais e a papéis específicos dentro de relações de dominação (QUIJANO, 2005, p. 117).

Diversos trabalhos, como de Hall (2019) e de Bauman (2005), discutem a questão da identidade do sujeito pós-moderno e apresentam diversas perspectivas e conceituações sobre o que é a identidade humana em si. Fugindo de tal presunção, não pretendemos conceituar e adentrarmos na complexa discussão sobre este tema, mas tendo como base as contribuições desses autores e de seus trabalhos, queremos, a partir disso, contextualizar essa questão tendo uma perspectiva decolonial de análise do fenômeno da linguagem. Quijano (2005, p. 127-129) ainda faz o uso de uma metáfora na qual ele explica que os americanos em geral (tanto nativos quanto descendentes de europeus, por exemplo), foram obrigados a se olhar a partir de um espelho eurocêntrico, o qual distorce o que reflete e ainda reflete de maneira parcial. Essa questão tornou-se tão problemática que, como o mesmo autor ainda apresenta, as pessoas que antes se identificavam como “astecas, maias, chimus, aimarás, incas, chibchas, etc.” passaram a

ter uma única identidade, denominada como "índios". O mesmo aconteceu com as pessoas trazidas do continente africano, que foram descaracterizadas como “achantes, iorubás, zulus, congos, bacongus, etc.” para serem identificadas como “negros”.

O processo de tornar o português a língua nacional não foi algo natural, mas sim, pensado, articulado e intencional. Rubim (2016, p. 83) destaca que: “No Brasil, houve um processo rápido de domínio político pelos colonizadores que tiveram como principal instrumento de dominação a igreja para que todos os povos existentes neste país fossem levados a abandonar suas línguas, suas culturas e suas tradições”. Ao retornarmos o que Garcia (2014, p. 90) apresenta, podemos destacar alguns fatos históricos como o caso do Nheengatu, língua originada do Tupi antigo, surgindo do contato, muitas vezes violento, entre colonizadores e indígenas no Brasil, especialmente na costa e na Amazônia. Ela foi a principal língua de comunicação no período colonial até ser proibida em 1758 pelo Marquês de Pombal, entrando em declínio desde então. O Nheengatu influenciou a toponímia do país, nomeando rios, cidades e paisagens. Hoje, poucos falantes ainda utilizam a língua, principalmente no norte do Brasil, como na região do Alto Rio Negro (GARCIA, 2014, p. 90).

Garcia (2014, p. 90-91) apresenta que, com base em dados fornecidos pelo censo de 2010 no Brasil, na época, existiam mais de 200 línguas sendo faladas no território nacional, o que demonstra a resistência desses povos apesar de toda a opressão e intenção de apagamento. O multilinguismo é uma realidade no Brasil muito pouco valorizada, e essa desvalorização, entre outros fatores, contribui para a diminuição ainda mais do número de línguas faladas em todo o mundo. Portanto, evidencia-se ainda mais a importância de políticas linguísticas nos âmbitos nacional e internacional, para que línguas como o Arutani não precisem ser documentadas às pressas antes de sua extinção.

2. Referencial teórico e metodológico

Neste capítulo, apresentaremos questões tanto teóricas, que embasam nossa pesquisa, quanto práticas, as quais referem-se às metodologias e fluxos de trabalho que utilizamos no processamento de nossos dados. Na questão teórica, apresentaremos conceitos e teorias que estão presentes em nossa pesquisa, desde conceitos basilares, como a definição de fone e fonema, até abordagens e teorias, como a teoria fonte-filtro que busca explicar como ocorre a produção dos sons da fala. De uma perspectiva mais prática, apresentaremos a metodologia de processamento e preparação dos dados para compor nosso *corpus* e, também, quais pontos da fonética e da fonologia do Arutani falado por Juanita serão descritos e analisados nesta pesquisa.

2.1 Contribuições da fonética articulatória, da fonética acústica e da fonologia

Tendo em mente a importância da definição de determinados conceitos, entendemos por fonética como a área da linguística que se dedica ao estudo dos sons da fala, podendo abordá-los sob diferentes perspectivas: do ponto de vista articulatório, investigando como os sons são formados pelo aparelho fonador; do ponto de vista acústico, analisando as suas características físicas durante a produção e transmissão; e do ponto de vista auditivo, focando-se na forma como os sons são percebidos pelo ouvinte (MATZENAUER, 2014, p. 11). Souza e Santos (2010a) apresentam o fone como a unidade de estudo da fonética. Os autores definem o fone como sendo o som da fala, o qual é segmentável e concreto. Segmentável por possuir fonação, ponto e modo de articulação e concreto pois são produzidos e percebidos acusticamente.

Já a fonologia, entendemos como a área que estuda os sistemas de sons das línguas do mundo, investigando a descrição, a estrutura e o funcionamento dessas línguas, analisando como se formam e organizam as sílabas, os morfemas, as palavras e as frases, bem como a ligação entre a mente e a língua, permitindo compreender como a comunicação ocorre de forma eficaz (MATZENAUER, 2014, p. 11). Assim como a fonética possui o fone como objeto de estudo, o objeto de estudo da fonologia é o fonema. Podemos definir os fonemas como sendo as “menores unidades linearmente segmentadas, não dotadas de significado, mas que permitem a distinção de significado” (SOUZA; SANTOS, 2010b, p. 35).

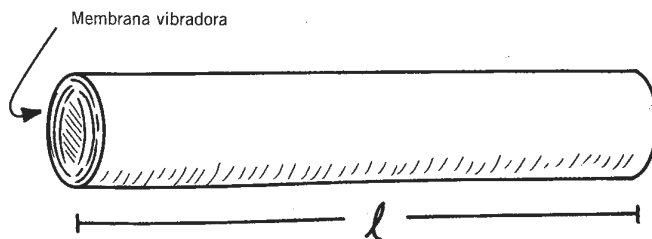
Em resumo, Matzenauer (2014, p. 12) apresenta que a fonética estuda todos os sons produzidos pelo aparelho fonador e que são utilizados na fala. Já, segundo a autora, a fonologia foca-se especificamente nos sons que têm a função de diferenciar significado, os fonemas, e na forma como estes se organizam e se combinam para formar estruturas linguísticas maiores, além de considerar as suas variações. Através do processo de comutação, que consiste na substituição de sons em contextos linguísticos semelhantes, torna-se possível identificar contrastes de significado e, assim, reconhecer os fonemas de uma determinada língua (MATZENAUER, 2014, p. 12).

Como já apresentado, existem três frentes dos estudos fonéticos: articulatória, perceptual¹⁷ e acústica. A fonética articulatória é muito importante para nossa pesquisa, pois ela permite que possamos compreender com o máximo de precisão como os segmentos da língua são articulados e produzidos, como, por exemplo, as diferenças articulatórias entre segmentos consonantais e vocálicos, os diferentes tipos de fonação, etc., para que saibamos reconhecer com o máximo de exatidão o que estamos transcrevendo. Além disso, a fonética articulatória contribui diretamente para os estudos acústicos, já que existem traços e características acústicas que são representações de correlatos articulatórios. Portanto, com a fonética articulatória nos ajudando na descrição dos segmentos e suprasegmentos, a acústica torna-se mais um recurso para uma descrição ainda mais precisa desses sons.

Uma das teorias basilares para esta pesquisa é a teoria fonte-filtro, a qual leva em consideração questões articulatórias e suas consequências acústicas. Kent e Read (2015, p. 37) apresentam essa teoria como “teoria linear fonte-filtro da produção da fala” e ressaltam o fato de ela ser importante para a compreensão das relações acústico-articulatórias, além de “fornecer fundamentos para muitos procedimentos necessários a uma análise acústica da fala e para métodos populares de síntese da fala”. Para melhor explicar a teoria, os autores usam o seguinte modelo:

¹⁷ Utilizaremos a fonética perceptual (também conhecida como auditiva ou oitiva) de maneira transversal em nossa pesquisa, já estamos sempre ouvindo os áudios da língua no momento da transcrição, mas confiamos mais na acústica como recurso para prova a existência de cada um dos sons encontrados no Arutani falado por Juanita.

Figura 4 – Um modelo simples de produção de vogais: tubo reto de seção transversal uniforme fechado em um lado (por uma membrana vibradora que simula as pregas vocais) e aberto no outro (correspondendo à abertura bucal).



Fonte: Kent e Read (2015, p. 41)

Com base nessa imagem os autores explicam que este modelo de trato é composto por uma membrana (vibrador) que é acoplada a um tubo reto. Já montado o “vibrador é esticado para se encaixar em um lado do tubo e o outro lado é deixado aberto. O vibrador é uma fonte de energia acústica que se propaga através do tubo. O tubo é um ressoador” (KENT; READ, 2015, p. 39). Fazendo um paralelo direto com o aparelho fonador, Barbosa e Madureira (2015, p. 77) explicam a **fonte**, dentro da teoria, como uma “geradora de energia para a produção de um som da fala”. De acordo com estes autores, essa fonte sonora pode localizar-se em qualquer ponto entre a glote e os lábios, podendo apresentar-se de forma periódica – como acontece com a vibração das cordas vocais – ou aperiódica, como no caso do ar que passa por uma zona estreita do trato vocal para formar uma consoante fricativa não-vozeada, como o som [s]. Além disso, também há a possibilidade da combinação entre periodicidade e aperiodicidade, como é o caso da produção de uma consoante fricativa vozeada como o [z] (BARBOSA; MADUREIRA, 2015, p. 82).

Com relação ao **filtro**, os mesmos autores explicam que o som produzido pela fonte é modificado ao atravessar o trato vocal, que funciona como um corpo ressoador. Esse trato potencializa determinadas faixas de frequência, dependendo das características transversais ao longo do eixo que vai da glote aos lábios. Assim, o trato vocal age como um filtro que favorece certas frequências sonoras, influenciando o som final da fala (BARBOSA; MADUREIRA, 2023, p. 82).

Observando a manifestação acústica dos segmentos nessa interface entre fonética articulatória e acústica, podemos chegar a análise acústica. Em nossa pesquisa, optamos por duas formas de extração de dados acústicos, a primeira é a LPC (*Linear Predictive Coding* – ou, Codificação Preditiva Linear). Barbosa e Madureira (2015, p. 154) apresentam que essa técnica é utilizada para codificar sinais de fala, reduzindo a quantidade de informação necessária para transmissão. Ela se baseia, segundo os autores, na previsão de uma amostra do sinal a partir de uma combinação linear de um número fixo de amostras anteriores. Silva *et al.* (2019, p. 76) explica que a “análise LPC aplicada a um sinal de fala permite obter a envoltória espectral e as frequências correspondentes aos formantes”, um recurso muito útil para pesquisas como esta, pois sua aplicabilidade vai desde fornecer informações para a distinção entre vogais com base nos três primeiros formantes até a distinção de segmentos oclusivos com base nos formantes de transição. Veremos a aplicação desses pontos nas seções de análise acústica.

A segunda forma de extração de dados que escolhemos foi a análise FFT (*Fast Fourier Transform* – em português, Transformada Rápida de Fourier) vem da análise de Fourier, que é a origem da representação espectral das ondas sonoras. A representação chamada de espectro de Fourier, proposta pelo matemático francês Joseph Fourier com base nas séries matemáticas que levam seu nome, permite determinar as amplitudes e as fases das ondas periódicas simples que, inevitavelmente, constituem uma onda periódica complexa. Esse procedimento de análise, segundo Barbosa e Madureira (2015, p. 149), pode ser aplicado a sinais digitalizados, utilizando amostras organizadas segundo a taxa de amostragem, e sua representação espectral final é chamada de transformada discreta de Fourier (DFT). Esse cálculo era muito demorado por volta dos anos 1970, porém descobriu-se que, caso o número de amostras na janela de análise fosse uma potência de 2, o número de operações seria reduzido pela metade, originando a Transformada Rápida de Fourier (FFT). Essa técnica é amplamente utilizada em *softwares* de análise acústica, como o Praat (BOERSMA; WEENINK, 2022), o qual permite realizar a análise FFT ao acessarmos a opção *View spectral slice* no Menu *Spectrum*. O cálculo considera uma janela temporal cuja extensão influencia a resolução espectral e pode ser ajustada pelo experimentador no menu *Spectrogram settings...*, na opção *Window length(s)*. A FFT é, assim,

essencial para viabilizar análises rápidas e precisas em estudos acústicos modernos (BARBOSA; MADUREIRA, 2015, p. 149).

Após observarmos a complexidade e a importância envolvendo tanto questões articulatórias quanto acústicas que levamos em consideração em nossa pesquisa, precisamos destacar um ponto importante que guia nosso trabalho desde o início, que é a relação direta de influência mútua entre a fonética e a fonologia. Hayes (2009) nos traz uma importante contribuição ao refletir sobre a relação direta entre os trabalhos da fonética e os trabalhos da fonologia. Ele apresenta a fonologia como uma ciência não apenas experimental, mas, também, preocupada com a análise formal e teorizações abstratas, além de defender que os dados iniciais basilares da fonologia são os dados fonéticos. O autor ainda apresenta que “O objetivo da fonologia é compreender o sistema tácito de regras que o falante utiliza para apreender e manipular os sons da sua língua” (HAYES, 2009, p. 1, tradução nossa), sons estes que estão no âmbito inicial da fonética.

Para a análise fonológica/fonêmica, temos o trabalho de Pike (1947), em que o autor apresenta a concepção do termo “fonêmica”. Para ele a fonêmica é uma forma prática de processar dados fonéticos brutos para o sistema da língua, ou seja, o caminho entre uma descrição fonética e uma análise fonológica. A fonêmica por tratar-se de um método de detecção de fonemas e alofones teve muito mais influência em nossa pesquisa no âmbito metodológico do que em comparação com o âmbito teórico. A partir disso, teremos como norte o roteiro, ou algoritmo, de análise fonêmica proposto em sua obra, adaptado por nós para a realidade do Arutani:

1. Montamos nosso *corpus* com diversas transcrições fonéticas detalhadas das palavras retiradas do Dicionário Arutani;
2. Organizamos os sons montando o quadro fonético completo do Arutani, baseado no idioleto de Junita;
3. Em seguida, organizamos os pares de sons foneticamente semelhantes, tidos como pares suspeitos;

4. Por fim, identificamos os pares mínimos e análogos, investigando questões de variação livre, distribuição complementar, neutralização e outros fenômenos fonológicos.

Tendo esse roteiro pikeano em mente, o resultado a ser encontrado, conseqüentemente, é de uma lista dos fonemas e alofones da língua, no caso, do idioleto de Juanita. É de extrema relevância a descrição dos fonemas de uma língua, tanto que Mori (2000, p. 155) ressalta a importância dos fonemas na diferenciação de significado nas palavras de uma língua em análise, como é no caso da identificação de pares mínimos e análogos. Portanto, “realizações fonéticas diferentes de um mesmo fonema não podem ocorrer em contraste”, dessa forma, “As diferentes realizações fonéticas de um fonema são conhecidas como alofones ou variantes fonéticas” (MORI, 2000, p. 155). Um ponto importante que o autor destaca é que o fenômeno da alofonia ocorre por uma questão de condicionamento de certos ambientes/contextos fonológicos, sendo chamado de distribuição complementar o modo utilizado para agrupar os alofones de um mesmo fonema. Ressaltamos que não é qualquer conjunto de sons que, em uma análise, serão tidos como suspeitos de alofonia, primeiro serão investigados os sons foneticamente semelhantes (SFS) e, em última hipótese, sons mais distintos. Quando não encontramos um contexto condicionante para a diferenciação de uso dos alofones chamamos esse fenômeno de variação livre, como sustenta Mori (2000, p. 157):

Pelo critério da distribuição complementar identificamos as variantes alofônicas de um mesmo fonema. Contudo, haverá casos em que as variações fonéticas não são contrastivas e nem estão em distribuição complementar. Nesses casos, o falante pode usar dois ou mais alofones no mesmo contexto sem destruir a identidade dos itens lexicais em questão. Quando isso ocorre, fala-se que os fones são variantes livres de um mesmo fonema (MORI, 2000, p. 157).

Esse conceito precisa estar bem claro em nossa pesquisa, pois quando analisamos o idioleto de uma pessoa, como é o caso de Juanita, precisamos entender se o que estamos analisando como alofones está sendo condicionado por um ambiente fonético-fonológico ou apenas uma variação livre na fala.

Nosso trabalho de análise fonética e fonológica do Arutani está sendo realizado no âmbito da palavra, por isso, apesar de o conceito de “palavra” ser bastante complexo, nos embasamos no trabalho de Barbosa (2019) sobre prosódia, quando ele conceitua palavra

fonológica, ou, mais especificamente, palavra prosódica. O autor define a palavra prosódica como “a unidade acima da sílaba que se pode considerar o máximo agrupamento que contém uma única sílaba portadora de acento primário e as sílabas átonas a ela associadas” (BARBOSA, 2019, p. 43). Outra contribuição para a conceituação de palavra fonológica que adotamos em nossa pesquisa está no trabalho de Aikhenvald e Dixon (2002, p. 13-14, tradução nossa), em que definem e caracterizam palavra fonológica como sendo:

[...] uma unidade fonológica maior que a sílaba (em algumas línguas pode ser no mínimo apenas uma sílaba) que tem pelo menos uma (e geralmente mais de uma) propriedade definidora fonológica escolhida entre as seguintes áreas: (a) Características segmentais – estrutura silábica interna e segmental; realizações fonéticas nesses termos; fenômenos de fronteira de palavras; fenômenos de pausa. (b) Características prosódicas – *stress* (ou acento) e/ou atribuição de tom; características prosódicas como nasalização, retroflexão, harmonia vocálica. (c) Regras fonológicas – algumas regras aplicam-se apenas dentro de uma palavra fonológica; outras (regras externas do sandhi) aplicam-se especificamente através do limite fonológico da palavra.

Teremos como norte esses dois conceitos que se complementam para o que estaremos definindo como palavras em Arutani dentro desta pesquisa.

A Fonologia Métrica é uma das principais teorias que embasam nosso trabalho, principalmente no que diz respeito ao nosso entendimento sobre o acento fonológico. Matzenauer (2014, p. 68-69) apresenta a Fonologia Métrica como um modelo teórico, o qual possibilitou uma análise mais apropriada do acento e uma nova forma de representação da sílaba, tendo como base a concepção hierárquica das estruturas linguísticas. Com base nessa teoria, existem três tipos básicos de acento que podem ser encontrados nas línguas do mundo: 1) acento primário, 2) acento secundário e 3) acento principal, sendo eles, respectivamente, segundo a autora, “o acento mais forte de uma palavra”, “o acento relativamente menos forte que o acento primário de uma palavra” e, por último, “o acento mais forte de uma sequência de palavras”. Com isso, é apresentado que para essa perspectiva o acento deixou de ser visto como propriedade do segmento e passou a ser propriedade da sílaba. Sendo assim, “somente uma sílaba pode ser portadora do acento primário”(MATZENAUER, 2014, p. 70). Dessa forma, “O acento passa a ter caráter relacional: não é mais um traço, mas uma proeminência que nasce da relação entre os elementos prosódicos: sílaba (σ), pé (Σ), palavra fonológica (ω)” (MATZENAUER, 2014, p. 70).

2.2 Origem dos dados e elaboração do *corpus*

Com relação aos áudios para a análise e descrição da fala de Juanita, todos eles foram retirados do dicionário Arutani (CHACON *et al.*, 2020). Desde 2017, o *corpus* cresceu significativamente para que o dicionário fosse produzido. As três viagens de campo feitas por Chacon foram todas no estado de Roraima, nos anos de 2017 em Boa Vista, 2019 na comunidade Saúba e 2020 em Tepequém, todas com diversas sessões de gravação com Juanita. Já temos cerca de 1.000 palavras catalogadas em Chacon *et al.* (2020) no dicionário Arutani. Temos cerca de 13 horas e 20 minutos de gravações de vídeo com dados de documentação linguística, que não serão utilizados nesta pesquisa, pois estamos trabalhando diretamente com áudios do dicionário, sem contar que, para uma descrição acústica, o uso desse tipo de arquivo não é muito adequado. Observe a listagem de horas de gravação em sessões feitas com Juanita presentes nos relatórios de campo em Chacon (2019, 2020b):

- **8h50min** de elicitación lexical e gramatical;
- **1h16min** de textos espontâneos, dos quais **1h10min** se referem a sessões que não foram gravadas como arquivos de áudio;
- **34min** de dados semiespontâneos; e
- **2h40min** de vídeos de reuniões com a comunidade Ninam em que Juanita está inserida.

Em Chacon (2019), temos um exemplo de como as sessões de gravação foram feitas e de como era o fluxo de trabalho em campo. Essas gravações foram feitas na comunidade Saúba, onde Juanita vive. Observe um de seus relatos com relação aos dados coletados e aos métodos de documentação linguística:

As sessões de gravação foram realizadas na casa de Júnior Sousa. Juanita saía de sua casa e vinha até nós nos dias que ela estava disponível. Trabalhávamos pela manhã, totalizando cerca de 4 horas, almoçávamos juntos e depois ela ia para sua casa. Pela tarde, os pesquisadores ficavam a cargo de organizar os dados, nomeá-los, transcrevê-los e analisá-los. Além das sessões de gravações, o coordenador do projeto andava com um caderninho de bolso e uma máquina fotográfica pequena que também fazia curtos filmes, de modo a permitir a documentação linguística em ambientes mais espontâneos criados a partir da metodologia de pesquisa-participativa (CHACON, 2019, p. 10).

O pesquisador ainda relata que essas sessões de gravação tinham dois principais objetivos: (1) a checagem de dados antigos coletados nos anos anteriores, e (2) a coleta de novos dados. Ele ainda apresenta que os dados gerados foram eliciações, algumas semiespontâneas e outras espontâneas, sendo utilizados os seguintes recursos metodológicos, segundo Chacon (2019, p. 10):

- **Dados de eliciação:** Itens lexicais, elementos gramaticais e fonológicos
 - Dados lexicais Arutani-Ninam coletados em 2017
 - Lista de palavras comparativa de 1.130 conceitos
 - Dicionário Visual
 - Fotografias, desenhos ou objetos do léxico cultural no contexto da comunidade
 - Corpo humano e carcaça de alguns animais
 - Questionário semântico-gramatical
- **Dados semiespontâneos:** Estímulos visuais com pequenas histórias
 - Pear story
 - Numeral Classifier stimuli
 - Totem field storyboards
 - Apprehensional morphology
 - Sequências de fotos documentando um conjunto de ações realizadas para um determinado fim
- **Dados espontâneos:** Narrativas, textos descritivos e textos instrucionais

Chacon (2019, p. 11, grifo nosso) relata que ocorreu um trabalho trilingue, que era realizado em equipe, tendo os seguintes colaboradores:

Juanita – consultora Arutani, traduzia palavras, frases e textos do Ninam para o Arutani e vice-versa; **Albino Xiriana** – consultor Ninam principal; traduzia palavras, frases e textos do português ao Ninam e vice-versa; preparou alguns questionários lexicais para a coleta de dados; **Júnior Sousa** – consultor Ninam; traduzia palavras, frases e textos do português ao Ninam e vice-versa; **Leandro** e **Alex** – pesquisadores indígenas, falantes de Ninam; faziam as transcrições das principais palavras; chegaram a preparar listas de palavras para eliciação lexical; e **Thiago Chacon** – linguista, preparava a metodologia de coleta de dados.

Com essas sessões de gravação, conseguimos construir um *corpus* composto por 345 áudios de palavras, os quais compõem o dicionário Arutani e foram utilizados para compor o *corpus* desta pesquisa. Como já mencionado, os principais fatores para esse número reduzido de dados são: a) o fato de Juanita, no começo das pesquisas em 2017, ter tido dificuldades para lembrar certas palavras por conta da falta de prática no uso da língua; b) a idade avançada de Juanita e, por consequência, o fato dela não aguentar sessões de gravação muito longas; e c) a baixa qualidade de isolamento acústico dos áudios, por conta de serem gravados em meio à comunidade, o que demanda muito mais tempo de análise para que se tenha uma descrição acústica mais precisa e, conseqüentemente, a existência de áudios inviáveis para uma análise acústica, seja por terem mais de um colaborador falando ao mesmo tempo ou pelo excesso de ruídos.

2.3 Processamento e preparação dos dados

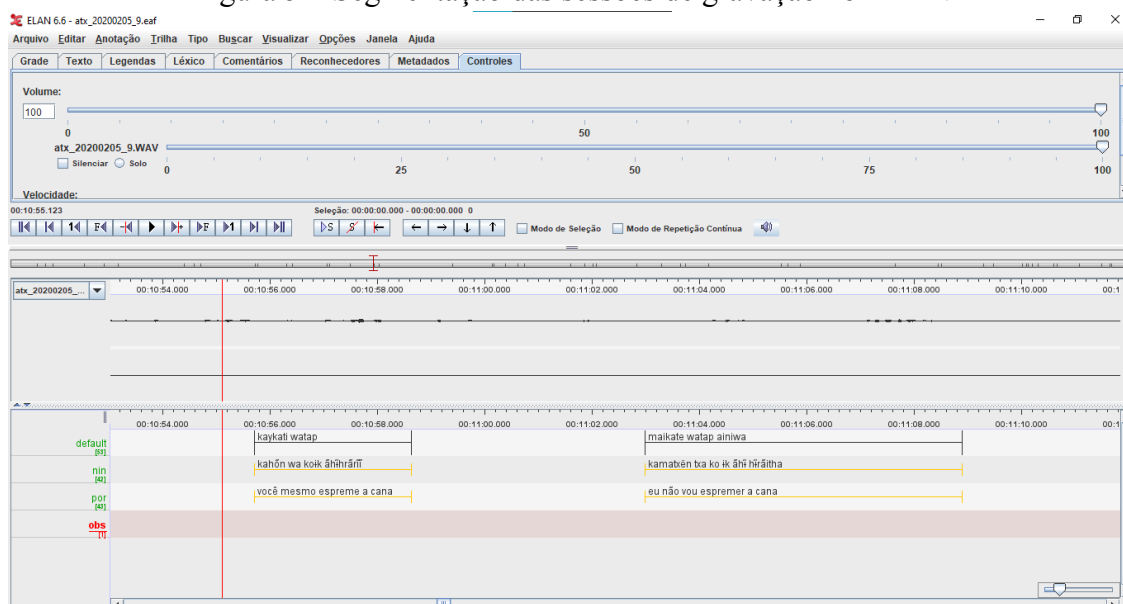
Todas essas questões conceituais e teóricas precisaram ser levadas em consideração quando estávamos ouvindo e fazendo o processamento e a preparação dos áudios produzidos nas sessões de gravação. Eles foram nomeados e catalogados com uma notação em que fosse possível ter várias informações sobre os áudios. Os metadados, como no exemplo “atx_20200408_tcc_1.wav”, eram os seguintes:

- atx: ISO da língua;
- 20200408: data da sessão (ano/mês/dia);
- tcc: responsável pela gravação (Thiago Costa Chacon, no caso);
- 1: número sequencial.

Essa lógica de nomeação é a encontrada em todos os áudios que compõem o dicionário Arutani. Já apresentamos como ocorreram as sessões de gravação com Juanita, a partir de agora, descreveremos as etapas de trabalho realizadas depois das viagens de campo. Essas etapas foram necessárias para chegarmos aos áudios finais que disponibilizamos no dicionário, os quais compõem tanto o dicionário quanto o *corpus* desta pesquisa. Para as transcrições das sessões de gravação, utilizamos alguns *softwares*. O primeiro deles foi o ELAN (2022), que foi usado para

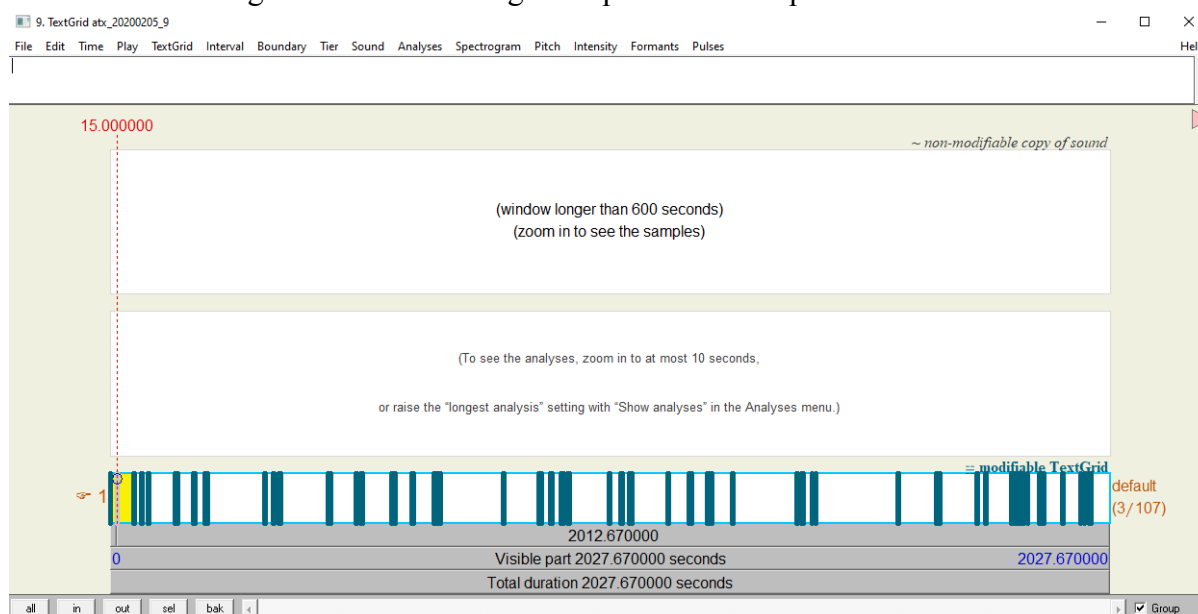
transcrições, traduções e notas iniciais. Essas transcrições foram feitas a partir de uma ortografia prática definida previamente. Segue um exemplo de como esse trabalho foi realizado:

Figura 5 – Segmentação das sessões de gravação no ELAN

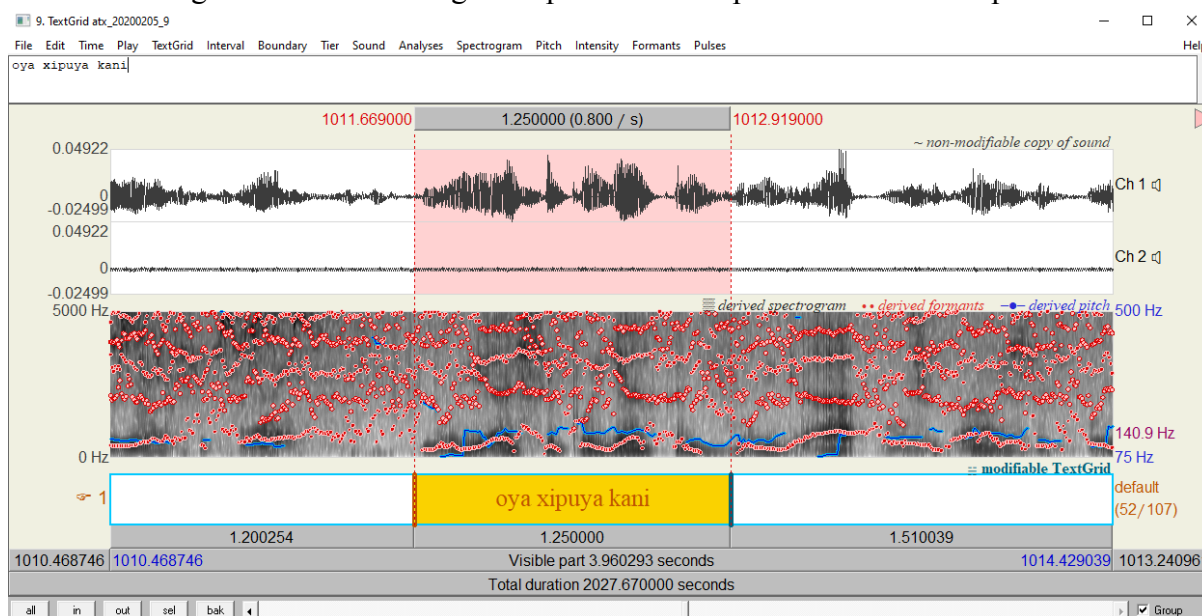


No ELAN (2022), como pode ser observado, é possível segmentar os trechos do áudio em que se tem momentos de fala da Juanita em Arutani. A primeira linha de transcrição intitulada “*default*” é onde fazemos a transcrição ortográfica da fala de Juanita. A segunda linha, intitulada “*nin*”, é a tradução para Ninam; a terceira, chamada “*por*”, é a tradução para Português; e a quarta com a sigla “*obs*” para algumas observações e anotações relevantes. Após essas transcrições iniciais, existem dois possíveis passos seguintes. O primeiro seria gerar um TextGrid – formato de arquivo que o Praat (BOERSMA; WEENINK, 2022) lê – com todas as informações do que foi transcrito na primeira linha para ser aberto no próprio Praat e em seguida salvar os diferentes trechos transcritos em arquivos de áudios diferentes em formato “.wav”. Assim sendo, os trechos transcritos em Arutani como *kaykati watap* ‘você mesmo espreme a cana’ e *maikate watap ainiwa* ‘eu vou espremer a cana’ corresponderão, respectivamente, a dois áudios diferentes. Tudo isso graças às ferramentas do Praat. Veja um exemplo desse mesmo áudio já exportado para o Praat:

Figura 6.1 – TextGrid gerado pelo ELAN exportado no Praat



Já com o zoom, conseguimos visualizar no Praat um dos vários trechos anteriormente segmentados no ELAN:

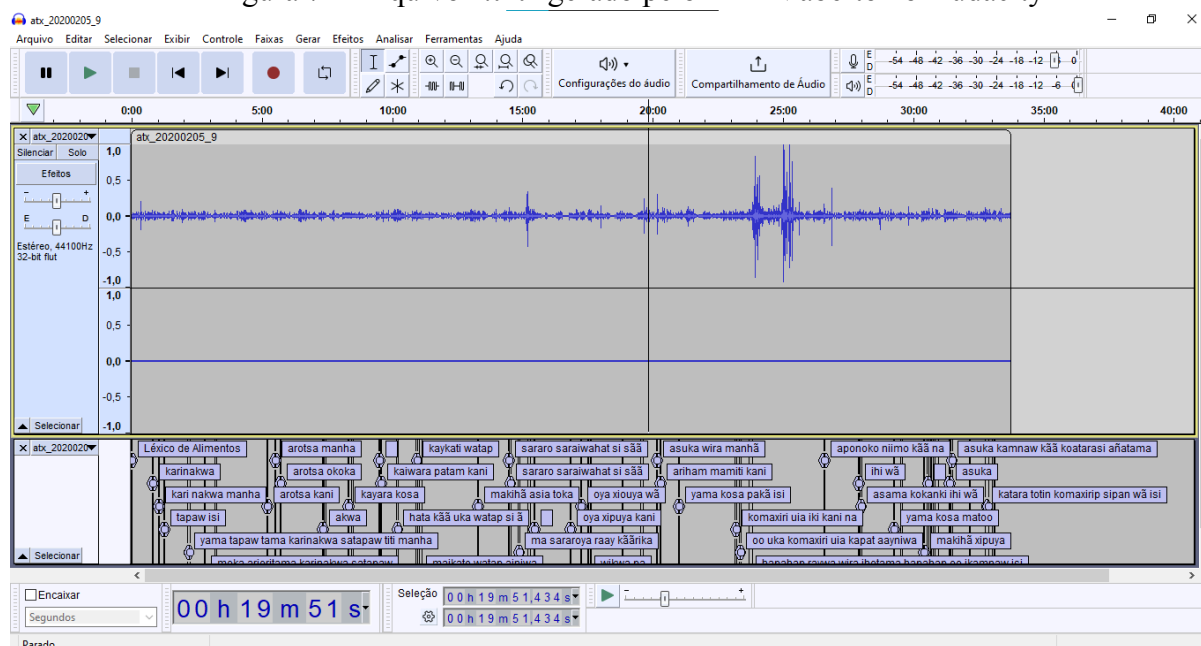
Figura 6.2 – TextGrid gerado pelo ELAN exportado no Praat ampliado¹⁸

¹⁸ No trecho em destaque temos Juanita falando *oya xipuya kani*. A palavra *oya* é o termo genérico para pimenta e *xipuya* a palavra para ‘pó’, *oya xipuya*, no entanto, seria o nome para a pimenta ‘juquitaia’. Já *kani* é um verbo de cópula que indica estado.

Após isso, utilizamos alguns *scripts* que fazem automaticamente a separação desses trechos e salvam cada um como clipes de áudio e arquivos de anotações para cada clipe no diretório que indicarmos. Esses *scripts* estão disponíveis no *GitHub* para o acesso público, o que contribui para a democratização desse tipo de ferramenta e, também, para o avanço de novas pesquisas na área. Os dois *scripts* que utilizamos estão disponíveis em Cesine (2006) e Cesine (2014).

Todavia, o Praat não consegue fazer a leitura de áudios muito longos, então, é necessário que pelo ELAN seja extraído um arquivo no formato “.txt” da linha “default”, que será, no caso, a linha que contém as transcrições que queremos utilizar. Com isso, precisamos apenas abrir o áudio no *Audacity* e importar o respectivo arquivo “.txt”. Veja como é a visualização no programa:

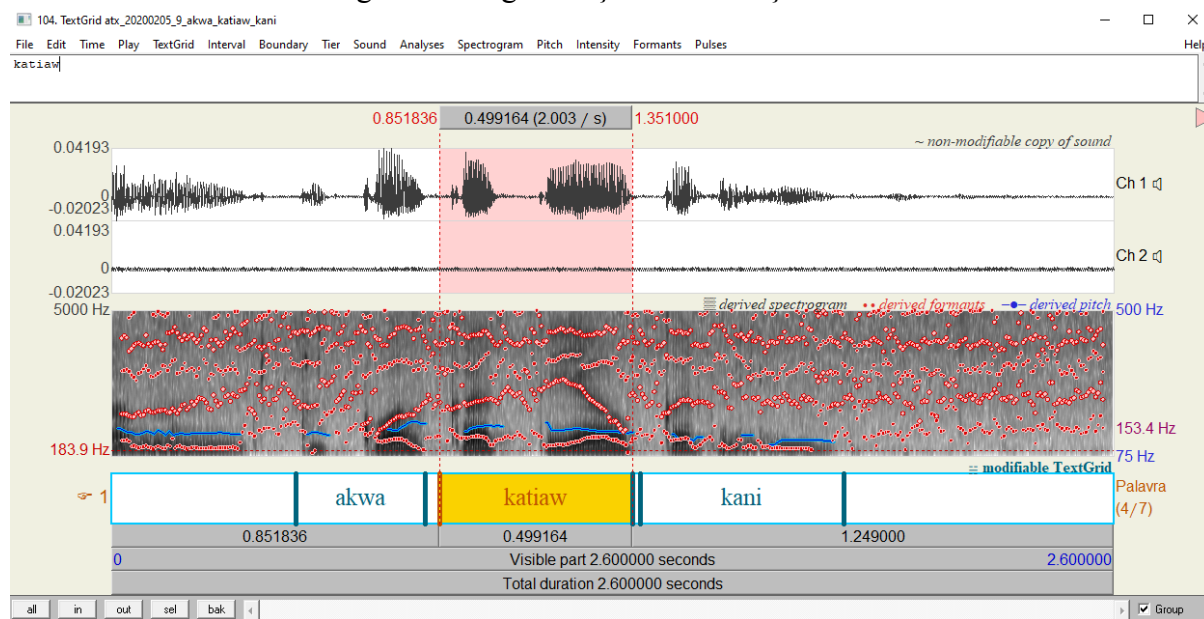
Figura 7 – Arquivo “.txt” gerado pelo ELAN aberto no Audacity



Após essa etapa, apenas exportamos esses áudios para seu respectivo diretório. Esse áudio será segmentado tendo como referência as segmentações feitas no ELAN. O resultado dessa segmentação será vários áudios de cada um dos trechos selecionados, ou seja, o produto final desse processo será o mesmo do feito no Praat com áudios menores. Após termos todas as

frases segmentadas, conseguimos segmentar e salvar áudios com palavras isoladas a partir desses áudios de frases. Para isso, foi preciso que abríssemos um desses áudios de frases segmentadas através do Praat e gerássemos um *TextGrid*. Em seguida, segmentamos as palavras que compunham essas frases. Observe:

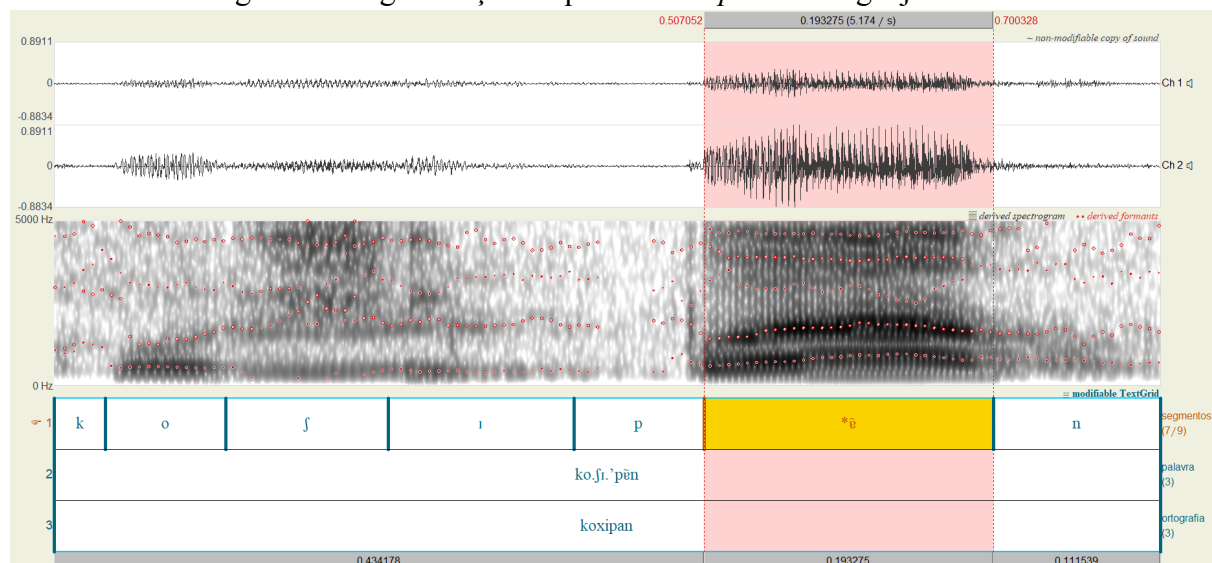
Figura 8 – Segmentação da sentença no Praat¹⁹



Na próxima etapa, salvamos cada uma das palavras selecionadas individualmente, através do próprio Praat e adicionamos esses áudios ao dicionário. Após isso, extraímos todos os áudios do dicionário e seguimos para a parte mais importante desse fluxo de trabalho: as transcrições fonéticas no Praat e a segmentação dos fones de cada uma dessas palavras segmentadas. Observe como são feitas essas transcrições:

¹⁹ O trecho segmentado, *akwa katiaw kani*, significa ‘há muitas batatas-doce’.

Figura 9 – Segmentação da palavra *koxipan* ‘caranguejo’ no Praat



Nesta imagem, podemos observar três linhas: 1) segmento: na qual segmentamos todos os fones encontrados na transcrição, em que destacamos com um * o seguimento tônico; 2) palavra: em que fizemos a transcrição com base no alfabeto da *International Phonetic Association* (IPA); e 3) ortografia: na qual seguimos as convenções do dicionário. Na segunda linha, tentamos ser o mais minuciosos possível, tanto na escolha dos fones com base no que ouvimos e no que pudemos analisar na imagem acústica (formato de onda e espectrograma), quanto com o uso de diacríticos.

Após a realização de cada uma dessas transcrições e segmentações no Praat, salvamos o áudio juntamente com seu respectivo *TextGrid*. Em seguida, copiamos as transcrições em uma tabela que também foi extraída do dicionário. Dessa forma, confeccionamos a seguinte tabela:

Figura 10 – Tabela extraída do Dicionário Arutani e com as transcrições feitas com base nas análises acústicas

1	Portal Japiim ProDoclin				
2	Tabela arutani_0002_100_palavras. Transcritor:				
3	Língua	Ortográfica	Fonética	Arquivo de Áudio	Glosa
39	Arutani	maxikito	mej.kt.'to	atx_20171207_tcc_47.wav	Cutiaia.
40	Arutani	mano	me.'nõ	atx_20171207_tcc_42.wav	Anta.
41	Arutani	marakupaw	me.re.ku'paw	atx_20171207_tcc_44.wav	Espécie de rã comestível.
42	Arutani	marakana	me.re.ke'ně	atx_20171207_tcc_43.wav	Maracanã.
43	Arutani	mariko	me.r.'ko:	atx_20171207_tcc_45.wav	Caxiri.
44	Arutani	oko	o.'ko	atx_20171207_tcc_48.wav	Zogue-zogue.
45	Arutani	pari pari	pe. r.pe.'ri	atx_20171207_tcc_50.wav	Araçari-Mulato.
46	Arutani	pasa	pe.'sa	atx_20171207_tcc_51.wav	Jundiá.
47	Arutani	pixa	pi.'ja	atx_20171207_tcc_52.wav	Cuxiú.
48	Arutani	amohã	e.mo.'hě	atx_20171207_tcc_4.wav	Curica.
49	Arutani	oti	o'ti	atx_20171207_tcc_49.wav	Urubu.
50	Arutani	xapi kuyta	fe. pi kuj.'ta	atx_20171209_tcc_21.wav	Flor.
51	Arutani	xapi aña	fe. pi.e.'ně	atx_20171209_tcc_30.wav	Folha de árvore.
52	Arutani	xomhay	fom.'haj	atx_20171209_tcc_31.wav	Acima.

A partir desta tabela conseguimos fazer diversas comparações e análises de caráter tanto fonético quanto fonológico. Foi possível, graças à elaboração desta tabela, a identificação dos fones presentes na fala de Juanita, a análise dos padrões silábicos e a realização do contraste entre os fones com o intuito de identificarmos os fonemas e outros fenômenos fonológicos.

2.4 Descrição e análise dos dados

Essa última etapa apresentada na seção anterior é, também, a primeira etapa relacionada diretamente a esta pesquisa, que é a de (1) **transcrições fonéticas e segmentação dos fones no Praat**. Mas essa, de longe, não é a etapa inicial do nosso trabalho, como já está claro, pois, para que esta pesquisa fosse possível, o trabalho vem ocorrendo desde o primeiro contato do pesquisador Dr. Thiago Chacon com Juanita em 2017. Com todas as transcrições finalizadas, elas foram inseridas no *corpus* desta pesquisa e, então, iniciamos a segunda etapa: (2) **identificação dos segmentos**. Nesta etapa, ocorreu a análise e identificação dos segmentos encontrados nas transcrições, com o objetivo de fazer a listagem das consoantes, vogais e encontros vocálicos. Esta etapa teve como foco a relação entre fonética articulatória e acústica, com uma ênfase maior na acústica. Para uma identificação mais precisa e que seguisse os mesmos parâmetros,

realizamos a extração de informações acústicas como duração, *pitch*, intensidade e os valores de frequência dos três primeiros formantes utilizando dois *scripts* (HIRST, 2012; HIRST, 2009 – disponibilizamos os dois *scripts* e seus respectivos *outputs* nos Anexos I, II, III e IV). Utilizamos dois *scripts* como uma forma de verificarmos se os dados extraídos pelo segundo *script* estavam próximos aos do primeiro, e vice-versa. Portanto, os dois *scripts* fizeram a extração dos dados acústicos por meio da análise LPC, apresentando pouca diferença nos valores extraídos, sendo utilizados objetivando um maior rigor metodológico.

Quanto à análise acústica dos segmentos, utilizamos o Praat como recurso para a realização de tal tarefa. Dessa forma, utilizamos três representações gráficas do sinal de fala, as quais são as mais relevantes para a descrição acústica dos sons das línguas naturais e nos fornecem diversas informações relevantes para o contraste entre segmentos, evidenciando os traços acústicos característicos dos diferentes tipos de sons. São elas: a forma de onda, que relaciona tempo (ms) com amplitude (variação de pressão relativa normalizada); o espectrograma, que faz a relação entre tempo (ms) e frequência (Hz); e o espectro, que relaciona a frequência (Hz) com a amplitude (dB) (cf. Silva *et al.*, 2019, p. 37). Vale ressaltar que o espectrograma também nos proporciona a visualização da intensidade, sendo tridimensional (tempo, frequência e intensidade), assim como é apresentado por Kent e Read (2015, p. 105).

Além disso, ao tratar-se da descrição das vogais orais, foram levadas em consideração as seguintes características acústicas: 1. **frequência dos formantes** – F1, F2 e F3; 2. ***pitch***; 3. **duração**; e 4. **intensidade**. A partir disso, foi possível elaborarmos o quadro vocálico da língua. Além das vogais orais, também descrevemos acusticamente as vogais nasais e nasalizadas, tendo como princípio de análise as seguintes características: formantes nasais e orais, *pitch*, intensidade, duração, antiformantes, amortecimento, entre outras características. Com relação aos ditongos, eles serão diferenciados, principalmente, com base na transição formântica e os hiatos por meio de estratégias de marcação de fronteira silábica.

Para a análise e descrição das consoantes, iremos levar em consideração todos os traços acústicos característicos de cada um dos grupos consonantais. Agrupamos as consoantes em subgrupos levando em consideração o modo de articulação, sendo eles, com base no idioleto de

Juanita: oclusivo, nasal, fricativo e *tepe*. Um ponto crucial é a comparação e o contraste entre os fones encontrados em cada um desses grupos, já que esta é a forma mais precisa para demonstrar que se trata de sons distintos. Com isso, partiremos para descrições e análises de caráter mais fonológico.

A seguinte etapa de trabalho foi a (3) **análise fonotática e padrões silábicos**, na qual analisamos e descrevemos as combinações e restrições entre os sons para a definição dos padrões silábicos do Arutani, introduzindo, assim, o início dos estudos fonêmicos/fonológicos. Na etapa (4) **análise fonêmica**, ocorreu a descrição e identificação de fonemas e alofones da língua, a partir da permutação de fones e a verificação da distinção de significado. Também realizamos a análise de alguns fenômenos fonológico. Apenas após todas essas etapas, foi possível partir para a (5) **descrição do acento**. Nesta etapa, o foco foi em uma perspectiva fonológica, já que as características acústicas do acento foram exploradas na parte sobre vogais em que já foi feita a comparação entre vogais tônicas e átonas levando em consideração duração, *pitch* e intensidade.

3. Análise Acústica

Os sons encontrados em nossas análises e que serão discutidos nas seções a seguir são fruto de análises perceptivas e, posteriormente, reanalisadas com base na análise acústicas da fala de Juanita. Lembrando que para a elaboração dos seguintes quadros estamos levando em consideração todos os fones que compõem nosso *corpus*. Eles não são os fonemas propriamente, os quais serão apresentados na seção sobre a fonologia da língua e que, como esperado, apresentam um número mais reduzido. Portanto, há uma tendência nas línguas do mundo e sempre se espera que haja mais fones do que fonemas em si.

Sendo assim, apresentaremos quadros dos segmentos consonantais, dos segmentos vocálicos orais, dos segmentos vocálicos nasais e dos ditongos existentes na fala de Juanita.

Tabela 2 – Quadro fonético das consoantes

	Tipo de Fonação	Labial	Alveolar	Pós-alveolar	Palatal	Velar	Velar Labializado	Velar Palatalizado	Glotal
Oclusiva	Não-vozeada	[p] [p̥]	[t] [t̥]			[k]	[kʷ]	[kʲ]	[ʔ]
	Vozeada	[b] [b̥]							
Nasal	Vozeada	[m]	[n]		[ɲ]				
Tepe	Vozeada		[r]						
Fricativa	Não-vozeada		[s]	[ʃ]					[h]
Aproximante	Vozeada	[w]			[j]				

Tabela 3 – Quadro fonético das vogais orais

	Anteriores	Centrais	Posteriores
Altas	i		
		ɪ	ʊ
Médias-altas	e		o
Médias-baixas			
		ɐ	
Baixas		a	

Tabela 4 – Quadro fonético das vogais nasais e nasalizadas

	Anteriores	Centrais	Posteriores
Altas	ĩ		
		ĩ	õ
Médias-altas	ẽ		õ
Médias-baixas			
		ẽ	
Baixas		ã	

Tabela 5 – Quadro fonético dos ditongos crescentes e decrescentes encontrados

Ditongos Crescentes		Ditongos Decrescentes
[ja]	[jo]	[aj] ~ [ãj]
[jæ]	[wa]	[ɐj]
[je] ~ [jẽ]	[wɐ] ~ [wẽ]	[aw]
[jo]	[wɪ]	[ɛw]
		[oj]

3.1 Segmentos Vocálicos

Em Arutani, mais especificamente na fala de Juanita, foram encontradas vogais orais e nasais, tanto em posição tônica quanto em posição átona, também encontramos em nosso *corpus* alguns ditongos (crescentes e decrescentes) e hiatos. Nas seções seguintes, iremos analisar acusticamente esses segmentos vocálicos com o objetivos de diferenciá-los entre si montarmos um quadro fonético das vogais da língua. Isso servirá de base para que seja elaborado, também, o quadro fonológico das vogais da língua. Sendo assim, iniciaremos com a descrição das vogais orais (tônicas e átonas), em seguida as nasais e, por fim, os ditongos e hiatos encontrados no idioleto de Juanita.

3.1.1 Tratamento dos Dados

Para a análise e descrição acústica das vogais utilizamos dois *scripts* (Anexo I e III), os quais realizaram a extração dos dados acústicos por meio do Praat. As informações acústicas extraídas pelos os dois *scripts* foram: duração, *pitch* e intensidade, os três primeiros formantes (F1, F2 e F3). Dessa forma, os dois *scripts* extraíram os mesmos dados, mas utilizamos os dois como segunda prova de que os dados extraídos estavam corretos, dessa forma, teríamos mais confiança e respaldo em nossa análise. Por fim, escolhemos o *script* que denominamos de *Script* 1, já que, como veremos, os dois *scripts* extraíram informações muito semelhantes. O produto desses *scripts* é uma tabela no formato “.txt” que pode ser convertida em um tabela no formato “.xlsx”, por exemplo.

Sobre os *scripts*, eles extraíram os dados com base numa análise LPC, a qual “aplicada a um sinal de fala permite obter a envoltória espectral e as frequências correspondentes aos formantes” (SILVA *et al.*, 2019, p. 76). Silva *et al.* (2019, p. 77) ainda destacam que a análise LPC baseia-se na previsibilidade parcial de cada amostra do sinal de fala a partir das amostras digitais imediatamente anteriores. Os autores explicam que ela representa o sinal como uma combinação linear dessas amostras passadas e modela apenas a envoltória espectral, desconsiderando o sinal de excitação do trato vocal. Essa análise é mais frequentemente aplicada a sinais periódicos, como sons vocálicos orais, como os que estamos analisando, devido à sua

maior regularidade. Para utilizá-la em sinais digitais de fala, é necessário determinar o número de formantes desejados, considerando a taxa de amostragem do sinal (SILVA *et al.*, 2019, p. 77) Lembramos que a análise LPC é a forma de análise padrão do Praat, o que também facilita todo o processo de extração de informações acústicas.

Após alguns ajustes de *layout*, temos o seguinte exemplo de um trecho de como a tabela ficou após um filtro para que fosse possível analisarmos apenas as vogais tônicas [a], a título de exemplo:

Tabela 6 – *Layout* e trecho da tabela produto do *Script* 1

ID	<i>extract</i>	<i>label</i>	<i>duration</i>	<i>pitch</i>	<i>intensity</i>	F1	F2	F3
323	atx_20171207_tcc_38	*a	221	193	77	876	1495	2996
1019	atx_20171209_tcc_25	*a	192	215	69	864	1775	2725
72	atx_20171206_tcc_14	*a	202	226	76	862	1420	2265
1976	atx_20200115_tcc_85	*a	124	204	72	861	1871	2661
966	atx_20171209_tcc_18	*a	210	252	67	858	1592	2117

Tendo essas informações para cada vogal, geramos gráficos em R, por meio do RStudio (disponível em McCloy, 2016), com os valores de F1 e F2 das vogais tônicas. Na Figura 11 temos o gráfico produto da compilação dos dados do *Script* 1 e na Figura 12 do *Script* 2.

Figura 11 – Quadro de dispersão inicial das vogais tônicas orais (*Script 1*)²⁰

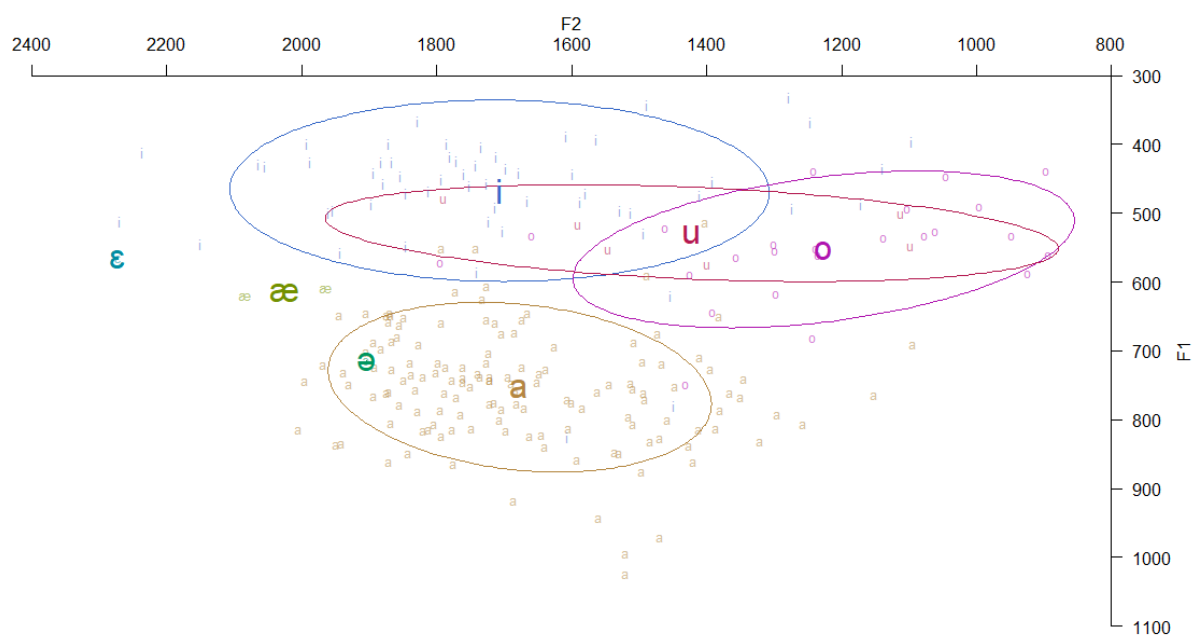
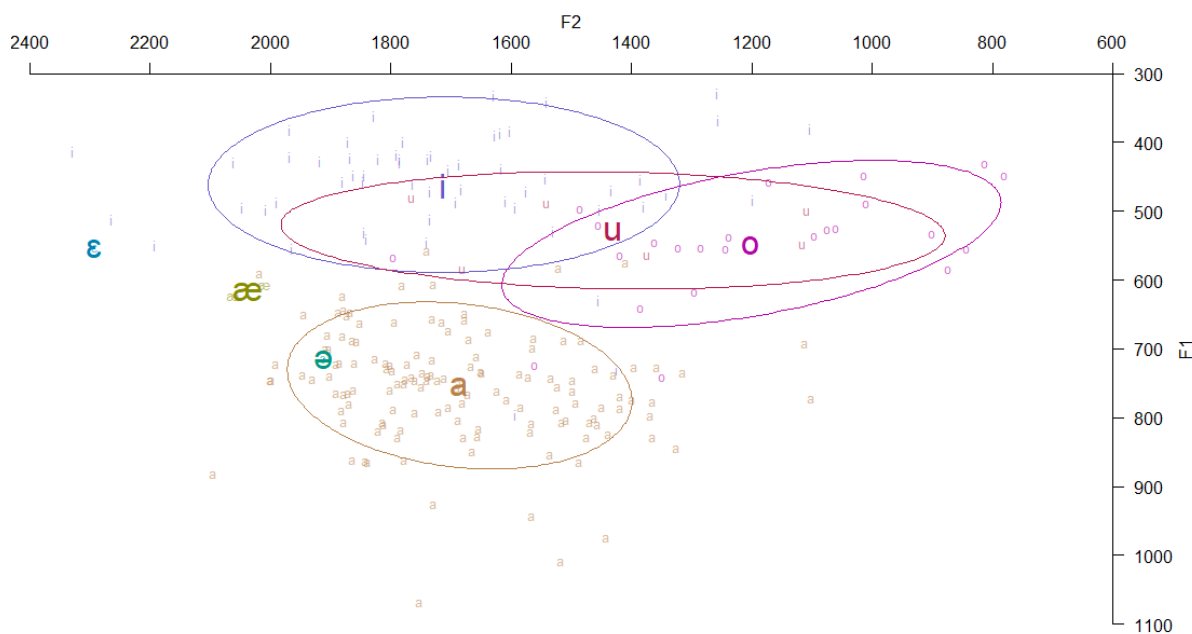


Figura 12 – Quadro de dispersão inicial das vogais tônicas orais (*Script 2*)



²⁰ Aparenta-se existir, tanto na Figura 11 quanto na Figura 12, dois de [a], entretanto, de uma perspectiva oitiva, não há nenhum fone mais preciso para o que está sendo analisado na fala de de Junita em contexto átono do que [a].

Observe que são gráficos muito parecidos com relação à dispersão das vogais. As elipses são a representação no plano do desvio-padrão dos valores de F1 e F2. Como se pode ver, algumas vogais estavam bem distantes de seu lugar esperado tendo como modelos o quadro vocálico do IPA e os símbolos fonéticos usados em nossa transcrição. Para avaliar essas inconsistências, fizemos uma seleção das vogais que estavam muito distantes da média e das que tivessem pouca ocorrência e/ou estivessem em regiões de muitas intersecções de elipses, como ilustrado nas Figuras 13 e 14.

Figura 13 – Quadro de dispersão inicial das vogais tônicas orais com marcações para reanálise (*Script 1*)

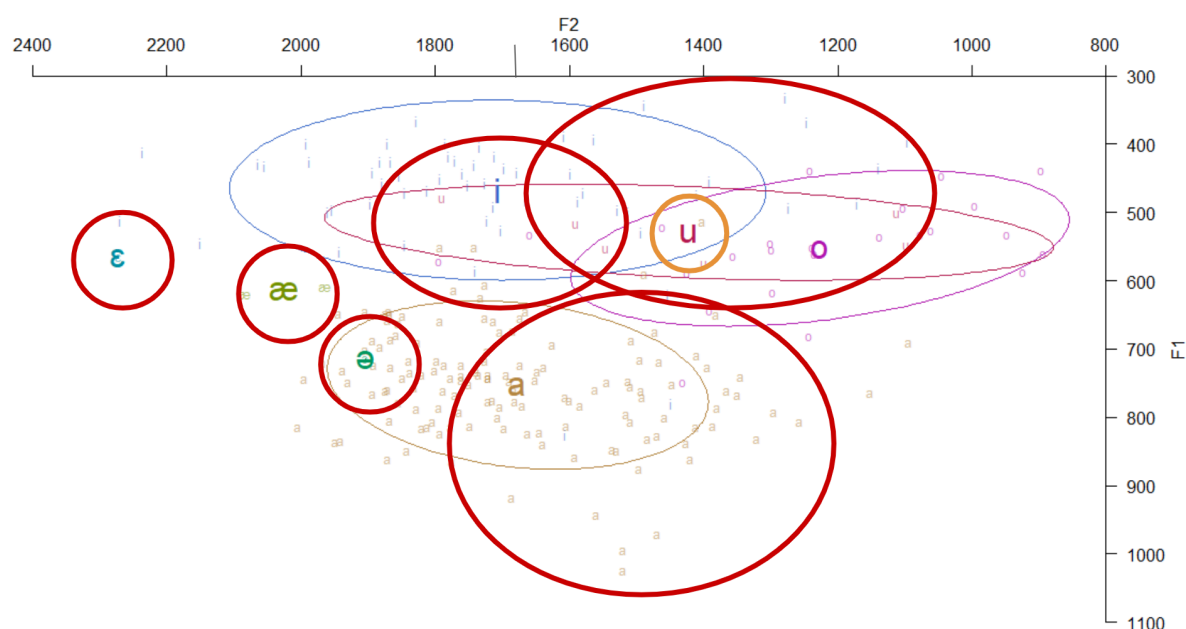
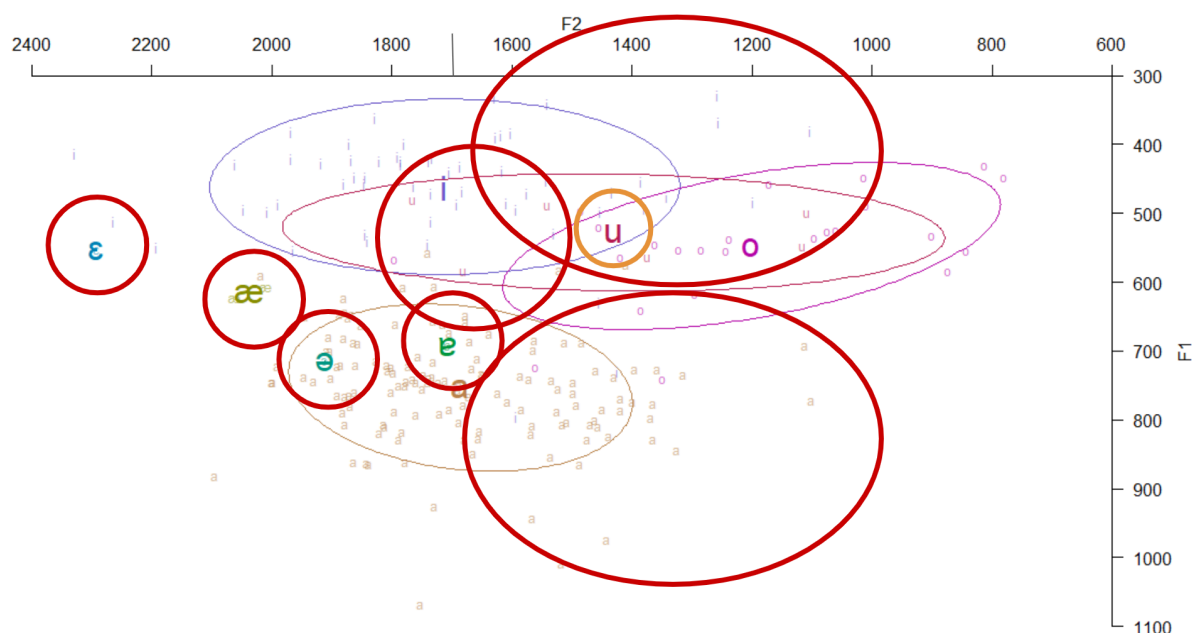


Figura 14 – Quadro de dispersão inicial das vogais tônicas orais com marcações para reanálise (*Script 2*)



Voltamos aos áudios das palavras que continham essas vogais e, após reescutá-los cuidadosamente, fizemos as seguintes correções:

- O que antes transcrevemos com [ə] foi mudado para [a];
- O [ɛ] na verdade fazia parte de um ditongo, e foi mudado para [jæ];
- Em uma das duas ocorrências de [æ] observamos que ela era na verdade um ditongo formado por [jæ] e a outra ocorrência era, na verdade, um hiato formado por [ɪ.æ]. Por esses motivos, não entram mais no grupo das vogais tônicas orais, mas dos ditongos e hiatos (esse processo será melhor explorado na seção 4.2.1); e
- Por fim, optamos por mudar o [u] para [ʊ], que apesar de serem sons próximos, o gráfico nos mostrou tratar-se de um vogal mais centralizada, apesar de na percepção não serem tão diferentes. Portanto, decidimos, objetivando uma maior precisão, que tanto de uma perspectiva acústica quanto perceptual o som mais próximo ao que foi encontrado na fala de Juanita seria o [ʊ].

Essas mudanças foram feitas em todas as tabelas. Entretanto, ainda havia alguns *outliers*, que, no gráfico, são observados como valores discrepantes de F1 e/ou F2, que estavam muito fora do esperado. Como já havíamos ouvido novamente os áudios e tínhamos certeza de estavam transcritos da maneira correta, estava na hora de fazermos uma análise mais minuciosa, que seria uma análise FFT, já que a análise LPC estava nos entregando valores discrepantes e fora do esperado.

Fizemos uma análise FFT com bases nos espectros gerados das vogais que tinham valores fora do esperado e muito distantes das médias. Esse foi nosso primeiro recurso para corrigir possíveis equívocos dos *scripts* utilizados na leitura dos formantes, já que a etapa anterior foi de escutar novamente os áudios e corrigir certas falhas que poderíamos ter cometido nas transcrições. As vogais selecionadas que analisamos novamente somam ao todo 90 vogais suspeitas, sendo 35 de [a], 32 de [i], 17 de [o] e 6 de [u]. Lembrando que, ainda, não havíamos utilizado nenhum recurso estatístico para a seleção dessas vogais, foi apenas uma análise exploratória dos valores muito altos ou muito baixos de F1 e/ou F2 e com base numa análise visual dos gráficos.

A partir disso, chegamos aos resultados da Tabela 7. Nela colocamos apenas uma amostra de apenas cinco exemplos de cada uma das vogais analisadas, apenas como forma de explicar e explicar de maneira mais visual a diferença dos resultados. Escolhemos mostrar apenas cinco de cada vogal por uma questão de objetividade e concisão. Na tabela, temos, por exemplo, a coluna “F1”, que representa os valores do *Script* 1 extraído pela análise LPC e em “F1 - FFT” temos os valores extraídos da análise FFT, fizemos o mesmo com os outros formantes.

Tabela 7 – Comparação dos valores de formantes extraído pela análise LPC e pela análise FFT

ID (wav;TextGrid)	label	F1	F1 - FFT	F2	F2 - FFT	F3	F3 - FFT
atx_20200115_tcc_92	a	716	769	1.376	1193	2313	1766
atx_20190808_tcc_2	a	650	640	1381	1066	2516	1492
atx_20200115_tcc_58	a	649	684	1873	1808	2905	3547
atx_20171207_tcc_19	a	646	599	1904	1957	2865	2996

atx_20200115_tcc_24	a	648	642	1943	1978	3079	2232
atx_20200115_tcc_18	i	396	387	1095	1172	2910	2911
atx_20171217_tcc_12	i	435	239	1138	1426	2874	2784
atx_20200115_tcc_114	i	486	215	1171	1183	2612	2602
atx_20200115_tcc_36	i	367	366	1246	960	2837	2890
atx_20200115_tcc_108	i	492	451	1273	1554	2448	2635
atx_20200115_tcc_119	o	548	557	1098	1129	2638	2423
atx_20200115_tcc_56	o	501	514	1111	1066	2845	2657
atx_20171208_tcc_23	o	575	514	1399	1151	2427	2169
atx_20171207_tcc_53	o	553	451	1545	1278	2807	2105
atx_20171207_tcc_24	o	517	447	1590	1557	2857	2132
atx_20171207_tcc_48	o	571	563	1794	1918	2820	3054
atx_20200115_tcc_39	o	532	535	1658	1532	3063	3123
atx_20171217_tcc_6	o	748	602	1431	1376	2846	2448
atx_20171207_tcc_37	o	644	684	1390	1214	2657	2529
atx_20171208_tcc_32	o	617	620	1296	1172	3097	3208

Com base na comparação entre os valores extraídos pela análise LPC e pela análise FFT, foi possível observar uma baixa diferença entre os valores encontrados. Isso sugere que os valores encontrados por meio da análise LPC com o *script* estão relativamente correspondentes ao da análise FFT, o que confirma Silva *et al* (2019, p. 77) quando explicam que “embora os algoritmos de análises LPC e FFT sejam distintos, os resultados obtidos através da aplicação de ambos são correspondentes”. Ainda, isso sugere que os valores muito discrepantes podem ter sido resultado da alta quantidade de ruídos em muitos dos áudios.

Sem muita diferença entre os valores, isso nos mostrou que, além de estarem corretos, precisaríamos buscar recursos de mineração de dados e estatística para a eliminação adequada desses possíveis erros de extração e de limitação de leitura do Praat para que pudéssemos chegar a resultados mais consistentes. A partir disso, geramos e analisamos histogramas de cada um dos grupos de vogais para entendermos o motivo de as elipses estarem com valores de desvio-padrão elevados. Nos dois exemplos das Figuras 15 e 16 temos, respectivamente, os histogramas de F1 e

F2 de [a] e de [i]. Observe que, com base em Silva *et al.* (2016, p. 39), o gráfico de F1 de [a] está com uma distribuição mais simétrica, enquanto o de F2 está com uma distribuição assimétrica à direita. Já em relação aos histogramas de [i], temos o gráfico de F1 com uma distribuição assimétrica à esquerda e de F2 com uma distribuição levemente assimétrica à direita F2.

Figura 15 – Histogramas de F1 e F2 de [a]

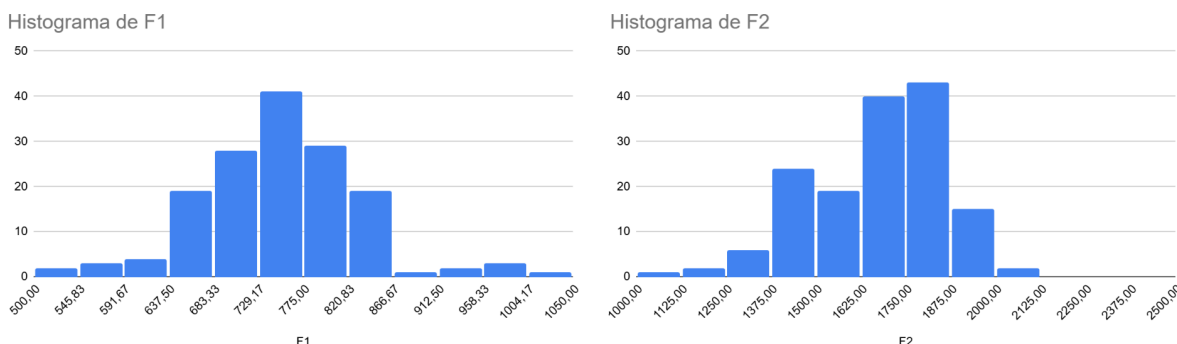
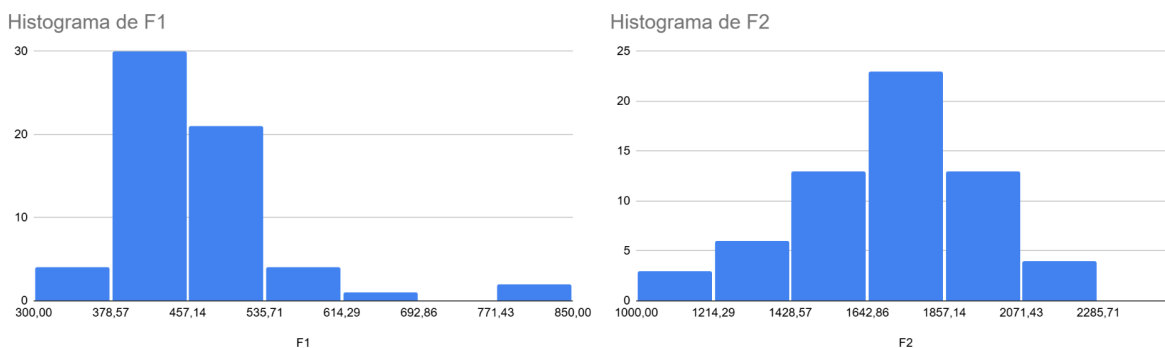
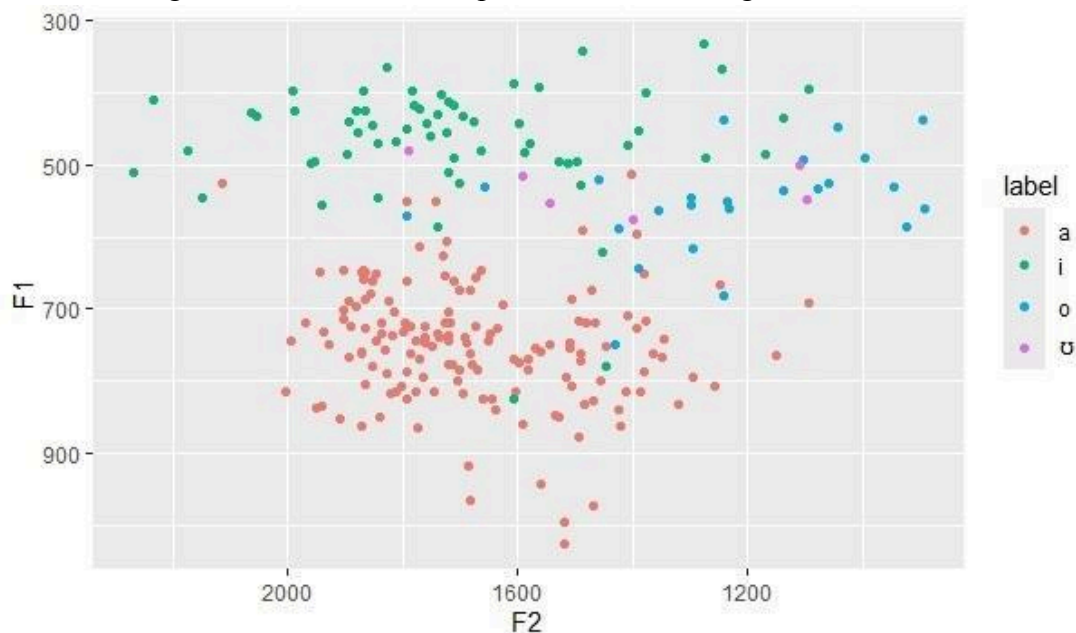


Figura 16 – Histogramas de F1 e F2 de [i]



Esses gráficos, juntamente com os das outras vogais, nos mostram que a distribuição de alguns valores de formantes estavam tendendo às extremidades. Fórmulas comuns à estatística e à mineração de dados nos ajudaram na identificação de *outliers* – conhecidos também como valores atípicos (MORETTIN; BUSSAB, 2017, p. 52) –, para que pudéssemos fazer uma análise sem a existência desses dados com valores equivocados. Antes de eliminarmos os *outliers* geramos o seguinte gráfico de dispersão para termos uma noção mais pontual, em outra perspectiva, da distribuição dos valores dos formantes.

Figura 17 – Gráfico de dispersão inicial das vogais orais tônicas



Em seguida, geramos *boxplots*²¹ para que pudéssemos identificar os *outliers*. Para isso, foi preciso a escrita de código em R (disponível no Anexo V). O código foi responsável por fazer o agrupamento dos valores dos dados em quartis e fazer o cálculo responsável por gerar os *boxplots* e identificar os *outliers*. O código identifica o IQR (*Interquatile Range* – Intervalo interquartil) calculado como $Q3 - Q1$; em seguida o LS (Limite Superior), calculado como $Q3 + 1.5 \times IQR$ e; o LI (Limite Inferior), calculado como $Q1 - 1.5 \times IQR$, em que Q3 e Q1 correspondem, respectivamente, ao primeiro e ao terceiro quartil (cf. Silva *et al.*, 2016; Bussab & Morettin, 2017).

²¹ Silva *et al.* (2016, p. 38) explica que: “Diagrama de caixa ou *Boxplot* é uma maneira popular de visualizar a distribuição dos dados. Incorpora o *resumo dos cinco números*, obtidos a partir da combinação dos quartis com valores de mínimo e máximo de um conjunto de valores, da seguinte maneira [...]: - Os pontos finais da caixa são os Q1 e Q3. O tamanho da caixa é também chamado de faixa interquartil (ou *interquatile range*, $IQR = Q3 - Q1$). - A mediana é marcada por uma linha dentro da caixa (Q2).- Barreiras de outlier, os traços no final das linhas transversais, são dadas por: barreira inferior = $Q1 - (1,5 * IQR)$; barreira superior = $Q3 + (1,5 * IQR)$. - Duas linhas nas transversais fora da caixa delimitadas pelas barreiras são chamadas bigodes”.

Figura 18 – *Boxplot* dos valores de F1 das vogais orais tônicas

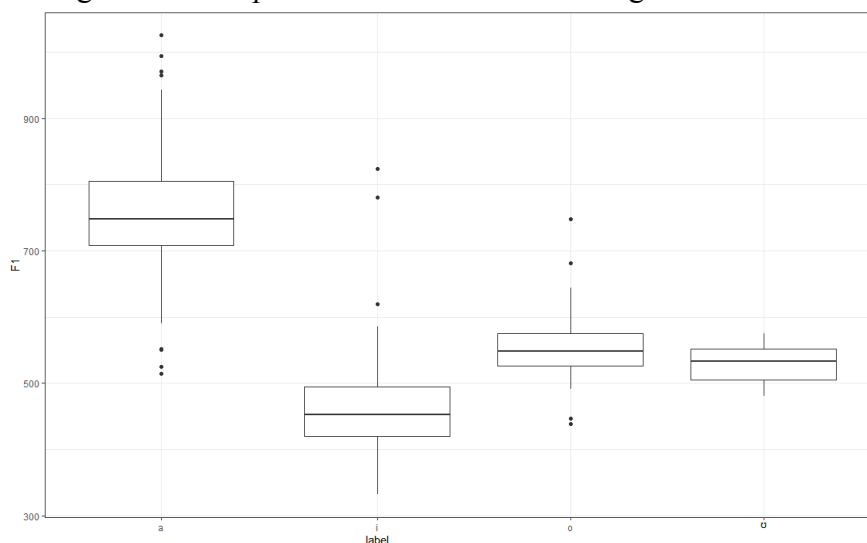
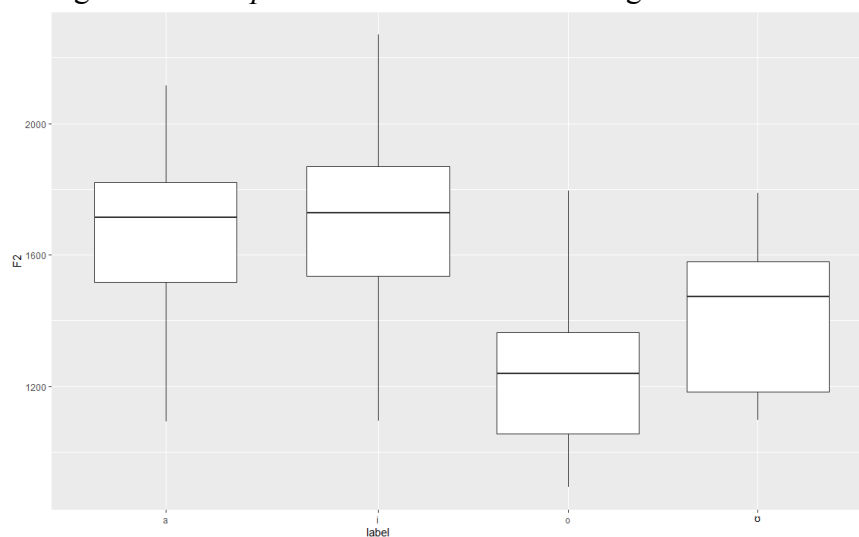


Figura 19 – *Boxplot* dos valores de F2 das vogais orais tônicas

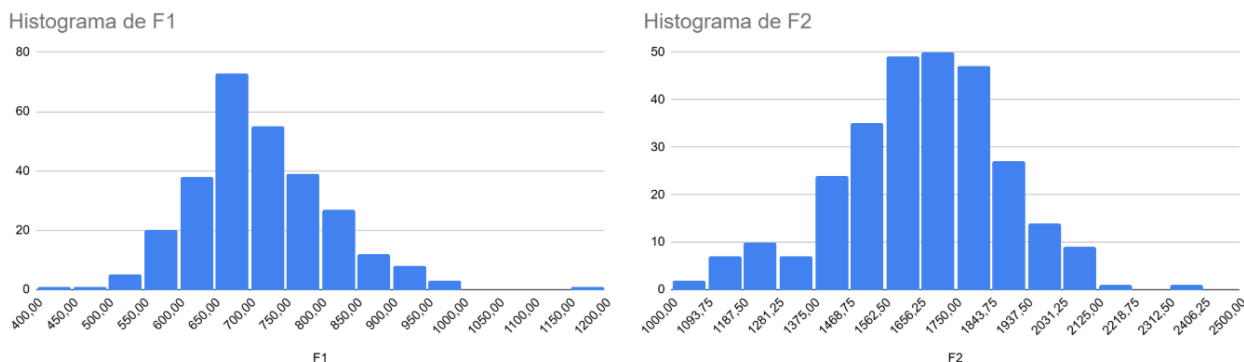


Com base nesses gráficos, o primeiro com a distribuição dos valores de F1 e o segundo com a distribuição dos valores de F2, identificamos e eliminamos os *outliers* (pontos nas extremidades das linhas) para que pudéssemos ter uma análise estatisticamente mais precisa. Para que os *outliers* fossem eliminados, nós identificamos cada um dos segmentos encontrados com um número de identificação (ID), e caso seus valores de F1 e/ou F2 não estivessem entre os valores de LS e LI, eles seriam descartados. Com isso, analisaremos apenas as vogais que

tiveram dados mais precisos que foram extraídos por meio do Praat. Na seção a seguir apresentaremos os gráficos resultantes dessa filtragem de *outliers* dos valores de formantes referentes às vogais tônicas orais.

Por conta do maior número de dados e ocorrências, não foi preciso fazer a eliminação de outliers na análise dos formantes das vogais átonas, já que a distribuição dos segmentos mostrou-se consistente com o esperado e os possíveis valores equivocados extraídos quase não tiveram interferência nos valores das médias. Os histogramas estavam com a distribuições dos valores bem distribuídos e dentro de uma dispersão já esperada, como pode ser observado nos histogramas de [e], por exemplo:

Figura 20 – Histograma de F1 e F2 de [e]



Nas seções a seguir trataremos de todos os segmentos encontrados na fala de Juanita. Dividimos o grupo das vogais em vogais orais tônicas e átonas. Nessas seções exploraremos as diferenças entre cada grupo vogais com base nos valores de F1 e F2, principalmente. Após isso, faremos a distinção entre esses dois grupos de vogais, tônicas e átonas, levando em consideração traços acústicos como duração, *pitch* e intensidade. Em seguida, apresentaremos as vogais nasais existentes na língua. Para comprovarmos a existência desse tipo de vogais, iremos utilizar alguns recursos acústicos, o principal deles será o contraste espectral. Por fim, ainda tratando de segmentos vocálicos, traremos os principais encontros vocálicos presentes no idioleto de Juanita.

3.1.2 Vogais Tônicas Orais

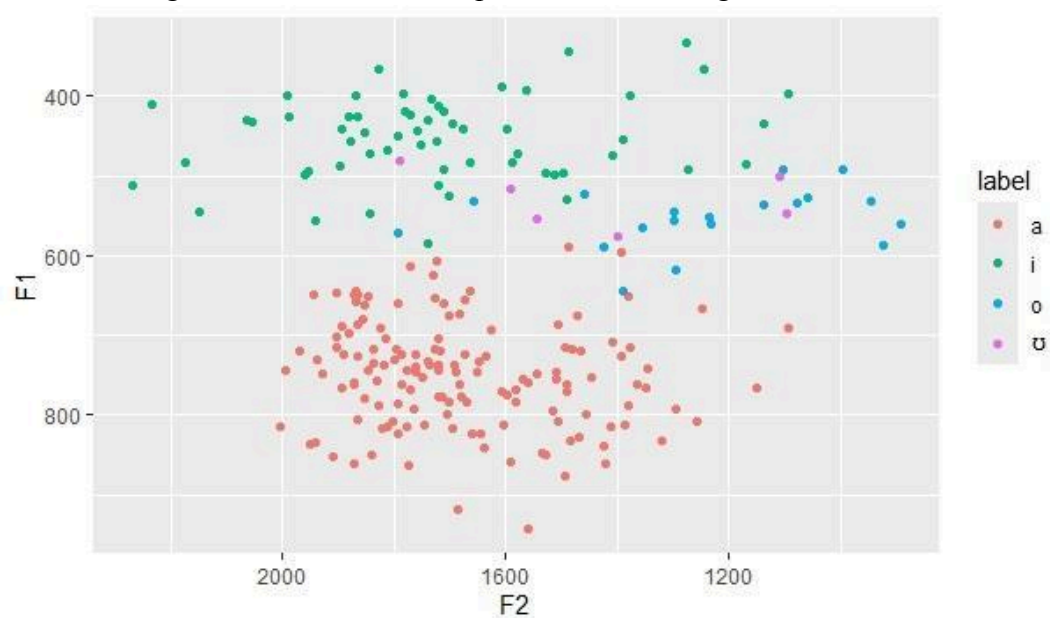
Após todas essas etapas, podemos chegar a uma proposta mais adequada do quadro fonético das vogais tônicas orais do Arutani falado por Juanita. Dessa forma, as vogais tônicas orais são: [a], [i], [o] e [u]. A seguir temos uma tabela em que é possível observar as médias dos valores da duração, do *pitch*, da intensidade e dos três primeiros formantes de todas as vogais tônicas orais encontradas em nosso *corpus* com seus respectivos valores de desvio-padrão.

Tabela 8 – Médias dos valores de duração, *pitch*, intensidade, F1, F2 e F3 das vogais tônicas orais

Vogal (Nº de Ocorrências)						
Desvio-Padrão	Duração ms	Pitch Hz	Intensidade dB	F1 Hz	F2 Hz	F3 Hz
[a] (144)	146 ms	215 Hz	72 dB	749 Hz	1.677 Hz	2.630 Hz
Desvio-Padrão em [a]	48 ms	29 Hz	8 dB	67 Hz	189 Hz	396 Hz
[i] (59)	154 ms	225 Hz	71 dB	452 Hz	1.716 Hz	2.880 Hz
Desvio-Padrão em [i]	51 ms	22 Hz	8 dB	53 Hz	264 Hz	148 Hz
[o] (19)	159 ms	233 Hz	76 dB	553 Hz	1.241 Hz	2.686 Hz
Desvio-Padrão em [o]	56 ms	28 Hz	5 dB	38 Hz	243 Hz	218 Hz
[u] (6)	161 ms	229 Hz	76 dB	529 Hz	1.422 Hz	2.708 Hz
Desvio-Padrão em [u]	59 ms	41 Hz	6 dB	36 Hz	276 Hz	165 Hz

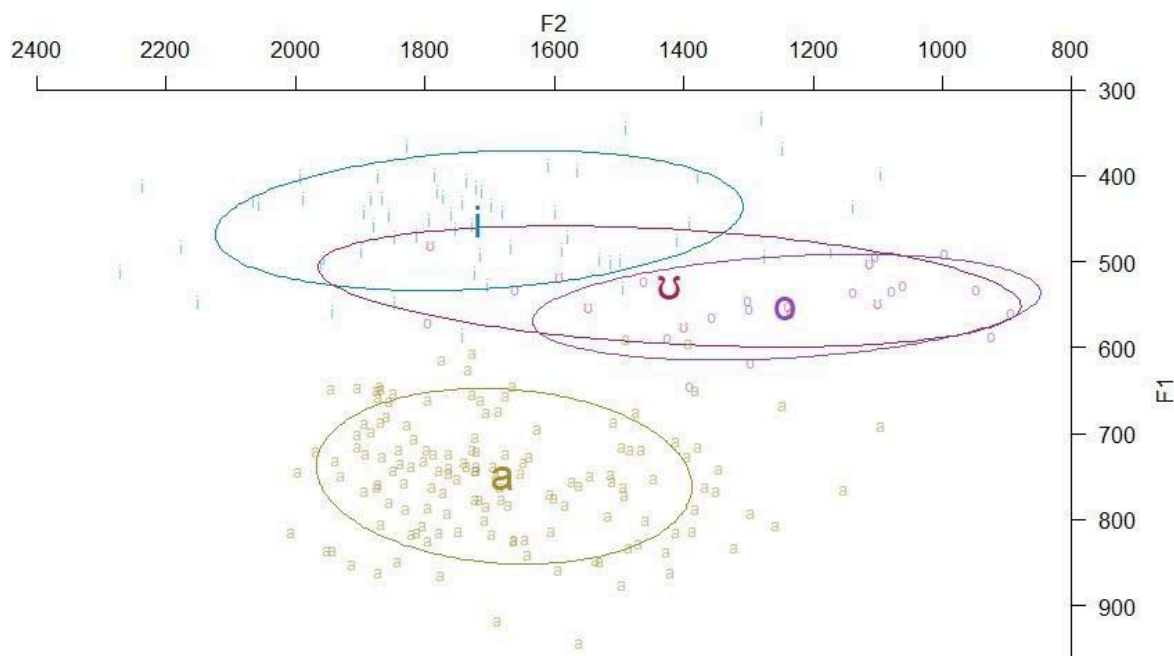
A partir desses dados, pudemos gerar gráficos para entendermos melhor a distribuição dos dados. Após a eliminação dos *outliers*, com base no gráfico da Figura 17, conseguimos observar que a dispersão da vogal [a] está mais em uma região inferior e central do gráfico, com algumas poucas sendo encontradas mais distantes, o que já era esperado pela questão da qualidade dos áudios. Já as outras vogais, [i], [o] e [u] possuem uma dispersão mais ampla e menos concentrada. A vogal [i] é a mais dispersa das três, mas podemos observar que existe uma relação direta entre maior dispersão e menor quantidade de ocorrência. Isso ocorre pelo fato de que, como são menos dados, acaba sendo mais difícil de observarmos uma nuvem de dispersão mais concentrada.

Figura 21 – Gráfico de dispersão final das vogais orais tônicas



Dessa forma, podemos chegar ao quadro vocálico com base nos dados acústicos (F1 e F2) do Arutani falado por Juanita. No gráfico a seguir, temos em destaque as vogais [a], [i], [o] e [u], elas estão em um tamanho maior e em negrito pois representam a posição em que o valor da média das respectivas vogais se encontra no gráfico, ou seja, o valor da média apresentado anteriormente na Tabela 8. Essas elipses são a representação gráfica do desvio-padrão de cada uma das vogais, apresentado, também, na Tabela 8.

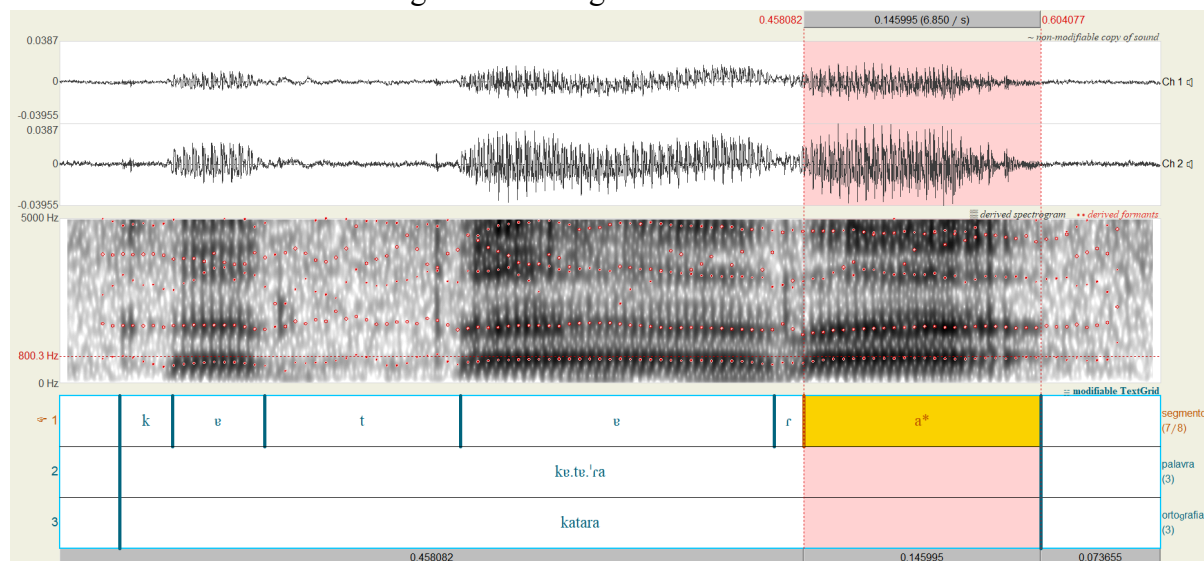
Figura 22 – Quadro de dispersão final das vogais tônicas orais



Vamos a seguir mostrar em maiores detalhes as propriedades acústicas das vogais tônicas orais. As palavras escolhidas para ilustrar cada vogal estão em (5) abaixo:

- (4)
- | | | |
|----|------------------|-------------|
| a. | <i>katara</i> | [kɐ.tɐ.'ra] |
| | ‘manicuera’ | |
| b. | <i>kayari</i> | [kɐ.jɐ.'ri] |
| | ‘pintura facial’ | |
| c. | <i>karo</i> | [kɐ.'rɔ] |
| | ‘mulher’ | |
| d. | <i>kuyuru</i> | [kɔ.jɔ.'rɔ] |
| | ‘tucumã’ | |

A vogal oral [a], assim como pôde ser observado no gráfico, é muito recorrente em posição tônica na fala de Juanita, ocorrendo em outras palavras como *mamusa* [mɐ.mɔ.'sa], que significa ‘maniva’; *kirixa* [kɪ.rɪ.'ja], que é palavra em Arutani para ‘inhame’; e, por fim, *katara* [kɐ.tɐ.'ra], que significa ‘manicuera’ em português, como já ilustrado em 5a. Observe a imagem acústica de [kɐ.tɐ.'ra].

Figura 23 – Imagem acústica de *kirixa*

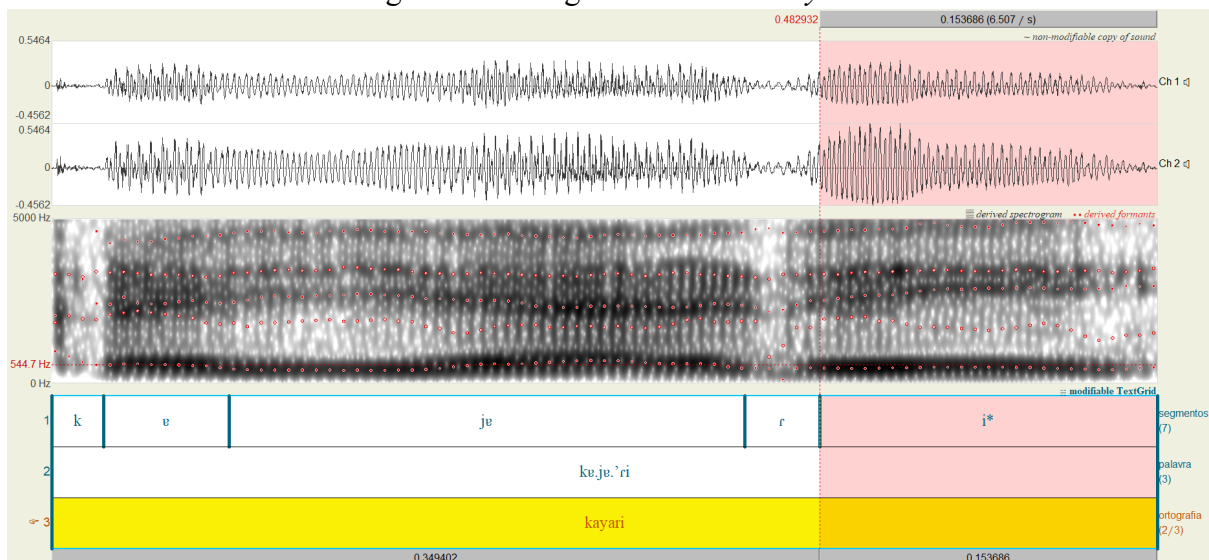
A vogal [a] em [kɐ.te.'ra] possui uma duração de aproximadamente 145 ms e os seguintes valores de formantes: F1: 738 Hz, F2: 1.693 Hz e F3: 3.287 Hz; dados retirados com base na análise LPC feita pelo próprio Praat com base no ponto médio da vogal. Os dados extraídos pelo *script* 1 são próximos, veja: F1: 724 Hz, F2: 1.673 Hz e F3: 3.210 Hz, evidenciando o fato de o [a] na fala de Juanita ser mais baixo com relação à altura da língua e do maxilar e mais central quanto ao avanço ou recuo da língua. A partir de agora seguiremos utilizando os dados extraídos pelo *Script* 1 já que eles foram os utilizados para elaborar os gráficos de análise apresentados na seção anterior. Dessa forma, além da duração, que já pode ser observada na figura anterior, temos os valores de *pitch* e intensidade de [a], que são, respectivamente, de 185 Hz e 52 dB. Os valores do [a] em [kɐ.te.'ra] estão próximos das médias gerais, apenas *pitch* e intensidade que estão um pouco abaixo da média geral, pois Juanita faz uma pronúncia com menos intensidade, por se tratar de um contexto de final de palavra e pelo fato de a intensidade ser, dentre as características acústicas aqui analisadas, a que mais pode sofrer variação²².

A segunda vogal oral tônica mais recorrente na fala de Juanita é o [i]. Ela ocorre em palavras como *kapti* [kɐp̃.'ti], que significa ‘ponta’; *awhi* [ɛw.'hi], que é nome em Arutani para a

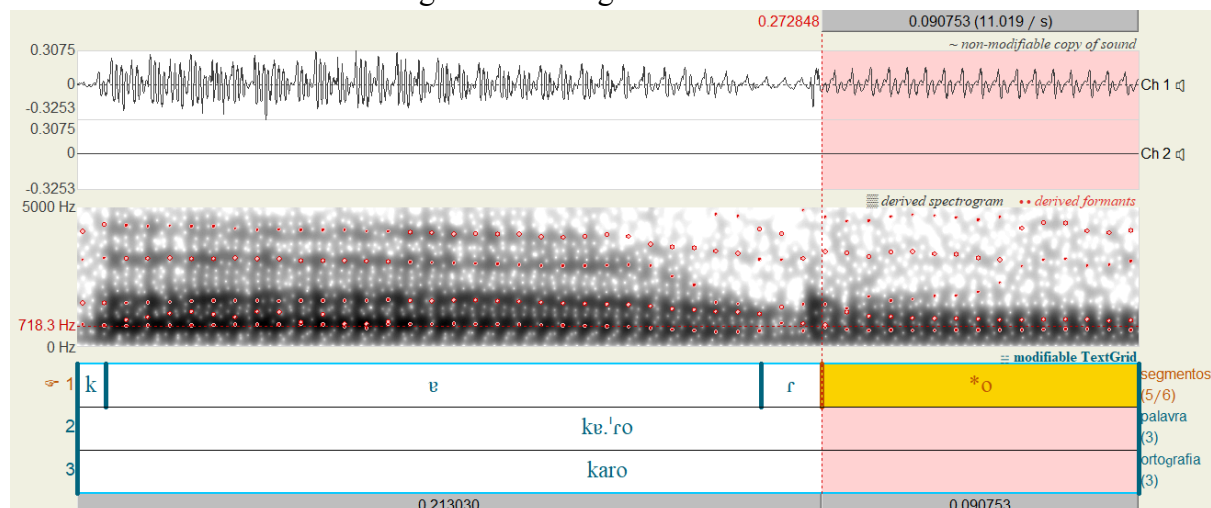
²² A intensidade é, dentre as características de duração, *pitch* e intensidade, a mesmo confiável, pois ela pode variar por conta de diversos aspetos que não estão dentro de traços estritamente acústicos, como por exemplo a alteração da distância entre o microfone do gravador e o falante quando o falante mexe a cabeça para os lados, por exemplo.

ave ‘sururina’ – ave pequena muito comum na região; e, como já apresentado em 5b, *kayari* [kə.jə.'ri], que é a palavra em Arutani para ‘pintura facial’. Em [kə.jə.'ri], temos a seguinte imagem acústica:

Figura 24 – Imagem acústica de *kayari*

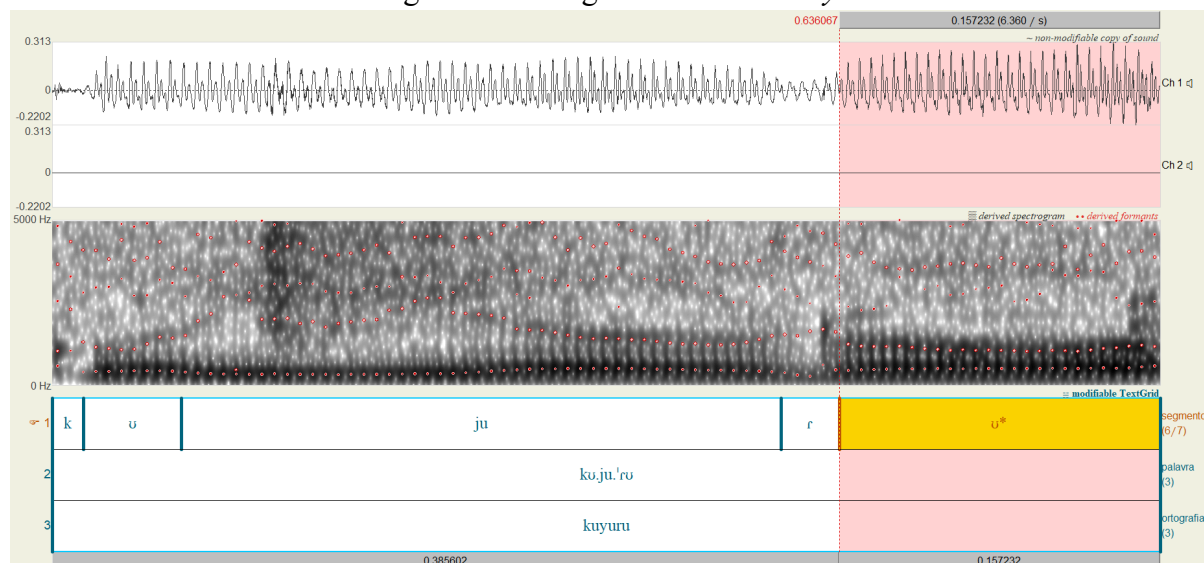


O [i] em [kə.jə.'ri] tem uma duração de aproximadamente 153 ms, um *pitch* de 219 Hz e 77 dB de intensidade. Os três primeiros formantes estão por volta dos seguintes valores: F1: 441 Hz, F2: 1.678 Hz e F3: 2.798 Hz, mostrando-se ser uma vogal alta e anterior em relação à vogal [a], pois o [i] possui um F1 mais baixo e um F2 mais alto. Esses valores estão muito próximos da média geral. Também temos a vogal tônica oral [o], que é a terceira mais recorrente na fala de Juanita. Ela ocorre em palavras como: *ako* [ə.'ko], que é nome dado ao cocho utilizado para fazer caxiri ou o nome para a massa de mandioca; *arako* [ə.rə.'ko], nome em Arutani para o ‘cará-roxo’; e *karo* [kə.'ro] que significa ‘mulher’, apresentado em 5c. Na imagem a seguir, observamos a vogal tônica [o], em destaque:

Figura 25 – Imagem acústica de *karo*

Nessa palavra o [o] possui uma duração de 90 ms os seguintes valores de formantes: F1: 578 Hz; F2: 961 Hz e; F3: 2.126, evidenciando o fato de o [o] ser uma vogal média e posterior. O F2 está um pouco abaixo da média dos outros [o] em contexto tônico, mas ainda está dentro do esperado. O [o] em [kɐ.'ro] possui um *pitch* de 195 Hz e uma intensidade de 66 dB, valores abaixo da média, mas que condizem com o modo menos intenso da fala de Juanita no áudio em análise.

Em Arutani, e na fala de Juanita, encontramos também a vogal oral tônica [ʊ]. Ela ocorre poucas vezes no *corpus* em contexto tônico, porém, podemos encontrá-la em palavras como: *pururu ata* [pʊ.rʊ.'rʊ ɐ.'ta], nome em Arutani da ave 'pipira-de-garganta-preta'; *yapururu* [jɐ.pʊ.rʊ.'rʊ], que é o nome de um tipo de banana, conhecida como 'banana-comprida'; e, como apresentado em 5d, *kuyuru* [kʊ.jʊ.'rʊ], que é a palavra para 'tucumã'. Em [kʊ.jʊ.'rʊ], o [ʊ] tônico possui a seguinte imagem acústica:

Figura 26 – Imagem acústica de *kuyuru*

Com uma duração de aproximadamente 157 ms, o [u] em destaque possui também um *pitch* de 170 Hz, um pouco abaixo da média geral, e uma intensidade de 71 dB. Ele possui os seguintes valores dos três primeiros formantes: F1: 501 Hz; F2: 1.111 Hz e; F3: 2.845 Hz. Estes valores evidenciam o fato de o [u] ser uma vogal alta e posterior, mas que é mais baixa e centralizada com relação às vogais mais extremas, como é o caso de [u] e [u].

3.1.3 Vogais Átonas Orais

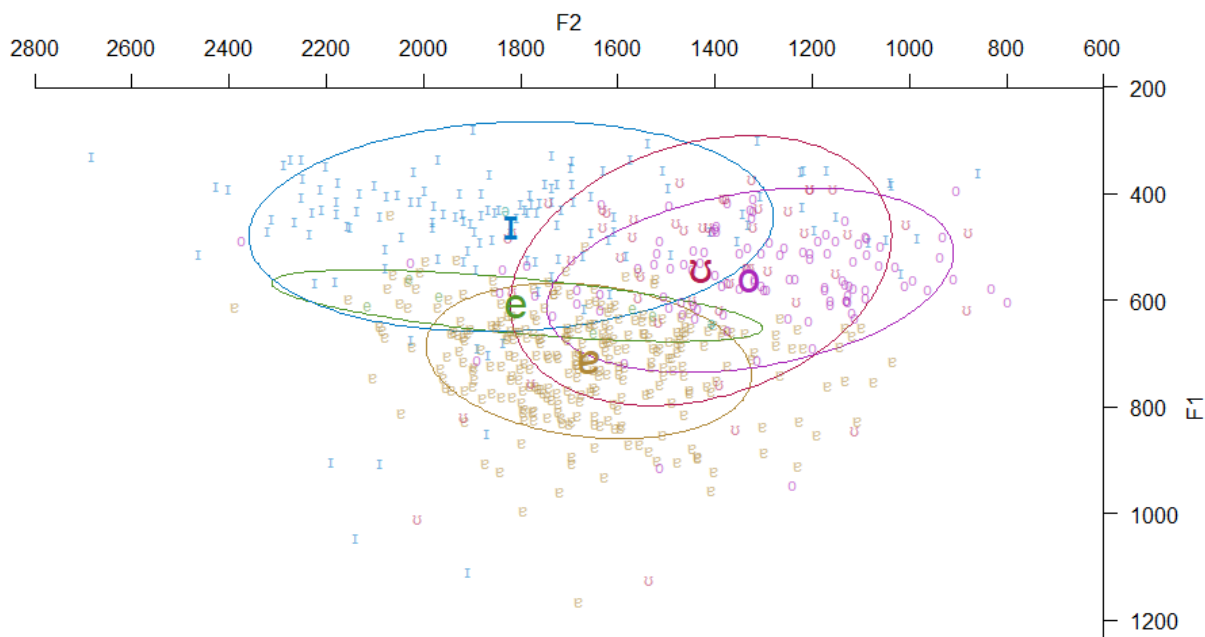
Quanto às vogais átonas orais encontradas na fala de Juanita, temos as seguintes cinco vogais orais átonas: [ɐ], [e], [ɪ], [o] e [u]. Para termos um panorama geral das características acústicas das vogais átonas da língua, apresentaremos as médias dos valores da duração, do *pitch*, da intensidade e dos três primeiros formantes de todas as vogais átonas orais existentes no nosso *corpus* com seus respectivos valores de desvio-padrão. Observe:

Tabela 9 – Médias dos valores de duração, *pitch*, intensidade, F1, F2 e F3 das vogais átonas orais

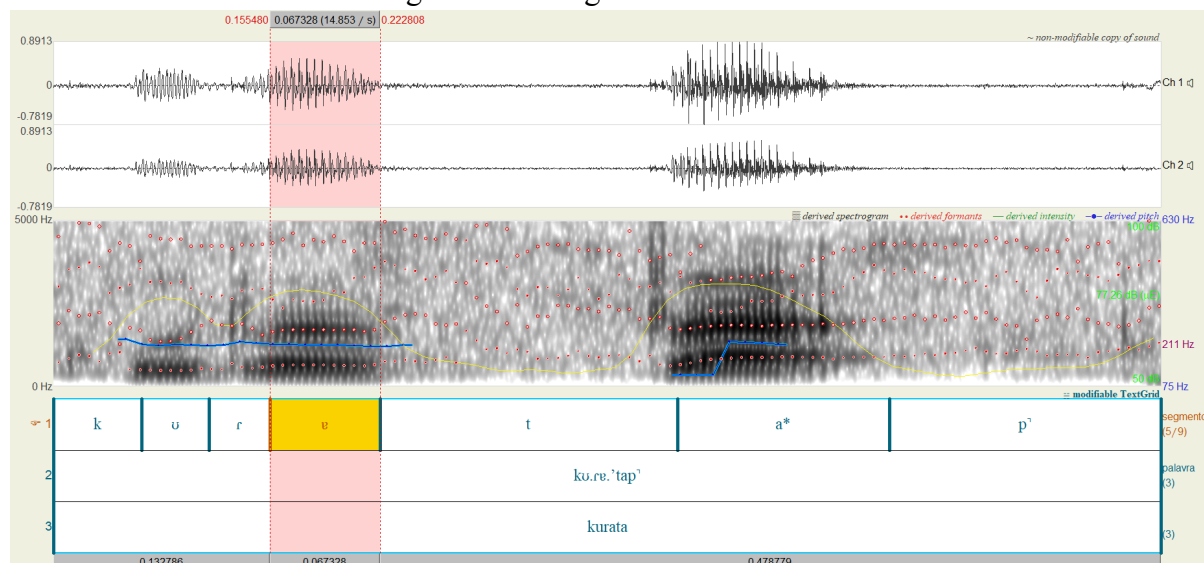
Vogais (Nº de Ocorrências)		Duração ms	<i>Pitch</i> Hz	Intensidade dB	F1 Hz	F2 Hz	F3 Hz
Desvio-Padrão							
[ɐ] (283)		99 ms	195 Hz	71 dB	713 Hz	1.660 Hz	2.693 Hz
Desvio-Padrão em [ɐ]		42 ms	24 Hz	9 dB	96 Hz	220 Hz	400 Hz
[e] (6)		91 ms	218 Hz	71 dB	609 Hz	1.807 Hz	2.505 Hz
Desvio-Padrão em [e]		25 ms	22 Hz	9 dB	34 Hz	256 Hz	191 Hz
[ɪ] (129)		87 ms	202 Hz	66 dB	460 Hz	1.819 Hz	2.854 Hz
Desvio-Padrão em [ɪ]		34 ms	28 Hz	10 dB	129 Hz	353 Hz	170 Hz
[o] (109)		90 ms	196 Hz	71 dB	561 Hz	1.327 Hz	2.753 Hz
Desvio-Padrão em [o]		40 ms	22 Hz	8 dB	113 Hz	273 Hz	302 Hz
[ʊ] (48)		69 ms	202 Hz	71 dB	543 Hz	1.429 Hz	2.718 Hz
Desvio-Padrão em [ʊ]		26 ms	22 Hz	7 dB	163 Hz	253 Hz	268 Hz

Com base nesses valores, podemos ter uma ideia de como esses segmentos se manifestam acusticamente e como variam quantitativamente. No gráfico a seguir, Figura 27, temos esses valores de média representados pelas vogais em destaque (em tamanho maior e em negrito) e ao redor, da mesma cor, as diferentes ocorrências desses segmentos. Novamente, as elipses representam os valores de desvio-padrão presentes na tabela anterior. Conseguimos observar algumas intersecções entre as elipses, isso evidencia o fato de que as vogais nem sempre são pronunciadas da mesma forma e que o ambiente em que elas estão inseridas pode influenciar diretamente nas suas características acústicas. No gráfico da Figura 27 das vogais átonas – ao compararmos com o gráfico das vogais tônicas da Figura 22 – podemos observar que os fones estão mais compactados dentro do plano, com isso, podemos deduzir que os valores formânticos das vogais átonas tendem ser mais próximos entre si, enquanto o das tônicas tendem a ter diferenças maiores.

Figura 27 – Quadro de dispersão das vogais átonas orais

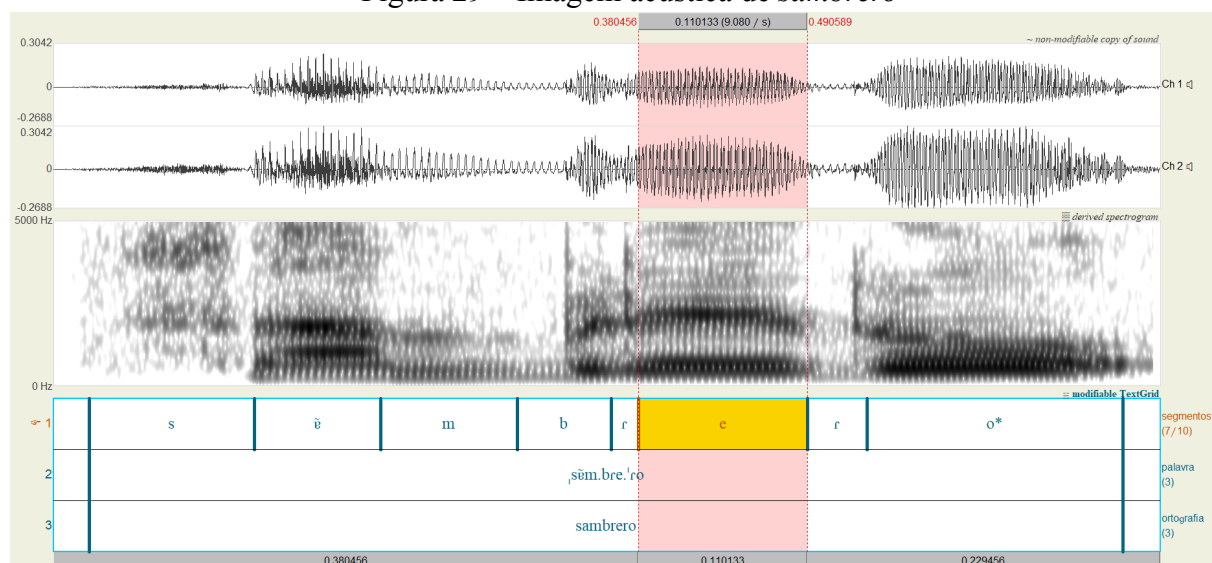


A seguir apresentaremos diversas palavras em que encontramos essas vogais átonas, mas faremos a análise acústica apenas das vogais em posição pré-tônica e ambientes o mais semelhantes possíveis, já que quase não existem vogais átonas pós-tônicas na língua, visto que o acento ocorre majoritariamente na última sílaba. Apesar disso, para gerarmos o gráfico na Figura 27 apresentado anteriormente, optamos pela seleção de todas vogais orais átonas, para que pudéssemos ter uma comparação mais geral de como elas se manifestam na língua e como ocorriam nos dados analisados. Assim sendo, encontramos a vogal átona oral [ɐ] como a mais produtiva e recorrente das vogais na fala de Juanita. Ela ocorre em palavras como *marakupaw* [mɐ.rɐ.kʊ.'paw], que é o nome em Arutani para uma espécie de ‘rã’ comestível; *parawa* [pɐ.rɐ.'wa], que é o nome em Arutani para o ‘Rio Paragua’; e *kurata* [kʊ.rɐ.'təp̄], que significa ‘zarabatana’. Observe a imagem acústica de [kʊ.rɐ.'təp̄]:

Figura 28 – Imagem acústica de *kurata*

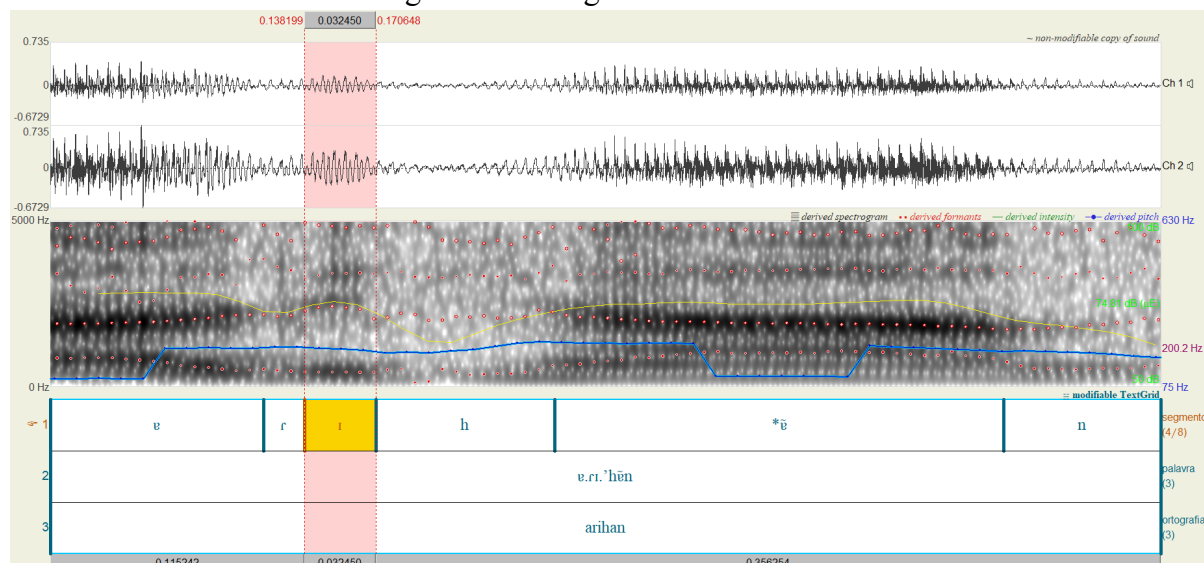
O [e] em destaque possui uma duração de aproximadamente 67 ms, um *pitch* de 211 Hz e uma intensidade de 77 dB. A vogal [e] em [ku.re.'tap̃] possui um F1 de 604 Hz, um F2 de 1.656 Hz e um F3 de 2.157 Hz, sendo que os valores de F1 e F3 estão um pouco abaixo da média. Esses valores de formantes evidenciam o fato de o [e] ser uma vogal com uma altura entre baixa e média-baixa e ser central, com relação à posição da língua no trato vocal.

Já a vogal átona oral [e] é muito menos comum, com aproximadamente 6 ocorrências, e não existe em posição tônica na fala de Juanita. Ela é encontrada em palavras como *kienawhas* [kĩ.e.nɐw.'hɛs], que é um quantificador, que pode ser traduzido como ‘pouco’; *emet* [e.'mɛt̃], que significa ‘casa’; e *sambrero* [ˌsɛm.bre.'ro] ‘chapéu’, que é um empréstimo do espanhol, com suas devidas adaptações para o padrão acentual da língua – que tende a acentuar de maneira *default* a última sílaba da palavra. Observe a imagem acústica de [ˌsɛm.bre.'ro]:

Figura 29 – Imagem acústica de *sambrero*

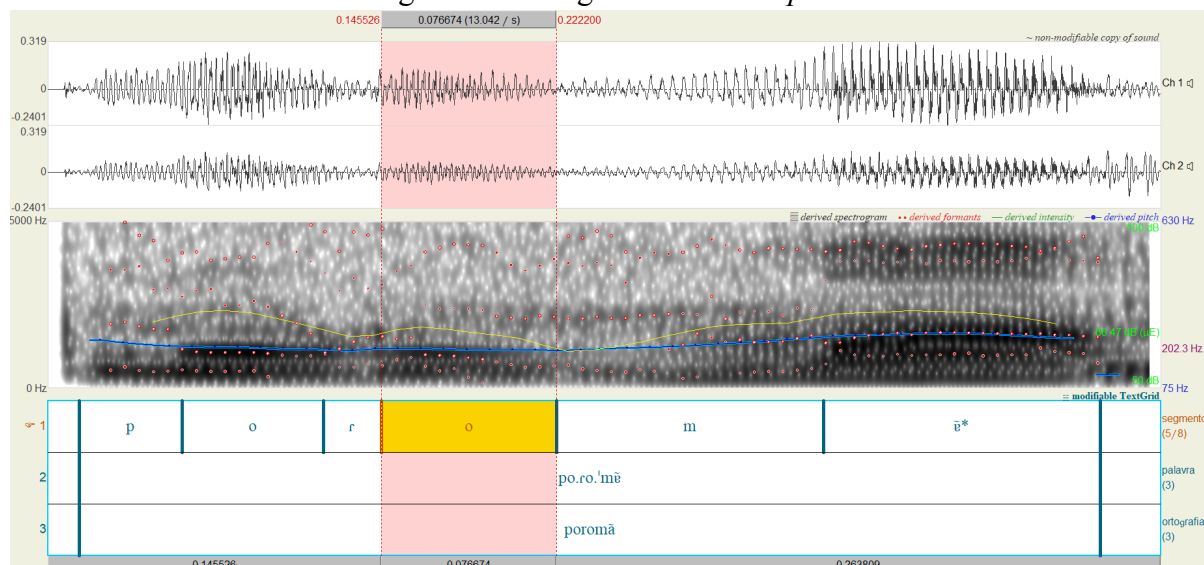
O [e] em [ˌsẽm.bre.'ro] apresenta 115 ms de duração, um *pitch* de 216 Hz e um intensidade de 73 dB, valores muito próximos da média. Nesse [e] identificamos os seguintes valores dos três primeiros formantes: F1: 558 Hz, F2: 2.027 Hz e F3: 2.451 Hz. O valores dos formantes, além de próximos da média geral, também nos provam que o que percebemos como [e], realmente se manifesta acusticamente como um [e], já que representam valores de uma vogal anterior com uma altura média-alta, se comparada com as vogais mais baixas e altas do gráfico na Figura 27.

Quanto à vogal átona [ɪ], temos coletadas aproximadamente 129 ocorrências dela na fala de Juanita, o que a torna segunda vogal mais produtiva. Ela ocorre em palavras como *pikoma* [pɪ.kõ.'ma] ‘pescoço’; *nikirio* [nɪ.kɪ.ɾɪ.'o] ‘abano’ (instrumento utilizado para abanar o fogo) e *arihan* [ɐ.ɾɪ.'hẽn], que significa ‘doce’. Observe a imagem acústica de [ɐ.ɾɪ.'hẽn]:

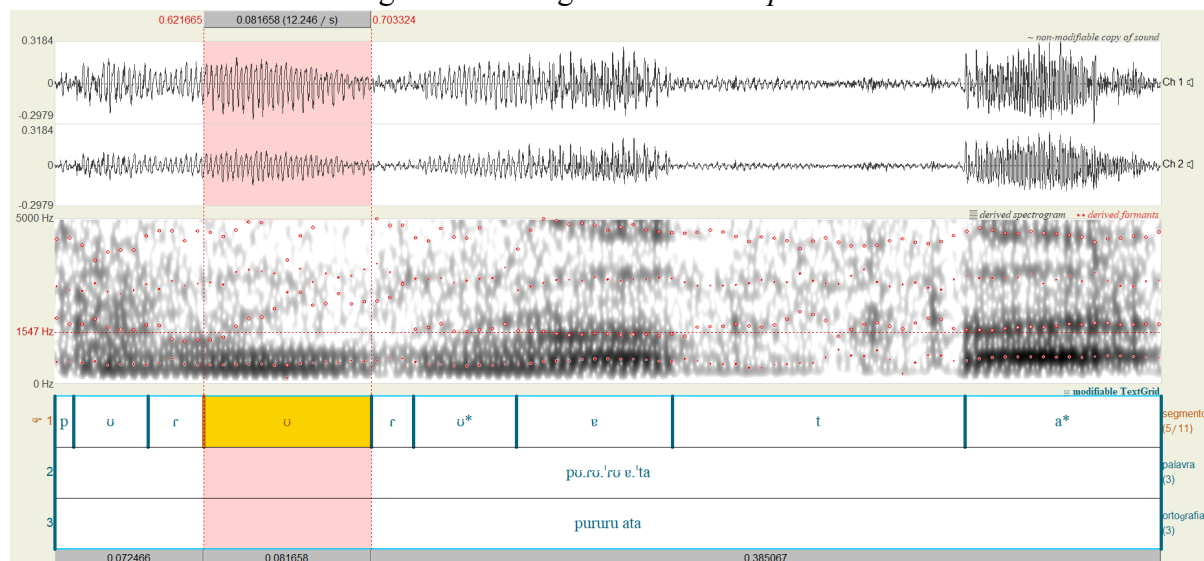
Figura 30 – Imagem acústica de *arihan*

Temos em [v.r.i.'hɛ̃n] um [i] de 32 ms de duração, que é bastante breve se comparado com a média geral. Seu *pitch* é de 194 Hz e sua intensidade é de 73 dB, a qual é um pouco acima do valor médio. O [i] em [v.r.i.'hɛ̃n] temos um F1 de 436 Hz, um F2 de 1.957 Hz e um F3 de 2.720 Hz, portanto, um valor de F1 mais baixo, característico de vogais altas, e um F2 mais alto, o que caracteriza uma vogal anterior. Seu valor de F3 também reforça o fato de o [i] ser uma vogal não-arredondada.

A terceira vogal átona mais recorrente na fala de Juanita é o [o]. Ele ocorre em palavras como *oso* [o.'so], que é um termo referencial para 'irmão'; *oo* [o.'ʔo] que pode significar 'criança' ou 'bebê'; e *poromã* [po.ro.'mẽ], que é o nome de uma árvore frutífera conhecida como 'cucura'. Vejamos a imagem acústica de [po.ro.'mẽ]:

Figura 31 – Imagem acústica de *poromã*

O [o] em destaque possui uma duração de 76 ms, um *pitch* de 202 Hz e uma intensidade de 67 dB, com os seguintes valores de formantes: F1: 568 Hz, F2: 1.130 Hz e F3: 2.176. Esses valores mais baixos de F1, F2 e F3 são característicos de vogais altas, posteriores e arredondadas. Por fim, temos o [ʊ] oral átono, ela é a vogal menos recorrente na fala de Juanita e é encontrada em palavras como *unaña* [ʊ.nɐ.'ɲa], que significa 'óleo' ou 'azeite'; *sui* [su.'i], que é o nome em Arutani de uma ave; e *pururu ata* [pʊ.rʊ.'rʊ ɐ.'ta], é o nome em Arutani para ave conhecida como 'pipira-garganta-preta'. Observe como ocorre o [ʊ] em [pʊ.rʊ.'rʊ ɐ.'ta] de uma perspectiva acústica:

Figura 32 – Imagem acústica de *pururu ata*

O segundo [u] em [pu.ru.'ru e.'ta] possui uma duração de 81 ms, um *pitch* de 205 Hz e uma intensidade de 71 dB, sendo a duração a única característica acima do valor médio geral. Os formantes extraídos possuem os seguintes valores: F1: 547 Hz, F2: 1.149 Hz e F3: 2.680 Hz. Esse valor mais baixo de F1 é característico de vogais mais altas e esse valor de F2 mais baixo é característico de vogais posteriores mais centralizadas.

3.1.4 Diferenças Acústicas entre Vogais Tônicas e Átonas Orais

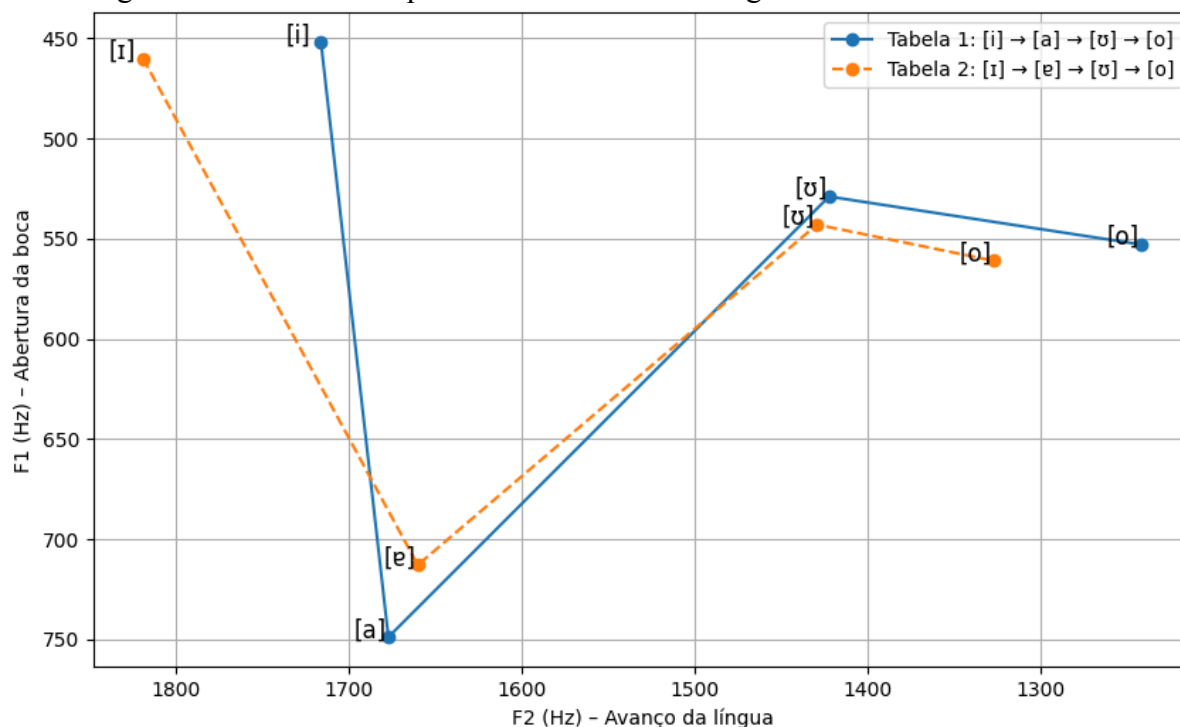
Em nossos dados, encontramos diferenças acústicas entre vogais átonas e tônicas quando levamos em consideração os parâmetros acústicos de duração, intensidade, *pitch*, F1, F2 e F3. Os traços acústicos mais adequados para o contraste entre vogais tônicas e átonas são a duração, o *pitch* e a intensidade, isso ocorre pelo fato de serem características referenciais para a demarcação do acento lexical (cf. Barbosa e Madureira, 2015). Entretanto, o contraste entre os valores dos formantes também pode nos fornecer certas evidências. Quando comparamos os valores de F1, F2 e F3 de cada uma das vogais tônicas orais com sua contraparte átona temos a seguinte tabela:

Tabela 10 – Contraste entre os valores de F1, F2 e F3 das vogais tônicas orais e das vogais átonas orais

Vogal	F1 Hz	F2 Hz	F3 Hz
[a] (tônica)	749 Hz	1.677 Hz	2.630 Hz
[ɐ] (átona)	713 Hz	1.660 Hz	2.693 Hz
[i] (tônica)	452 Hz	1.716 Hz	2.880 Hz
[ɪ] (átona)	460 Hz	1.819 Hz	2.854 Hz
[o] (tônica)	553 Hz	1.241 Hz	2.686 Hz
[ɔ] (átona)	561 Hz	1.327 Hz	2.753 Hz
[ʊ] (tônica)	529 Hz	1.422 Hz	2.708 Hz
[ʊ] (átona)	543 Hz	1.429 Hz	2.718 Hz

Quando transformamos esses valores médios em um gráfico, temos o seguinte quadro:

Figura 33 – Gráfico comparativo de F1 e F2 das vogais tônicas orais e átonas orais



A partir desse gráfico, percebemos que as vogais átonas e tônicas podem se distinguir pelo grau de dispersão ou compactação no trato vocálico. Podemos observar, também, que as vogais orais átonas [ɐ], [ʊ] e [o] são produzidas de forma mais centralizada no trato vocal de Juanita, portanto, estão mais compactadas tendendo ao centro do gráfico. Esse padrão está dentro do esperado. Todavia, o fone [i] e sua contraparte átona [ɪ] estão fora do esperado, já que [i] está mais centralizado do que [ɪ]. Essa discrepância pode ser explicada pela baixa ocorrência do fone [i] em nosso *corpus*, fazendo com que valores menores de F2 tendam a ter um efeito maior de centralização do valor médio em um quantitativo de dados menor.

Entretanto, ao nos atentarmos à duração, ao *pitch* e à intensidade dessas vogais e contrastarmos tônicas e átonas, encontramos uma tendência de que a média geral dos valores dessas unidades de medida serão maiores nas vogais tônicas do que em suas contrapartes átonas, ou seja, a duração, o *pitch* e a intensidade sempre serão maiores nas vogais tônicas do que em suas correlatas átonas. Veja:

Tabela 11 – Comparação duração, *pitch* e intensidade entre as vogais tônicas orais e suas contrapartes átonas orais

Vogal	Duração ms	Pitch Hz	Intensidade dB
[a] (tônica)	146 ms	215 Hz	72 dB
[ɐ] (átona)	99 ms	195 Hz	71 dB
[i] (tônica)	154 ms	225 Hz	71 dB
[ɪ] (átona)	87 ms	202 Hz	66 dB
[o] (tônica)	159 ms	233 Hz	76 dB
[ʊ] (átona)	90 ms	196 Hz	71 dB
[ɔ] (tônica)	161 ms	229 Hz	76 dB
[ʌ] (átona)	69 ms	202 Hz	71 dB

Ao analisarmos os valores da Tabela 11, fica clara essa tendência e observamos, também, que ela é consistente em todos os casos. Entretanto, precisamos encontrar qual é a medida mais relevante para a distinção de proeminência do acento. Para isso, precisamos fazer uma normalização dos dados. Silva *et al.* (2016, p. 44) explicam que as características descritivas de

um conjunto de dados representam valores que ajudam a identificar e distinguir cada amostra, no entanto, segundo os autores, esses valores podem ter naturezas distintas. É frequente, por exemplo, que diferentes sensores sejam utilizados para coletar informações sobre um determinado fenômeno, sendo que cada sensor realiza medições específicas, registradas em unidades distintas (SILVA *et al.*, 2016, p. 44). Esta é a nossa realidade neste momento, temos três traços acústicos diferentes: duração (ms), *pitch* (Hz) e intensidade (dB), os quais possuem grandezas diferentes.

Uma das alternativas para a normalização desse tipo de conjunto de dados com unidades de medidas diferentes é o uso da normalização *z-score*. Silva *et al.* (2016, p. 45) explicam que nesse tipo de transformação de dados “os valores são reorganizados pela média e desvio-padrão do conjunto de valores original”. Sobre o *z-score*, os mesmo autores ainda apresentam que este tipo de padronização possibilita a uniformização dos valores. Contudo, essa transformação é realizada com base na distribuição dos dados de cada atributo. Por meio dessa técnica, cada valor é ajustado conforme sua posição relativa aos demais, considerando uma distribuição simétrica (SILVA *et al.*, 2016, p. 45). Para esse cálculo, utilizamos um *script* em R, o qual está disponível no Anexo VI.

Após a normalização de todos os dados de cada uma das vogais orais segmentadas, calculamos a média dos valores de duração, *pitch* e intensidade de cada vogal. Após isso, dividimos as vogais em seus respectivos grupos de tônicas e átonas, como pode ser observado na Tabela 12:

Tabela 12 – Normalização dos valores de duração, *pitch* e intensidade das vogais orais por meio de *z-score*

Vogais	Duração (ms)	<i>Pitch</i> (Hz)	Intensidade (dB)
[a] (tônica)	2,39	0,47	0,32
[ɐ] (átona)	-0,42	-0,31	0,10
[i] (tônica)	0,89	0,74	0,32
[ɪ] (átona)	-0,50	-0,19	-0,34
[o] (tônica)	1,15	0,96	0,88
[ɔ] (átona)	-0,48	-0,35	0,10
[u] (tônica)	1,04	0,89	0,66
[ʊ] (átona)	-0,87	-0,13	0,21

Com base na Tabela 12, é possível fazer uma análise comparativa das diferenças de duração, *pitch* e intensidade entre vogais tônicas e átonas normalizadas. O objetivo é identificar qual dessas medidas tem maior relevância para a definição do acento – ou seja, qual apresenta a maior diferença sistemática entre as formas tônicas e átonas.

Tabela 13 – Análise comparativa das diferenças de duração, *pitch* e intensidade após normalização

Vogais (tônicas ~ átonas)	Diferença de Duração	Diferença de <i>Pitch</i>	Diferença de Intensidade
[a] ~ [ɐ]	2,8	0,78	0,22
[i] ~ [ɪ]	1,38	0,93	0,67
[o] ~ [ɔ]	1,62	1,31	0,78
[u] ~ [ʊ]	1,9	1,02	0,45

Ao calcularmos a média das diferenças para cada parâmetro temos os seguintes resultados:

Tabela 14 – Média das diferenças dos valores de duração, *pitch* e intensidade após normalização

Duração (ms)	<i>Pitch</i> (Hz)	Intensidade (dB)
1,93 z	1,01 z	0,53 z

Portanto, ao interpretarmos esses resultados, temos que a duração é a medida que apresenta a maior média de diferença entre vogais tônicas e átonas (1,93 z), indicando que é o parâmetro mais relevante na distinção entre sílabas acentuadas e não acentuadas, dentro do nosso conjunto de dados. O *pitch* é a segunda medida mais relevante (1,01 z), enquanto a intensidade é a menos marcante (0,53 z).

Sendo assim, podemos compreender que os valores dos três primeiros formantes foram importantes para o contraste entre tônicas e átonas. Porém, as medidas de duração, *pitch* e intensidade, por serem referenciais na demarcação do acento lexical, foram mais eficazes para esse propósito. Dessa forma, não excluimos a importância do contraste entre os valores formânticos para a diferenciação entre vogais tônicas e átonas, apenas entendemos, ao seguirmos nossas bases teóricas, que unidades de medida como duração, *pitch* e intensidade, nos fornecem informações mais precisas quando visamos identificar o acento no âmbito da palavra.

3.1.5 Vogais Nasais e Nasalizadas

Na fala de Juanita, a partir do nosso *corpus*, encontramos vogais nasais. A depender do contexto, podemos diferenciá-las em vogais nasais e nasalizadas. Na seção de fonologia exploraremos as diferenças entre vogais nasais e nasalizadas, entretanto, nesta seção já será possível perceber certas pistas que diferenciam esses dois grupos. Os contextos são os seguintes: (i) vogais nasalizadas (Ñ) ocorrendo quando há consoante nasal (N) em ataque silábico (.NÑ.)²³; (ii) Ñ com N adjacente, seja na mesma sílaba em coda ou ataque silábico na sílaba seguinte (.CÑN. / .CÑ.N)²⁴; (iii) Ñ tende a ocorrer quando anteceder ou preceder por [h] em ataque silábico, ou seja, na mesma sílaba ou na sílaba seguinte (.hÑ./ .CÑ.h). Para o caso (i) temos o exemplo de *poromã* [po.ro.'mẽ] (.NÑ.), que é o nome em Arutani para a planta ‘cucura’. Para

²³ Os pontos (.) marcam as margens das sílabas.

²⁴ C será usado para representar qualquer consoante do inventário da língua que não seja uma consoante nasal.

(ii) podemos citar os casos de *ariran* [ɐ.ɾɪ.ˈrẽɳ] (.C \tilde{V} N.), que corresponde a ‘lustrado’, e *mapaña* [mɐ.pẽ.ˈɲa] (.C \tilde{V} .N), que é a palavra para ‘mamão’. Por fim, para (iii) temos os exemplos de *atakhã* [ɐ.tɛk.ˈhẽ] (.h \tilde{V} .), que é o nome em Arutani para ‘irmã mais velha’, e *akohã* [ɐ.kõ.ˈhẽ] (.C \tilde{V} .h), que é a palavra para ‘água’.

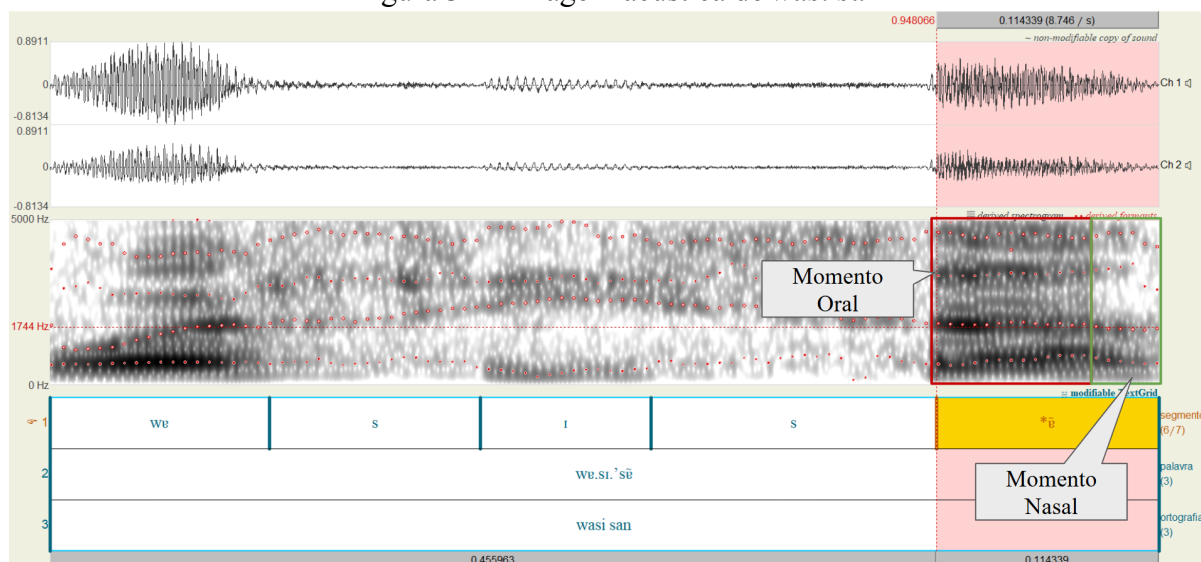
Esse processo de nasalização é irregular, ou seja, é uma tendência, mas não uma regra. Existem casos como os do contexto (i), como em *imo* [ɪ.ˈmo], que pode ser traduzido como ‘pajé’, em que há uma ambiente nasal, mas não há \tilde{V} . Quanto ao contexto (ii) temos os exemplos de *anhĩ* [ɐɳ.ˈhĩ], que significa ‘nádega’, e *konihã* [ko.ɳɪ.ˈhẽ], palavra para ‘camarão’. Por fim, temos o contexto (iii), temos os casos de *kakohap* [kɐ.ko.ˈhɑ̃p̃], que significa ‘caroço’, e o exemplo repetido de *konihã* [ko.ɳɪ.ˈhẽ]. Exploraremos o *status* fonêmico da nasalidade no Arutani falado por Juanita na seção 4.4.

Esse caso em que a fricativa glotal não-vozeada [h] induz à nasalidade é conhecido como rinoglotofilia, um fenômeno e uma tendência recorrente nas línguas do mundo. Matisoff (1975, p. 265) é um dos primeiros a falar sobre este assunto, e explica a rinoglotofilia como “uma afinidade entre a característica da nasalidade e o envolvimento articulatório da glote” (tradução nossa). Sobre o tema, Parker (2006, p. 28) apresenta que “consoantes articuladas na laringe (especialmente /h/) apresentam tendência a induzir nasalização (possivelmente irregular) em vogais adjacentes...” (tradução nossa). Como visto, a rinoglotofilia é uma fenômeno que também ocorre no idioleto de Juanita e, apesar de produtivo, é irregular, como vimos no contraste entre [ɐ.kõ.ˈhẽ] e [kɐ.ko.ˈhɑ̃p̃].

Há também casos raros de vogais nasais que ocorrem em contextos totalmente orais, como são os casos de *wasi san* [wɐ.sɪ.ˈsẽ], *ira ã* [ɪ.ˈrɐ.ˈẽ] e *xapi ã* [ʃɐ.ˈpi.ˈẽ]. Não há nenhuma consoante ou qualquer outro tipo de segmento que esteja motivando a nasalidade das vogais, por isso, por estarem inseridas em ambientes completamente orais que as diferenciamos das vogais nasalizadas. A palavra *wasi sã* [wɐ.sɪ.ˈsẽ] é o nome em Arutani para a ave ‘alma-de-gato’, *ira ã* [ɪ.ra.ˈẽ] é o nome em Arutani para a palmeira conhecida como ‘patauá’ e *xapi ã* [ʃɐ.ˈpi.ˈẽ] é o termo genérico para ‘fruta’. Quando observamos suas imagens acústicas, conseguimos identificar

claramente os dois momentos acústicos das vogais nasais, momento oral, que ocorre primeiro, e momento nasal, que ocorre posteriormente (cf. Silva *et al.*, 2019, p. 120). Veja:

Figura 34 – Imagem acústica de *wasi sã*



O esmaecimento que ocorre no espectrograma entre o segundo e o terceiro formante é consequência dos antiformantes. Por conta de as vogais nasais e nasalizadas serem produzidas com o ar vindo do pulmão sendo ejetado pela cavidade oral e nasal, os antiformantes são a consequência da ressonância simultânea que ocorre nas duas cavidades. Silva *et al.* (2019, p. 117) explica que “...devido ao acoplamento do tubo acústico da cavidade nasal, os componentes de frequência que estão próximos às frequências de ressonância do trato vocal cancelam-se, tornando-se antirressonâncias (antiformantes)”. Com essa saída de ar pelos dois tratos, encontramos também dois tipos de formantes das vogais nasais, os formantes orais (F1, F2, F3...) e os formantes nasais (FN1, FN2, FN3...). Os formantes nasais são, segundo Silva *et al.* (2019, p. 112), “picos de ressonância adicionais [...] que ocorrem devido à saída de ar nasal pelas fossas nasais”.

Podemos observar esses picos de ressonância quando comparamos vogais nasalizadas com vogais orais em ambientes análogos. Para uma melhor apresentação dos dados e comparação, iremos apresentar primeiramente quais são os segmentos vocálicos nasais e nasalizados encontrados na fala de Juanita, os quais estão divididos entre tônicos e átonos.

Depois apresentaremos as vogais nasais que possuem correlatos orais para uma comparação espectral para melhor visualização dos formantes nasais e antiformantes. Veja as tabelas a seguir em que temos na primeira coluna de cada uma delas os segmentos vocálicos seguidos entre parênteses pelo número de ocorrências, seguida pelas colunas de duração, *pitch*, intensidade, F1, F2 e F3:

Tabela 15 – Dados acústicos das vogais tônicas nasalizadas

Vogal (Nº de Ocorrências)	Duração ms	<i>Pitch</i> Hz	Intensidade dB	F1 Hz	F2 Hz	F3 Hz
[ẽ] (37) ~ [ã] (6)	143 ms	222 Hz	72	761 Hz	1660 Hz	2617 Hz
[ē] (1)	128 ms	263 Hz	75	547 Hz	2104 Hz	2332 Hz
[ī] (4)	197 ms	241 Hz	75	454 Hz	1549 Hz	2817 Hz
[ō] (1)	156 ms	225 Hz	80	656 Hz	1179 Hz	2489 Hz
[õ] (2)	164 ms	257 Hz	72	578 Hz	1412 Hz	2331 Hz

Tabela 16 – Dados acústicos das vogais átonas nasalizadas

Vogal (Nº de Ocorrências)	Duração ms	<i>Pitch</i> Hz	Intensidade dB	F1 Hz	F2 Hz	F3 Hz
[ẽ] (34)	113 ms	204 Hz	68 dB	718 Hz	1677 Hz	2786 Hz
[ē] (1)	73 ms	232 Hz	76 dB	622 Hz	1476 Hz	2755 Hz
[ī] (7)	115 ms	233 Hz	70 dB	444 Hz	1639 Hz	2773 Hz
[ō] (17)	81 ms	188 Hz	69 dB	671 Hz	1412 Hz	2832 Hz

Como podemos observar, as átonas nasais e nasalizadas tendem a ter valores menores de duração, *pitch* e intensidade quando comparadas com as correlatas tônicas.

Tendo em mente que essas são todas as vogais que possuem nasalidade encontradas na fala de Juanita, podemos fazer algumas generalizações sobre como vogais nasais/nasalizadas e orais contrastam em Arutani. Primeiramente, podemos observar uma tendência de as vogais nasais/nasalizadas serem menos frequentes no idioleto de Juanita quando comparadas com as vogais orais. Observe a diferença na frequência quando contrastamos tônicas orais e tônicas que possuem nasalidade:

Tabela 17 – Quantitativo das vogais orais e suas contrapartes nasais e nasalizadas

Vogal	Nº de Ocorrências
[i]	59
[ĩ]	4
[a]	144
[ẽ] ~ [ã]	43
[o]	6
[õ]	2
[o]	19
[õ]	1

Outra generalização seria que a diferença acústica entre vogais orais e nasais/nasalizadas pode ser demonstrada a partir do valor dos formantes, como resumido na tabela a seguir, ao compararmos os valores de F1, F2, F3 e a soma de F2+F3 de vogais orais e nasais/nasalizadas. Kent e Read (2015, p. 225) apresentam algumas tendências de medidas a serem esperadas nesse tipo de contraste entre vogais orais e nasais. Com relação a tendências dos valores dos formantes, os autores explicam que se espera que vogais nasais tendam a ter F1 maior do que vogais orais e que a soma dos valores de F2 e F3 tenda a ser menor em vogais orais do que em nasais. É exatamente isso o que encontramos na Tabela 18.

Tabela 18 – Comparação dos valores de F1, F2 e F3 entre vogais orais e nasais

Vogal	F1 (Hz)	F2 (Hz)	F3 (Hz)	F2 + F3
[i]	452 Hz	1716 Hz	2.880 Hz	4.596 Hz
[ĩ]	454 Hz	1549 Hz	2817 Hz	4.366 Hz
[a]	749 Hz	1677 Hz	2.630 Hz	4.307 Hz
[ẽ] ~ [ã]	761 Hz	1660 Hz	2617 Hz	4.277 Hz
[o]	529 Hz	1422 Hz	2.708 Hz	4.130 Hz
[õ]	578 Hz	1412 Hz	2331 Hz	3.743 Hz
[o]	553 Hz	1241 Hz	2.686 Hz	3.927 Hz
[õ]	656 Hz	1179 Hz	2489 Hz	3.668 Hz

Por outro lado, temos, também, os formantes nasais, os quais podem ser encontrados por meio de comparações espectrais a partir de análises FFT. Vamos começar pela comparação entre as palavras *xipirin ti* [ʃi.pi.'rĩn 'ti] – nome para ‘banana-casca-verde’ – e *kari* [kɐ.'ri] – palavra para ‘mulher’ em Arutani – para o contraste entre [ĩ] e [i]. Vejamos primeiramente a imagem acústica de cada uma dessas palavras:

Figura 35 – Imagem acústica de *xipirin ti*

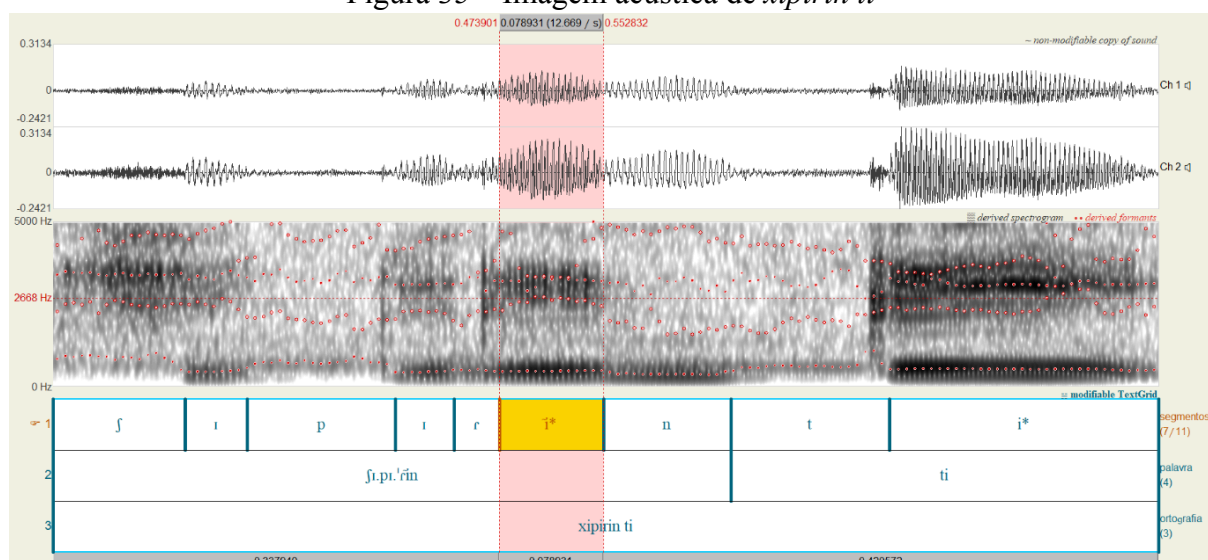
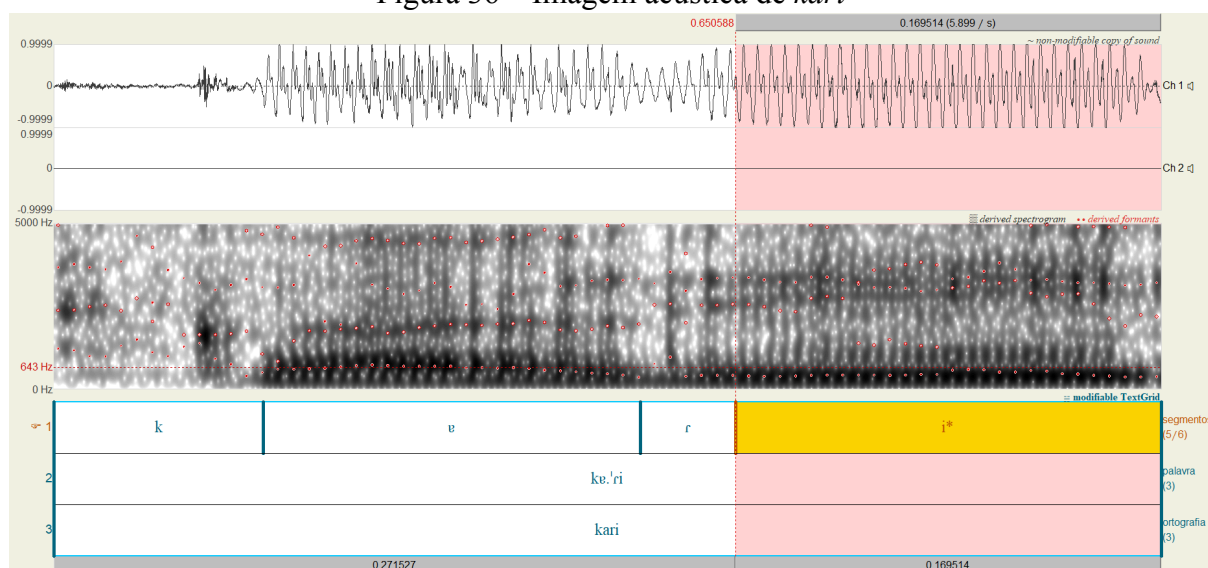
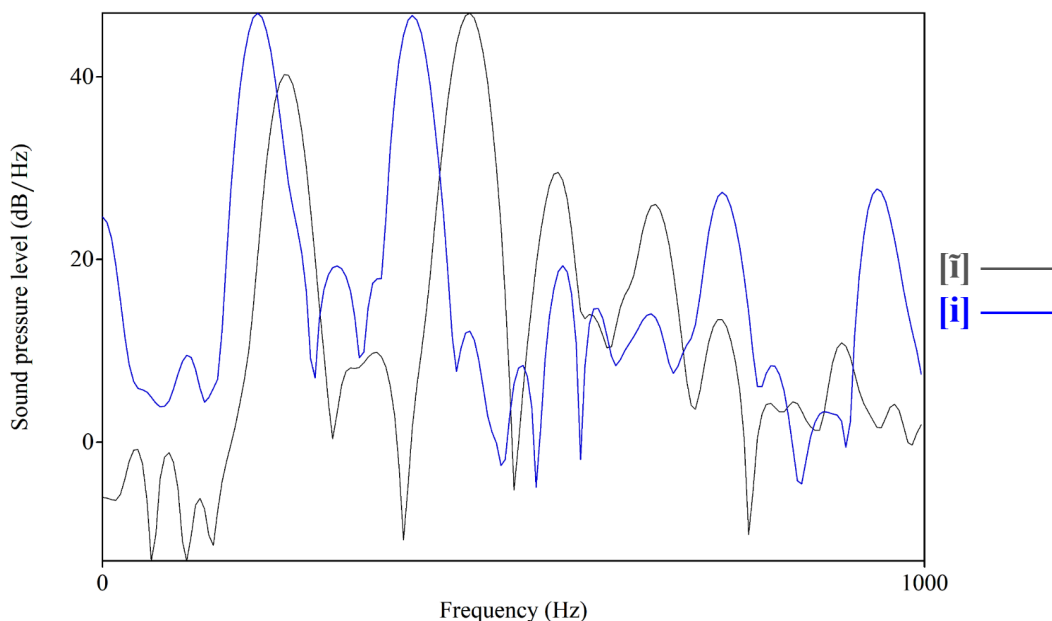


Figura 36 – Imagem acústica de *kari*



Dessa forma, ao compararmos as vogais tônicas [ĩ] e [i] dessas palavras, encontramos as seguintes configurações espectrais:

Figura 37 – [ĩ] vs. [i]



A partir desse gráfico podemos observar traços acústicos como os formantes nasais e os antiformantes. Os formantes nasais são, assim como apresentado por Silva *et al.* (2019, p. 112), os picos de ressonância que surgem e que não são encontrados no espectro da vogal oral. Da mesma forma, os antiformantes podem ser identificados pelos vales espectrais acentuados que não estão presentes no segmento oral. Para [ẽ] e sua contraparte oral [a] podemos extrair essas informações a partir de palavras como *ariran* [ɐ.ɾɪ.'rẽn] – que é a palavra em Arutani para ‘lustrado’ – e *katara* [kɐ.tɐ.'ra] – palavra na língua para ‘manicuera’. Veja as imagens acústicas dessas palavras:

Figura 38 – Imagem acústica de *ariran*

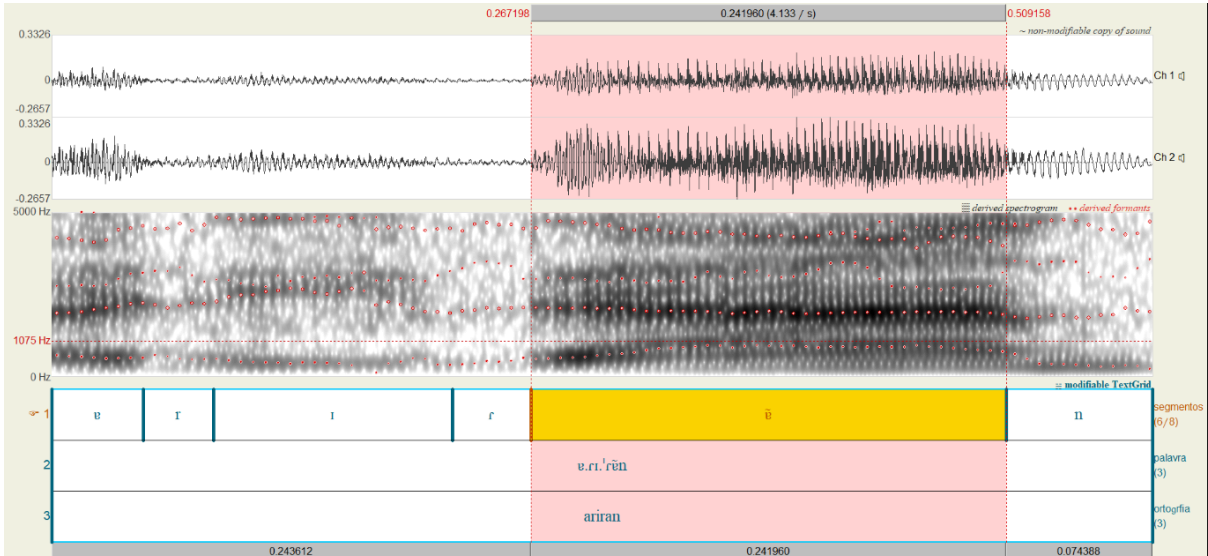
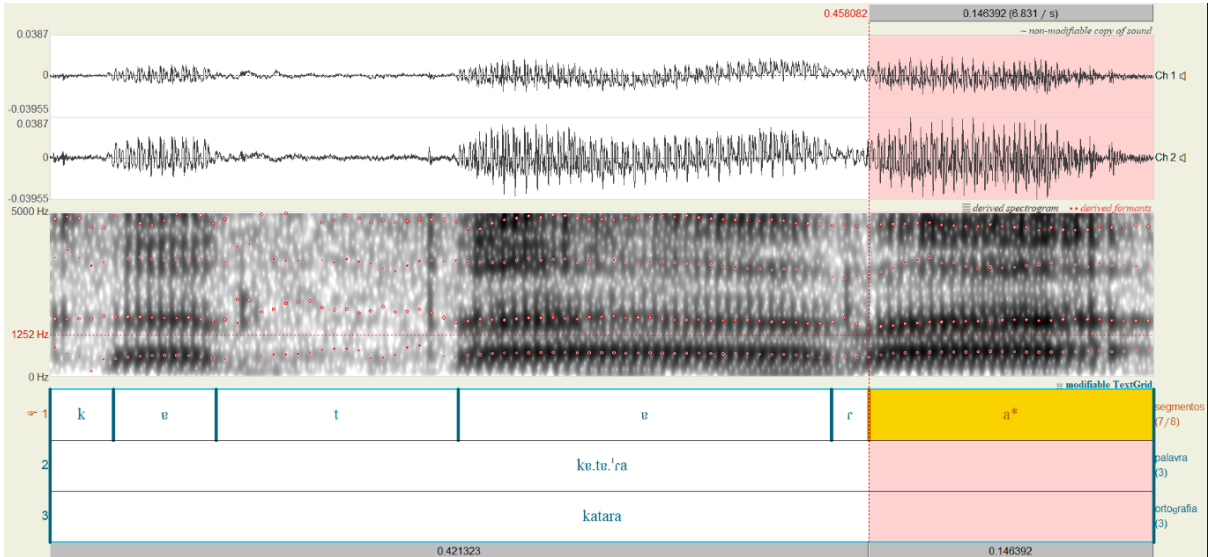
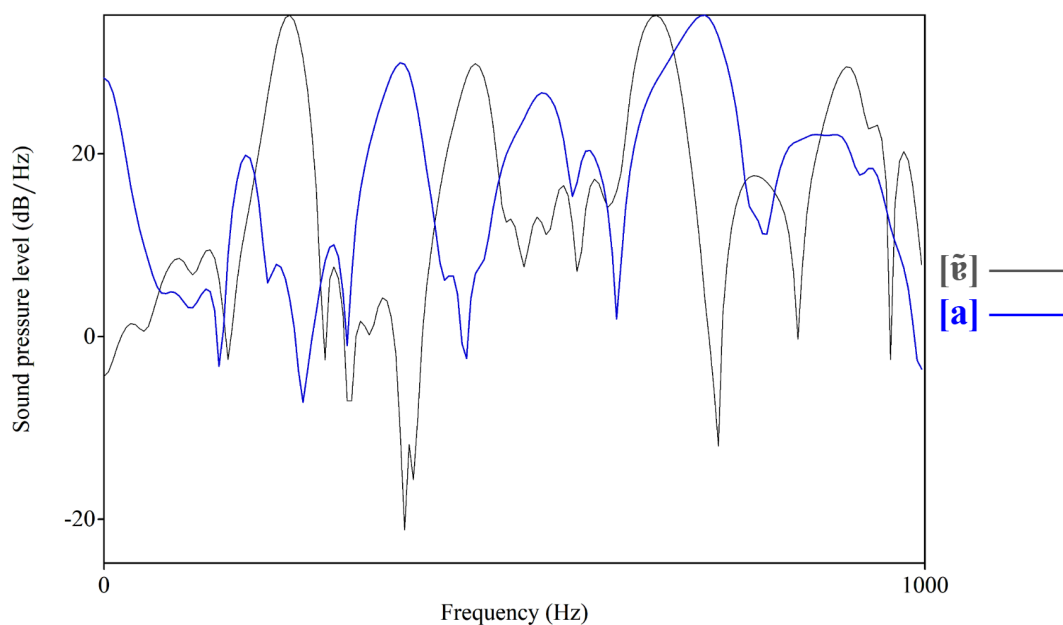


Figura 39 – Imagem acústica de *katara*

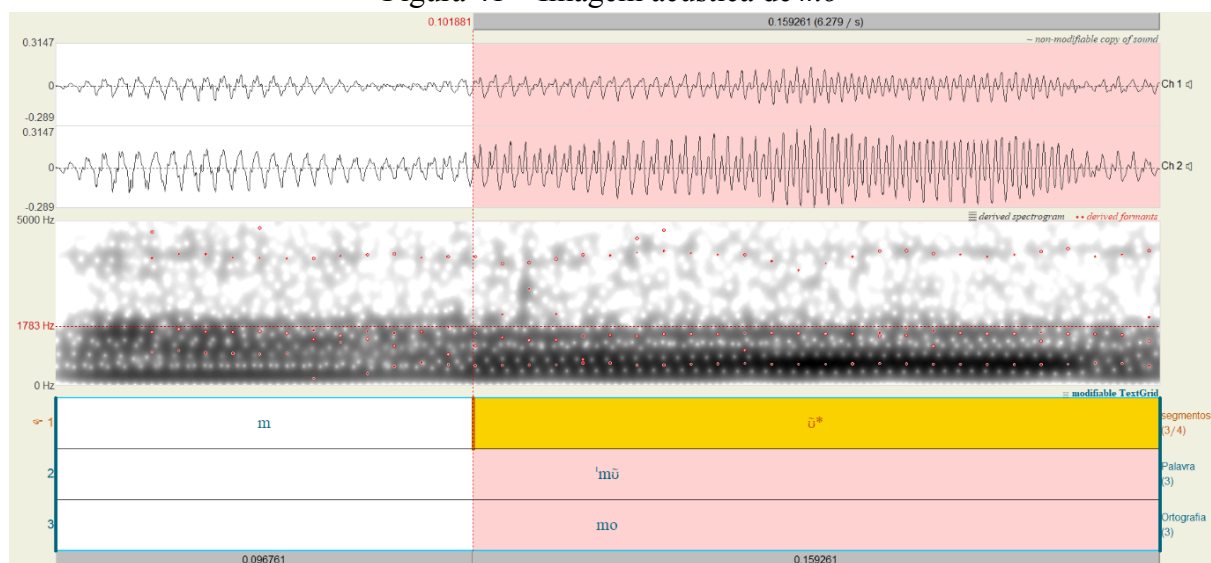
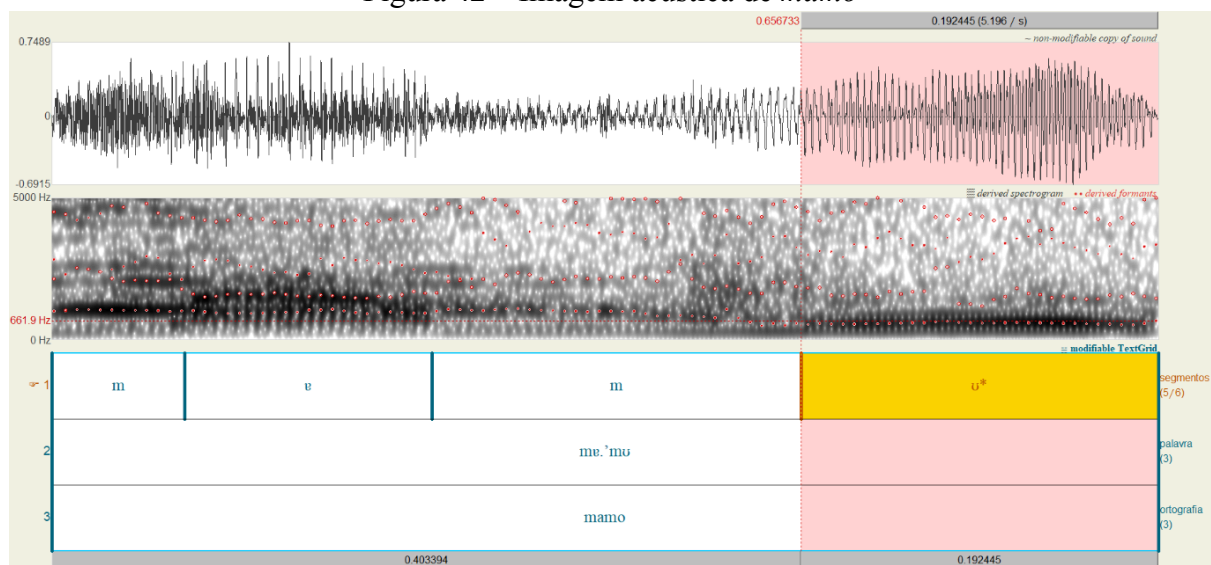


Quando comparamos os espectros de [ẽ] e [a] obtemos o seguinte:

Figura 40 – [ẽ] vs. [a]

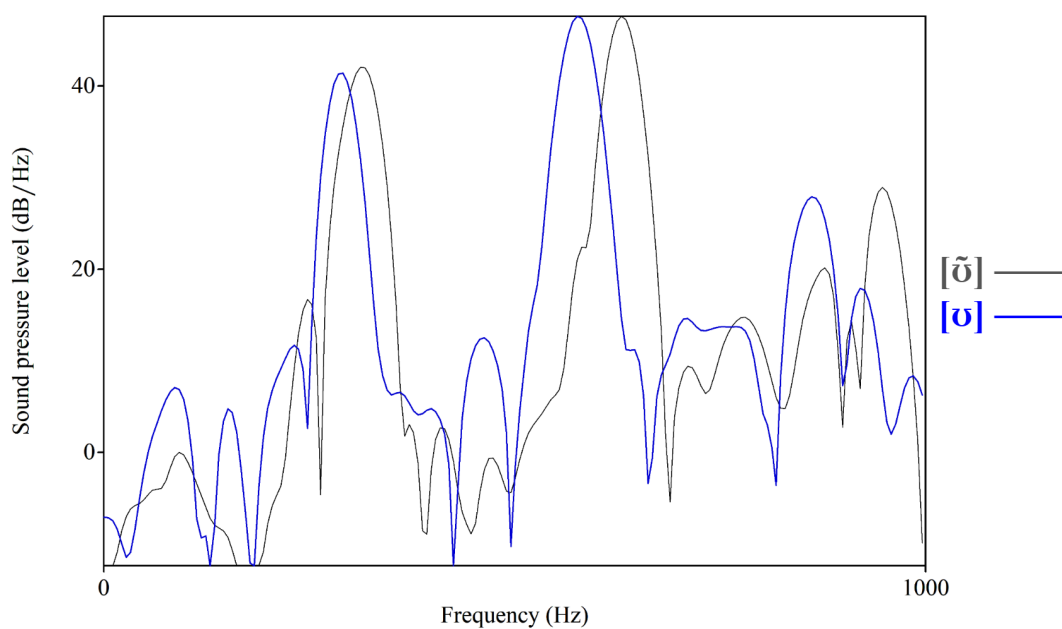


Observe como os antiformates estão bem claros na forma de vales espectrais e os formantes nasais como picos onde não encontramos picos na vogal correlata oral. Para o contraste entre [õ] e [u] temos como exemplos as palavras *mo* ['mõ] e *mamo* [mɐ.'mɔ]. A primeira significa 'osso' e a segunda é o nome em Arutani para 'roça de maniva'. Acusticamente elas têm as seguintes imagens:

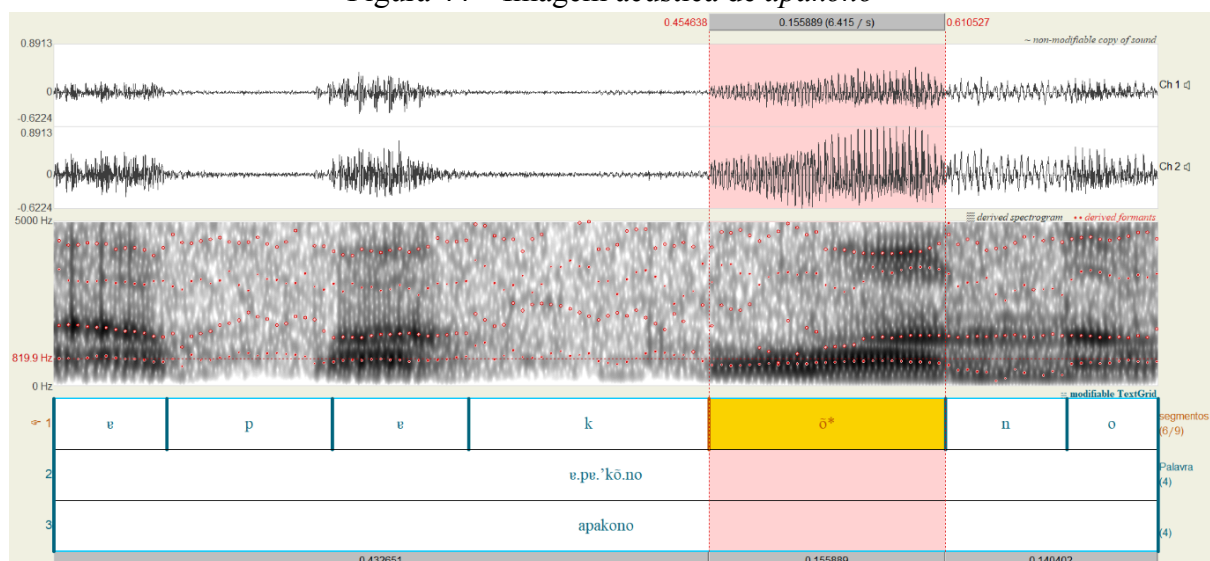
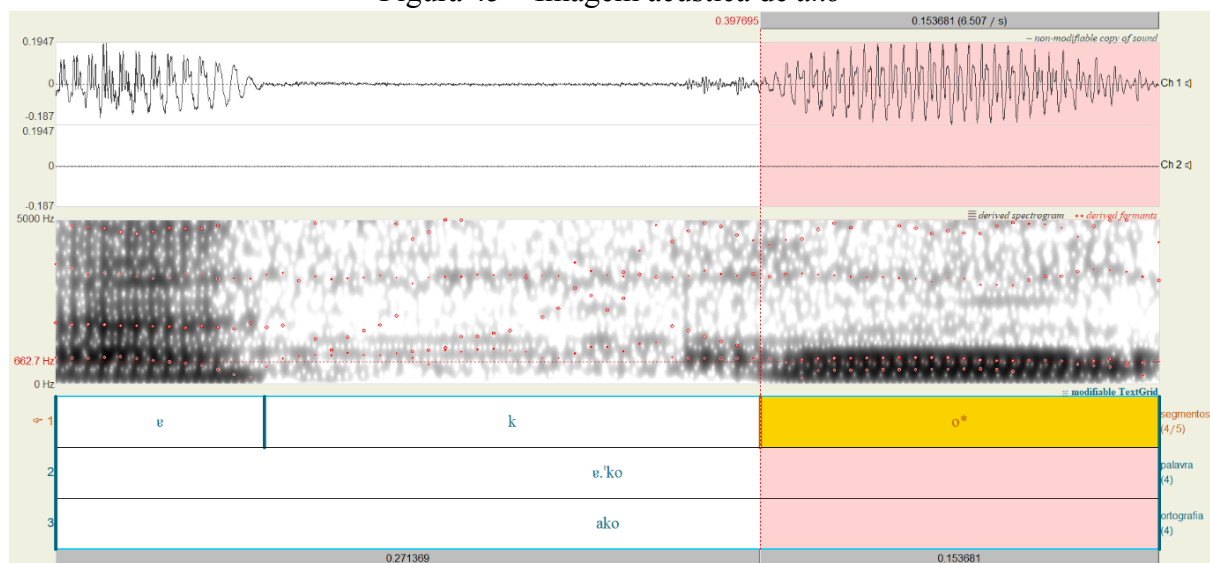
Figura 41 – Imagem acústica de *mo*Figura 42 – Imagem acústica de *mamo*

Com relação à comparação entre os espectros das duas vogais temos o seguinte:

Figura 43 – [õ] vs. [o]

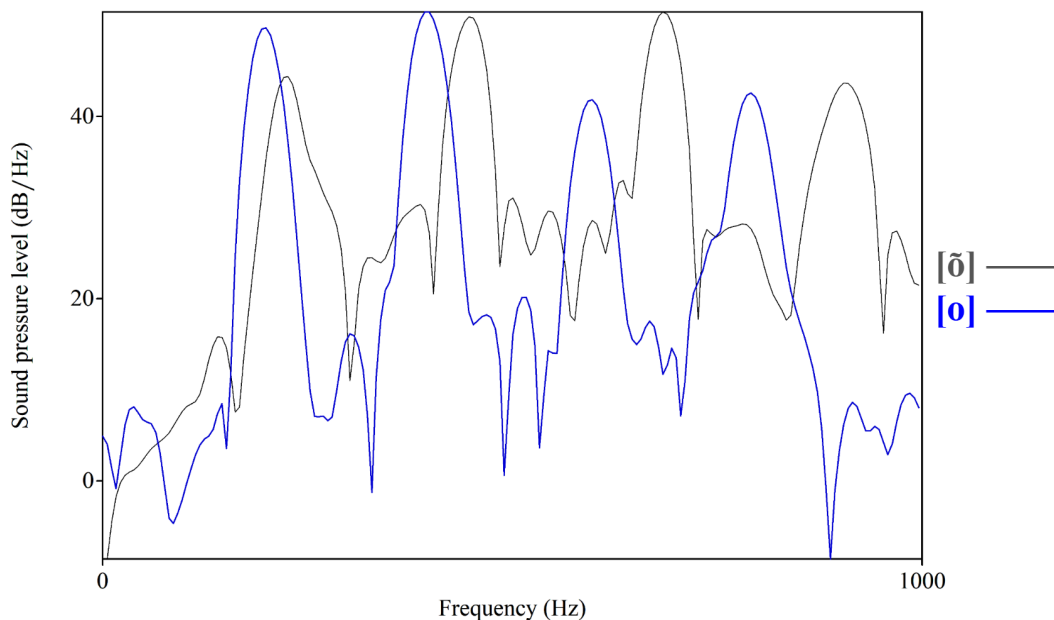


Mais uma vez fica clara a distinção entre os formantes nasais e orais, por conta dos picos espectrais que não existem na vogal correlata [o]. O vale no espectro próximo ao 0 do eixo x do gráfico mostra-se ser referente aos antiformantes. Por último, temos a vogal nasalizada [õ] e sua correlata oral [o]. Elas são encontradas em palavras como *apakono* [ɐ.pɐ.'kõ.no] e *ako* [ɐ.'ko]. A palavra *apakono* significa ‘falecidos’ ou ‘mortos’ e *ako* é o nome que se dá para o ‘cocho de caxiri’. Essas duas palavras têm as seguintes imagens acústicas:

Figura 44 – Imagem acústica de *apakono*Figura 45 – Imagem acústica de *ako*

Quando comparamos o espectro das vogais [õ] e [o] a partir de uma análise FFT temos o seguinte gráfico:

Figura 46 – [õ] vs. [o]



Nesse contraste espectral, temos vales pouco acentuados, mas que ainda são a representação gráfica dos antiformantes. Quanto aos formantes nasais, os picos que os caracterizam já estão mais evidentes ao longo do espectro.

A análise feita nesta seção mostrou a existência de vogais que sofrem nasalidade na fala de Juanita por meio de diversas pistas acústicas que vão desde traços como diferença nos valores de formantes orais até comparações espectrais. Com isso, seguimos para a análise e descrição dos ditongos e hiatos encontrados em nosso *corpus*.

3.2 Glides

Os *glides* encontrados na fala de Juanita são [j] e [w]. Este tipo de segmento é sempre produzido acompanhando alguma vogal. De uma perspectiva fonética, podemos dizer que eles formam ditongos quando acompanham apenas uma vogal. Entretanto, isso pode mudar de uma perspectiva da fonologia e de acordo com o sistema fonotático das línguas. Nesta seção, chamaremos esse tipo de encontro de ditongo, mas veremos na seção 4.3, “Segmentos Ambivalentes”, qual o real *status* deles na língua no âmbito fonológico.

Os ditongos, tanto orais quanto nasais, encontrados na fala de Juanita formam 16 combinações possíveis entre os segmentos vocálicos e os *glides*, os quais podem formar ditongos crescentes (*glide* + vogal → GV) e decrescentes (vogal + *glide* → VG). Esses tipos de ditongos são extremamente comuns e produtivos na fala de Juanita. Veja as combinações possíveis:

Tabela 19 – Ditongos crescentes e decrescentes

Ditongos Crescentes		Ditongos Decrescentes
[ja]	[jo]	[aj] ~ [ãj]
[jæ]	[wa]	[ej]
[je] ~ [jẽ]	[wɐ] ~ [wẽ]	[aw]
[jo]	[wi]	[ɐw]
		[uj]

Cada uma dessas combinações ocorre em palavras como:

- (5)
- | | |
|--|--|
| <i>oya</i> [o.'ja] ‘apimentado’ | <i>mariya</i> [mɐ.ɾi.'jæ] ‘faca’ |
| <i>yama kosa</i> [jɐ.'ma ko.'sa] ‘mingau de mandioca’ | <i>oyañã</i> [o.jẽ.'jẽ] ‘espinho’ |
| <i>mamitiri ayo</i> [mɐ.mi.'ti.ɾi ɐ.'jo] ‘nós’ (inclusivo) | <i>kuyupi</i> [ko.jo.'pi] ‘muchiva’ |
| <i>ãrãwa</i> [ɐ.ɾɐ.'wa] ‘traíra’ | <i>awarika</i> [ɐ.wɐ.ɾi.'ka] ‘quando’ |
| <i>sawan</i> [sɐ.'wɛn] ‘cutia’ | <i>awini</i> [ɐ.wi.'ni] ‘miçanga’ |
| <i>xaray</i> [ʃɐ.'raj] ‘pente’ | <i>anay</i> [ɐ.'nãj] ‘tia’ |
| <i>ãyhã</i> [ɐj.'hẽ] ‘jacamim’ | <i>marakupaw</i> [mɐ.,ɾɐ.ko.'paw] ‘rã’
(espécie comestível) |
| <i>awhi</i> [ɐw.'hi] ‘sururina’ | <i>xipa kuyta</i> [ʃɐ.,pi koj.'ta] ‘flor’ |

Acusticamente, os ditongos são caracterizados pela transição formântica e pela duração, segundo Silva *et al.* (2019, p. 129), sendo a duração um traço pouco confiável e preciso, por isso, iremos dar mais ênfase para a transição formântica. De acordo com estes autores, a alteração nos ditongos, em termos de formantes, manifesta-se principalmente nos dois primeiros formantes (F1 e F2), essa variação pode ser visualizada nos espectrogramas, onde se nota uma mudança contínua na trajetória das barras escuras horizontais correspondentes aos formantes.

Para uma análise mais objetiva, iremos comparar a transição formântica dos dois tipos de ditongos encontrados na fala de Juanita, crescentes e decrescentes. Escolhemos como exemplos os ditongos [ja] e [aj]. Esses ditongos não ocorrem em ambientes idênticos na fala de Juanita, portanto, iremos analisar palavras em que os ditongos ocorrem em sílaba tônica. Os ditongos a serem analisados estão presentes nas palavras *xipuya* [ʃi.pu.'ja], que significa ‘pó’ e *xaray* [ʃɐ.'raj] que significa ‘pente’. Observe suas imagens acústicas, respectivamente.

Figura 47 – Imagem acústica de *xipuya*

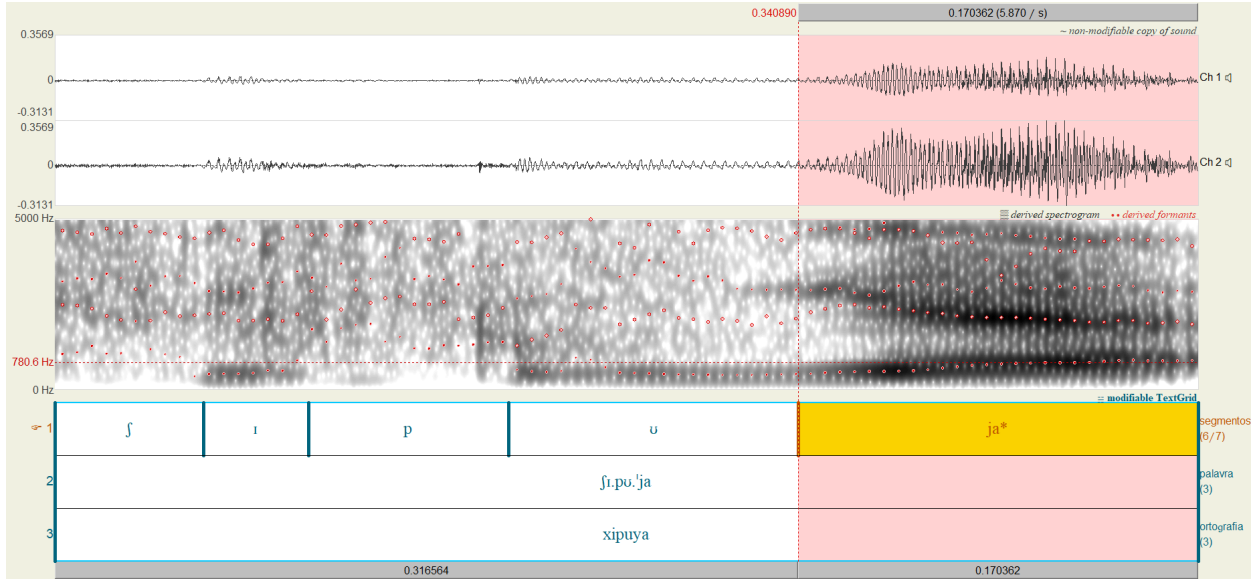
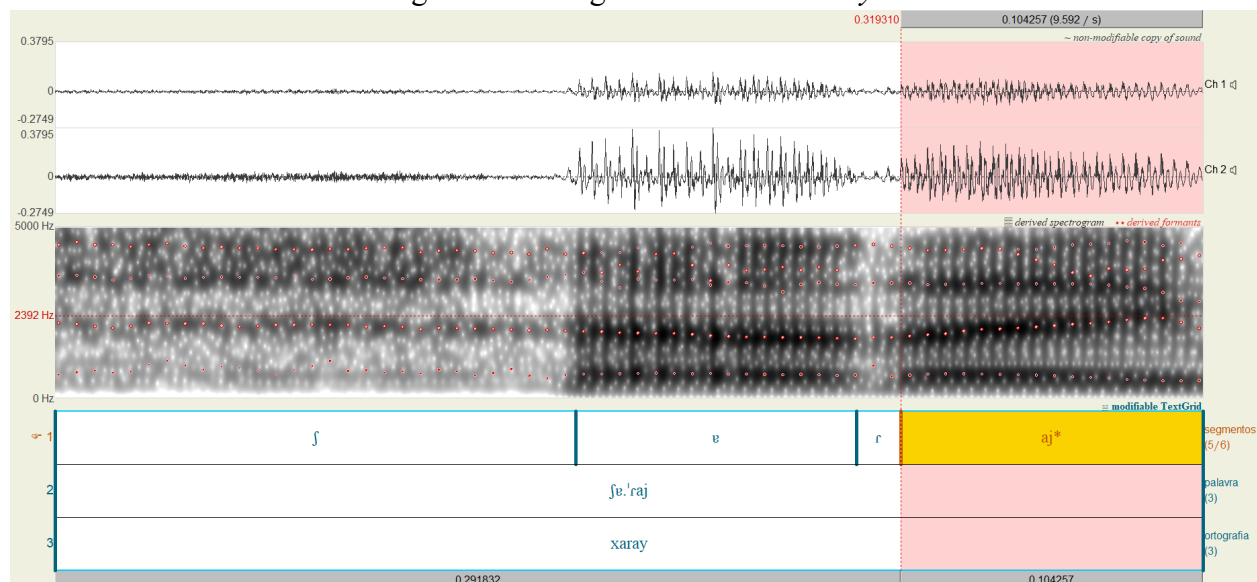


Figura 48 – Imagem acústica de *xaray*

Quanto à duração dos ditongos, temos [ja] com 170 ms e [aj] com 104 ms. Observe o sentido decrescente de F2 e convergente entre F1 e F2 em [ja], o que nos mostra uma tendência já esperada. O contrário ocorre em [aj], já que temos os formantes em sentido divergente. Isso reflete correlatos articulatórios importantes.

Com relação ao F1 em [ja] ele tende a aumentar, já que articulatoriamente a língua sai de uma posição mais próxima ao palato para uma posição mais baixa. Já em [aj] o F1 tende a diminuir, pois a língua muda de uma posição mais baixa e tende a ser deslocada para uma direção mais alta, rumo ao palato. Pelo fato de o *glide* [j] ser articulado numa região mais anterior do trato vocal ele tende a ter um F2 mais alto, o qual vai diminuindo ao se coarticular com o [a] na pronúncia de [ja], tendo em vista que [a] possui um valor de F2 mais baixo.

O inverso acontece com [aj], o F2 de [a] tende a aumentar gradativamente de acordo com que Juanita altera a posição de sua língua, com o mesmo fluxo de ar, para produzir o *glide* [j]. Articulatoriamente, o F2 representa em [ja] a língua saindo de uma posição mais anterior e movendo-se para uma posição mais centralizada. Observe essas mudanças articulatórias ocorrendo na diminuição ou no aumento dos valores iniciais e finais de [ja] e [aj].

Tabela 20 – Valores de F1 (inicial vs. final) e F2 (inicial vs. final) na transição formântica

[ʃi.pu.'ja] → [ja]	Inicial	F1:	492 Hz
		F2:	2.453 Hz
	Final	F1:	827 Hz
		F2:	1.976 Hz
[ʃe.'raj] → [aj]	Inicial	F1:	650 Hz
		F2:	1.792 Hz
	Final	F1:	488 Hz
		F2:	2.302 Hz

Uma característica dos *glides* intervocálicos encontrados na fala de Juanita é a ambissilabidade, ou seja, *glides* que “transitam” ou estão presentes em duas sílabas: ocorrendo em ataque em ditongos crescentes e em coda em ditongos decrescente. Isso pode ser ilustrado na palavra *muya*, que significa ‘buraco’, a qual pode ser transcrita como: [mu.'ja] ou [moj.'ja]. Veja as diferenças nas respectivas imagens acústicas:

Figura 49 – Imagem acústica de [mu.'ja]

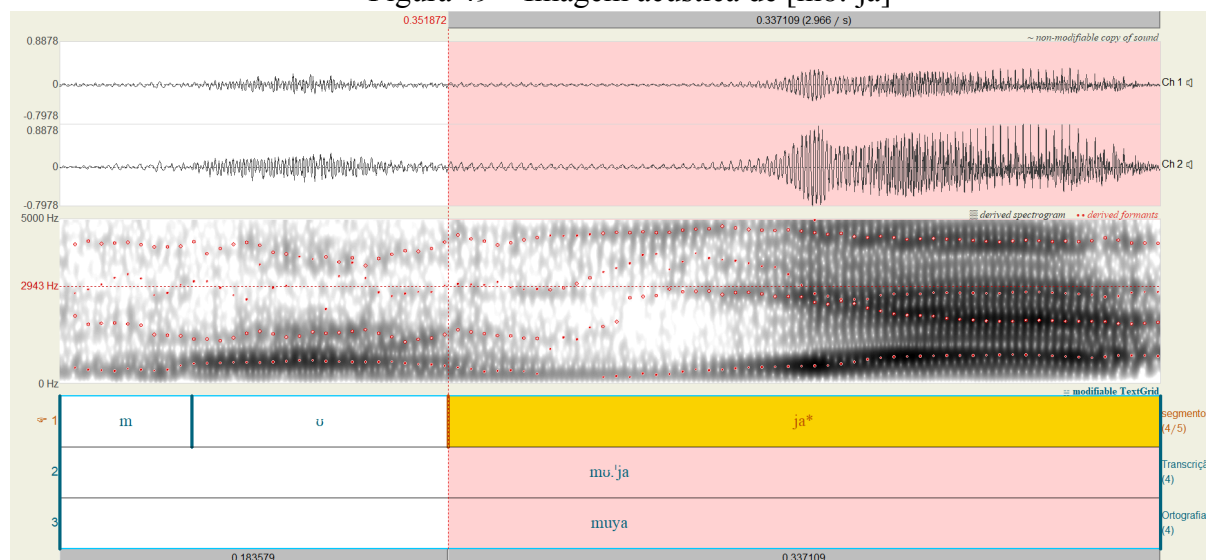
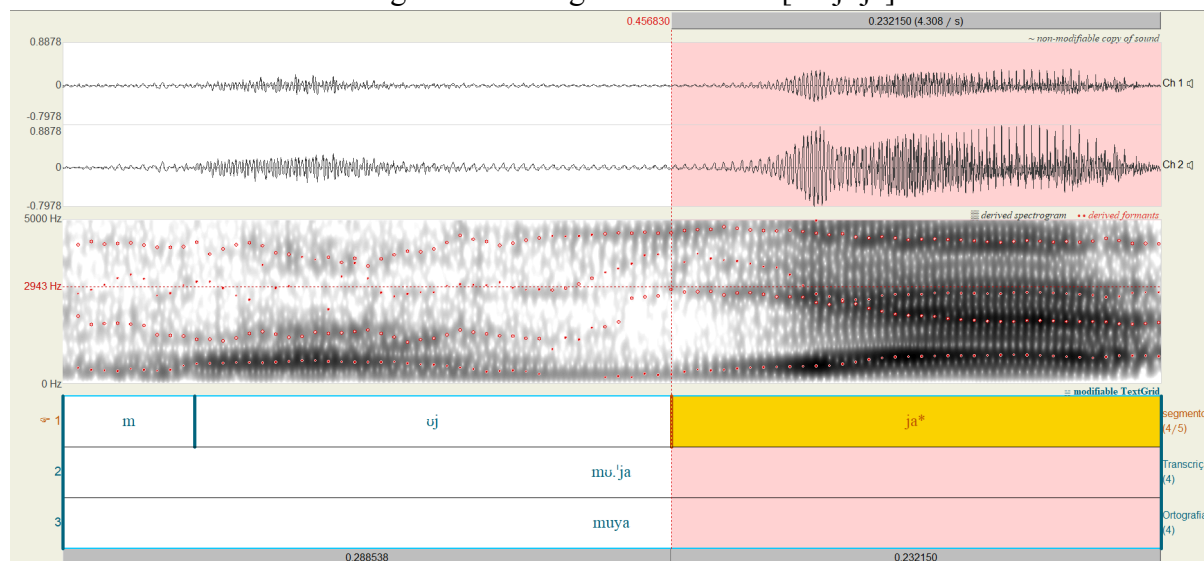
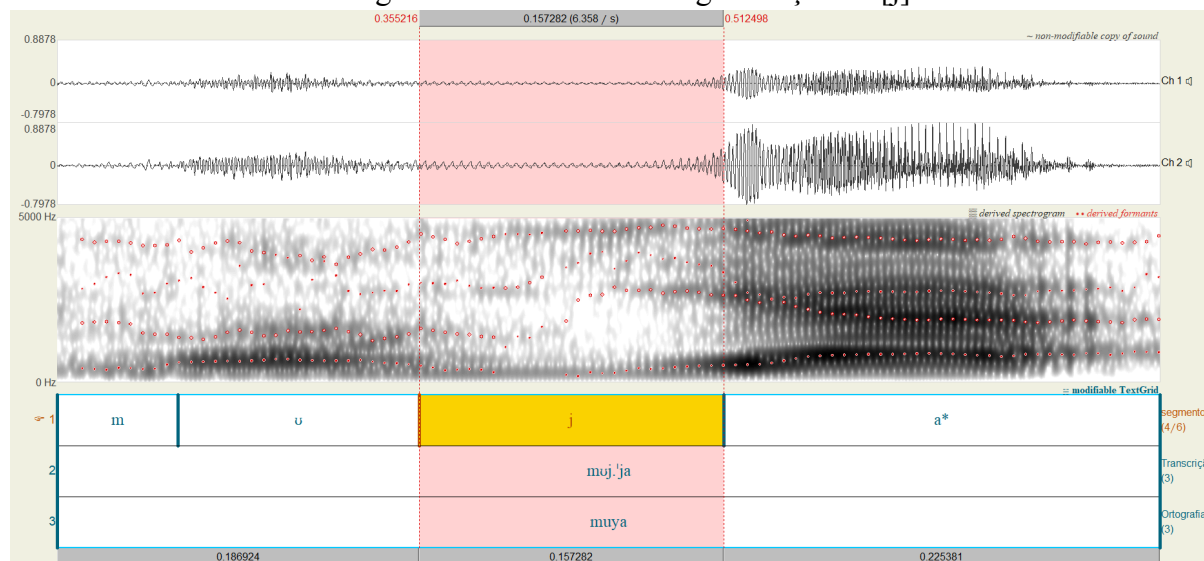


Figura 50 – Imagem acústica de [moj.'ja]



Dessa forma, fica difícil segmentar a posição exata do *glide* [j], mas podemos identificar o momento acústico em que ele ocorre plenamente onde quase não se escuta as vogais [u] ou [a]. Essa tentativa de segmentação pode ser observada na Figura 51.

Figura 51 – Tentativa de segmentação de [j]



O *glide* [j] só pode estar adjacente aos fonemas /a/ e /o/, e seus alofones na mesma sílaba, enquanto [w] só ocorre adjacientemente na mesma sílaba com os fonemas /a/ e /i/. Portanto, estruturas como *[ji], *[ij], *[ij] e *[jɪ] não ocorrem na língua, talvez por se trata de um

“redundância articulatória”, a qual está indiretamente relacionada com o *Obligatory Contour Principle* (OCP – Princípio do Contorno Obrigatório; cf. Matzenauer, 2014, p. 66). O mesmo acontece com *[wo], *[ow], *[wɔ] e *[ɔw]. Com isso, agora, podemos explorar encontros entre vogais em sílabas diferentes e, também, vogais seguidas de *glides* em sílabas distintas.

3.3 Hiatos e estratégias para a demarcação de fronteira silábica

Na fala de Juanita encontramos 15 tipos de sequências heterossilábicas com estratégias para a demarcação da fronteira silábica. Encontramos alguns hiatos (V.V) e algumas vogais seguidas por *glides* (V.G). Veja na seguinte tabela quais são eles:

Tabela 21 – Hiatos existentes na língua

Hiatos		
[ɐ.a]	[ɪ.j]	[o.ɪ]
[ɐ.j]	[ĩ.õ]	[o.j]
[ɪ.a]	[ɪ.ʊ]	[o.o]
[ɪ.æ]	[o.a]	[ʊ.i]
[ɪ.e]	[o.i]	[w.ĩ]

Os hiatos em Arutani são geralmente demarcados pela presença de uma pausa glotal completa ou pela laringalização da voz na vogal átona e/ou na transição entre as duas vogais ou até mesmo pelo uso de *glides*. Um exemplo da ocorrência de uma pausa glotal bastante enfática é a que ocorre na palavra *oi* [oʔ.ˈi] ‘cunhado’, Figura 52. Já o caso de uma pausa glotal menos enfática ocorre na palavra *sui* [soʔ.ˈi] ‘inambu’ (passáro), Figura 53. Outra estratégia é a laringalização com voz crepitante, a qual pode ser vista na palavra *aay* [ɛ.ˈaj] ‘negativo, não’, na Figura 54. A última estratégia é o uso de um *glide* para demarcar a fronteira entre as vogais, e pode ser visto na palavra *toropia* [to.ro.pi.ˈja] ‘chifre’, Figura 55.

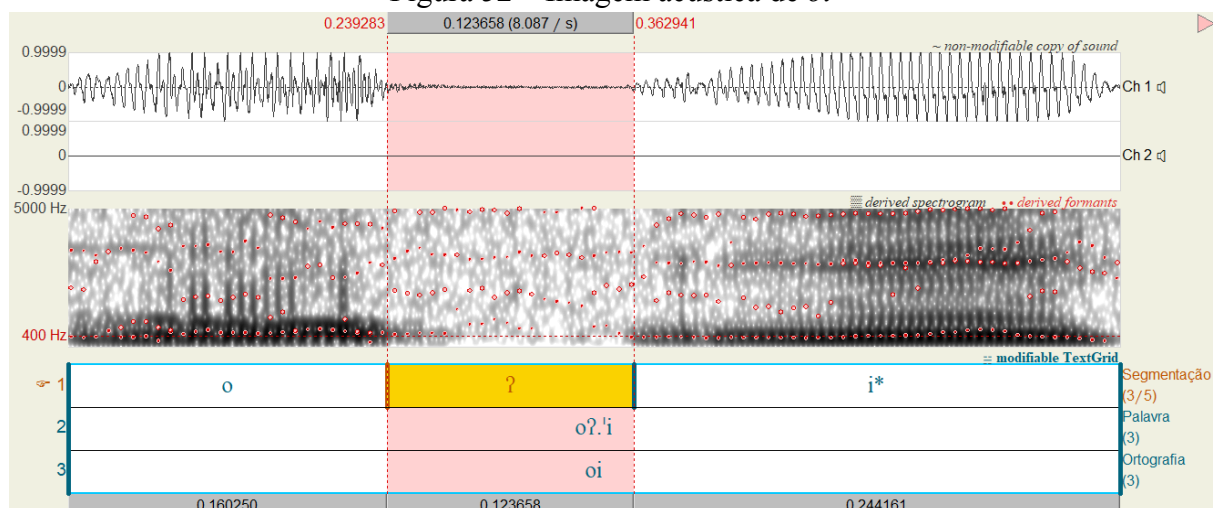
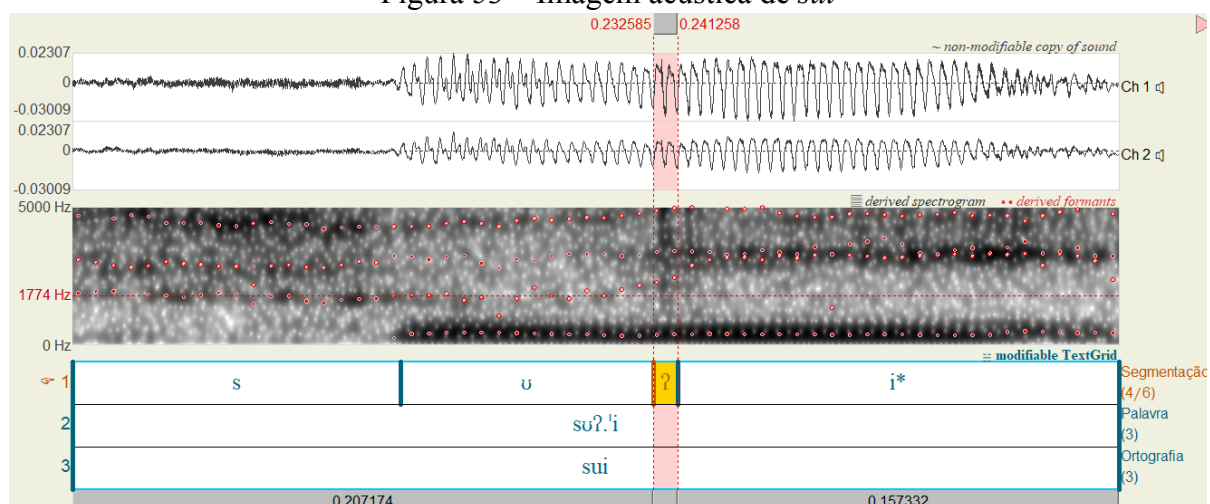
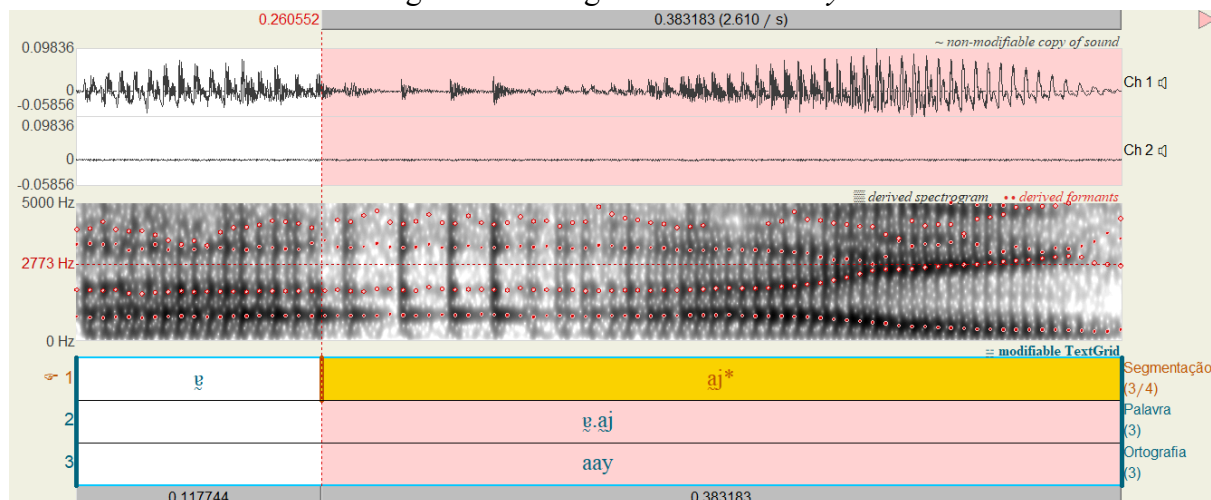
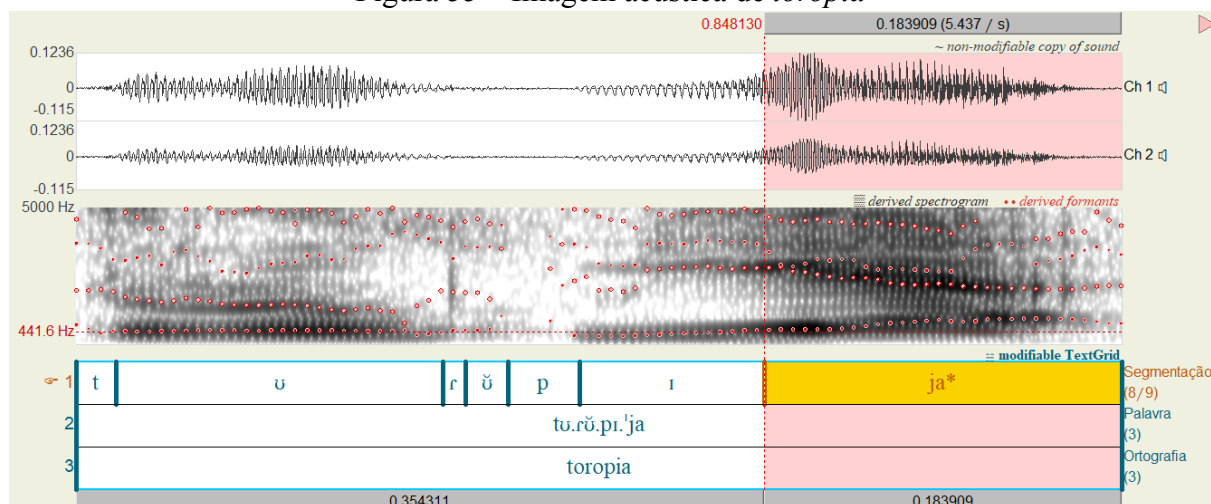
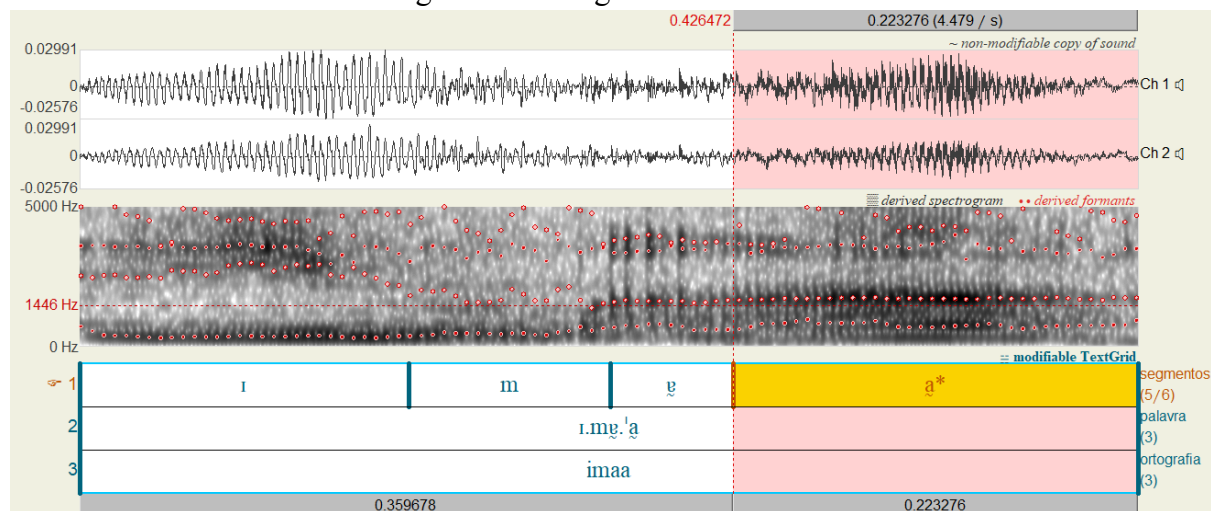
Figura 52 – Imagem acústica de *oi*Figura 53 – Imagem acústica de *sui*

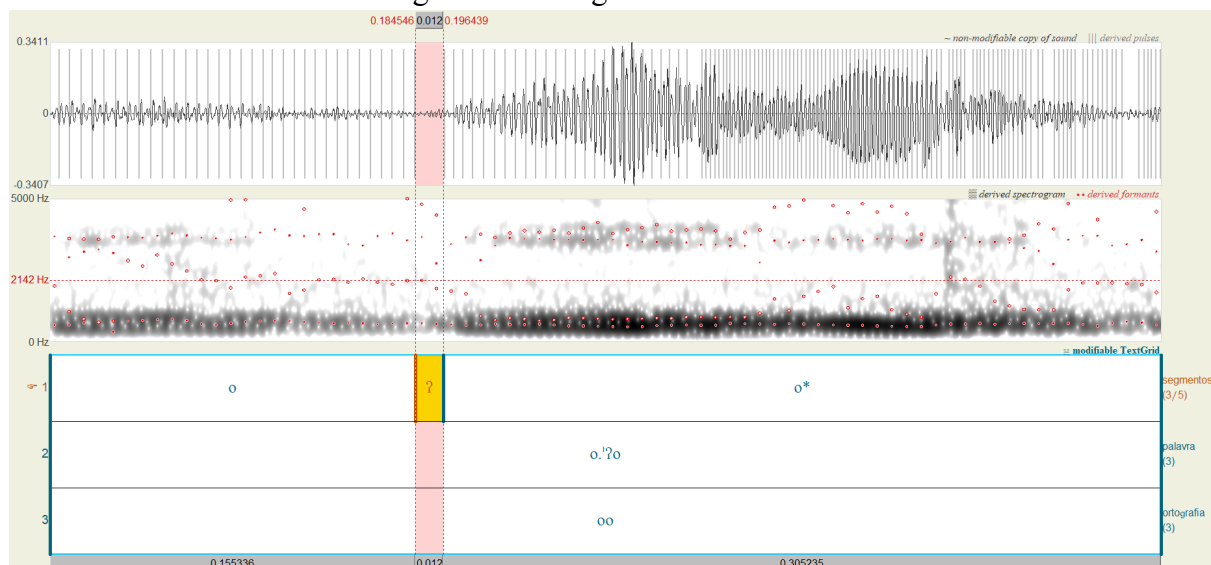
Figura 54 – Imagem acústica de *aay*Figura 55 – Imagem acústica de *toropia*

Muitos dos hiatos do Arutani ocorrem em diversas línguas, como no próprio Português do Brasil. Entretanto, destacamos três desses hiatos que chamaram bastante nossa atenção em nossas análises e que foram de difícil compreensão e distinção. São eles, os encontros [ɐ.a], [o.o] e [ɪ.j]. O principal motivo deles terem chamado nossa atenção é pelo simples fato de serem sons bastante semelhantes, e até mesmo idênticos. No caso de [ɐ.a], temos o exemplo na palavra *imaa* [ɪ.mɔ̃.'a], que se refere à ‘parte interior da boca’ ou ‘cavidade bucal’, nela acontece o mesmo fenômeno que ocorre em *aay* [ɐ.'aj] ‘negativo, não’, que é a laringalização, conhecido também como *creek voice*. Observe a seguinte imagem acústica:

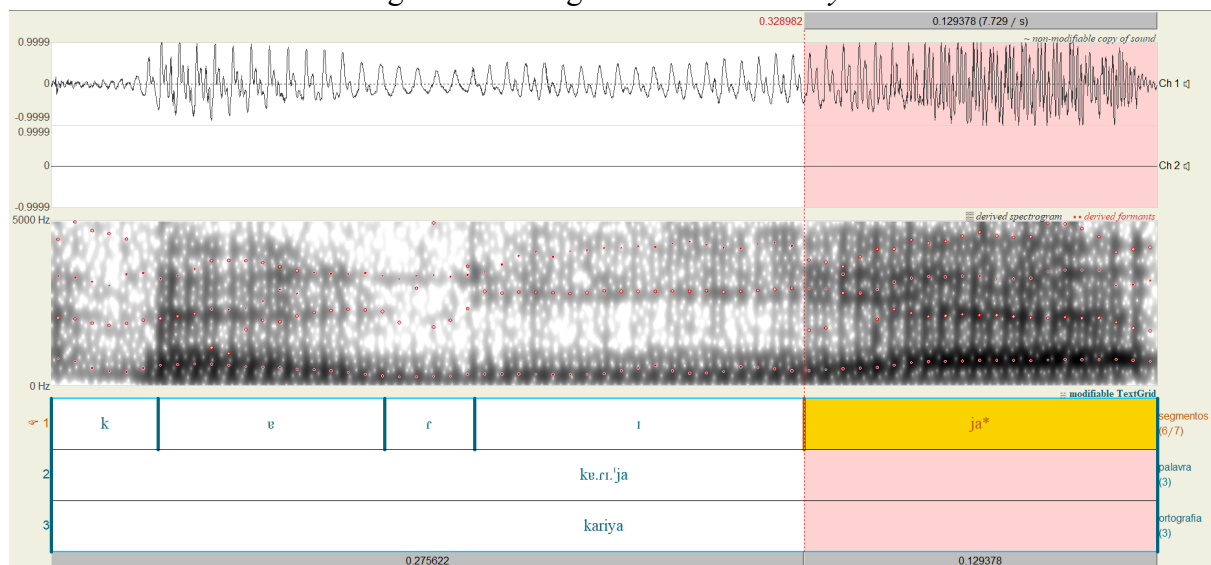
Figura 56 – Imagem acústica de *imaa*

Um ponto que respalda a segmentação adequada é quando medimos os formantes no ponto central, principalmente o F1, que é o traço mais distintivo entre [ɐ] e [a] laringalizados. Dessa forma, [ɐ] possui um F1 de 701 Hz e um F2 1.609 Hz, enquanto [a] possui um F1 de 776 Hz e um F2 1.684 Hz, provando que [ɐ] é realmente uma vogal mais fechada e centralizada com relação à [a]. Em [i.mɐ.'a] ‘cavidade bucal’, observamos a laringalização tanto na vogal tônica [a], quanto na vogal átona [ɐ], semelhantemente ao que ocorre em *aay* [ɐ.'aj], em que as duas vogais que compõem o hiato são laringalizadas.

Em *oo* [o.'ʔo], ‘criança’ ou ‘bebê’, há a presença breve da oclusiva glotal [ʔ]. Com auxílio do recurso *pulse* do Praat, foi possível, de forma mais clara, observar a realização dessa pausa. Neste caso, temos uma oclusiva glotal [ʔ] bem breve, menos enfática – mais um dos recursos de demarcação entre segmentos vocálicos heterossilábicos –, com uma duração de 12 ms, tendo o primeiro [o] 155 ms e o segundo [o] 305 ms de duração.

Figura 57 – Imagem acústica de *oo*

Um terceiro caso bastante interessante é o que ocorre na palavra *kariya* [kɐ.ɾi.'ja], que é o termo de parentesco que pode ser traduzido como ‘irmã mais nova de alguém’. Este é mais um exemplo de uso de *glide* para delimitar a fronteira entre as vogais. Veja a imagem acústica:

Figura 58 – Imagem acústica de *kariya*

Vale ressaltar que não é qualquer tipo de *glide* que pode ocorrer com essa função de fronteira entre vogais. Sílabas terminadas em [ɪ], por exemplo, neste contexto de fronteira

silábica, só podem ser seguidas pelo *glide* articulatoriamente mais próximo, no caso, o [j]. O mesmo valeria para [ʊ.w] e [o.w], caso esse tipo de encontro ocorra na fala de Juanita. Um recurso para a identificação, diferenciação e segmentação entre [ɪ] e [j] é a observação do formato de onda e a realização de uma análise oitiva precisa. Ao ouvirmos o áudio de *kariya* [kɐ.ɾɪ.ˈja] fica nítido o momento em que Juanita altera sua pronúncia de [ɪ] para [j].

Um último recurso que ajuda na segmentação são os formantes no espectrograma. Observe que os dois primeiros formantes de [ɪ], na Figura 58, estão bem claros no espectrograma, estão em linhas quase retas, enquanto, no trecho de segmentação em que há o *glide* [j], há uma perturbação grande e uma imprecisão na demarcação dos formantes, característica comum em *glides* (cf. Silva *et al.*, 2019, p. 132). Uma das hipóteses para isso é que, por conta da rápida produção do *glide* e por tratar-se de uma aproximante, o Praat não consegue identificar precisamente onde esses formantes estão ocorrendo e estão sendo produzidos. Portanto, o que na verdade é uma limitação do *software*, passou a ser um recurso que fornece mais uma pista acústica para a distinção de segmentos semelhantes.

3.4 Segmentos Consonantais

3.4.1 Consoantes Oclusivas

Em nosso *corpus* foram encontrados as seguintes consoantes oclusivas [p] (bilabial não-vozeada), [p̚] (bilabial não-vozeada não-explodida), [t] (alveolar não-vozeada), [t̚] (alveolar não-vozeada não-explodida), [k] (velar não-vozeada), [b] (bilabial vozeada) e [b̚] (bilabial vozeada não-explodida). De uma perspectiva quantitativa, é interessante sabermos quais desses segmentos oclusivos são mais produtivos na língua. Sendo assim, temos os seguintes quantitativos:

Tabela 22 – Quantitativo de consoantes oclusivas

Consoantes Oclusivas	Nº de Ocorrências
[k]	200
[t]	93
[t̃]	5
[p]	58
[p̃]	16
[b]	4
[b̃]	1

Dessa forma, pode-se perceber o [k] como sendo o fone mais produtivo, ocorrendo 200 vezes, sendo que esse é número de vezes em que ele ocorre, e não o número de palavras em que ele é produzido, ou seja, em palavras como *kikayu* [kɪ.kɐ.'jʊ], termo de parentesco que significa ‘avô dele’ ou ‘sogro dele’, temos duas ocorrências de [k]. A segunda mais produtiva é o [t], seguida pelo [p]. Os menos produtivos são o [b] e [b̃], os quais ocorrem em raros casos e em determinados contextos, o que nos leva a questionar seu *status* fonológico, o que será explorado na seção 4.1.4.

Antes de tudo, precisamos compreender quais seriam as principais características acústicas que os segmentos oclusivos possuem e que nos evidenciam que estamos certos em sua classificação de uma perspectiva tanto articulatória quanto acústica. Barbosa e Madureira (2015, p. 173) apresentam essas duas classificações como complementares e mostram uma visão não apenas acústica, mas, também, articulatória do fenômeno da oclusão, no qual as consoantes oclusivas são caracterizadas “[...] articulatoriamente por uma obstrução total, formada pelo contato entre articuladores, o que impede a passagem da corrente de ar, seguida de liberação momentânea dessa corrente, o que ocasiona uma plosão”. Os autores ainda destacam que, de uma perspectiva acústica, as obstruções criadas durante a produção de sons não-vozeados resultam em um silêncio e quando o ar é liberado, ocorre um ruído breve e intenso, conhecido como plosão ou ruído transiente. Já nas oclusivas vozeadas, a fase de obstrução apresentará uma

"barra de vozeamento", que é a representação da vibração das pregas vocais (BARBOSA; MADUREIRA, 2015, p. 173).

Barbosa e Madureira (2015, p. 173), portanto, explicam as seguintes características acústicas das consoantes oclusivas: “silêncio de oclusão/barra de vozeamento, ruído transiente e transição até o início da vogal”. Complementarmente, o autor traz o conceito de VOT (*Voice Onset Time*) e o apresenta como delimitador do intervalo de interesse para o estudo do vozeamento e que, quanto à sua ocorrência ele “[...] pode ocorrer entre o início do ruído transiente e o início da vogal que o segue (VOT positivo) ou entre o início do ruído transiente e o final da vogal que o precede (VOT negativo)”. Dessa forma, os autores diferenciam os sons não-vozeados dos vozeados, tendo os segmentos oclusivos não-vozeados um VOT positivo e os vozeados um VOT negativo.

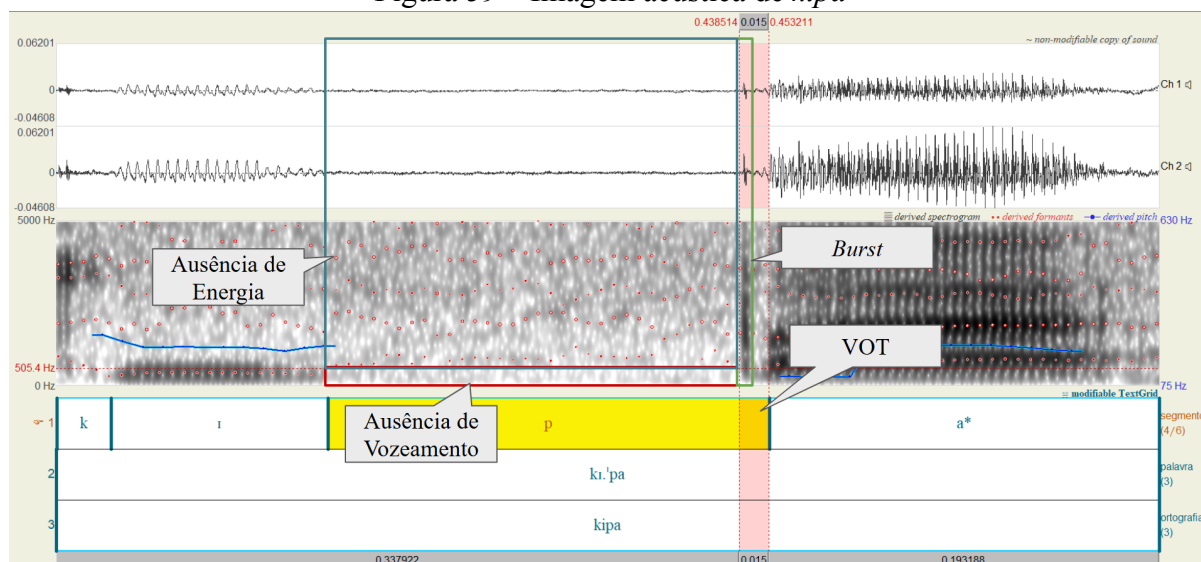
De maneira complementar, Silva *et al.* (2019) apresenta seis características acústicas das consoantes oclusivas, as quais são: “1. ausência de energia; 2. barra de vozeamento ou barra de sonoridade; 3. ruído transiente ou soltura da oclusão (*burst*); 4. VOT; 5. *loci* acústico consonantal ou F2 de transição; 6. configuração espectral da soltura da oclusão”.

Em nossa análise das oclusivas em Arutani presentes na fala de Juanita iremos utilizar esses seis critérios. A escolha desses parâmetros foi feita tendo como base pontos em comum entre os autores apresentados e a relevância para a análise de traços em comum (modo de articulação) e distintivos (ponto de articulação). Ademais, apresentaremos apenas as oclusivas com relação aos ambientes e contextos em que algumas ocorrem, pelo fato de ocorrerem em ambientes distintos e terem uma produção de difícil percepção e distinção, como é o caso do [p], que, a depender do contexto, apenas algumas de suas características acústicas podem ser observadas. Novamente fica evidente a interface entre a fonética e a fonologia (fonotática), trazendo à tona o fato de elas se complementarem, permitindo uma noção mais interrelacional de língua.

Com relação aos ambientes em que o [p] ocorre na língua, iremos iniciar apresentando este fone ocorrendo no meio de duas vogais em posição de início de sílaba, ou seja, em ataque

silábico, visto que, como destacam Barbosa e Madureira (2015, p. 174), “A segmentação da oclusiva no início de um enunciado ou seguida de pausa silenciosa fica prejudicada, pois não se sabe o quanto é silêncio por interrupção da fala ou silêncio referente à obstrução dos articuladores”. Vejamos, então, a imagem acústica do [p] em *kpa* [kɪ.'pa], que é a palavra para ‘urina’. Veja:

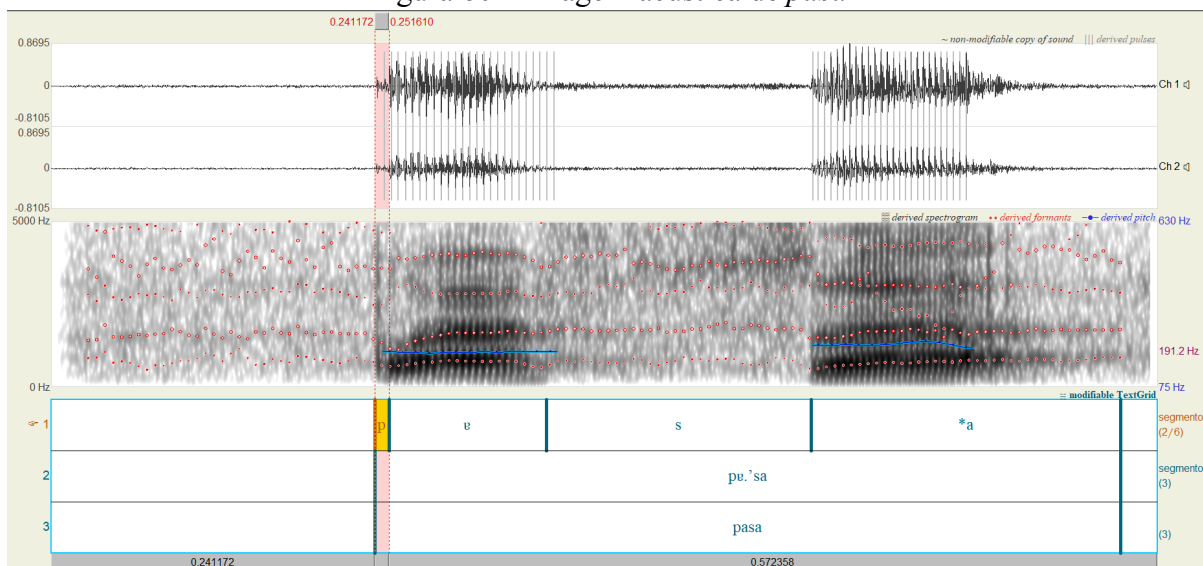
Figura 59 – Imagem acústica de *kpa*



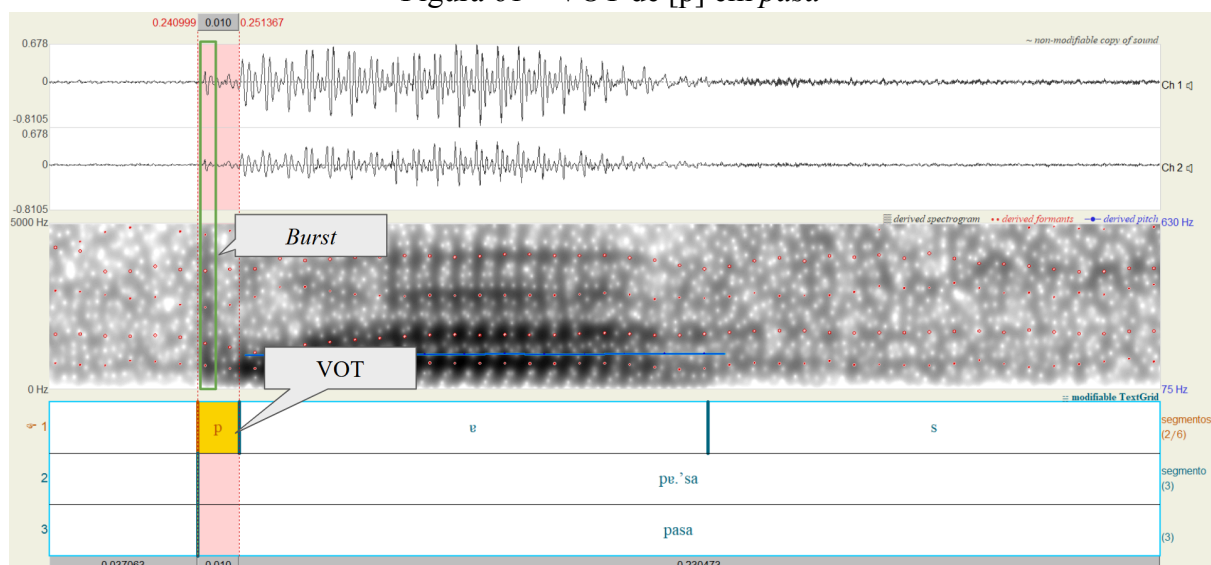
Na Figura 59 podemos observar o [p] com uma duração de aproximadamente 210 ms. Quanto às características de uma consoante oclusiva não-vozeada podemos observar o período de silêncio de oclusão (destacado pelo retângulo vermelho) ocorrendo entre o [ɪ] e o momento da plosão (destacado pelo retângulo verde), evidenciado pela ausência de energia no formato de onda e no espectrograma. O silêncio de oclusão do [p] possui 200 ms, mostrando que o silêncio de oclusão é a maior parte em duração de um segmento oclusivo. Não há barra de vozeamento no espectrograma, o que nos prova que o fone que realmente ouvimos como uma oclusiva não-vozeada está adequado com seus parâmetros acústicos. Em seguida podemos observar o ruído transiente, também conhecido como *burst*, traço conhecido por tender a ser bastante breve. Por fim, temos o VOT. Podemos observar o um VOT positivo de 15 ms, característico de oclusivas não-vozeadas.

No início de palavras, mais especificamente no começo de enunciação, não é possível ser feita uma análise precisa da ausência de energia e, por conta do [p] ser não-vozeado, também não há barra de vozeamento, porém, podemos analisar o *burst* e o VOT. Observe esses traços estando presentes na palavra *pasa* [pe.'sa], que é nome em Arutani para o peixe ‘jundiá’. Vejamos:

Figura 60 – Imagem acústica de *pasa*

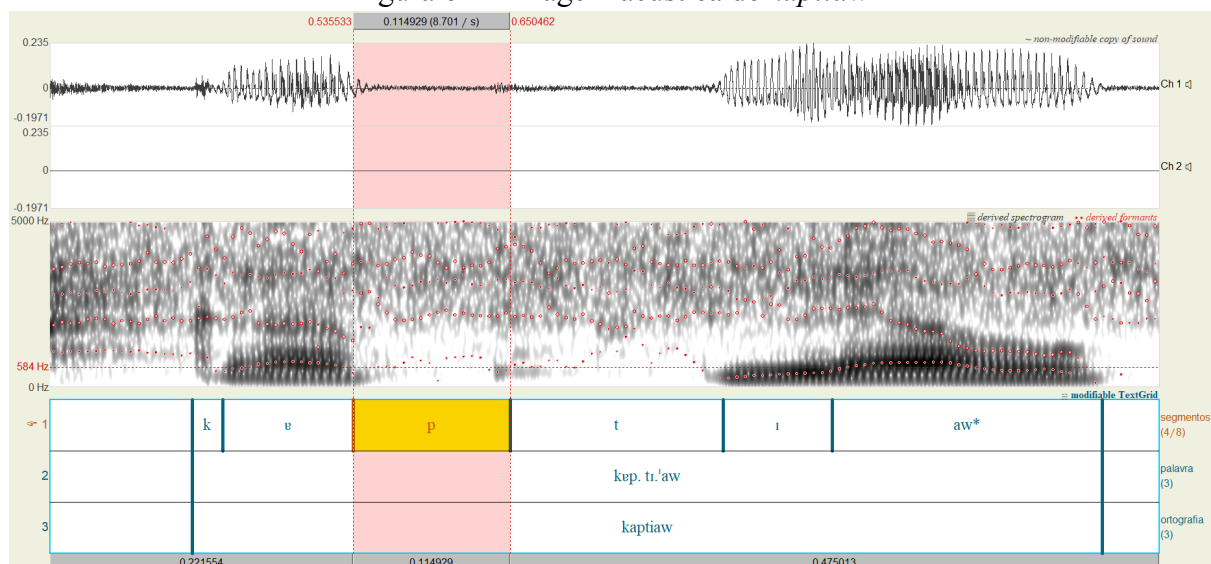


Na Figura 60, podemos observar o momento do *burst* de duas formas, no formato de onda e no espectrograma. Destacamos aqui o menor comprimento da segmentação e a menor duração do [p], visto que não é possível fazer a delimitação com precisão devido ao ambiente em que ele ocorre. Quando analisamos mais de perto este [p], torna-se possível a medição do VOT.

Figura 61 – VOT de [p] em *pasa*

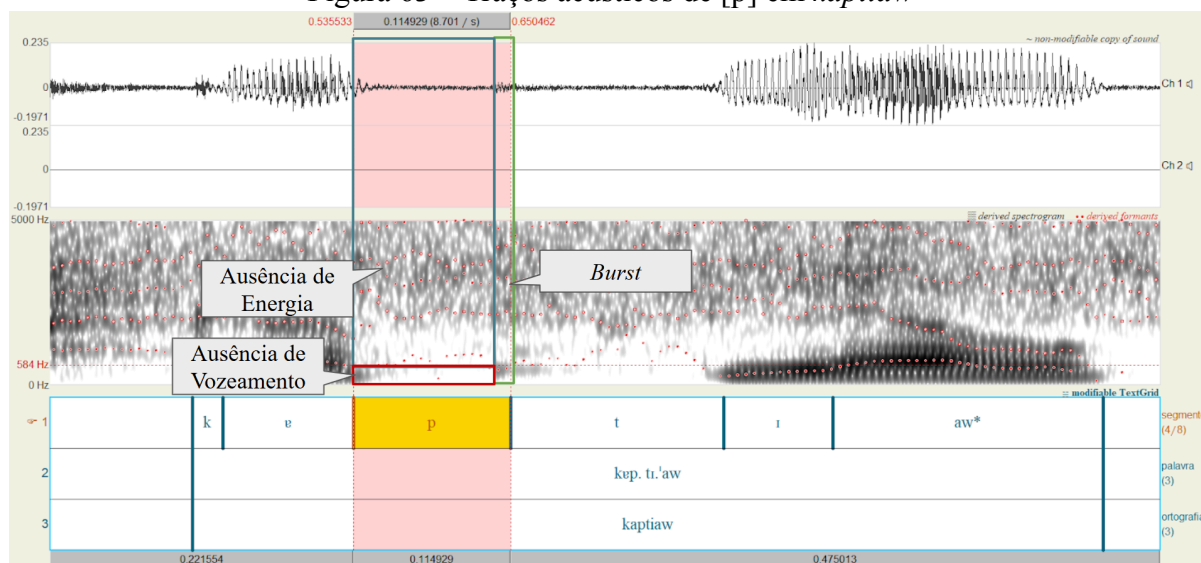
Com base na Figura 61, podemos observar um VOT positivo de 10 s, que corresponde, neste caso, à duração inteira do [p] possível de ser medida.

Um outro contexto em que o [p] ocorre é em posição de final de sílaba (ou coda silábica), como na palavra *kaptiaw* [kep.ti.'aw], que é um 'quantificador' usado para 'quantidades maiores ou iguais a 3'.

Figura 62 – Imagem acústica de *kaptiaw*

Agora, na Figura 62, podemos observar apenas o silêncio de oclusão e o ruído transiente. Não é possível fazer a medição do VOT, visto que não há início de vozeamento, já que o [p] está seguido por outra consoante, que também é não-vozeada. Na Figura 63, a seguir, pode-se observar de vermelho o momento de ausência de energia, no qual há o momento de transição entre o vozeamento da vogal [e] e o desvozeamento do [p], destacado pelo retângulo vermelho, e, em seguida, destacado em verde, está um leve *burst*. Juanita, nessa palavra, produz oclusivas ([p] e [t]) com plosões mais fracas, ou seja, ela as produz com o trato vocal mais relaxado, isso é refletido na imagem acústica, em que os ruídos transientes são observados com uma baixa frequência tanto na forma de onda quanto do espectrograma.

Figura 63 – Traços acústicos de [p] em *kaptiaw*

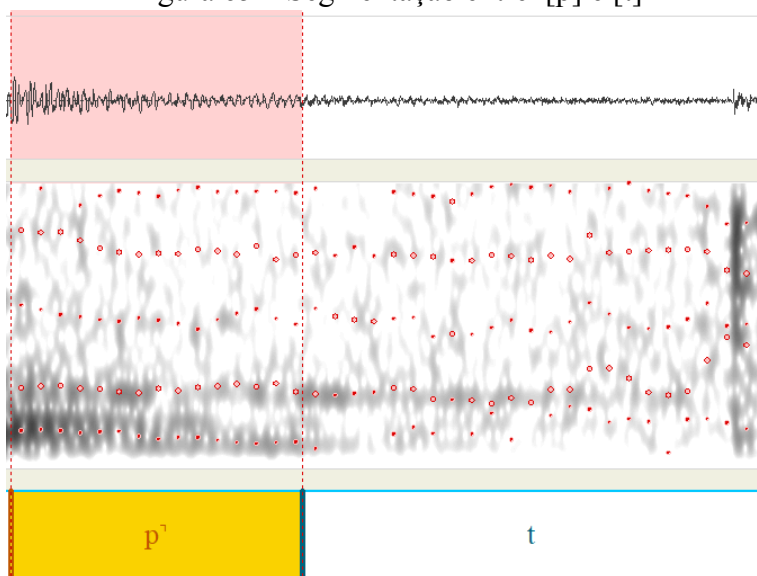


O [p] em coda silábica em meio de palavra também pode ocorrer sem a presença do *burst*, sendo transcrito como [p̄], com uso do diacrítico. Veja a imagem acústica de [p̄] na Figura 64 na qual se pode observar a palavra *kapti* [kep̄. 'ti], que significa ‘ponta’.

Figura 64 – Imagem acústica de *kapti*

O [p̌] em [kəp̌.'ti] possui as mesmas características acústicas do [p] em [kəp.ti.'aw], com exceção do ruído transiente. Tanto que se torna difícil a delimitação exata entre o final da realização do [p̌] e o início do [t] por conta da ausência da plosão. Para isso, utilizamos como recurso a mudança dos formantes e, conseqüentemente, da frequência no espectrograma e uma leve mudança no formato de onda.

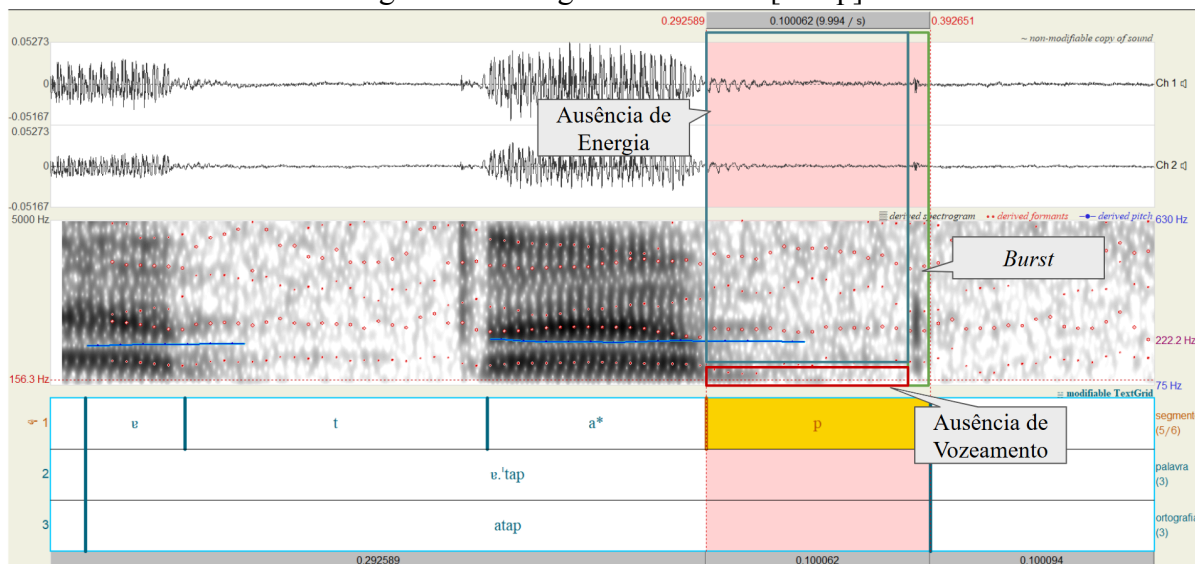
Figura 65 – Segmentação entre [p̌] e [t]



Nos casos anteriores vimos o [p] e o [p̃] ocorrendo no final da sílaba, porém, no meio da palavra. Agora iremos observar o [p], ainda em coda silábica, mas ocorrendo no contexto de final de palavra. A análise do [p] fica mais complexa nesse ambiente, pois ocorre uma certa variação na descrição do fonema tanto de uma perspectiva de coarticulação quanto de vozeamento. Podemos observar esses dois pontos em apenas uma palavra, a qual é grafada como *atap* ‘lua’, porém em três ocorrências diferentes foi transcrita de três formas diferentes, como: [v.ˈtap], [v.ˈtap̃] e [v.ˈtab̃]. Portanto, a partir de agora analisaremos esses três segmentos, respectivamente: [p], [p̃] e [b̃].

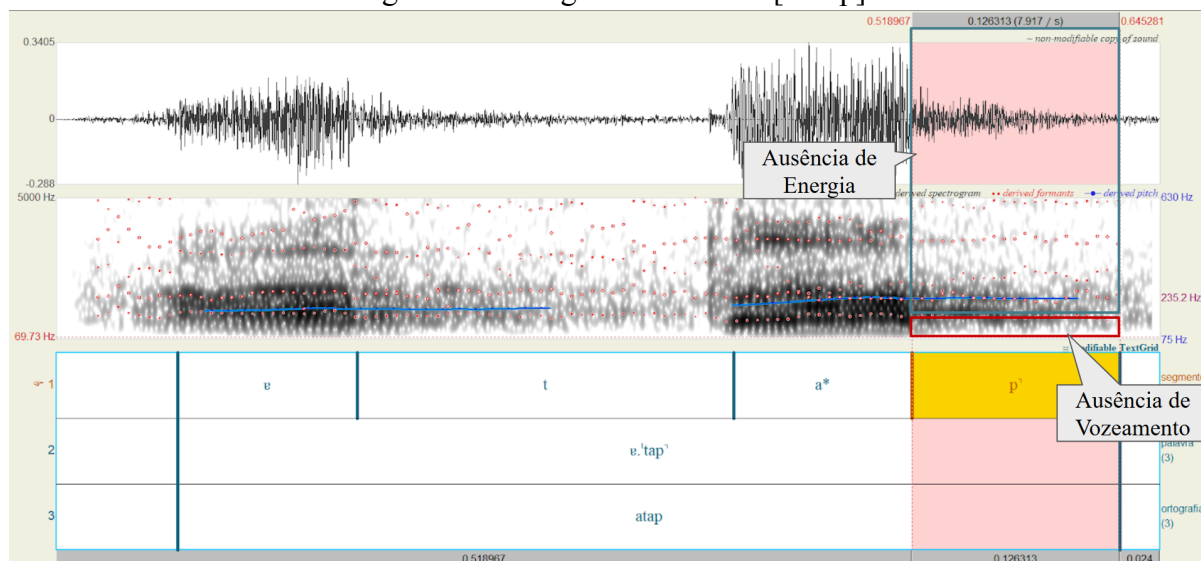
Em [v.ˈtap] podemos observar um [p], o qual é explodido e possui o momento de ausência de energia, sem a presença de uma barra de vozeamento. Note que na Figura 66, a seguir, não há barra de vozeamento no espectrograma por se tratar de uma oclusiva não-vozeada, mas notamos um *burst* nítido no espectrograma. Não há momento de VOT, pois não existe nenhuma vogal após o [p] pelo simples fato de ele estar em final de palavra. O que vemos após a segmentação do [p] são apenas ruídos.

Figura 66 – Imagem acústica de [v.ˈtap]



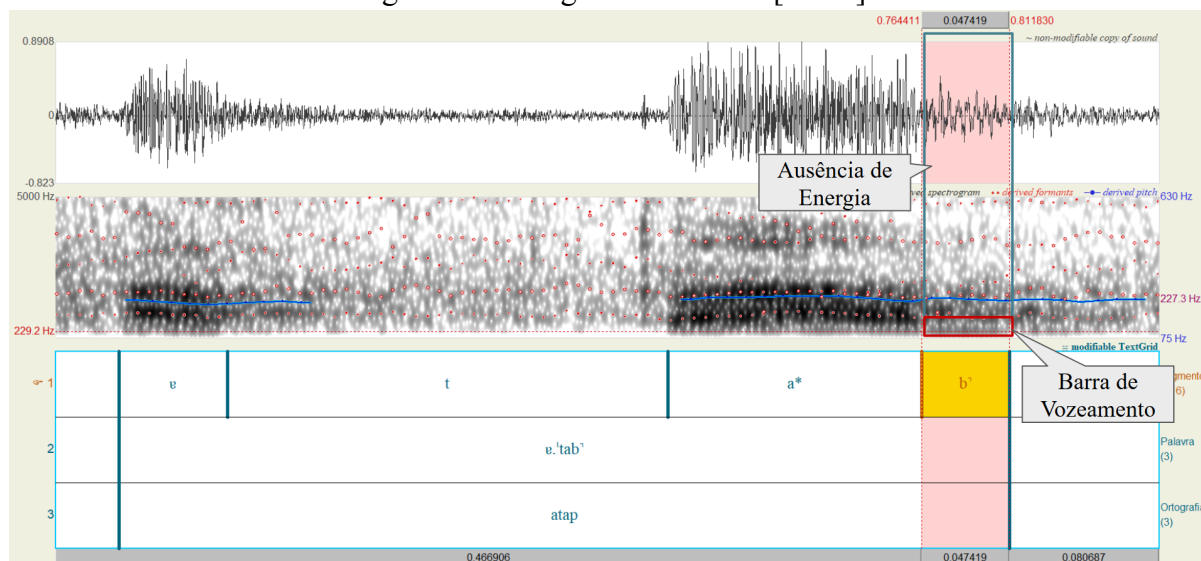
Uma outra forma que Juanita tem de pronunciar *atap* é sem a plosão no [p], portanto ela acaba produzindo o som que pode ser transcrito como [p̃]. Essa variação não muda o sentido da palavra. Vejamos sua imagem acústica com o formato de onda e o espectrograma:

Figura 67 – Imagem acústica de [v. 'tap̃]



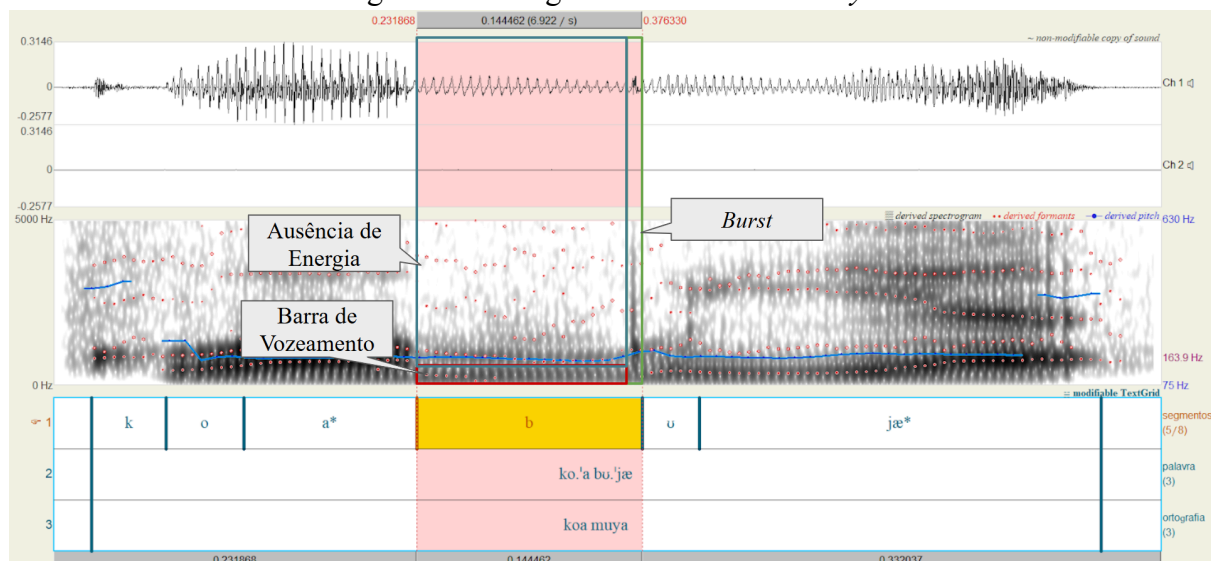
Observa-se que é um [p̃], pois possui as mesmas características acústicas visíveis do [p], exceto pelo fato de não ocorrer plosão. Ou seja, não há barra de vozeamento por ser uma oclusiva não-vozeada, podendo ser observada tanto de uma perspectiva perceptual quanto acústica. O que há de forma de onda na área indicada com ausência de energia é apenas um leve eco por conta do ambiente em que o áudio foi gravado. Ademais, a medição da duração do [p̃] é apenas aproximada, tendo em vista que o que normalmente delimita o final de sua produção é a plosão, por isso, utilizamos os formantes como guia para a demarcação. O fenômeno de oclusivas sem plosão é muito comum no Arutani, e ocorre em muitas línguas, como o Inglês na palavra *back* que é muitas vezes produzida e transcrita como [bæk̃]. Outra forma que encontramos de Juanita pronunciar *atap* foi com o uso de [b̃]. Veja:

Figura 68 – Imagem acústica de [v. 'tab̃]



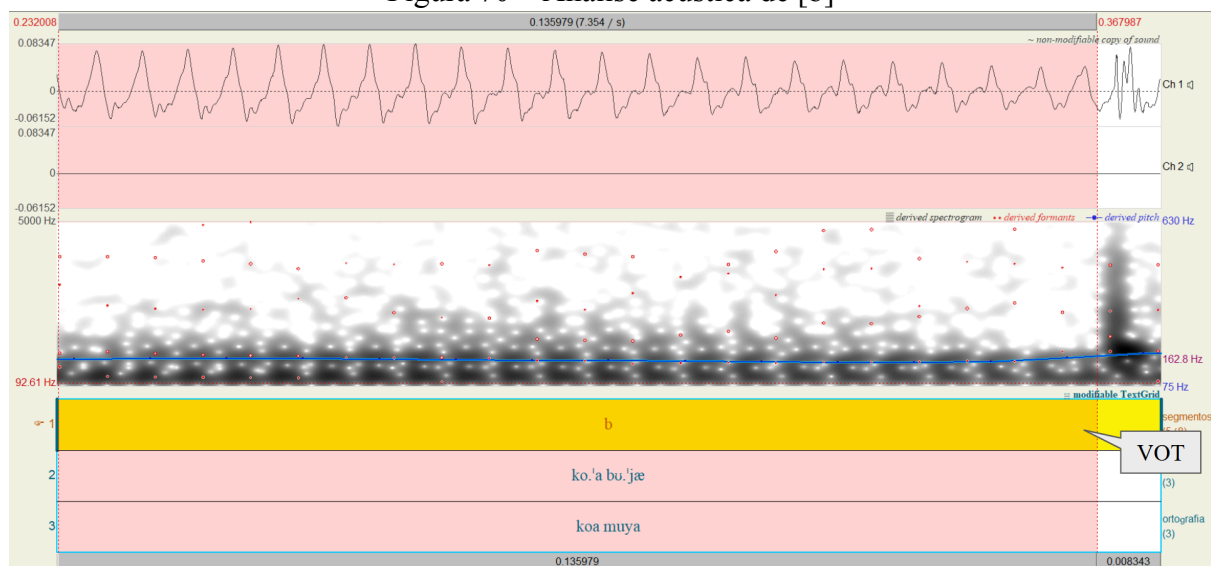
Neste caso, o único traço acústica que irá diferenciar o [b̃] do [p̃] será a barra de vozeamento que está muito mais preenchida em [b̃] do que em [p̃], portanto, a vibração ou ausência de vibração das cordas vocais, também chamado de F0. Para a delimitação da duração do [b̃], utilizamos como guia a constância da barra de vozeamento, já que não há o *burst* para auxiliar em uma delimitação mais adequada. Além disso, o que pode ser observado como frequência no formato de onda dentro da área indicada como ausência de energia, na verdade, são ruídos advindos do ambiente. Portanto, o que há em comum entre esses dois segmentos é o silêncio de oclusão, que não fica tão evidente por conta dos ruídos do ambiente, e a ausência de plosão, sendo [b̃] e [p̃] diferenciados apenas pela barra de vozeamento.

O [b] ocorre de forma pouco produtiva na fala de Juanita, ocorrendo apenas em alguns contextos, sendo eles de variação de pronúncia e em empréstimos de outras línguas que possuem o [b] como um fone com maior ocorrência. Temos o caso da palavra *rawkap* [rɐw. 'kɛ̃b̃] ‘jacupemba’, uma espécie de pássaro, em que o [b] é produzido sem plosão em contexto de coda silábica e final de palavra. Entretanto, o [b] pode ocorrer em outros contexto como na palavra *koa muya* [ko. 'a b̃u. 'jæ], que pode ser traduzida como ‘narina’ ou, literalmente, como ‘buraco do nariz’. Podemos observar algumas de suas características acústicas:

Figura 69 – Imagem acústica de *koa muya*

Na Figura 69, fica claro que se trata de um [b] de duração de 144 ms, pois podemos ver nitidamente a barra de vozeamento, característica das oclusivas vozeadas, além do *burst* mais evidente no espectrograma, mas que também é possível de ser observado no formato de onda. Há, também, a ausência de energia entre o [a] o momento da plosão e podemos também medir o seu VOT que será de aproximadamente -135 ms. Veja:

Figura 70 – Análise acústica de [b]



A palavra [ko.'a bu.'jæ] é uma palavra composta, podendo ser dividida em *koa* [ko.'a] ‘nariz’ e *muya* [mʊ.'ja] ~ [mɔj.'ja] ‘buraco’. Veja sua respectivas imagens acústicas:

Figura 71 – Imagem acústica de *koa*

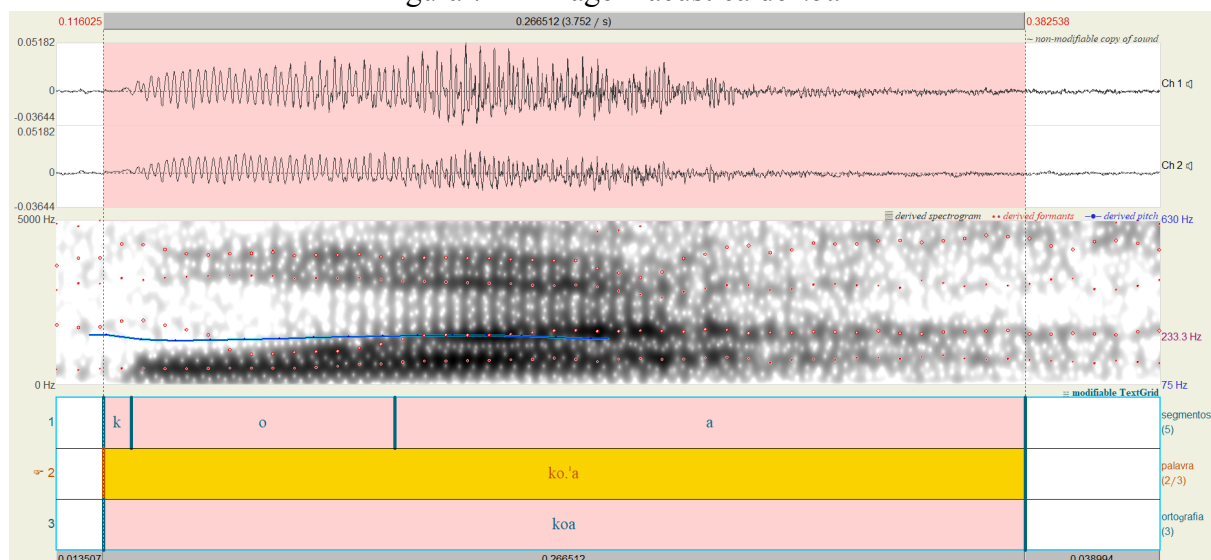
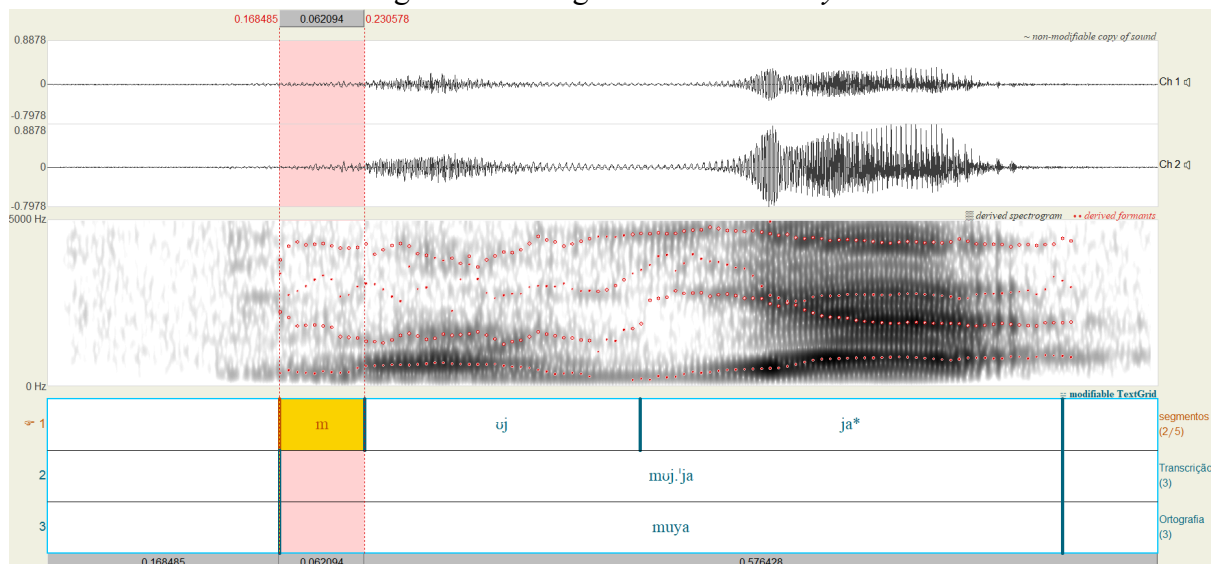


Figura 72 – Imagem acústica de *muya*



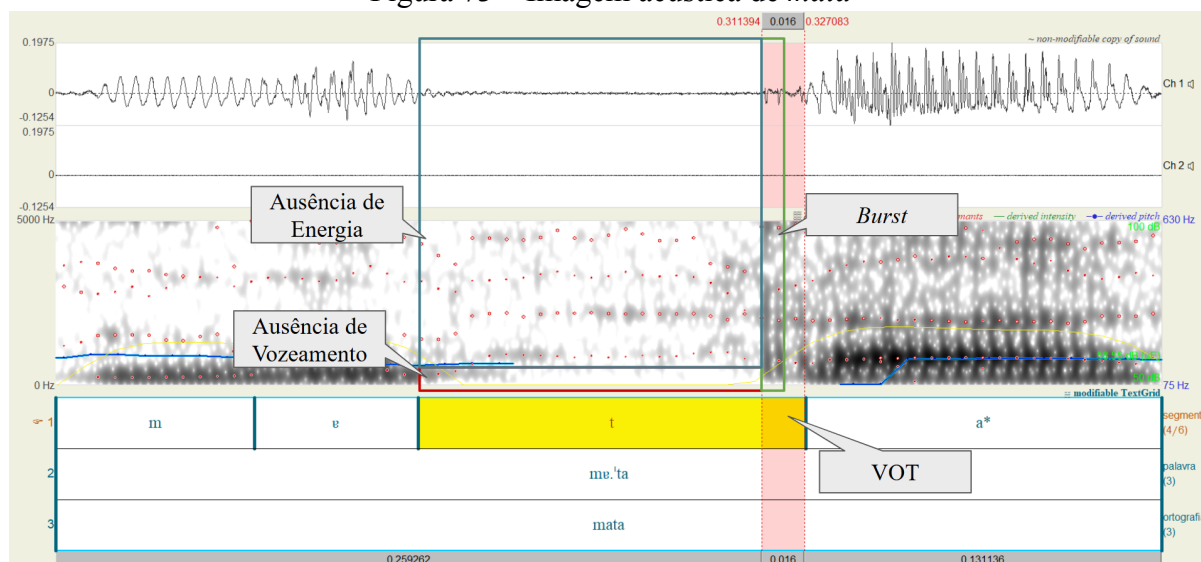
Outra ocorrência do [b] é na palavra *sambbrero* [ˌsẽm.bre.'ro] ‘chapéu’ que vem de empréstimo do Espanhol, cujo [b] é um fone produtivo, que possui a palavra *sombrero*, mas foi adaptada às regras de acentuação do Arutani. Portanto, temos ao todo apenas quatro ocorrência

de [b] no *corpus*: 1) variação entre [p̄] e [b̄] ([p̄] ~ [b̄]); 2) a variação entre [b] e [m] ([m] ~ [b]); e 3) empréstimo do Espanhol. O real *status* do [b] será melhor desenvolvido na seção 4.1.4, mas apenas com a questão da baixa ocorrência na língua e de ocorrer em alguns contextos de variação com fones produtivos, cujos de ponto ou modo de articulação são similares, já nos traz algumas pistas sobre como classificá-lo. Com base em nossa análise, fica claro, também, que não tratamos [p] e [p̄] ou [b] e [b̄] como fonemas distintos, apenas como dois fones aos quais foram inseridos o mesmo diacrítico [̄], veremos mais também sobre essa questão quando tratarmos de variação e alofonia na seção sobre a fonologia da língua.

Uma outra consoante oclusiva que ocorre na fala de Juanita, e que é muito produtiva, é o [t], o qual também sendo produzido de forma não explodida em posição de coda silábica e ocorre em ambientes iguais ao [p]. Para nossa análise, não diferenciaremos o [t] do [t̄] como fonemas diferentes, mas sim que o [t] em determinados contextos tende a receber o diacrítico [̄] por questões ambientais. Dessa forma, analisaremos primeiramente o [t] ocorrendo entre vogais, por ser o melhor ambiente de análise de seus traços acústicos; em seguida, observaremos o [t] em início de palavra e, por fim, a ocorrência de [t] ou [t̄] em final de palavra e coda silábica.

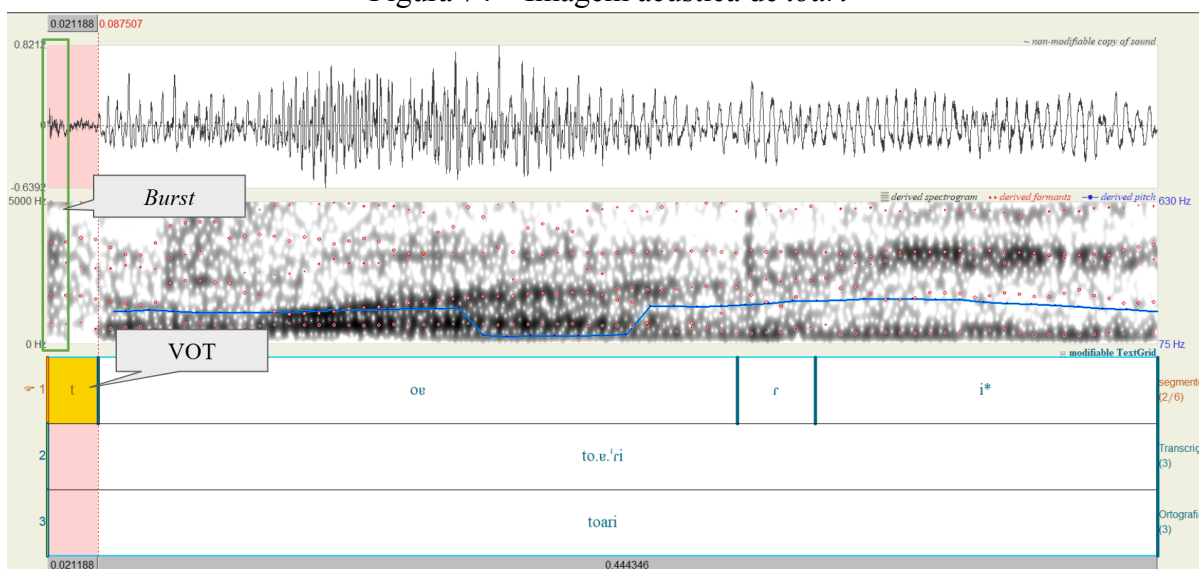
Vejamos o [t] ocorrendo entre vogais na palavra *mata* [mɐ.'ta], que significa ‘minha mãe’:

Figura 73 – Imagem acústica de *mata*



Temos na imagem um [t] de 141 ms de duração, no qual podemos observar nitidamente o silêncio de oclusão entre o primeiro [e] e o momento da plosão, que está mais claro de ser visualizado no espectrograma. Também é possível perceber que se trata de uma oclusiva não-vozeada por conta da ausência da barra de vozeamento nessa camada inferior do espectrograma na região segmentada. A partir dessa imagem podemos observar a marcação em destaque do VOT positivo de aproximadamente +16 ms. Podemos observar o *burst* característico das consoantes oclusivas que é mais evidente no espectrograma do que no formato de onda. Dessa forma, podemos observar o [t] em seu ambiente mais apropriado, visto que se torna possível uma análise de suas principais características acústicas. Entretanto, esse não é o único ambiente em que o [t] pode ocorrer em Arutani, mais especificamente no idioleto de Junita, pois o [t] também ocorre em início de palavra, como em *toari* [to.e.'ri], palavra para ‘cachorro’ na língua.

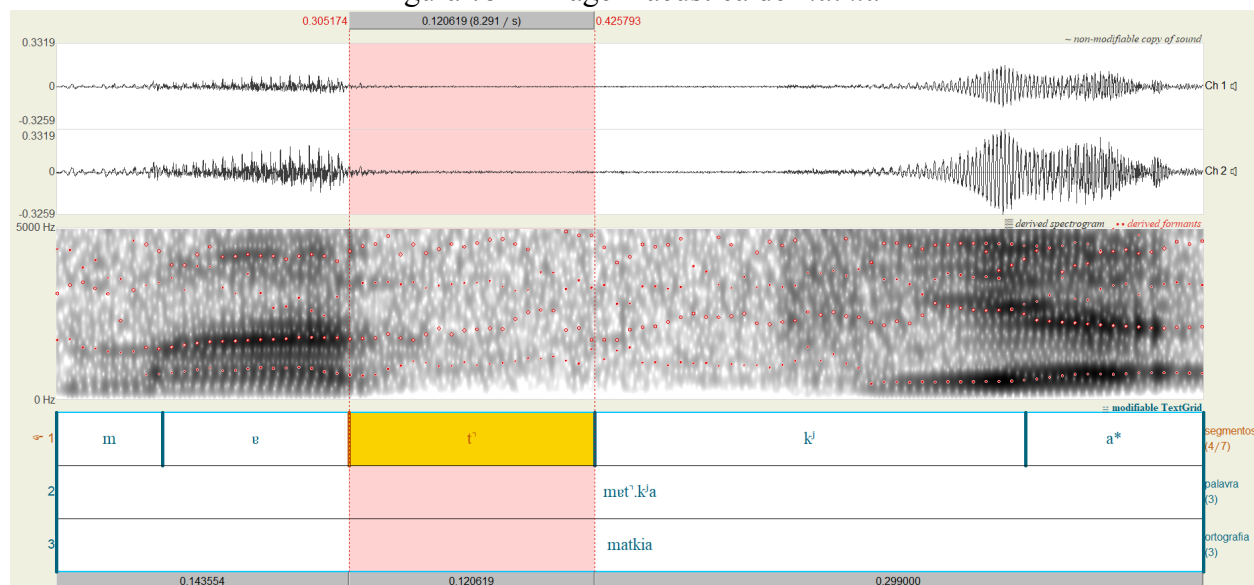
Figura 74 – Imagem acústica de *toari*



Pode-se ver, na Figura 74, a ausência de barra de vozeamento, dando-nos a evidência de uma consoante não-vozeada. Em destaque, temos um VOT de aproximadamente +21 ms e ao lado esquerdo, ainda dentro da segmentação do [t], temos o ruído transiente. O [t] também ocorre em coda silábica, como na palavra *matkia* [mæt̚.kia], que significa ‘homem’. Nessa palavra, temos um grande desafio de análise e segmentação, já que temos um [t̚], ou seja, uma oclusiva

sem *burst*, e em seguida um [k]. Dessa forma, fica quase que impossível definir o final de [t̥] e o começo de [k], já que temos um momento de ausência de energia entre eles. Outra observação que pode ser percebida na Figura 75 é que o Praat não identificou o *burst* do [k], mas ele pode ser ouvido.

Figura 75 – Imagem acústica de *matkia*

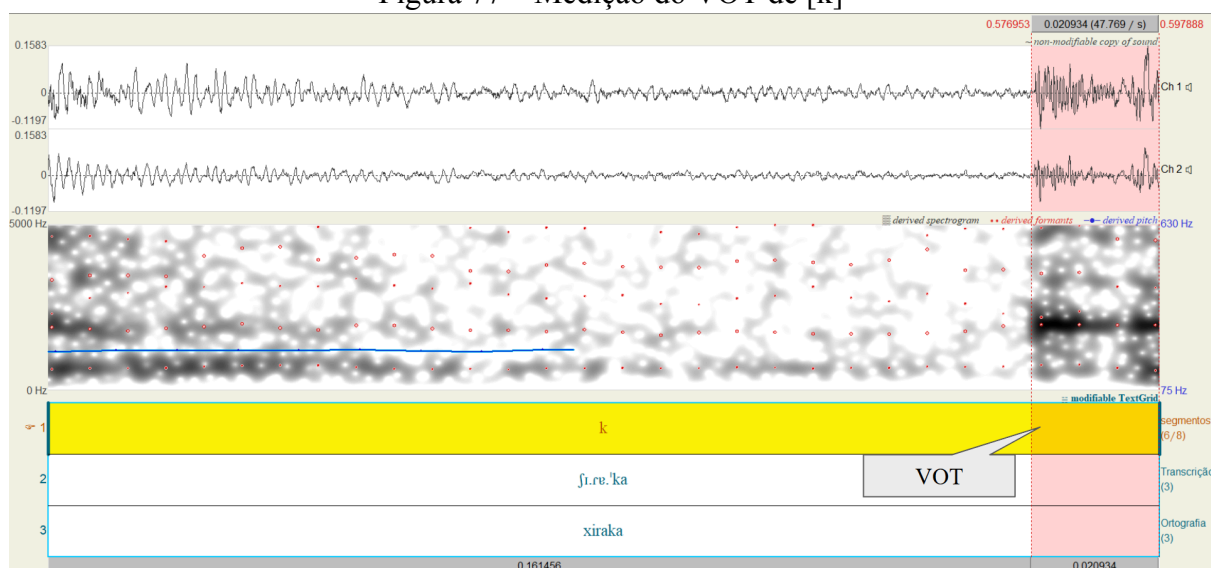


O último segmento oclusivo que iremos analisar é o [k], uma das consoantes mais produtivas da língua. Assim como foi feito com as outras oclusivas, iremos observar as propriedades acústicas do [k] em alguns contextos. O primeiro será entre vogais, para que seja possível a análise do silêncio de oclusão, da barra de vozeamento, do ruído transiente e o VOT. Para isso, escolhemos a palavra *xiraka* [ʃi.rɐ.'ka], que podemos traduzir como ‘cesto’ ou ‘aturá’.

Figura 76 – Imagem acústica de *xiraka*

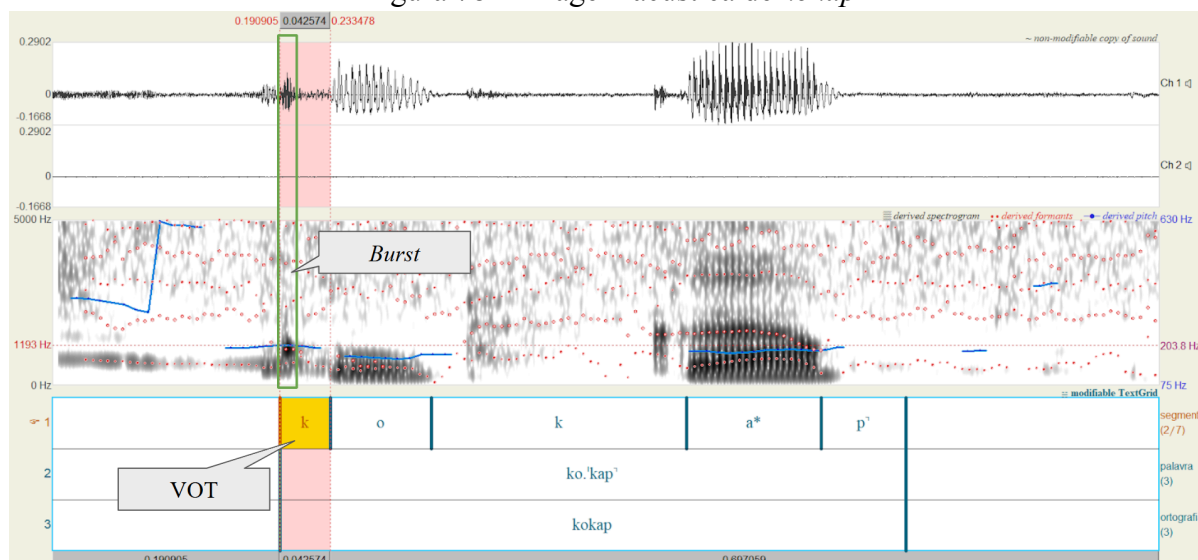
Nesta imagem pode-se observar o [k] em um ambiente ideal de análise, no qual ele apresenta aproximadamente 178 ms de duração. Dessa forma, o trecho marcado pelo retângulo vermelho, refere-se ao momento de ausência de energia e, em verde, o *burst*. Pelo fato de ser um segmento não-vozeado, não há barra de vozeamento, apenas um trecho em branco onde espera-se que haja a frequência fundamental, ou F0. Ao analisarmos o VOT encontramos um valor de + 20 ms.

Figura 77 – Medição do VOT de [k]



O próximo contexto em que analisaremos o [k] será no início de palavra, como no caso de *kokap* [ko.'kap̃], que significa ‘semente’. Veja:

Figura 78 – Imagem acústica de *kokap*



Neste caso, o [k] poderá ser analisado apenas levando em consideração sua plosão e seu VOT, já que são os traços acústicos possíveis de serem observados neste contexto. Na imagem, tem-se um VOT de +42 ms e o momento do *burst*. A plosão fica mais clara quando a observamos conjuntamente o espectrograma e o formato de onda. Por conta da duração, ainda é possível ver bem clara a ausência de vozeamento na parte inferior do espectrograma no trecho segmentado em que ocorre o [k]. Ressaltamos, que não encontramos [k] ocorrendo em coda silábica, nem mesmo sua contraparte não explodida *[k̚].

Até agora, exploramos características acústicas que são comuns às consoantes oclusivas, dessa forma, foi possível identificarmos o modo de articulação e de fonação das consoantes oclusivas que compõem nosso *corpus*. Agora, veremos traços acústicos que distinguem o ponto de articulação. Silva *et al.* (2019, p. 148) explica que “O ponto de articulação das consoantes oclusivas pode ser caracterizado acusticamente pelo VOT e pelas configurações espectrais do ruído transiente que é relacionado com a soltura da explosão”. Para a distinção das oclusivas na fala de Juanita iremos observar o ponto de articulação que as difere, sendo eles, bilabial, velar e alveolar. Silva *et al.* (2019, p. 148) explica que:

Para as oclusivas velares, o deslocamento da língua cria um espaço maior à frente do ponto de obstrução. À medida que o deslocamento da língua se anterioriza, o espaço à frente da língua diminui. Como consequência, em sons com espaços anteriores maiores - como as oclusivas velares -, a pressão do ar será maior do que nos sons com cavidades anteriores menores - como as oclusivas alveolares e bilabiais. Em razão da maior pressão de ar na cavidade que se forma na frente do ponto de obstrução para as consoantes velares, o VOT será mais longo para [k], por exemplo, do que para [p] e [t]. Em outras palavras, quanto maior o tamanho da cavidade oral à frente do ponto de obstrução, maior a pressão e maiores serão os valores de VOT.

Tendo isso em mente, já sabemos o que esperar dos valores de VOT das oclusivas em Arutani. Tomando como base os exemplos de palavras apresentadas anteriormente, nas quais as consoantes ocorrem entre as mesmas vogais, temos os seguintes valores:

Tabela 23 – Valores de VOT das consoantes oclusivas

Consoantes Oclusivas	VOT
[p]	15 ms
[t]	16 ms
[k]	20 ms
[b]	-135 ms

Com isso, cumpre-se o esperado, que seria o valor de [p] sendo o menor, seguido do valor de [t] e por último o [k], ou seja, $[k] > [t] > [p]$, assim como representado em Silva *et al.* (2019, p. 149). O valor de [b] é negativo por conta de ser uma consoante vozeada e, portanto, a forma com que seu VOT é medido.

A mesma relação é observada nos valores de F2 de transição, com expectativa de que $[k] > [t] > [p]$. Também conhecido como *locus* acústico consonantal, esse traço acústico também serve para distinguir o ponto de articulação entre as consoantes oclusivas. Dessa forma, o F2 de transição, que corresponde à região de maior energia na passagem do segundo formante entre uma oclusiva e a vogal seguinte, é um indicador do ponto de articulação das consoantes oclusivas, sendo medido nesse momento de transição e influenciado pelo tamanho da cavidade posterior à obstrução, de forma que cavidades menores resultam em valores mais altos de F2

(SILVA *et al.*, 2019, p. 150). Para uma descrição ainda mais precisa, vamos utilizar os dois primeiros formantes levando em consideração altura da língua (F1) e posição da língua (F2) das consoantes [p], [t], [k] e [b]. Esses valores podem ser observados na seguinte tabela:

Tabela 24 – Valores de F1 e F2 de transição das consoantes oclusivas

Consoantes Oclusivas	F1 de transição	F2 de transição
[p]	600 Hz	1.426 Hz
[t]	640 Hz	1.941 Hz
[k]	588 Hz	1.944 Hz
[b]	333 Hz	1050 Hz

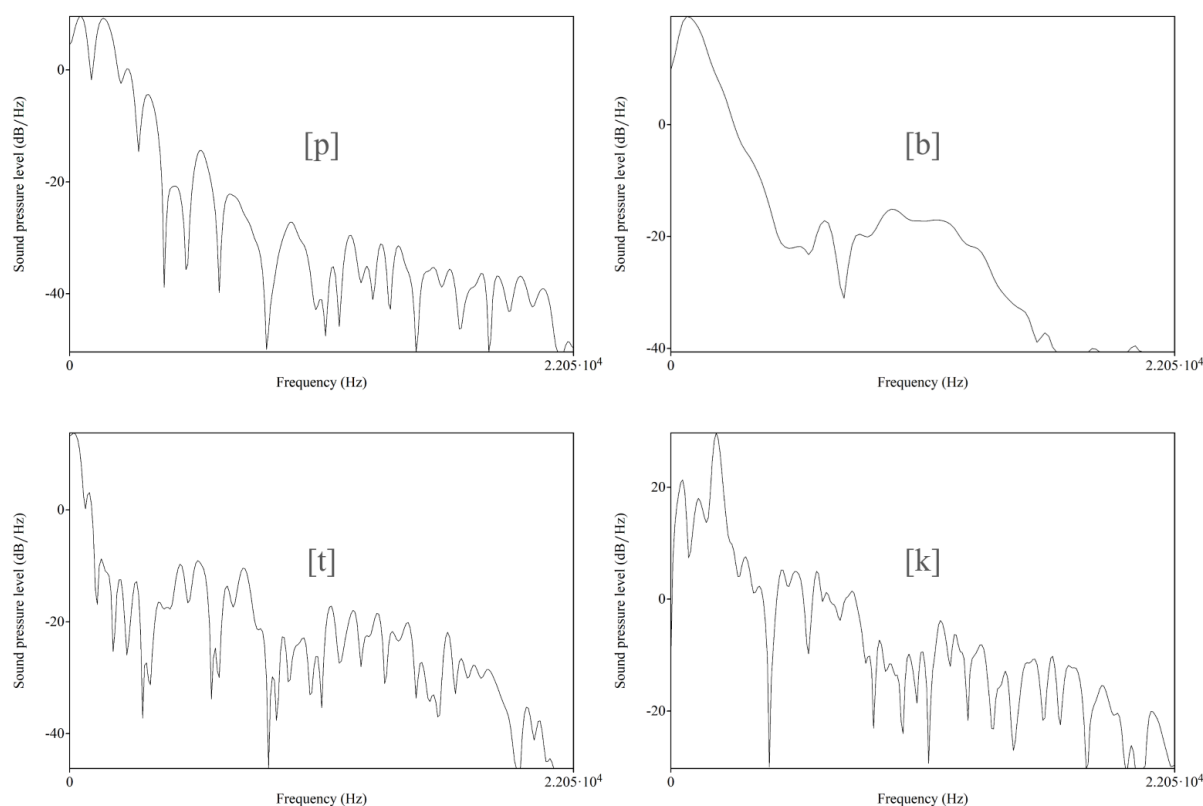
Os valores de F2 de transição mostram justamente o esperado e postulado por Silva *et al.* (2019), ou seja, [k], o qual possui uma cavidade menor, apresentou o maior valor de F2 de transição, 1.944 Hz, dentre as oclusivas não-vozeadas analisadas, as quais, todas, estavam precedendo a mesma vogal [a] em posição tônica. De forma semelhante ocorre com [t], com uma cavidade maior que [k], e menor que [p], a oclusiva alveolar, possui um F2 de transição de 1.941 Hz, sendo menor que [k], porém maior do que de [p]. Por fim, temos [p], o qual possui a maior cavidade dentre as oclusivas não-vozeadas e, conseqüentemente, o menor valor de F2 de transição, 1.426 Hz. Com isso, permanece a lógica $[k] > [t] > [p]$ com relação aos valores do F2 de transição.

Os valores dos formantes de transição de [b] variam mais por conta da influência da vogal seguinte. Como [b] foi analisado seguido de [u] já é esperado que ele irá possuir esses valores mais baixos quando comparados com os valores das outras oclusivas que foram analisadas diante de [a], justamente por influência da coarticulação da vogal seguinte. Destacamos que não existe [b] ocorrendo antes de [a], pelo fato dele ser um segmento muito raro na língua, ocorrendo apenas nos casos já citados anteriormente, portanto, não foi possível analisar a transição formântica ocorrendo no mesmo ambiente dos demais segmentos oclusivos.

A última característica acústica que iremos apresentar é a configuração espectral do *burst*, que é o terceiro recurso para a identificação e distinção dos pontos de articulação dos sons oclusivos. Silva *et al.* (2019, p. 152), ao citarem o trabalho de Blumstein e Stevens (1979), apresentam que as oclusivas bilabiais [p] e [b] exibem um padrão espectral descendente ou plano, denominado difuso descendente, com picos em frequências baixas, enquanto as oclusivas alveolares [t] e [d] apresentam um padrão difuso ascendente, com frequências distribuídas ao longo do espectro e picos em torno de 800 - 1.600 Hz e 1.800 Hz, e que as oclusivas velares [k] e [g] possuem um padrão compacto, com picos de energia em frequências médias entre 1.200 e 3.500 Hz.

Antes de gerarmos os espectros, foi preciso, após selecionarmos os *burst*, trocar a largura do comprimento de janela de leitura do Praat de banda estreita, que é o *default* do Praat, para banda larga, que é o tipo de banda mais adequado para a análise FFT, pois com ela é possível observar mais informações e dados acústicos do que quando fazemos uma análise FFT a partir de um espectrograma com banda larga. Observe as diferenças entres os espectros das oclusivas presentes na fala de Juanita gerados por meio da análise FFT:

Figura 79 – Análise espectral do *burst* das consoantes oclusivas



Ao analisarmos esses gráficos obtemos algumas informações acústicas importantes que vão além da tendência de inclinação dos espectros. A primeira dessas características é o centro de gravidade, que representa a média dos valores; a segunda é a variância, que é o valor da distribuição de energia em torno do valor da média, equivale ao desvio-padrão; a terceira é a distorção, que representa o valor da inclinação espectral; a quarta, e última, é a curtose, que nos permite observar quão pontiagudo ou quão achatado é o espectro. Na tabela a seguir apresentamos cada um desses valores dos *bursts* das respectivas consoantes oclusivas:

Tabela 25 – Análise dos espectros dos *bursts*

Oclusivas	Duração	Centro de Gravidade	Variância	Distorção	Curtose
[p]	6 ms	1.130 Hz	802 Hz	2	26
[b]	5 ms	943 Hz	598 Hz	5	90
[t]	8 ms	396 Hz	1.093 Hz	8	86
[k]	9 ms	1.862 Hz	918 Hz	5	63

Observe que quanto à duração do *burst*, seguimos a seguinte lógica já esperada: [k] > [t] > [p], tendo [b] a menor duração. Com relação ao centro de gravidade, [k] possui o maior valor, que é de 1.862 Hz, com um desvio-padrão de 918 Hz. Em oposto temos o [t] com um valor de 396 Hz de centro de gravidade com uma variância de 1.093 Hz. Já [p] e [b] possuem valores mais próximos e menos extremos, como pôde ser observado, isso tende a ocorrer pelo fato de terem o mesmo ponto e o mesmo modo de articulação. A distorção de todos está mais voltada para a direita, por conta dos valores positivos, dessa forma, observamos que o pico espectral das oclusivas tende a ser à esquerda. Por fim, quando tratamos da curtose, observamos que a oclusiva com o menor valor é [p], portanto, trata-se do espectro mais achatado, enquanto [b] e [t] possuem os valores mais altos, sendo os espectros mais pontiagudos.

As consoantes oclusivas no idioleto de Juanita mostraram-se muito recorrentes, e sua análise e sua descrição acústica corroboram o fato de que o que foi escutado no momento da transcrição estava correto. O uso da fonética acústica torna mais precisa a descrição do Arutani e colabora para os trabalhos futuros de descrição da língua, apesar da quase inevitável extinção do Arutani. A seguir, prosseguiremos com a descrição acústica de outro grupo de segmentos consonantais encontrado em nosso *corpus*, as consoantes nasais.

3.4.2 Consoantes Oclusivas Velares com Articulação Secundária

Na fala de Juanita, além das consoantes oclusivas, encontramos, também, oclusivas com articulação secundária: [k^w] (oclusiva velar não-vozeada labializada) e [k^j] (oclusiva velar

não-vozeada palatalizada). Observe na tabela a seguir quantas vezes esses segmentos foram encontrados em nosso *corpus*.

Tabela 26 – Quantitativos das consoantes oclusivas velares com articulação secundária

Consoantes Oclusivas	Nº de Ocorrências
[k ^w]	19
[k ^j]	4

Ladefoged e Maddieson (1995, p. 354) explicam que, de maneira geral, a articulação secundária se trata de uma articulação primária – oclusiva, no caso do Arutani – acompanhada de uma articulação em que há um menor grau de estreitamento articulatorio, que seria a articulação secundária – labializada ou palatalizada, no caso da fala de Juanita. Os autores ainda destacam que, de uma perspectiva fonética, as articulações secundárias terão sempre uma natureza aproximante ou semelhante a vogais quando observamos o grau de estreitamento (p. 355).

Assim como em alguns exemplos apresentados em Ladefoged e Maddieson (1995, p. 354), as oclusivas com articulação secundária no Arutani possuem um caráter muito mais fonológico e relevante para a fonotática da língua do que propriamente para a fonética em si, dada a dificuldade para a medição, precisão e comparação desses traços secundários. Ladefoged e Maddieson (1995, p. 355-356) explicam ainda que, dos tipos de articulação secundária, a labialização é a mais comum nas línguas do mundo, ocorrendo quando há a adição de um gesto articulatorio de arredondamento dos lábios.

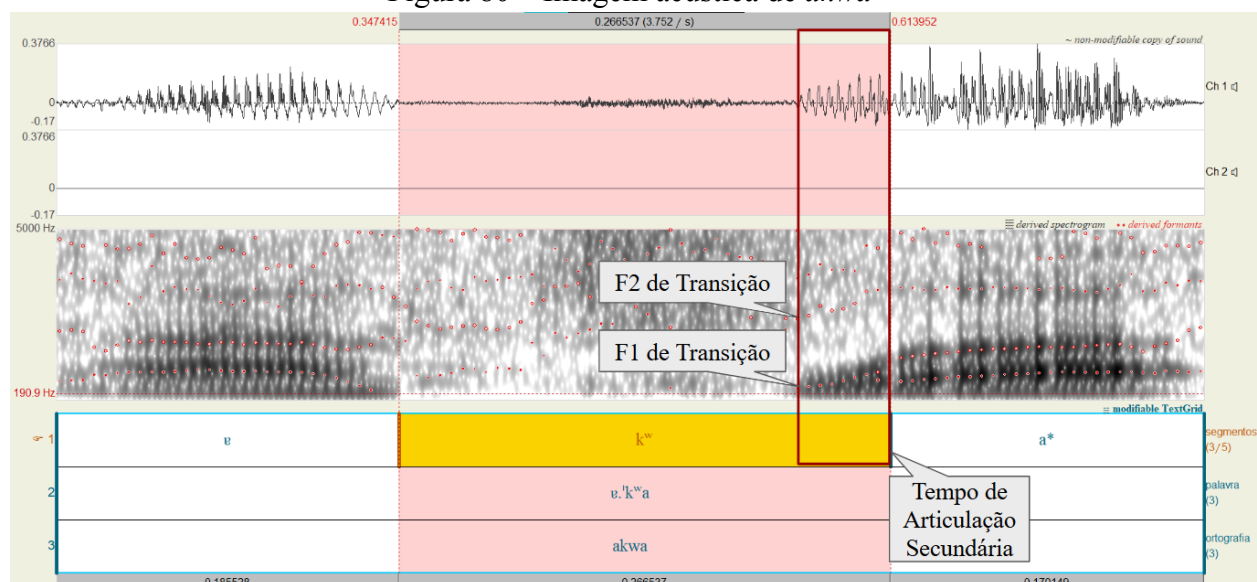
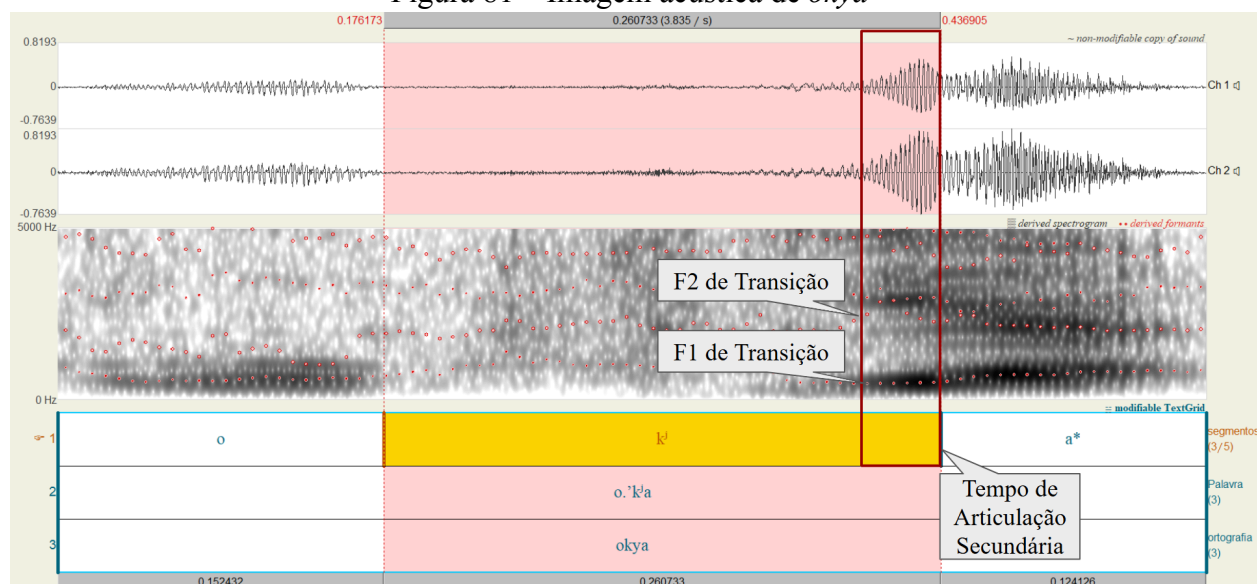
Em muitas das línguas que conhecemos, nota-se que o impacto acústico da articulação labial tende a ser mais acentuado no momento em que se libera o estreitamento principal de uma consoante labializada, do que no início desse movimento (LADEFOGED; MADDIESON, 1995, p. 357), por isso, iremos medir o F1 e F2 de transição após o VOT da articulação primária, já que este é o momento final da produção dos sons oclusivos e de liberação do estreitamento. Além disso, poderemos observar o F1, representando a altura da língua, e o F2, representando quão

posterior a língua está, para que seja possível constatar se há alguma sobreposição dos gestos articulatórios de [k] e [w] no momento de soltura do ar.

Com relação à palatalização como um processo de articulação secundária, Ladefoged e Maddieson (1995, p. 363-364) explicam que, articulatoriamente, ela se realiza a partir da “sobreposição de uma elevação da parte anterior da língua em direção a uma posição semelhante à do [i] em um gesto primário” (tradução nossa). Sendo este processo, segundo os mesmo autores, “frequentemente mais aparente na liberação do que na formação de uma constrição primária” (tradução nossa). Por sua realização ser mais visível no momento de *burst*, iremos investigar, também, o F1 e o F2 de transição da oclusiva velar, com o intuito de observarmos se há uma coarticulação ocorrendo na produção da consoante oclusiva, que seria justamente a elevação e a anterioridade da língua produzidas juntamente com a oclusiva velar, movimento influenciado pela articulação secundária [j].

Uma outra característica, a qual não foi explorada pelos autores, é a questão da duração quando comparadas com vogais plenas, articulatoriamente correlacionadas, como [ʊ], no caso das labializadas, e [i], no caso das palatalizadas. Portanto, [w] e [j] tendem a ter uma duração menor do que [ʊ] e [i], respectivamente. Dessa forma, faremos algumas comparações, como [k^w] vs. [kʊ] e [k^j] vs. [ki], esperando sempre uma duração maior nas sequências sem a articulação secundária.

Esses dois tipos de comparações evidenciam que não estamos tratando de uma sequência CV, por exemplo, mas, sim, de uma articulação secundária, já de [w] e [ʊ], assim como [j] e [i], são fones articulatoriamente semelhantes, sendo alofones em muitas línguas. Para questões de altura e posterioridade da língua no momento de transição, iremos medir os valores de F1 e F2 nas palavras *akwa* [ɐ.'k^wa] ‘batata-doce’ (espécie) e *okya* [o.'k^ja] ‘pupunha’. Esperamos encontrar valores baixos de F1 tanto em [k^w] quanto em [k^j], por conta da altura da língua. Entretanto, espera-se que [k^w] tenha um F2 menor, pela articulação secundária ser mais posterior, e [k^j] um F2 maior, pela articulação secundária ser mais anterior. Observe a seguir as imagens acústicas de [ɐ.'k^wa] e [o.'k^ja] respectivamente e seus valores de F1 e F2 de transição na Tabela 27:

Figura 80 – Imagem acústica de *akwa*Figura 81 – Imagem acústica de *okya*Tabela 27 – Valores de F1 e F2 de transição em [k^w] e [k^j]

Fones	F1 de Transição (Hz)	F2 de Transição (Hz)
[k ^w]	343 Hz	2.386 Hz
[k ^j]	451 Hz	2.581 Hz

Com base nesses dados, imagens acústicas e valores formânticos, podemos observar a influência das articulações secundárias, labializadas e palatalizadas, na coarticulação dos segmentos $[k^w]$ e $[k^j]$. Portanto, encontramos valores mais baixos de F1 de transição e um valor de F2 de transição mais baixo em $[k^w]$ quando comparado com o de $[k^j]$.

Um outro recurso que propomos é a medição da duração quando comparamos as coarticulações $[w]$ e $[j]$ com $[u]$ e $[i]$, respectivamente. Para esta comparação, observe a duração das imagens acústicas das palavras *arako* $[v.rv.'k\text{u}]$ ‘cará-roxo’ e *iki* $[i.'ki]$ ‘crescer’, as quais comparamos com as imagens acústica de $[v.'k^wa]$ e $[o.'k^ja]$ apresentadas anteriormente.

Figura 82 – Imagem acústica de *arako*

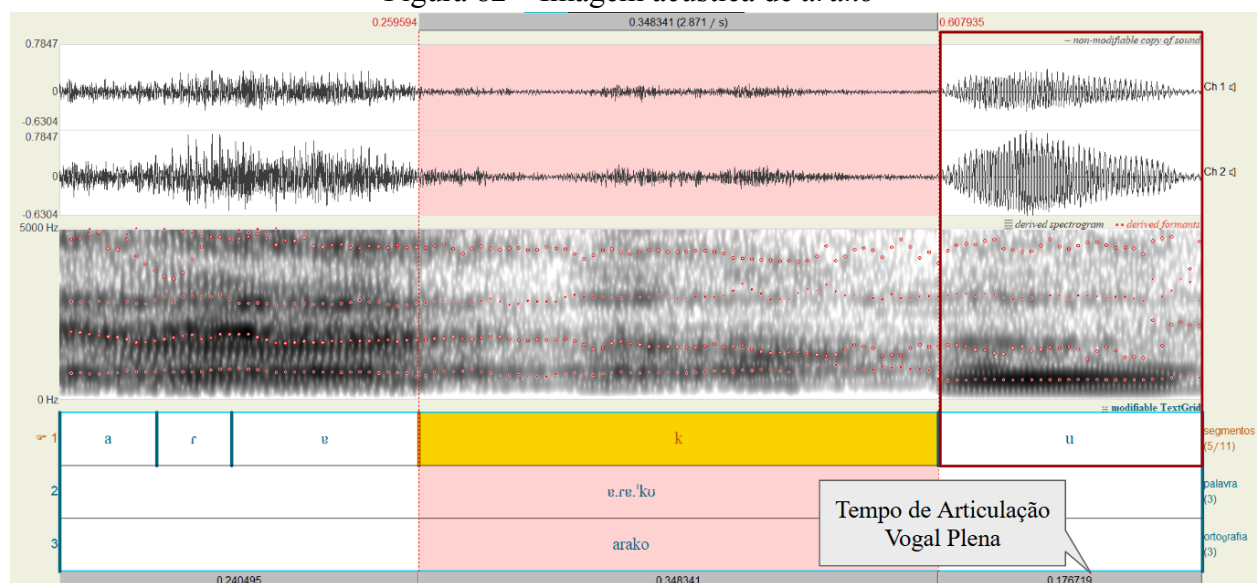
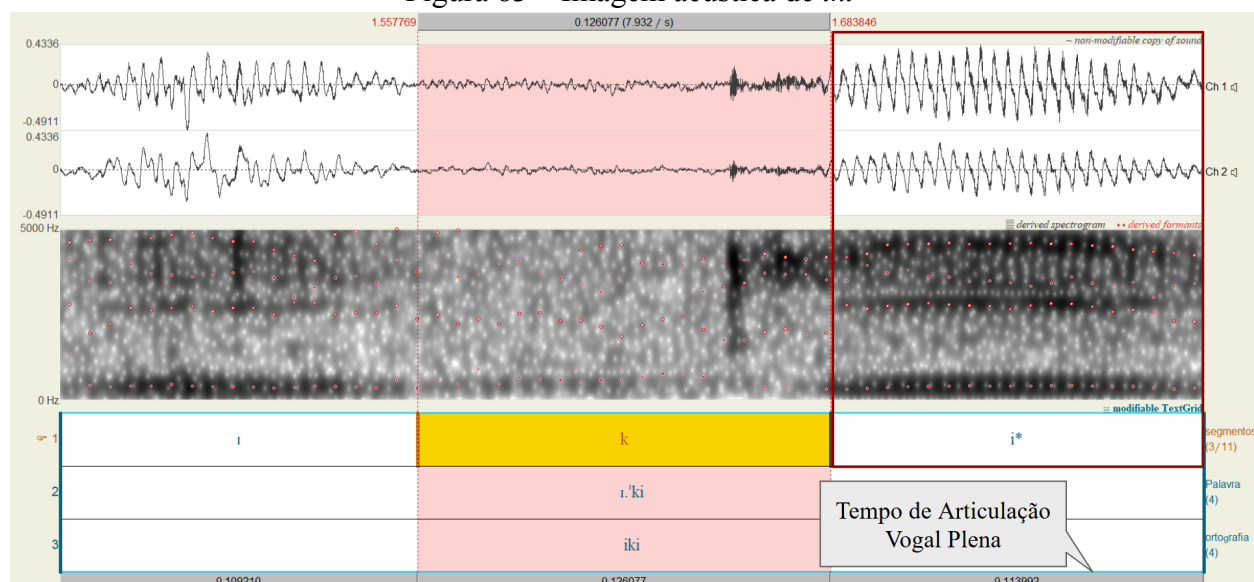


Figura 83 – Imagem acústica de *iki*

Apenas comparando as imagens acústicas já fica clara a diferença na duração, sendo [u] e [i] muito mais longos que os momentos de articulação secundária de [w] e [j] em [k^w] e [k^j]. Veja, agora, a comparação desses valores na Tabela 28, evidenciando ainda mais o contraste nas imagens acústicas.

Tabela 28 – Comparação da duração entre [w] vs. [u] e [j] vs. [i]

Articulação Secundária vs. Vogal Plena	Duração (ms)
[w]	42 ms
[u]	176 ms
[j]	33 ms
[i]	113 ms

Com base nos valores da Tabela 28, observamos que [u] é aproximadamente 319,05% mais longo que [w], sendo a duração de [u] cerca de 4,19 vezes maior que [w] em [k^w]. Quanto à [i] vs. [j], temos que [i] é aproximadamente 303,03% mais longo que [j], com uma duração cerca de 4,03 vezes maior. Quando comparamos esses dois grupos, observamos que os valores de porcentagem e de diferença são bastante próximos, mesmo tratando-se de articulações secundárias diferentes.

Com isso, veremos na seção 4.3 o *status* fonológico de [kʷ] e [kʲ]. Nela ficará ainda mais evidente o real papel das oclusivas velares com articulação secundária na língua, pois a existência desse tipo de segmento no contexto em que ocorrem reforça questões que englobam a existência ou não encontros consonantais na língua e evidenciam o *status* fonêmico dos *glides* na fala de Juanita, de forma que o sistema silábico do Arutani passa a ser explicado da maneira mais otimizada possível.

3.4.3 Consoantes Nasais

Na fala de Juanita, foram encontradas três consoantes nasais: [m] (bilabial), [n] (alveolar) e [ɲ] (palatal). Elas ocorrem tanto em posição de ataque quanto de coda silábica e são fones muito produtivos e recorrentes em nosso *corpus*. As consoantes nasais possuem algumas características acústicas específicas, o que nos permite fazer uma identificação precisa desse grupo de segmentos, indo além da simples percepção. Nesta seção, analisaremos essas três consoantes no contexto mais adequado para a sua análise, que é entre vogais. Não observaremos os demais ambientes em que esses segmentos ocorrem, pois, na análise acústica das consoantes nasais, existem traços acústicos mais relevantes a serem explorados e que demandam uma descrição mais acurada, como a comparação com as vogais adjacentes e a diferenciação de segmentos semelhantes, como é caso das aproximantes.

Quanto à produtividade e recorrência de uso das três consoantes nasais encontradas na fala de Juanita e em nosso *corpus*, é possível fazer uma escala de ocorrência, na qual a consoante nasal bilabial [m] é a mais produtivo, seguida da consoante alveolar [n] e, por último, e menos produtiva, a consoante nasal palatal [ɲ]. Em números, temos o seguinte:

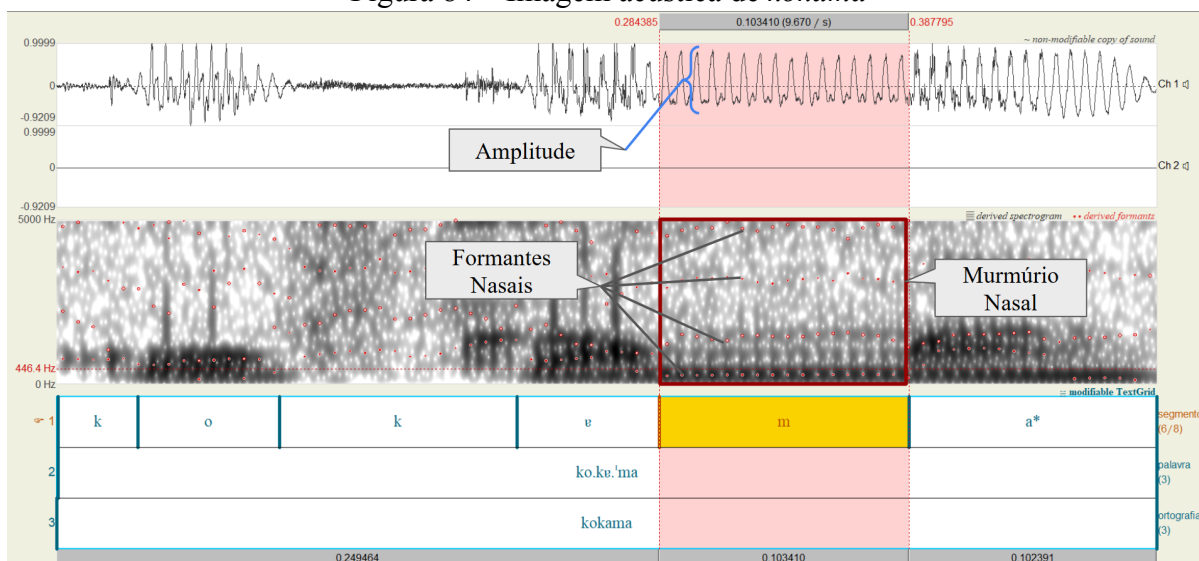
Tabela 29 – Quantitativos das consoantes nasais

Consoantes Nasais	Nº de Ocorrência
[m]	94
[n]	63
[ɲ]	16

Barbosa e Madureira (2015, p. 177) fazem uma relação direta entre uma perspectiva articulatória e acústica da produção desses fones ao explicarem que esse tipo de segmento é produzido por meio da “obstrução total entre articuladores na cavidade bucal e com abaixamento do véu palatino, o que proporciona uma passagem livre para a corrente de ar através da cavidade nasal”. Dessa forma, os autores complementam expondo que a anatomia da cavidade nasal favorece a diminuição de energia de algumas frequências com relação à produção de consoantes, o que geram os antiformantes – ou antirressonância – que são a principal característica para a identificação do espectro nasal. Outro traço possível de ser percebido é a menor amplitude da forma de onda com relação aos sons vocálicos próximos/adjacentes e a ocorrência de maior energia no primeiro formante com relação aos demais (BARBOSA; MADUREIRA, 2015, p. 177). Silva *et al.* (2019, p. 154), de maneira complementar, apresentam as seguintes características acústicas das consoantes nasais: 1. murmúrio nasal; 2. formantes nasais; 3. amplitude; e 4. antiformantes.

Portanto, tendo em mente todos esses traços que caracterizam os segmentos consonantais nasais, podemos observar essas características presentes no fone [m] na palavra *kokama* [ko.kə.'ma], que pode ser traduzida como ‘filho de alguém’. Vejamos a imagem acústica dessa palavra:

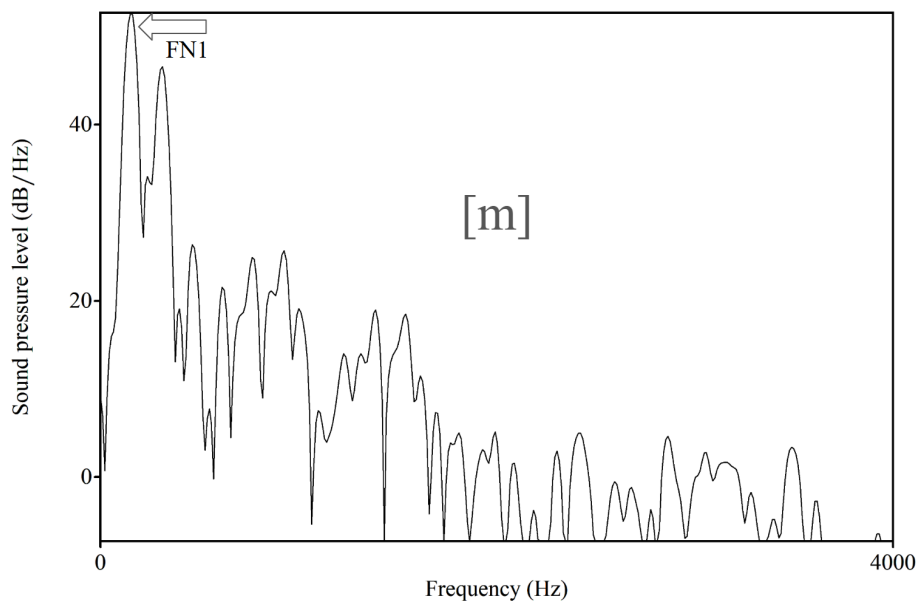
Figura 84 – Imagem acústica de *kokama*



Nesta palavra temos um [m] de aproximadamente 103 ms em contexto de sílaba tônica e entre um [ɐ] à esquerda e uma [a] à direita. Como em destaque, o murmúrio nasal refere-se a toda essa parte selecionada na segmentação. Podemos observar também a diminuição da amplitude no formato de onda em relação às vogais adjacentes e uma maior concentração de energia no FN1 quando comparado com os demais formantes nasais. Sobre isso, Silva *et al.* (2019, p. 158) explica que as consoantes nasais só possuem formantes nasais, o que as difere das vogais nasais, as quais possuem formantes orais e nasais.

Os formantes nasais, representados como FN1, FN2 e FN3, são traços “que podem ser identificados em espectros de frequência” (SILVA *et al.*, 2019, p. 158). Sendo assim, ao gerarmos um gráfico (Figura 86) do [m] em [ko.kɐ.'ma] por meio da análise de FFT obtemos um espectro, no qual podemos observar um pico que, segundo os autores, representa o primeiro formante nasal. Segundo Silva *et al.* (2019, p. 158-159), o murmúrio nasal distingue-se por apresentar um formante nasal primário (FN1) com intensidade elevada em comparação com os outros formantes nasais do espectro. Além disso, a redução da intensidade desses formantes ocorre devido à presença de antiformantes – áreas de frequência onde a energia é bastante reduzida, visíveis como depressões acentuadas no espectro – que surgem do acoplamento entre o trato oral e o trato nasal (SILVA *et al.*, 2019, p. 159). Dessa forma, podemos observar que existe uma relação direta de influência entre os antiformantes e o murmúrio nasal, na qual os antiformantes são os responsáveis pela ocorrência do murmúrio, que é acusticamente expresso e evidenciado pela baixa amplitude das frequências, atenuação da energia e amortecimento gerados pelos antiformantes. No espectro a seguir, temos em destaque o FN1:

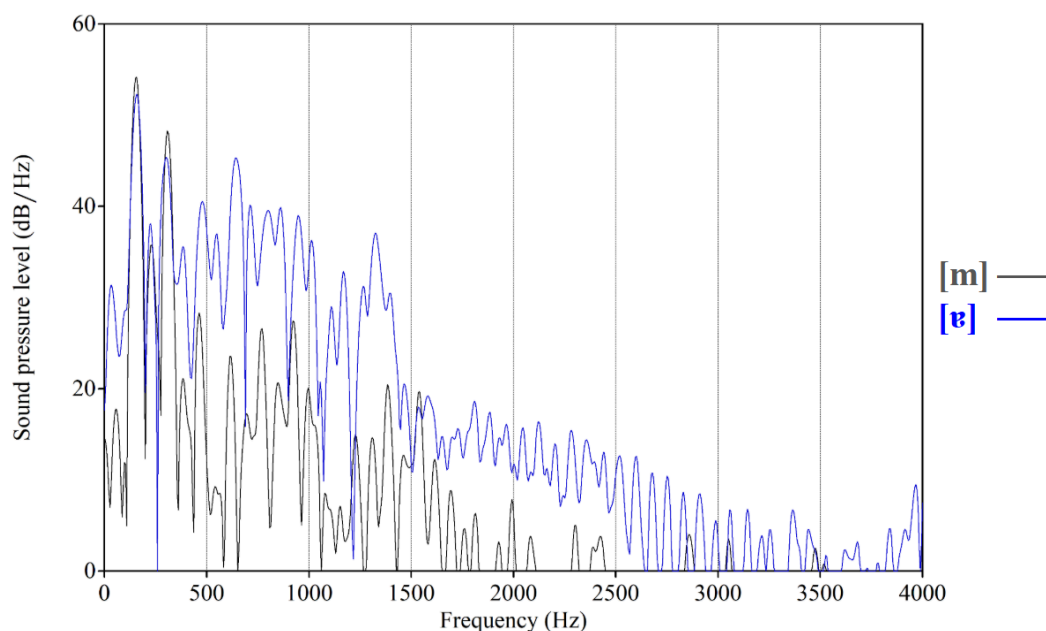
Figura 85 – Espectro de [m]



Observamos nesse espectro um “FN1 de grande intensidade quando comparado com os demais formantes nasais presentes no espectro”, uma das principais características das consoantes nasais segundo Silva *et al.* (2019, p. 158).

Além disso, o espectro auxilia na indicação de “que a energia é menor nas consoantes nasais quando comparadas com as vogais adjacentes”. Portanto, espera-se que os valores de amplitude das consoantes nasais sejam menores se comparados com os valores de amplitude das vogais adjacentes (SILVA *et al.*, 2019, p. 159). E é exatamente isso que encontramos quando comparamos os espectros de [m] e de [ɐ] em [ko.kɐ.'ma]:

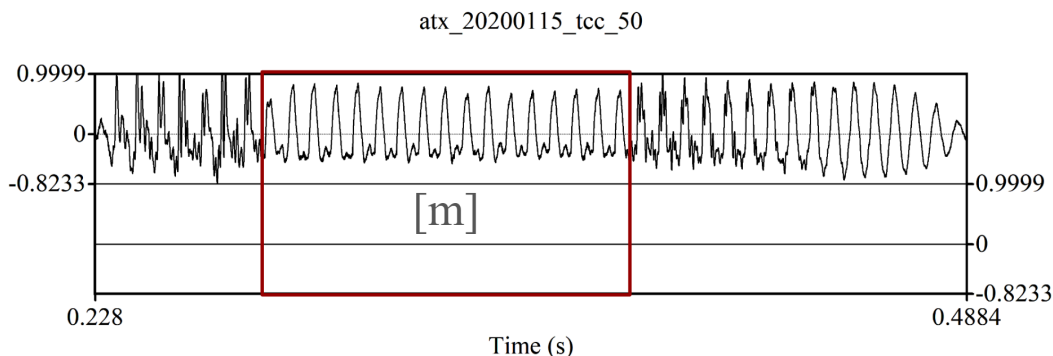
Figura 86 – [m] vs. [v]



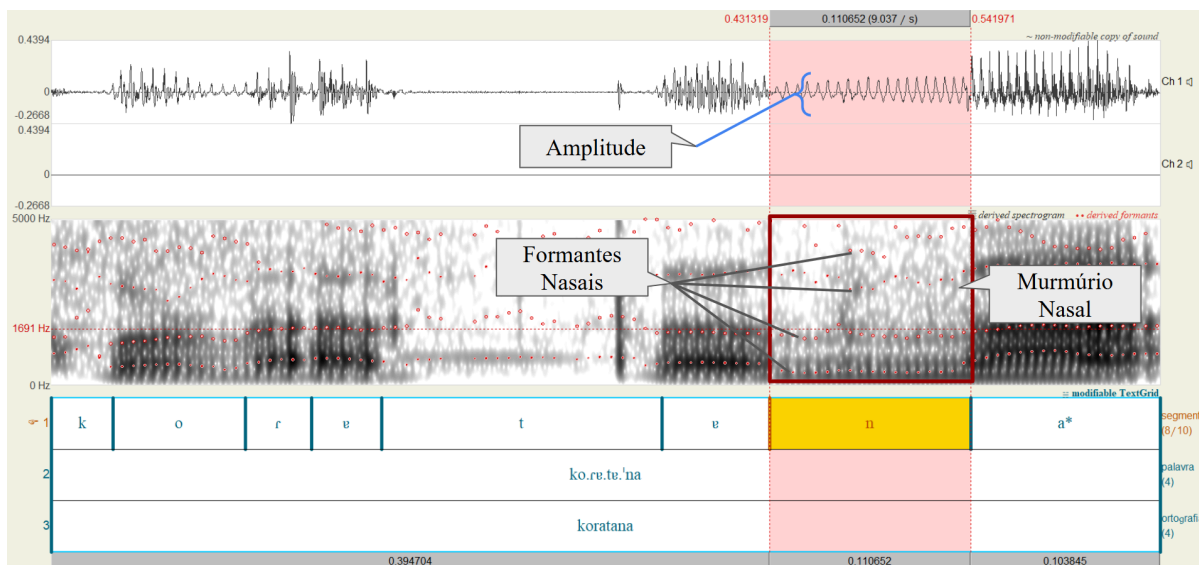
Em azul temos o espectro de [v] que possui uma frequência maior quando comparada com a frequência de [m] em preto. Se colocarmos isso em números e ao compararmos os valores obtidos pelo *Script 1* que utilizamos na extração dos dados acústicos, veremos que há uma diminuição na frequência com relação à passagem da vogal [v] para a consoante bilabial nasal vozeada [m]. Em [v] encontramos um F1 de 651 Hz, enquanto em [m] são encontrados um FN1 de 289 Hz.

Sobre o amortecimento da amplitude, traço também característico das consoantes nasais, podemos observá-lo não apenas no espectrograma ou no espectro obtido pela análise FFT, mas também na própria forma de onda. Na Figura 89, a seguir, fica clara a diferença entre a amplitude da onda do [m] (parte selecionada) e as vogais adjacentes, ainda na palavra [ko.kv.'ma].

Figura 87 – Formato de onda de [m]



Quanto à nasal alveolar [n], observe a imagem acústica de *koratana* [ko.rv.te.'na], que é a palavra para ‘banana’ em Arutani:

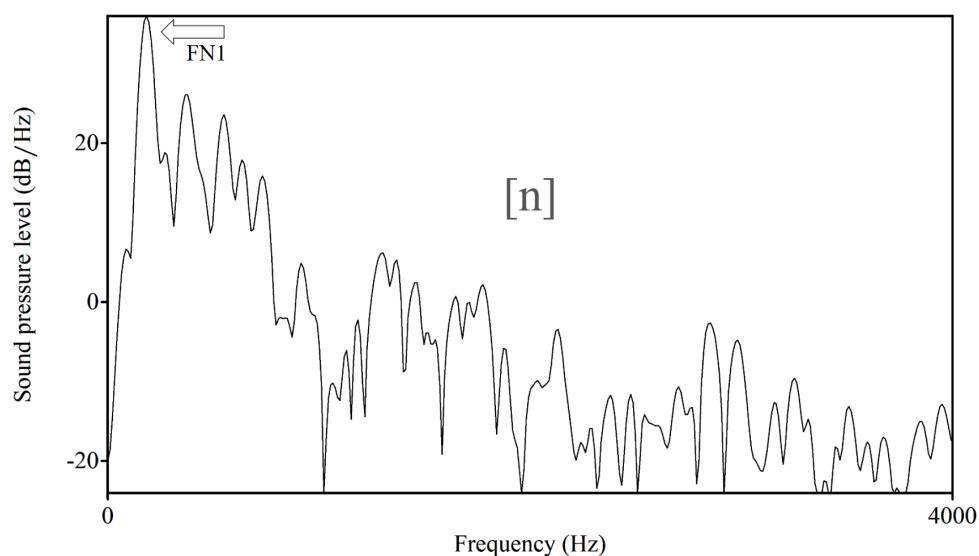
Figura 88 – Imagem acústica de *koratana*

Podemos observar nitidamente a diminuição da amplitude da consoante no formato de onda, evidenciado pela região mais clara do espectrograma na área segmentada. Em seguida, temos uma outra característica acústica do [n] que é murmúrio nasal, o qual é responsável pela definição do modo de articulação, segundo Silva *et al.* (2019, p. 156). O murmúrio do [n] pode ser observado na região dentro do retângulo vermelho, na Figura 90, caracterizado pela região mais clara do espectrograma, reflexo acústico da baixa amplitude. Outra característica acústica

do [n] como consoante nasal são os formantes nasais, indicado na Figura 90 pelas linhas e os pontos vermelhos.

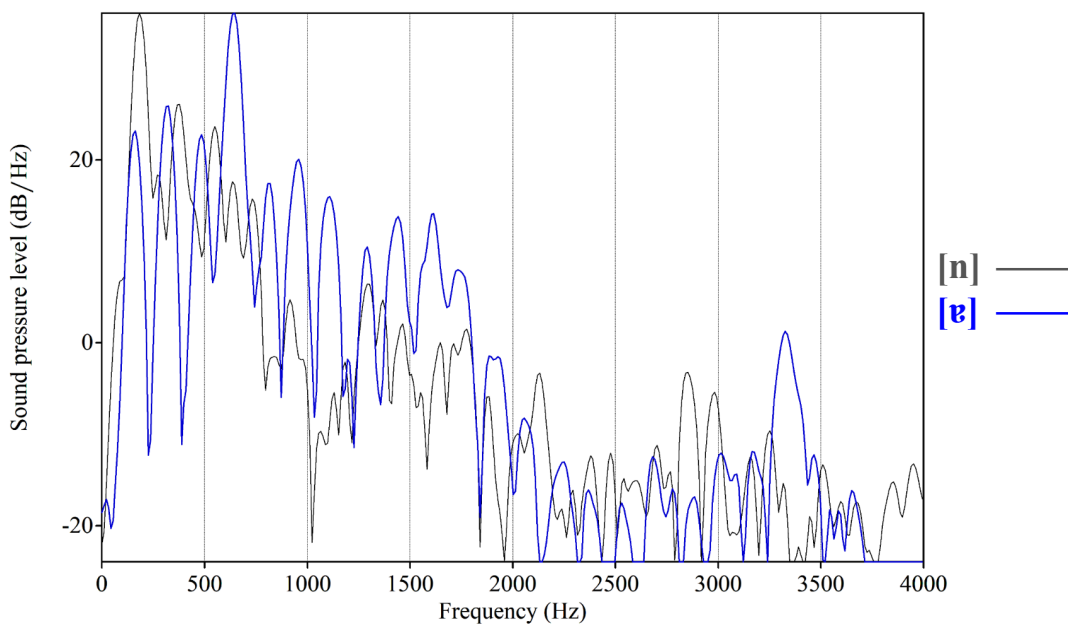
Na Figura 91 a seguir, podemos observar o FN1 do [n] com maior amplitude em relação aos outros formantes nasais:

Figura 89 – Espectro de [n]



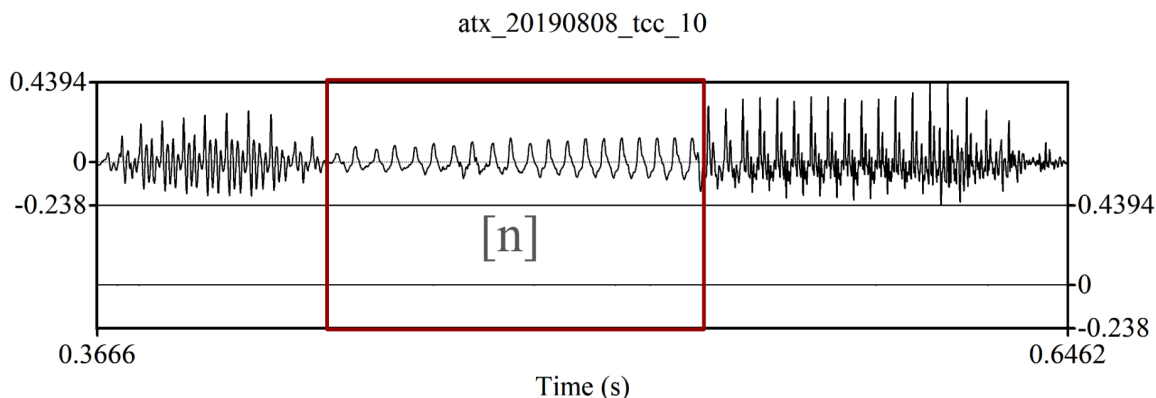
Segundo a literatura já citada, espera-se que os picos de amplitude de [n] sejam menores do que os de sua vogal adjacente. Com o fim de seguirmos os mesmo parâmetros e ambientes que utilizamos para a análise do [m] anteriormente, selecionamos o [n] (em preto) e o [v] (em azul) em [ko.rɐ.tɐ.'na] para a comparação a partir da análise FFT. Dessa forma, temos o seguinte espectro:

Figura 90 – [n] vs. [ɐ]



Quando comparamos a análise FFT com as informações acústicas extraídas pelo *Script 1*, encontramos dados numericamente próximos com relação aos formantes encontrados nas duas análises. Temos um F1 de 631 Hz para [ɐ], enquanto que para [n] temos um FN1 de 414 Hz. Essa diminuição da frequência é o que Silva *et al.* (2019, p. 158-159) chamam de amortecimento, que é causado pelos antiformantes. Observe, ao darmos um *zoom* no formato de onda, fica mais clara a visualização da diminuição dos picos de frequência e da amplitude entre a consoante nasal [n] e suas vogais adjacentes em [ko.rɐ.tɐ.'na]:

Figura 91 – Formato de onda de [n]



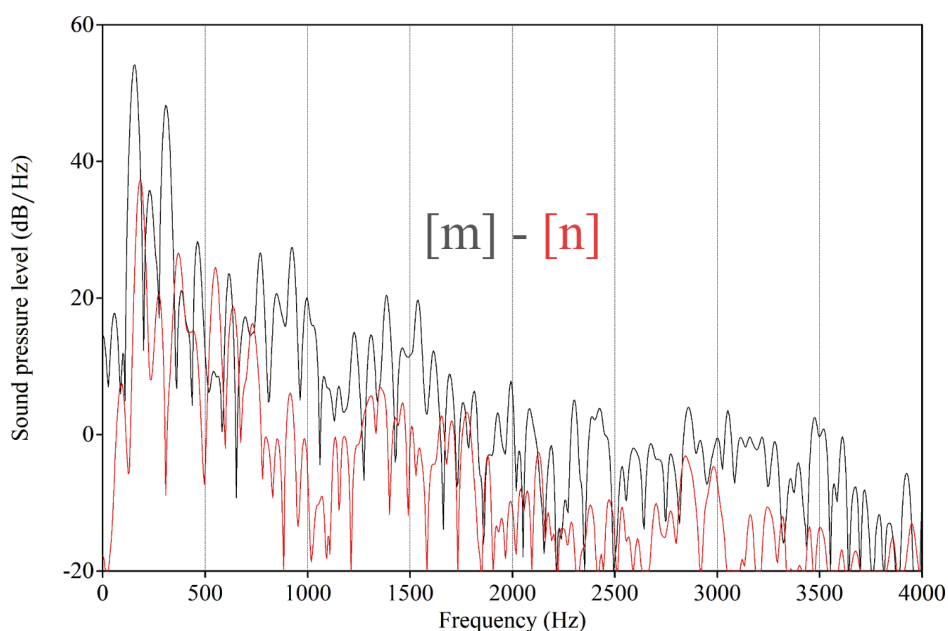
Como vimos, o FN1, quando de elevada intensidade, ao ser comparado com os demais formantes, assim como demonstrado anteriormente, é a principal característica do murmúrio nasal. Já os FN2 e FN3 são responsáveis pela diferenciação do ponto de articulação. Silva *et al.* (2019, p. 164) explica que: “Para a diferenciação do ponto de articulação das consoantes nasais [m,n], é necessário considerar FN2 e FN3, que apresentam valores mais baixos para [m] do que para [n]”. Observe exatamente isso ocorrendo na tabela a seguir ao compararmos os valores dos formantes nasais de [m] em [ko.kɐ.'ma] e o [n] em [ko.rɐ.tɐ.'na].

Tabela 30 – Comparação entres os valores dos formantes nasais de [m] e [n]

Segmentos	FN1 (Hz)	FN2 (Hz)	FN3 (Hz)
[m]	286	1.414	2.845
[n]	414	1.438	2.896

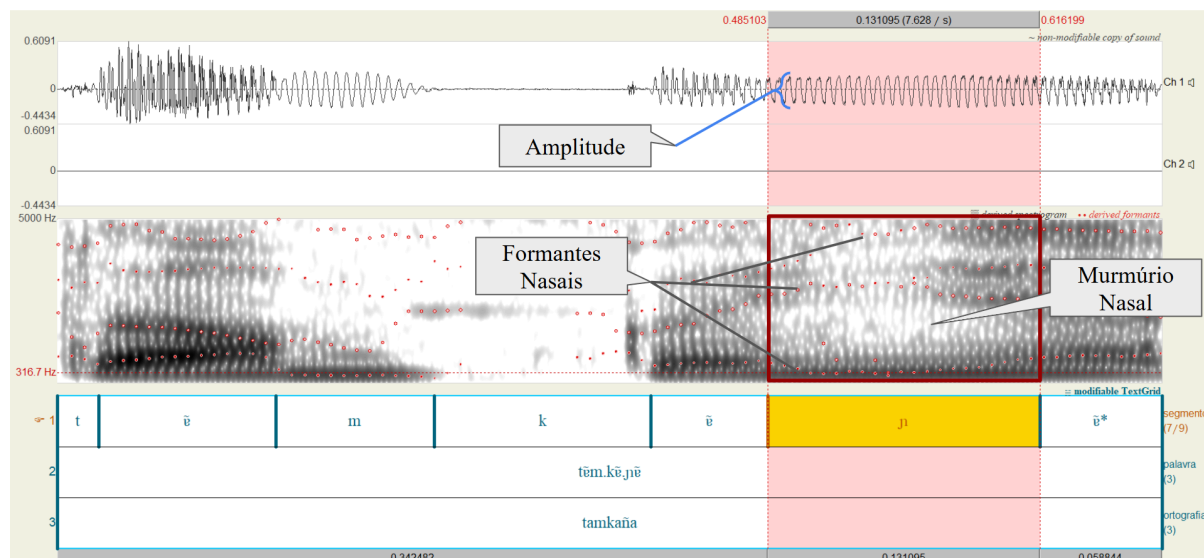
Essa é a representação quantitativa do que observamos na Figura 94 ao compararmos [m] (em preto) e [n] (em vermelho).

Figura 92 – [m] vs. [n]



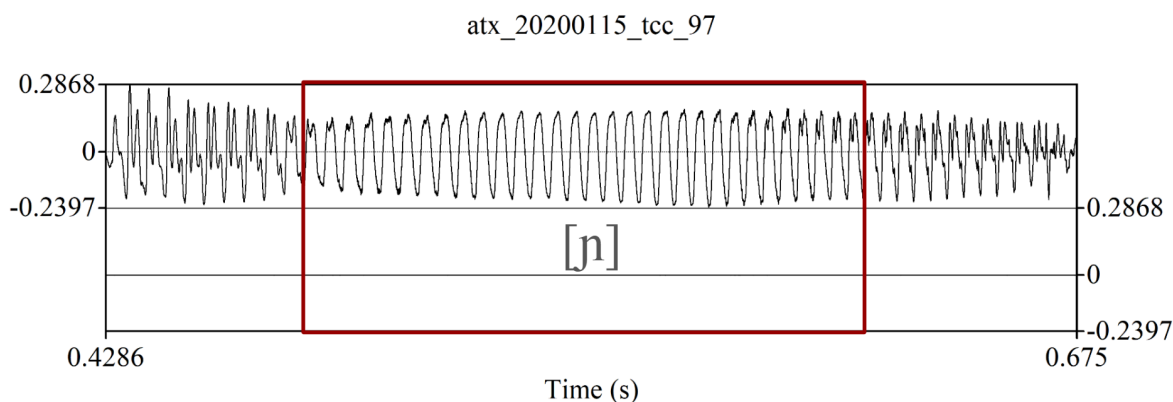
Dessa forma, foi possível analisarmos o [m] e o [n] existentes na fala de Juanita e provar que os sons que foram transcritos a partir de uma análise oitiva também podem ser analisados e comprovados também a partir de uma perspectiva acústica. Ressaltamos que os parâmetros e as características evidenciadas nas imagens acústicas anteriores foram os mesmos levados em consideração para a análise das outras ocorrências desses segmentos nasais. Além disso, conseguimos mostrar suas características acústicas esperadas e evidenciar traços que distinguem entre si essas consoantes nasais.

Entretanto, ainda há a necessidade de fazermos a análise e a descrição da consoante nasal palatal [ɲ]. O [ɲ] ocorre menos em Arutani quando comparado com as outras consoantes nasais, porém ainda continua sendo um fone produtivo no idioleto de Juanita. Essa consoante pode ser encontrada em palavras como *tamkaña* [tẽm.kẽ.ɲẽ], que significa ‘ombro’. Na imagem acústica de [tẽm.kẽ.ɲẽ], podemos ver indicados o murmúrio nasal, os formantes nasais e a amplitude:

Figura 93 – Imagem acústica de *tamkaña*

Os formantes nasais e o murmúrio estão dentro do esperado para a imagem acústica, porém a forma de onda não, segundo a leitura do Praat, não sofreu tanto amortecimento quanto o esperado, mas fica claro tratar-se de uma consoante nasal quando observamos essa região embranquecida entre o FN1 e o FN2, que é justamente a região mais característica do murmúrio nasal, evidenciando, assim, a existência dos antiformantes.

Figura 94 – Formato de onda de [ɲ]



Esse baixo amortecimento pode ocorrer por dois motivos neste contexto. O primeiro seria por estar em um ambiente de vogais nasais e, com isso, o amortecimento causado pelos

antiformantes tenham amortecido toda a amplitude das ondas não somente na consoante, mas, também, nas próprias vogais adjacentes. Essa hipótese pode ser respaldada até mesmo pelo fato de as vogais adjacentes terem já uma baixa frequência, entretanto, essa diferença fica clara na comparação espectral. O segundo motivo seria pelo contexto de final de palavra, que, apesar ser em sílaba tônica, pode ocorrer essa baixa amplitude das vogais pelo fato de ser um ambiente em que há uma tendência de produção de sons de maneira mais relaxada e, até mesmo, com alguns apagamentos ou desvozeamentos.

Observe a seguir o espectro de [ɲ] e a comparação entre o espectro da vogal adjacente [ẽ] e consoante nasal:

Figura 95 – Espectro de [ɲ]

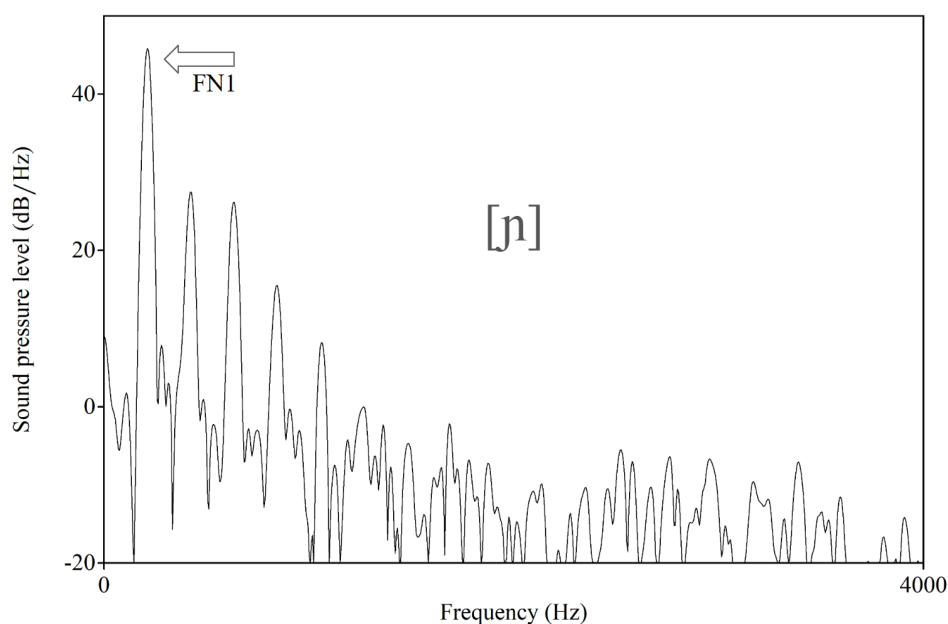
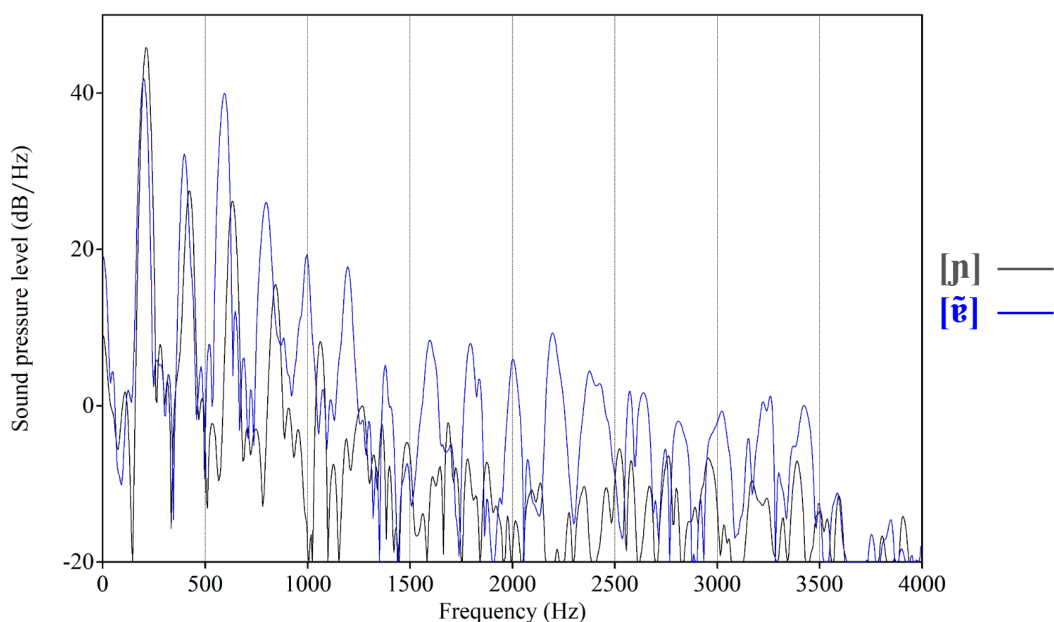


Figura 96 – [ɲ] vs. [ẽ]



Neste caso da análise espectral fica evidente o que não pôde ser observado tão claramente na amplitude do formato de onda. Portanto, na análise espectral é possível perceber uma amplitude maior da vogal adjacente [ẽ] em comparação com o espectro da consoante nasal [ɲ]. O amortecimento fica ainda mais claro quando comparamos os valores de frequência desses dois segmentos. Em [ẽ] temos um F1 de 568 Hz que cai para 282 Hz quando medimos o FN1 de [ɲ].

Quando comparamos o espectro das três consoantes, observamos uma diferença clara entre suas frequências. Isso pode ser percebido a partir do centro de gravidade, que seria a média da distribuição da frequência do segmento. Em [m] encontramos um centro de gravidade de 211 Hz, em [n] de 239 Hz e um [ɲ] de 224 Hz, evidenciando a diferenciação entre [m] e [n] a partir de valores menores em [m]. O fone [ɲ] não entra nessa comparação pelo fato de que não encontramos estudos suficientes que apresentem tendências ou máximas de características acústicas que permitam a diferenciação de [ɲ] das demais consoantes nasais existentes na língua

3.4.4 Consoante Fricativas

A partir de agora, faremos uma análise acústico-articulatória das fricativas do Arutani: [s], [ʃ] e [h]. Barbosa e Madureira (2015, p. 176) definem as fricativas como sons gerados por estreitamentos no trato vocal, produzindo ruído contínuo. Entre as seis características acústicas destacadas por Silva *et al.* (2019, p. 168) para identificação dessas consoantes estão: ruído, barra de vozeamento, amplitude, pico espectral, F2 de transição e duração. O ruído é visualizado no espectrograma por hachuras verticais e se manifesta como aperiodicidade na forma de onda (SILVA *et al.*, 2019, p. 170).

A barra de vozeamento é um critério fundamental para distinguir as fricativas vozeadas das não-vozeadas. No entanto, a análise do *corpus* do Arutani falado por Juanita revela apenas fricativas não-vozeadas. A amplitude é outro parâmetro relevante, variando conforme o ponto de articulação: fricativas mais anteriores, como as labiodentais, exibem menor amplitude do que as alveopalatais, que são mais posteriores (SILVA *et al.*, 2019, p. 173). Essa hierarquia de amplitude reflete o grau de abertura do trato vocal, sendo esperada na análise dos dados do nosso *corpus*.

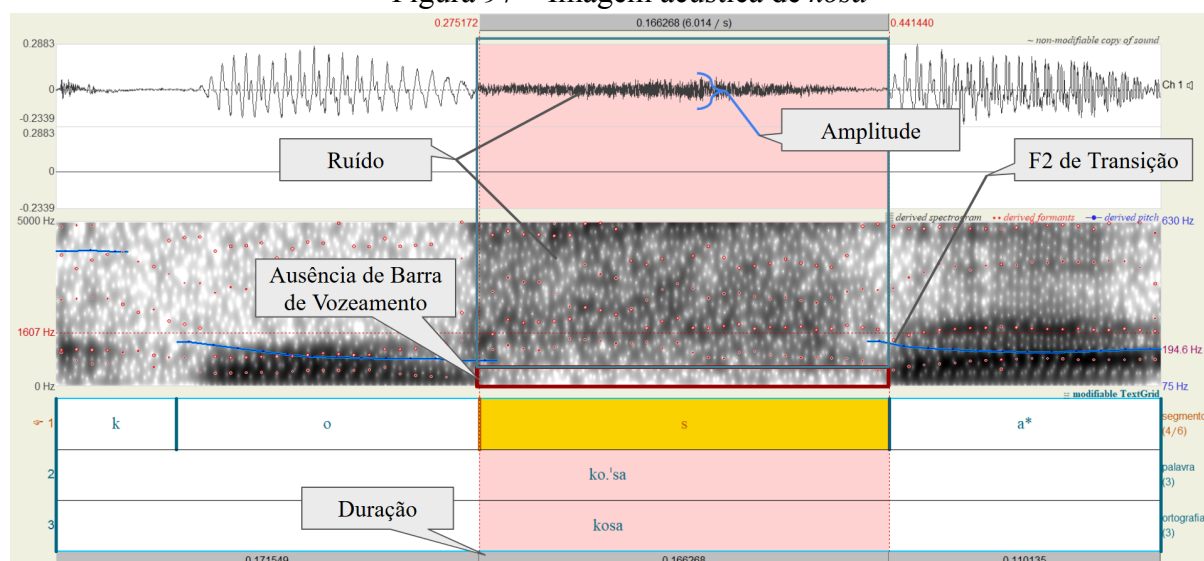
A diferenciação entre os sons [s] e [ʃ] deve ser baseada em dois parâmetros acústicos cruciais: o pico espectral e o F2 de transição. Conforme Silva *et al.* (2019, p. 175), fricativas com articulação mais anterior, como [s], possuem pico espectral mais alto, enquanto [ʃ], sendo mais posterior, exibe pico mais baixo. Além disso, a cavidade posterior menor de [ʃ] em comparação a [s] afeta diretamente o F2 de transição, permitindo distinguir essas consoantes no espectro acústico. Por fim, embora a duração das fricativas seja um traço útil para diferenciar vozeadas de não-vozeadas, esse parâmetro tem pouca relevância no presente estudo, dada a ausência de fricativas vozeadas no *corpus*.

Em uma escala de produtividade podemos ordenar as fricativas da língua da seguinte forma:

Tabela 31 – Quantitativo de consoantes fricativas

Consoantes Fricativas	Nº de Ocorrências
[s]	68
[ʃ]	63
[h]	44

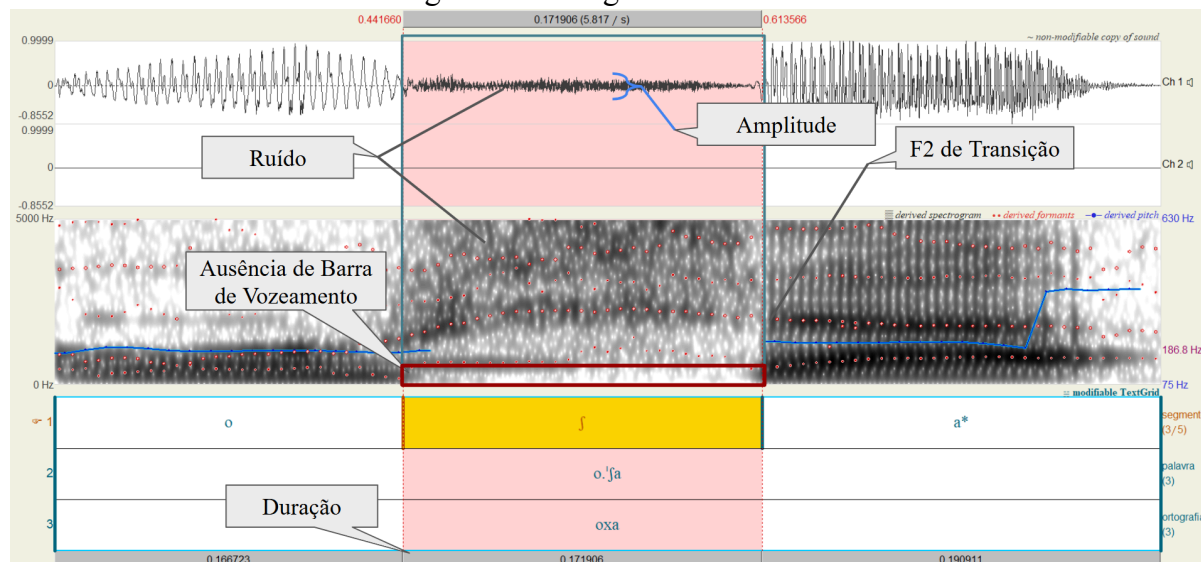
As fricativas do nosso *corpus* ocorrem majoritariamente em ataque silábico, por isso, não as analisaremos em outros contextos. Além disso, a partir de um viés comparativo, iremos focar, nesta seção, mais na distinção acústica entre as fricativas, para que seja possível uma melhor distinção entre elas, ao invés de observarmos suas diferentes propriedades a partir de ambientes distintos. Iniciaremos nossa descrição com a consoante fricativa alveolar [s], na palavra *kosa* [ko.'sa], que significa ‘seio’.

Figura 97 – Imagem acústica de *kosa*

Na Figura 99, encontramos um [s] de 166 ms de duração sem a presença de barra de vozeamento, o que evidencia o fato de ser uma consoante não-vozeada. Além disso, é possível observar o ruído, principal característica das fricativas, tanto no espectrograma quanto na forma de onda, isso fica mais evidente no espectrograma já que nele encontramos as hachuras verticais. O F2 de transição de [s] é de aproximadamente 1.370 Hz, esse valor será importante para

compararmos com o valor encontrado em [ʃ], fricativa pós-alveolar. Para uma melhor comparação, escolhemos a palavra *oxa* [o.ʃa], já que as fricativas encontram-se no mesmo ambiente. A palavra *oxa* em Arutani significa ‘pelo’ ou ‘cabelo’. Observe sua imagem acústica:

Figura 98 – Imagem acústica de *oxa*



Em [o.ʃa] encontramos um [ʃ] de 171 ms de duração. Pode-se observar, também, na Figura 100, que há ausência da barra de vozeamento, evidenciando tratar-se de uma consoante não-vozeada. Ademais, é possível observar as hachuras no espectrograma o que evidencia o ruído característico das fricativas. No espectrograma também conseguimos medir o F2 de transição, que neste caso é de aproximadamente 1.980 Hz.

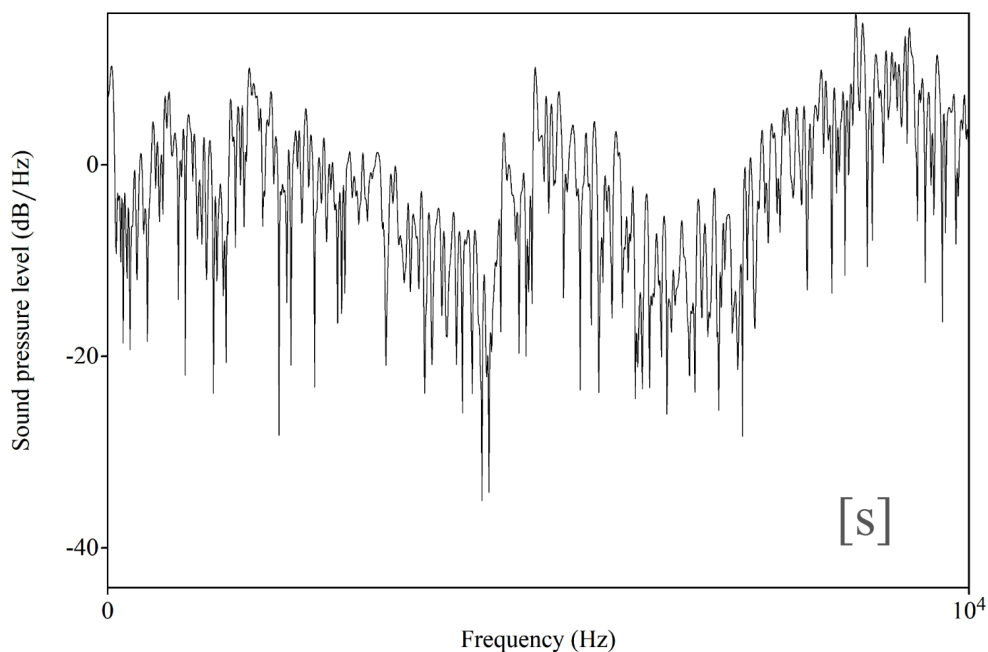
Quando comparamos o F2 de transição de [s] (~1.485 Hz) e de [ʃ] (~1.949 Hz) comprovamos o que Silva *et al.* (2019, p. 175-176) apresentam sobre a relação entre o tamanho da cavidade no momento da constricção em [ʃ] ser menor do que em [s] e isso é refletido diretamente na diferença dos valores de F2 de transição. sendo o valor de [ʃ] maior do que o de [s]. Quando fazemos uma análise mais precisa com os dois primeiros formantes de transição essa diferença fica mais clara, ainda mais que é levado em consideração não apenas o avanço ou recuo da língua, mas também a sua altura com relação à abertura no maxilar.

Tabela 32 – Valores de F1 e F2 de transição de [s] e [ʃ]

Consoantes Fricativas	F1 de transição	F2 de transição
[s]	485 Hz	1.408 Hz
[ʃ]	509 Hz	1.949 Hz

Outra característica das fricativas que leva em consideração a cavidade anterior à constrição é o pico espectral, que serve também para a distinção entre modo de articulação entre [s] e [ʃ]. Silva *et al.* (2019, p. 175) apresentam que, com relação ao pico espectral, “quanto menor a cavidade anterior à constrição maior será a frequência do pico espectral”, ou seja, esperamos que o pico espectral de [s] ocorra em uma frequência maior do que em [ʃ]. Dessa forma, como veremos nos espectros, espera-se que o pico espectral de [ʃ] seja mais à esquerda e de [s] mais à direita. Vejamos o pico espectral da fricativa [s]:

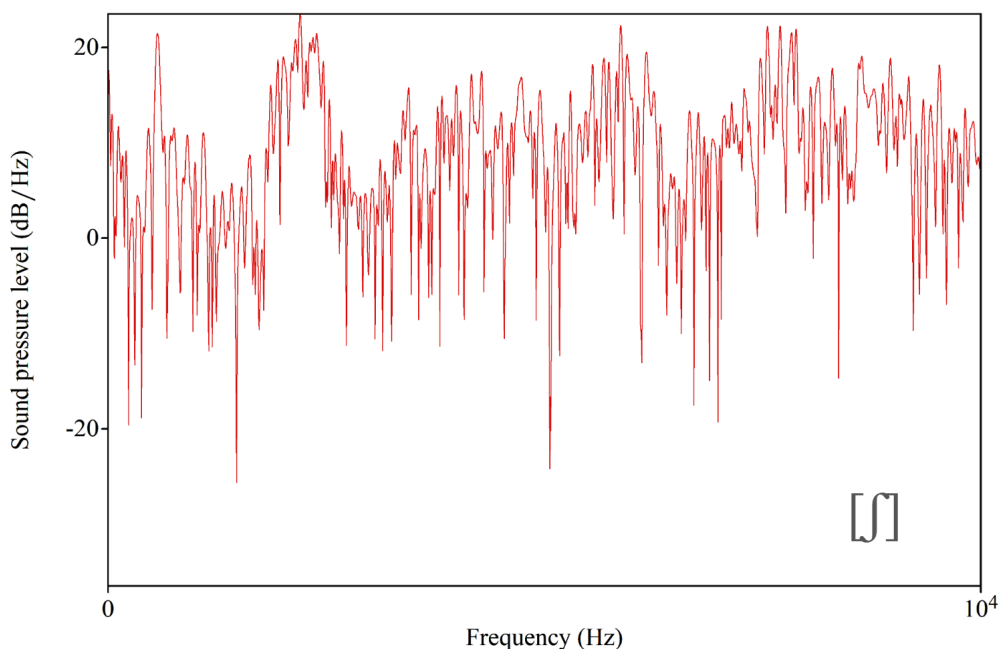
Figura 99 – Espectro de [s]



O [s] em [ko.'sa] tem um centro de gravidade de 9.046 Hz, que é a média da distribuição da frequência, representando onde há maior concentração. A partir da média é possível observarmos uma variância de 9.381 Hz e uma distorção de - 0,70, que nos mostra um espectro

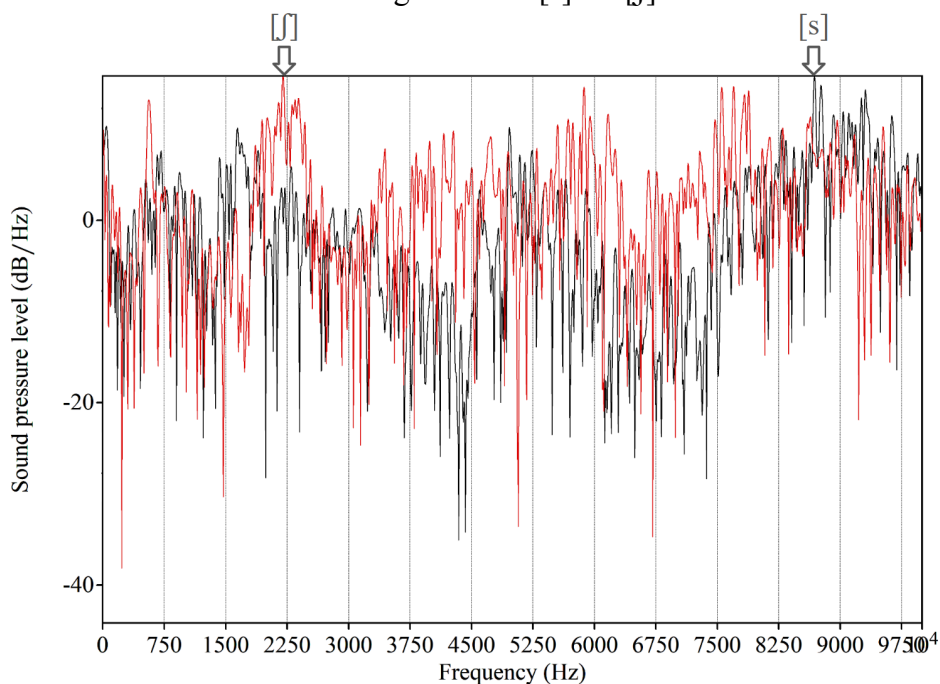
com valores de frequência mais concentrados à direita. Por fim, temos uma curtose de 0,32. Para o [ʃ] em [o.'ʃa], temos o seguinte espectro:

Figura 100 – Espectro de [ʃ]



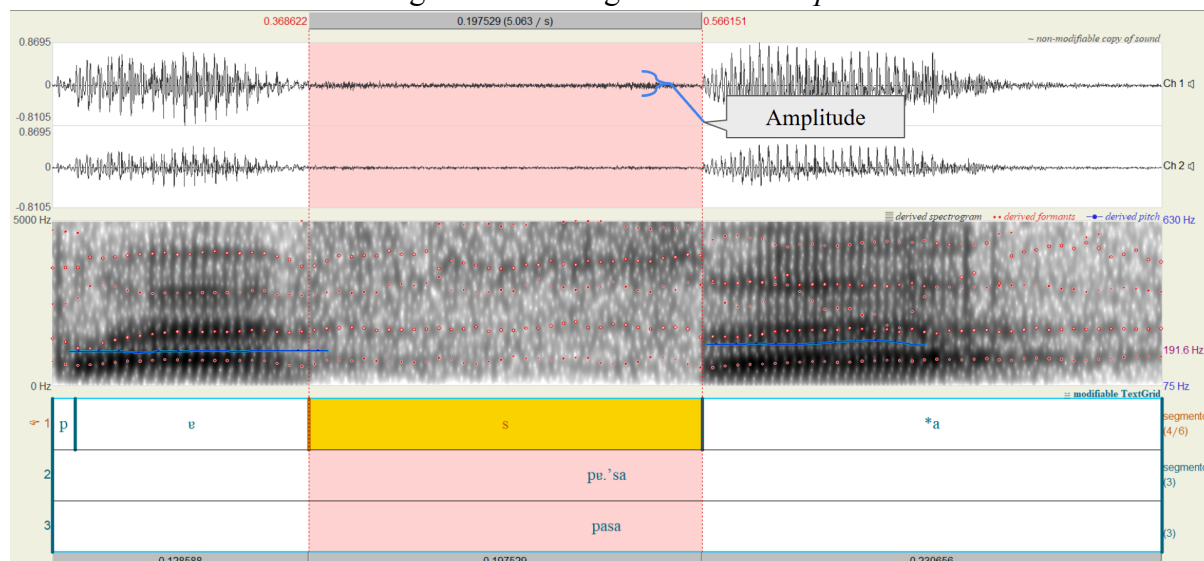
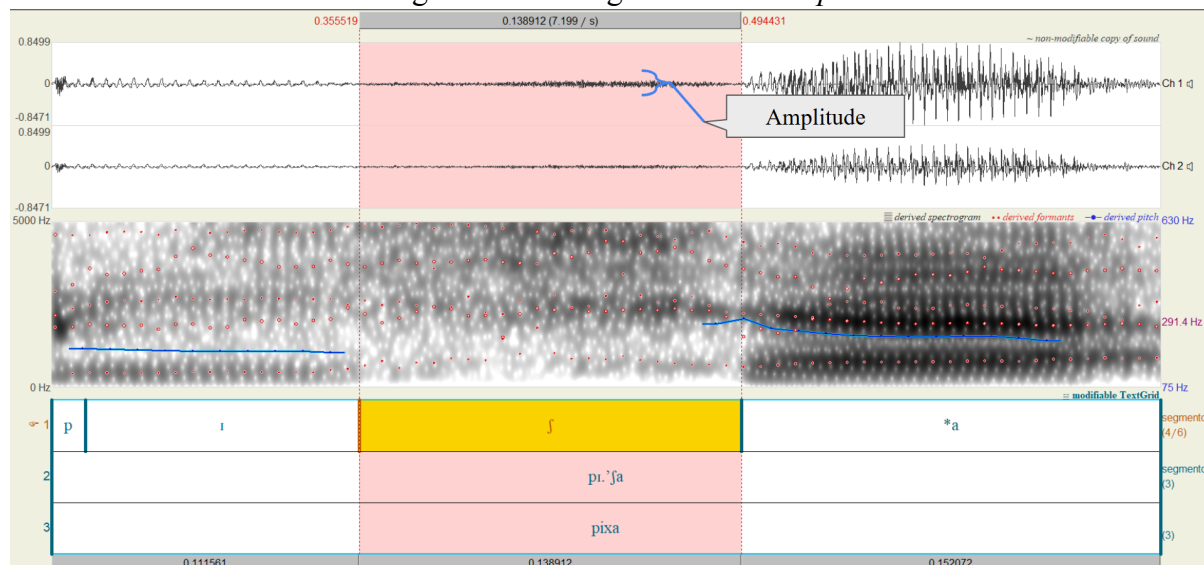
Em [ʃ] encontramos um centro de gravidade de 6.077 Hz, o que nos mostra que ele está mais concentrado à esquerda em relação ao [s]. A distribuição de energia em torno da média é de 3.439 Hz e a distorção é de 0,34, evidenciando o fato de que a energia está mais concentrada à esquerda. Com relação à curtose, encontramos um valor de - 0,12. Quando sobrepomos os dois espectros e identificamos seus picos espectrais, fica ainda mais claro que estamos tratando de segmentos distintos e fica nítida a diferença acústica entre [s] e [ʃ] quanto ao ponto de articulação. Veja:

Figura 101 – [s] vs. [ʃ]



As setas acima do espectro indicam o pico espectral da respectiva fricativa. O pico de [s] ocorre entre 8.250 Hz e 9.000 Hz, enquanto o de [ʃ] ocorre próximo a 2.250 Hz. Isso nos mostra justamente a questão que Silva *et al.* (2019, p. 175) destaca quanto a relação entre o tamanho da cavidade com a frequência dos picos, em outras palavras, com o pico mais à direita temos a produção de um [s] já que seu correlato articulatório seria uma cavidade menor, possuindo, conseqüentemente, uma frequência maior. Essa relação é inversamente proporcional, quanto maior a frequência, menor será a cavidade, e vice-versa. Portanto, como na produção de [ʃ] a tamanho da cavidade anterior à constrição é maior, encontramos uma frequência mais baixa. Isso fica mais evidente ainda quando comparamos os valores de centro de gravidade, em que o de [s] ocorre numa frequência de 9.046 Hz e de [ʃ] numa frequência de 6.077 Hz.

Algo inesperado e que se repete em nossos dados é que o fone [ʃ], na fala de Juanita, não possui uma amplitude de onda muito distinta da de [s], já que o esperado é que a amplitude de [ʃ] fosse consideravelmente maior que a de [s]. Isso pode ser observado quando comparamos esses dois fones em outras palavras como *pasa* [pɛ.'sa] – que é o nome em Arutani para o peixe ‘jundiá’ – e *pixa* [pi.'ʃa] – nome em Arutani para a espécie de macaco ‘cuxiú’.

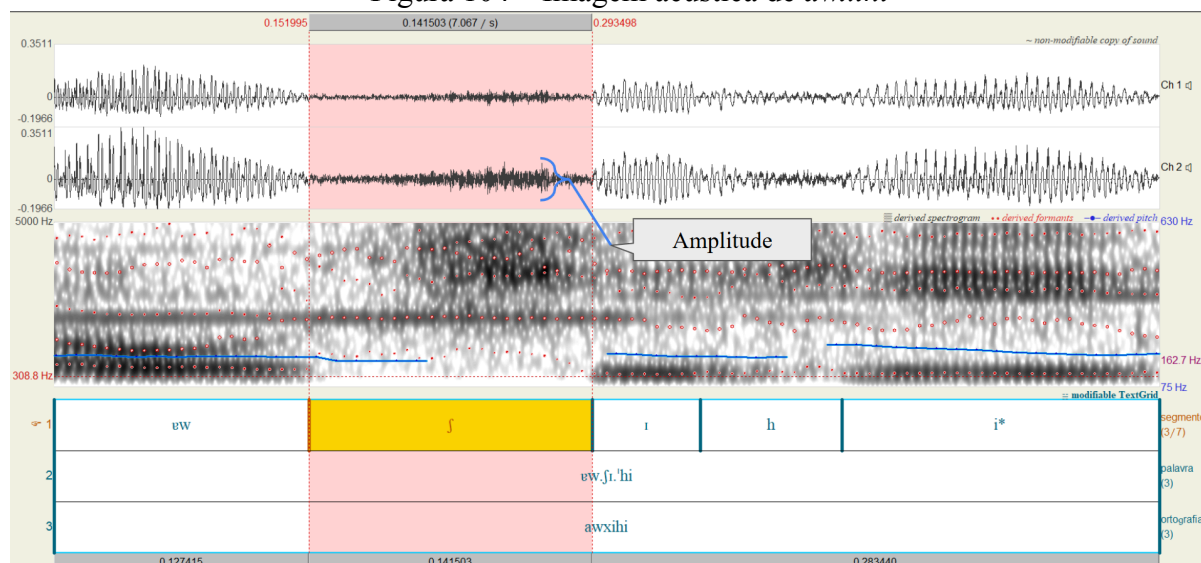
Figura 102 – Imagem acústica de *pasa*Figura 103 – Imagem acústica de *pixa*

Com base nessas duas imagens é possível perceber apenas uma amplitude levemente maior em [x] quando comparada com a amplitude apresentada em [s]. Portanto, a amplitude não é uma das melhores formas de diferenciação de fricativas na fala de Juanita. Dessa forma, para isso, os melhores recursos de distinção são a análise e a comparação espectral, levando em consideração os picos espectrais, e o F2 de transição, já que não há fricativas vozeadas e a barra de vozeamento e a duração servem mais para compararmos entre vozeadas e não-vozeadas.

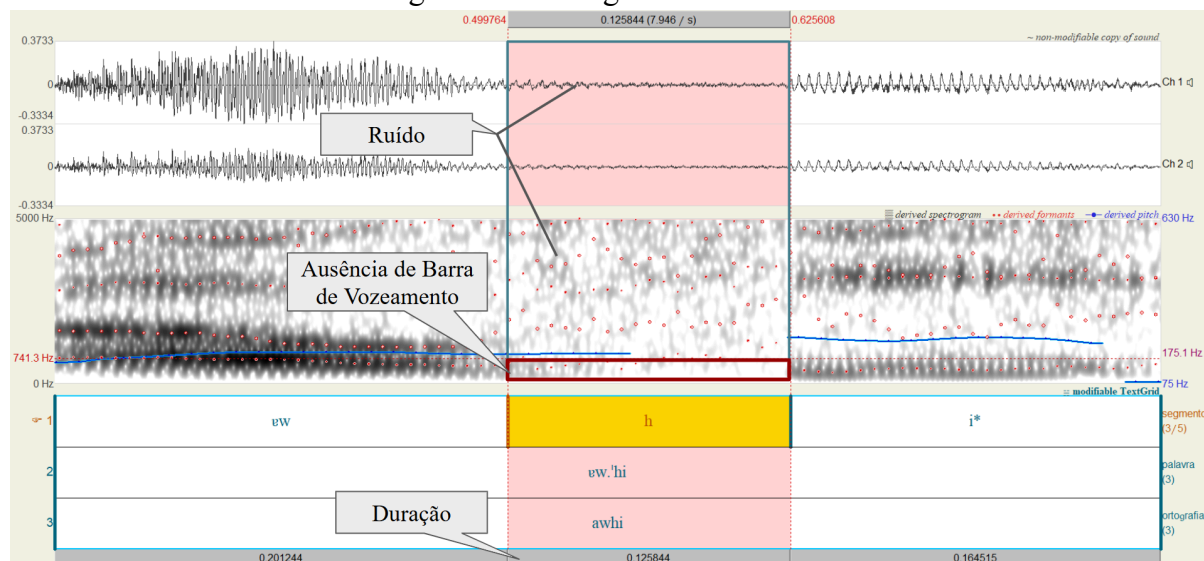
Além disso, a percepção sempre será um recurso para esse tipo de distinção, pois são claramente ouvidos como [s] e [ʃ], respectivamente.

Entretanto, ainda assim, encontramos algumas consoantes fricativas alveopalatais não-vozeadas [ʃ] que possuem uma frequência maior se comparadas com a ocorrência de outras fricativas alveolares [s]. Um exemplo está na Figura 106, a seguir, na palavra *awxihi* [ɐw.ʃi.'hi] que significa ‘ubim’, espécie de palmeira. Observe a frequência mais elevada captada pelo canal 2 (Ch 2) na forma de onda de [ʃ]:

Figura 104 – Imagem acústica de *awxihi*

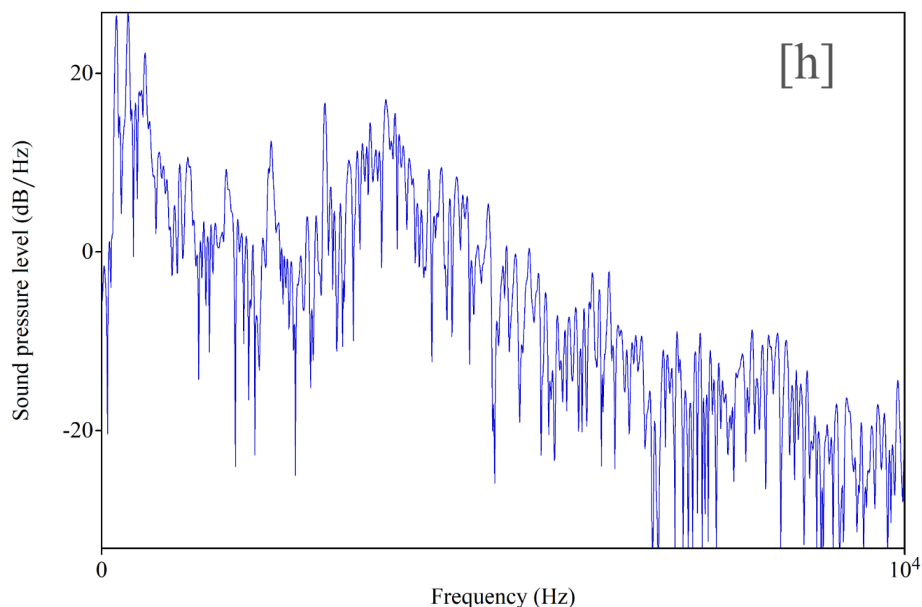


Como pode-se observar, em [ɐw.ʃi.'hi], assim como em *awhi* [ɐw.'hi] – nome em Arutani para a ave ‘sururina’ –, temos a ocorrência da consoante fricativa glotal não-vozeada [h]. Em [ɐw.'hi] temos um áudio com menos ruídos externos e, portanto, é possível analisar melhor essa fricativa posterior, que é a única desse tipo de fricativa encontrada no idioleto de Juanita. Essa consoante é evidenciada pelo ruído, visível principalmente no espectrograma, e a ausência de vozeamento em F0, o que comprova tratar-se de uma fricativa não-vozeada. O [h] em [ɐw.'hi] possui uma duração de aproximadamente 125 ms, ou seja, possui uma duração adequada com o que se espera de uma fricativa não-vozeada dentro dos valores encontrados nas outras fricativas da língua.

Figura 105 – Imagem acústica de *awhi*

Em uma comparação com as fricativas anteriores, Silva *et al.* (2019, p. 185) explica que “o principal correlato acústico das fricativas posteriores é o ruído decorrente da constrição”, que foi destacado na segmentação da figura anterior. Além disso, os autores destacam que as hachuras verticais espaçadas no espectrograma são a representação acústica do ruído e que “fricativas posteriores apresentam concentração de energia em baixas frequências, na região dos dois primeiros formantes”, o que também é evidenciado na Figura 107. Tanto que, em relação aos formantes de transição, encontramos valores baixos de frequência, sendo 403 Hz para F1 e 1.919 Hz para F2. Quando fazemos uma análise FFT e geramos um espectro do [h] em [ew.'hi] encontramos o seguinte:

Figura 106 – Espectro de [h]



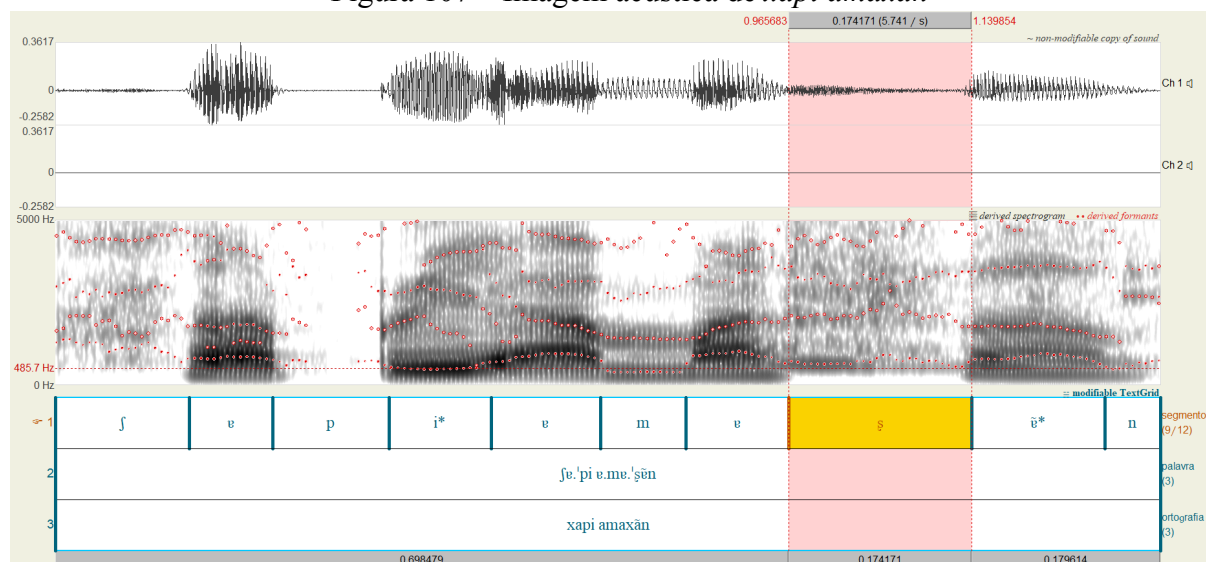
No espectro de [h] temos um centro de gravidade por volta de 1.217 Hz e uma variância de 1.452 Hz. Uma distorção de 1,47, nos mostrando que a energia está mais concentrada à esquerda. Por último, uma curtose de 2,9, o que nos mostra ser um espectro mais pontiagudo do que os encontrados em [s] e [ʃ].

Com base no que foi apresentado em Chacon *et al.* (2020), o Arutani, baseado na fala de Juanita teria quatro consoantes fricativas [s], [h] e uma variação entre [ʃ] e [ʂ]. Entretanto, revisando os dados não é bem isso o que encontramos. Levando em consideração a hipótese inicial, de uma variação entre [ʃ] e [ʂ], portanto, tratando-se de dois fones distintos, compramos-a com trabalhos antigos, como os de Migliazza (1967; 1978), e percebemos a existência apenas de [s] e [ʃ], o que nos fez revisar a hipótese levantada.

Se pensarmos no trato vocal como um *continuum*, perceberemos que um leve recuo da língua pode fazer com que certos segmentos mudem e sejam transcritos de uma maneira diferente. Dessa forma, ao escutarmos um som fricativo pós-alveolar não-vozeado [ʃ] e um fricativo retroflexo não-vozeado [ʂ] percebemos justamente isso, apenas um leve recuo da língua em direção à parte anterior do trato vocal. O que percebemos é que o que estava sendo transcrito

no dicionário como [ʃ], na verdade, tratava-se de uma espécie de ceceo lateral²⁵ (cf. Leite *et al.*, 2008, p. 31) presente na fala de Juanita, talvez por conta da idade, ou não, mas que todos nós estamos suscetíveis em nossa fala diária. A única transcrição em que encontramos o segmento [ʃ] em Chacon *et al.* (2020) é no nome *xapi amaxãn* [ʃa'pi ama'ʃẽn], que significa ‘tábua’. Observe sua imagem acústica com correções apenas na transcrição das vogais de [a] para [e]:

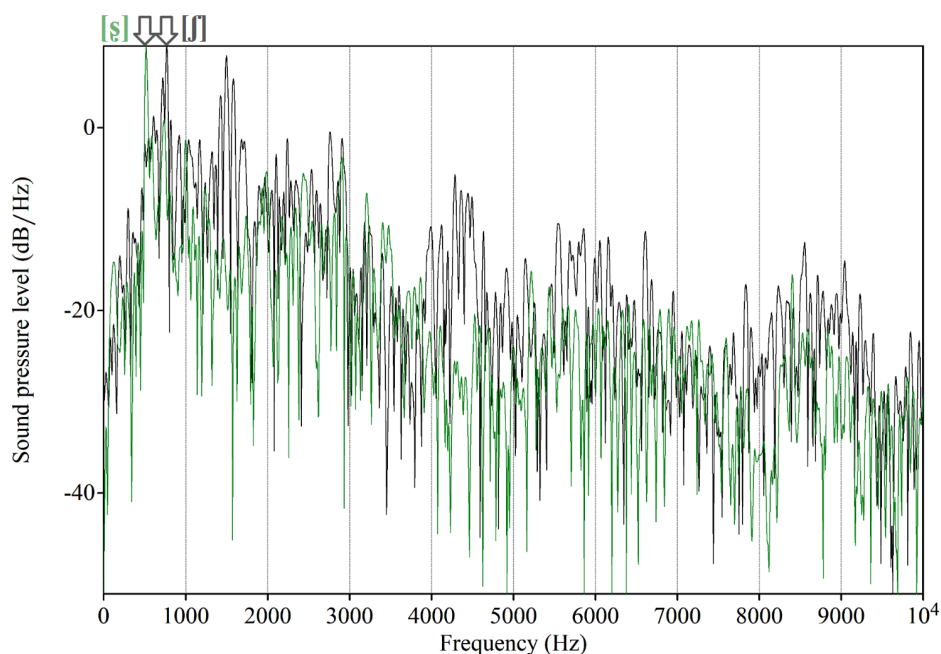
Figura 107 – Imagem acústica de *xapi amaxãn*



De uma perspectiva perceptual e oitiva, os sons [ʃ], na primeira sílaba, e [ʃ], na sílaba tônica, parecem realmente ser iguais e a imagem acústica, tanto no formato de onda quanto no espectrograma, nos faz perceber que se trata de segmentos fricativos. Quando extraímos os espectros de cada um deles com base na análise FFT encontramos o seguinte:

²⁵ As autoras apresentam que: “Os fonemas fricativos línguo-alveolares /s/ e /z/ são produzidos com constrição da corrente aérea entre a ponta da língua e o alvéolo inferior. É necessário que a borda da língua assuma forma e posição específica e precisa para que o fluxo aéreo possa ser direcionado para os incisivos inferiores. Quando ocorre alteração dessas condições o som é produzido de forma incorreta. O termo ceceo se refere à incorreta produção dos fricativos línguo-alveolares” (LEITE *et al.*, 2008, p. 31). Sobre o ceceo lateral elas o descrevem como quando “ocorre quando a ponta da língua faz pressão sobre os incisivos centrais superiores ou no alvéolo, a corrente de ar é dividida e observa-se um escape lateral do ar pela cavidade oral”.

Figura 108 – [ʃ] vs. *[ɣ]



Os valores dos picos espectrais são próximos também. Em [ʃ] o pico está por volta de 750 Hz e de [ɣ] por volta de 500 Hz, portanto, uma diferença muito pequena. O pico espectral é um dos correlatos acústicos do ponto de articulação, evidenciando o fato de serem segmentos com o ponto de articulação bem próximo. Quando analisamos alguns dados desses espectros fica ainda mais claro que se trata de segmentos muito próximos. Observe a tabela a seguir:

Tabela 33 – Comparação dos dados espectrais de [ʃ] vs. *[ɣ]

Segmentos	Centro de Gravidade	Variância	Distorção	Curtose
[ʃ]	1.766 Hz	1.804 Hz	3,9	19
[ɣ]	1.341 Hz	1.498 Hz	3,6	20

Esses valores estão bem próximos e podem variar por influência do ambiente e sons adjacentes. Um outro ponto importante fora da análise acústica para nossa hipótese de que não há [ɣ] na língua é o fato de que não há mais nenhum outro segmento retroflexo na língua, ponto que já nos faz gerar certa desconfiança. Portanto, tanto pela questão do possível ceceo lateral quanto

pela análise espectral, não faz sentido postularmos a existência de [ʂ] no quadro fonético da fala do Arutani baseado no idioleto de Juanita.

O papel difícil do linguista, em momentos assim, é de ter a sensibilidade e indícios suficientes para que seja possível diferenciar o que realmente pode ser classificado como **fone** e o que se trata de questões que surgem por conta do envelhecimento, alguma afasia articulatória ou mesmo o resultado de uma fala mais “relaxada”. Isso tudo precisa ser bem definido e bem esclarecido, pois pode gerar problemas futuros em análises e descrições de **fonemas**.

3.4.5 Tepe

A consoante conhecida como tepe [ɾ] aparece 101 vezes em nossos dados, mostrando-se muito produtiva e recorrente no idioleto de Juanita. O tepe é um segmento consonantal, normalmente, produzido por meio de “um único movimento, rápido e pontual, que causa uma breve obstrução da passagem da corrente de ar na região alveolar” (SILVA *et al.*, 2019, p. 190). Estes autores apresentam que o tepe possui três características acústicas: 1. ausência de energia; 2. ruído transiente (*burst*); e 3. vogal de apoio. Porém a vogal de apoio só aparece em contexto de encontro consonantal, o que não é o caso do Arutani, já que o tepe [ɾ] só ocorre em posição de ataque silábico, a não ser no único caso que é um empréstimo, em ‘sambbrero’.

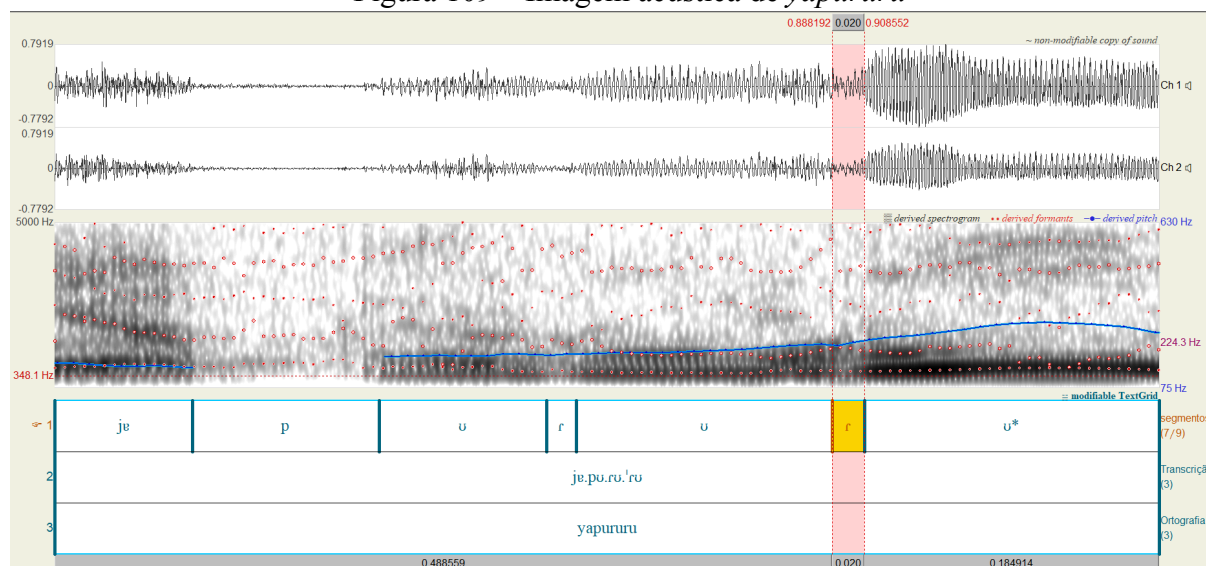
A produção do [ɾ] ocorre em dois momentos: primeiro, um fechamento passageiro que interrompe o fluxo de ar, seguido por uma abertura igualmente rápida, permitindo o restabelecimento da passagem do ar. Essa obstrução temporária é definida como uma oclusão muito breve, destacada pela natureza transitória do som (SILVA *et al.*, 2019, p. 190-191).

A ausência de energia é observada tanto no espectrograma quanto na forma de onda. No espectrograma, segundo Silva *et al.* (2019, p. 191), a identificação do [ɾ] será possível por meio da visualização de uma região mais clara quando comparada com os sons adjacentes. Já na forma de onda, segundo os mesmo autores, a ausência de energia é observada por meio da baixa amplitude das ondas. O *burst*, entretanto, é de mais difícil visualização, pois é possível de ser observado mais claramente apenas no espectrograma e, como destacado por Silva *et al.* (2019, p. 192) ele “pode ser muito breve ou mesmo não ocorrer como correlato acústico”. Um outro ponto

destacado pelos autores é que o tepe tende a possuir barra de vozeamento visível no espectrograma.

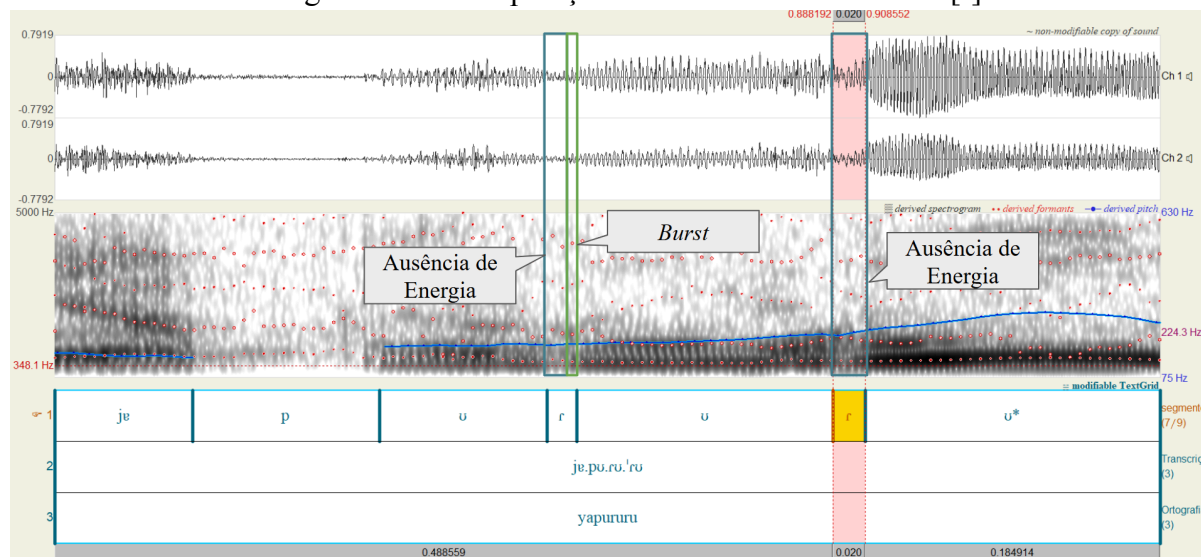
Dessa forma, vamos explorar, na fala de Juanita, apenas duas características principais do tepe, a ausência de energia e a ocorrência, ou não, do *burst* em sua produção. O [r] pode ser encontrado na palavra *yapururu* [jɐ.ɸʊ.rʊ.'rʊ] ‘banana-comprida’ em posição de sílaba tônica e pretônica:

Figura 109 – Imagem acústica de *yapururu*



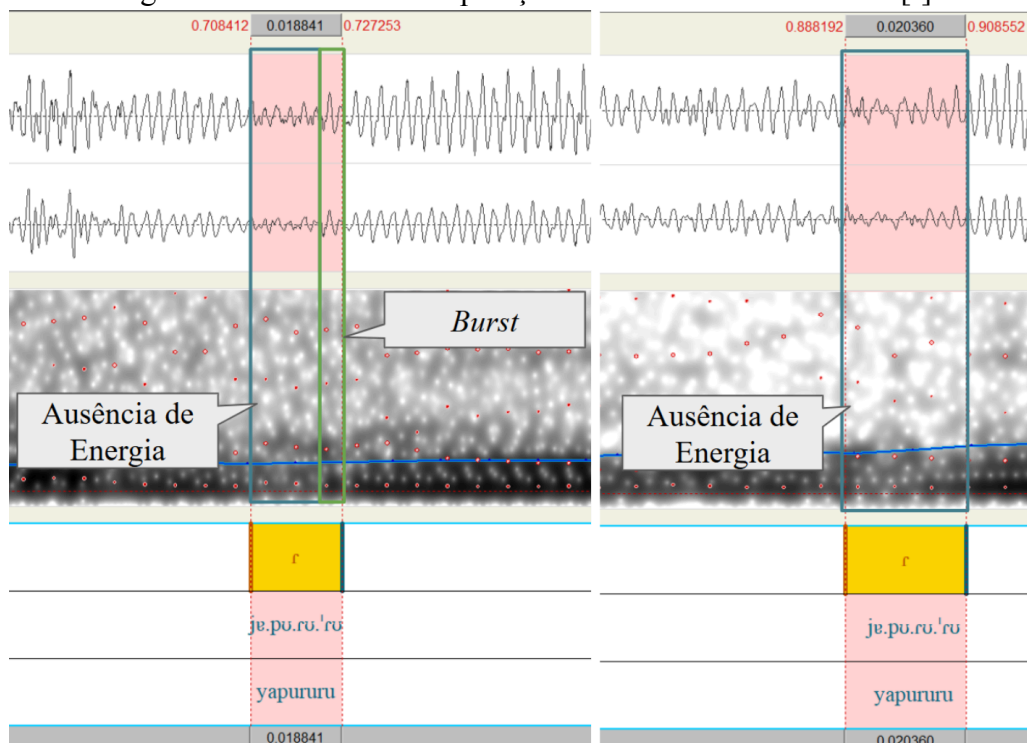
Para uma segmentação mais adequada, utilizamos tanto o espectrograma quanto a forma de onda para nos guiarmos. O [r] da sílaba pretônica é menos proeminente do que o [r] em sílaba tônica, sendo que o primeiro possui aproximadamente 18 ms de duração, enquanto o segundo possui aproximadamente 20 ms de duração. Como já apresentado, o *burst* pode ser visível ou não na imagem acústica e é mais visível no espectrograma, e isso pode ser observado quando comparamos os dois tepes. Lembrando que o tepe pode ser dividido em dois momentos principais: ausência de energia e *burst*, e que nem todo tepe terá necessariamente um *burst* visível como correlato acústico, observamos o seguinte na Figura 112.

Figura 110 – Comparação entre duas ocorrências de [r]

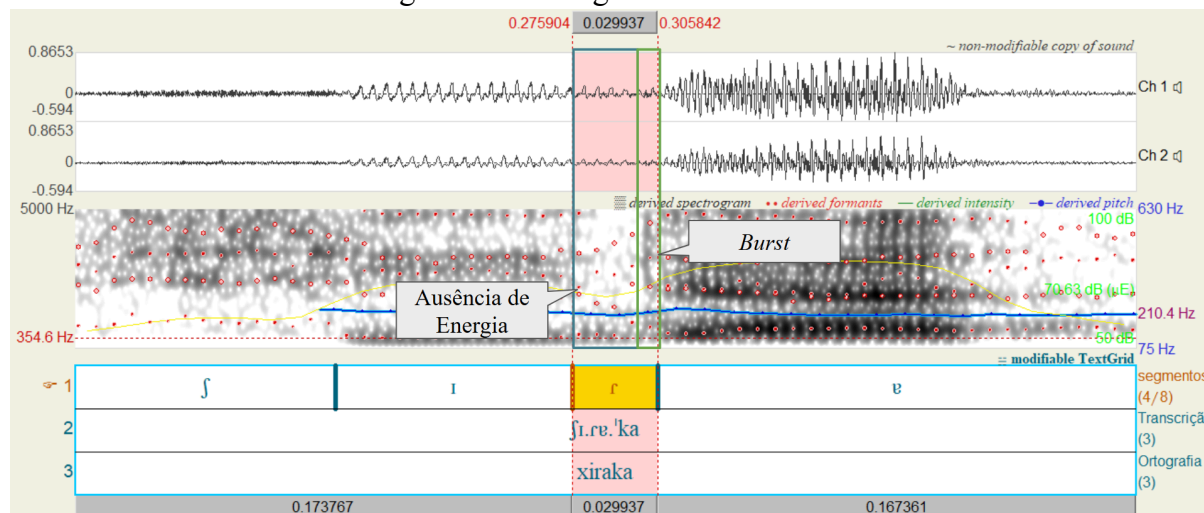


Na Figura 112, essas questões ficam comprovadas ao observarmos que no [r] em sílaba pretônica há o momento de ausência de energia e um breve um *burst* e no [r] em sílaba tônica é possível observar apenas o momento de ausência de energia. Essa diferença fica mais clara quando observamos a forma de onda e o espectrograma mais de perto. Na Figura 113, a seguir, temos à esquerda o [r] pretônico e oônico à direita. Na imagem do tepe pretônico, observe que há um leve escurecimento no espectrograma na região verde quando comparado com o espectrograma na região azul. Além disso, veja como há uma diferença na frequência no formato de onda, em que a frequência da onda dentro da região verde é mais elevada quando comparada com a frequência dentro do retângulo azul. Esse escurecimento no espectrograma não ocorre na imagem à direita e não há muita diferença entre as frequências na região azul, evidenciando o fato de que na fala de Juanita existem tepes com e sem *burst*.

Figura 111 – Zoom na comparação entre duas ocorrências de [r]



Apesar disso, ainda é possível encontrar na fala de Juanita tepes prototípicos com as características esperadas mais marcadas. Isso ocorre em *xiraka* [ʃi.rɐ.'ka], palavra que significa ‘cesto’ e que já foi analisada anteriormente quando estávamos tratando da consoante oclusiva [k]. O [r] nesta palavra possui ausência de energia bem clara e o *burst* bem evidente. Veja:

Figura 112 – Imagem acústica de *xiraka*

A seguir, iniciaremos uma análise de caráter mais fonológico do idioleto de Juanita. Portanto, a partir de agora nosso foco nas seguintes seções será na definição dos fonemas vocálicos e consonantais, no padrão acentual, na fonotática e em alguns fenômenos fonológicos.

4. Análise Fonológica

Após a definição, descrição e comprovação por meio da análise acústica de todos os fones encontrados no Arutani falado por Juanita, podemos partir para uma análise de caráter mais fonológico. Dessa forma, podemos observar a importância que os estudos fonéticos têm quando tratamos da descrição da fonologia de uma língua. Com isso, nesta seção, veremos, também, como a fonologia explica certos fenômenos que só podem ser compreendidos no âmbito fonológico.

Já apresentamos que, para autores como Hayes (2009), há uma relação direta entre fonética e fonologia, ao explicarem, por exemplo, que os dados que dão base para estudos fonológicos são de origem fonética. Matzenauer (2014, p. 11-12) explica que a fonologia é responsável pelo estudo da organização sistemática dos sons em uma língua, enquanto a fonética analisa a realidade física desses sons sob três perspectivas: articulatória, acústica e auditiva. A autora explica que, embora relacionadas, a fonética investiga todos os sons da fala, enquanto a fonologia se concentra nos fonemas, os quais distinguem significados, e na forma como se combinam para formar unidades linguísticas maiores. A identificação dos fonemas ocorre pelo processo de comutação, que verifica a substituição de sons em contextos semelhantes para detectar diferenças de significado (MATZENAUER, 2014, p. 11-12).

Como vimos, a partir de uma perspectiva fonética, identificamos os seguintes fones consonantais: [p], [p̃], [b], [b̃], [t], [t̃], [k], [kʷ], [k̃], [ʔ], [m], [n], [ɲ], [s], [ʃ], [h] e [r]. Também identificamos as seguintes vogais: [a], [ɐ], [e], [i], [ɪ], [o] e [u], todas com sua contraparte nasal, além dos *glides* [w] e [j]. Levando esses fones em consideração, iremos iniciar nossa análise a partir do agrupamento de sons foneticamente semelhantes (SFS), retomando o roteiro de análise que é apresentado em Pike (1947) e Mori (2000).

Sendo assim, iniciaremos primeiramente nossa análise contrastando os fones dos grupos dos segmentos consonantais, inicialmente agrupando-os e contrastando-os de acordo com traços articulatórios em comum. Em seguida, faremos um contraste entre as vogais, levando em consideração se ocorrem em contexto tônico ou átono e se estes ambientes promovem a

ocorrência de distribuição complementar. No caso dos *glides*, também conhecidos como aproximantes ou segmentos ambivalentes, iremos investigar seu *status* dentro da língua, ou seja, se tratam-se de vogais ou consoantes fonológicas.

Para a análise fonológica, o *corpus* utilizado foi maior do que o *corpus* da análise acústica. Além do *corpus* que montamos com as transcrições e segmentações a partir das análises acústicas, também utilizamos os relatos escritos do dicionário Arutani, disponível em Chacon *et al.* (2020). Utilizamos as transcrições do próprio dicionário pelo fato de que no caso de algumas palavras não foi possível identificar o áudio, apenas o registro escrito das transcrições. Algumas outras palavras encontramos sendo pronunciadas em áudios de frases do dicionário e utilizamos esses áudios para identificá-las e transcrevê-las. A maior parte desses dados estão disponíveis no Dicionário Arutani em Chacon *et al.* (2020).

Após uma análise fonética aprofundada que nos deu base para analisarmos a fonologia da língua, chegamos aos seguintes fonemas consonantais e vocálicos:

Tabela 34 – Quando fonológico das consoantes

	Labial	Alveolar	Pós-alveolar	Palatal	Velar	Velar Labializado	Velar Palatalizado	Glotal
Oclusiva	/p/	/t/			/k/	/k ^w /	/kʲ/	
Nasal	/m/	/n/		/ɲ/				
Tepe		/ɾ/						
Fricativa		/s/	/ʃ/					/h/
Glide	/w/			/j/				

Tabela 35 – Quando fonológico das vogais

	Anterior	Central	Posterior
Alta	/i/		
Média-alta			/o/
Baixa		/a/	

4.1 Consoantes

Os fonemas consonantais presentes no Arutani falado por Juanita são: /p/, /t/, /k/, /kʷ/, /kʲ/, /m/, /n/, /ɲ/, /ɾ/, /s/, /ʃ/, /h/, /w/ e /j/. Para a nossa análise, agrupamos primeiramente os sons foneticamente semelhantes, em seguida, agrupamos alguns fones que não possuem muitos traços em comum, mas que ainda assim reforçam questões de contraste. Para evidenciar a existência desses fonemas iremos contrastá-los, sempre que houver a possibilidade, em ambientes idênticos e/ou análogos. Nesta seção, exploraremos, também, questões de alofonia por variação livre e por distribuição complementar, além da variação da pronúncia de algumas palavras.

4.1.1 Oclusivas não-vozeadas /p/, /t/, /k/

[p]	vs.	[t]	vs.	[k]
[ɐ.'pa]		[ɐ.'ta]		[ɐ.'ka]
‘pai’ (vocativo e referencial)		‘irmã mais velha’		‘dente’
		[o.'to]		[o.'ko]
		‘peixe’ (espécie)		‘zogue-zogue’
[kɐ.'pa]		[kɐ.'ta]		
‘mastigar’		‘ingá’		

❖ Fonemas distintos: /p/, /t/ e /k/

➤ Fones [p], [t] e [k] ocorrendo em ambiente idêntico e contrastando significado.

[p]	vs.	[p̰]	vs.	[b̰]
[ɐ.'tap]		[ɐ.'tap̰]		[ɐ.'tab̰]
‘lua’		‘lua’		‘lua’

❖ Alofones de /p/ em variação livre:

➤ Fones [p], [p̰] e [b̰] variando sem contrastar significado.

Os fones [p] e [p̰] variam livremente em coda silábica, com uma tendência maior para a ocorrência de [p̰], principalmente em final de palavra. Há também uma espécie de distribuição complementar, já que [p] pode ocorrer em *onset* e em coda, mas [p̰] só ocorre em coda, sendo o

contexto de final de sílaba condicionante para a ausência da oclusão. O fone [ḃ] ocorre apenas no caso de [v. 'tab̃] ‘lua’ e *rawkap* [ɾew. 'kēḃ] ‘pacupemba’.

[t]	vs.	[ṭ]
[mɐ. 'ta]		[mɐṭ. 'kia]
‘minha mãe’		‘homem’
[tɐ. 'ka]		[v. 'kat̃]
‘mentir’		‘papagaio-de-cabeça-roxa’

❖ Alofones /t/ em distribuição complementar:

- Os fones [t] e [ṭ] variando sem contrastar significado;
- Ocorrem em ambientes diferentes: [t] apenas em *onset* e [ṭ] apenas em coda, sendo o contexto de final de sílaba condicionante para a ausência da oclusão.

O fonema /k/ ocorre apenas em ataque, não havendo nenhum tipo de variação livre ou distribuição complementar com uma contraparte não explodida. Sendo assim, ele não ocorre em posição de coda, diferentemente dos fonemas oclusivos /p/ e /t/.

4.1.2 Oclusiva glotal [ʔ]

[ʔ]		
[o. 'ʔo]	[o. 'ʔi]	[sɐ. 'ʔaw]
‘criança’	‘cunhado(a)’	‘inambu’ (ave)

- ❖ Trata-se de um recurso de marcação de fronteira, tanto no âmbito segmental (entre segmentos) quanto no âmbito suprasegmental (fronteira silábica). Ocorre na separação de hiatos (ver capítulo anterior).

4.1.3 Oclusivas velares com articulação secundária: /kʷ/ e /kʲ/

[kʷ]	vs.	[k]
[v. kʷa]		[v. 'ka]
‘batata-doce’ (espécie)		‘dente’

[v.'k ^w ap]	[v.'kat̃]
‘galho’	‘papagaio-de-cabeça-roxa’
[v.'k ^w a]	[v.r̃v.'ko]
‘batata-doce’ (espécie)	‘cará-roxo’
[.'k ^w a]	[ko.'aj]
‘nariz’	‘capivara’

❖ Fonemas distintos: /k^w/ e /k/

- Fones [k^w] e [k] ocorrendo em ambiente idêntico e análogo, contrastando significado.

[kⁱ]	vs.	[k]
[k ⁱ ṽ.'ta]		[k̃ṽ.'ta]
‘pênis’		‘ingá’ (espécie)
[o.'k ⁱ a]		[o.'ko]
‘pupunha’		‘zogue-zogue’ (macaco)
[o.'k ⁱ a]		[ko.'ka]
‘pupunha’		‘piolho’

❖ Fonemas distintos: /kⁱ/ e /k/

- Fones [kⁱ] e [k] ocorrendo em ambiente idêntico e análogo, contrastando significado.

[k^w]	vs.	[kⁱ]
[v.k ^w a]		[o.'k ⁱ a]
‘batata-doce’ (espécie)		‘pupunha’

❖ Fonemas distintos: /k^w/ e /kⁱ/

- Fones [k^w] e [kⁱ] ocorrendo em ambiente idêntico e contrastando significado.

[k^w]	vs.	[p]
[v.k ^w a]		[v.'pa]
‘batata-doce’ (espécie)		‘pai’ (vocativo e referencial)

❖ Fonemas distintos: /k^w/ e /p/

- Fones [k^w] e [p] ocorrendo em ambiente idêntico e contrastando significado.

Cabe notar que um importante argumento para análise /k^w/ e /kⁱ/ como fonemas co-articulados foneticamente vem de ordem fonotática, em especial a estrutura silábica, como será discutido na seção 4.5.

4.1.4 Nasais: /m/, /n/, /ɲ/

[m]	vs.	[n]
[mɐ.'mo]~[mɐ.'mo]		[mɐ.'nõ]
‘roça de maniva’		‘anta’
[ko.,mɐ.ʃi.'ri]		[ko.nɐ.'ʃi]
‘damorida’ (prato típico)		‘anzol’
[ko.ʃi.'nẽm]		[ko.ʃi.'pẽn]
‘Apagar, extinguir-se’		‘caranguejo’
[ɐ.tɐ.'mã]		[ko.rɐ.tɐ.'na]
‘ralar’		‘banana’

❖ Fonemas distintos: /m/ e /n/

- Fones [m] e [n] ocorrendo em ambiente análogo, contrastando significado.

[m]	vs.	[ɲ]
[ɐ.'ma]		[ɐ.'ɲa]
‘caminho’		‘folha’
[mɐ.'ma]		[mẽ.'ɲã] ²⁶
‘membro superior (braços, mãos e dedos)’		‘o quê’ (pronome indefinido)

❖ Fonemas distintos: /m/ e /ɲ/

²⁶ Trataremos da questão da nasalidade e de seu status fonêmico na seção 5.4.4.

- Fones [m] e [ɲ] ocorrendo em ambiente idêntico e ambiente análogo, contrastando significado.

[n]	vs.	[ɲ]
[ẽ.'ni]		[ẽ.'ɲa]
‘lenha’		‘folha’
[ẽ.'nẽj]		[ẽ.'ɲẽ]
‘mãe’ (vocativo e referencial)		‘folha’
[mẽ.nẽ.'ri]		[mẽ.'ɲã]
‘peneira’		‘o quê’ (pronome indefinido)

❖ Fonemas distintos: /n/ e /ɲ/

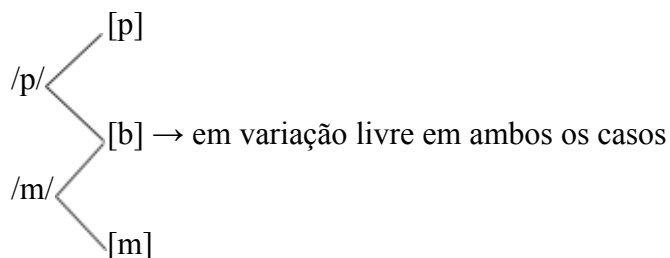
- Fones [n] e [ɲ] ocorrendo em ambientes idênticos e análogos, contrastando significado;
- [ɲ]: só ocorre diante de /a/ com os alofones [a], [ẽ], [ẽ] - majoritariamente em sílaba tônica (restrição fonotática).

4.1.5 [b] como alogone de /m/ e /p/

Esses dois fones são bilabiais, por isso são SFS e precisam ser investigados. Para entendermos esse contraste precisamos iniciar observando as seguintes palavras:

[ko.'a]	[mɔ.'ja]
‘nariz’	‘buraco’

A palavra para ‘narina’ ou, literalmente, ‘boraco do nariz’ seria, então, *koa muya* [ko.'a bo.'jæ] ‘narina’. Dessa forma, temos [m] e [b] variando entre si, sem nenhuma motivação ambiental clara. Podendo ser pronunciado das duas formas.



Temos, também, o caso de *tiponkin* [tɪ.põn.'kĩn] ‘brancos’ ou ‘não indígenas’ sendo pronunciada como, também, [tɪ.bõn.'kĩn]. Portanto, o fone [b] não é um fonema, o que ocorre, na

verdade, é um *overlapping* em que [b] pode ser tanto uma alofone de [p] quanto de [m] – o que o aproxima de [p] é o fato de terem o mesmo modo e ponto de articulação (oclusivo e bilabial) e de [m] pelo fato de terem o mesmo ponto de articulação e mesmo modo de fonação (bilabial e vozeado). Há uma exceção [b] em caso de empréstimo na palavra *sambrero* [ˌsẽm.bre.'ro] – que significa ‘chapéu’, neste caso, o [b] passa a ser um fone da língua, mas não um nome típico da língua, pois ocorre por conta de empréstimo. Não é um fone propriamente dito da língua neste contexto de *sambrero*, pois surge a partir de um empréstimo.

Além disso, existem casos em que a alternância entre [m] e [p] contrastam significado. O que nos prova que se trata de fonemas diferentes. Como ocorre em:

[m]	vs.	[p]
[ɐ.'ma]		[ɐ.'pa]
‘caminho’		‘pai’ (vocativo)
[mi.'sa]		[pi.'ʃa]
‘carne’		‘cuxiú’

❖ Fonemas distintos: /m/ e /p/

➤ Fones [m] e [p] ocorrendo em ambiente idêntico e análogo, contrastando significado.

4.1.6 Fricativas: /s/ vs. /ʃ/ vs. /h/

[s]	vs.	[ʃ]	vs.	[h]
[ko.'sa]		[ko.'ʃa]		[ko.'haj]
‘seio’		‘queimar’		‘velho’
[ko.si'pẽn]		[ko.ʃi'pẽn]		
‘cheio’		‘caranguejo’		
[i.'so]				[i.'ho]
‘matar com timbó’				‘panela’

❖ Fonemas distintos: /s/, /ʃ/ e /h/

➤ Fones [s], [ʃ] e [h] ocorrendo em ambientes idênticos e análogos, contrastando significado.

4.1.7 [s] e [ʃ] em variação livre:

[s]	vs.	[ʃ]
[v.'saw]		[v.'ʃaw]
‘formiga’		‘formiga’
[sv.'ko]		[ʃv.'ko]
‘tucano’		‘tucano’

O contraste entre [s] e [ʃ] é semelhante ao de [t] e [tʃ] em português. Em certos momentos eles contrastam significado, como observamos em *kosa* ‘seio’ e *koxa* ‘queimar’, assim como em português temos *tal* [taw] e *tchau* [tʃaw]; e *teco* [ˈtɛ.ko] e *tcheco* [ˈtʃɛ.ko], palavras em que [t] e [tʃ] contrastam significado. Enquanto, em palavras como *tia* [ˈti.ɐ] ~ [ˈtʃi.ɐ], não contrasta significado e variam dialetalmente pelo Brasil, enquanto [s] e [ʃ], no Arutani falado por Juanita, esses fones variam idioletalmente. Nos casos apresentados em que há variação livre entre [s] e [ʃ], os formatos de onda são bem parecidos, a distinção é feita de uma perspectiva oitiva, um contraste espectral através de uma análise FFT, no qual é possível compararmos o pico espectral, e o F2 de transição, características acústicas que servem para a diferenciação do ponto de articulação de segmentos fricativos. Esses fones representam o fonema /s/, pois ele acaba sendo o fone mais produtivo na língua e observamos uma tendência de Juanita estar pronunciando o fone [s] cada vez mais palatal.

4.1.8 Palatais: /j/ e /ɲ/

[j]	vs.	[ɲ]
[kɐ.'ja]~[kɐ.'jæ]		[v.'ɲa]
‘onça’		‘folha’
[mɔ.'ja]~[mɔj.'ja]		[mẽ.'ɲã]
‘buraco’		‘o quê’ (pronome indefinido)

❖ Fonemas distintos: /j/ e /ɲ/

- Fones [j] e [ɲ] ocorrendo em ambientes análogos e contrastando significado.

4.1.9 Glides: /w/ e /j/

[w]	vs.	[j]
[wɐ. 'ta]		[jɐ. 'ta]
‘acabar’		INTENSIFICADOR
[ẽw̃. 'hã]		[ẽj̃. 'hã]
‘viver’ ou ‘morar’		‘venha cá’ (verbo)
[ɐw. 'to]		[ɐj. 'ta]
‘peixe’ (espécie)		‘ali’ (locativo)
[wɐ. 'si]		[jɐ. 'sa]
‘esperar’		‘cipó’
[kɪ 'wa]		[kɐ 'ja]
‘língua’		‘onça’

❖ Fonemas distintos: /w/ e /j/

➤ Fones [w] e [j] ocorrendo em ambientes idênticos e análogos, contrastando significado.

4.1.10 Contraste entre /t/ vs. /s/ vs. /ʃ/

[t]	vs.	[ʃ]
[kɐ. 'ta]		[kɐ. 'ʃa]
‘Espécie de ingá’		‘Pedaços/peças de algum bicho ou coisa’ (verbo)
[o 'ti]		[o 'ʃi]
‘urubu’		‘sol’

❖ Fonemas distintos: /t/ e /ʃ/

➤ Fones [t] e [ʃ] ocorrendo em ambiente idêntico e contrastando significado.

[t]	vs.	[s]
[ɐ. 'ta]		[ɐ. 'sa]
‘mãe’ (referencial)		‘pé de árvore’

❖ Fonemas distintos: /t/ e /s/

- Fones [t] e [s] ocorrendo em ambiente idêntico e contrastando significado.

4.1.11 Outros contrastes:

[n]	vs.	[r]	vs.	[t]
[kɐ.'ni]~[kẽ.'ni]		[kɐ.'ri]		[kɐ.'ti]
‘cópula’		‘irmã mais nova’		‘irmã(o) mais novo(a)’
		‘interrogativo’		

❖ Fonemas distintos: /n/, /r/ e /t/

- Fones [n], [r] e [t] ocorrendo em ambiente idêntico e contrastando significado.

[n]	vs.	[t]
[e.'mẽn]		[e.'mẽt̃]
‘casa’		‘casa’

❖ Alofones de /t/ em variação livre:

- Fones [n] e [t] ocorrendo em ambiente idêntico e não contrastando significado;

Os fonemas /n/ e /t/ (alveolares) contrastam em *onset*, mas não em coda ou final de palavra. O ambiente nasal promove esta alofonia em final de palavra em que [t] > [n].

[p]	vs.	[n]
[v.'pi]		[v.'ni]
‘vespa’		‘fogo’
[v.'pap]		[v.'pẽn]
‘fraco’		‘quebrar’

❖ Fonemas distintos: /p/ e /n/

- Fones [p] e [n] ocorrendo em ambiente idêntico e em ambiente análogo, contrastando significado.

[p]	vs.	[n]
[v.'pa]		[v.'na]

‘pai’ (vocativo)

‘folha’

❖ Fonemas distintos: /p/ e /ɲ/

➤ Fones [p] e [ɲ] ocorrendo em ambiente idêntico e contrastando significado.

[k]

vs.

[s]

[ko. 'ka]

[ko. 'sa]

‘piolho’

‘seio’ ‘clado, líquido’

❖ Fonemas distintos: /k/ e /s/

➤ Fones [k] e [s] ocorrendo em ambiente idêntico e contrastando significado.

[k]

vs.

[n]

vs.

[p]

[kɪ. 'ka]

[kɪ. 'na]

[kɪ. 'pa]

‘passado’

‘pessoa’

‘urina’

❖ Fonemas distintos: /k/, /n/ e /p/

➤ Fones [k], [n] e [p] ocorrendo em ambiente idêntico e contrastando significado.

[k]

vs.

[j]

[ʃɐ. 'ko]

[ʃɐ. 'jo]

‘tucano’

‘fazer’

[ʃɐ.ɾɪ. 'ko]

[ʃɐ.ɾɪ. 'jo]

‘mosca’ (espécie)

‘girino’

❖ Fonemas distintos: /k/ e /j/

➤ Fones [k] e [j] ocorrendo em ambiente idêntico e contrastando significado.

[m]

vs.

[ɾ]

vs.

[k]

[ɪ. 'ma]

[ɪ. 'ra]

[ɪ. 'ka]

‘interior de algo’

‘bacaba’

‘jacaré’

❖ Fonemas distintos: /m/, /ɾ/ e /k/

- Fones [m], [r] e [k] ocorrendo em ambiente idêntico e contrastando significado.

[m]	vs.	[h]	vs.	[s]
[i. 'mo]		[i. 'ho]		[i. 'so]
‘pajé’		‘panela’		‘matar com timbó’

❖ Fonemas distintos: /m/, /h/ e /s/

- Fones [m], [h] e [s] ocorrendo em ambiente idêntico e contrastando significado.

[r]	vs.	[h]
[rɐ 'ta]		[hɐ 'ta]
‘falar’, ‘dizer’ ou ‘contar’		‘esse’ ou ‘aquele’

❖ Fonemas distintos: /r/ e /h/

- Fones [r] e [h] ocorrendo em ambiente idêntico e contrastando significado.

4.2 Vogais

Nesta seção, iremos apresentar os fonemas vocálicos existentes na fala de Juanita, sendo eles: /a/, /i/ e /o/. Para chegarmos em tal resultado, contrastamos os fones vocálicos ocorrendo em ambiente idêntico e/ou análogo. A partir disso, será possível observarmos os fonemas vocálicos e seus alofones. Nossa análise iniciará com os fones que ocorrem em ambiente tônico, em seguida, partiremos para os fones que ocorrem em ambiente átono e, por fim, apresentaremos o contraste entre fones que ocorrem apenas em contexto tônico e fones que ocorrem apenas em contexto átono para observarmos se há ou não distribuição complementar entre eles.

4.2.1 Contraste das vogais em ambiente tônico:

[a]	vs.	[i]
[o. 'fa]		[o. 'fĩ]
‘pelo’, ‘cabelo’		‘sol’
[i. 'ka]		[i. 'ki]

‘jacaré’	‘crescer’
[kɐ.ˈta]	[kɐ.ˈti]
‘ingá’ (espécie)	‘irmã(o) mais novo(a)’
[ɐ.ˈpa]	[ɐ.ˈpi]
‘pai’ (vocativo)	‘vespa’
[ki.ˈna]	[kɐ.ˈni]
‘pessoa’	‘cópula’

❖ Fonemas distintos: /a/ e /i/

- Alofones tônicos dos fonemas /a/ e /i/, respectivamente, contrastando significado em ambiente idêntico e em ambiente análogo.

[a]	vs.	[æ]²⁷
[ɐw.ɐ.ˈja]		[kɐ.ˈjæ]
‘jibóia’		‘onça’
[kɐ.ɾi.ˈja]		[mɐ.ɾi.ˈjæ]
‘irmã mais nova’		‘faca’
[ʃi.pɔ.ˈja]		[ko.ˈa bu.ˈjæ]
‘pó’		‘narina’
[o.ˈja]		[ki.ro.ˈjæ]
‘apimentado’		‘cobra’ (termo genérico)

❖ Alofones de /a/ em variação livre.

O fone [æ] ocorre apenas após consoantes palatais ou palatalizadas, porém, trata-se de uma tendência e não de uma distribuição complementar, pois /a/ diante de /j/ tende a se transformar em [æ], mas ainda há casos em que permanece como [a]. Esta tendência trata-se de um **alçamento vocálico**, um tipo de assimilação, na qual a vogal [a] tende a assimilar um traço palatal em comum, que, articulatoriamente, tem relação com o ponto de articulação palatal ou, da perspectiva da vogal, com a altura da língua, ou seja, quando a língua está em uma posição mais

²⁷ Não apresentamos o fone [æ] na análise acústica separadamente, pois ele ocorre majoritariamente na sílaba [jæ] e, como já apontamos, não é possível fazer a segmentação de ditongos, pois não é clara a fronteira de término e início dos segmentos.

próxima do palato. Esse fenômeno só ocorre em sílaba tônica, sempre em final de palavra. O mesmo não ocorre com ditongos decrescentes como [aj].

[a]	vs.	[o]
['ʃa]		['ʃo]
‘peidar’		‘vá’
[ɐ. 'ka]		[ɐ. 'ko]
‘dente’		‘cocho para caxiri’ ou ‘massa de mandioca’
[mɐ. 'ma]		[mɐ. 'mo]~[mɐ. 'mɔ]
‘membro superior’ (braços, mãos e dedos)		‘roça de maniva’
[ɪ. 'ma]		[ɪ. 'mo]~[ɪ. 'mɔ]
‘interior de algo’		‘pajé’
[ko.rɐ. 'ka]~[kʊ.rɐ. 'ka]		[korɐ 'to]
‘cuatá’		‘jacaré-açú’

❖ Fonemas distintos: /a/ e /o/

- Alofoneônico do fonema /a/ e alofone de /o/, respectivamente, contrastando significado em ambiente idêntico e em ambiente análogo.

[a]	vs.	[o]
[kɐ. 'ja]~[kɐ. 'jæ]		[kɐ. 'jo]
‘onça’		‘tamanduá-bandeira.’
[mɐ. 'ma]		[mɐ. 'mo]~[mɐ. 'mɔ]
‘membro superior’ (braços, mãos e dedos)		‘roça de maniva’

❖ Fonemas distintos:

- Alofoneônico do fonema /a/ e alofone de /o/, respectivamente, contrastando significado em ambiente idêntico.

[i]	vs.	[o]
[kɐ. 'ri]		[kɐ. 'ro]

‘irmã mais nova’ ou ‘interrogativo’	‘mulher’
[tɪ. 'ti]	[to 'to]
‘também’	‘irmão mais velho’
[oʔ. 'i]	[oʔ. 'o]
‘cunhado(a)’	‘bebê’
[mɐ. 'ni]~[mɐ. 'nĩ] ²⁸	[mɐ. 'nõ]
‘sapo’	‘anta’

❖ Fonemas distintos:

- Alofone tônico do fonema /i/ e alofone de /o/, respectivamente, contrastando significado em ambiente idêntico.

[i]	vs.	[o]
[ko.jo.'pi]~[kɔ.jo.'pi]		[kɔ.jo.'ro]
‘besouro’ ou ‘muchiva’		‘tucumã’

❖ Fonemas distintos:

- Alofone tônico do fonema /i/ e alofone de /o/, respectivamente, contrastando significado em ambiente idêntico.
- Ressaltamos a baixa ocorrência [ɔ] e a difícil existência de pares mínimos.

[o]	vs.	[ɔ]
[ɪ. 'mo]		[ɪ. 'mɔ]
‘pajé’		‘pajé’
[mɐ. 'mo]		[mɐ. 'mɔ]
‘roça de maniva’		‘roça de maniva’
[ɐ.rɐ. 'ko]		[ɐ.rɐ. 'kɔ]
‘cará-roxo’		‘cará-roxo’

❖ Alofones em variação livre:

- Fonema /o/ que varia entre [o] e [ɔ] sem contraste.

²⁸ Trataremos da questão da nasalidade e de seu *status* fonêmico na seção 4.4.

Escolhemos /o/ como representação do fonema pelo fato de [o] ser mais produtivo na língua, ocorrendo diversas vezes, enquanto [u], em posição tônica, ocorre apenas seis vezes, como, por exemplo, nas palavras acima em: *yapururu* [jɐ.pu.ru.'ru] ‘banana-comprida’, *kuyuru* [ku.ju.'ru] ‘tucumã’ e *wasuu* [wɐ.si.'u] ‘guariba’ (macaco).

4.2.2 Contraste das vogais em ambiente átono:

[ɐ]	vs.	[ɪ]
[ɐ.'ka]		[ɪ.'ka]
‘dente’		‘jacaré’ ou ‘fígado’ (homônimas)
[ɐ.rɐ.'ko]		[ɐ.ɾɪ.'ko]
‘cará roxo’		‘mosca’ (espécie)
[pɐ.'sa]		[pɪ.'ʃa]
‘jundiá’		‘cuxiú’

❖ Fonemas distintos:

- Alofones átonos dos fonemas /a/ e /i/, respectivamente, contrastando significado em ambiente idêntico e em ambiente análogo.

[ɐ]	vs.	[o]
[ɐ.'ko]		[o.'ko]
‘cocho para caxiri’ ou ‘massa de mandioca.’		‘zogue-zogue’

❖ Fonemas distintos:

- Alofone átono do fonema /a/ e alofone do fonema /o/, respectivamente, contrastando significado em ambiente idêntico.

[ɐ]	vs.	[ʊ]
[kɐ.'ju]		[ku.'ju]
‘tamanduá-bandeira’		‘grande’
[ɐ.'ka]		[ʊ.'ka]
‘dente’		clítico de caso instrumental e ergativo

❖ Fonemas distintos:

- Alofone átono do fonema /a/ e alofone do fonema /o/, respectivamente, contrastando significado em ambiente idêntico.

[ɪ]	vs.	[o]
[ɪ.'so]		[o.'so]
‘matar com timbó’		‘irmão’ (termo referencial)
[ɪ.ko.'ʃa]		[o.ko.'ka]
‘parte superior da coluna vertebral’		‘cozinhar’

❖ Fonemas distintos:

- Alofone átono do fonema /i/ e alofone do fonema /o/, respectivamente, contrastando significado em ambiente idêntico e em ambiente análogo.

[ɪ]	vs.	[ʊ]
[ɪ.'ka]		[ʊ.'ka]
‘jacaré’		clítico de caso instrumental e ergativo

❖ Fonemas distintos:

- Alofone átono do fonema /i/ e alofone do fonema /o/, respectivamente, contrastando significado em ambiente idêntico.

[o]	vs.	[ʊ]
[oj.'ja]~[oj.'jæ]		[ʊj.'ja]~[ʊj.'jæ]
‘apimentado’		‘apimentado’
[mɐ.mo.'sa]		[mɐ.mʊ.'sa]
‘maniva’		‘maniva’
[ko.rɐ.'ka]		[kʊ.rɐ.'ka]
‘cuatá’		‘cuatá’
[ko.jʊ.'pi]		[kʊ.jʊ.'pi]
‘besouro’ ou ‘muchiva’		‘besouro’ ou ‘muchiva’

❖ Alofones em variação livre:

- Fonema /o/ que varia entre [o] e [ʊ] sem contraste.

4.2.2.1 Sobre o estatuto fonêmico de [e] e /i/:

Quanto ao fone [e], não foi possível contrastá-lo com nenhum outro fone, por conta da sua baixa ocorrência e, conseqüentemente, a difícil existência de pares mínimos ou análogos. Dessa forma, sincronicamente, ou seja, na fala de Juanita, não fica claro o seu *status*.

Para averiguar seu possível *status* na língua, seja de fonema ou de alofone, resolvemos expandir o corpus e observar outros registros do Arutani. O primeiro doculeto está registrado em Armellada e Matallana (1942) e o segundo em Migliazza (1978). Para tal comparação, elaboramos a Tabela 36:

Tabela 36 – Comparação de [e] vs. /i/

Arutani (Juanita)	Arutani (Armellada e Matallana, 1942)	Uruak (Migliazza, 1978)	Tradução
<i>maikáti</i> [mɛj. 'ka.tɪ]	<i>akiramat</i>	<i>maykate/ ma- / tsa-</i>	‘eu’
<i>kaikáti</i> [kɛj. 'ka.tɪ]	-	<i>kaykate/ ka-</i>	‘você’
<i>kakwati</i> [kɛ.kʷɛ. 'ti]	<i>kikuati-usa</i>	<i>kwate</i>	‘cabeça’
<i>xati</i> [ʃɛ. 'ti]	<i>achati</i>	<i>xate</i>	‘pé’
<i>kirakoti</i> [kɪ. rɛ.ko. 'ti]	<i>kirajute</i>	<i>kirakote</i>	‘coração’
<i>tosin</i> [to.sɪn]	-	<i>tose</i>	‘noite’
<i>tasiri</i> [tɛ. 'si.ɾɪ]	-	<i>taseri</i>	‘bom’
<i>misa</i> [mɪ. 'sa]	<i>imesa</i>	-	‘carne’
<i>karo simhõ</i> [kɛ. ro.sɪm. 'hõ]	<i>semhu</i>	-	‘vagina’
<i>emet</i> [e. 'mɛt̃]	<i>imat</i>	-	‘casa’

Podemos observar que as palavras correspondentes para ‘eu’ e ‘você’ na fala de Juanita são, respectivamente, *maikáti* [mɛj. 'ka.tɪ] e *kaikáti* [kɛj. 'ka.tɪ], ambas terminadas em [ɪ] átono. Quando comparamos com suas correspondentes em Migliazza (1978), temos *maykate* e *kaykate*, ambas registradas pelo autor com <e> onde encontramos [ɪ] na fala de Juanita. Ao contrastarmos os três registros – o Arutani falado por Juanita; o Arutani registrado por Armellada e Matallana (1942); e o Uruak registrado por Migliazza (1978) – e seus respectivos correspondentes para

‘cabeça’, temos, respectivamente, as seguintes ocorrências e variações *kakwati* [kɐ.kʷɐ.'ti] ocorrendo com [i], *kikuati-usa*, sendo grafada com <i> e *kwate*, grafada com <e>. Já para a palavra para ‘coração’, encontramos, respectivamente, *kirakoti* [kɪ.ɾɐ.ko.'ti], *kirajute* e *kirakote*. Neste contexto, temos apenas na fala de Juanita ocorrendo [i] (tônico), enquanto nos dois doculetos há a ocorrência de <e>. Em todos os outros casos seguintes na fala de Juanita encontramos /i/, enquanto nos outros registros estará ocorrendo <e>, exceto em *emet* [e.'mẽt̃] ‘casa’, em que encontramos [e] e [ẽ] no idioleto de Juanita e no registro de Armellada e Matallana (1942) está ocorrendo <i> e <a> em *imat*.

Sendo assim, fica evidente que nos casos em que os doculetos tendem a ocorrer <e>, que inferimos tratar-se de um possível fone [e] com base na ortografia, na fala de Juanita está sendo produzido como /i/. O caso de *emet* vs. *imat*, em que esta tendência é invertida, nos dá mais pistas quanto ao possível caráter alofônico de [e].

A partir disso, percebemos que se trata de um possível alofone de /i/ que foi sendo cada vez menos utilizado nas variedades linguísticas analisadas, pois ele também é extremamente raro nos doculetos e na fala dos Arutani e dos Uruak nas décadas de 1940 e 1970, mas que ainda está presente na fala de Juanita em contextos átonos. Esta restrição de uso/ocorrência em apenas contextos átonos, juntamente com a sua baixa ocorrência e a variação de uso/ocorrência entre os diferentes registros das variedades linguísticas reforçam ainda mais seu *status* de alofone.

Na fala de Juanita e nos dois doculetos o uso de /i/ ou <i> é muito mais recorrente. Assim sendo, podemos dizer que se trata de um alofone de um fonema /i/ em variação livre, variando cronologicamente e entre os diferentes dialetos (fala de Juanita e os dois doculetos/variedades da língua que foram analisados).

4.2.3 Contraste das vogais em ambienteônico vs. átono

[a]	vs.	[ɐ]
[sɐp.'sa]		[ɐ.ɾɐ.'sa]
‘rato’		‘pica-pau’ (espécie)
[ɐ.ɾɐ.'wa]		[ɐ.'ka ɐ.nɐ.'ma]
‘traíra’		‘dentes molares e caninos’

❖ Alofones em distribuição complementar:

- Fonema /a/ com [a] sempre ocorrendo em posição tônica e [ɐ] em posição átona.

[i]	vs.	[ɪ]
[ʃi.'ri ʃi.'ri]		[ni.kɪ.'ʃi]
‘araguaí’		‘tipiti’
[ɪ.'ki]		[ɪ.mi.'ki]
‘crescer’		‘nome’

❖ Alofones em distribuição complementar:

- Fonema /i/ com [i] sempre ocorrendo em posição tônica e [ɪ] em posição átona.

4.3 Segmentos ambivalentes

Quando falamos de segmentos ambivalentes estamos nos referindo aos *glides*, também conhecidos como aproximantes, sendo eles [j] e [w]. Estes segmentos são conhecidos como ambivalentes, pois, nas línguas do mundo, variam entre o *status* de consoante ou o de vogal. Ademais, de uma perspectiva fonética, principalmente articulatória, esses fones possuem traços que os aproximam mais às vogais, enquanto, de uma perspectiva fonológica, eles acabam tendo características que os aproximam mais das consoantes. Santos e Souza (2010, p. 24) explicam que as aproximantes, do ponto de vista fonético, permitem a passagem do ar de forma livre, sem bloqueios ou fricção, com ressonância no centro do trato vocal e com um espaço de passagem do ar mais estreito do que nas vogais correspondentes. Fonologicamente, segundo os mesmo autores, as aproximantes funcionam como consoantes, pois não ocupam o núcleo da sílaba nem recebem acento.

Na fala de Juanita, encontramos os *glides* /j/ [j] e /w/ [w] ocorrendo sempre acompanhados por uma vogal, como apresentado na seção 3.2. Os encontros/sílabas GV presentes na fala de Juanita são os seguintes: [ja], [jɐ] ~ [jẽ], [jæ], [jo], [ju], [wa], [wɐ] ~ [wẽ] e [wi]. Também encontramos sílabas VG como: [aj] ~ [ãj], [ɛj], [aw], [ɐw] e [ɔj].

Para tratarmos sobre os *glides* na língua, iremos adiantar um pouco algumas questões da fonotática do Arutani, mas as retomaremos e traremos um aprofundamento maior sobre elas na seção 3.2 Em Arutani, encontraremos palavras que possuem apenas uma vogal (V) ocupando o núcleo silábico, como em *oxi* [o.'ʃi] ‘sol’ e *imo* [ɪ.'mo] ‘pajé’, por exemplo. Não havendo encontros consonantais, CC. Porém, é muito comum encontrarmos sílabas GV, VG e CVG. Também há um caso de GVC. Observe alguns exemplos a seguir:

(6) **GV:**

parawa [pɐ.rɐ.'wa] ‘rio paragua’ *kiroya* [ki.ro.'jæ] ‘cobra’ (termo genérico)
watika [wɐ.tɪ.'ka] ‘orelha’ *makayu* [mɐ.kɐ.'jo] ‘meu avô’ ou ‘meu sogro’

(7) **VG:**

oaw [o.'aw] ‘criança’ *awxihi* [ɐw.ʃi.'hi] ‘ubim’
ayta [ɐj.'ta] ‘ali’ *koay* [kɔ.'aj] ‘capivara’

(8) **CVG:**

tahay [tɐ.'haj] ‘peixe’ (espécie da família do mandi) *xipiray* [ʃip̃.'raj] ‘pacu’
marakupaw [mɐ.,rɐ.ko.'paw] ‘rã’ (espécie comestível) *asaw* [ɐ.'saw] ‘formiga’
asaraw [ɐ.sɐ.'raw] ‘molhado’

(9) **GVC:**

kaya rawin [kɐ.'ja rɐ.'win] ‘sussuarana’

Algumas estruturas não são encontradas na fala de Juanita por conta das oclusivas velares labializadas e palatalizadas [k^w] e [kʲ], as quais representam apenas um fone e apenas um fonema. Portanto, estruturas como CGV, CGVG e CGVC, não são encontradas, por conta de os *glides* só ocorrerem diante de consoantes quando estão em contexto de articulação secundária em [k^w] ou [kʲ]. Dessa forma, encontros que poderiam ser analisados como CG, passam a ser analisadas apenas como C. Observamos isso nos casos de palavras como: CGV > CV: *kwaso* [k^wɐ.'so] ‘joaninha’; CGVG > CVG: *koay* ã [k^waj 'ẽ] ‘buriti’; e CGVC > CVC: *akwap* [ɐ.'k^wap] ‘galho’, *kwakhaw* [k^wɛk.'haw] ‘cabaça’ e *xokwan* [ʃo.'k^wɛn] ‘cupim’. Com isso, estruturas como [k^wap] e [k^wɛk], passam a ser analisadas como CVC, assim como a sílaba [tap] em *atap* [ɐ.'tap] ‘lua’.

Com isso, podemos dizer que, em Arutani, os *glides* ocorrem sempre em ataque ou em coda silábica, acompanhando uma vogal. Esse tipo de característica é típica das consoantes da língua, assemelhando-se a consoantes como /p/, /t/, /k/, /m/ e /n/, que também podem ocupar tanto o ataque quanto a coda silábica. Portanto, por não estarem dentro do núcleo silábico, por não receberem acento e por ocuparem posições iguais a consoantes prototípicas da língua que podemos classificar os *glides* como consoantes em Arutani, sendo representados, a partir de agora, como C na estrutura silábica.

Ademais, outro ponto que reforça o *status* dos *glides* /j/ e /w/ como consoantes na língua é que, com isso, o sistema fonotático da língua fica mais sintético e conciso. Isso ocorre, pois há a diminuição e a simplificação das possibilidades de estruturas silábicas, permitindo um modelo (C)V(C), no qual não há a possibilidade para encontros consonantais e nem encontros vocálicos na mesma sílaba. Além disso, impede a existência de encontros vocálicos, os quais seriam mais uma estrutura que complexificaria a fonotática do Arutani.

4.4 Nasalidade Vocálica

A nasalidade já foi introduzida na seção de acústica sobre as vogais nasais quando apresentamos suas características acústicas e os contextos em que elas podem ocorrer. Sendo assim, retomaremos essas questões e apresentaremos os mesmos exemplos e contextos, mas agora de uma perspectiva fonológica. Nesta seção, também pretendemos deixar clara a diferença entre vogais nasais e nasalizadas. Por fim, tentaremos hipotetizar sobre o *status* fonológico da nasalidade na língua, para isso, apresentaremos questões sobre irregularidade desse fenômeno na fala de Juanita e faremos alguns contrastes entre pares em ambientes idênticos.

Na seção 3.1.5, apresentamos os contextos em que ocorre nasalização na fala de Juanita. São eles:

- I. Em contexto de vogais nasalizadas (Ñ) quando há consoante nasal (N) em ataque silábico (.NÑ.);

A. *poromã* [po.ro.'mẽ] ‘cucura’

- II. Em contexto de \tilde{V} com N posposta, seja na mesma sílaba em coda ou ataque silábico na sílaba seguinte (.C \tilde{V} N. / .C \tilde{V} .N);
- A. *ariran* [ɐ.ɾɪ.'**rẽn**] 'litrado'
- B. *mapaña* [mɐ.**pẽ**.'**ja**] 'mamão'
- III. Contexto de rinoglotalia em que uma \tilde{V} tende a ocorrer quando anteceder ou precedida por [h] em ataque silábico, ou seja, na mesma sílaba ou na sílaba seguinte (.h \tilde{V} ./C \tilde{V} .h);
- A. *atakhã* [ɐ.tɛk.'**hẽ**] 'irmã mais velha'
- B. *akohã* [ɐ.kõ.'**hẽ**] 'água'
- IV. Contexto de vogais nasais que ocorrem em contextos totalmente orais.
- A. *wasi sã* [wɐ.sɪ.'**sẽ**] 'alma-de-gato'
- B. *ira ã* [ɪ.'ra.'**ẽ**] 'patauá'
- C. *xapi ã* [ʃɐ.'pi.'**ẽ**] 'fruta' (termo genérico)

Observamos em I, II e III vogais nasalizadas, ou seja, vogais que sofrem o processo de nasalização por conta do ambiente em que estão. Nesses três casos, o contexto propício para a nasalização está relacionado com a influência das consoantes /m/, /n/, /ɲ/ e /h/ que estão adjacentes às vogais nasalizadas, anteceder ou precedendo as vogais, seja na mesma sílaba ou na seguinte. Portanto, fica evidente que estamos tratando de um processo de nasalização por assimilação, no qual as vogais assimilam o traço nasal das consoantes, com exceção da consoante glotal /h/, já que ela na verdade induz a nasalização por um processo de tendência articulatória, o que não deixa de ser um tipo processo de assimilação de traço articulatorio, mas não se trata da assimilação do traço nasal.

Em IV, encontramos vogais nasais, grupo diferente das vogais nasalizadas, pois não têm nenhum tipo de assimilação por influência do ambiente em que ocorrem, já que estão inseridas em um contexto totalmente oral. No caso de *wasi sã* [wɐ.sɪ.'**sẽ**] 'alma-de-gato', podemos hipotetizar um possível apagamento da consoante [n] por conta do contexto de final de palavra, mas como trata-se de uma palavra pouco recorrente na língua (nome de uma ave), não temos outra ocorrência dela para comprovarmos tal hipótese. Portanto, vale ressaltar que o fenômeno

de uma vogal nasal em contexto oral como o de *wasi sã* [wɐ.sɪ.'sẽ] é extremamente raro na língua.

Entretanto, temos os casos de *ira ã* [ɪ.'ra. 'ẽ] ‘patauá’ e *xapi ã* [ʃɐ.'pi 'ẽ] ‘fruta’ (termo genérico), em que temos na verdade um morfema *ã* [ẽ]. Sua função e sua classificação ainda não estão muito claras em nossa análise, mas podemos dizer, até o atual momento, que esse morfema parece ser produtivo dentro no campo lexical botânico, tendo a função de algo como um classificador. Portanto, de uma perspectiva mais geral, podemos interpretá-lo como sendo um classificador de seres inanimados, sendo seu uso frequente no léxico botânico. Outros exemplos da ocorrência de *ã* [ẽ] que encontramos são: *akôm ã* [ɐ.'kôm 'ẽ] ‘jatobá’ e *koay ã* ['kʷaj 'ẽ] ‘buriti’. Com isso, temos, além de um termo genérico para fruta, dois nomes de árvores/plantas, todos fazendo parte do campo lexical botânico e referindo-se a seres inanimados.

Outro ponto que também já foi introduzido na seção 3.1.5 foi a questão da irregularidade da nasalização. O que temos no Arutani falado por Juanita, na verdade, é uma tendência à assimilação da nasalidade, tendo como catalisadores para tal fenômeno os ambientes/contextos I, II e III. Trata-se de uma tendência pelo fato de que existem algumas exceções, apesar de raras, como é o caso de *imo* [ɪ.'mo] ‘pajé’ em que nenhuma das vogais adjacente ao [m] são nasalizadas. Também temos os outros exemplos apresentados na seção 3.1.5 em que há contextos propícios para a nasalização, mas que as vogais permanecem orais, como em: *anhĩ* [ɛn.'hĩ] ‘nádega’, *konihã* [ko.nɪ.'hẽ] ‘camarão’ e *kakohap* [kɛ.ko.'hap̃] ‘caroço’. Nesses quatro casos temos contextos em que há N adjacentes às vogais e elas não assimilam o traço nasal, o mesmo ocorre com o ambiente de rinoglotofilia.

No dicionário de Arutani (Chacon *et al.*, 2020) há um caso suspeito de contraste fonêmico entre uma vogal oral e nasal com os verbos *ha* [ha] ‘fluir’ e *hã* [hã] ‘levar’. O verbo *hã* [hã] ‘levar’ é encontrado dentro da seguinte frase: *kanawa hãhi kani akohãtama* ‘eles estão levando a canoa para o rio’. Vê-se que a vogal nasal está sendo, possivelmente, nasalizada por estar em um contexto muito propício para a nasalização, que é o de rinoglotofilia, já que temos duas consoantes glotais [h] ocorrendo à sua esquerda e à sua direita. Além disso, quando tratamos da semântica dos dois verbos, *ha* [ha] ‘fluir’ e *hã* [hã] ‘levar’, observamos que têm em

comum o sentido de ‘deslocamento’. Portanto, podemos suspeitar de seja um mesmo caso de variação fonética nasal e oral, e que se trata de um mesmo verbo que possui uma variação semântica a depender do contexto em que ocorre, podendo ser nasalizado a depender do ambiente fonético-fonológico.

Com isso, podemos interpretar a nasalidade vocálica com um fenômeno que tende a ocorrer na fala de Juanita, mas que não chega a ter um *status* fonêmico, ou seja, ela não contrasta significado. Ressaltamos que até mesmo em casos de vogais tidas nasais não há o contraste com um par mínimo oral. As comparações a seguir nos mostram a não existência de contraste entre vogal oral e vogal nasal em ambientes idênticos, sempre havendo uma assimilação.

(10)	[a]	vs.	[ẽ]
	[ko.'sa] ‘seio’		[ko.'sẽn] ‘cavar’
	[ko.'ʃa] ‘queimar’		[ko.'ʃẽn] ‘não maduro’
	[ɐ.'ma] ‘caminho’		[ɐ.'mẽn] ‘suave’

Os exemplos de contraste entre [a] e [ẽ] em posição tônica não foram escolhidos por acaso, pois eles são exemplos do único caso de alofonia por conta de nasalidade na língua, não ocorrendo com as demais vogais, seja em contextoônico ou átono. A tendência na fala de Juanita é de que para que uma vogal seja nasalizada precisa haver algum dos motivadores de nasalidade da língua, ocorrendo, nesses casos, sempre por assimilação.

Na fala de Juanita ocorre com o fonema /a/ uma assimilação da nasalidade em contextoônico. O processo é o seguinte: [a]ônico tende a ser produzido como [ẽ] ao ser nasalizado. Há casos em que [a] passa a ser produzido como [ã], mas são muito raros, por isso postulamos esta alofonia como uma tendência. Casos o de *koay ã* ['kʷaj 'ẽ] ‘buriti’, reforçam tal proposta, pois nesta palavra podemos observar o morfema ã [ẽ] que varia de [ẽ] (mais regular e mais produtivo) para [ã] (menos regular e menos produtivo). Com isso, temos que o fonema /a/ é

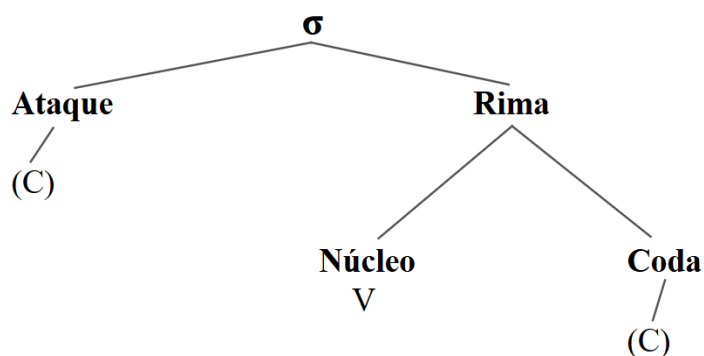
produzido como [a] em posição tônica e como [ɐ] em posição átona – esse processo é regular na língua e já foi descrito anteriormente – e que /a/ tônico nasalizado tende a ser produzido como [ẽ], mas pode ocorrer também como [ã]. Já [ɐ], por ser átono, ao nasalizar-se passa apenas a ser produzido como [ẽ].

Quanto às outras vogais, elas seguem a lógica da língua, uma tendência a se nasalizarem se estiverem inseridas nos contextos I, II e/ou III. Portanto:

- /i/ quando produzido como seu alofone [e] tende a [e] > [ẽ];
- /i/ tende a [i] (tônico) > [ĩ] (tônico);
- /i/ tende a [ɪ] (átono) > [ĩ] (átono);
- /o/ (átono e tônico) tende a [o] > [õ] ~ [ʊ] > [ũ].

4.5 Sílabas e Fonotática

Em Arutani o modelo silábico da língua é (C)V(C), no qual V é observado ocorrendo como uma única vogal. Com isso, observe o modelo de árvore com o molde silábico da língua:



A partir desse modelo podemos listar as estruturas silábicas possíveis e palavras em que elas ocorrem. Em sílabas compostas por apenas por uma vogal (V), por exemplo, temos a palavra *ata* [ɐ.'ta] ‘irmã mais velha’ (vocativo) e *akwa* [ɐ.'kʷa] ‘batata-doce’ (espécie). Essas palavras também ilustram sílabas CV, bem como a palavra *pori* [po.'ri] ‘mapinguari’ ou ‘espírito’. Sílabas como VC, ocorrem em palavras como *koay* [kɔ.'aj] ‘capivara’, *amhi* [am.'hĩ:] ‘trovão’. Por fim, temos a estrutura silábica mais complexa presente na fonologia do Arutani falado por

Juanita, que é a CVC. Sílabas CVC são encontradas em palavras como *atap* [v.'**ta**p] ‘lua’, *arihan* [v.r.ɪ.**hē**n] ‘doce’ e xipiray [ʃɪp̃.**ra**j] ‘pacu’.

A partir desse esquema e desses exemplos, apresentaremos os fonemas que ocorrem nas posições de ataque e coda, no âmbito da palavra e da sílaba, e, por fim, a distribuição e organização das vogais, que representam o núcleo silábico, na língua.

4.5.1 Ataque e coda

O fonema /p/ ocorre em ataque e em coda silábica, sendo que na posição de coda ocorre a variação livre entre [p] e [p̃]. Com isso, podemos explicar a fonotática do /p/ a partir da seguinte regra:

$$(11) \quad /p/ \rightarrow [p] / \sigma _ V \\ [p] \sim [p̃] / (C)V _ \sigma$$

Entretanto, ocorre algo diferente com o fonema /t/, pois [t] e [t̃] são alofones, mas em distribuição complementar, na qual [t] só ocorre em ataque e [t̃] só ocorre em coda, possuindo, assim, a seguinte regra:

$$(12) \quad /t/ \rightarrow [t] / \sigma _ V \\ [t̃] / (C)V _ \sigma$$

Além disso, /t/ possui uma peculiaridade, já que em ambiente nasal, ele pode variar com [n] em contexto de coda ([t] ~ [n] / (C) _ \sigma).

Esta mesma regra se aplica aos fonemas nasais /m/ e /n/ e aos *glides* /w/ e /j/, ambos ocorrem em ataque e em coda. Podemos representar os ambientes em que eles ocorrem com as regras:

$$(14) \quad \begin{array}{ll} \text{a.} & /m/ \rightarrow [m] / \sigma _ V _ \sigma \\ \text{b.} & /n/ \rightarrow [n] / \sigma _ V _ \sigma \\ \text{c.} & /w/ \rightarrow [w] / \sigma _ V _ \sigma \end{array}$$

$$d. \quad /j/ \rightarrow [j] / \sigma _ V _ \sigma$$

Quanto aos fonemas /r/, /s/, /ʃ/, /h/, /ɲ/, /k/, /k^w/ e /k^j/ todos ocorrem sempre em ataque silábico. A ocorrência de cada um deles dentro da estrutura silábica pode ser representada da seguintes maneira:

$$(15) \quad a. \quad /r/ \rightarrow [r] / \sigma _ V$$

$$b. \quad /s/ \rightarrow [s] / \sigma _ V$$

$$c. \quad /ʃ/ \rightarrow [ʃ] / \sigma _ V$$

$$d. \quad /h/ \rightarrow [h] / \sigma _ V$$

$$e. \quad /ɲ/ \rightarrow [ɲ] / \sigma _ V^{29}$$

$$f. \quad /k/ \rightarrow [k] / \sigma _ V$$

$$g. \quad /k^w/ \rightarrow [k^w] / \sigma _ V$$

$$h. \quad /k^j/ \rightarrow [k^j] / \sigma _ V$$

Cabe observar que os fonemas /k^w/ e /k^j/ foram propostos em consonância com nossa teoria silábica. Seguindo nossa proposta, os ditongos crescentes na língua são resultado da sequência entre um *glide* no ataque e uma vogal no núcleo, mas não encontramos ditongos crescentes quando o ataque está ocupado por uma consoante (*C+*glide*+V), salvo nas sequências com oclusiva velar (e.g. [kwa], [kja]). Logo o estatuto fonêmico de /k^w/ e /k^j/ torna sequências *C+*glide*+V como simplesmente CV, sendo, então, uma análise preferível por explicar uma exceção na fonotática da língua ao mesmo tempo em que se conforma a uma estrutura silábica mais econômica.

Vejamos, agora, exemplos de ocorrência dos fonemas consonantais em seus diferentes contextos e suas respectivas estruturas silábicas. Com base no molde silábico do Arutani, já

²⁹ Lembramos que /ɲ/ - [ɲ] só ocorre diante de /a/ com os alofones [a], [ɐ], [ẽ] – na maioria das vezes em casos de sílaba tônica.

apresentamos o que pode ocorrer em contexto de núcleo silábico, agora, em posição de **ataque** podem ocorrer os seguintes fonemas com seus respectivos alofones (caso existam):

- ❖ O fonema /p/, ao ocorrer em ataque, é produzido sempre como [p]:
 - Exemplos: *pori* [po.'ri] ‘mappinguari’ (espírito); *kapixaw* [kɐ.pɪ.'ʃaw] ‘xexéu’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /t/, ao ocorrer em ataque, é produzido sempre como [t]:
 - Exemplos: *tawaro* [tɐ.wɐ.'ro] ‘bacurau’; *soti* [so.'ti] ‘paca’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /k/, ao ocorrer em ataque, é produzido sempre como [k]:
 - Exemplos: *kiwa* [kɪ.'wa] ‘nome’; *turaka* [tɔ.rɐ.'ka] ‘besouro de chifre’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /m/, ao ocorrer em ataque, é produzido sempre como [m]:
 - Exemplos: *manari* [mɐ.nɐ.'ri] ‘peneira’; *komaxiri* [ko.,mɐ.ʃɪ.'ri] ‘damorida’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /n/, ao ocorrer em ataque, é produzido sempre como [n]:
 - Exemplos: *nikixi* [nɪ.kɪ.'ʃi] ‘tipiti’; *ani kati* [ɐ.'ni kɐ.'ti] ‘carvão’.
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /ɲ/, ao ocorrer em ataque, é produzido sempre como [ɲ]:
 - Exemplos: *unaña* [ɔ.nɐ.'ɲa] ‘azeite’ ou ‘óleo’; *tamkaña* [tɛm.kɛ.ɲɛ] ‘ombro’.
 - Ocorrendo apenas em meio de palavra e, principalmente, em sílaba tônica.
- ❖ O fonema /r/, ao ocorrer em ataque, é produzido sempre como [r]:
 - Exemplos: *rata* [rɐ.'ta] ‘dizer’ ou ‘falar’; *asaraw* [ɐ.sɐ.'raw] ‘molhado’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /s/, ao ocorrer em ataque, é produzido como [s], variando em alguns casos com [ʃ]:
 - Exemplos: *arasa* [ɐ.rɐ.'sa] ‘pica-pau’; *sawan* [sɐ.'wɛn] ‘cutia’; *sako* [sɐ.'ko] ~ [ʃɐ.'ko] ‘tucano’.
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /ʃ/, ao ocorrer em ataque, é produzido sempre como [ʃ]:

- Exemplos: *kirixa* [kɪ.ɾɪ.'ʃa] ‘inhame’; *xipuya* [ʃɪ.pɔ.'ja] ‘pó’.
- Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /h/, ao ocorrer em ataque, é produzido sempre como [h]:
 - Exemplos: *hani* [hɐ'ni] ‘de novo’; *awxihi* [ɐw.ʃɪ.'hi] ‘ubim’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /kʷ/, ao ocorrer em ataque, é produzido sempre como [kʷ]:
 - Exemplos: *kwaso* [kʷɐ.'so] ‘joaninha’; *akwap* [ɐ.'kʷap] ‘galho’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /kɪ/, ao ocorrer em ataque, é produzido sempre como [kɪ]:
 - Exemplos: *kyeta* [kɪɐ.'ta] semente; *okia* [o.'kɪa] ‘pupunha’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /w/, ao ocorrer em ataque, é produzido sempre como [w]:
 - Exemplos: *wahap* [wɐ'hap] ‘bater com pau’; *kiwa* [kɪ.'wa] ‘nome’; *sawika* [sɐ.wɪ.'ka] ‘cujubim’;
 - Ocorrendo tanto em começo quanto em meio da palavra.
- ❖ O fonema /j/, ao ocorrer em ataque, é produzido sempre como [j]:
 - Exemplos: *yakoxa* [jɐ.ko.'ʃa] ‘nuca’; *kaya* [kɐ.'jæ] ‘onça’;
 - Ocorrendo tanto em começo quanto em meio da palavra.

Já em posição de **coda** silábica, encontramos os seguintes fonemas:

- ❖ O fonema /p/, ao ocorrer em coda, pode ser produzido como [p], como [p̄] ou como [b̄]:
 - Exemplos: *kaptiaw* [kɐp.tɪ.'aw] (quantificador usado para quantidades maiores ou iguais a 3); *atap* [ɐ.'tap] ~ [ɐ.'ta^{b̄}] ‘lua’; *ihap* [ɪ.'ha^{p̄}] ‘grilo’; *kapti* [kɐ^{p̄}.ti] ‘ponta’;
 - Ocorrendo tanto em final quanto em meio da palavra.
- ❖ O fonema /t/, ao ocorrer em coda, pode ser produzido como [t̄] ou como [n] quando estiver em ambiente nasal:
 - Exemplos: *akat* [ɐ.'ka^{t̄}] ‘papagaio-de-cabeça-roxa’; *matkia* [mɐ^{t̄}.kɪa] ‘homem’; *emet* [ɐ.'mɛ^{t̄}] ~ [ɐ.'mɛn] ‘casa’;

- Ocorrendo tanto em final quanto em meio da palavra.
- ❖ O fonema /m/, ao ocorrer em coda, é produzido sempre como [m]:
 - Exemplos: *akõm* ã [ɐ.'kõ**m** 'ẽ] ‘jatobá’; *xomhay* [ʃom.'haj] ‘acima’; *xomni* [ʃom.'ni] ‘carrapato’;
 - Ocorrendo tanto em final quanto em meio da palavra.
- ❖ O fonema /n/, ao ocorrer em coda, é produzido sempre como [n]:
 - Exemplos: *xokwan* [ʃo.'kʷẽ**n**] ‘cupim’; *anhĩ* [ɐ**n**.'hĩ] ‘bunda’;
 - Ocorrendo tanto em final quanto em meio da palavra.
- ❖ O fonema /w/, ao ocorrer em ataque, é produzido sempre como [w]:
 - Exemplos: *asaw* [ɐ.'saw] ‘formiga’ (termo genérico); *rawkap* [ɾɐw.'kẽb̃] ‘pacupemba’;
 - Ocorrendo tanto em final quanto em meio da palavra.
- ❖ O fonema /j/, ao ocorrer em ataque, é produzido sempre como [j]:
 - Exemplos: *tahay* [tɐ.'haj] ‘peixe’ (família do mandi); *muya* [moj.'ja] ‘buraco’;
 - Ocorrendo tanto em final quanto em meio da palavra.

Vemos, assim, que as codas em Arutani somente permitem *glides*, nasais e oclusivas (exceção de /k/), mas não fricativas, tepe e /h/.

4.5.2 Núcleo silábico

Com relação aos fonemas vocálicos /a/, /i/ e /o/ e seus alofones, todos ocorrem sempre em posição de núcleo silábico. Os fonemas /a/ e /i/, possuem os alofones [a] e [i] respectivamente, os quais só ocorrem em sílaba tônica e os alofones [ɐ] e [ɪ], que só ocorrem em posição átona. Já quanto aos *glides*, eles sempre vão ocupar a posição de periferia silábica no molde (C)V(C), ou como coda ou como ataque. Portanto, o núcleo silábico só poderá ser composto por uma vogal fonêmica (V).

4.5.3 Sequências Consonantais Heterossilábicas

Essa questão da distribuição de fonemas em posição de ataque e/ou coda e o molde para os padrões silábicos do Arutani é muito importante quando surge a dúvida se há ou não

segmentos africados na língua. Encontramos na língua alguns encontros <ps> e <ts>, como nas palavras: *mapsa* [mɐp.'sa] ‘irmã mais velha dele’ (termo referencial), *irapsa* [ɪ.rɐp̃.'sa] ‘mutum’, *ipsa* [ip̃.'sa] ‘avós’ ou ‘netos’ (vocativo) e *atsia* [ɐ̃t̃.sl.'æ] ‘intestino’. Existem dois principais fatores que nos fazem postular a não existência de encontros como [ps̃] ou africadas como [ts̃], que são: (i) não é produtivo postularmos mais dois novos fones, sendo que a fonotática da língua apresenta oclusivas ocorrendo em coda e [s] ocorrendo em ataque, esse tipo de encontro seria improdutivo para o sistema linguístico, gerando exceções desnecessária a regras já bem estabelecidas e recorrentes; e (ii) a existência de oclusivas não explodidas como [p̃] e [t̃] antes de [s] também nos mostra que esses encontros não se tratariam nem mesmo de africadas plenas, enfraquecendo ainda mais a falsa ideia de africadas na fala de Juanita.

4.5.4 Ambissilabidade

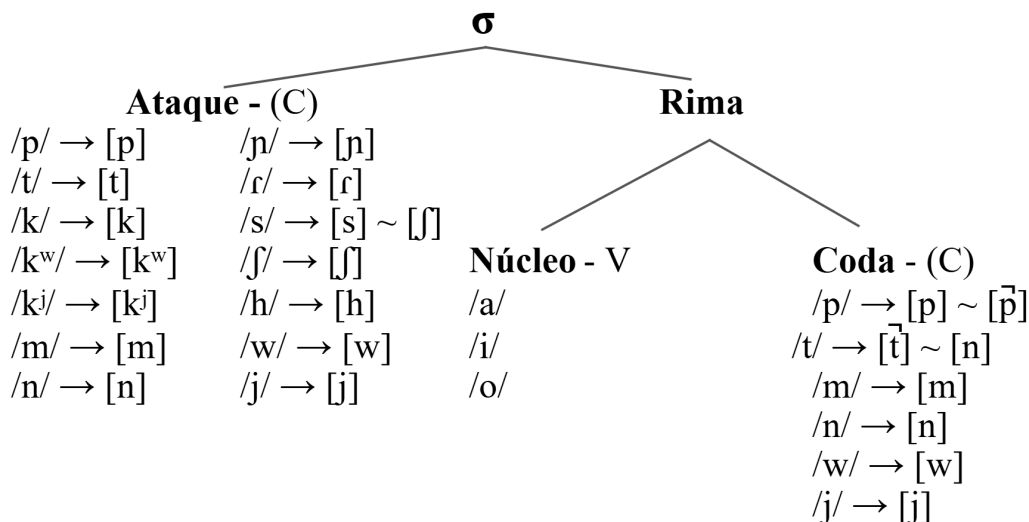
Um outro fenômeno que ocorre na âmbito fonotático da língua é a questão da ambissilabidade na ocorrência de *glides*, questão que já foi introduzida na seção de acústica. Tal fenômeno ocorre quando *glides* passam a “transitar” entre sílabas que ocorrem seguidas. Portanto, os *glides*, representados por /j/ e /w/, em alguns casos, podem transitar da coda de uma sílaba, para o ataque da sílaba seguinte, quando estiverem entre vogais (V_V). tratando-se de uma tendência e não de uma regra da língua. Exemplos dessa **ambissilabidade** estão em palavras como:

- (16) *oya* [oj.'ja] ~ [oj.'a] ‘apimentado’ *awaya* [ɐw.wɐ.'ja] ~ [ɐ.wɐ.'ja] ‘jiboia’

4.5.5 Resumo da estrutura silábica

A partir de todo esse panorama, podemos apresentar quais fonemas e fones podem ocupar as posições de ataque (*onset*), núcleo e/ou coda no Arutani falado por Juanita, tomando como base o seguinte modelo:

Figura 113 – Resumo da estrutura silábica do Arutani



4.6 Acento

Já vimos a questão do acento do Arutani falado por Juanita a partir de uma perspectiva acústica, na qual comparamos duração, *pitch* e intensidade e verificamos que a duração é o traço mais relevante para a marcação do acento, o *pitch*, o segundo mais relevante e a intensidade, o menos relevante entre os três. Agora, analisaremos o acento de uma perspectiva fonológica.

Podemos dizer com confiança que o padrão do Arutani falado por Juanita é o de acentuar a última sílaba da palavra, tendo, assim, um padrão oxítono, com raras exceções em que o acento recai na penúltima sílaba, como em: *maykati* [mɛj.ˈka.tɪ] ‘1SG’, *kaikáti* [kɛj.ˈka.tɪ] ‘2SG’ e *tasiri* [tɛ.ˈsi.ɾɪ] ‘bom’. Portanto, levando em consideração esta forte tendência oxítona e os raros casos de palavras paroxítonas, já fica evidente que o acento não possui um valor fonêmico de contraste na língua, como ocorre no português, por exemplo.

Levando em consideração o que Matzenauer (2014, p. 69) trata sobre a Fonologia Métrica, podemos identificar três tipos de acento em Arutani: primário, secundário e principal. Ao retomarmos estes conceitos, com base na mesma autora, podemos entender o acento primário como sendo o mais proeminente dentro de uma palavra, o acento secundário como sendo menos proeminente que o primário, mas ainda sendo considerado proeminente quando comparado com

as sílabas átonas, e o acento principal como o acento mais proeminente em uma sequência de palavras. Veja cada um desses tipos de acento ocorrendo na fala de Juanita:

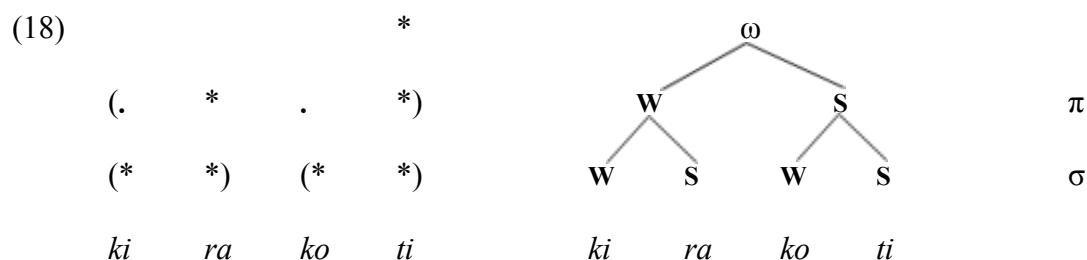
- **Acento primário:** *oso* [o.'so] 'irmão' (termo referencial); e *konaxi* [,ko.nɐ.'ʃi] 'anzol';
- **Acento secundário:** *orikitima* [,o.ɾi.,kɪ.tɪ.'mẽ] 'mata' ou 'floresta'; e *kirakoti* [kɪ.,rɐ.ko.'ti] 'coração';
- **Acento principal:** *pururu ata* [,pɔ.rɯ.'rɯ ɐ.'ta] 'pipira-garganta-preta' e 'rabadilha amarela'; e *yama kosa* [jɐ.'ma ko.'sa] 'mingau de mandioca'.

Ainda seguindo alguns pressupostos da Fonologia Métrica – com base no que é proposto por Hayes (1995) e Collischonn (2014) – quanto ao acento e ao ritmo encontrados no Arutani falado por Juanita, temos: um pé iâmbico, com proeminência final, tendo o elemento mais forte sempre à direita; uma direção de segmentação ou construção dos pés da direita para a esquerda (D/E); e uma segmentação iterativa, um dos motivos pelos quais encontrarmos acento secundário na língua. O Arutani parece permitir pés degenerados, como vemos em algumas palavras monossilábicas como *pa* ['pa] 'fazer'.

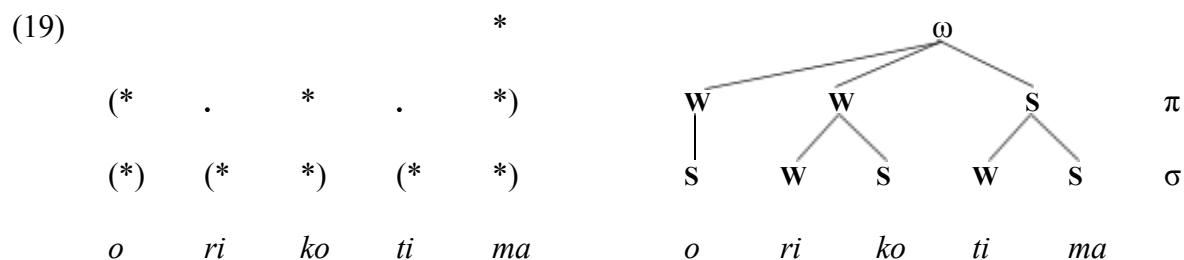
Começemos por ilustrar esse sistema com a palavra *oso* [o.'so] 'irmão', em que temos uma estrutura dissilábica, composta por duas sílabas (σ), com um pé (π) e uma palavra fonológica (ω), sendo representado pela seguinte estrutura de grade:

(17)	*	ω
(.	*)	π
*	*	σ
<i>o</i>	<i>so</i>	

O mesmo padrão ocorre em palavras compostas por quatro sílaba, como *kirakoti* [kɪ.,rɐ.ko.'ti], quando então também observamos o secundário. Observe a seguir a estrutura de grade e de árvore juntas na análise de *kirakoti*, em que **S** representa uma sílaba forte (*strong*) e **W** uma fraca (*weak*). Notemos, especialmente, o ritmo alternante fraco-forte da palavra Arutani, fruto de seus pés iâmbicos iterativos



Por ser uma língua de segmentação iterativa, em que a construção de pés segue até o final da palavra, podemos observar sílabas fortes além da tônica, surgindo assim acentos secundários, como é o caso da sílaba *ra* [ˌrɐ]. Entretanto, há, também, casos de pés degenerados em posição acento secundário, como ocorre em *orikotima* [ˌo.ɾɪ.kɪ.tɪ.ˈmẽ]:



No caso de *orikotima* [ˌo.ɾɪ.kɪ.tɪ.ˈmẽ], o constituinte degenerado passa ser a primeira sílaba, *o* [o], por conta tanto do ritmo acentual (iambo) quanto da direção da construção dos pés e da segmentação iterativa (D/E).

5. Considerações finais

A partir da análise e da descrição da fonética e da fonologia do Arutani falado por Juanita Martinez, conseguimos ampliar e detalhar aspectos de uma língua que era pouco descrita e estudada. As línguas ameríndias, em geral, carecem de um foco ainda maior por parte dos linguistas, devido não só à quantidade de estudos que vêm sendo realizados, mas, também, por possuírem fenômenos ainda pouco investigados na linguística em geral. Com isso, observamos que estudos acústicos sobre línguas indígenas ainda são pouco frequentes e, esperamos, com esta pesquisa, motivar futuros trabalhos para que utilizem ferramentas modernas e tecnológicas como recursos para que as descrições das línguas indígenas sejam ainda mais precisas e embasadas.

Buscamos, também, com esta pesquisa, evidenciar a infeliz realidade que o Brasil e o mundo vivem atualmente, que é a extinção e o silenciamento constante de diversas línguas. Quando uma língua morre a linguística não está simplesmente perdendo mais objeto de estudo, mas a humanidade está tornando-se menos diversa, povos estão perdendo sua identidade e conhecimentos acumulados durante milhares de anos estão deixando de ser repassados. As consequências dessa desvalorização vão além da perda simbólica, pois muitos povos têm deixado de existir e línguas têm deixado de ser faladas por conta do desmatamento e da exploração ilegal de recursos florestais, causando, assim, exodus e morte de diversas comunidades indígenas.

A partir desta pesquisa foi possível constatar que o Arutani falado por Juanita possui 19 fones consonantais: [p], [p̃], [b], [b̃], [t], [t̃], [k], [kʷ], [k̃], [ʔ], [m], [n], [ɲ], [s], [ʃ], [h], [ɾ], [w] e [j], representados por 10 fonemas: /p/, /t/, /k/, /kʷ/, /k̃/, /m/, /n/, /ɲ/, /ɾ/, /s/, /ʃ/, /h/, /w/ e /j/, sendo que [ʔ] é encontrado apenas como um recurso demarcação de fronteira silábica, não possuindo valor fonêmico. Além disso, existem 7 fones vocálicos, [a], [ɐ], [e], [i], [ɪ], [o] e [ʊ], representados pelos fonemas /a/, /i/ e /o/. Todas as vogais possuem sua contra parte nasalizada, entretanto, a nasalidade não possui valor fonêmico na fala de Juanita.

A língua possui um molde silábico (C)V(C), sendo que o núcleo só pode ser ocupado por vogais. Fonemas consonantais oclusivos, nasais e *glides* ocupam tanto posição de ataque quanto

de coda – exceto /k/, que ocorre apenas em ataque. Todos os fonemas consonantais podem ocupar a posição de ataque. O acento na língua tende sempre a recair sobre a última sílaba, com um ritmo iambo, segmentação iterativa (D/E), permitindo a existência de pés degenerados e acento secundário.

Portanto, dar o nosso melhor e fazer o melhor trabalho possível com a realidade e o contexto de pesquisa que temos e em que estamos inseridos deve ser o papel de cientistas que vivem em países como o Brasil, nos quais é escasso o incentivo à pesquisa e o acesso à recursos necessários. O contexto perfeito não existe, mas podemos desenvolver as melhores análises e pesquisas possíveis com base nos nossos dados e objetos de estudo. A ciência só avança quando passamos a ter um olhar científico e fazer ciência apesar das circunstâncias adversas. É principalmente em momentos assim, de contextos difíceis – como não ter um laboratório e um ambiente ideal, visto que a realidade do mundo nunca será a ideal – que avanços significativos surgiram na ciência.

6. Referências

- AIKHENVALD, A. Y.; DIXON, R. M. W. (eds). *Word: A cross-linguistic Typology*. Cambridge: Cambridge University Press, 2002.
- ARMELLA, C.; MARTALLANA, B. Exploración Del Paragua. *Boletín de La Sociedad Venezolana de Ciencias Naturales* 53, 61-110, 1942.
- BAGNO, Marcos. *Nada na língua é por acaso: por uma pedagogia da variação linguística*. – São Paulo. Parábola Editorial, 2007. .
- BARBOSA, P. A. *Prosódia*. [coordenação Tommaso Raso, Celso Ferrarezi 3], - 1, ed. - São Paulo: Parábola, 2019.
- BARBOSA, Plínio A.; MADUREIRA, Sandra. *Manual de fonética acústica experimental: aplicações a dados do português*. – 1. ed. – São Paulo : Cortez Editora, 2015.
- BAUMAN, Zygmunt. *Identidade: entrevista a Benedetto Vecchi*. Tradução: Carlos Alberto Medeiros – Rio de Janeiro: Jorge Zahar Ed., 2005.
- BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer*. 2022. Disponível em: <https://www.fon.hum.uva.nl/praat/>. Acesso em: 29 dez. 2022.
- BORTONI-RICARDO, S. M. *Manual de sociolinguística*. Editora Contexto, 2014.
- BRASIL. [Constituição (1988)]. *Constituição da República Federativa do Brasil de 1988*. Brasília, DF: Presidência da República,. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 11 set. 2022.
- BRASIL. *Decreto nº 7.387*, Brasília, 9 de dezembro de 2010. Institui o Inventário Nacional da Diversidade Linguística e dá outras providências. 2010.
- BRASIL. *Lei nº 9.394, de 20 de dezembro de 1996*. Estabelece as diretrizes e bases da educação nacional. Diário Oficial da União, Brasília, DF, 23 dez. 1996.
- BRASIL. Ministério da Educação. *Plano Nacional de Educação - PNE: Lei nº 13.005, de 25 de junho de 2014*. Aprova o Plano Nacional de Educação - PNE e dá outras providências. Diário Oficial da União, Brasília, DF, 26 jun. 2014.
- CABRAL, Diego. *Nomes e o sintagma nominal em Arutani*. Relatório final de Programa de Iniciação Científica da Universidade de Brasília. Brasília, 2022.
- CAMPBELL, L. *Historical Linguistics: an introduction*. Edinburgh University Press, 2013.
- CAMPBELL, Lyle. Language isolates and their history, or, what's weird, anyway?. In: *Annual Meeting of the Berkeley Linguistics Society*. 2010. p. 16-31. doi: <https://doi.org/10.3765/bls.v36i1.3900>.

CESINE. *save_intervals_to_wav_sound_files.praat*, 2006.

_____. *save_labeled_intervals_to_wav_sound_files.praat*, 2014.

CHACON, T. C. *Produto 2 [Manuscrito]*. ProDoclin - Museu do Índio, 2019.

_____. *Produto 4 [Manuscrito]*. ProDoclin - Museu do Índio, 2020a.

_____. *Produto 6 [Manuscrito]*. ProDoclin - Museu do Índio, 2020b.

CHACON, T. C.; SEIFERT, A. P. ; BERTO, F. F. ; PEREIRA, G. R. ; SILVERSTRIN, M. ; GARCIA, M. V. C. *Guia de Pesquisa e Documentação para o INDL (I)*. 1. ed. Brasília: IPHAN, v. 2. 95p. 2014.

CHACON, Thiago C.; LABRADA, Jorge E. R. How isolate can a language be? The case of Arutani from Northeast Amazonia. *Amazônicas IX*, 9-05-2023, Universidad de los Andes, Bogotá.

CHACON, Thiago. C. (compilador), Jorge. E. ROSÉS LABRADA (tradutor), Juanita MARTINEZ (pesquisadora), Albino XIRIANA (pesquisador), G.F. SOARES (pesquisador), G. F. BARROS (pesquisador), D. CABRAL (pesquisador). *Dicionário Multimídia Arutani-Ninam(-Português-Inglês-Espanhol)*. Versão 1.0. Rio de Janeiro: PRODOCLIN/Museu do Índio. 2020. Disponível em: <https://japiim.museudoindio.gov.br/dic/arutani/>. Acesso em: 20 abr. 2024.

COLLISCHONN, G. Acento em português. In: BISOL, Leda. *Introdução a estudos de fonologia do português brasileiro* [recurso eletrônico]. (Org.) Leda Bisol. – 5.ed. – Dados eletrônicos. – Porto Alegre : EDIPUCRS, 2014.

COPPENS, Walter. Los Uruak (Arutani). In W. COPPENS, M. Á. Perera, R. Lizarralde & H. Seijas (eds.) *Los aborígenes de Venezuela*. Volume 2, 747-770. Caracas: Fundación La Salle/Monte Avila Editores/Ediciones IVIC/Instituto Caribe de Antropología y Sociología, 2008.

ELAN (Versão 6.4) [Software de computador]. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive, 2022. Disponível em: <https://archive.mpi.nl/tla/elan>. Acesso em: 29 dez. 2022.

FERREIRA, H. P.; MACHADO, A. M. A.; SENRA, E. B. *As Línguas Yanomami no Brasil: Diversidade e Vitalidade*. São Paulo, Instituto Socioambiental, 2019.

GARCIA, M. V. C. Retratos | A diversidade linguística como patrimônio cultural. In: IPEA. *Desafios do Desenvolvimento: Quem somos, aonde vamos*. 11. ed. Brasília - DF, p. 88-92, 2014.

HALL, Stuart. *A identidade cultural na pós-modernidade*. Tradução de Tomaz Tadeu da Silva & Guacira Lopes Louro – 12 ed. – Rio de Janeiro: Lamparina, 2019.

HAYES, Bruce. *Introductory Phonology*. Blackwell Publishing, 2009.

HAYES, Bruce. *Metrical stress theory: Principles and case studies*. University of Chicago press, 1995.

HIRST, Daniel. *analyse_intervals.praat* [script]. 2009.

HIRST, Daniel. *analyse_tier.praat* [script]. 2012.

INE. *Censo Nacional de Población y Vivienda 2011*. Empadronamiento de La Población Indígena - Caracas. 2015.

JOLKESKY, Marcelo Pinho De Valhery. *Estudo arqueo-ecolinguístico das terras tropicais Sul-Americanas*. 2016. 834 f., il. Tese (Doutorado em Linguística)—Universidade de Brasília, Brasília, 2016.

KENT, R. D.; READ, C. *Análise acústica da fala*. Tradução: Alessandro Rodrigues Meireles. – 1. ed. – São Paulo: Cortez, 2015.

LABRADA, Jorge E. R.; CHACON, Thiago C.; MEDINA, Francia. Understanding the Near Extinction of Arutani and Sapé (Venezuela and Brazil). Special Workshop “The Sociolinguistics of Language Endangerment”, *20th International Congress of Linguists*, 2-6-07-2018. Cape Town, South Africa.

LABRADA, Jorge Emilio Rosés; CHACON, Thiago; MEDINA, Francia. Arutani (Venezuela and Brazil) – Language Snapshot. In Peter K. Austin (ed.) *Language Documentation and Description 17*, 170-177. London: EL Publishing, 2020. Disponível em: . Acesso em: 20 set. 2022.

LADEFOGED, P.; MADDIESON, I. *The Sounds of the World's Languages*. Oxford: Blackwell Publishers, 1995.

LEITE, A. F., SILVA, S. B., BRITTO, A. T. B. D. O., & DI NINNO, C. Q. D. M. S. Caracterização do ceceo em pacientes de um Centro Clínico de Fonoaudiologia. *Revista da Sociedade Brasileira de Fonoaudiologia*, v. 13, p. 30-36, 2008.

LOUKOTKA, Č. *Classification of South American Indian languages*. Los Angeles: Latin American Center, University of California; Caracas: Sucre, 1968. 453 f. (Reference Series, v. 7.)

MATISOFF, James A. Rhinoglottophilia: the mysterious connection between nasality and glottality. In: *Nasálfest: Papers from a symposium on nasals and nasalization*. Stanford: Language Universals Project, Department of Linguistics, Stanford University, 1975. p. 265-287.

MATZENAUER, C. M. B. Introdução à teoria fonológica. In: BISOL, Leda. *Introdução a estudos de fonologia do português brasileiro* [recurso eletrônico]. (Org.) Leda Bisol. – 5.ed. – Dados eletrônicos. – Porto Alegre : EDIPUCRS, 2014.

McCLOY, Daniel R. *Normalizing and plotting vowels with phonR 1.0.7*, University of Washington, 2016.

MIGLIAZZA, Ernest C. Maku, Sape and Uruak languages current status and basic lexicon. *Anthropological Linguistics* 20(3), 133-140, 1978.

MIGLIAZZA, Ernest. *Grupos lingüísticos do Território Federal de Roraima*. Atas do Simpósio sobre Biota Amazônica, Antropologia, vol. 2, pp. 153-173, 1967.

MORETTIN, Pedro A.; BUSSAB, Wilton O. *Estatística básica*. - 9. ed. - São Paulo: Saraiva, 2017.

MORI, A. C. Fonologia. In: *Introdução à linguística: domínios e fronteiras*, Fernanda Mussalim e Anna Christina Bentes (orgs). Volume 1. 4a edição. Editora Cortez: São Paulo. 2000. 147-180.

PARKER, Steve. La rinoglotofilia y el gruñido de afirmación: una tendencia universal. *Lengua y Sociedad*, [S. l.], v. 8, n. 1, p. 27-56, 2006. DOI: 10.15381/lengsoc.v8i1.26504. Disponível em: <https://revistasinvestigacion.unmsm.edu.pe/index.php/lenguaysociedad/article/view/26504..> Acesso em: 7 jan. 2025.

PIKE, K.L. *Phonemics: A Technique for Reducing Languages to Writing*. Ann arbor: mich. Univ. Press, 1947.

PISEMSKAYA, Natalia B. Problemática lingüístico-educativa de los pueblos indígenas de Venezuela. *Íkala, revista de lenguaje y cultura*, v. 14, n. 23, p. 65-84, 2009.

_____. *Lenguas minoritarias de Venezuela: consideraciones desde la perspectiva ecolingüística*. *Revista de Filología y Lingüística de la Universidad de Costa Rica*, p. 175-189, 2010.

QUIJANO, Aníbal. Colonialidade do poder, eurocentrismo e América Latina. In: LANDER, Edgardo (Org.). *A colonialidade do saber: eurocentrismo e ciências sociais - perspectivas latino-americanas*. Ciudad Autónoma de Buenos Aires, Argentina: Clacso, p. 107-30, 2005.

RODRIGUES, A. D. Tarefas da lingüística no Brasil. *Estudos Lingüísticos* (Revista Brasileira de Lingüística Teórica e Aplicada), vol. 1, n. 1, p. 4-15, 1966.

RODRIGUES, Aryon D. Diversidade Lingüística na Amazônia. *Anais do Seminário Internacional sobre Meio Ambiente, Pobreza e Desenvolvimento da Amazônia*. Belém, Sindamazônia, 1992.

RUBIM, Altaci Corrêa. *O reordenamento político e cultural do povo Kokama: a reconquista da língua e do território além das fronteiras entre o Brasil e o Peru*. 2016. 319 f., il. Tese (Doutorado em Linguística)—Universidade de Brasília, Brasília, 2016.

SILVA, Leandro A.; PERES, Sarajane M.; BOSCARIOLI, Clodis. *Introdução à mineração de dados: com aplicações em R*. - 1. ed. - Rio de Janeiro: Elsevier, 2016.

SILVA, Thaïs C.; SEARA, Izabel; SILVA, Adelaide; RAUBER, Andreia S.; CANTONI, Maria. *Fonética Acústica: os sons do português brasileiro*. Editora Contexto, 2019.

SOUZA, P.C.; SILVA, R.S. Fonética. In.: FIORIN, J. L. *Introdução à Linguística II: princípios de análise*. São Paulo: Contexto, 2010a.

SOUZA, P.C.; SILVA, R.S. Fonologia. In.: FIORIN, J. L. *Introdução à Linguística II: princípios de análise*. São Paulo: Contexto, 2010b.

SPINASSÉ, K. P. Os conceitos Língua Materna, Segunda Língua e Língua Estrangeira e os falantes de línguas alóctones minoritárias no Sul do Brasil. *Contingentia*, Porto Alegre, Brasil., v. 1, n. 1, 2008. Disponível em: <https://seer.ufrgs.br/index.php/contingentia/article/view/3837>. Acesso em: 5 jun. 2023.

ZAMPONI, Raoul; ROGERS, Chris. 16 Máku. In: DARQUENNES, Jeroen; EPPS, Patience. *Language Isolates II: Kanoé to Yurakaré: An International Handbook*, p. 767-806, 2023.

Anexo I

O *Script 1* (Hirst, 2009) é executado no Praat e realiza a análise acústica automática de intervalos marcados em arquivos "TextGrid", os quais estão associados a arquivos de áudio ".wav". Sua função principal é analisar todos os intervalos da *tier* “segmentos” em múltiplos pares de arquivos “.TextGrid” e “.wav”, extraíndo e organizando diversas medidas acústicas para cada intervalo. As informações extraídas para cada intervalo são: duração (em milissegundos); *pitch* médio (F0); intensidade média; formantes (F1, F2 e F3): média no intervalo e valor no ponto médio do intervalo; valor máximo e mínimo do *pitch* (F0) no intervalo; amplitude máxima da intensidade; e variação de *pitch* (f0range).

Os dados (*output*) foram organizados da seguinte forma pelo *script*: os resultados são impressos na janela de informações (*Info window*) do Praat, cada linha representando um intervalo analisado e cada coluna contendo uma medida específica, incluindo também: nome do arquivo de som, *label* do intervalo e código de vogal (inseridos como parâmetros). Há casos de valores indefinidos, estes são representados por "NA".

Script 1 para a extração de valores de duração, *pitch*, intensidade e formantes (F1, F2 e F3)::

```
#praat script
script_name$ = "analyse_intervals.praat"
#author Daniel Hirst
#email daniel.hirst@lpl.univ-aix.fr
version$ = "[2009:01:29]"
date$ = date$()

#purpose Analyse a folder of Sound files and a folder of TextGrid files
# for each interval on selected tier calculate
#       - duration, mean pitch, intensity, f1, f2 f3
#output to Info window

#define parameters used in the script
form analyse intervals
  word subject sub7
  word vowelp 1
  sentence Sound_folder \Users\HP\Desktop\audios_e_transcricoes - Copia\
  sentence TextGrid_folder \Users\HP\Desktop\audios_e_transcricoes - Copia\
  word Analysis_tier phones
  positive Time_step 0.01
  boolean automatic_max_min yes
  natural Min_pitch 50
```

```

    natural Max_pitch 700
    natural Number_of_formants 5
    natural Maximum_formant 5500
    positive Window_length 0.025
    positive Pre_emphasis 50
    word Undefined_value NA
endform

clearinfo

printline 'tab$'extract'tab$'label'tab$'subject'tab$'vp'tab$'duration'tab$'pitch'tab$'intensity
... 'tab$'F1'tab$'F2'tab$'F3'tab$'f1mid'tab$'f2mid'tab$'f3mid'tab$'stim'tab$'MaxInt'tab$'f0max'tab$'f0min'tab$'f0range

#Read in list of sound files
myStrings = Create Strings as file list... sounds 'sound_folder$'/*.wav
nSounds = Get number of strings
item = 0

#check if TextGrid file exists for each sound and call treatment
for iSound from 1 to nSounds
    select myStrings
    sound_name$ = Get string... iSound
    textGrid_name$ = sound_name$ - "wav" + "TextGrid"
    sound$ = sound_folder$ + "/" + sound_name$
    textGrid$ = textGrid_folder$ + "/" + textGrid_name$
    if fileReadable(textGrid$)
        call treatment
    endif
endfor

#subroutine treatment
procedure treatment
    Read from file... 'sound$'
    mySound = selected("Sound")
    name$ = selected$("Sound")
    Read from file... 'textGrid$'
    myTextGrid = selected("TextGrid")
    nTiers = Get number of tiers
#find number of analysis tier
    tier = 0
    for iTier from 1 to nTiers
        tier_name$ = Get tier name... iTier
        if tier_name$ = analysis_tier$
            tier = iTier
        endif
    endfor
    if tier
#create analysis objects
        select mySound
        myPitch = To Pitch... time_step min_pitch max_pitch
        if automatic_max_min

```

```

q25 = Get quantile... 0 0 0.25 Hertz
q75 = Get quantile... 0 0 0.75 Hertz
min_pitch = 0.75*q25
max_pitch = 1.5*q75
Remove
select mySound
myPitch = To Pitch... time_step min_pitch max_pitch
#printline # min_pitch 'min_pitch:0'; max_pitch 'max_pitch:0'
endif
select mySound
myIntensity = To Intensity... min_pitch time_step Yes
select mySound
myFormant = To Formant (burg)... time_step number_of_formants maximum_formant
window_length pre_emphasis

#Get time values of beginning and end of each interval for treatment
select myTextGrid
nIntervals = Get number of intervals... tier

for iInterval from 1 to nIntervals
    select myTextGrid
    label$ = Get label of interval... tier iInterval
    #type$ = Get label of interval... 4 iInterval
    ##senPosition$ = Get label of interval... 4 iInterval
    ##tones$ = Get label of interval... 5 iInterval
    #stim$ = Get label of interval... 1 iInterval
    subject$ = subject$
    vp$ = vowelp$
    start = Get starting point... tier iInterval
    end = Get end point... tier iInterval
    mid = (end-start)/2
    if label$ != "" and label$ != "_" and label$ != "#"
#calculate parameters for each non empty interval
        item = item+1
#duration
        duration = 1000*(end - start)
        call set_undefined duration
        duration$ = value$
#pitch
        select myPitch
        meanPitch = Get mean... start end Hertz
        call set_undefined meanPitch
        meanPitch$ = value$
        maxTime= Get time of maximum... start end Hertz Parabolic
        f0max= Get maximum... start end Hertz Parabolic
        f0min= Get minimum... start end Hertz Parabolic
        f0range=f0max-f0min
        #f0minPre = Get minimum... start maxTime Hertz Parabolic
        #f0minPost = Get minimum... maxTime end Hertz Parabolic

#intensity
        select myIntensity

```

```

meanIntensity = Get mean... start end energy
max_int = Get maximum... start end Parabolic
call set_undefined meanIntensity
meanIntensity$ = value$

#formants

select myFormant
f1 = Get mean... 1 start end Hertz
f1_mean = Get value at time... 1 mid Hertz Linear
call set_undefined f1
f1$ = value$
f2 = Get mean... 2 start end Hertz
f2_mean = Get value at time... 2 mid Hertz Linear
call set_undefined f2
f2$ = value$
f3 = Get mean... 3 start end Hertz
f3_mean = Get value at time... 3 mid Hertz Linear
call set_undefined f3
f3$ = value$

#print out results

printline 'item"tab$"name$"tab$"label$"tab$"subject$"tab$"vp$"tab$"duration$'
...'tab$"meanPitch$"tab$"meanIntensity$"tab$"f1$'

...'tab$"f2$"tab$"f3$"tab$"f1_mean"tab$"f2_mean"tab$"f3_mean"tab$"stim$"tab$"max_int"tab$"f0max"tab$"f0min"t
ab$"f0range'

endif
endfor
#Remove analysis files
select myPitch
plus myIntensity
plus myFormant
Remove
else
#print warning if TextGrid has no analysis tier
printline ###TextGrid 'name$' has no tier 'analysis_tier$'
endif
#Remove Sound and TextGrid
select mySound
plus myTextGrid
Remove
endproc

#Remove file list
select myStrings
Remove

procedure set_undefined value
if value = undefined
value$ = undefined_value$
else
value$ = "value:0"
endif

```

endproc

#version history

#[2009:01:29] added automatic max min for pitch

#[2007:10:30] first version calculates duration pitch intensity & formants on intervals

Anexo II

Trecho do *output* do *Script* 1:

	extract	label	duration	pitch	intensity	F1	F2	F3
1	atx_20170901_tcc_1	ɐ	172	214	77	822	1619	2688
2	atx_20170901_tcc_1	t	124	NA	68	816	1670	2687
3	atx_20170901_tcc_1	a*	137	233	83	807	1508	2352
4	atx_20170901_tcc_1	ḃ	126	224	74	753	1568	2452
5	atx_20170901_tcc_2	ɐ	230	192	69	868	1575	2705
6	atx_20170901_tcc_2	t	108	270	61	833	1637	2635
7	atx_20170901_tcc_2	a*	202	231	73	800	1457	2205
8	atx_20170901_tcc_2	p̄	32	234	56	749	1539	2246
9	atx_20170901_tcc_3	m	168	241	70	798	1786	2485
10	atx_20170901_tcc_3	ɐ	99	202	82	815	1683	3225
11	atx_20170901_tcc_3	r	35	207	77	680	1772	2809
12	atx_20170901_tcc_3	ɪ	145	210	78	444	1808	2994
13	atx_20170901_tcc_3	j	121	214	76	508	1620	2909
14	atx_20170901_tcc_3	æ*	209	233	80	770	1600	2367
15	atx_20170901_tcc_4	o	161	207	61	600	1127	2566
16	atx_20170901_tcc_4	ʔ	16	208	58	588	1264	2332
17	atx_20170901_tcc_4	o*	300	260	74	560	893	2542
18	atx_20170901_tcc_5	t	63	NA	65	690	1665	2751
19	atx_20170901_tcc_5	o	76	210	76	626	1733	2791
20	atx_20170901_tcc_5	ɐ	192	212	80	633	1507	2768
21	atx_20170901_tcc_5	r	43	236	78	469	1747	2548
22	atx_20190807_tcc_2	i*	166	168	69	332	1278	3000

Anexo III

Este *Script 2* (Hirst, 2012) também é executado no *software* Praat. O *script* realiza a análise automática de intervalos ou pontos marcados em uma *tier* específica de arquivos “TextGrid”, utilizando os respectivos arquivos de áudio “.wav”. Ele percorre uma pasta com esses arquivos e, para cada segmento anotado, extrai informações acústicas. O *script* calcula a duração de cada intervalo ou janela de análise. No caso desta pesquisa, extraímos os valores de: frequência fundamental (*pitch*): valores mínimo, médio e máximo; intensidade: valores mínimo, médio e máximo; e formantes: F1, F2 e F3, com base em análise LPC.

Esses dados (*output*) são organizados em uma tabela, impressa na janela de informações (*Info window*) do Praat e formatada para fácil exportação, sendo salva como arquivo “.txt”. Cada linha da tabela corresponde a um intervalo ou ponto analisado, contendo o nome do arquivo, o rótulo do segmento, a duração e as medidas acústicas solicitadas. Os valores indefinidos são substituídos por “NA”.

Script 2 para a extração de valores de duração, *pitch*, intensidade e formantes (F1, F2 e F3):

```
#praat script
script_name$ = "analyse_tier.praat"
#author Daniel Hirst
#email daniel.hirst@lpl.univ-aix.fr
version$ = "[2012-10-30]"
date$ = date$()

#purpose Analyse a folder of Sound files
#           and a folder of TextGrid files
#           The folders can be selected by giving their path
#           or by using the browser

#           For each interval or point on selected tier
#           the script calculates
#           - duration
#           and, depending on the options selected,
#           -mean/min/max pitch (using Hertz/octaves/semitones and 1, 100, 200, 440 or median
#           as reference)
#           intensity, f1, f2 f3
#           the results are output to the Info window
#           and can be saved as a .txt file which can be read directly by
#           a statistics package like R

#define parameters used in the script
```

```

form analyse tier
sentence investigator <Put your name here>
comment Give the path to the folders or leave empty to select them with the browser
sentence Sound_folder E:\Audios\
sentence TextGrid_folder E:\Audios\
word Analysis_tier segmento
word Sound_extension .wav
word TextGrid_extension .TextGrid
real Time_step 0 (= automatic)
word Undefined_value NA
boolean calculate_pitch yes
boolean automatic_min_max yes
natural min_pitch 60
natural max_pitch 750
optionmenu pitch_units: 1
    option Hertz
    option octaves
    option semitones
comment Dor log scales:
optionmenu pitch_reference: 1
    option 1
    option 100
    option 200
    option 440
    option speaker_median
boolean calculate_intensity yes
comment For formants
boolean calculate_formants yes
natural Number_of_formants 5
natural Maximum_formant 5500
positive Window_length 0.05
positive Pre_emphasis 50
comment For point tier
positive Analysis_window 0.1 (= seconds)
endform

clearinfo
default_minimum_pitch = 60
default_maximum_pitch = 750

if sound_folder$ = ""
    sound_folder$ = chooseDirectory$("Choose the folder containing the Sound files")
endif
if textGrid_folder$ = ""
    textGrid_folder$ = chooseDirectory$("Choose the folder containing the TextGrid files")
endif

#Read in list of sound files
mySounds = Create Strings as file list... sounds
... 'sound_folder$/*'sound_extension$'
nSounds = Get number of strings
item = 0

```

```

myTextGrids = Create Strings as file list... TextGrids
... 'textGrid_folder$'/*'textGrid_extension$'
nTextGrids = Get number of strings

if nSounds > 0 and nTextGrids > 0
    nSounds = Get number of strings
#print header of output file
    printline #File created by 'script_name$'
    ... version 'version$'
    printline #Author: Daniel Hirst <daniel.hirst@lpl-aix.fr>
    printline #Analysis carried out by ['investigator$']
    ... on 'date$' on tier ['analysis_tier$']
    printline #
    printline #Parameters:
    if calculate_pitch
        if automatic_min_max
            printline # 'tab$pitch: Automatic min max
        else
            printline # 'tab$pitch:
            ... min='min_pitch'; max='max_pitch'
        endif
        printline # Pitch units: 'pitch_units$'
        if pitch_units != 1
            printline Pitch_reference: 'pitch_reference$'
        endif
    endif
    if calculate_intensity
        printline # intensity:
    endif
    if calculate_formants
        print # 'tab$formants: n='number_of_formants'
        ... , max='maximum_formant', window='window_length',
        ... pre-emphasis='pre_emphasis'
    endif
    printline #
#print column names
    print 'tab$file 'tab$'label 'tab$'duration
    if calculate_pitch
        print 'tab$f0_min 'tab$f0_mean 'tab$f0_max
    endif
    if calculate_intensity
        print 'tab$int_min 'tab$int_mean 'tab$int_max
    endif
    if calculate_formants
        printline 'tab$F1 'tab$F2 'tab$F3
    endif
    printline
#check if TextGrid file exists for each sound
# and call treatment
    for iSound from 1 to nSounds
        select mySounds
        sound_name$ = Get string... iSound
        textGrid_name$ = sound_name$ - sound_extension$

```

```

        ... + textGrid_extension$
        sound$ = sound_folder$ + "/" + sound_name$
        textGrid$ = textGrid_folder$ + "/" + textGrid_name$
        if fileReadable(textGrid$)
            call treatment
        else
            printline # Cannot find TextGrid file
            ... ['textGrid_name$']
        endif
    endfor
else
    if nSounds = 0
        printline Folder ['sound_folder$'] doesn't contain
        ... any files with extension ['sound_extension$']
    elseif nTextGrids = 0
        printline Folder ['textGrid_folder$'] doesn't contain
        ... any files with extension ['textGrid_extension$']
    endif
endif
#Remove file list
select mySounds
plus myTextGrids
Remove

exit

#subroutine treatment
procedure treatment
    Read from file... 'sound$'
    mySound = selected("Sound")
    name$ = selected$("Sound")
    sound_duration = Get total duration
    Read from file... 'textGrid$'
    myTextGrid = selected("TextGrid")
    textGrid_duration = Get total duration
    if sound_duration != textGrid_duration
        plus mySound
        Scale times
        printline # TextGrid and Sound have different durations
        printline # TextGrid has been scaled to the duration of Sound
    endif
#find number of analysis tier
select myTextGrid
call find_analysis_tier
if tier
#create analysis objects
select mySound
printline # file : 'name$'

# - pitch
    if calculate_pitch
        if automatic_min_max
            myPitch = To Pitch... time_step
            ... default_minimum_pitch default_maximum_pitch
        endif
    endif
endif

```

```

        call automatic_min_max_pitch
    else
        myPitch = To Pitch... time_step
        ... min_pitch max_pitch
    endif
    median_pitch = Get quantile... 0 0 0.5 Hertz
    printline # min_pitch: 'min_pitch:0';
    ... median_pitch: 'median_pitch:0'
    ... max_pitch 'max_pitch:0';
    if pitch_reference$ = "speaker_median"
        pitch_reference = median_pitch
    else
        pitch_reference = 'pitch_reference$'
    endif
endif
# - intensity
if calculate_intensity
    select mySound
    myIntensity = To Intensity... min_pitch
    ... time_step Yes
endif
if calculate_formants
    select mySound
    myFormants = To Formant (burg)... time_step
    ... number_of_formants maximum_formant
    ... window_length pre_emphasis
endif
#Get time values of beginning and end of each interval
select myTextGrid
if isIntervalTier
    nIntervals = Get number of intervals... tier
else
    nIntervals = Get number of points... tier
endif
for iInterval from 1 to nIntervals
    select myTextGrid
    if isIntervalTier
        label$ = Get label of interval... tier
        ... iInterval
        start = Get starting point... tier
        ... iInterval
        end = Get end point... tier iInterval
    else
        label$ = Get label of point... tier
        ... iInterval
        mid = Get time of point... tier iInterval
        start = mid - analysis_window/2
        if start < 0
            start = 0
        endif
        end = mid + analysis_window/2
        if end > sound_duration
            end = sound_duration
        endif
    endif
endfor

```

```

        endif
    endif
    #calculate parameters for each non empty interval
    if label$ != "" and label$ != "_" and
    ... label$ != "#"
        item = item+1

# - duration
        duration = 1000*(end - start)
        call set_undefined duration
        duration$ = value$

#- pitch
        if calculate_pitch
            select myPitch
            f0_min = Get minimum... start end
            ... Hertz Parabolic
            call convert_f0 f0_min
            f0_min = value
            call set_undefined f0_min
            f0_min$ = value$
            f0_mean = Get mean... start end Hertz
            call convert_f0 f0_mean
            f0_mean = value
            call set_undefined f0_mean
            f0_mean$ = value$
            f0_max = Get maximum... start end
            ... Hertz Parabolic
            call convert_f0 f0_max
            f0_max = value
            call set_undefined f0_max
            f0_max$ = value$
        endif

#- intensity
        if calculate_intensity
            select myIntensity
            intensity_min = Get minimum... start
            ... end Parabolic
            call set_undefined intensity_min
            intensity_min$ = value$
            intensity_mean = Get mean... start end
            ... energy
            call set_undefined intensity_mean
            intensity_mean$ = value$
            intensity_max = Get maximum... start
            ... end Parabolic
            call set_undefined intensity_max
            intensity_max$ = value$
        endif

#- formants
        if calculate_formants
            select myFormants
            f1 = Get mean... 1 start end Hertz
            call set_undefined f1
            f1$ = value$

```

```

        f2 = Get mean... 2 start end Hertz
        call set_undefined f2
        f2$ = value$
        f3 = Get mean... 3 start end Hertz
        call set_undefined f3
        f3$ = value$
    endif
#Print out results
    print 'item' 'tab$"name$" 'tab$"label$"
    ... 'tab$"duration$"
    if calculate_pitch
        print 'tab$"f0_min$" 'tab$"f0_mean$"
        ... 'tab$"f0_max$"
    endif
    if calculate_intensity
        print 'tab$"intensity_min$"
        ... 'tab$"intensity_mean$"
        ... 'tab$"intensity_max$"
    endif
    if calculate_formants
        print 'tab$"f1$" 'tab$"f2$"
        ... 'tab$"f3$"
    endif
    printline
endif
endfor
#Remove objects
    if calculate_pitch
        select myPitch
        Remove
    endif
    if calculate_intensity
        select myIntensity
        Remove
    endif
    if calculate_formants
        select myFormants
        Remove
    endif
else
#Print warning if TextGrid has no analysis tier
    printline ###TextGrid ['name$'] has no tier
    ... ['analysis_tier$']
endif
#Remove Sound and TextGrid
    select mySound
    plus myTextGrid
    Remove
endproc

procedure convert_f0 value
    value = value/pitch_reference
    if pitch_units$ = "octaves"

```



```

        value = log2(value)
    elseif pitch_units$ = "semitones"
        value = 12*log2(value)
    endif
endproc

procedure set_undefined_value
    if value = undefined
        value$ = undefined_value$
    elseif pitch_units = 1
        value$ = "value:0"
    else
        value$ = "value:3"
    endif
endproc

procedure automatic_min_max_pitch
    q25 = Get quantile... 0 0 0.25 Hertz
    q75 = Get quantile... 0 0 0.75 Hertz
    min_pitch = 0.75*q25
    max_pitch = 1.5*q75
    Remove
    select mySound
    myPitch = To Pitch... time_step min_pitch
    ... max_pitch
endproc

procedure find_analysis_tier
    nTiers = Get number of tiers
    tier = 0
    for iTier from 1 to nTiers
        tier_name$ = Get tier name... iTier
        if tier_name$ = analysis_tier$
            tier = iTier
            isIntervalTier = Is interval tier... tier
        endif
    endfor
endproc

```

Anexo IV

Trecho do *output* do *Script 2*:

	file	label	duration	f0_min	f0_mean	f0_max	int_min	int_mean	int_max	F1	F2	F3
# file : atx_20170901_tcc_1												
ch: 164; median_pitch: 224 max_												
1	atx_20170901_tcc_1	ɐ	172	206	215	223	65	77	83	829	1608	2683
2	atx_20170901_tcc_1	t	124	NA	NA	NA	62	67	74	793	1659	2680
3	atx_20170901_tcc_1	a*	137	195	233	249	79	83	85	809	1566	2519
4	atx_20170901_tcc_1	ḃ	126	186	219	233	66	74	78	752	1532	2224
# file : atx_20170901_tcc_2												
ch: 145; median_pitch: 201 max_												
5	atx_20170901_tcc_2	ɐ	230	184	192	197	54	69	75	872	1612	2871
6	atx_20170901_tcc_2	t	108	201	201	203	51	61	73	804	1605	2697
7	atx_20170901_tcc_2	a*	202	203	231	240	59	73	76	800	1461	2103
8	atx_20170901_tcc_2	p̄	32	231	232	236	53	56	57	838	1558	2145
# file : atx_20170901_tcc_3												
ch: 155; median_pitch: 212 max_												
9	atx_20170901_tcc_3	m	168	209	238	285	60	70	76	885	1810	2724

10	atx_20170901_tcc_3	ɐ	99	196	202	205	77	82	84	816	1682	3220
11	atx_20170901_tcc_3	ɾ	35	200	206	209	76	77	78	599	1792	2952
12	atx_20170901_tcc_3	ɪ	145	207	210	214	74	78	81	443	1809	3011
13	atx_20170901_tcc_3	j	121	205	214	235	73	76	80	465	1602	2765
14	atx_20170901_tcc_3	æ	209	204	233	241	68	80	83	770	1623	2272
# file : atx_20170901_tcc_4												
ch: 160; median_pitch: 243 max_												
15	atx_20170901_tcc_4	o	161	198	210	228	52	61	64	620	1072	2393
16	atx_20170901_tcc_4	ʔ	16	169	192	210	51	57	61	572	1159	2293
17	atx_20170901_tcc_4	o*	300	162	437	616	55	74	80	555	843	2523
# file : atx_20170901_tcc_5												
ch: 153; median_pitch: 214 max_												
18	atx_20170901_tcc_5	t	63	270	270	270	62	65	69	667	1664	2770
19	atx_20170901_tcc_5	o	76	192	208	283	74	76	79	639	1787	2850
20	atx_20170901_tcc_5	ɐ	192	197	213	225	75	80	83	633	1459	2698
21	atx_20170901_tcc_5	ɾ	43	219	237	247	76	78	81	434	1360	2308
22	atx_20170901_tcc_5	i*	203	155	222	251	72	77	80	332	1629	2914

Anexo V

Este código em R tem como objetivo principal analisar dados acústicos de vogais da língua encontradas em Arutani, focando nas frequências dos formantes F1 e F2. O processo começa com a verificação e o carregamento automático dos pacotes necessários, como “readxl” para leitura de planilhas, “tidyverse” para manipulação de dados, MASS, e os pacotes “reshape2” e “reshape” para reestruturação dos dados. Em seguida, os dados são importados de uma planilha do Excel. A partir desse arquivo, o *script* seleciona apenas as colunas relevantes – o rótulo das vogais (*label*), e os valores de F1 e F2 –, e adiciona uma coluna auxiliar chamada “id” para numerar as observações.

Com os dados preparados, o *script* gera três gráficos: dois *boxplots*, um para F1 e outro para F2, organizados por vogal, que permitem visualizar a dispersão e a presença de possíveis *outliers*; e um gráfico de dispersão entre F2 e F1, com os eixos invertidos — o que segue a convenção tradicional de representação do espaço vocálico. Após essa visualização inicial, o *script* realiza um cálculo manual dos quartis e do intervalo interquartil (IIQ) para uma vogal específica (no caso, a vogal [a]). Com base nesses valores, o código determina os limites superior e inferior que serão usados para detectar *outliers* de acordo com a regra padrão de 1,5 vezes o IIQ.

Na etapa seguinte, os dados são reorganizados em formato longo (*long format*), o que permite aplicar a mesma lógica de identificação de *outliers* a todas as vogais e para os dois formantes simultaneamente. Para cada combinação de vogal e variável (F1 ou F2), o código calcula os quartis, o IIQ, os limites de corte e marca os valores que se situam fora desses limites como *outliers*. Por fim, o *script* cria uma nova base que identifica quais observações possuem ao menos um valor considerado *outlier* e, a partir dela, gera um novo gráfico de F2 por F1 excluindo os outliers. Esse gráfico final mostra o espaço vocálico de maneira mais limpa, eliminando valores extremos que poderiam distorcer a análise fonético-fonológica.


```
BASE_COMUM_ini <- BASE_OUTLIER %>%  
  group_by(id) %>%  
  summarize(outlier = any(outlier)) # Adiciona colunas originais F1 e F2  
  
BASE_PLOT_OUTLIER = left_join(BASE, BASE_COMUM_ini)  
  
ggplot(BASE_PLOT_OUTLIER %>% filter(outlier == FALSE), aes(x = F2, y = F1)) +  
  geom_point(aes(color = label)) +  
  scale_y_reverse() +  
  scale_x_reverse()
```

Anexo VI

Este código em R realiza a normalização de dados acústicos por meio do cálculo do *z-score*. Ele começa listando os nomes das colunas do “dataframe” e, em seguida, define uma função chamada “*z_score*”, que padroniza os valores de uma variável subtraindo a média e dividindo pelo desvio-padrão – método para tornar os dados comparáveis em uma mesma escala.

Depois disso, o código cria uma cópia dos dados originais chamada “dados_norm” e aplica a função de “*z_score*” às colunas “duração.ms”, “*pitch*.Hz” e “intensidade.dB”, substituindo os valores originais por suas versões normalizadas. Por fim, ele exibe as primeiras linhas da nova base normalizada com “head()” e salva os resultados em um arquivo “.csv” chamado “dados_normalizados.csv”.

Script em R para a normalização *z-score* unidade de medida de duração (ms), *pitch* (Hz) e intensidade (dB):

```
names(dados)
# Função para calcular o Z-score
z_score <- function(x) {
  return((x - mean(x, na.rm = TRUE)) / sd(x, na.rm = TRUE))
}

# Criar um novo dataframe com os valores normalizados
dados_norm <- dados # Copiar os dados originais
dados_norm$duração.ms <- z_score(dados$duração.ms)
dados_norm$pitch.Hz <- z_score(dados$pitch.Hz)
dados_norm$intensidade.dB <- z_score(dados$intensidade.dB)

# Ver os primeiros resultados
head(dados_norm)
write.csv(dados_norm, "dados_normalizados.csv", row.names = FALSE, quote = FALSE, dec = ".")
```