

**UNIVERSIDADE FEDERAL DE PELOTAS**

**Centro de Letras e Comunicação**

**Programa de Pós-Graduação em Letras**



**Dissertação**

**Emergência dos gestos articulatórios na aquisição do *tap* em sílaba CVC:  
aspectos acústicos**

**Patrícia Pereira Melcheque**

Pelotas, 2021

**Patrícia Pereira Melcheque**

**Emergência dos gestos articulatórios na aquisição do *tap* em sílaba CVC:  
aspectos acústicos**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal de Pelotas como requisito parcial à obtenção do título de Mestre em Letras.

Orientador: Profa. Dr. Giovana Ferreira Gonçalves

Pelotas, 2021

Universidade Federal de Pelotas / Sistema de Bibliotecas  
Catalogação na Publicação

M518e Melcheque, Patrícia Pereira

Emergência dos gestos articulatórios na aquisição do *tap* em sílaba CVC : aspectos acústicos / Patrícia Pereira Melcheque ; Giovana Ferreira-Gonçalves, orientadora. — Pelotas, 2021.

254 f. : il.

Dissertação (Mestrado) — Programa de Pós-Graduação em Letras, Centro de Letras e Comunicação, Universidade Federal de Pelotas, 2021.

1. Aquisição. 2. *Tap*. 3. Contrastes encobertos. 4. Fonologia articulatória. I. Ferreira-Gonçalves, Giovana, orient. II. Título.

CDD : 418

**Patricia Pereira Melcheque**

**“EMERGÊNCIA DOS GESTOS ARTICULATÓRIOS NA AQUISIÇÃO DO TAP  
EMSÍLABA CVC: ASPECTOS ACÚSTICOS”**

Dissertação aprovada, como requisito parcial, para obtenção do grau de Mestre em Letras, Área de Concentração Linguagem, Texto e Imagem, do programa de Pós-Graduação em Letras, da Universidade Federal de Pelotas.

Pelotas, 20 de agosto de 2021.

Banca examinadora:



---

Profa. Dra. Giovana Ferreira Gonçalves  
Orientadora/Presidente da banca  
Universidade Federal de Pelotas



---

Profa. Dra. Márcia Keske Soares  
Membro da Banca  
Universidade Federal de Santa Maria



---

Profa. Dra. Mirian Rose Brum de Paula  
Membro da banca  
Universidade Federal de Pelotas

Ao meu filho Douglas, razão da minha vida.

## **Agradecimentos**

Uma pesquisa de mestrado é uma jornada longa que inclui um caminho permeado por inúmeras adversidades, incertezas, alegrias, tristezas e muitos percalços durante o seu desenvolvimento. No entanto, não passei por esse processo de forma solitária, pois estava rodeada de pessoas especiais que me deram muito apoio, energia e força quando eu achava que não teria forças para continuar. Em um ano atípico, perpassado por uma pandemia, o pessimismo tornou-se um inimigo a ser derrotado. Por conta disso, eu gostaria de agradecer às pessoas que me acompanharam e foram fundamentais para que eu não desistisse.

À minha orientadora, Profa. Giovana Ferreira Gonçalves, por sempre ter acreditado e depositado sua confiança no meu trabalho. Agradeço pelas orientações pautadas por um elevado e rigoroso nível científico e por uma visão crítica e exigente que contribuiu para enriquecer todas as etapas deste trabalho. Agradeço também pela generosidade, carinho e paciência que sempre teve comigo. Se hoje eu cheguei ao fim deste curso é porque a senhora sempre me guiou pelo melhor caminho.

À professora Mirian Rose Brum de Paula, que me apresentou à área de fonética e fonologia quando abriu as portas do LELO para mim, em 2015, oportunizando o trabalho com a minha primeira bolsa de iniciação científica. Fazer parte desse grupo é um orgulho e uma honra para mim!

À professora Márcia Keske Soares, por fazer parte da banca de qualificação e de defesa desta pesquisa. Obrigada pela leitura atenciosa e por ter trazido contribuições valiosas a esta dissertação.

Ao professor Felipe Bilharva da Silva, por gentilmente ter aceitado compor a banca de qualificação. Suas sugestões pertinentes e precisas enriqueceram o meu texto de uma forma grandiosa.

Ao meu marido, Felipe Vieira Melcheque, pelo apoio incondicional, amor, carinho e companheirismo. Agradeço pelo cuidado, por estar ao meu lado o tempo todo e por me amar sempre. O sentimento que nos une é mais do que um feliz acaso, é uma história de amor desde a nossa adolescência. A ti, meu amor, minha eterna gratidão. Eu te amo!

Ao meu filho, Douglas Pereira Melcheque, que amo incondicionalmente. Sempre agradeço a Deus por tê-lo colocado em meu caminho, me dando motivos para

acreditar na vida e nos sonhos. Espero poder compensar as horas de atenção, brincadeiras e momentos em família que lhe devo. Tu foste o meu maior estímulo durante esta caminhada.

Aos meus pais, Jozi dos Santos Bilhalva e Gilmar Garcia Pereira, por serem responsáveis pela minha educação e por sempre me incentivarem a estudar. Hoje, meu interesse pelo conhecimento vem de vocês.

À minha irmã, Raquel Pereira Villela, pela nossa convivência durante a infância, por sempre acreditar em mim e me apoiar em todos os momentos delicados pelos quais passei. Obrigada pela vida do meu sobrinho, Pedro Henrique, que amo como se fosse meu filho.

Aos meus irmãos, Pedro Gomes Pereira e Daniel Gomes Pereira. Que vocês sigam sempre a trilha dos estudos e do conhecimento. À minha madrasta Kelly Gomes Pereira, por ser uma pessoa muito especial e amiga.

Aos meus sogros, Gladis Vieira Melcheque e Felizardo da Silva Melcheque, por serem a minha rede de apoio e cuidarem tão bem do meu filho enquanto eu estava em aula ou no laboratório. Em vocês, sempre tive o apoio de verdadeiros pais.

Aos meus avós, Maria Clara (in memorian), Eromeu, Celi e Enildo (in memorian) que foram o meu porto seguro e sempre tiveram orgulho de mim, mesmo quando eu ainda nada havia conquistado.

Aos amigos que fiz no LELO durante essa caminhada, Bruna, Laís, Brenda, Misael e Thalena. Agradeço pela parceria, companheirismo e pelo apoio motivacional durante todo o desenvolvimento deste trabalho. Vocês fizeram esta experiência ser mais leve e divertida. Obrigada por todas as risadas, conselhos, pomodoros, reuniões de estudo, ainda que virtualmente, na nossa sala FUZ, durante este último ano.

Aos amigos que fiz durante o curso de Letras, Carlos Ossanes, Amanda Alves, Gilson Lopes, Ana Cláudia, Bruna Cavalheiro, Paola, Aline e Cristiane. Obrigada por compartilharem comigo conhecimentos e sorrisos.

À minha amiga de infância Daniela e aos meus padrinhos Glaci e Vicente, pois, mesmo distantes, sempre estiveram presentes no meu coração. Vocês são parte da minha família.

Aos meus padrinhos, Enilda e Genecildo, pelo apoio, carinho, cuidado e pelo colo que, por muitas vezes, acolheu as minhas angústias.

À minha psicóloga Daniela Figueiredo que foi fundamental para me ajudar a conciliar a vida de mãe, filha, esposa, dona de casa e pesquisadora. Graças a ela superei medos e inseguranças, tornando possível, assim, que eu concluisse esta pesquisa dando o melhor de mim.

Aos meus pequenos informantes e seus pais, pela sua disposição, confiança e paciência durante as coletas de dados.

À CAPES, pela bolsa de mestrado que possibilitou minha dedicação exclusiva a este trabalho.

Aos meus familiares, amigos e compadres que compreenderam a minha ausência, mas sempre torceram por mim e me enviaram energias positivas.

Aos meus afilhados Matheus, Yan, Otávio, Arthur, Lívia e Pedro Henrique. Eu amo muito vocês!

A Deus pelo dom da vida e pelo Seu amor infinito. Obrigada por me permitir trilhar o meu caminho nesta vida ao lado de pessoas tão maravilhosas.

Por fim, a todos aqueles que, mesmo não citados aqui, contribuíram, direta ou indiretamente, para a realização desta dissertação, o meu sincero agradecimento.

E que venham novos sonhos e novos obstáculos a superar.

*“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes”.*

*(Martin Luther King)*

## Resumo

MELCHEQUE, Patrícia Pereira. **Emergência dos gestos articulatórios na aquisição do *tap* em sílaba CVC: aspectos acústicos**. Dissertação (Mestrado em Letras). Programa de Pós-Graduação em Letras, Universidade Federal de Pelotas, 2021.

O presente trabalho busca investigar o processo de aquisição do *tap* em sílaba CVC, em posições medial e final de palavra, com base na Fonologia Articulatória (BROWMAN e GOLDSTEIN, 1986). O [r] é um dos últimos segmentos a surgir na fala da criança, apresentando diversas alterações durante o seu período de aquisição. Trabalhos de base gerativa analisaram a aquisição de estruturas silábicas e dos fonemas no português brasileiro. Apesar das contribuições trazidas por essas pesquisas, o detalhe fonético, que não é percebido auditivamente, não era considerado como pista para a marcação de possível contraste fonológico. Trabalhos mais recentes, de base emergentista, como Rodrigues (2007), consideram o gesto articulatório como primitivo de análise. Assim, é possível representar o detalhe fonético no nível gramatical. Neste sentido, o presente estudo investiga a presença de contrastes encobertos nas produções do *tap* na aquisição fonético-fonológica do português brasileiro. Foram analisados os dados orais de um sujeito controle, de quatro sujeitos transversais e de um sujeito longitudinal. As coletas foram gravadas com um gravador digital *Zoom H4n* em uma cabine acústica no Laboratório Emergência da Linguagem Oral (UFPel). O rótico se encontrava em sílaba tônica, em contexto de [a], [i] e [u], em posição medial e final de palavra. Foi elaborado um *corpus* de tríades de palavras, com a presença do rótico, sua omissão e a presença do ditongo palatal, apresentado por meio de um instrumento de nomeação de imagens. A análise acústica revelou 28 possíveis contrastes encobertos para 129 possibilidades de produção do *tap* na fala dos informantes transversais. Na fala do informante longitudinal, foi identificada a presença de 51 possíveis contrastes encobertos nas 142 possibilidades de ocorrência do segmento. A variante rótica mais empregada na fala das crianças foi o *tap*, porém, também foram identificadas produções de vibrante, aproximante retroflexa, aproximante alveolar, fricativa glotal e *tap* fricatizado. Ao não produzirem o *tap*, foi possível inferir que as crianças utilizaram principalmente estratégias de redução gestual e substituição de gestos, ainda que gestos intrusivos também tenham ocorrido. A inspeção visual dos espectrogramas e a análise da trajetória formântica foram fundamentais para evidenciar possíveis contrastes encobertos em andamento na fala das crianças. Ao observar o comportamento dos elementos vocálicos, a partir da comparação entre esses e a porção final de uma vogal nuclear, constatou-se que apenas os elementos vocálicos produzidos em contexto de [u] tendem à centralização. Tais resultados indicam, portanto, uma relação entre o elemento vocálico e a vogal nuclear em sílaba CVC. Sendo assim, espera-se que este trabalho tenha contribuído para evidenciar que produções gradientes, previstas pela Fonologia Articulatória, são detectadas no processo de emergências dos segmentos róticos. A análise dos dados, sob a perspectiva de um modelo teórico que considere o caráter dinâmico da linguagem, viabiliza a interpretação de produções gradientes como pistas para a emergência da aquisição do sistema de contrastes da língua.

Palavras-chave: Aquisição. *Tap*. Contrastos encobertos. Fonologia Articulatória

## Abstract

MELCHEQUE, Patrícia Pereira. **Emergence of articulatory gestures in the acquisition of tap in CVC syllable: acoustic aspects.** Dissertation (Masters in Languages). Postgraduate Program in Languages, Federal University of Pelotas, 2021.

The present work seeks to investigate the process of acquisition of tap in CVC syllable, in medial and final position of word, based on Articulatory Phonology. The [r] is one of the last segments to emerge in the child speech, presenting several changes during their acquisition period. Generative-based studies analyzed the acquisition of syllabic structures and phonemes in Brazilian Portuguese. Despite the contributions brought by those studies, the phonetic detail, which is not perceived auditorily, was not considered as a clue for the markedness of potential phonological contrast. More recent works, emergentist-based, such as Rodrigues (2007), consider articulatory gesture as primitive analysis. Thus, it is possible to represent phonetic detail at the grammatical level. In this sense, the present study investigates the presence of covert contrasts on tap productions in the phonetic-phonological acquisition of Brazilian Portuguese. Oral data productions of one control subject, four cross-sectional subjects and a longitudinal subject were analyzed. The data collections were recorded with a Zoom H4n digital recorder in an acoustic booth at the Emergency of Oral Language Laboratory (UFPel). The rhotic was in stressed syllable, in the context of [a], [i] and [u], in medial and final position of the word. A corpus of word triads was elaborated, with the presence of the rhotic, its omission and the presence of the palatal diphthong, presented by means of an instrument for naming images. Acoustic analysis revealed 28 possible contrasts for 129 tap production possibilities in the speech of cross-sectional informants. In the longitudinal informant's speech, the presence of 51 possible covert contrasts in 142 possibilities of occurrence of the segment was identified. The rhotic variant most used in children speech was the tap, but productions of trill, retroflex approximant, alveolar approximant, glottal fricative and fricatized tap were also identified. When not being able to produce the tap, it was possible to infer that the children used strategies of gestural reduction, gesture substitution, while intrusive gestures may also have occurred. Visual inspection of spectrograms and analysis of the formant trajectory were fundamental to evidence possible covert contrasts in progress on children's speech. When observing the behavior of vocalic elements from the comparison between them and the final portion of a nuclear vowel, it was found that only vocalic elements produced in the context of [u] tend to centralize. Thus, the results indicate a relation between the vocalic element and the nuclear vowel in the CVC syllable. Thus, it is expected that this work has contributed to evidence that gradient productions, expected by Articulatory Phonology, are detected in the emergence process of rhotic segments. The analysis of the data, from the perspective of a theoretical model that considers the dynamic character of the language, enables the interpretation of gradient productions as clues to the emergence of the acquisition of the language contrast system.

Keywords: Acquisition. Tap. Covert Contrasts. Articulatory Phonology

## Lista de Figuras

Figura 1 - Comparação entre a vibrante em [ərə] e o <i>tap</i> em [ərə]. A vibrante apresenta várias vibrações, o <i>tap</i> apenas uma. Fonte: THOMAS (2011, p.130) .....	38
Figura 2 – Momento de máxima constrição do gesto de ponta de língua para o <i>tap</i> [r]. Fonte: Adaptado do site <i>fonologia.org</i> de CRISTÓFARO-SILVA; YEHIA (2009) .....	39
Figura 3 - <i>Frames</i> com a trajetória da ponta da língua na sequência [ara]. Fonte: a autora .....	40
Figura 4 - Dados de ultrassom para a sequência /iri/ produzida por um falante de Catalão. Fonte: RECASENS (2016, p. 36) .....	41
Figura 5 - Forma de onda e espectrograma da plosiva, vogal e <i>tap</i> das palavras <i>carta</i> e <i>par</i> . Fonte: adaptado de SILVA et al. (2019, p. 196) .....	42
Figura 6 - Teste-T do par de médias das produções do elemento vocálico em CCV (traçados pretos) e dos 25% da fase inicial da vogal nuclear em CV (traçados vermelhos). Linha preta superior apresenta a <i>outlier da estatística</i> . Fonte: MELCHEQUE; FERREIRA-GONÇALVES; BRUM-DE-PAULA (2021, p. 250) .....	44
Figura 7 - Teste-T do par de médias das produções do elemento vocálico em CVC e 25% da fase final da vogal nuclear em CV. À esquerda CV (traçados vermelhos) e CVC medial (traçados verdes); à direita CV (traçados vermelhos) e CVC final (traçados azuis). Caixas vermelhas indicam os pontos de diferenças com valores de <i>p</i> entre 0.01 e 0.05. Linha preta superior apresenta a <i>outlier da estatística</i> . Fonte: MELCHEQUE; FERREIRA-GONÇALVES; BRUM-DE-PAULA (2021, p. 251) .....	45
Figura 8 - Teste-T do par de médias das produções do elemento vocálico em CCV e em CVC. À esquerda, CCV (traçados pretos) e CVC medial (traçados verdes); à direita, CCV (traçados pretos) e CVC final (traçados azuis). Caixas vermelhas indicam os pontos de diferenças com valores de <i>p</i> entre 0.01 e 0.05. Linha preta superior apresenta a <i>outlier da estatística</i> . Fonte: MELCHEQUE; FERREIRA-GONÇALVES; BRUM-DE-PAULA (2018) .....	45
Figura 9 - Espaço acústico vocálico apresentando os valores de F1 e F2 da VA (vocal de apoio) e VN (vocal núcleo) seguidos das vogais [a], [i] e [u]. À esquerda dados do informante masculino e à direita dados da informante feminina. Fonte: adaptado de SILVEIRA; SEARA (2008, p. 44 e 45) .....	46
Figura 10 - Variáveis do trato e representação do trato vocal com a localização das respectivas variáveis. Fonte: BROWMAN; GOLDSTEIN (1989, p. 73) .....	64

Figura 11 - Pautas gestuais da palavra <i>palm</i> , do inglês. Fonte: BROWMAN; GOLDSTEIN (1989, p. 76) .....	66
Figura 12 - Pauta gestual para o <i>tap</i> em cudas. Fonte: SILVA; CLEMENTE; NISHIDA (2006, p. 22) .....	67
Figura 13 - Interior da cabine com isolamento acústico, utilizada para coletas acústicas e articulatórias, LELO/UFPel. Fonte: a autora .....	77
Figura 14 - Gravador <i>Zoom H4n</i> e cartões utilizados para o instrumento de coleta..	78
Figura 15 - Tela do formulário online com um exemplo de tarefa de identificação dos juízes.....	81
Figura 16 - Tela do software <i>Praat</i> em S0_circo. Fonte: a autora .....	83
Figura 17 - Tela do software <i>Excel</i> , contendo as observações da inspeção visual das produções em contexto de [u] final produzidos na segunda coleta da informante longitudinal. ....	85
Figura 18 - Exemplo dos gráficos de Trajetória Formântica de F1, F2 e F3 na vogal [a] que antecede o rótico na produção da palavra <i>carta</i> , por S0. Fonte: a autora.....	87
Figura 19 - Forma de onda e espectrograma de banda larga da produção de ['tur.ku] do alvo "turco" por S0. Em destaque pela seta a produção do <i>tap</i> . ....	107
Figura 20 - Forma de onda e espectrograma de banda larga da produção de ['kur.tə] do alvo "curta" por S0. Em destaque pela setas a produção da vibrante. ....	108
Figura 21 - Forma de onda e espectrograma de banda larga da produção de ['a.zar] do alvo "azar" por S0. Em destaque pelas setas a produção da vibrante. ....	109
Figura 22 - Forma de onda e espectrograma de banda larga da produção de ['lu.ga.ʃ] do alvo "lugar" por S0. Em destaque pelas setas a trajetória formântica. ....	110
Figura 23 - Forma de onda e espectrograma de banda larga da produção de ['si] do alvo "círculo" produzido na primeira coleta por SL1. As setas indicam a presença do contraste encoberto.....	137
Figura 24 – Forma de onda e espectrograma de banda larga da podução de [si], do alvo "circo" produzido na segunda coleta por SL1. As setas indicam a presença do contraste encoberto.....	138
Figura 25 - Forma de onda e espectrograma de banda larga da podução de [vi], do alvo "vir" produzido na segunda coleta por SL1. A seta indica a presença do contraste encoberto. ....	139

Figura 26 - Forma de onda e espectrograma de banda larga da produção de ['a.tuj] do alvo "Artur" produzido por SL1 na terceira coleta.....	140
Figura 27 - Forma de onda e espectrograma de banda larga da produção de ['kuj] do alvo "curta" por ST1. Em destaque pela seta o contraste encoberto .....	172
Figura 28 - Forma de onda e espectrograma de banda larga da produção de ['si] do alvo "círculo" por ST2. Em destaque pela seta o contraste encoberto .....	173
Figura 29 - Forma de onda e espectrograma de banda larga da produção de ['si] do alvo "circo" por ST3. Em destaque pela seta o contraste encoberto .....	175
Figura 30 - Forma de onda e espectrograma de banda larga da produção de ['ka.pe] do alvo "carpa" por ST4. Em destaque pela seta o contraste encoberto.....	176
Figura 31 - Forma de onda e espectrograma de banda larga da produção de ['ga] do alvo "lugar" por ST1. Em destaque pelas setas o contraste encoberto .....	177
Figura 32 - Forma de onda e espectrograma de banda larga da produção de ['tuɾ] do alvo "Artur" por ST3.....	178

## Lista de Gráficos

Gráfico 1 - Produção de 'r-fraco' versus posição silábica. Fonte: MIRANDA (1996, p. 102) .....	50
Gráfico 2 - Aquisição da líquida não-lateral em coda. Fonte: MEZZOMO (2004, p.145) .....	51
Gráfico 3 - Recursos utilizados na aquisição do /r/ em coda final. Fonte: MEZZOMO (2004, p. 119) .....	52
Gráfico 4 - a1, a2 e a3) intervalos de confiança para os contornos médios de língua; b1, b2 e b3) diferença nos contornos médios entre [r] e [l] para todas as palavras nos grupos AT, CT e CA. Em vermelho os traçados de língua para [r] e em azul os traçados para [l]. FONTE: BARBERENA (2016, p.118) .....	58
Gráfico 5 - a) intervalo de confiança para os contornos médios de língua para alvos [r] produzidos como [l] e para alvos [l] – CA; b) diferenças para os contornos médios entre esses sons. Em vermelho, os traçados para as substituições de [r] para [l] e, em azul, os traçados para as produções de [l]. FONTE: BARBERENA (2016, p. 158) .....	59
Gráfico 6: a) intervalo de confiança para os contornos médios de língua para alvos [r] produzidos como [j] e para alvos [j] – CA e b) diferenças para os contornos médios entre esses sons. Em vermelho, os traçados para as substituições de [r] para [j] e, em azul, os traçados para as produções de [j]. FONTE: BARBERENA (2016, p. 159)...	60
Gráfico 7 - Índices de produções do <i>tap</i> e recursos utilizados por SL1, em posição medial, de acordo com a ordem de coleta.....	92
Gráfico 8 - Índices de produções do <i>tap</i> e recursos utilizados por SL1, em posição final, de acordo com a ordem de coleta.....	93
Gráfico 9 - Médias das produções do <i>tap</i> e recursos utilizados nas três coletas por SL1 para posição medial (à esquerda) e para posição final (à direita).....	94
Gráfico 10 - Índices de produção do <i>tap</i> e recursos utilizados por ST1 para posição medial (azul) e para posição final (laranja).....	96
Gráfico 11 - Índices de produção do <i>tap</i> e recursos utilizados por ST2 para posição medial (azul) e para posição final (laranja).....	97
Gráfico 12 - Índices das produções do <i>tap</i> e recursos utilizados por ST3 para posição medial (azul) e para posição final (laranja).....	98
Gráfico 13 - Índices de produção do <i>tap</i> e recursos utilizados por ST4 para posição medial (azul) e para posição final (laranja).....	99

Gráfico 14 - Médias das produções do <i>tap</i> e recursos utilizados pelos informantes transversais para posição medial (à esquerda) e para posição final (à direita) .....	100
Gráfico 15 - Médias das produções do <i>tap</i> e recursos utilizados pelo informante longitudinal (roxo) e pelos informantes transversais (verde) para posição medial ..	101
Gráfico 16 - Médias das produções do <i>tap</i> e recursos utilizados pelo informante longitudinal (roxo) e pelos informantes transversais (verde) para posição final .....	102
Gráfico 17 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico articulado como <i>tap</i> (verde) das produções de S0, para a posição medial ...	113
Gráfico 18 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico articulado como vibrante (verde) das produções de S0, para a posição medial .....	114
Gráfico 19 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico articulado como vibrante (verde) das produções de S0, para a posição final	116
Gráfico 20 – Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [a] que antecede o rótico na produção de ['kar.tə] por S0 .....	118
Gráfico 21 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [a] de ['ka.tə] por S0.....	119
Gráfico 22 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) do ditongo [aj] na produção de ['kaj.ʃə] por S0 .....	119
Gráfico 23 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] que antecede o rótico na produção de ['sir.ku] por S0 .....	121
Gráfico 24 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] de ['ʃi.ku] por S0 .....	122
Gráfico 25 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u] que antecede o rótico na produção de ['kur.tu] por S0.....	124
Gráfico 26 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u] na produção de ['ku.kə] por S0 .....	124
Gráfico 27 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) do ditongo [uj] na produção de ['kuj.də] por S0 .....	125
Gráfico 28 - Espaço acústico de [a], [i] e [u] para a vogal nuclear (preto) e o elemento vocálico (vermelho) produzidos por S0 .....	128

Gráfico 29 - Espaço acústico de [a], [i] e [u] para o ponto médio da porção final vogal CV (preto) e elemento vocálico produzidos por S0 – contexto de [a] (azul), [i] (verde) e [u] (amarelo) .....	131
Gráfico 30 - Espaço acústico de [a], [i] e [u] para o ponto médio da porção final vogal CV (preto) e para a média do ponto médio do elemento vocálico produzidos por S0 – contexto de [a] (azul), [i] (verde) e [u] (amarelo).....	132
Gráfico 31 – Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1, para a posição medial .....	144
Gráfico 32 – Médias de duração relativa da vogal em contexto de [a] (laranja), [i] (cinza) e [u] (verde) em produções de SL1 para a posição medial .....	145
Gráfico 33 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita) em alvos CVC medial .....	146
Gráfico 34 - Médias de duração absoluta da sílaba (azul) e da vogal/ditongo (vermelho) das produções de SL1, para a posição medial.....	149
Gráfico 35 - Médias de duração relativa do ditongo em contexto de [a] (laranja) e [u] (cinza) produzido por SL1, para a posição medial .....	150
Gráfico 36 - Médias de duração relativa da vogal/ditongo em contexto de [a] (laranja) e [u] (cinza) produzidos por S0, para a posição medial.....	151
Gráfico 37 - Médias de duração absoluta da sílaba (azul) e do ditongo/vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial .....	152
Gráfico 38 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1, para a posição final .....	155
Gráfico 39 - Médias de duração relativa da vogal em contexto de [a] (laranja), [i] (cinza) e [u] (verde) produzida por SL1, para a posição final.....	156
Gráfico 40 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita), para a posição final .....	157
Gráfico 41 - Médias de duração absoluta da sílaba (azul) e da vogal/ditongo (vermelho) das produções de SL1, para a posição final.....	160

Gráfico 42 - Médias de duração relativa de produções semivocalizadas em contexto de AR e IR (laranja) e de produções do ditongo [aj] (cinza) produzido por SL1, para a posição final .....	161
Gráfico 43 - Médias de duração absoluta da sílaba (azul) e do ditongo/vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita), para a posição final .....	162
Gráfico 44 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u], resultante da omissão do rótico em ['ku.tu] da segunda coleta de SL1 .....	164
Gráfico 45 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u], em contexto CV, da produção de ['ku.tu], da segunda coleta de SL1.164	
Gráfico 46 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i], resultante da omissão do rótico em ['si.ku.lu] da segunda coleta de SL1 .....	165
Gráfico 47 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i], em contexto CV, da produção de ['si.ku], da segunda coleta de SL1 ...166	
Gráfico 48 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de S0, ST2 e ST3 em contexto vocálico de AR, para a posição medial .....	181
Gráfico 49 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de S0, ST2 e ST3 em contexto vocálico de UR, para a posição medial .....	182
Gráfico 50 - Médias de duração relativa do rótico de S0 (preto), ST2 (cinza) e ST3 (vermelho), em contexto de AR e UR, para posição medial.....	183
Gráfico 51 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2 e ST4, em contexto de AR (com omissão do <i>tap</i> ) e A, para a posição medial .....	187
Gráfico 52 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja) e ST4 (verde), em contexto de AR e A, para posição medial .....	188
Gráfico 53 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2, ST3 e ST4, em contexto de IR (com a omissão do <i>tap</i> ) e I, para a posição medial .....	189

Gráfico 54 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja), ST3 (cinza) e ST4 (verde), em contexto de IR (com omissão do <i>tap</i> ) e I, para posição medial .....	190
Gráfico 55 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2 e ST4, em contexto de UR (com omissão do <i>tap</i> ) e U, para a posição medial .....	191
Gráfico 56 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja) e ST4 (verde), em contexto de UR (com omissão do <i>tap</i> ) e U, para posição medial .....	192
Gráfico 57 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de AR, IR e UR (com omissão do <i>tap</i> ), das produções de ST1 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial .....	193
Gráfico 58 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de AR, IR e UR, (com omissão do <i>tap</i> ) das produções de ST2 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial .....	194
Gráfico 59 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de IR, das produções de ST3 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial .....	195
Gráfico 60 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de AR, IR e UR (com omissão do <i>tap</i> ), das produções de ST4 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial .....	196
Gráfico 61 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de AR e AI, para a posição medial .....	199
Gráfico 62 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de AR e AI, para posição medial .....	200
Gráfico 63 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de UR e UI, para a posição medial .....	201
Gráfico 64 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de UR e UI, para posição medial .....	202
Gráfico 65 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de ST2, para a posição final .....	205
Gráfico 66 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de ST3, para a posição final .....	206

Gráfico 67 - Médias de duração relativa do rótico de ST2 (laranja) e ST3 (cinza), em contexto de AR, IR e UR, para posição final .....	206
Gráfico 68 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de AR e A, para a posição final .....	210
Gráfico 69 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de AR e A, para posição final.....	211
Gráfico 70 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de IR e I, para a posição final.....	212
Gráfico 71 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de IR e I, para posição final.....	213
Gráfico 72 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2, ST3 e ST4, em contexto de UR e U, para a posição final	214
Gráfico 73 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja), ST3 (cinza) e ST4 (verde), em contexto de UR e U, para posição final .....	215

## **Lista de Quadros**

Quadro 1 - Listas de palavras utilizadas no estudo de Barberena (2016) .....	57
Quadro 2 - Conjuntos de articuladores e parâmetros dinâmicos associados, conforme proposta de Browman e Goldstein (1989, p. 74). Fonte: adaptado de BARBIERI (2019, p. 45) .....	65
Quadro 3- Relação das palavras utilizadas no instrumento da entrevista .....	73
Quadro 4 - Coletas realizadas com o sujeito longitudinal.....	74
Quadro 5 - Coletas com os sujeitos transversais. .....	75
Quadro 6 - Ocorrência do <i>tap</i> e recursos utilizados pelos informantes no dia da entrevista. As produções estão divididas por informantes e por posição silábica .....	76
Quadro 7 - Coleta com o sujeito controle. .....	76
Quadro 8 - Relação dos pares de palavras utilizadas no instrumento de coleta .....	78
Quadro 9 - Relação de autores e de palavras utilizadas no instrumento de coleta...	80

## Lista de Tabelas

Tabela 1 - Índices de produções do <i>tap</i> e recursos utilizados por SL1, de acordo com a ordem de coleta e a posição do <i>tap</i> na sílaba, médias e desvio padrão .....	91
Tabela 2 - Índices de produções do <i>tap</i> e recursos utilizados por ST1, ST2, ST3 e ST4, de acordo com a posição da sílaba, médias e desvio padrão .....	95
Tabela 3 - Produção das variantes róticas de acordo com o ambiente vocálico e a posição na palavra por S0.....	106
Tabela 4 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como <i>tap</i> ) produzidos pelo sujeito controle em posição medial, conforme contexto vocálico.....	112
Tabela 5 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como vibrante) em posição medial do sujeito controle, conforme contexto vocálico .....	114
Tabela 6 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como vibrante) em posição final produzidos pelo sujeito controle, conforme contexto vocálico.....	116
Tabela 7 - Médias e desvio-padrão para os valores de F1, F2 e F3 nos trechos de início, meio e fim da trajetória formântica de [a], em contexto [ar], [a] e de [aj], nas produções de S0 .....	120
Tabela 8 - Médias e desvio-padrão para os valores de F1, F2 e F3 nos trechos de início, meio e fim da trajetória formântica de [i], em contexto [ir] e [i], nas produções de S0 .....	123
Tabela 9 - Médias e desvio-padrão para os valores de F1, F2 e F3 nos trechos de início, meio e fim da trajetória formântica de [u], em contexto [ur], [u] e de [uj], nas produções de S0 .....	126
Tabela 10 - Médias gerais de F1 e F2 do ponto médio da vogal nuclear e do elemento vocálico, em contexto de [a], [i] e [u], produzidos por S0 .....	128
Tabela 11 - Diferenças dos valores médios de F1 e F2 entre as vogais nucleares (VN) e os elementos vocálicos (EV), produzidos por S0 .....	129
Tabela 12 – Proporção do elemento vocálico em relação à vogal nuclear CV, por S0 .....	130

Tabela 13 – Ponto médio de F1 e F2 das vogais protótipos (extraídos da porção final e dos elementos vocálicos produzidos por S0 .....	130
Tabela 14 – Possibilidades e ocorrências de contrastes encobertos a partir das características acústicas do <i>tap</i> , sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala da informante longitudinal, em CVC medial.....	134
Tabela 15 – Possibilidades e ocorrências de contrastes encobertos a partir das características acústicas do <i>tap</i> , sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala da informante longitudinal, para CVC final .....	135
Tabela 16 – Possibilidades e ocorrências de prováveis contrastes encobertos na fala da informante longitudinal, de acordo com a posição silábica e o ambiente vocálico em posições medial e final .....	136
Tabela 17 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos, com o recurso de omissão, pela informante longitudinal, conforme contexto vocálico .....	142
Tabela 18 - Médias de duração absoluta e relativa dos alvos CV em posição medial da informante longitudinal, conforme contexto vocálico .....	143
Tabela 19 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos, com o recurso da semivocalização, pela informante longitudinal, conforme contexto vocálico .....	147
Tabela 20 - Médias de duração absoluta e relativa dos alvos CVV que contêm ditongo em posição medial produzidos pela informante longitudinal, conforme contexto vocálico .....	148
Tabela 21 - Médias de duração absoluta e relativa da sílaba e da vogal dos alvos CVC em posição final, produzidos, com o recurso de omissão, pela informante longitudinal, conforme contexto vocálico .....	153
Tabela 22 - Médias de duração absoluta e relativa da sílaba e da vogal dos alvos CV em posição final produzidos pela informante longitudinal, conforme contexto vocálico .....	154
Tabela 23 - Médias de duração absoluta e relativa dos alvos CVC em posição final, produzidos, com o recurso da semivocalização, pela informante longitudinal, conforme contexto vocálico .....	158

Tabela 24 - Médias de duração absoluta e relativa dos alvos CVV em posição final produzido pela informante longitudinal, conforme contexto vocálico.....	159
Tabela 25 - Produção das variantes róticas de acordo com o ambiente vocálico e posição na palavra para cada informante transversal .....	168
Tabela 26 - Possibilidades e ocorrências de contrastes encobertos a partir das características do <i>tap</i> , sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala dos sujeitos transversais, para posição medial .....	169
Tabela 27 - Possibilidades e ocorrências de contrastes encobertos a partir das características do <i>tap</i> , sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala dos sujeitos transversais, para posição final .....	170
Tabela 28 - Lista de possibilidades e ocorrências de prováveis contrastes encobertos na fala dos sujeitos transversais, de acordo com a posição silábica e o contexto silábico em posições medial e final .....	171
Tabela 29 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como <i>tap</i> ) em posição medial dos sujeitos transversais, conforme contexto vocálico .....	180
Tabela 30 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos com o recurso de omissão pelos sujeitos transversais, conforme contexto vocálico .....	184
Tabela 31- Médias de duração absoluta e relativa da vogal dos alvos CV em posição medial produzidos pelos sujeitos transversais, conforme contexto vocálico .....	185
Tabela 32 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos com o recurso da semivocalização pelos sujeitos transversais, conforme contexto vocálico .....	197
Tabela 33 - Médias de duração absoluta e relativa os alvos CVV que contêm ditongo em posição medial produzidos pelos sujeitos transversais, conforme contexto vocálico .....	198
Tabela 34 - Médias de duração absoluta e relativa dos alvos que contêm rótico (articulados como <i>tap</i> ) em posição final produzidos pelos sujeitos transversais, conforme contexto vocálico .....	203

Tabela 35 - Médias de duração absoluta e relativa dos alvos CVC em posição final, produzidos, com o recurso da omissão, pelos sujeitos transversais, conforme contexto vocálico .....	208
Tabela 36 - Médias de duração absoluta e relativa dos alvos CV em posição final dos sujeitos transversais, conforme contexto vocálico.....	209
Tabela 37- Médias de duração absoluta e relativa dos alvos CVC em posição final, produzidos com o recurso da semivocalização, pelos sujeitos transversais, conforme contexto vocálico .....	216

## **Lista de abreviaturas e siglas**

AT	adultos típicos
C <sub>1</sub>	primeira consoante do encontro consonantal
C <sub>2</sub>	segunda consoante do encontro consonantal
CA	crianças atípicas
CCV	consoante – consoante – vogal
CT	crianças típicas
CV	consoante – vogal
CVC	consoante – vogal – consoante
CVV	consoante – vogal – vogal
DP	desvio padrão
EV	elemento vocálico
F1	primeiro formante
F2	segundo formante
F3	terceiro formante
FAAR	fonologia acústico-articulatória
FAR	fonologia articulatória
FonGest	Fonologia Gestual
HI	hipótese de investigação
Hz	hertz
ms	milissegundos
PB	português brasileiro
S <sub>0</sub>	sujeito controle
SL	sujeito longitudinal
ST	sujeito transversal
TCLE	Termo de Consentimento Livre e Esclarecido
V	vogal
VA	vogal de apoio
VN	vogal nuclear
VS	vogal suarabácti

## Sumário

<b>1 Introdução .....</b>	<b>29</b>
<b>2 Fundamentação Teórica .....</b>	<b>35</b>
2.1 Descrição acústica e articulatória do <i>tap</i> .....	35
2.2 Aquisição do <i>tap</i> do Português Brasileiro.....	48
2.3 Fonologia Articulatória.....	62
<b>3 Metodologia .....</b>	<b>71</b>
3.1 Caracterização da amostra .....	71
3.2 Os sujeitos.....	73
3.2.1 Sujeito longitudinal .....	74
3.2.2 Sujeitos transversais .....	75
3.2.3 Sujeito controle.....	76
3.3 Coletas de dados.....	77
3.4 Análise de oitiva .....	80
3.5 Análise acústica .....	82
3.5.1 Organização e segmentação dos dados acústicos .....	82
3.5.2 Critérios de análise acústica.....	84
<b>4 Resultados e Discussão .....</b>	<b>90</b>
4.1 Análise de oitiva .....	90
4.1.1 Sujeito longitudinal .....	90
4.1.2 Sujeitos transversais .....	94
4.1.3 Comparação entre dados longitudinais e transversais .....	100
4.2 Análise acústica .....	105
4.2.1 Sujeito controle.....	105
4.2.1.1 Inspeção visual do espectrograma.....	105
4.2.1.2 Aspectos duracionais .....	111
4.2.1.3 Trajetória formântica.....	117
4.2.1.4 Natureza do elemento vocálico .....	127
4.2.2 Sujeito longitudinal .....	133
4.2.2.1 Inspeção visual do espectrograma.....	133
4.2.2.2 Aspectos duracionais .....	141
4.2.2.3 Trajetória Formântica .....	163
4.2.3 Sujeitos Transversais .....	167

4.2.3.1 Inspeção Visual do Espectrograma .....	167
4.2.3.2 Aspectos Duracionais.....	179
4.2.3.3 Trajetória formântica.....	218
<b>5 Conclusão .....</b>	<b>226</b>
<b>Referências Bibliográficas .....</b>	<b>232</b>
<b>Apêndices .....</b>	<b>241</b>
<b>Anexos .....</b>	<b>247</b>

## 1 Introdução

Na classe dos róticos, são encontrados alguns dos segmentos que apresentam maior dificuldade de produção por parte das crianças e, consequentemente, são de aquisição mais tardia. Por esse motivo, são recorrentes as alterações na produção ao longo do processo de aquisição. Tais alterações não ocorrem de forma aleatória, mas distribuem-se de forma sistemática nas produções da criança. Sendo assim, é importante buscar um padrão acústico e articulatório dos segmentos róticos<sup>1</sup> para que seja possível compreender como esses segmentos se estabilizam na fala da criança e como essas alterações contribuem para o surgimento dessa classe de sons. Conforme Goldstein *et al.* (2007), essas alterações, descritas categoricamente como apagamentos ou substituições, na verdade, podem revelar que unidades menores que os segmentos atuam durante essa etapa de produção das crianças.

O presente trabalho tem como foco principal a investigação da emergência dos gestos articulatórios na aquisição do segmento *tap* em sílaba CVC, em posição medial e final de palavra, por cinco crianças em desenvolvimento fonológico típico de fala.

Com base em princípios emergentistas, optou-se por investigar o desenvolvimento linguístico, a partir de dados de quatro sujeitos transversais e um longitudinal, de aquisição do Português Brasileiro (PB) como língua materna. Dados longitudinais viabilizam o acompanhamento do desenvolvimento da linguagem da criança no transcorrer de um período específico, possibilitando que as etapas de aquisição de um determinado aspecto linguístico sejam reveladas mais claramente. Isso permite compreender, por exemplo, como se dá a emergência dos gestos articulatórios ao longo do processo de aquisição. A fim de comparar as produções dos sujeitos em período de aquisição, dados de um sujeito controle também foram coletados. Sendo assim, a presente pesquisa conta com a participação de seis informantes.

Dentre os diferentes sistemas linguísticos, o rótico tem sido reportado como uma das consoantes com o domínio mais tardio e complexo pela criança. Esse fato é evidenciado pela maior parte dos trabalhos. Nesse sentido, Lindau (1985) aponta que

---

<sup>1</sup> Na classe dos róticos há dois segmentos do ponto de vista fonológico, distintos tanto em aspectos acústicos quanto articulatórios: a fricativa velar e o *tap* alveolar. É importante ressaltar que esses dois segmentos estabelecem contrastes lexicais em português. No entanto, com base no foco do presente estudo, apresentado mais adiante, a palavra *rótico* será utilizada para fazer referência aos sons articulados como *tap* alveolar e vibrante múltipla, não incluindo os sons fricativos.

um dos principais motivos que podem contribuir para a complexidade fonológica dos segmentos róticos é a grande variação fonética que existe entre esses sons nas diferentes línguas do mundo. Maddieson (1984) afirma que 75% dos idiomas existentes apresenta um som de *r*, embora em apenas 18% haja atividade contrastiva entre dois ou três róticos.

Jakobson (1968) afirma que, em várias línguas, o *tap* é o último segmento a ser adquirido na fala infantil e que, em casos de afasia, é o primeiro a ser perdido. Goldstein & Cintrón (2001), em seu estudo sobre o Espanhol Porto-riquenho, afirmam que as nasais são as primeiras consoantes em final de sílaba a serem adquiridas e os róticos<sup>2</sup> são as últimas. Os três informantes do estudo tinham idades entre 1:10 e 2:5 e produziram o *tap* e a vibrante de forma menos precisa. Entre esses dois segmentos, levando em consideração todas as posições silábicas, o *tap* foi o mais omitido, em 85% dos casos. Salidis e Johnson (1997), em um estudo longitudinal, acompanharam uma criança, falante nativa do inglês, dos 10 aos 20 meses de idade. Os autores observaram que, aos 11 meses, a maior parte da produção dos alvos com coda não apresentava a presença da consoante. Aos 17 meses, a porcentagem de omissão das líquidas ainda era alta, cerca de 75% de omissão. Ainda em inglês, Kehoe e Stoel-Gammon (2001) analisaram as produções de crianças com idades entre 1:3 e 2:0. As autoras encontraram, em coda, primeiramente, a produção de plosivas desvozeadas, após, fricativas surdas e nasais. Na sequência, surgiram as produções de plosivas e fricativas vozeadas, sendo a líquida /r/ o próximo segmento em coda adquirido, e, por último, o /l/. A lateral não foi encontrada nas produções das crianças até 2:0 anos.

A complexidade na aquisição desses sons pode ainda aumentar se forem consideradas as diferentes posições que ocupam na sílaba. No Português, em relação aos padrões silábicos, as crianças produzem inicialmente, de forma concomitante, as estruturas CV e V. A estrutura CVC é a terceira a ser adquirida, seguida da estrutura CCV (LAMPRECHT *et al.*, 2004; MEZZOMO, 2004; BONILHA, 2005). Assim, a produção de um som em estrutura CV, por exemplo, não significa sua emergência em outros contextos, como o CVC. A estrutura silábica, portanto, pode apresentar papel importante na aquisição dos segmentos.

---

<sup>2</sup> É importante ressaltar que, em espanhol, a vibrante múltipla /r/ é produzida em início absoluto de palavra, como em *razón* e seu papel de distinção de significado opera em contexto intervocálico, como em *pero/perro*, o primeiro significa *mas* e é articulado como *tap*, já o segundo significa *cachorro* e é articulado como vibrante. Isso significa que o rótico articulado como vibrante preenche o espaço ocupado pelas fricativas velares no português.

As características articulatórias dos róticos podem ser a principal causa para que esses segmentos sejam uns dos últimos a surgirem na fala das crianças (HERNANDORENA e LAMPRECHT, 1997). Estudos envolvendo a aquisição dos róticos, por crianças brasileiras, com desenvolvimento fonológico típico de fala, apontam um surgimento e aquisição mais tardios dessa classe de som em final de sílaba, como Lamprecht (1990), Miranda (1996), Mezzomo (2004) e Bonilha (2005). Esses trabalhos, dentre outros, trouxeram relevantes contribuições para a investigação acerca da aquisição fonológica do português, contudo, a maior parte dessas pesquisas descreveu os dados por meio de julgamentos de oitiva e desenvolveu suas análises com base em modelos teóricos gerativos, nos quais os objetos de análise são os fonemas ou traços distintivos. No entanto, novas perspectivas de análises surgem à medida que as pesquisas avançam, aliando-se, assim, a novos modelos teóricos. Um exemplo é a inspeção acústica, que se apresenta como uma ferramenta acurada, permitindo a observação do sinal da fala, tornando possível, assim, a descrição e explicação de determinados detalhes fonéticos até então não registrados pela análise de oitiva.

O presente estudo utiliza a análise acústica, pois essa permite identificar o papel de produções gradientes durante o processo de aquisição da linguagem, tornando possível formular hipóteses sobre a representação de um determinado fenômeno na gramática da criança.

São poucos os trabalhos que, sob o olhar da Fonologia Articulatória – também reportada como Fonologia Gestual (FonGest) –, são voltados para a aquisição do *tap* no português brasileiro. Barberena (2016) investiga os gestos articulatórios de [r], [l] e [jj], em adultos, crianças típicas e atípicas. Vassoler (2016) e Barbieri (2019) investigam o padrão de coordenação gestual imbricado na produção dos padrões silábicos do tipo C<sub>1</sub>C<sub>2</sub>V – em que C<sub>2</sub> é constituído por [r] – em oposição à CV de crianças com aquisição típica e atípica.

Mediante o exposto, a presente pesquisa, com base em pressupostos de um modelo dinâmico de produção da fala, analisa a emergência do *tap* e sua coordenação gestual sob o enfoque da Fonologia Articulatória.

A utilização da Fonologia Articulatória na análise dos dados viabiliza a discussão de um aspecto importante, relativos ao *tap*, que têm recebido especial atenção da literatura da área: o segmento vocálico adjacente ao *tap*, constatado nas inspeções acústicas (SILVA, 1996, 2002).

Barbosa e Madureira (2015, p.536) afirmam que, na produção do rótico, “cada vez que a ponta da língua deixa os alvéolos abaixando-se, um elemento vocálico é produzido”. Discussões no que concerne à natureza do elemento vocálico (EV) identificado na produção do segmento, na fala adulta, estão presentes em Silva (1996; 2002), Silva, Clemente e Nishida (2006), Silveira e Seara (2008), Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2018; 2021), Bilharva-da-Silva (2019), dentre outros. Conforme Silva, Clemente e Nishida (2006), no português, esse elemento apresenta a mesma natureza da vogal núcleo quando aparece à esquerda do *tap* em sílabas CCV, como se fosse “entrecortado” pela vogal, mas comporta-se como um *schwa*, quando à direita, compondo sílabas (C)VC. Bilharva-da-Silva (2019) afirma que a correlação entre a vogal núcleo da sílaba e o elemento vocálico ocorre de maneira mais intensa na posição pré-vocálica; já em posição pós-vocálica, essa correlação não é tão forte e os formantes (F1 e F2) do elemento vocálico tendem a se centralizar, porém, não é possível afirmar que a qualidade dessa vogal se assemelha a uma vogal neutra. No trabalho de Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2018; 2021), medidas articulatórias revelaram semelhanças entre a vogal núcleo da sílaba e o elemento vocálico em sílaba CCV, no entanto, quando o *tap* apresentava-se no final da sílaba, as medidas de contorno de língua evidenciaram diferenças entre o elemento vocálico e a vogal núcleo.

Modelos teóricos que levam em consideração o caráter dinâmico da linguagem permitem que o detalhe fonético apresentado anteriormente seja representado na língua portuguesa a nível gramatical. Por levar em conta que o gesto articulatório é composto por uma extensão temporal, a Fonologia Articulatória permite explicar fenômenos, como a produção do elemento vocálico, a partir da sobreposição gestual e da variável tempo, característico do primitivo de análise. Segundo Steriade (1990), o elemento vocálico representa um fragmento da vogal-núcleo da sílaba e, graças a não sobreposição temporal do gesto das consoantes, desponta no interior do cluster. Tal representação não seria possível a partir de um modelo gerativo, pois, devido à natureza categórica dos primitivos de análise, nenhum segmento poderia ser “entrecortado” por outro. A descrição desse elemento é fundamental para a análise da coordenação gestual apresentada na produção do *tap* em sílabas CVC e CCV no português.

Com base nas considerações aqui tecidas, estabeleceu-se o seguinte objetivo geral do trabalho: descrever e analisar, do ponto de vista acústico, a emergência dos

gestos articulatórios do segmento *tap* em sílaba CVC, em posição medial e final de palavra, por cinco crianças em fase de aquisição do português brasileiro.

Assim, considerando o objetivo geral desta pesquisa, as justificativas expostas, o aporte teórico escolhido e as ferramentas metodológicas utilizadas, foram delineados os seguintes objetivos específicos:

- (i) descrever acusticamente o *tap* e suas variações, presentes em sílaba CVC, produzido em fase de aquisição da fonologia;
- (ii) investigar quais recursos as crianças utilizam quando ainda não conseguem coordenar os gestos articulatórios necessários para produzir o *tap*;
- (iii) analisar a presença de prováveis contrastes encobertos, ainda que esses não sejam percebidos de oitiva;
- (iv) descrever a natureza do elemento vocálico adjacente à produção do *tap*.

Associadas aos objetivos propostos para este trabalho, foram tecidas as seguintes questões de investigação (Q1, Q2, Q3 e Q4), acompanhadas da hipótese de investigação (HI):

*Q1: A produção do rótico apresentará o padrão de tap, embora algumas variações possam ser identificadas?*

HI: Sim, embora a maioria dos róticos produzidos nessas posições seja do padrão de *tap* (MEZZOMO, 1999, 2004; MIRANDA, 1996), algumas variações do segmento podem ser observadas. É necessário destacar que o trabalho de Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2021) encontrou a produção de vibrantes alveolares em sujeitos adultos. Sendo assim, esse tipo de segmento pode também aparecer na fala das crianças.

*Q2: A criança usará principalmente a estratégia de redução gestual a fim de alcançar a produção de um tap?*

HI: Sim, espera-se que as crianças utilizem, na maior parte dos casos, a estratégia de redução da magnitude dos gestos, visto que, no trabalho de Vassoler (2016), as crianças típicas valeram-se dessa estratégia a fim de produzir o *tap* presente na sílaba CCV. Já as crianças atípicas apresentaram uma maior realização de gestos intrusivos, apresentando dificuldades em identificar os gestos de ponta e de

dorso de língua. A autora infere que esse fato ocorre em virtude dos desvios e atrasos de controle dessas regiões do corpo da língua.

*Q3: A análise acústica será capaz de revelar contrastes encobertos que não foram percebidos na análise de oitiva?*

HI: Sim, pois, conforme constatou Rodrigues (2007), a trajetória formântica foi um dos principais parâmetros que revelou a presença desses contrastes no processo de aquisição do rótico pelas crianças. Barbieri (2019) também encontrou contrastes encobertos na produção do *tap* em sílaba CCV na fala infantil.

*Q4: Existe relação entre a natureza formântica do elemento vocálico e a vogal núcleo da sílaba?*

HI: Sim, ainda que o “elemento vocálico que sucede o *tap* em cudas de final de palavra parece tender à centralização, assemelhando-se a um schwa” (SILVA; CLEMENTE; NISHIDA, 2006, p. 22), Bilharva-da-Silva (2019) não encontra indícios fortes para afirmar que a qualidade do elemento vocálico seja de uma vogal neutra.

Com essa investigação, o trabalho busca contribuir para a descrição do fenômeno de aquisição da estrutura CVC e dos processos gradientes presentes na emergência do *tap* na fala infantil. Ainda, poderá ampliar as discussões relativas à natureza do elemento vocálico que acompanha o *tap* nas produções do português. Para que se alcance esses objetivos, o uso de coletas longitudinais aliadas à utilização da inspeção acústica são ferramentas importantes para descrever e analisar a construção gradual da sílaba CVC.

## 2 Fundamentação Teórica

### 2.1 Descrição acústica e articulatória do *tap*

A produção dos segmentos róticos no português tem sido tópico de investigação em diferentes áreas, como aquelas que abordam aspectos voltados à aquisição, variação e à análise fonético-fonológica. Estudos acústicos envolvendo esses segmentos no português brasileiro podem ser encontrados em trabalhos como Silva, Clemente e Nishida (2006), Silveira e Seara (2008), Nishida (2009), Barbosa e Madureira (2015), Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2021), Bilharvada-Silva (2019), Silva *et al.* (2019), dentre outros.

A classe dos segmentos que podem ser considerados róticos consiste em um grupo amplo de sons. O Alfabeto Fonético Internacional (em inglês IPA - *International Phonetic Alphabet*) possui 7 diferentes símbolos para os sons representados pela letra *r* (WIESE, 2001). De modo mais geral, em português, o meio de contraste lexical se dá pela posição do segmento na sílaba ou na palavra. As produções dos róticos podem ser distintas, a depender da região do País, pois a variedade linguística tem uma influência direta em sua produção e, dependendo do dialeto, este pode ser articulado como: *tap* alveolar, vibrante múltipla, retroflexa alveolar, fricativa velar vozeada e desvozeada, fricativa glotal vozeada e desvozeada ou aproximante roticizada (*r* caipira). Por conta dessa grande variedade de articulações, faz-se necessária a caracterização desses segmentos em dialetos específicos. O *tap* alveolar e a vibrante múltipla são as variantes mais encontradas na fala produzida no sul do Brasil (MONARETTO, 1996).

Conforme dados do *corpus* CETENFolha<sup>3</sup> (CIRIGLIANO *et al.*, 2005), os quais tratam sobre a contagem do número de ocorrências dos fones, o *tap* [r] corresponde a 4,6% dos sons do PB, sendo a segunda consoante mais frequente. A fricativa [s] ocupa o posto de primeira consoante mais produzida no *corpus*, 6,5%. Contudo, vale ressaltar que o seu papel de morfema de plural exerce uma força importante para essa distribuição. Quanto à realização do segmento em coda, segundo Callou, Moraes e

<sup>3</sup> O CETENFolha, (Corpus de Extratos de Textos Eletrônicos NILC/Folha de S. Paulo) é formado a partir das 365 edições do ano de 1994 do jornal Folha de São Paulo, incluindo os cadernos não-diários. O *corpus* possui cerca de 24 milhões de palavras em português brasileiro. Foi criado pelo projeto *Processamento computacional do português* que visa aumentar a qualidade do processamento computacional da língua portuguesa.

Leite (1996), com dados de entrevistas realizadas ainda da década de 1970, a frequência do uso do *tap* é superior a 80% em São Paulo capital e Porto Alegre. Cerca de trinta anos depois, com dados coletados entre 2006 e 2007, Botassini (2011) investigou quais variantes dos róticos em posição de coda silábica são utilizadas na fala de moradores da cidade de Porto Alegre. Os resultados continuam sendo semelhantes. Com 176 produções do rótico em coda silábica, 152 ocorrências foram produzidas como *tap* e 24 como vibrante múltipla. Ou seja, o *tap* ocorreu em 86% das produções dos porto-alegrenses, enquanto a vibrante teve um índice de produção pequeno, apenas 14%.

Nas três capitais da região Sul, Porto Alegre (RS), Florianópolis (SC) e Curitiba (PR), Monaretto (1997) investiga a vibrante alveolar e aponta que, em posição pós-vocálica, o *tap* é o segmento mais produzido com um total de 40% no número total de ocorrências. É observado também um alto índice de apagamento, principalmente em verbos, ao passo que há um número pequeno de produções como variantes retroflexa, vibrante e fricativa velar.

A literatura geralmente relaciona os segmentos representados, na escrita, pelo grafema *r* a duas classes fonético-fonológicas: líquidas não-laterais e róticos. O primeiro é mais frequentemente utilizado em trabalhos sob a ótica de teorias gerativas, já o último acompanha trabalhos apoiados em outras bases teóricas.

A classificação das consoantes como líquidas, em uma perspectiva gerativa, se dá porque os segmentos /l/, /ʎ/, /r/ e /R/ fazem parte da mesma classe natural, ou seja, compartilham alguns traços distintivos. A classe das líquidas divide-se em dois grupos: as líquidas laterais e as líquidas não-laterais.

No que diz respeito ao modo de articulação, com bases articulatórias, esses sons são produzidos pela liberação da passagem do ar pelas partes laterais da boca. Segundo Silva et al. (2019, p.30), “canais laterais são formados pelo abaixamento e estreitamento do corpo da língua e permitem a propagação lateral da corrente de ar”, o que propicia a produção das consoantes líquidas laterais [l] e [ʎ]. Porém, quando o obstáculo da passagem de ar é realizado com uma ou várias batidas da língua nos alvéolos, são produzidos os *taps* e vibrantes ([r] e [r]), esses pertencentes à classe das líquidas não-laterais.

O termo rótico é utilizado para os sons fricativos, aproximantes, vibrantes e *taps* (THOMAS, 2011). Esse é usado sem motivação fonética, visto que, com pontos e modos de articulação distintos, os sons de *r* apresentam-se com qualidades fonéticas

diferentes. Ladefoged e Maddieson (1996) afirmam que, por certo, o uso desse vocábulo tem motivação na escrita, pois, de fato, as palavras grafadas com R representam as variações dos sons de *r*. “Cada membro da classe dos róticos se assemelha a outro membro em relação a alguma propriedade, mas não é a mesma propriedade que constitui a semelhança de todos os membros de uma classe” (LINDAU, 1985, p. 166, tradução nossa).<sup>4</sup> Em outras palavras, não há uma semelhança articulatória específica que seja capaz de unir todos os róticos.

Acusticamente, o *tap* pode ser identificado por dois momentos. O primeiro é a ausência de energia que ocorre durante a produção do *tap*, devido ao contato da ponta/lâmina da língua com os alvéolos. Essa apresenta-se como um espaçamento em branco no espectrograma e reflete o “breve momento de obstrução que corresponde à fase de fechamento do trato vocal” (SILVA et al., 2019, p. 191). Já o segundo, o ruído transiente, refere-se ao “momento do afastamento dos articuladores correspondente à fase de abertura da passagem da corrente de ar na articulação do tepe, e ocorre imediatamente após a soltura da oclusão” (Ibid., p. 200) e pode não ocorrer em todas as produções do segmento. Lindau (1985) afirma que alguns *taps* apresentam, durante o seu fechamento, uma mínima quantidade de energia, porém, na maior parte das vezes, o espectrograma apresenta-se sem formantes durante esse breve período. A autora também afirma que esses valores se modificam não só de acordo com a língua, mas entre os falantes de um mesmo idioma.

Partindo para uma análise contrastiva, entre a vibrante e o *tap*, é clara a diferença desses sons no espectrograma. O *tap* produz apenas uma vibração no espectrograma, gerada a partir de uma batida da ponta da língua no céu da boca, enquanto a vibrante tem, no mínimo, duas marcas visíveis dessa ocorrência.

A Figura 1 apresenta um espectrograma de banda larga, no qual há a produção de uma vibrante [r], sendo possível observar quatro espaçamentos em branco, sinalizados pelas setas. É possível inferir que esses espaços sejam uma consequência acústica das batidas da ponta da língua nos alvéolos; enquanto na produção o *tap* [r] apenas um único espaçamento é observado, sendo possível inferir que há uma única batida.

---

<sup>4</sup> No original: *each member of the rhotic class resembles some other member with respect to some property, but it is not the same property that constitutes the resemblance for all members of a class.*

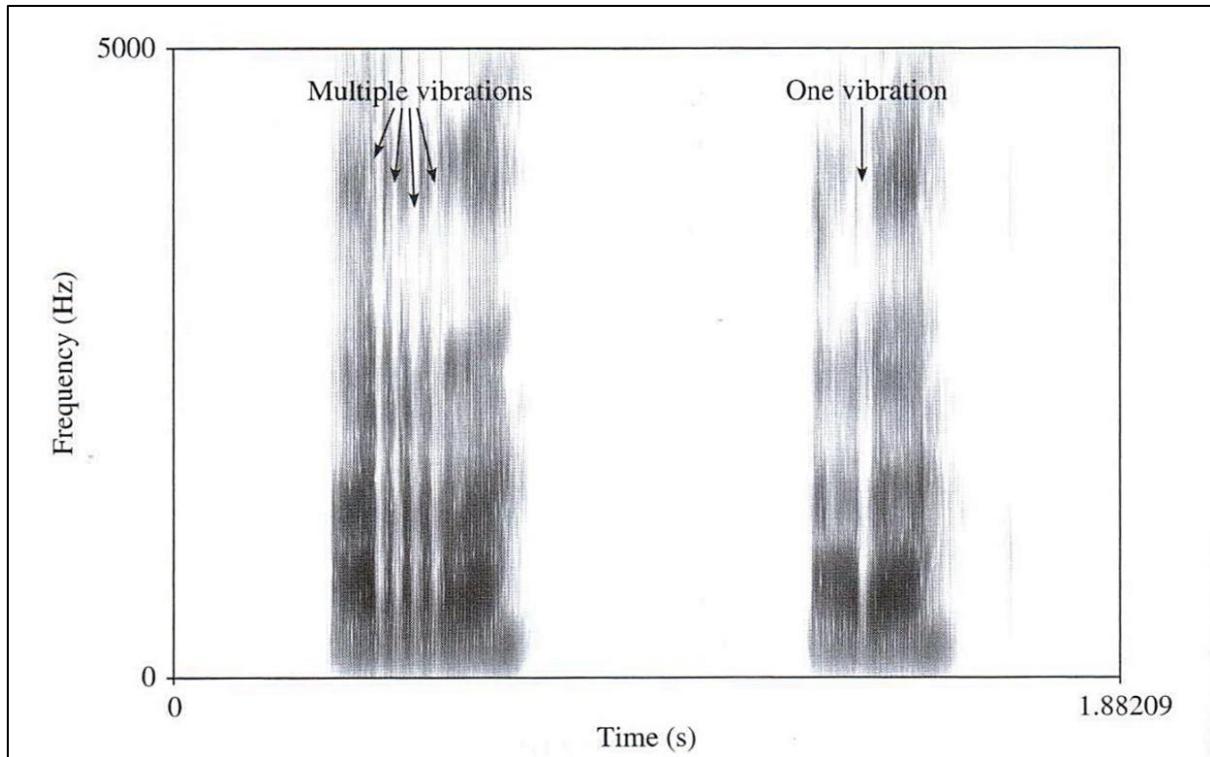


Figura 1 - Comparação entre a vibrante em [ərə] e o *tap* em [ərə]. A vibrante apresenta várias vibrações, o *tap* apenas uma. Fonte: THOMAS (2011, p.130)

Barbosa e Madureira (2015) afirmam que o *tap* é o som mais curto em português, já que seus valores de duração giram em torno de 30 ms, ao contrário da vibrante múltipla, que pode ter duração absoluta de 194 ms. No que diz respeito à produção de *tap* e vibrantes, os autores afirmam que:

A produção como vibrante múltipla diz respeito ao som produzido pela batida repetida de um articulador ativo de pequena massa (ponta ou lâmina da língua ou úvula) contra um articulador passivo (palato duro). A produção como *tap* diz respeito a uma batida rápida da ponta ou lâmina da língua no palato duro, com o articulador ativo na forma convexa ou côncava (no caso do retroflexo) (BARBOSA e MADUREIRA, 2015, p. 535).

Ladefoged e Maddieson (1996) caracterizam o *tap* como um som em que um breve contato entre os articuladores é feito pelo movimento de um articulador ativo diretamente ao céu da boca. Isto quer dizer que a ponta/lâmina da língua toca os alvéolos. Para ilustrar essa descrição, a Figura 2 mostra o momento de máxima constrição do gesto de ponta de língua para o segmento. O *tap* é produzido com uma oclusão total e rápida do fluxo de ar nas cavidades orais. O véu do palato está levantado, impedindo a passagem do ar pelas cavidades nasais.

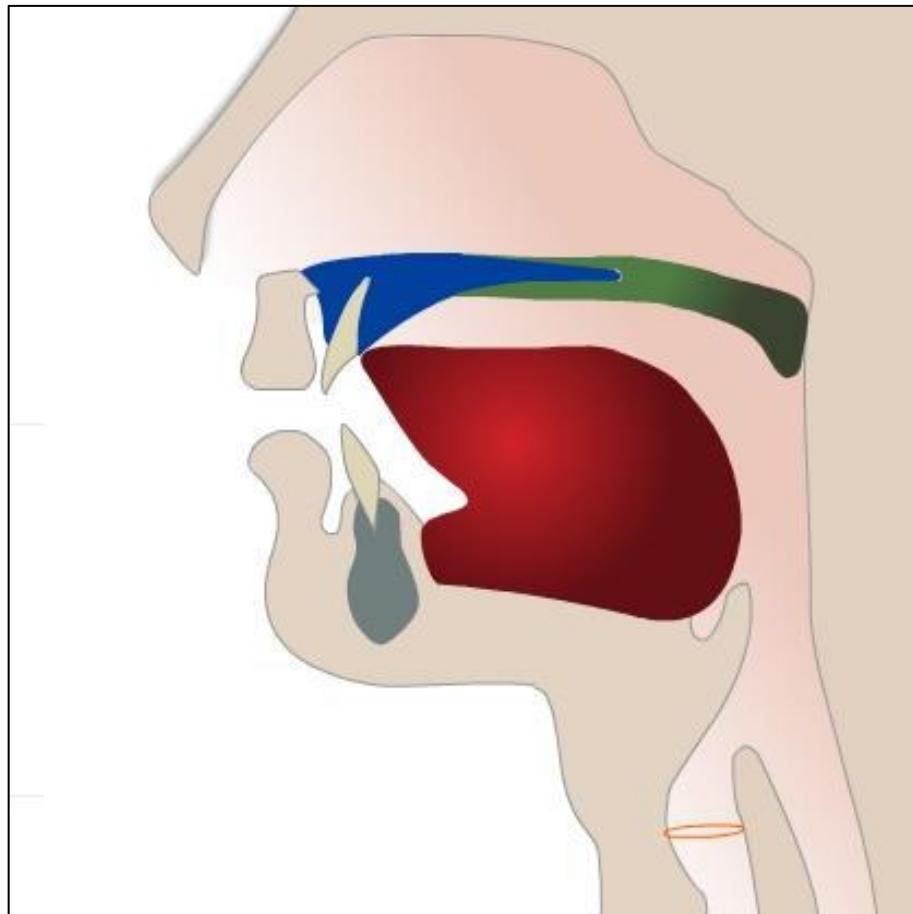


Figura 2 – Momento de máxima constrição do gesto de ponta de língua para o *tap* [r]. Fonte: Adaptado do site *fonologia.org* de CRISTÓFARO-SILVA; YEHIA (2009)

O movimento articulatório para a produção do *tap* é veloz, sendo assim, esse segmento apresenta uma duração curta se comparado a outros segmentos. A Figura 3 apresenta um corte sagital que permite a visualização da extensão do corpo da língua, abrangendo as principais regiões, como o dorso e a ponta. É possível observar que são apresentadas três imagens ultrassonográficas do movimento da língua para a produção do *tap*, sendo possível observar o ponto máximo de contrição. A região posterior do trato oral encontra-se à direita nas imagens. A seta azul sinaliza o dorso e a amarela a ponta da língua. Abaixo dos *frames* é possível observar o espectrograma das produções com uma linha preta que marca o momento de produção para cada *frame*.

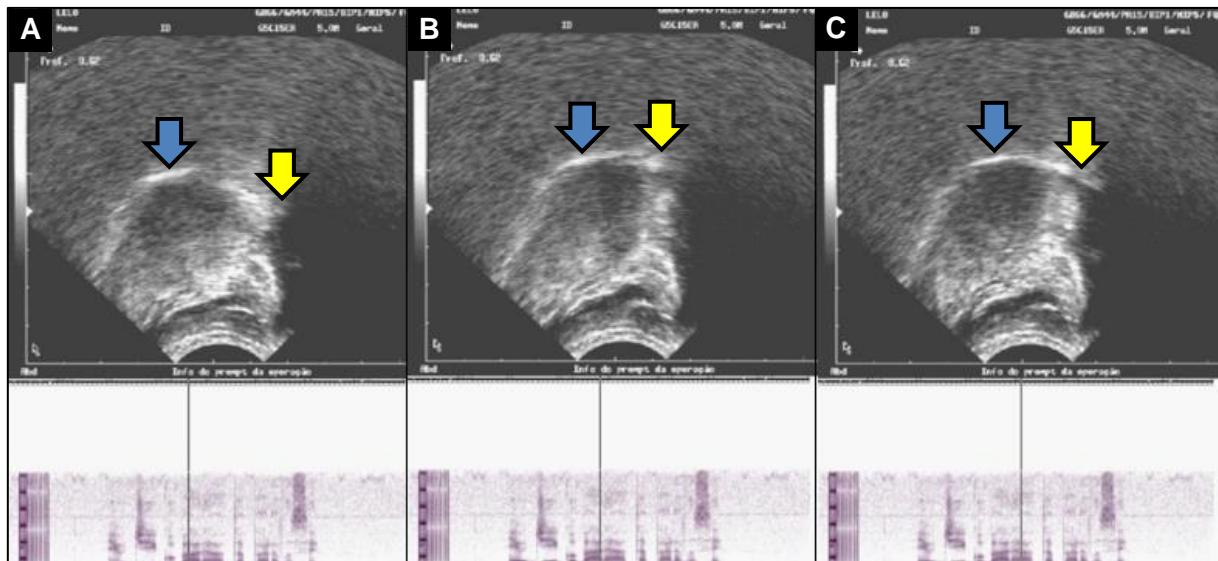


Figura 3 - Frames com a trajetória da ponta da língua na sequência [ara]. Fonte: a autora

No *frame* A, é possível observar a ponta da língua em uma posição mais baixa. Em seguida, no *frame* B, ocorre um movimento de elevação da ponta. Na imagem do espectrograma, logo abaixo, nota-se que a elevação da ponta da língua coincide com um breve espaçamento em branco. Posteriormente, no *frame* C, inicia-se o abaixamento da ponta da língua. A partir dessas imagens articulatórias pode-se inferir o breve momento em que a ponta da língua atinge o seu ponto máximo de constrição e toca os alvéolos, confirmando, assim, a curta duração que esse segmento apresenta.

Há na literatura divergências a respeito de quantos gestos articulatórios compõem alguns segmentos, dentre eles o *tap*. Proctor (2011), para o espanhol e o russo, afirma haver dois gestos linguais para o *tap*, assim como Gick et al. (2007) para o inglês. A identificação de dois gestos independentes também é defendida para o *tap* em Português Brasileiro, em trabalhos como Silva (2002), Barberena, Keske-Soares e Berti (2014) e Barberena (2016). Recasens (2016) questiona a classificação dada pela literatura à vibrante e ao *tap* como contendo dois gestos articulatórios, na qual o *tap*, além do gesto de ponta, é produzido com algum abaixamento e retração do corpo da língua. Observe-se a Figura 4:

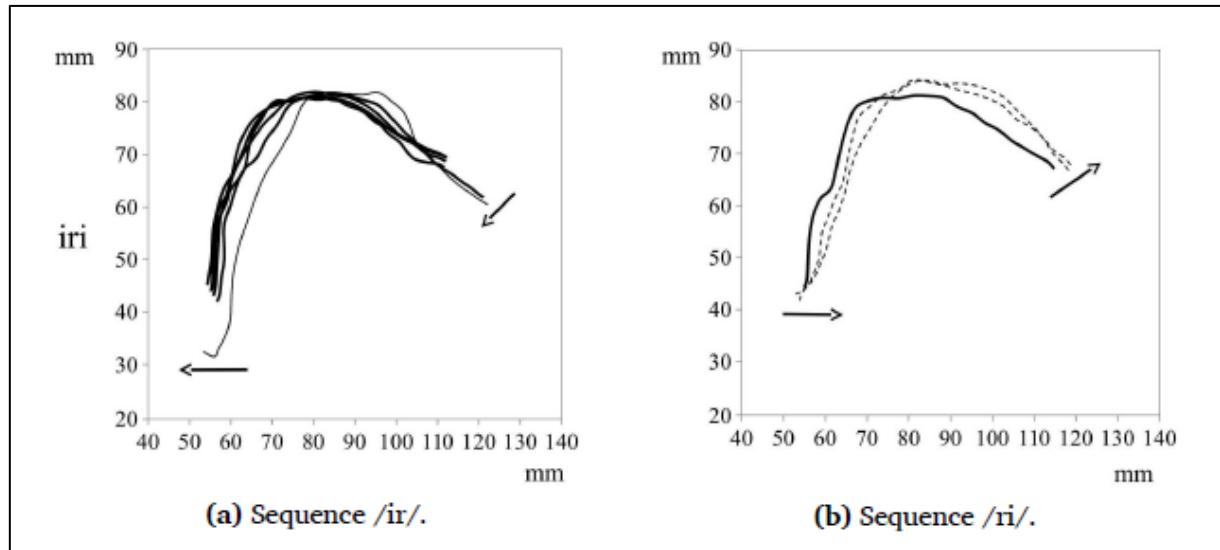


Figura 4 - Dados de ultrassom para a sequência /iri/ produzida por um falante de Catalão. Fonte: RECASENS (2016, p. 36)

Na Figura 4, conforme o autor, os dados de ultrassom produzidos por um falante nativo do Catalão, para a sequência /iri/, revelam:

movimento de retração do corpo da língua e um abaixamento do pré-dorso, conforme ocorre o avanço de V1 para o ponto médio da vibrante [figura 4 (a)], e o movimento oposto da língua, isto é, anteriorização do corpo da língua e elevação do pré-dorso, do último contato apical para a vibrante (linha grossa) até a soltura de V2 [figura 4 (b)] (RECASENS, 2016, p. 36 tradução nossa<sup>5</sup>).

Para o autor, os movimentos da língua e suas configurações, no que diz respeito às vibrantes alveolares, não estão ligados à existência de um movimento independente do corpo da língua, mas a mecanismos articulatórios necessários à sua produção, ou seja, o gesto primário, de ponta, desencadeia o que seria um gesto secundário, de corpo. De acordo com Recasens, assim, ao invés de um segmento complexo, essas consoantes configurariam, portanto, um segmento de um único gesto de constrição. Ainda, Recasens pontua que, em sua perspectiva, as consoantes complexas que envolvem gestos da parte anterior e do dorso da língua deveriam apenas incluir consoantes palatalizadas e velarizadas/faringalizadas. A análise articulatória, por meio do aparelho de ultrassom, pode, pois, contribuir para discussões acerca da trajetória articulatória apresentada pelo *tap* no português.

Segundo Barbosa e Madureira (2015, p.536), “cada vez que a ponta da língua deixa os alvéolos abaixando-se, um elemento vocálico é produzido”. O elemento

<sup>5</sup> No original: *some tongue body backing and predorsum lowering as we proceed from V1 onset to about the trill midpoint (fig. 3 (a), and the opposite tongue motion, i.e., tongue body fronting and tongue predorsum raising, from the last apical contact for the trill (thick line) to V2 offset (fig. 3 (b).*

vocálico de apoio foi previamente encontrado em pesquisas como Silva (2002), Silva, Clemente e Nishida (2006) e Silveira e Seara (2008). No oscilograma, é possível observar suas manifestações acústicas à esquerda em sílabas CCV e à direita em sílabas CVC – embora, nesta última posição, alguns trabalhos apontem que nem sempre esse elemento encontra-se presente no PB (MELCHEQUE, FERREIRA-GONÇALVES E BRUM-DE-PAULA, 2018, 2021; BILHARVA-DA-SILVA, 2019). Em sílabas CV o elemento vocálico não pode ser identificado em razão do posicionamento do *tap* entre vogais. Na Figura 5, é possível observar a produção do *tap* nas palavras *carta*, à esquerda, e *par*, à direita. Destacado pelas linhas vermelhas, tanto na forma de onda como no espectrograma, é possível identificar a ocorrência do *tap*. As setas apontam à produção do elemento vocálico, na forma de onda.

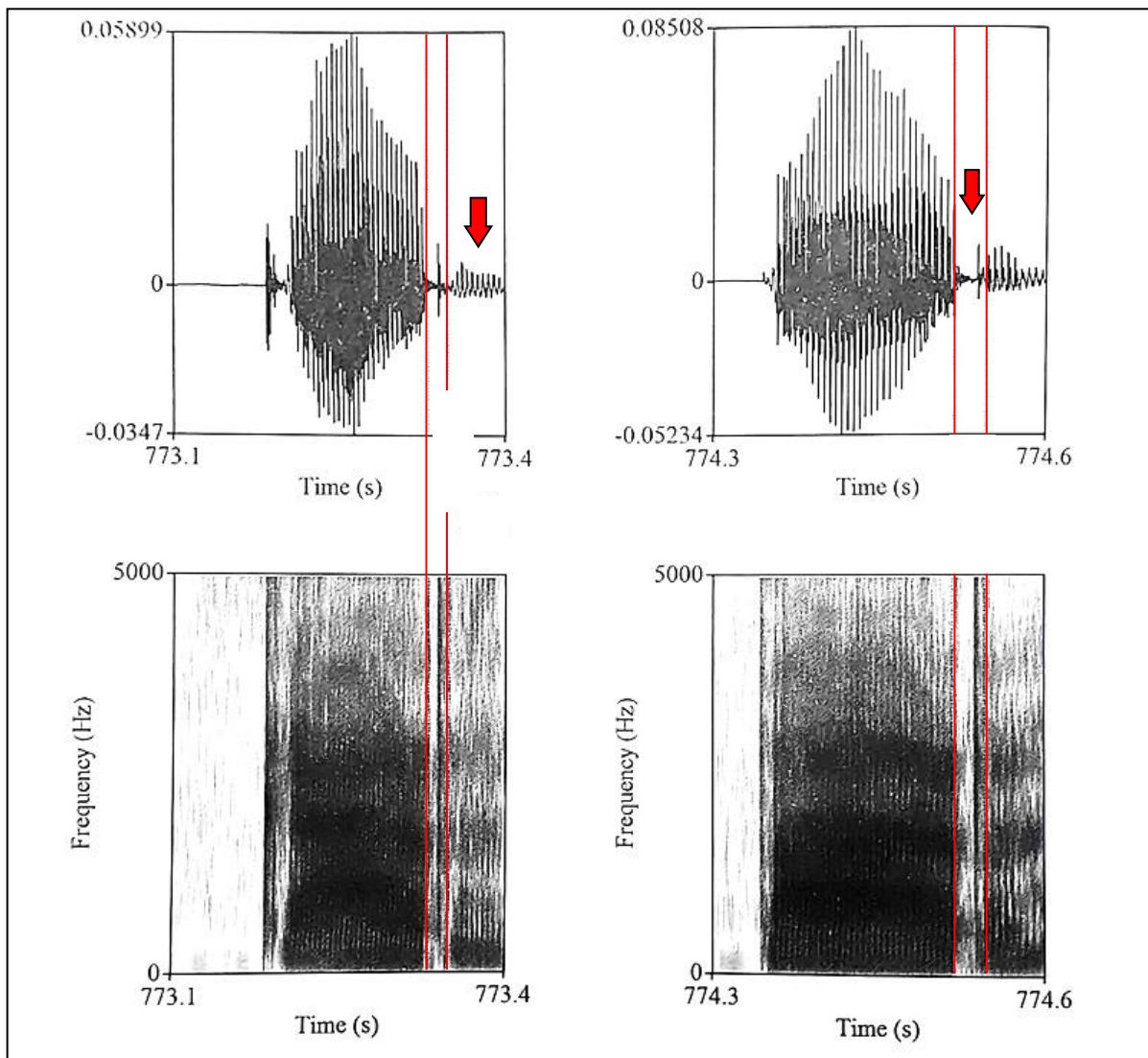


Figura 5 - Forma de onda e espectrograma da plosiva, vogal e *tap* das palavras *carta* e *par*. Fonte: adaptado de SILVA et al. (2019, p. 196)

Conforme Silva, Clemente e Nishida (2006), no português, o elemento vocálico apresenta a mesma natureza da vogal núcleo da sílaba quando aparece à esquerda do *tap* em sílabas CCV, mas comporta-se como um *schwa*, quando à direita, compõndo sílabas (C)VC. Tem, assim, natureza variável com base na posição silábica que este ocupa.

Os trabalhos de Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2018, 2021) buscaram investigar acústica e articulatoriamente a produção do segmento rótico em sílabas CV, CCV e CVC, bem como o elemento vocálico adjacente presente nas sílabas do tipo CrV e CVr. As amostras analisadas pertencem a cinco informantes adultas, do sexo feminino e naturais da cidade de Pelotas/RS. Os resultados acústicos dos estudos indiciam que o rótico produzido segue o padrão de um *tap* em sílaba CV e CCV. Em relação à sílaba CVC, foram encontrados, ainda, produções de vibrantes múltiplas, padrão não esperado para o dialeto pelotense. No que concerne ao elemento vocálico, este foi encontrado em sílabas CCV e CVC, porém, em sílabas CVC, apresentou-se instável e, algumas vezes, não foi detectado. As imagens ultrassonográficas evidenciaram a presença de gesto de ponta, no ápice do movimento articulatório, indicando o movimento anterior da língua, conforme previsto pela literatura, ou seja, elevação crescente da ponta e do corpo da língua, na sequência *ar*, e abaixamento da ponta e do corpo, na sequência *ra*, como constatado no estudo de Recasens (2016).

Com o objetivo de investigar a configuração gestual do elemento vocálico, medidas articulatórias foram obtidas a partir de um teste estatístico, o *Teste-T*, a partir do software AAA<sup>6</sup>. Esse comparava o par de médias das produções de um mesmo informante. Para CCV, as medidas articulatórias não revelaram diferenças entre o elemento vocálico e os 25% da fase inicial da vogal nuclear em CV. A partir desse resultado é possível inferir que o elemento vocálico, para os dados da informante analisada, teria a mesma natureza da vogal núcleo da sílaba CCV. É possível observar, na Figura 6, que não há pontos de diferença estatisticamente significativos ao longo do traçado da borda da língua. A linha preta na parte superior da figura

---

<sup>6</sup> O software foi utilizado para traçar as linhas para as bordas de língua. Após ser gerada as médias das produções e definido o valor de *p*, o Teste-T criava uma tabela com as regiões de diferença estatística e as sinalizava com um asterisco (\*) ao lado. As regiões da língua são indicadas pelas numerações presentes na primeira coluna que inicia da região anterior do trato para a posterior.

apresenta a *outlier* das produções. À direita da imagem há um quadro com os valores de  $p$ .

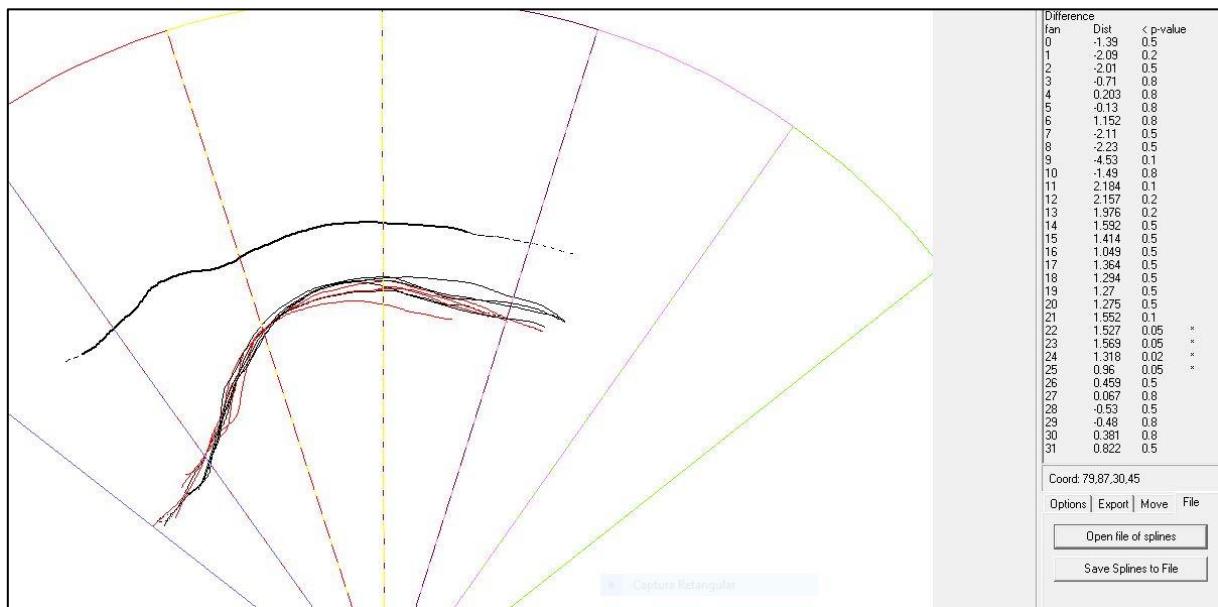


Figura 6 - Teste-T do par de médias das produções do elemento vocálico em CCV (traçados pretos) e dos 25% da fase inicial da vogal nuclear em CV (traçados vermelhos). Linha preta superior apresenta a *outlier da estatística*. Fonte: MELCHEQUE; FERREIRA-GONÇALVES; BRUM-DE-PAULA (2021, p. 250)

Para analisar o elemento vocálico produzido no final da sílaba, as autoras compararam os pares de médias de CVC e os 25% da fase final da vogal nuclear em CV. Os resultados revelaram diferenças entre esses pares. Na Figura 7, à esquerda, é possível observar uma grande zona de pontos de diferenças estatisticamente significativas para o par CV e CVC medial. Essa grande região engloba a parte do dorso até a ponta da língua e os valores de  $p$  estão entre 0.001 e 0.01. À direita, para os pares CV e CVC final, também são apresentadas diferenças estatísticas para as regiões de dorso e ponta de língua, com a região da ponta da língua apresentando valores de  $p$  de 0.001.

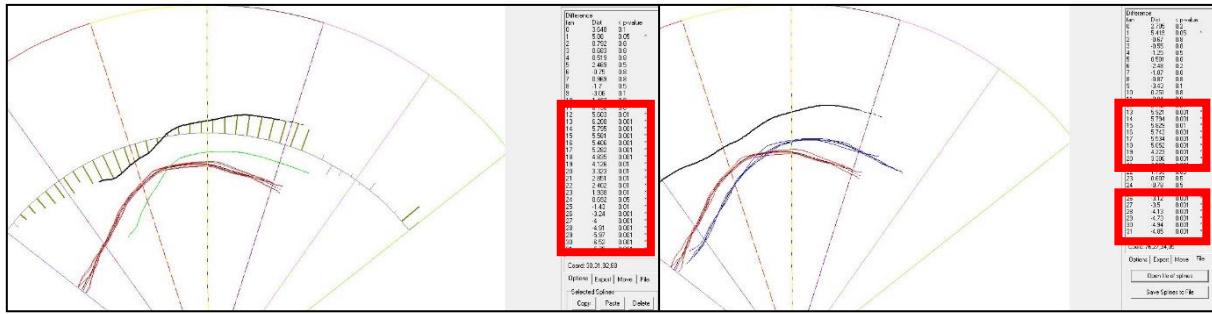


Figura 7 - Teste-T do par de médias das produções do elemento vocálico em CVC e 25% da fase final da vogal nuclear em CV. À esquerda CV (traçados vermelhos) e CVC medial (traçados verdes); à direita CV (traçados vermelhos) e CVC final (traçados azuis). Caixas vermelhas indicam os pontos de diferenças com valores de  $p$  entre 0.01 e 0.05. Linha preta superior apresenta a *outlier da estatística*.  
Fonte: MELCHEQUE; FERREIRA-GONÇALVES; BRUM-DE-PAULA (2021, p. 251)

O par de médias que comparava os traçados da língua do elemento vocálico em CCV e em CVC revelou diferenças do movimento de ponta de língua para ambas as posições, como pode ser observado na Figura 8.

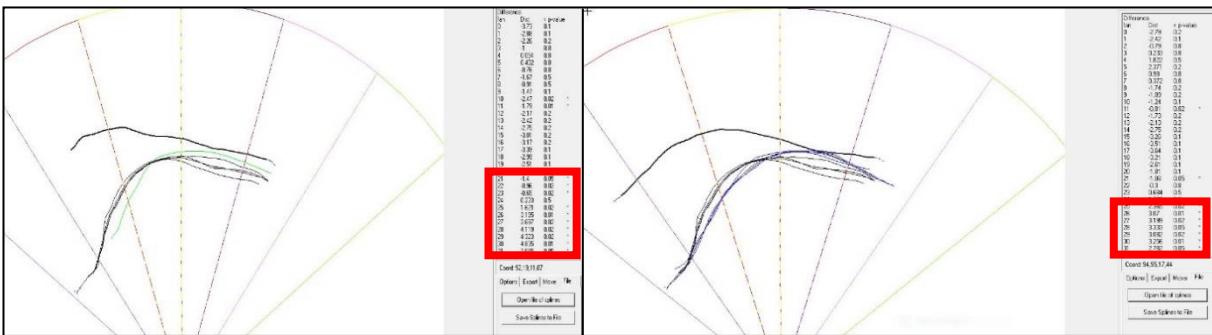


Figura 8 - Teste-T do par de médias das produções do elemento vocálico em CCV e em CVC. À esquerda, CCV (traçados pretos) e CVC medial (traçados verdes); à direita, CCV (traçados pretos) e CVC final (traçados azuis). Caixas vermelhas indicam os pontos de diferenças com valores de  $p$  entre 0.01 e 0.05. Linha preta superior apresenta a *outlier da estatística*.  
Fonte: MELCHEQUE; FERREIRA-GONÇALVES; BRUM-DE-PAULA (2018)

Portanto, o trabalho das autoras indica que o elemento vocálico adjacente ao *tap* se comporta como parte da vogal eclipsada apenas em sílaba CCV. Quando esse encontra-se no final de sílaba, nos casos de CVC, seu comportamento é diferente da vogal núcleo da sílaba na qual está inserido.

Quanto ao estudo acústico de Silveira e Seara (2008), as autoras buscaram caracterizar a vogal de apoio, assim denominada, presente nas sílabas CCV produzidas por dois informantes (um masculino e um feminino) naturais de Florianópolis/SC. O trabalho contou com um *corpus* de palavras compostas por encontros consonantais tônicos que apresentavam as plosivas /p/, /b/, /t/, /d/ e as fricativas /f/, /v/, seguidas do rótico /r/ e das vogais altas /i/, /u/ e /a/. Os resultados do

informante masculino apontaram diferenças entre os dois primeiros formantes quando a vogal de apoio é produzida juntamente às vogais [a] e [i]. Para a informante do sexo feminino, além das diferenças presentes nos dados do informante masculino, foram observadas diferenças para os valores do segundo formante da vogal [u]. Posto isto, as autoras concluem que as vogais de apoio produzidas por seus informantes “possuem uma qualidade acústica significativamente diversa daquela mostrada pela vogal núcleo, tendo uma forte tendência a centralizar-se” (SILVEIRA; SEARA, 2008, p. 43). Na Figura 9, é possível visualizar que, no espaço acústico, há uma centralização das vogais de apoio, pois há uma diminuição do distanciamento entre as vogais, tanto no que concerne à altura, quanto ao grau de anterioridade e posterioridade.

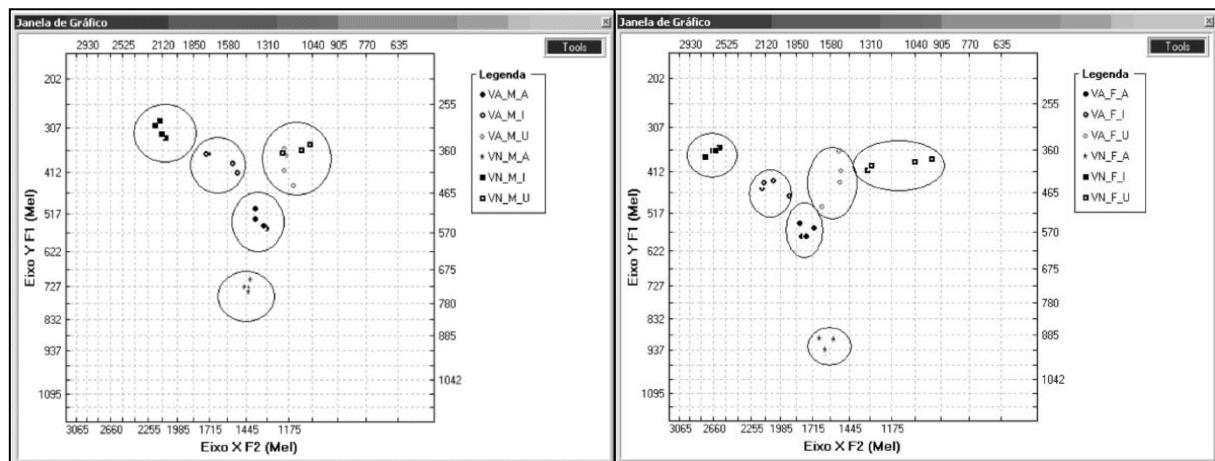


Figura 9 - Espaço acústico vocálico apresentando os valores de F1 e F2 da VA (vogal de apoio) e VN (vogal núcleo) seguidos das vogais [a], [i] e [u]. À esquerda dados do informante masculino e à direita dados da informante feminina. Fonte: adaptado de SILVEIRA; SEARA (2008, p. 44 e 45)

Bilharva-da-Silva (2019) investiga a natureza do elemento vocálico – chamado de vogal suarabácti – presente em encontros consonantais tautossilábicos e heterossilábicos, com base na perspectiva gestual de Browman e Goldstein (1986, 1989a, 1992). A coleta dos dados consistiu em amostras de produções de dez informantes – cinco bilíngues e cinco monolíngues –, com idade superior a 42 anos. As falantes bilíngues (português/pomerano) eram residentes do município de São Lourenço do Sul/RS. O grupo das monolíngues foi formado por moradoras da cidade de Pelotas/RS.

Os resultados, provenientes da aplicação de três experimentos – (i) nomeação de palavras relativas a figuras presentes no instrumento do livro *Avaliação Fonológica da Criança* (YAVAS; HERNANDORENA; LAMPRECHT, 1991); (ii) nomeação de

palavras, em uma frase-veículo, motivadas pela apresentação de figuras e (iii) nomeação das mesmas palavras do segundo experimento, motivadas pela apresentação da forma ortográfica dos itens lexicais – evidenciaram que o grupo dos bilíngues apresentou uma taxa menor de produções do elemento vocálico, ainda que, para ambos os grupos, a presença desse elemento tenha sido elevada. Destaca-se também o fato de a suarabácti apresentar uma maior duração quando produzida pelas informantes bilíngues. Conforme o autor, esse fato pode refletir um diferente padrão de coordenação gestual quando se comparam o português e o pomerano. Em relação à natureza do elemento vocálico, o autor constatou uma correlação mais forte com a vogal núcleo da sílaba quando o elemento vocálico estava em posição pré-vocálica. Para a posição pós-vocálica os dois primeiros formantes tendem a centralizar-se, no entanto, o autor afirma que não é possível efetivamente comprovar que a qualidade acústica desse elemento se assemelha a uma vogal neutra.

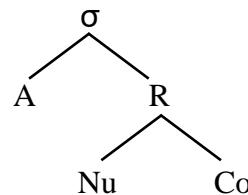
De modo geral, os trabalhos apresentados nesta seção buscaram discutir a caracterização do *tap* do ponto de vista acústico e articulatório, bem como a natureza do elemento vocálico que geralmente é constatado nas produções.

## 2.2 Aquisição do *tap* do Português Brasileiro

Esta pesquisa busca investigar a emergência dos gestos articulatórios do *tap* em sílabas CVC, tendo como teoria base a Fonologia Articulatória<sup>7</sup>. No entanto, é necessário que se faça um panorama de trabalhos que, a partir dessa e de outras bases teóricas, contribuíram para o avanço nos estudos sobre a aquisição dos róticos e das estruturas silábicas no PB. Pesquisas como Lamprecht (1990), Miranda (1996), Hernandorena e Lamprecht (1997) e Mezzomo (2004), por exemplo, partiram de teorias que se apoiam em princípios gerativistas, como a Fonologia Natural (STAMPE, 1973) e a Teoria de Princípios e Parâmetros (CHOMSKY, 1981, 1986).

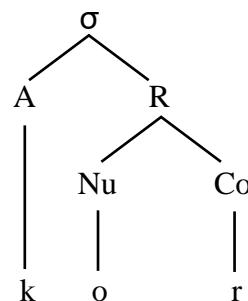
As pesquisas citadas fazem uso da teoria métrica da sílaba. De acordo com esse modelo, o menor constituinte prosódico, a sílaba, possui uma estrutura interna, dividida em unidades menores. A sílaba pode ser, então, composta por ataque (A) e rima (R) – essa última divide-se em núcleo (Nu) e coda (Co) (SELKIRK, 1982). Qualquer componente pode ser vazio, exceto o núcleo. Selkirk (1982), proponente do modelo, baseou-se em propostas anteriores feitas por Pike e Pike (1947) e Fudge (1969). Posto isto, a representação de sílaba seria a disposta em (1):

(1)



A representação em (2), da palavra *cor*, apresenta uma sílaba com a coda sendo ocupada pela consoante /r/.

(2)



<sup>7</sup> O modelo teórico será abordado na sessão 1.3.

A consoante de coda tem uma relação mais forte com a vogal núcleo do que a consoante de *onset*, pois fazem parte da rima, embora, essa consoante não esteja presente em todas as sílabas. Uma vez que essa não é um elemento obrigatório da estrutura, pois não são todos os segmentos que podem ocupar essa rima ramificada. Em português, existem fortes restrições aos sons que a ocupam. Na mesma direção, não existe uma unanimidade por parte dos teóricos sobre quais sons podem ocupar essa posição. Para Câmara Jr. (1953, 1977), Bisol (1999) e Cagliari (2002), as consoantes /S/, /R/, /l/ e /N/ são permitidas em coda. No entanto, Alvarenga e Oliveira (1997) e Mateus e d'Andrade (2000) não consideram o /N/ como um segmento possível para a posição de coda no português.

Collischonn (1997) aponta a existência de cinco padrões de sílabas com coda simples, tais como: VC – **ar, ás, um, alto**; CVC – **lar, parte, paz, pasta, gol**; CVVC – **mais**; CCVC – **três, frasco, trem, planta, flor, fralda**; CCVVC – **claustro**; e três padrões de sílabas com codas complexas. Tais estruturas complexas consistem na sequência *soante* + S, como: VCC – **instante**; CVCC – **monstro, perspicaz, solstício** – e CCVCC – **transporte**.

Embora existam codas com mais de uma consoante, o alvo do presente estudo são as codas simples, uma vez que as complexas constituídas por róticos no português ocorrem raramente na língua. As estruturas silábicas com coda têm, no processo de aquisição do português, emergência mais tardia, após o surgimento de CV e de V (LAMPRECHT, 1990; SANTOS, 1990; ROSA, 1992; BONILHA, 2005).

A pesquisa de Lamprecht (1990) acompanhou o desenvolvimento fonológico de 12 crianças com idades entre 2:9 e 5:5. Para a estrutura CVC, a autora observou que a consoante nasal, no final da sílaba, é adquirida muito antes que a fricativa /S/, em torno de 1:6 a 2:0. Após, surge a líquida lateral /l/ – que não apresenta dificuldades por ser, em geral, semivocalizada em final de sílaba no português –, ao passo que a líquida não lateral /r/ é a última a ser adquirida. Os dados revelaram também que, dentro da classe natural das líquidas, o /l/ se estabelece antes de /r/ e o /R/, em geral, surge antes de /r/.

Hernandorena e Lamprecht (1997) analisaram a fala de 310 crianças com idades entre 2:0 e 4:0 (Banco de Dados AQUIFONO) e observaram, durante a aquisição dos sons de *r*, uma grande distância cronológica entre esses segmentos, dependendo da posição silábica em que os mesmos se encontram. Em posição de onset simples, o /R/ é dominado aos 3:4, enquanto o *tap* /r/ é dominado quase um ano

depois, aos 4:2. Entre as demais posições silábicas, também há um grande distanciamento cronológico, pois, em coda, o segmento é dominado aos 3:10, em contrapartida, em onset complexo, apenas aos 5 anos.

Na mesma direção das pesquisas já citadas, Miranda (1996), em seu estudo transversal, analisou a aquisição do *r* por grupos de idade e posição silábica. A autora observa que o *tap* surge primeiro em coda final aos 2 anos, mantendo percentual de produção acima das demais posições silábicas até a faixa etária dos 2 anos e 9 meses, como é possível observar no Gráfico 1:

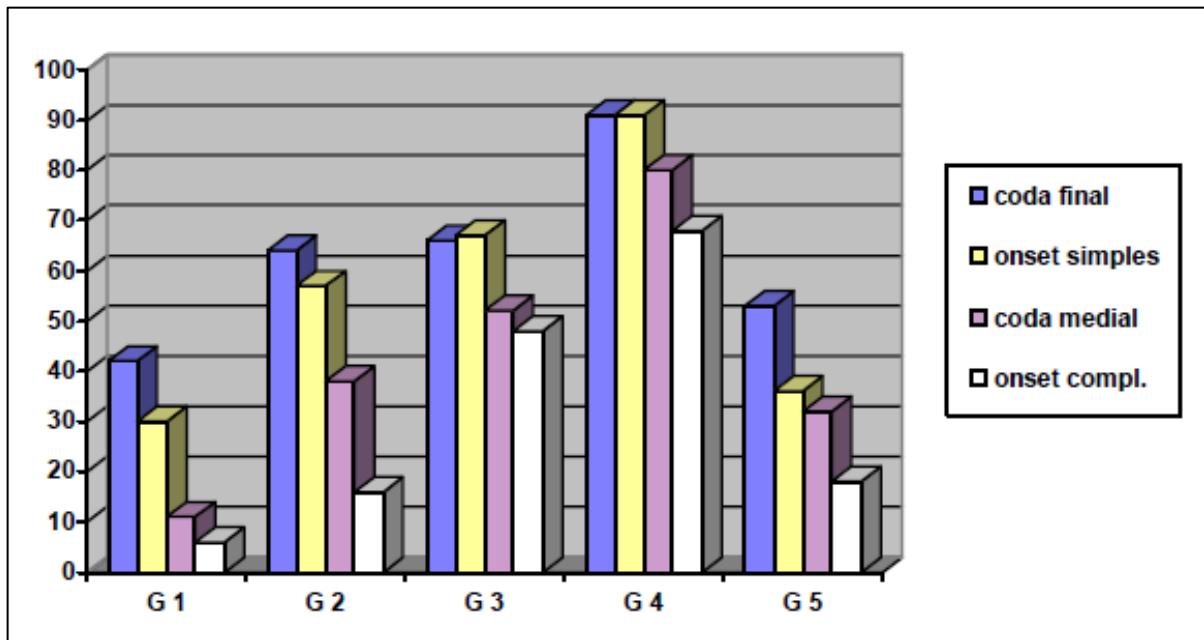


Gráfico 1 - Produção de 'r-fraco' versus posição silábica. G1 2:0 - 2:1; G2 2:2 - 2:3; G3 2:4 - 2:5; G4 2:6 - 2:7 e G5 2:8 - 2:9. Fonte: MIRANDA (1996, p. 102)

A autora, ao analisar a variável: posição na palavra versus tonicidade, verificou que, em 305 possíveis alvos de realização, o rótico foi produzido em 60% dos casos, sendo 287 em posição tônica. Já em coda medial, ao se considerar a tonicidade, dos 1209 alvos possíveis, 41% dos róticos foram produzidos, sendo 607 tônicos. Com base nessas informações, a pesquisadora afirma que

pode-se inferir, conforme já sugerido por Hernandorena (1990), que a conjunção tonicidade e posição de final absoluto é a responsável pela produção precoce de líquida não-lateral em coda final (MIRANDA, 1996, p. 63).

Assim, a posição da coda dentro da palavra também apresenta diferenças cronológicas de aquisição, pois, em coda final, os segmentos surgem e se estabilizam

primeiro, só então a coda medial se estabiliza. Isso ocorre independentemente do tipo de segmento (HERNANDORENA, 1990; MIRANDA, 1996).

Mezzomo (2004) investigou amostras de fala a partir de dois bancos de dados, o AQUIFONO e o INIFONO – com dados de crianças das cidades de Porto Alegre e Pelotas, ambas do Rio Grande do Sul. A autora também aponta o surgimento do *tap* em coda final antes de coda medial. O primeiro ocorre aos 1:11 e o segundo, aos 2:2. No Gráfico 2, é possível observar o percurso de aquisição da líquida não-lateral.

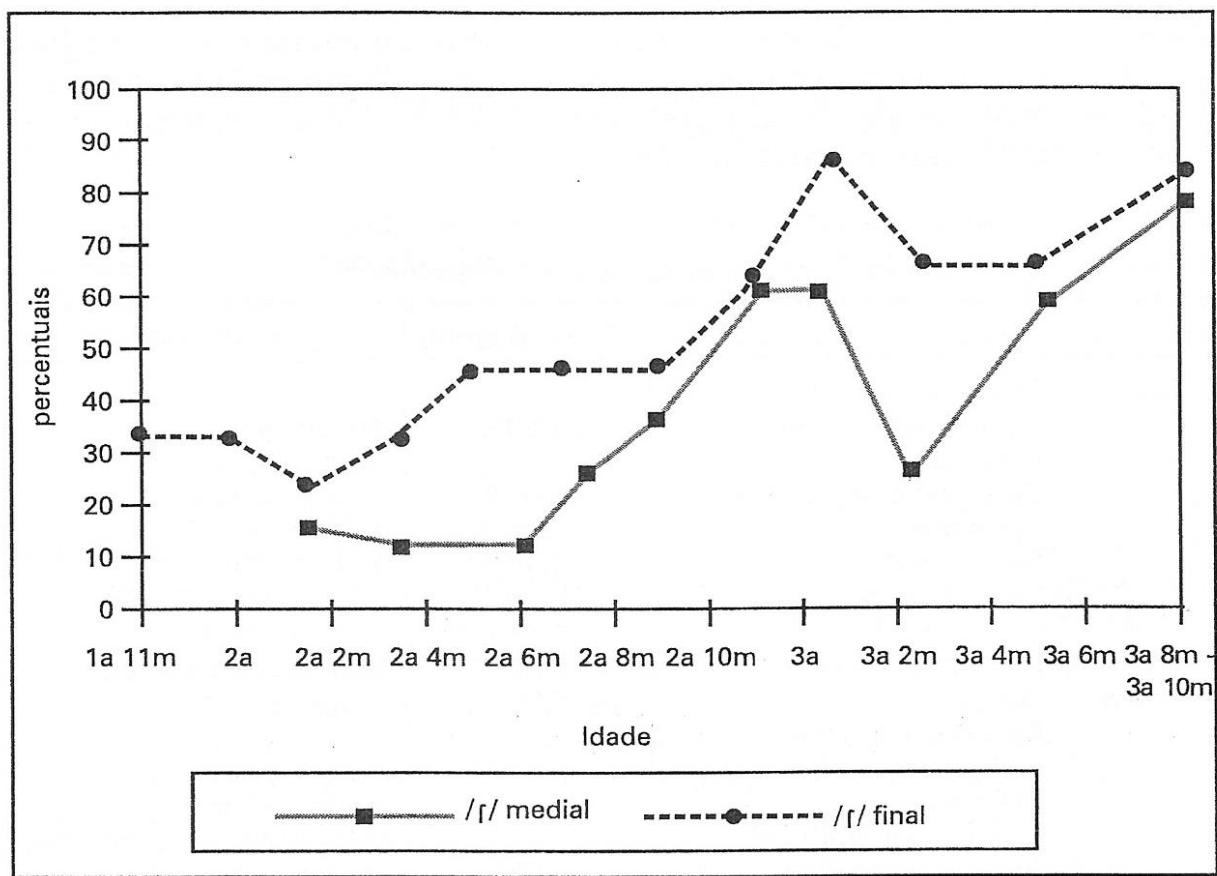


Gráfico 2 - Aquisição da líquida não-lateral em coda. Fonte: MEZZOMO (2004, p.145)

Conforme pode ser visualizado no Gráfico 2, é possível observar uma queda no percentual de produção, em coda final, entre 2:2 e 3:6; em coda medial, entre 2:4 e 2:6 e entre 3:2 e 3:4. Esse fenômeno foi intitulado por Strauss (1982) como *Curva em U*. Essa observação de regressão de produções corretas também é observada nos estudos de Hernandorena (1990), Mezzomo (1999) e Miranda (1996). A aquisição completa do segmento em ambas posições, porém, dar-se-á aos 3:10.

No que concerne à não realização correta do segmento, as autoras já citadas referem-se a essas produções como resultantes do emprego de *estratégias de reparo*.

Essas consistem em inserir, apagar ou substituir segmentos. Mezzomo (2004) afirma que as crianças optam por estratégias diferentes, a depender da posição do *tap* na sílaba. Para coda final, as crianças lançam mão de mais estratégias de reparo do que para coda medial. Em meio de palavra, a autora observou que o caso de omissões do segmento é de 56,59%, ou seja, as crianças optam por não produzir o *tap* nessa posição. Foi notado, ainda, um alto número de metáteses, com a ocorrência de 17 casos. Já para a coda final, a autora observou 20 ocorrências de epênteses, dessas, 16 ocorreram juntamente com a substituição do alvo, por exemplo, a produção da palavra *trator* como [ta'tolu], e apenas 4 com somente o recurso da epêntese, como a palavra *colher*, sendo produzida como [ku'eri]. Semivocalizações também foram observadas, assim como foram registrados 2 casos de metátese e 23,53% de omissões. O Gráfico 3 apresenta as estratégias de reparo para a coda final, registradas pela autora.

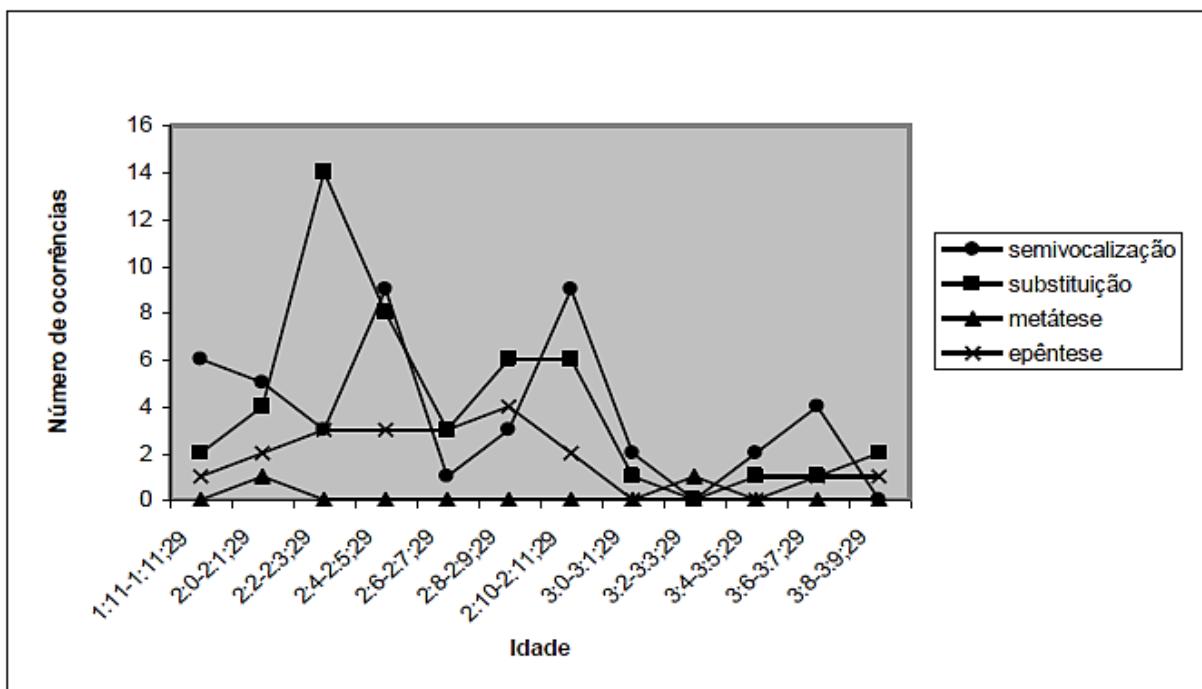


Gráfico 3 - Recursos utilizados na aquisição do /r/ em coda final. Fonte: MEZZOMO (2004, p. 119)

Hernadorena e Lamprecht (1997, p.19) afirmam que “/l/ é a consoante protótipo da classe das líquidas: é a primeira a ser adquirida, é aquela capaz de substituir – durante o processo de aquisição da fonologia – todas as outras”. Esse segmento é, também, um dos mais encontrados nas línguas do mundo. Em cerca de 317 línguas estudadas, o /l/ é encontrado em 79% delas, sendo um dos 8 mais frequentes segmentos produzidos nessas línguas. (Maddieson, 1984 *apud*

Hernandorena e Lamprecht, 1997). Esse fato pode explicar o motivo pelo qual Mezzomo (2004) encontrou a maioria dos casos de substituição de /r/ por /l/.

As pesquisas até aqui reportadas apontam uma produção mais tardia do segmento rótico quando comparado às demais consoantes que ocupam a posição de coda no português. A criança faz uso, então, de estratégias para adequar sua produção ao segmento que ela ainda não produz e/ou à estrutura silábica ainda não conhecida. Ressalta-se, no entanto, que, considerando as diferentes posições que o *tap* pode ocupar na silaba, a silaba CVC final parece favorecer o processo de aquisição, conforme Miranda (1996).

É possível observar que os estudos citados entendem que os processos fonológicos ocorrem de maneira categórica, ou seja, se o informante não produziu o segmento alvo, ou houve a realização de um apagamento ou a realização de uma substituição. No entanto, mesmo que esses trabalhos tenham contribuído para o campo de pesquisas sobre aquisição fonológica, um modelo teórico dinâmico pode proporcionar importantes subsídios acerca do desenvolvimento da linguagem da criança e seus desdobramentos até a produção de um segmento alvo.

O trabalho de Bates, Dale e Thal (1997) apresenta a importância da variação para a compreensão do desenvolvimento da linguagem. Segundo os autores, ao longo da história, o caráter variável foi ignorado pelos pesquisadores, pois o foco principal dos estudos se concentrava em atentar às semelhanças na ampla maioria das crianças que estavam adquirindo a linguagem. Com isso, as diferenças individuais tornavam-se ruídos banais nas pesquisas. Porém, essas variações podem apresentar caráter substancial e estável, trilhando um curso evolutivo próprio. Sendo assim:

Por ser substancial, esta variação é fundamental para se definir o limite entre o desenvolvimento normal e anormal; por ser estável, ela fornece uma janela para os correlatos e (por inferência) para as causas da mudança evolutiva; e por possuir seu próprio curso evolutivo, ela pode ser utilizada para localizar transições evolutivas críticas que formam a base das teorias da aprendizagem e mudança (BATES; DALE; THAL, 1997, p. 87)

Os autores afirmam ainda que as muitas diferenças individuais que a linguagem inicial apresenta, no seu ritmo inicial de aquisição, desafiam e limitam uma falsa noção de que existiria um cronograma maturacional universal para o desenvolvimento linguístico. Sendo assim, as teorias que buscam explicar os processos de aquisição da linguagem necessitam levar em conta o caráter variável da linguagem e considerar a importância das diferenças individuais durante esse período.

Em uma abordagem emergentista, é possível aproximar pressupostos teóricos e as inferências feitas a partir de dados empíricos, possibilitando novas explicações para fenômenos que, antigamente, careciam de atenção. Nessa perspectiva, a aquisição de padrões gramaticais fonológicos, sejam relativos à emergência segmental ou de estruturas silábicas, por exemplo, está fundamentalmente vinculada à própria emergência lexical, pois a gramática emerge do léxico. Considera-se, assim, o caráter dinâmico, variável e não-linear da linguagem, dado que o processo de aquisição é influenciado pela interação com o ambiente do qual a criança faz parte e pela maturação do seu sistema cognitivo.

Com base em uma perspectiva emergentista da aquisição da linguagem, Rodrigues (2007) investiga a aquisição dos róticos, a partir da Fonologia Acústico-Articulatória (FAAR)<sup>8</sup>. A pesquisadora considerou os dados de duas crianças do sexo masculino, com idades entre 5:7 e 5:8, ambas com queixa fonoaudiológica. Os informantes foram acompanhados por cinco meses, tendo sido realizadas três coletas de dados. Além disso, duas vezes por semana, durante esse período, os sujeitos, nomeados como LF e MD, participaram de sessões de terapia fonoaudiológica. Com a intenção de comparar as produções desses informantes com as de um sujeito de aquisição típica da fala, cujo o período de estabilização dos sons já estivesse concluído, foi selecionada uma criança de 5:6 como sujeito-controle.

O foco principal da pesquisa baseava-se na investigação do papel dos acertos gradientes durante o processo de estabilização de contrastes fônicos nas produções dos informantes. Em outras palavras, isso significa que a autora investigou a natureza dos acertos e erros de produção, buscando a presença de contrastes encobertos. Os contrastes investigados pautavam-se entre duas estratégias de reparo utilizadas pelos informantes selecionados: o apagamento do *tap* em posição de encontro consonantal nas produções de LF e MD, por exemplo, ['patu] para ['pratu] – o informante LF ainda omitia o segmento em posição de início de sílaba, por exemplo, ['pa:] para ['para] ou ['kaw] para ['karu] – e a semivocalização do rótico ocorrida nas produções de MD em

---

<sup>8</sup> O modelo teórico denominado FAAR (Fonologia Acústico-Articulatória), proposto por (ALBANO, 2001), está filiado à Fonologia Articulatória, doravante FAR (BROWMAN; GOLDSTEIN, 1987, 1988, 1989a, 1989b, 1990a, 1990b, 1992, 1993). Para a FAR, a unidade de análise é o gesto articulatório, isto é, o movimento dos articuladores que produzem constrições no trato vocal, sendo que seu objetivo é o cumprimento de uma tarefa. No entanto, para a FAAR, tanto o aspecto articulatório como o acústico constituem o gesto, havendo a necessidade de incorporar à definição de gesto os aspectos acústicos da produção. Nessa direção, as regiões acústico-articulatórias definem os gestos e devem ser indicadas nas representações formais dos segmentos, as pautas gestuais.

posições de início de sílaba, por exemplo, ['kaja] para ['kara] e, em final de sílaba<sup>9</sup>, por exemplo, ['kajta] para ['kaṭa] ou ['paj] para ['paj].

A autora relaciona a não produção “correta” do segmento a possíveis contrastes encobertos, ou seja, enquanto uma análise de oitiva afirmaria que o sujeito produziria uma semivocalização ao invés do segmento alvo, a análise acústica permitiria comprovar que existem produções gradientes e que os contrastes distintivos estão presentes na fala das crianças, embora ainda não sejam percebidos auditivamente.

Para tanto, a investigação acústica abrangeu alguns parâmetros de análise, como a inspeção visual dos espectrogramas, investigação da trajetória formântica, cálculos de medidas de duração e análise estatística dos valores obtidos. A primeira coleta de LF apresentou a omissão do segmento na análise de oitiva, porém, a trajetória formântica evidenciou uma distinção nos valores dos dois primeiros formantes para os pares [cara] e [sala], e apenas no segundo formante para [cárie] e [pai]. Os dados de oitiva da segunda e terceira coleta permitiram identificar a produção de um *tap* e, segundo a pesquisadora, “esse fato pode ser levado em consideração para a conclusão de que os resultados da primeira coleta, quanto à presença de contraste encoberto, pareceram ser de fato fidedignos”. (RODRIGUES, 2007, p. 51). A inspeção visual do espectrograma evidenciou dois modos de articulação do rótico: um com características de som plosivo e outro como som aproximante. Os valores de duração relativa apontaram que a segunda coleta apresentou valores de duração relativa entre 3% e 17%, evidenciando um período de maior instabilidade, ou seja, um novo padrão estava em processo de estabilização. A terceira coleta evidenciou produções mais estáveis, com valores de duração relativa em torno de 8,5% a 16%,

Com relação ao processo de aquisição de [r] e [l] do informante MD, a análise de oitiva da primeira coleta identificou que, para as possíveis produções do *tap*, a semivocalização ocorreu em todos os casos. Contudo, medidas de trajetória formântica revelaram uma distinção entre a semivogal [j] e o *tap* [r], evidenciando uma tentativa de distinção entre esses segmentos por parte da criança. Valores mais altos do primeiro formante para o item lexical com rótico, nos pares [caro] e [caio], sugerem

---

<sup>9</sup> A autora pontua que a aproximante retroflexa [l] era esperada em final de sílaba, por ser a variedade linguística de Campinas/SP, cidade onde o informante MD residia. Todavia, produções com o *tap* [r] também poderiam ser realizadas pela criança, uma vez que sua mãe era argentina e seu pai, russo.

a existência de uma diferenciação articulatória, provavelmente, relacionada à altura da língua.

Na segunda sessão de coleta de dados, ainda que a análise de oitiva identificasse produções semelhantes à semivogal, a trajetória formântica não era a mesma da primeira coleta. Estatisticamente, as distinções para o terceiro formante nos pares [cara] e [saia], como também para [cárie] e [caí], revelaram uma produção gradiente com características de um som retroflexo. É necessário salientar que, em posição final de sílaba, a aquisição do rótico retroflexo [ɻ] ocorreu a partir da segunda coleta de dados, visto que os valores do terceiro formante foram significativamente diferentes para os pares [cuɻta] e [cuida], e para [paɻ] e [pai]. Sendo assim, infere-se que o surgimento da retroflexa em final de sílaba tenha motivado a variação na produção do *tap* em posição silábica inicial.

Por fim, na terceira coleta, a criança apresentava produções estabilizadas para o *tap* em início de sílaba. Contudo, as produções em final de sílaba, que pareciam estar estabilizadas como um som retroflexo, apresentaram variações entre aproximante retroflexa e o *tap*. A pesquisadora assegura que as “mudanças relacionadas a um som dessa classe acarretaram desestabilizações e emergência de novos padrões relacionadas a outro som em aquisição”. (RODRIGUES, 2007, p. 125).

Em suma, os valores de duração relativa revelaram variações nas produções, isto quer dizer que o sujeito está experimentando diversas formas para produzir seu alvo. Por meio de medidas de trajetórias formânticas, a autora conseguiu identificar contrastes em andamento – principalmente nos valores de F2 –, porém, não percebidos na análise de oitiva. Esses apontam que os informantes produziam constrições no trato oral, isto é, a análise acústica tornou possível inferir a presença de gestos de ponta e/ou dorso de língua. Isso significa que “o contraste é sacrificado em termos auditivos, embora preservado em termos articulatórios” (ALBANO, 2007, p. 143). Sendo assim, Rodrigues (2007) afirma que os sujeitos apresentam dificuldades frente à regulação da magnitude dos gestos e, assim, o movimento no trato oral não é percebido pelos ouvintes.

Outro estudo de perspectiva emergentista, Barberena (2016), apoiou-se na Fonologia Gestual para caracterizar os gestos articulatórios do *tap* [ɾ] e seus pares [ɻ] e [ɻ̪]. Sua pesquisa contou com 30 sujeitos adultos (AT), 30 crianças com desenvolvimento típico de fala (CT) e 30 crianças atípicas (CA). A pesquisadora utilizou a ultrassonografia (USG) como técnica para coleta/análise de dados e como

instrumento para intervenção fonoaudiológica. O trabalho contou com três estudos: o primeiro buscou caracterizar e comparar a variação dos gestos de língua e da produção dos sons entre os três grupos de sujeitos; o segundo realizou comparações entre os gestos linguais dos grupos CT e CA, conforme sexo e idades, assim como comparou as produções de [l] e [j] àquelas em que o [r] foram substituídas por [l] ou [j]; por fim, o terceiro estudo realizou uma proposta de intervenção em CA ,com o uso da USG como auxílio para o tratamento.

A coleta de dados ocorreu com base em duas listas, contendo pares de palavras com os segmentos [r], [l] e [j], nos contextos de /a/, /i/ e /u/. A primeira lista continha pares com itens lexicais dissílabos com o alvo em sílaba átona, por exemplo, enquanto que a segunda lista era composta por pares formados de palavras trissílabas com o segmento alvo das produções em posição tônica. O Quadro 1 apresenta as listas de palavras utilizadas pela autora.

PARES DE PALAVRAS DISSÍLABAS ÁTONAS			
	[r]	[l]	[j]
[a]	[‘karə]	[‘kalə]	[‘kajə]
[i]	[‘kari]	[‘kali]	[‘kaj]
[u]	[‘karu]	[‘kalu]	[‘kaju]
PARES DE PALAVRAS TRISSÍLABAS TÔNICAS			
	[bə’rate]	[sə’ladə]	–
[a]	[bə’rate]	[sə’ladə]	–
[i]	[fi’ridə]	[fe’lidə]	–
[u]	[ʃə’rutu]	[kabe’ludu]	–

Quadro 1 - Listas de palavras utilizadas no estudo de Barberena (2016)

Para que fosse possível observar os movimentos da língua durante as produções, foi utilizado um aparelho de ultrassom e um transdutor posicionado logo abaixo do queixo dos informantes. Assim, era possível captar as imagens da borda da língua durante as produções. Embora o trabalho da pesquisadora conte com minuciosas descrições e comparações, aqui serão reportados os principais resultados dos dois primeiros estudos, uma vez que o foco da presente pesquisa não está voltado para a terapia fonoaudiológica.

Dessa forma, o primeiro estudo evidenciou que, para todos os grupos, foram identificados dois gestos distintos para as produções de [r] e [l]: retração da raiz da língua e a elevação, em direção aos alvéolos, da ponta da língua. O grupo dos adultos típicos apresentou, na raiz da língua, diferença significativa entre [r] e [l]. O Gráfico 4

apresenta os contornos médios da borda da língua para as produções de [r] e [l] de cada grupo de informantes.

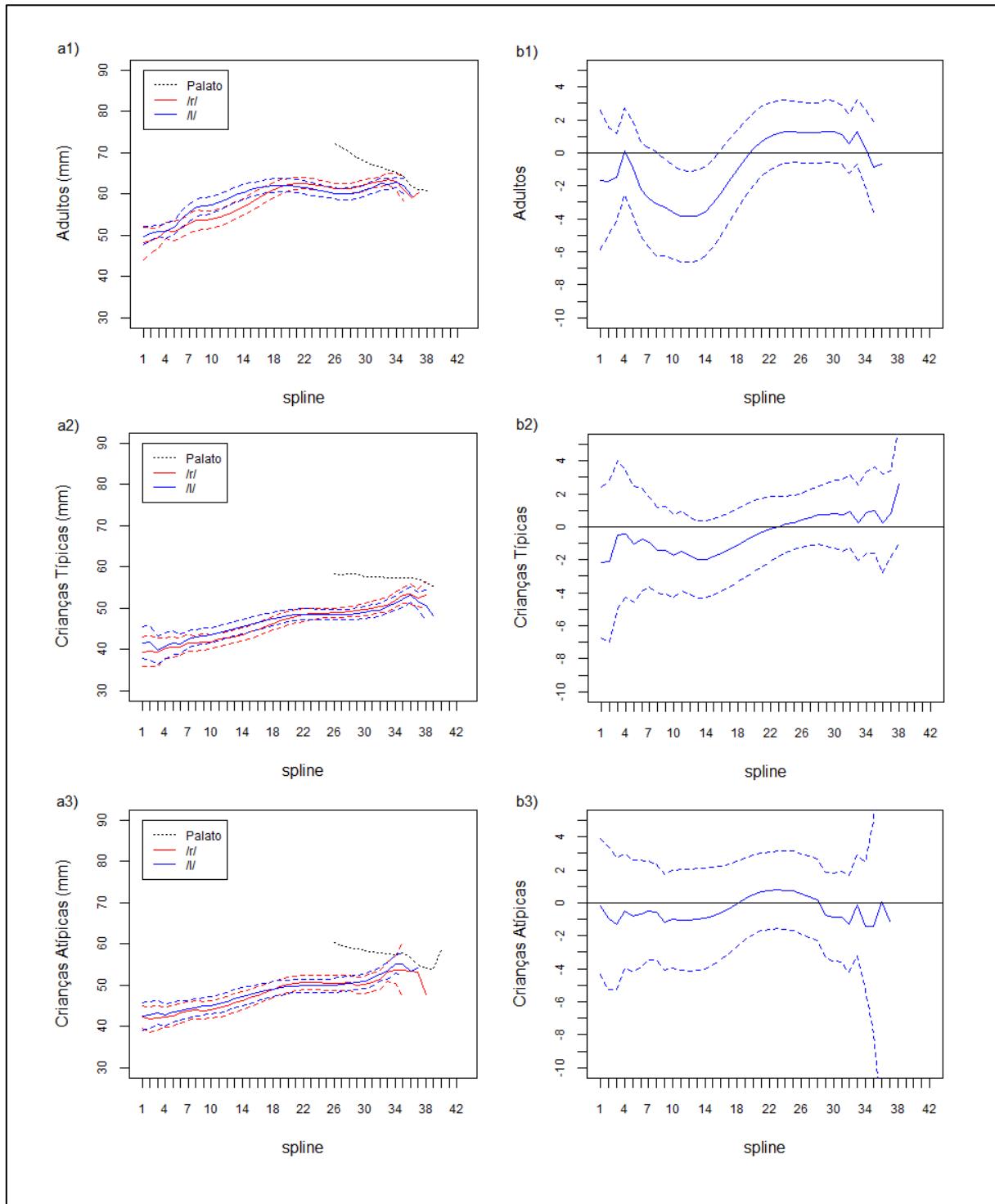


Gráfico 4 - a1, a2 e a3) intervalos de confiança para os contornos médios de língua; b1, b2 e b3) diferença nos contornos médios entre [r] e [l] para todas as palavras nos grupos AT, CT e CA. Em vermelho os traçados de língua para [r] e em azul os traçados para [l]. FONTE: BARBERENA (2016, p.118)

É possível observar que os contornos da borda da língua dos dois grupos de crianças aproximaram-se do padrão gestual dos adultos. As crianças típicas tendem a diferenciar os gestos de [r] e [l] na raiz da língua, com o *tap* apresentando maior retração nessa região, ao passo que a ponta da língua exibe uma sobreposição ao [l], apresentando uma leve elevação no [r]. As crianças atípicas sobrepõem os gestos de [r] e [l] na região da raiz da língua, porém, a ponta da língua apresentou, na região dos alvéolos, movimentos ora para baixo, ora para cima, evidenciando, assim, tentativas de produção do *tap*. Barberena (2016) ainda afirma que, apesar de qualitativamente ser permitido afirmar que há um maior refinamento nos gestos do grupo CT em relação ao grupo CA, a análise estatística não apresentou diferenças significativas entre os traçados de língua para os segmentos [r] e [l].

Os resultados do segundo estudo apontam que, das 30 crianças atípicas, 29 apresentavam a produção de [l]. Porém, quando o segmento alvo era o [r], 18 substituíram por [l]. No Gráfico 5, são apresentados os traçados da produção de [l] e das substituições de [r] por [l] do grupo CA.

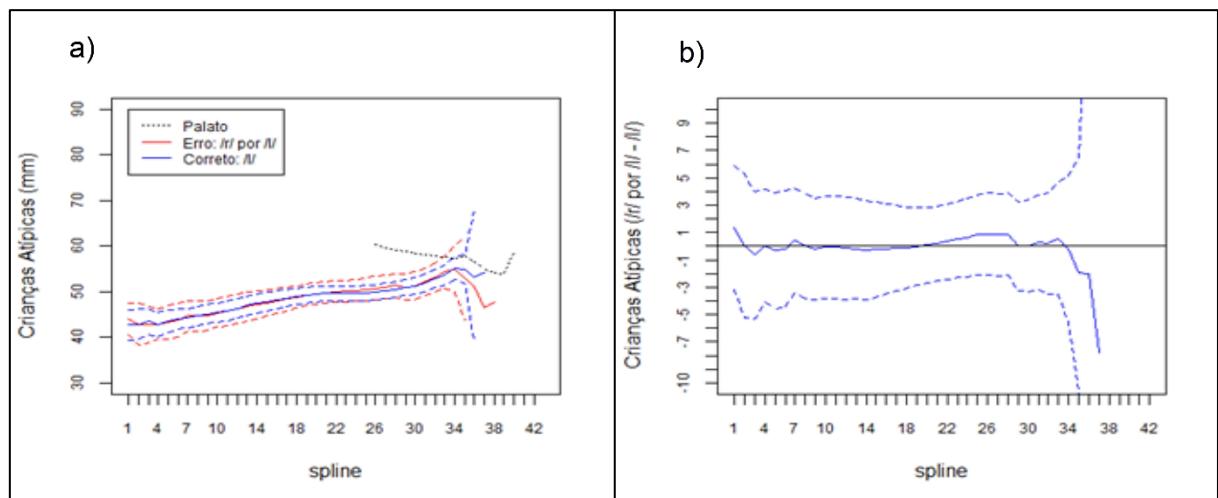


Gráfico 5 - a) intervalo de confiança para os contornos médios de língua para alvos [r] produzidos como [l] e para alvos [l] – CA; b) diferenças para os contornos médios entre esses sons. Em vermelho, os traçados para as substituições de [r] para [l] e, em azul, os traçados para as produções de [l]. FONTE: BARBERENA (2016, p. 158)

É possível observar que os traçados para a produção correta de [l] e para as substituições de [r] por [l] apresentaram proximidade no que concerne à região do dorso da língua. Apesar disso, quando ocorreram as substituições, tanto na raiz quanto na ponta da língua, há pistas para possíveis tentativas de produção do [r]. Essa maior variação de movimentos “sugerem acertos gradientes na fala das CA, que pela

análise perceptivo-auditiva assemelhou-se ao [l]” (BARBERENA, 2016, p.158). Sendo assim, a autora identificou tentativas de acertos graduados nas produções dos sujeitos atípicos. Enquanto a oitiva identificava uma substituição do *tap* por [l], os contornos médios da raiz e ponta da língua indicavam as tentativas de produção do segmento alvo.

Acerca das substituições do *tap* [r] pela semivogal [j], 9 crianças substituíam o segmento pela semivogal [j]. O Gráfico 6 apresenta os traçados da produção de [j] e das substituições de [r] por [j] do grupo CA.

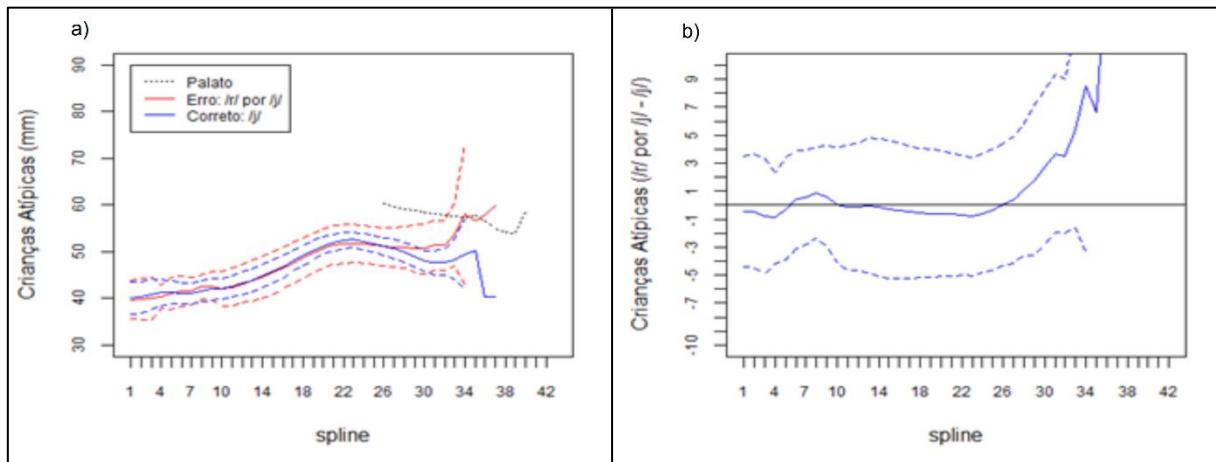


Gráfico 6: a) intervalo de confiança para os contornos médios de língua para alvos [r] produzidos como [j] e para alvos [j] – CA e b) diferenças para os contornos médios entre esses sons. Em vermelho, os traçados para as substituições de [r] para [j] e, em azul, os traçados para as produções de [j]. FONTE: BARBERENA (2016, p. 159)

Os traçados permitem observar que os gestos de ponta de língua para [r], produzido como a semivogal [j], demonstram maior magnitude quando comparado com as produções em que o [j] era o alvo das produções. Em algumas tentativas, a ponta da língua atinge a região alveolar. Isso demonstra claras tentativas de produção do *tap* quando ocorreram as substituições, ou seja, a pesquisadora observou tentativas de atingir a região alveolar para a produção do alvo.

Conforme exposto, os trabalhos baseados em oitiva não permitem a observação do detalhe fonético. Esses detalhes ajudam a melhor compreender os fatos fônicos do processo de aquisição. As teorias dinâmicas de produção da fala oferecem suporte para que os processos graduados possam ser identificados e analisados, por exemplo, os contrastes encobertos. A distinção entre o fonético e o fonológico é desfeita quando o aspecto temporal é incorporado ao sistema. Sendo assim:

A interpretação desse fenômeno é favorecida pela premissa da FonGest de não separar a tarefa motora que a criança realiza e a representação desta tarefa. Dessa forma, torna-se possível atribuir a nuances articulatórias – apreendidas por meio do detalhe fonético – um estatuto de contraste em construção e, assim, investigar as produções apresentadas por uma criança que se revelam intermediárias a duas categorias fônicas. Ao invés de interpretar um contraste como presente ou ausente, a FonGest abre a possibilidade de pesquisar a sua construção gradual (FREITAS; ALBANO, 2012, p. 198).

Albano (2007) ressalta a importância que os achados acústicos apresentam para essas teorias, pois a descoberta de sutis ou encobertas regularidades responde mais do que a pergunta “saber o que”, mas traz a resposta de “saber como”. O uso da ultrassonografia também pode auxiliar na investigação de produções intermediárias de determinados segmentos, pois a ferramenta viabiliza a inspeção articulatória, resgatando, de forma direta, aspectos que constituem a gramática fonológica.

### 2.3 Fonologia Articulatória

A Fonologia Articulatória surgiu nos Estados Unidos durante os anos 80. Seu propósito inicial era construir um protótipo de síntese de fala, hoje disponibilizado como TADA (*Task Dynamic Application*)<sup>10</sup>. Na construção desse modelo teórico, pesquisadores como Carol Fowler, Elliot Saltzman, Catherine Browman e Louis Goldstein, ao perceberem que, para a produção de um som, vários articuladores precisariam ser coordenados no tempo, decidiram investigar quais seriam os “átomos” básicos da linguagem, menores unidades linguísticas, que constituem as estruturas fonológicas.

Compostos por ações primitivas dos articuladores no trato vocal (oral, laríngeo ou vélico), os autores os chamam de gesto, ou seja, uma constrição no trato que pode ser formada pela ação de um dos três conjuntos relativamente independentes (lábios, ponta e corpo de língua). Os gestos têm um tempo intrínseco (*timing*) que se associa aos movimentos através do espaço e ao longo do tempo.

Com essa perspectiva, a fonologia seria como as “moléculas” que se reúnem a partir desses gestos, portanto, o primitivo de análise, o gesto articulatório, é capaz de possuir concomitantemente informações físicas e abstratas. O elemento físico é formado por unidades de ação, ao passo que o elemento abstrato contém unidades de informação contrastivas. Assim, a teoria não separa fonética e fonologia, uma vez que a unidade discreta de ação é composta por informações fonéticas e fonológicas, como estruturas físicas e cognitivas de um único sistema.

Feitas essas observações, é importante ressaltar que Fowler (1980) preconizou um modelo dinâmico para analisar os fenômenos fonológicos e Browman e Goldstein (1986, 1989) materializaram essa proposta a partir da Fonologia Gestual. Tal proposta dinâmica utiliza uma equação para descrever a relação temporal de um elemento no espaço. De acordo com Albano (2012), as equações dinâmicas descrevem a variação de algo no tempo, sendo, portanto, importantes para a implementação do modelo.

Um ponto importante a ser destacado é o fato de segmentos e gestos articulatórios não apresentarem uma correspondência entre si:

---

<sup>10</sup> TADA é uma implementação, em um *software*, do modelo *Task Dynamic* de coordenação. O sistema gera modelos articulatórios de enunciados em inglês a partir da entrada de texto fonético ou ortográfico e pode gerar arquivos de síntese de fala. O sistema TADA também foi projetado para permitir o desenvolvimento de modelos articulatórios em outros idiomas através do uso de arquivos de dicionário específicos de idiomas e bancos de dados gestuais. Link para acesso: [http://www.haskins.yale.edu/tada\\_download/index.php](http://www.haskins.yale.edu/tada_download/index.php)

O gesto não corresponde ao segmento ou ao traço, ficando a meio caminho entre os dois. Um [b], por exemplo, envolve apenas a coordenação de um gesto de oclusão labial com um gesto de adução das pregas vocais, enquanto um [m] envolve os mesmos gestos coordenados ao gesto vélico de abertura, cujo início e fim não precisam coincidir inteiramente com os daqueles (ALBANO, 2001, p.54).

Durante o processo de aquisição da linguagem existem as unidades pré-linguísticas que surgem antes de qualquer comportamento linguístico e maturação do sistema cognitivo da criança. Os padrões fonológicos de uma língua são criados a partir de unidades de ações discretas, que se tornam unidades básicas da fonologia. Desse modo, segundo Browman e Goldstein (1989), as representações abstratas seriam adquiridas por meio de rotinas articulatórias e, posteriormente, se tornariam unidades de contrastes. Indícios desse desenvolvimento podem ser observados no balbucio dos bebês, os quais apresentam gestos simples e grosseiros de constrição e que, posteriormente, irão desenvolver padrões de coordenação intergestual, capazes de diferenciar e afinar os gestos.

O sistema abstrato da criança desperta a partir do momento em que ela utiliza o conjunto de gestos disponíveis no seu inventário fonético para formar unidades complexas com valor distintivo. Segundo Browman e Goldstein (*op cit*), existem dois processos relevantes que separam a fala da criança da do adulto: diferenciação/afinação dos gestos individuais e coordenação desses gestos. No entanto, o mesmo primitivo de análise pode ser usado para descrever o sistema fonológico de adultos e crianças, pois as estruturas fonológicas são padrões de unidades gestuais organizadas.

Um gesto é determinado, especificamente, por um conjunto de equações que formam um modelo de *dinâmica de tarefas* (SALTZMAN e KELSO, 1987). Essas equações determinam os movimentos dos articuladores no tempo para que uma tarefa dinâmica específica seja alcançada. Para Browman e Goldstein (1989), na dinâmica de tarefas, a meta de um determinado gesto é especificada pelas variáveis do trato, cujo valor é determinado pelos movimentos de um conjunto de articuladores que as envolvem. Em outras palavras, o gesto articulatório é formado a partir da ação de um sistema neuromotor encarregado de coordenar inúmeros músculos e articuladores que formam, no trato vocal, as constrições. Por exemplo, para realizar o gesto de abertura labial (LA), o qual corresponde à distância vertical entre os lábios, três articuladores são acionados concomitantemente para alterar o valor de LA: mandíbula, lábio inferior e superior, contando com a participação de quinze músculos

para a realização da constrição labial. O movimento desses articuladores irá interferir no valor da variável LA, pois, para realizar essa tarefa gestual, esses precisam se descolar ao longo do tempo. A Figura 10 especifica cada variável e seus respectivos articuladores.

tract variable	articulators involved
LP	lip protrusion
LA	lip aperture
TTCL	tongue tip constrict location
TTCD	tongue tip constrict degree
TBCL	tongue body constrict location
TBCD	tongue body constrict degree
VEL	velic aperture
GLO	glottal aperture

Figura 10 - Variáveis do trato e representação do trato vocal com a localização das respectivas variáveis. Fonte: BROWMAN; GOLDSTEIN (1989, p. 73)

Os gestos são caracterizados pelos valores conferidos aos parâmetros dinâmicos e contínuos de cada variável do trato, esses são: grau de constrição, localização e rigidez. É necessário ressaltar que um som pode empregar as mesmas variáveis do trato oral, porém, o caráter distintivo é definido pelo ajuste dos valores desses parâmetros para cada gesto. O Quadro 2 apresenta os conjuntos de articuladores e os parâmetros dinâmicos associados.

CONJUNTO DE ARTICULADORES	PARÂMETROS DINÂMICOS ASSOCIADOS		
<b>Lábios (LIPS)</b>	Grau de constrição	Local de constrição	Rigidez
<b>Ponta de língua (TT – <i>Tongue tip</i>)</b>	Grau de constrição	Local de constrição	Formato de constrição
<b>Corpo de língua (TB – <i>Tongue body</i>)</b>	Grau de constrição	Local de constrição	Formato de constrição
<b>Raiz de língua (TR – <i>Tongue root</i>)</b>	Grau de constrição	Local de constrição	Rigidez
<b>Véu palatino (VEL – <i>Velum</i>)</b>		Grau de constrição	
<b>Glote (GLO – <i>Glottis</i>)</b>	Grau de constrição	Local de constrição	Rigidez

Quadro 2 - Conjuntos de articuladores e parâmetros dinâmicos associados, conforme proposta de Browman e Goldstein (1989, p. 74). Fonte: adaptado de BARBIERI (2019, p. 45)

Cada parâmetro dinâmico possui um conjunto de descritores associados. Para o grau de constrição, existem os seguintes descritores: fechado, crítico, estreito, médio, aberto; já para o local de constrição, são utilizados: protruso, labial, dental, alveolar, pós-alveolar, palatal, velar, uvular e faringal.

Os gestos de um determinado segmento são organizados em uma constelação gestual, a qual é representada por uma pauta gestual que especifica os conjuntos de parâmetros dinâmicos para cada gesto e os intervalos temporais em que cada um está ativo, obedecendo a princípios de faseamento. São esses que determinam a coordenação gestual no espaço e no decorrer do tempo. Uma vantagem da abordagem gestual é que é possível representar, de maneira intrínseca, o aspecto temporal – *timing* – de um gesto, ou seja, como os gestos se coordenam temporalmente, tendo em vista que as teorias fonológicas outras postulavam o aspecto temporal como um componente extrínseco.

Em suma, a pauta gestual é uma ferramenta utilizada para formalizar a descrição de fenômenos fonológicos, pois apresenta as ações realizadas pelos articuladores e os aspectos temporais durante a produção da fala. Cada gesto é apresentado por caixas bidimensionais, estando representado, no eixo horizontal, o tempo de ativação de cada gesto; já no eixo vertical, é representada a magnitude do gesto produzido. Sendo assim, as variações nas proporções da caixa equivalem às variações desses parâmetros. Na Figura 11, é possível observar dois exemplos elencados por Browman e Goldstein (1989).

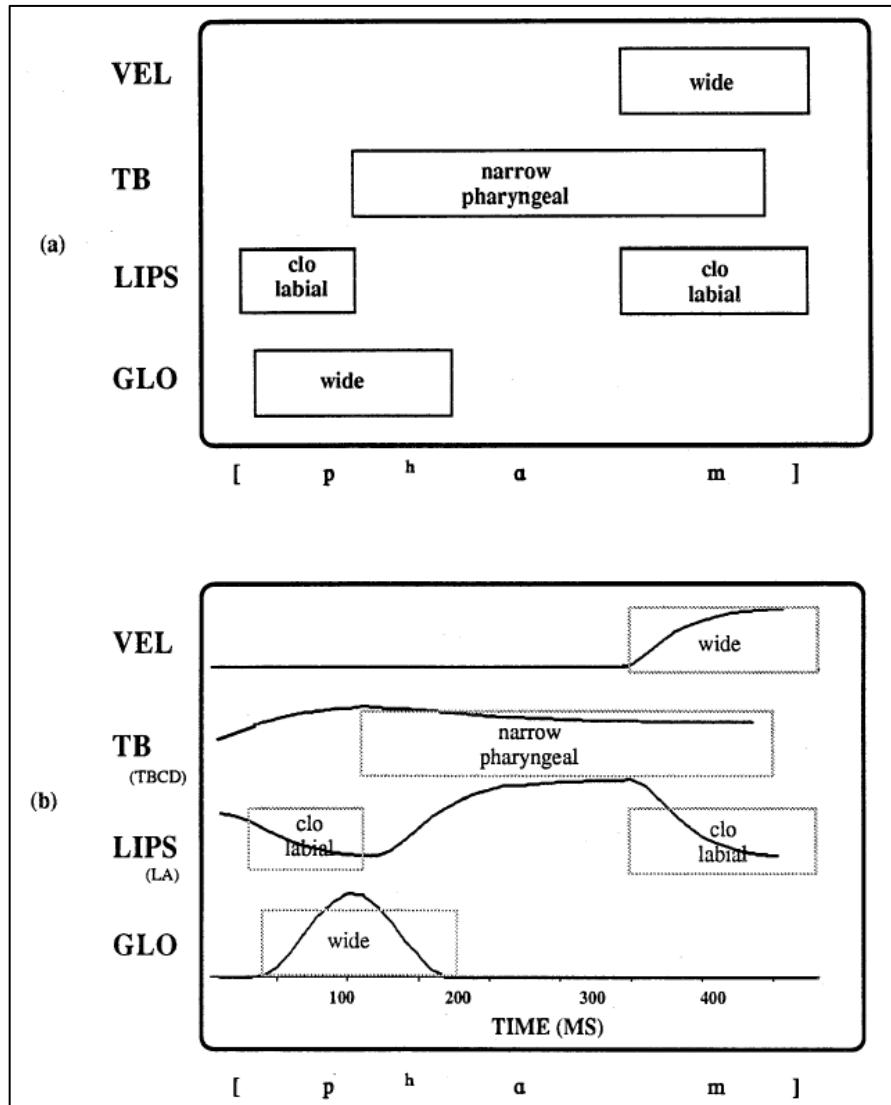


Figura 11 - Pautas gestuais da palavra *palm*, do inglês. Fonte: BROWMAN; GOLDSTEIN (1989, p. 76)

Conforme Silva (2002), em acordo com Browman e Goldstein (1986, 1989, 1992), um evento esperado nos sistemas dinâmicos, e que também pode ser responsável por muitos fenômenos fonológicos, é a sobreposição de gestos. Em outras palavras, a presença de encadeamento dos sons ou efeitos de coarticulação são resultados da sobreposição entre gestos. Sendo assim, para representar as possíveis sobreposições gestuais na Figura 11, em (a), é possível observar que cada *tier* (linha) evidencia o conjunto de articuladores utilizado na produção dos gestos. Os descritores gestuais estão localizados dentro das caixas e, quando pertinente, é apresentada a informação referente aos valores dinâmicos, por exemplo, grau e local de constrição. Em (b), além das informações contidas em (a), tem-se a inclusão de curvas. Essas exibem a variação dos articuladores durante suas trajetórias. Com

base nas curvas, é possível observar os intervalos de ativação que, por sua vez, registram a duração de cada gesto.

Com base no modelo proposto por Albano (2001) – FAAR –, Silva, Clemente e Nishida (2006) apresentam uma pauta gestual para a produção do *tap* em coda final no PB, conforme a Figura 12.

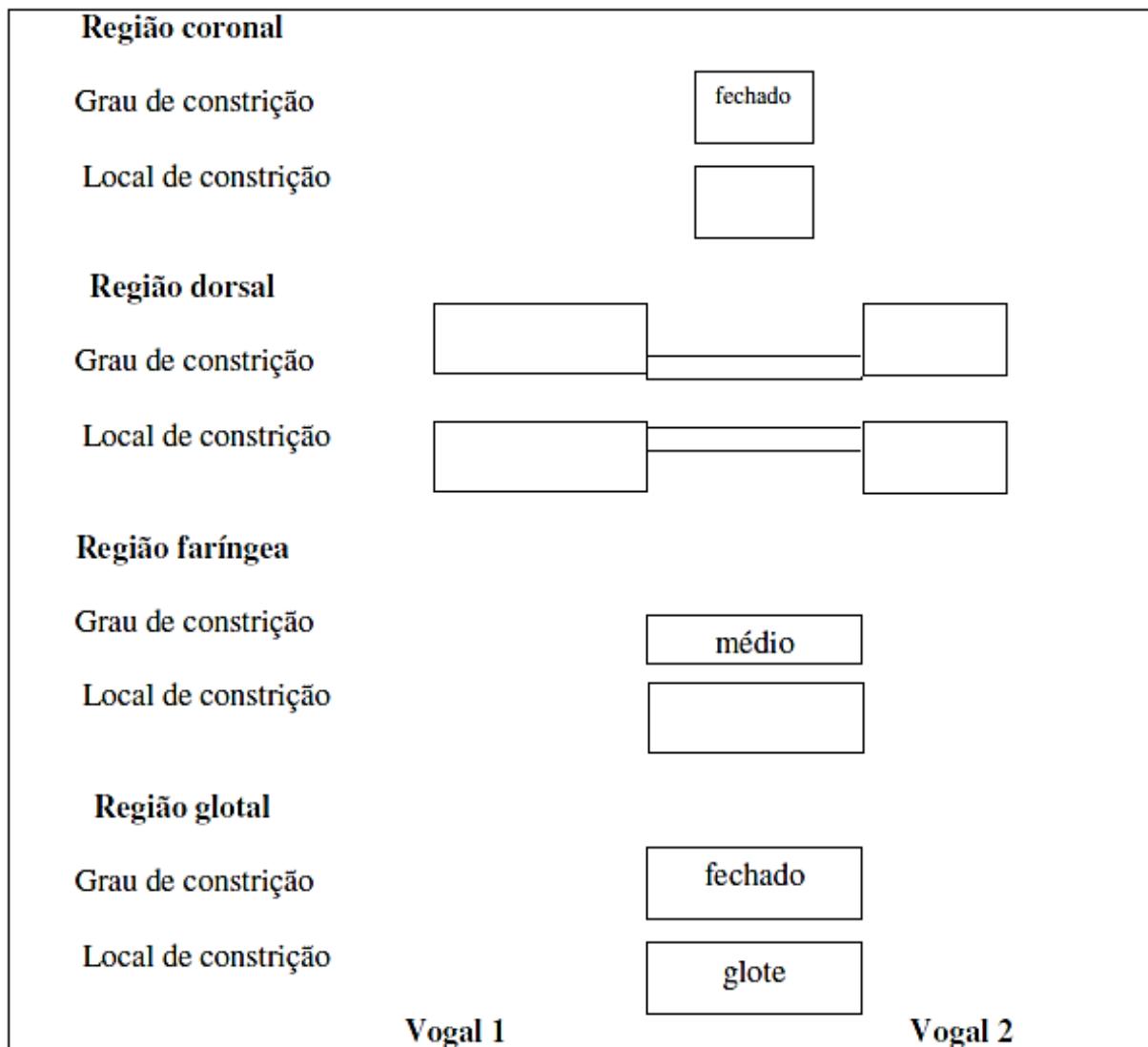


Figura 12 - Pauta gestual para o *tap* em codas. Fonte: SILVA; CLEMENTE; NISHIDA (2006, p. 22)

É possível observar, na pauta proposta pelos autores, uma relação entre o gesto de ponta com o gesto de dorso de língua envolvidos na produção do *tap*. Parte do gesto dorsal é encoberto pelo gesto coronal de maior magnitude, que tem um tempo de ativação menor se comparado com os demais gestos. Isso se dá por conta do movimento balístico da ponta da língua que realiza uma oclusão rápida no trato vocal. Na pauta gestual, esse movimento é caracterizado no primeiro conjunto de articuladores, sendo representado pelo descritor gestual *fechado*. O elemento vocálico

adjacente ao *tap*, que ocorre em coda final, à direita do segmento, é representado pelo gesto dorsal. Seu tempo de ativação é menor se comparado com a vogal núcleo da sílaba. O terceiro conjunto de articuladores apresenta um descritor gestual *médio* para a constrição faringal. Já o último conjunto de descritores representa a vibração das pregas vocais. Por meio da pauta gestual, os autores evidenciam que, quando o elemento vocálico estiver à direita, terá natureza e duração diferentes se comparado ao elemento vocálico que se encontra à esquerda, como nas sílabas CCV.

Uma representação gestual dos róticos do PB, em início de palavra, foi proposta por Silva (2002). A autora discute os processos de alofonia sofridos por esses segmentos, buscando identificar a existência de um contínuo físico na produção de /r/ e o papel do contexto prosódico. Discorre, ainda, sobre qual tipo de aporte teórico dá conta de representar esses fatos na gramática da língua. Seu trabalho aponta a existência de um condicionamento da estrutura prosódica na produção do rótico, pois, em fronteiras prosódicas fortes, versões fortalecidas do rótico tendem a ocorrer. O inverso ocorre em fronteiras prosódicas fracas, nas quais versões enfraquecidas do rótico são encontradas. A pesquisadora afirma que a alofonia dos segmentos róticos é gradiente e não categórica, somente sendo possível sua representação por meio de uma teoria que leve em conta o caráter dinâmico da linguagem. Sendo assim, “há um contínuo físico, ao longo do qual é possível encontrar ‘n’ variantes que se caracterizam por carregarem, concomitantemente, características das duas categorias” (SILVA, *op. cit.*, p.197). Esse estudo também identifica, na tentativa de identificar as variantes dos róticos, a ativação de gestos em mais de uma região articulatória.

No que concerne à estrutura silábica, a partir da perspectiva da Fonologia Articulatória, Browman e Goldstein (1988) exploraram duas abordagens. A primeira estaria relacionada a uma organização local, na qual gestos individuais ligam-se a outros gestos individuais. Já a segunda afirma que os gestos constituiriam conglomerados maiores, formando, assim, uma organização global. O estudo dos autores baseou-se em vocábulos do inglês e propõe que haveria, para as consoantes presentes no final de sílaba (VC), uma organização local, pois o gesto da vogal seria coordenado com o gesto da consoante, mantendo sua unidade temporal. Relação similar é encontrada em início de sílaba (CV), no entanto, o alinhamento da consoante, nesse caso, dá-se com o início da vogal, apresentando, assim, uma única unidade de tempo. Quando uma sequência de consoantes está presente no início da sílaba,

haveria uma organização global e essa seria organizada por uma medida: o *C-center* (centro-C), por meio desse centro, a consoante se coordenaria com o gesto da vogal.

Considerando que os gestos têm particularidades distintas, assim como sua extensão temporal, esses aspectos podem explicar as sobreposições gestuais durante a produção da fala. Assim, é possível explicar, sob uma outra ótica, fenômenos que antes eram caracterizados como apagamentos, substituições, dentre outros.

Segundo (GOLDSTEIN *et al.*, 2007), a partir da observação de dados de cinemática articulatória, sob uma abordagem de sistemas dinâmicos, são proporcionados registros pertinentes para a análise de erros de fala, bem como é possível revelar a natureza e existência das unidades de ação existentes na fala.

Para os autores, os erros apresentam um padrão de ocorrência e distribuição. Um exemplo disso ocorre na sequência *coffee pot* que poderia ser pronunciada como *poffee pot* ou *poffee cot*, contudo, a sequência *cottee poff* não seria produzida. Com base nesse exemplo, os autores pressupõem que os erros de fala estão sujeitos às leis da fonologia e desenvolvem dois experimentos com adultos. Os informantes deveriam repetir uma sequência de palavras dissílabas, aumentando a velocidade da fala e intercalando a consoante do início da sílaba, por exemplo em *tip kip* e *cop top*.

Os erros de produção foram organizados conforme as seguintes informações gestuais: (i) intrusão; (ii) redução e (iii) substituição. O erro de intrusão ocorre quando há “a adição de um gesto não produzido naquele local temporal em uma produção normal e sem erros”, de acordo com Goldstein *et al.* (*op. cit.*, pág. 6, tradução nossa<sup>11</sup>). O erro de redução gestual dá-se quando há “uma redução inadequada na magnitude de um movimento articulatório pretendido, por exemplo, uma menor elevação do dorso da língua durante a produção de /k/” (GOLDSTEIN *et al.*, *op. cit.*, pág. 6, tradução nossa<sup>12</sup>). Os erros de intrusão e de redução podem ocorrer de forma total ou parcial, já o erro de substituição surge quando acontece uma intrusão e redução de gestos simultaneamente. Os resultados do estudo dos autores apontam uma maior ocorrência de erros de intrusão, pois cerca de 28,2% dos casos foram registrados

---

<sup>11</sup> No original: *the addition of a gesture not produced at that temporal location in a normal, non-errorful production.*

<sup>12</sup> No original: *an inappropriate reduction in the magnitude of an intended articulatory movement, for instance a smaller tongue dorsum raising during /k/.*

como tal. Os erros de substituição ocorrem em 4,3% dos dados e, por último, cerca de 3,3% dos erros são classificados como redução gestual.

Em suma, é possível perceber que modelos com um primitivo de análise categórico não dão conta de representar fenômenos de natureza gradiente. Sendo assim, apoiar-se em modelos dinâmicos de produção de fala tem sido o recurso amplamente empregado para caracterizar e descrever esses fenômenos na gramática de uma língua. É importante ressaltar que, a partir dessas representações, é possível identificar como os padrões de coordenação gestual surgem e se estabilizam na fala da criança.

### **3 Metodologia**

Este capítulo descreve o método empregado em cada uma das etapas necessárias para a realização do presente estudo. Inicialmente, serão apresentadas as informações referentes aos informantes, cujo dados de produção oral constituem o *corpus* a ser investigado. Na sequência, serão detalhados os procedimentos realizados para as coletas de dados, bem como os instrumentos utilizados. Por fim, serão especificados os parâmetros adotados para as análises de oitiva e acústica.

#### **3.1 Caracterização da amostra**

Tratando-se de uma pesquisa prospectiva e descritiva, com amostra limitada por conveniência, a seleção dos sujeitos seguiu alguns critérios previamente definidos. Contudo, para todos os informantes, definiu-se um critério de exclusão: não ter sido submetido a tratamento fonoaudiológico anterior, tampouco no momento do desenvolvimento dessa pesquisa.

Ainda que as crianças não tenham sido submetidas a uma avaliação fonoaudiológica, estas foram consideradas dentro do padrão típico de desenvolvimento da linguagem. Durante as interações com as crianças, essas seguiram as instruções dadas sem qualquer dificuldade, mantiveram atenção na tarefa proporcionada e ouviram bem. Não somente a interação com a pesquisadora se deu de forma satisfatória, como não foi observada nenhuma troca frequente de ponto de articulação ou troca de sonoridade. Sendo assim, uma vez que as produções dos sons pelas crianças ocorreram conforme o esperado para as suas respectivas idades (LAMPRECHT *et al.*, 2004; ILHA CERON *et al.*, 2019), os sujeitos foram considerados aptos a contribuírem com o trabalho.

Esse primeiro contato com os possíveis informantes aconteceu por meio de uma entrevista que propiciou à criança uma produção de fala próxima ao natural e espontânea. Aos responsáveis foi apresentado um documento (Apêndice A), elucidando como aconteceria a entrevista e apresentando os próximos passos do trabalho, caso a criança fosse considerada apta a participar da pesquisa. Para que fosse autorizado o contato com o menor e a gravação deste primeiro momento, os responsáveis assinaram uma autorização (Apêndice B) que permitia a realização da entrevista.

Nessas entrevistas, foi preenchida uma ficha de cadastro (Apêndice C) que auxiliou no cumprimento dos critérios de inclusão, apresentados a seguir: (i) as crianças deveriam ser naturais da cidade de Pelotas/RS e nunca terem habitado outra cidade, com o intuito de evitar a influência de outro grupo linguístico nas produções orais; (ii) os informantes deveriam ser monolíngues, tendo o português brasileiro como língua materna e não apresentarem contato com qualquer língua estrangeira; (iii) as crianças deveriam estar em processo de aquisição do *tap* em final de sílaba e palavra, apresentando, assim, produções variáveis ou com a presença de um glide ou da líquida lateral ocupando a posição que deveria ser do *tap*. Esse critério justifica-se uma vez que o propósito do trabalho é investigar como esse segmento emerge na fala infantil. Salienta-se, no entanto, que o sujeito controle deveria já ter superado a fase de aquisição, produzindo todos os sons, em todas as posições silábicas, visto que seus dados serão usados como parâmetro de comparação com os dados dos informantes em fase de aquisição.

Durante as entrevistas, foi utilizado um gravador digital, modelo *Zoom H4n*, a fim de gravar a conversa com as crianças. Para provocar a produção de determinados itens lexicais, foram apresentadas figuras (Anexo A) na tela de um *notebook* da marca *Samsung*, modelo *Expert X30*. Esses foram produzidos de forma isolada, sem controle de repetições. Em relação às palavras apresentadas aos entrevistados, foram utilizados os vocábulos presentes em Melcheque, Ferreira-Gonçalves e Brum-de-Paula, (2021) e alguns do trabalho de Barbieri (2019). Em ambos, os itens lexicais foram exibidos aos informantes por meio de imagens, nos quais o *tap*, em sílaba tônica, ocupava as posições de CV e CVC medial, CCV inicial e CVC final, conforme Quadro 3.

	<b>Melcheque; Ferreira-Gonçalves; Brum-de-Paula (2021)</b>
<b>CV</b>	barata – siri - sururu
<b>CCV</b>	prato – frita - frutas
<b>CVC Medial</b>	carta – circo – curto
<b>CVC Final</b>	mar – vir – cur
<b>Distratores</b>	bolsa – gato – casa – chinelo – galinha - sapato
	<b>Barbieri (2019)</b>
<b>CCV</b>	prego – preso – broa – frota

Quadro 3- Relação das palavras utilizadas no instrumento da entrevista

A pesquisadora apresentava as imagens e fazia perguntas para as crianças de forma a propiciar a produção do alvo, para que fosse permitido observar a presença ou ausência, na fala das crianças, do fenômeno de investigação alvo deste estudo. Além disso, foi possível obter essas produções de forma mais natural, estabelecendo assim, uma conversa entre ambas as partes. Com isso, era possibilitado a observação de produções outras das crianças, a fim de saber em qual fase de aquisição essa encontrava-se. Paralelamente, essas produções permitiram que fossem observadas possíveis ocorrências de desvios de fala.

É necessário destacar que os áudios das entrevistas apresentavam ruídos externos por conta do seu local de gravação, impossibilitando uma análise acústica confiável. No entanto, por mais que esses áudios não sejam submetidos a qualquer tipo de análise quantitativa ou qualitativa, essa etapa foi fundamental para a seleção dos informantes que fazem parte do *corpus* investigativo desta pesquisa.

### 3.2 Os sujeitos

Participaram da pesquisa seis crianças com desenvolvimento típico de fala. Cinco crianças em fase de aquisição do *tap* em final de sílaba e uma com a fase de aquisição superada e com o segmento alvo completamente adquirido. Esses informantes foram divididos em três grupos: (i) sujeito longitudinal (SL); (ii) sujeitos transversais (ST) e (iii) sujeito controle (S0). Os responsáveis legais pelas crianças assinaram um termo de consentimento livre e esclarecido (Apêndice D), autorizando que os achados dessa pesquisa, resultantes das produções dos menores, fossem

analisados e apresentados de forma pública, porém, preservando a identidade dos envolvidos. Esse termo foi desenvolvido com base em Barbieri (2019).

### 3.2.1 Sujeito longitudinal

O sujeito longitudinal (SL1) é uma menina que tinha 3 anos e 8 meses, na data da primeira coleta de dados e atendia a todos os critérios de inclusão. No dia da entrevista com a criança, com base em análise de oitiva, foi observado que ela ainda não produzia o *tap* em nenhuma posição silábica, no entanto, produzia o som de [l] em algumas posições silábicas nas quais o alvo principal deveria ser o [r]. Em sílabas CV, o alvo *aranha* foi produzido com a substituição do *tap* pela lateral [l], por exemplo. Quando o segmento era alvo de encontros consonantais, nenhuma produção ocorreu para o segmento alvo. Para as sílabas CVC, o *tap* foi omitido na posição medial, como em [si.ku], para *circo*. Para a posição final de palavra, ocorreram dois fenômenos: o primeiro foi a substituição pela semivogal [j], como em [maj] para *mar*, e a substituição pela líquida [l] mais a inserção da vogal [i], como em [a.ba.Zu.li] para *abajur*. O Quadro 4, apresenta a relação das coletas realizadas com a informante.

	<b>Idade</b>	<b>Data da coleta</b>
1	3;08;09	20/01/2020
2	3;08;24	04/02/2020
3	3;09;17	28/02/2020

Quadro 4 - Coletas realizadas com o sujeito longitudinal

Estavam previstas coletas quinzenais com a informante durante um período de cinco meses, o que totalizaria dez amostras de dados. No entanto, com o surgimento da doença Covid-19<sup>13</sup>, provocada pelo coronavírus SARS-CoV-2, as coletas precisaram ser suspensas. Sendo assim, ocorreram três coletas de dados com a criança.

<sup>13</sup> O primeiro caso no Brasil surgiu dia 26/02/2020. A OMS (Organização Mundial de Saúde) declarou estado de pandemia em 11/03/2020. Com isso, a Universidade Federal de Pelotas suspendeu suas atividades no dia 16/03/2020.

### 3.2.2 Sujeitos transversais

Os quatro sujeitos transversais realizaram apenas uma coleta. A primeira informante (ST1) é uma menina com 3 anos e 2 meses. Na entrevista, a criança apresentava produção com *tap* em palavras como *caro*, no entanto, nos encontros consonantais, nenhum alvo foi produzido. Para itens lexicais com o segmento alvo em posição final de sílaba, este foi substituído pela semivogal [i], como nas palavras *curto* e *mar*. O Quadro 5 apresenta detalhes sobre os informantes: sexo, as datas das coletas e a idade das crianças no dia das gravações.

O segundo sujeito (ST2) é um menino com 3 anos e 4 meses. Durante a entrevista, o menino apresentou produções em CV, como em *barata*; para os alvos em CCV, o *tap* não foi produzido, como em *bruxa*. Porém, quando o segmento se apresentava no final da sílaba, por vezes era omitido, como em *carta* e por vezes era produzido, como em *azar*.

O terceiro sujeito (ST3) é um menino com idade de 3 anos e 11 meses. No dia da entrevista com a criança, foi observado que ela apresentava produções variáveis do *tap*, pois foi produzido o segmento na palavra *carta*, mas, no alvo *garfo*, a produção não aconteceu, embora, em ambos vocábulos, o segmento alvo estivesse em posição CVC medial.

Por fim, a quarta informante (ST4) é uma menina com 4 anos e 5 meses. No dia da entrevista, foi observado que a criança produzia o *tap* em sílaba CV, como na palavra *Maria*; em posição CCV, o elemento era omitido, como ocorreu no vocábulo *quatro*. Para as produções em CVC, foram observadas produções variáveis, para ambas as posições do *tap* na sílaba, medial ou final. Por exemplo, no alvo *curta*, o segmento foi substituído pela semivogal [i] e, para a palavra *lugar*, o alvo foi omitido.

Sujeito	Sexo	Idade	Data da coleta
ST1	Feminino	3;02;16	28/02/2020
ST2	Masculino	3;04;10	15/03/2020
ST3	Masculino	3;11;18	02/12/2019
ST4	Feminino	4;05;21	03/03/2020

Quadro 5 - Coletas com os sujeitos transversais.

O Quadro 6 sintetiza as informações acerca da ocorrência do *tap* e quais os tipos de recurso o informante utilizava quando não alcançava a produção do alvo, no dia da entrevista. As produções estão divididas por informantes e por posição silábica.

	ST1				ST2				ST3				ST4			
	CV	CVC <sub>M</sub>	CVC <sub>F</sub>	CCV	CV	CVC <sub>M</sub>	CVC <sub>F</sub>	CCV	CV	CVC <sub>M</sub>	CVC <sub>F</sub>	CCV	CV	CVC <sub>M</sub>	CVC <sub>F</sub>	CCV
'r'	S	V	N	N	S	V	V	N	S	V	V	V	S	N	V	N
ø	-	V	V	-	-	V	V	-	-	V	V	V	-	V	V	-
'p'	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	-
'j'	-	V	V	-	-	-	-	-	-	-	-	-	-	V	V	-

Quadro 6 - Ocorrência do *tap* e recursos utilizados pelos informantes no dia da entrevista. As produções estão divididas por informantes e por posição silábica. S: sempre; N: não, V: variável.

Para todos os sujeitos, foram observadas as características previstas dentro dos critérios de inclusão e nenhum fator que permitisse a exclusão foi identificado. Todos os informantes, portanto, sempre produziam o *tap* em sílaba CV, apresentavam produções variáveis em coda medial e final (à exceção de ST1 para CVC final) e não realizavam nunca o *tap* em sílaba CCV (à exceção de ST3).

### 3.2.3 Sujeito controle

Assim como os sujeitos transversais, os dados do sujeito controle (S0) foram coletados em um único encontro. O referido sujeito é um menino e atende a todos os critérios de inclusão, dentre eles o que trata da fase de aquisição completamente superada, uma vez que apresenta, inclusive, a produção de todos os encontros consonantais. No Quadro 7, é possível observar a data da coleta e a idade da criança na respectiva data.

Idade	Data da coleta
6;07;08	30/01/2020

Quadro 7 - Coleta com o sujeito controle.

Na sua entrevista, nenhuma troca de som alvo foi observada, assim como ausência de produções. A criança permaneceu atenta durante toda a gravação e mostrou-se bem comunicativa.

### 3.3 Coletas de dados

Com um instrumento previamente delineado para esta fase, segundo critérios linguísticos que serão apresentados posteriormente, os informantes tiveram suas produções gravadas individualmente, em uma cabine com isolamento acústico, localizada no Laboratório Emergência da Linguagem Oral (LELO), na Universidade Federal de Pelotas. A utilização da cabine é indispensável para preservar a qualidade do áudio, minimizando os ruídos exteriores. A Figura 13 exibe o interior da cabine.

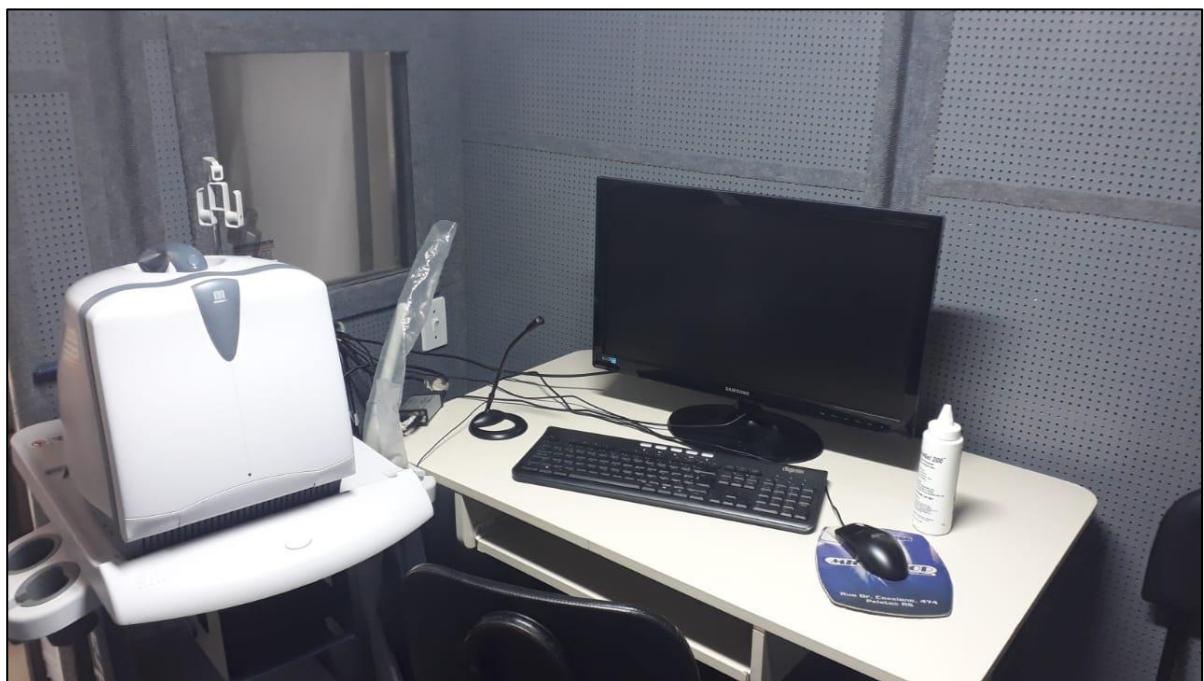


Figura 13 - Interior da cabine com isolamento acústico, utilizada para coletas acústicas e articulatórias, LELO/UFPel. Fonte: a autora

Para a captação dos dados, foi utilizado um gravador digital, modelo *Zoom H4n* (o mesmo aparelho da seleção dos informantes) e cartões com as imagens (Anexo B) que deveriam ser nomeadas pelos informantes. Os dados de áudio foram gravados em formato *WAV*, com taxa de amostragem de 44.000 Hz. Para a confecção dos cartões, foram pesquisadas imagens, no *site* de busca Google, que correspondessem direta ou indiretamente aos alvos a serem nomeados. Após a escolha, as figuras foram editadas no *software Photoscape*, versão 3.7, a fim de uniformizá-las para a impressão em papel *Couchê 300g*. Cada imagem foi impressa três vezes, com o intuito de proporcionar ao informante três repetições de modo aleatório.



Figura 14 - Gravador Zoom H4n e cartões utilizados para o instrumento de coleta.

Para a realização das coletas, elaborou-se um *corpus* de tríades – seguindo Rodrigues (2007) – que apresentavam, em sílaba tônica, o *tap*, a sua omissão e a presença de ditongo constituído pelo glide palatal. Para o contexto consonantal, deu-se preferência para palavras que possuíssem oclusivas e segmentos desvozeados, visando diminuir os efeitos de coarticulação e facilitar a identificação do som no momento da segmentação dos dados. O alvo de investigação encontrava-se em posição medial ou final de palavra. Para o ambiente vocálico, foram escolhidas as vogais das extremidades do triângulo vocálico: [a], [i] e [u]. O instrumento conta com um total de 26 palavras que foram repetidas três vezes, assim, totalizando 78 *tokens* a cada coleta. No Quadro 8, é possível observar a relação das palavras utilizadas no instrumento de coletas da segunda etapa.

CONTEXTO TÔNICO	[a]	[i]	[u]
<b>CVC MEDIAL</b>	carta x cata x caixa farpa carpa	circo x chico círculo –	curto x cuca x cuida turco curta
<b>CVC FINAL</b>	mar x má x pai lugar azar	vir x vi Zair Nadir	cur x tatu abajur Artur

Quadro 8 - Relação dos pares de palavras utilizadas no instrumento de coleta

Sobre a ausência de alvo lexical constituído pelo ditongo [uj], salienta-se que, ao elaborar a metodologia da presente pesquisa, houve uma falha na seleção da palavra que, provavelmente, transcorreu de um erro de digitação. O item lexical que

foi gravado para ser o par dos alvos semivocalizados em posição final foi *rubi*, no entanto, o correto seria a palavra *Rui*.

Devido à dificuldade de se encontrar vocábulos que atendessem a todos esses critérios, e após todas as opções serem esgotadas, algumas exceções foram admitidas. Apesar disso, não foi permitido o uso de não-palavras e buscou-se a máxima aproximação com palavras que fossem familiares às crianças com as idades alvo desse estudo. É importante considerar que, das 26 palavras presentes no instrumento, 15 são representadas diretamente pelas imagens, ao passo que 11 não correspondem assertivamente aos alvos. Assim, precisava ser explicado às crianças qual alvo determinada imagem representava. Por exemplo, o cartão da palavra ['kur]<sup>14</sup> corresponde a uma imagem de um cachorro. Era, então, ensinado à criança que o nome daquele cachorrinho era *cur*. Isso ocorria quando se tornava inviável que, por meio de uma mímica, por exemplo, se induzisse a criança a descobrir o alvo de produção, uma vez que a palavra não fazia parte do vocabulário do informante. É evidente que, nesses casos, a pesquisadora precisava dizer qual era a palavra e, após mostrar outras figuras, solicitava que a criança repetisse a palavra apresentada anteriormente. Para não ocorrer interferências durante a coleta de dados, devido a problemas de identificação das imagens, foi realizado um momento de habituação com o informante antes de cada coleta.

Ainda que alguns vocábulos do instrumento evidenciem disparidades no que concerne aos contextos linguísticos prefixados e a familiaridade das crianças com as palavras, estas não foram alteradas pelas seguintes razões: (i) não foi encontrado exemplares que, além de abrangerem os contextos selecionados, fossem possíveis de ilustrar por meio de imagens; (ii) as palavras abrangem a maior parte dos critérios linguísticos estabelecidos e (iii) não foi configurada necessariamente uma repetição de palavras, uma vez que a pesquisadora ensinava a palavra e pedia que a criança a pronunciasse depois da nomeação de outras imagens. Além disso, existem estudos que afirmam não haver diferenças entre instrumentos de nomeação ou imitação de palavras utilizados com crianças (WERTZNER; PAPP; GALEA, 2006).

Para a presença do *tap*, os vocábulos presentes em Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2021) foram mantidos; as palavras referentes à omissão

---

<sup>14</sup> Catahoula Cur é uma raça de cachorro que surgiu em Louisiana, nos Estados Unidos. No Brasil é chamado de Cão Leoparda da Catahoula. Fonte: <https://www.racasdecachorro.com.br/catahoula-cur/>

do segmento foram elaboradas pela presente pesquisadora; no que concerne às palavras com ditongos, foram utilizadas algumas do trabalho de Rodrigues (2007). No Quadro 9, é possível observar a relação dos itens lexicais e as pesquisas que os utilizaram para a investigação acerca dos róticos.

	<b>Melcheque et al (2021)</b>	<b>Da presente pesquisa</b>	<b>Rodrigues (2007)</b>
<b>CVC Medial</b>	carta – circo – curto	cata – caixa – carpa – farpa círculo – cuca – turco – curta	Chico – cuida
<b>CVC Final</b>	mar – vir – cur	azar – má – lugar – Nadir – vi Zair – abajur – Artur – tatu	pai

Quadro 9 - Relação de autores e de palavras utilizadas no instrumento de coleta

### 3.4 Análise de oitiva

A análise de oitiva descreverá o percentual de possibilidades e ocorrências de produções do *tap* e de suas variáveis. Para tanto, realizou-se a transcrição fonética das produções antes da análise acústica. Os índices de produções do *tap* e os recursos utilizados pelos informantes quando a produção não era atingida foram registrados previamente em quadros. Tendo em vista que o segmento-alvo está sendo analisado apenas em sílaba tônica, produções como [a.'tur] foram contabilizadas como ocorrência, pois, embora não produzido na primeira sílaba, o alvo encontrava-se na sílaba tônica, ou seja, no final da palavra, alvo deste estudo.

Essa análise foi aplicada para as palavras que abarcam o *tap* como alvo, essas são: carta, farpa, carpa, circo, círculo, curto, turco e curta – para o contexto medial – e mar, lugar, azar, vir, Zair, Nadir, cur, abajur e Artur – para o contexto final.

Para os casos de contrastes encobertos que foram detectados após a inspeção visual dos espectrogramas ou que a oitiva de uma palavra não estava clara, causando dúvida na hora da transcrição, os áudios foram submetidos à apreciação de juízes. O objetivo central dessa etapa do trabalho foi validar a presença de contrastes encobertos que podem ser encontrados na análise acústica, porém, não são percebidos auditivamente.

Para tanto, foi formado um grupo de três juízes com experiência em análises de oitiva e acústica de dados. Esse foi composto por três doutores ou doutorandos em

linguística, vinculados à Universidade Federal de Pelotas. Eles receberam uma carta convidando-os para participar da pesquisa e com orientações acerca de como proceder durante a análise dos dados. Em seguida, assinaram um Termo de Consentimento Livre e Esclarecido, no qual eles concordaram em participar do trabalho. Tendo essa etapa sido concluída, os juízes receberam os áudios produzidos pelos informantes em formato *WAV*. Em um formulário online, elaborado na plataforma *Microsoft Forms*, os juízes tiveram três opções de transcrições para analisar e escolher, conforme é possível observar na Figura 15. Um campo também foi disponibilizado para outras transcrições ou observações, caso os juízes julgassem necessário. Considerando o número expressivo de dados a serem apreciados pelos juízes, optou-se por dividir os alvos em quatro seções.

**Seção 1**

Para essa sessão, você irá abrir os áudios presentes na pasta "Seção 1" e reproduzi-los quantas vezes for necessário. Logo, marque uma das opções de transcrição fonética disponíveis ou faça a transcrição que julgar necessário na opção "outra".

1. Alvo 1

[aba.'ʒuj]  
 [aba.'ʒu.li]  
 [aba.'ʒul]  
 Outra

Figura 15 - Tela do formulário online com um exemplo de tarefa de identificação dos juízes.

Após o recebimento das avaliações feitas pelos juízes, foram elaborados os quadros finais, com as possibilidades de ocorrências do *tap* e seus índices de efetivas produções. É necessário enfatizar que, durante a elaboração dos referidos quadros, diversas verificações dos dados foram realizadas a fim de assegurar uma descrição sem equívocos.

O propósito desse tipo de análise é investigar se há algum contraste encoberto, que não é percebido auditivamente, na fala da criança durante a fase de aquisição, mas que, ainda assim, é evidenciado durante a análise acústica. Nesse sentido, atribui-se à análise de oitiva um papel secundário, pois a inspeção acústica dos dados não é empregada neste trabalho com o intuito de confirmar ou refutar a inspeção auditiva.

### 3.5 Análise acústica

A análise acústica dos dados foi feita por meio do software *Praat*, versão 6.1.16 (BOERSMA; WEENINK, 2019). Como a taxa de frequência ideal para a captação de dados linguísticos gira em torno de 22.000 Hz, é preciso que, para análise, a taxa de amostragem seja o dobro do valor da frequência máxima do sinal de fala (SILVA *et al.*, 2019). Sendo assim, foi empregada uma taxa de amostragem de 44.000 Hz para análise. Os valores das frequências dos formantes foram retirados por meio da análise LPC (*Linear Predictive Coding*). A visualização de um traçado LPC confiável exige valores adequados para *Maximum formant (hz)* e *number of formants*. Para dados de crianças, como é o caso da presente pesquisa, “o mais seguro é contar o número de faixas escuras até a máxima frequência de análise e usar esse número como ponto de partida” (BARBOSA; MADUREIRA, 2015, p. 159).

Nas próximas seções, serão detalhados os procedimentos adotados para organização e segmentação dos dados, bem como serão apresentados os critérios empregados nas análises acústicas.

#### 3.5.1 Organização e segmentação dos dados acústicos

Todas as coletas contaram com uma planilha que registrou suas informações e particularidades. Por meio deste documento, foi possível registrar prováveis alterações morfológicas (produção de *cartas* ao invés de *carta*) e motivações para descarte de dados sem qualidade para análise (como a presença de ruído sobre o alvo produzido, por exemplo).

Os arquivos, contendo os áudios, foram organizados em pastas diferentes de acordo com sujeito e posição do *tap* na palavra. Os áudios dispõem das seguintes informações: (i) identificação do informante – sujeito transversal (ST), longitudinal (SL) ou controle (S0) –; (ii) palavra alvo e (iii) número da repetição, por exemplo: *ST1\_carta\_1*. Para os dados do sujeito longitudinal, incluiu-se também a informação de ordem da coleta, como em: *SL1\_#1\_carta\_1*.

Posteriormente à organização dos áudios, estes foram convertidos do modo *stereo* para *mono*, pois o gravador digital faz a gravação por meio de dois canais e, dessa forma, aparece na tela do *Praat* dois oscilogramas, cada um representando um canal. Com o intuito de evitar confusão na hora da anotação do dado acústico, que

ora pode ocorrer no primeiro canal, ora no segundo, foi feita a conversão para unificar os dois canais de áudio, padronizando, assim, as marcações.

Logo que a conversão foi feita, deu-se início a tarefa de segmentação e anotação no software *Praat*. Primeiro, foi feita uma única *TextGrid* para o áudio de cada coleta, com uma linha de anotação, a da palavra. Em seguida, um *script*<sup>15</sup> (Anexo C) foi rodado para extrair e salvar as palavras segmentadas em formato *wav*. Esses áudios foram salvos com a identificação mencionada anteriormente. Para segmentar cada dado acústico, foram marcadas, na forma de onda, as fronteiras entre os segmentos a partir do cruzamento de zero do ciclo glotal. Desse modo, em cada *tier* foi segmentado: (i) a palavra, (ii) a sílaba; (iii) rótico e (iv) a vogal núcleo da sílaba. Quando o rótico foi produzido com um padrão *tap*, delimitou-se também as fronteiras: (iv) do elemento vocálico adjacente ao *tap* e (v) do *tap* incluindo o elemento vocálico, conforme a Figura 16. Com o controle do segmento vocálico adjacente ao *tap*, pode-se chegar à forma como este se apresenta na sílaba. Os valores de duração do *tap* são capazes de indicar se esse segmento é um *tap* ou uma vibrante, tendo em vista que estes têm valores de duração visivelmente diferentes, como já foi descrito nas seções anteriores.

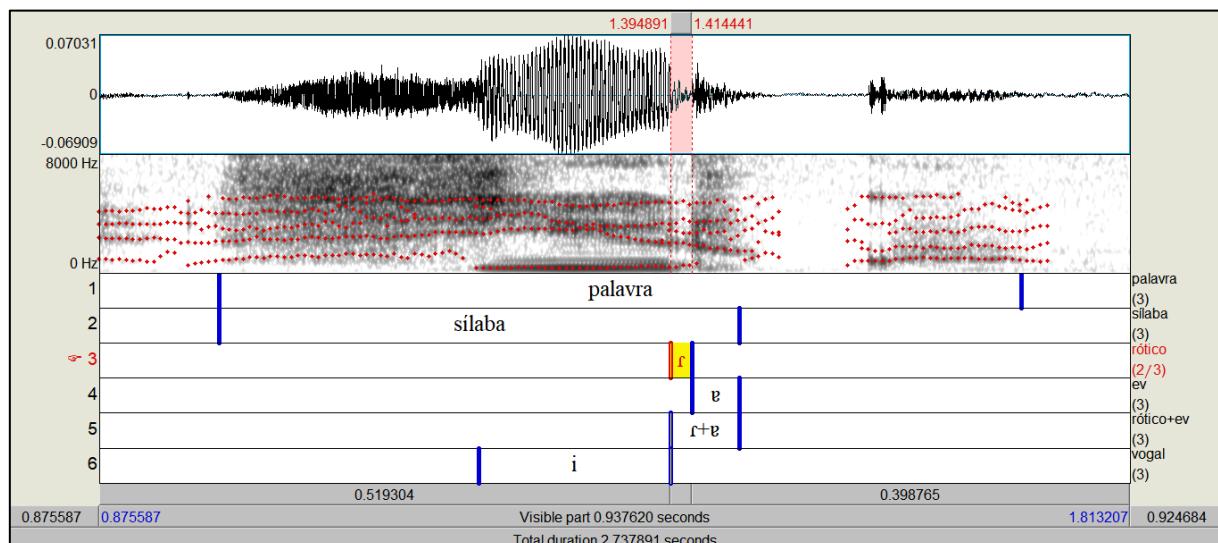


Figura 16 - Tela do software *Praat* em S0\_circo. Fonte: a autora

<sup>15</sup> *Extract\_and\_save\_intervals*. Wendy Elvira-García, (2015). Fonte: <http://www.wendyelvira.ga/scripts.html>

### 3.5.2 Critérios de análise acústica

Com base em Rodrigues (2007), foram adotados os seguintes critérios acústicos para análise: (i) inspeção visual; (ii) aspectos duracionais e (iii) trajetória formântica.

O primeiro critério justifica-se pela possibilidade de reconhecer as características acústicas dos segmentos por meio da inspeção visual do sinal acústico. Estas são capazes de identificar e distinguir vogais e consoantes.

Nesse sentido, Silva *et al.* (2019) elencam três principais características para o *tap*: (i) ausência de energia; (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico.

A primeira, apresenta como correlato visual um breve espaçamento em branco no espectrograma, este é causado pela queda de intensidade em relação aos elementos adjacentes ao *tap*. Essa região mais clara representa o breve momento de fechamento do trato vocal para produção do segmento, sendo assim, característica fundamental para a identificação do *tap* no espectrograma.

A segunda característica é identificada por meio de um sinal aperiódico, na forma de onda; já no espectrograma, são reveladas estrias verticais e estreitas. Essa não é uma qualidade constante do segmento e são observadas produções nas quais não ocorrem esse correlato acústico.

Por fim, a terceira característica elencada pelos autores é a presença de um elemento vocálico adjacente ao *tap*. Esse é identificado em encontros consonantais e em final de sílaba. É possível observá-lo, na forma de onda, após a ocorrência do *tap*, e sua natureza formântica será discutida neste trabalho. Assim como o ruído transiente, o elemento vocálico não é identificado em todas as produções quando o segmento se encontra em final de sílaba.

Em suma, estes três elementos apresentam-se como pistas visuais que são capazes de identificar a produção de um *tap*. Sendo assim, foi realizada uma inspeção visual no espectrograma e na forma de onda de todas as produções, para que fosse possível observar como essas características ocorriam na fala das crianças. Tais observações foram anotadas em uma planilha do software Excel 2016, versão 19.09, conforme a Figura 17.

				FINAL U				ESPECTROGRAMA				TAP?
	OITIVA	ACÚSTICA	CONTRASTE ENCOBERTO	Ausência de Energia	Ruído Transiente	Soltura da Oclusão	Elemento Vocálico	Ausência de Energia	Ruído Transiente	Soltura da Oclusão	Elemento Vocálico	
CUR1	[ku.lu]	[ku.lu]	SIM	SIM	NÃO	NÃO	SIM	SIM	NÃO	NÃO	SIM	NÃO
CUR2	[kul]	[kul]	SIM	SIM	NÃO	NÃO	SIM	SIM	NÃO	NÃO	SIM	NÃO
CUR3	[kul]	[kul]	SIM	SIM	NÃO	NÃO	SIM	SIM	NÃO	NÃO	SIM	NÃO
ABAJUR1	[aba.'3u.li]	[aba.'3u.li]	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO
ABAJUR2	[aba.'3uj]	[aba.'3uj]	TALVEZ	SIM	NÃO	NÃO	NÃO	SIM	NÃO	NÃO	NÃO	NÃO
ABAJUR3	[aba.'3u.li]	[aba.'3u.li]	TALVEZ	SIM	NÃO	NÃO	NÃO	SIM	NÃO	NÃO	NÃO	NÃO
ARTUR1	[a.'tu.li]	[a.'tu.li]	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO
ARTUR2	[a.'tuj]	[a.'tuj]	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO
ARTUR3	[a.'tuj]	[a.'tuj]	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO

Figura 17 - Tela do software Excel, contendo as observações da inspeção visual das produções em contexto de [u] final produzidos na segunda coleta da informante longitudinal.

Partindo para o segundo critério de análise, esse tem como objetivo principal estabelecer possíveis diferenças e semelhanças entre os pares de palavras presentes no *corpus* do trabalho. Assim, foram analisadas as medidas de duração relativa em relação à sílaba. A preferência por valores relativos deu-se para minimizar os efeitos da variação da taxa de elocução entre os falantes. A escolha da medida em relação à sílaba foi feita, visto que alguns vocábulos diferem no número de sílabas, tornando impossível o uso da medida de duração relativa em relação à palavra. Para tanto, os valores de duração absoluta, em milissegundos, foram retirados por meio de um *script*<sup>16</sup> (Anexo D) e dispostos em uma planilha no software Excel. O programa realizou os cálculos de duração relativa, baseado na fórmula abaixo:

$$duração\ relativa\ (%) = \frac{duração\ absoluta\ do\ alvo\ (ms)}{duração\ absoluta\ da\ sílaba\ (ms)} * 100$$

As medidas organizadas na planilha foram: (i) duração absoluta da sílaba; (ii) duração absoluta do *tap*, vogal ou ditongo; (iii) duração absoluta da vogal núcleo da sílaba; (iv) duração absoluta do elemento vocálico, (v) duração relativa do *tap* em relação à sílaba e (vi) duração relativa da vogal nuclear em relação à sílaba. Para os casos que o alvo era o vocábulo com o *tap*, mais dois elementos foram inseridos: (vi) soma da duração absoluta do *tap* mais o elemento vocálico; (vii) soma da duração absoluta da vogal núcleo mais o elemento vocálico e (viii) soma da duração absoluta do elemento vocálico, *tap* e a vogal núcleo. Além disso, foram calculados os valores médios de duração, dentro de uma mesma palavra, de todas as produções de uma

<sup>16</sup> *calculate\_segment\_durations*. Mietta Lennes, (2002). Fonte: <https://lennes.github.io/spect/>

mesma coleta, para cada informante. O *software* calculou os valores de média utilizando a seguinte fórmula:

$$\text{média (ms)} = \frac{\text{soma dos valores de duração absoluta (ms)}}{\text{total de produções (n)}}$$

Após, foi feita uma média geral de produção de cada alvo para todos os informantes, essa foi calculada a partir da média de cada sujeito. O *software* utilizado disponibiliza funções para que esses cálculos sejam realizados.

O último critério a ser analisado foi a trajetória formântica. Cabe ressaltar que, assim como no trabalho de Rodrigues (2007), esse parâmetro foi utilizado para os pares de palavras:

- (1) carta x cata x caixa
- (2) circo x Chico
- (3) curta x cuca x cuida
- (4) mar x má x pai
- (5) vir x vi
- (6) cur x tatu

Com o resultado da análise de oitiva, foram analisadas as palavras que continham o *tap* como alvo e que, segundo os juízes, foram produzidos sem o *tap* ou com um ditongo em seu lugar. Para realizar essa análise, os formantes foram obtidos por meio do *software Praat*. A partir da segmentação feita na *TextGrid*, levando em consideração cada tipo de vocábulo, foi solicitado, por meio do comando *formant listing*, a lista de formantes que surgem na segmentação da vogal selecionada. O programa gerou uma lista em arquivo .txt, a qual foi salva e, posteriormente, aberta no *software Bloco de Notas*, versão 20H2, para a substituição de pontos (.) por vírgulas (,). Após, foi novamente salva. Este arquivo foi aberto no *software Excel* por meio do assistente de importação de texto. Na planilha, foram mantidos os valores de tempo e dos três primeiros formantes (F1, F2 e F3). Baseados nesses valores, foram gerados gráficos de linhas com marcadores, pois esses apresentam dados contínuos ao longo do tempo. Os valores dos formantes foram dispostos no eixo vertical (em *Hertz*), enquanto os valores de tempo foram ordenados no eixo horizontal (em segundos), como é possível visualizar na Figura 18:

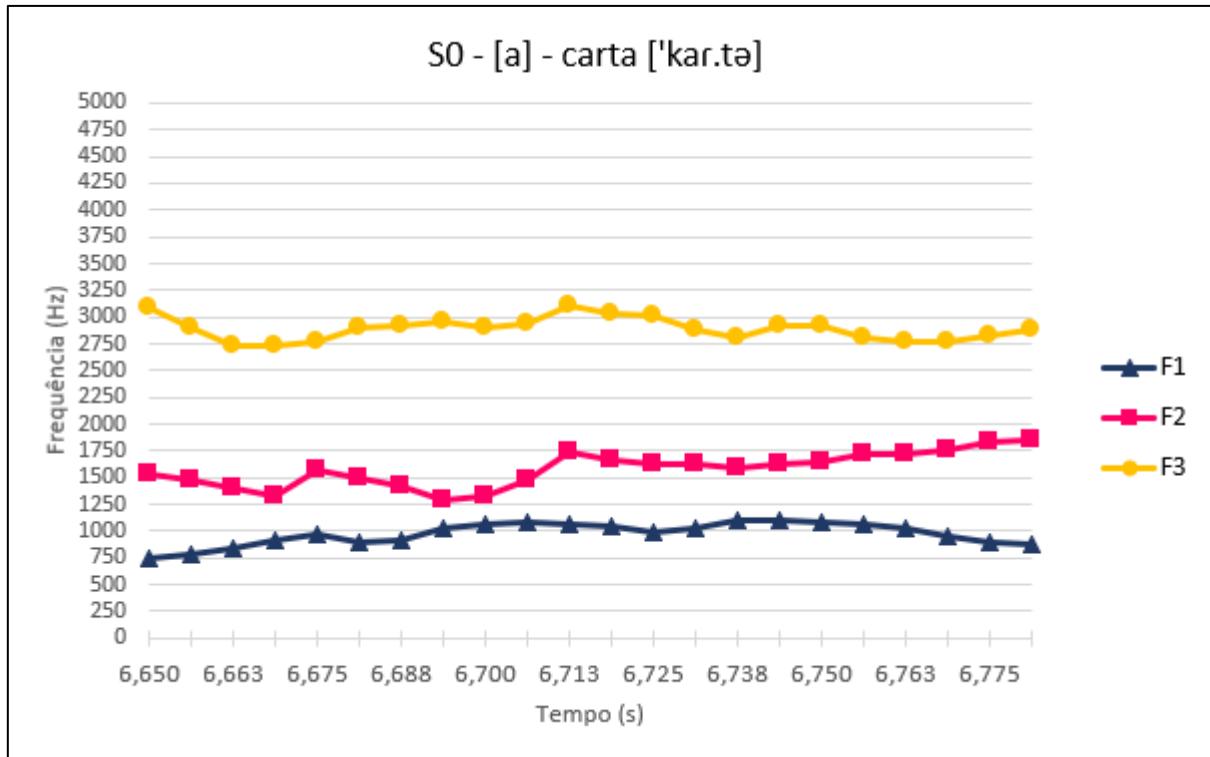


Figura 18 - Exemplo dos gráficos de Trajetória Formântica de F1, F2 e F3 na vogal [a] que antecede o rótico na produção da palavra *carta*, por S0. Fonte: a autora

Esse critério foi relevante para comparar a trajetória formântica das palavras que tinham como alvo somente a vogal ou o ditongo, por exemplo, *cata* e *caixa*, e aquelas que tinham o *tap* como alvo, mas que sofreram alterações na hora de sua produção.

Em síntese, esses três critérios de análise são meios possíveis para descrever e caracterizar as produções dos informantes. Também é necessário salientar que se esperava, por meio desses parâmetros, investigar possíveis contrastes encobertos que, auditivamente, não fossem percebidos, embora presentes no detalhe fonético.

Para investigar a natureza do elemento vocálico, foram realizados dois tipos de análises: (i) comparar as características formânticas do elemento vocálico com as da vogal nuclear da sílaba CVC, com base em Silva, Clemente e Nishida (2006) e (ii) comparar as características do elemento vocálico com as características formânticas da porção final de uma vogal CV, utilizando-se a mesma medida duracional de EV em termos proporcionais.

Para ambas análises, foram examinados apenas pares de palavras produzidas pelo sujeito controle, considerando-se o mesmo ponto e modo de articulação da primeira consoante da sílaba CVC. Sendo assim, foram investigados os elementos

vocálicos presentes nos alvos: *carpa*, *carta*, *circo*, *círculo*, *curta* e *curto*. Apenas os dados em posição medial foram considerados, uma vez que S0 articulou o *tap* apenas nessa posição, pois, em posição final, foram observados a produção de variantes róticas do tipo vibrante, aproximante retroflexa, fricativa glotal e *tap* fricatizado<sup>17</sup>.

Sendo assim, para a primeira análise, foram extraídos de forma manual os valores de F1 e F2 do ponto médio da vogal núcleo (VN) da sílaba CVC e do elemento vocálico (EV). Para calcular as médias, os valores foram dispostos em uma planilha no software Excel. Esses valores foram plotados por meio de um *script*<sup>18</sup> (Anexo E).

A segunda análise se justifica pela hipótese de que, em sílaba CVC, a vogal nuclear seja entrecortada pelo *tap* – como já constatado para sílaba CCV (SILVA, CLEMENTE e NISHIDA, 2006; BILHARVA-DA-SIVA, 2019; MELCHEQUE, FERREIRA-GONÇALVES e BRUM-DE-PAULA, 2021) –, apresentando, portanto, natureza formântica próxima de uma vogal CV. Nesse sentido, foram selecionadas palavras com as vogais [a], [i] e [u], presentes em sílabas CV, a saber: *cata*, *chico* e *cuca*. As referidas vogais foram comparadas, em termos de duração e valores de frequências, com cada elemento vocálico. Desse modo, com base no valor de duração de cada EV, foi calculado o quanto ele ocupava, em termos proporcionais, da vogal CV, a partir de um cálculo de duração relativa, considerando a seguinte fórmula:

$$\text{duração relativa (\%)} = \frac{\text{duração absoluta do elemento vocálico (ms)}}{\text{duração absoluta da vogal nuclear CV (ms)}} * 100$$

Após saber qual era a proporção de EV em relação à vogal CV, foram extraídos os valores de F1 e F2 da vogal CV. Para descobrir em qual ponto da vogal seria feita a extração desses valores, utilizou-se uma regra de três simples, baseada na fórmula abaixo:

$$\frac{100 (\%)}{\text{proporção de EV (\%)}} = \frac{\text{duração de CV (ms)}}{x}$$

Assim, por exemplo, ao aplicar a fórmula acima para o elemento vocálico em *carpa* 1, constatou-se uma duração de EV de 23 ms. Considerando-se que a vogal

<sup>17</sup> Para saber mais sobre a produção de *tap* fricatizado, ver Bilharva-da-Silva (2019), capítulo 6.

<sup>18</sup> *plot\_vowels*. Ricardo Bion, (2006)

em CV, na palavra *cata*, apresenta duração de 115 ms, o elemento vocálico ocupa 20% da vogal nuclear. Logo, foram extraídos os valores formânticos do ponto médio de EV e dos 23 ms finais da vogal em CV. Esse cálculo foi feito pelo *software* Excel para cada EV analisado.

As frequências extraídas do ponto médio da proporção de CV e do ponto médio do elemento vocálico foram dispostas em uma planilha no *software* Excel. Esses valores foram plotados por meio do último *script* apresentado (Anexo E).

## 4 Resultados e Discussão

Este capítulo apresenta os resultados das análises de oitiva e acústica. A Seção 4.1 exibe a descrição dos dados de oitiva dos sujeitos e a Seção 4.2 discute os resultados da inspeção visual dos espectrogramas, dos aspectos duracionais dos segmentos, das trajetórias formânticas e da natureza do elemento vocálico.

### 4.1 Análise de oitiva

A análise presente nessa seção tem por base a descrição de oitiva das produções dos informantes. Primeiramente, será apresentado o comportamento geral dos informantes, por meio de percentuais de produções do *tap* e de suas variáveis. Logo, serão exibidos os principais padrões individuais de cada informante. Por fim, será estabelecida uma análise comparativa entre os resultados do sujeito longitudinal e os resultados dos sujeitos transversais. Essa análise foi aplicada para as palavras que abarcam o *tap* como alvo, essas são: *carta*, *farpa*, *carpa*, *circo*, *círculo*, *curto*, *turco* e *curta* – para o contexto medial – e *mar*, *lugar*, *azar*, *vir*, *Zair*, *Nadir*, *cur*, *abajure* e *Artur* – para o contexto final. Logo, os vocábulos do *corpus* que não possuem o *tap* e os alvos que são formados com o ditongo palatal não fizeram parte dessa avaliação.

#### 4.1.1 Sujeito longitudinal

A informante SL1 conta com três coletas de dados, as quais ocorreram com intervalos de 20 dias aproximadamente entre as gravações. Quando a criança não conseguia produzir o segmento alvo, utilizava oito diferentes estratégias<sup>19</sup> para a produção do alvo. Esses foram classificados como:

- (i) não realização: ['ka.pə] para ['kar.pə];
- (ii) semivocalização com o glide [j]: ['zaj] para [za.'ir];
- (iii) semivocalização com o glide [w] ['paw.pə] para ['far.pə];
- (iv) substituição por [l]: [lu.'gal] para [lu.'gar];
- (v) substituição por [l] mais acréscimo da vogal [i]: [aba.'ʒu.li] para [a.ba.'ʒur];

---

<sup>19</sup> Considerando que a atual sessão descreverá os dados por meio de análise de oitiva, ainda que a presente pesquisa tenha como base uma abordagem dinâmica, optou-se por manter o uso de alguns termos tradicionais empregados em trabalhos que apoiavam-se em teorias gerativas, como Yavas, Hernadorena e Lamprecht (1991).

- (vi) substituição por [l] mais acréscimo da vogal [u]: ['ku.lu] para ['kur];
- (vii) metátese CV mais acréscimo da vogal [u]: ['tu.ku.ru] para ['tur.ku];
- (viii) substituição por [l] mais metátese CV: ['si.ku.lo] para ['sir.ku].

A Tabela 1 apresenta em detalhes os índices e médias das produções do *tap* e suas variações, de acordo com a ordem de coleta e a posição que o segmento se encontrava na sílaba.

Sujeito	CVC	r	Ø	j	w	-	+ i (cv)	+ u (cv)	metátese (cv) + u	+ metátese (cv)
SL1 #1	medial	-	14/20 70%	1/20 5%	1/20 5%	-	-	-	2/20 10%	2/20 10%
	final	-	7/25 28%	4/25 16%	-	8/25 32%	6/25 24%	-	-	-
SL1 #2	medial	1/20 5%	17/20 85%	-	-	-	-	-	-	2/20 10%
	final	1/24 4%	8/24 33%	8/24 33%	-	2/24 9%	4/24 17%	1/24 4%	-	-
SL1 #3	medial	-	22/24 92%	-	-	-	-	-	1/24 4%	1/24 4%
	final	-	11/24 46%	4/24 17%	-	5/24 20%	4/24 17%	-	-	-
MÉDIA	medial	1/64 1%	53/64 83%	1/64 1%	1/64 1%	-	-	-	3/64 5%	5/24 9%
	final	1/73 1%	26/73 36%	16/73 22%	-	15/73 21%	14/73 19%	1/73 1%	-	-
DP	medial	-	0,11	-	-	-	-	-	0,05	0,03
	final	-	0,09	0,10	-	0,12	0,04	-	-	-

Tabela 1 - Índices de produções do *tap* e recursos utilizados por SL1, de acordo com a ordem de coleta e a posição do *tap* na sílaba, médias e desvio padrão

No que se refere a posição medial, é possível observar que a criança utiliza o recurso de omissão do *tap* em 70% dos casos na primeira coleta. Esse percentual aumenta quando a segunda gravação é analisada, pois, em 85% dos casos, a criança priorizou a omissão do segmento. Do mesmo modo, a terceira coleta apresenta um alto índice de emprego desse recurso, 92%. É necessário destacar que, na presente amostra, a vogal baixa foi o ambiente que mais favoreceu a não produção do segmento. Os alvos *carpa* e *carta* não apresentaram o rótico em nenhuma coleta. O Gráfico 7 permite identificar a tendência à omissão do *tap* ao longo das coletas em posição medial.

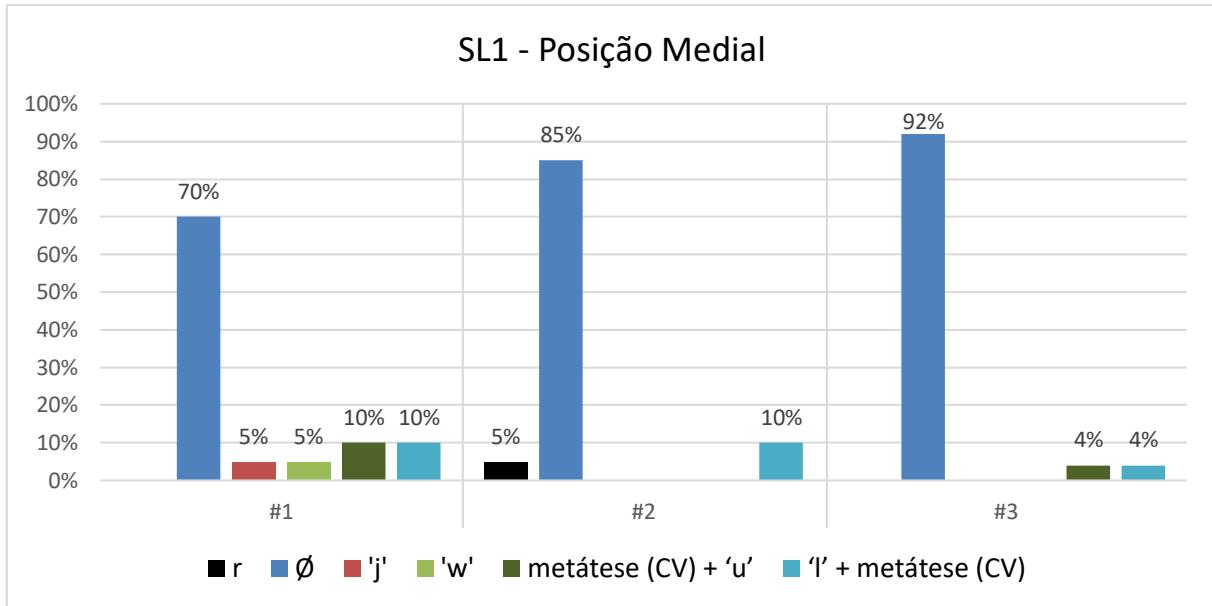


Gráfico 7 - Índices de produções do *tap* e recursos utilizados por SL1, em posição medial, de acordo com a ordem de coleta

Os casos de semivocalização ocorreram apenas na primeira coleta, representando 5% das produções articuladas com o glide [j] e 5% com o glide [w]. O recurso de semivocalização com o glide [j] foi empregado para o alvo *curta*, que foi produzido como ['kuj.tə]. Já o único caso de semivocalização com o glide posterior [w] ocorreu para o alvo *farpa*, que foi produzido como ['paw.pə]. O recurso de substituição por [l] mais metátese CV foi adotado para o alvo *circo*, que foi articulado como ['si.ku.lo]. Embora o *tap* não tenha sido articulado no final da sílaba, sua articulação ocorreu por meio de metátese mais o acréscimo da vogal [u] para duas palavras: *circo* e *turco*. Essas foram articuladas como: ['si.ku.ro] e ['tu.ku.ru]. Em posição medial, portanto, as estratégias de omissão e de metátese são as mais recorrentes. Interessante destacar que a ocorrência da metátese sinaliza para a preservação do segmento, já realizado em posição CV, o que, a princípio, não corrobora os resultados de Miranda (1996) e Hernandorena e Lamprecht (1997), que constatam a emergência do *tap*, primeiramente, em coda, não em *onset* simples.

Para a posição final de palavra, conforme Gráfico 8, o recurso de omissão do segmento ocorre em menor proporção, com valores de 28%, 33% e 46%. Ainda assim, seu baixo índice de desvio padrão indica que esse é o único recurso que apresenta uma utilização regular quando comparados seus percentuais aos valores das demais estratégias de reparo.

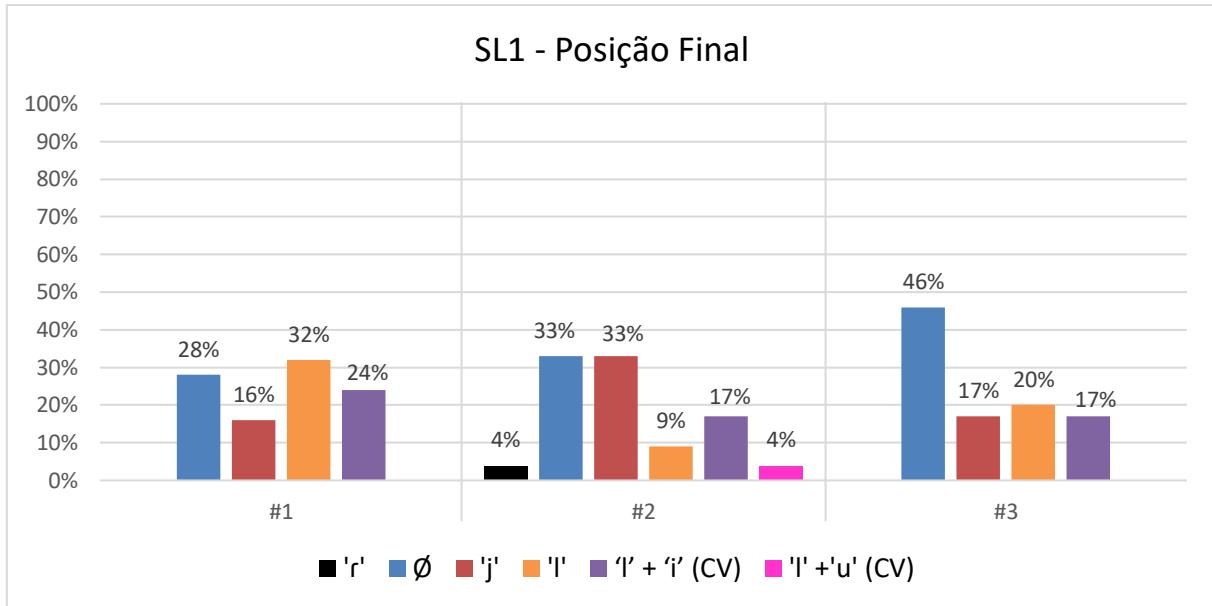


Gráfico 8 - Índices de produções do *tap* e recursos utilizados por SL1, em posição final, de acordo com a ordem de coleta

Na primeira coleta, a omissão chega a atingir percentual menor do que a substituição por [l], sendo o recurso de aplicação predominante apenas na terceira coleta. A produção do segmento ocorreu uma única vez, na segunda coleta. O Gráfico 8 possibilita a inferência de que há um maior avanço nas tentativas de realização do segmento, pois as substituições por [l], [l] mais acréscimo da vogal [i], [l] mais acréscimo da vogal [u] e a semivocalização [j] apresentam percentuais que distribuem-se entre 4% e 33%.

Percebe-se, portanto, um comportamento distinto na emergência do *tap*, tendo por base as posições medial e final, parecendo, a posição final ser mais propícia à aquisição do segmento, o que corrobora os resultados de Miranda (1996) e Mezzomo (2004).

Traçado o percurso longitudinal das coletas, o Gráfico 9 apresenta as médias das produções ocorridas nas três coletas. À esquerda, é possível visualizar os resultados para a posição medial e, à direita, os resultados para a posição final.

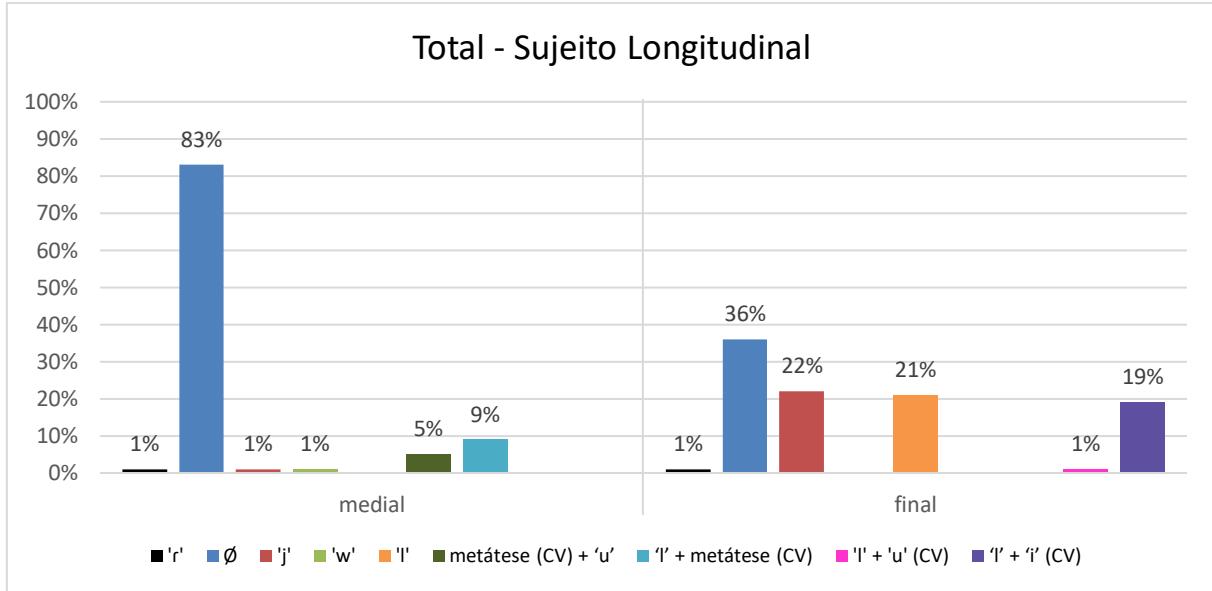


Gráfico 9 - Médias das produções do *tap* e recursos utilizados nas três coletas por SL1 para posição medial (à esquerda) e para posição final (à direita)

Os resultados, para a posição medial, apontam que a informante realiza a omissão do segmento em 83% dos casos, sendo constatadas, ainda que em grau bem menor, metáteses e substituições. O alto índice de não ocorrências do *tap* em posição medial de sílaba também é detectado em Miranda (1996) e Mezzomo (2004).

Já em posição final de sílaba, os percentuais são mais bem distribuídos, entre os casos de semivocalização e substituição por lateral, compondo ou não sílaba CV.

#### 4.1.2 Sujeitos transversais

Os resultados dos informantes transversais correspondem a uma coleta de dados de cada um. No momento em que as crianças não alcançavam a produção do segmento alvo, empregaram, assim como o sujeito longitudinal, oito recursos diferentes durante em suas produções. Esses foram classificados como:

- (i) não realização: [a.'za] para [a.'zar];
- (ii) semivocalização com o glide [j]: ['kuj.tu] para ['kur.tu];
- (iii) semivocalização com o glide [w]: ['viw] para ['vir];
- (iv) substituição por [l]: ['mal] para ['mar];
- (v) metátese CCV: ['fa.prə] para ['far.pə]
- (vi) metátese CV mais acréscimo da vogal [u]: ['tu.ku.ru] para ['tur.ku];
- (vii) substituição por [l] mais metátese CV: ['si.ku.lo] para ['sir.ku];
- (viii) metátese CV: ['si.ku.ro] para ['sir.ku.lo]

A Tabela 2 apresenta em detalhes os índices e médias das produções do *tap* e suas variações, de acordo com o informante e a posição em que o segmento se encontrava na sílaba.

Sujeito	CVC	r	ø	j	w	-	metátese (ccv)	metátese (cv) + u	I + metátese (cv)	metátese (cv)
ST1	medial	—	3/23 13%	15/23 65%	—	—	—	2/23 9%	—	3/23 13%
	final	—	18/25 72%	6/25 24%	1/25 4%	—	—	—	—	—
ST2	medial	8/24 33%	10/74 42%	—	—	—	4/24 17%	—	—	2/24 8%
	final	17/19 89%	2/19 11%	—	—	—	—	—	—	—
ST3	medial	17/23 74%	3/23 13%	—	—	—	—	1/23 4%	2/23 9%	—
	final	22/23 87%	2/23 9%	—	—	1/23 4%	—	—	—	—
ST4	medial	—	7/22 33%	13/22 59%	—	—	—	1/22 4%	—	1/22 4%
	final	2/24 8%	16/24 67%	1/24 4%	—	5/24 21%	—	—	—	—
MÉDIA	medial	25/92 27%	23/92 25%	28/92 31%	—	—	4/92 4%	4/92 4%	2/92 2%	6/92 7%
	final	39/91 43%	38/91 42%	7/91 8%	1/91 1%	6/91 6%	—	—	—	—
DP	medial	0,29	0,15	0,04	—	—	—	0,03	—	0,05
	final	0,46	0,34	0,14	—	0,14	—	—	—	—

Tabela 2 - Índices de produções do *tap* e recursos utilizados por ST1, ST2, ST3 e ST4, de acordo com a posição da sílaba, médias e desvio padrão

Com base nos resultados reportados na Tabela 2, observa-se que, à exceção de ST1, todos os sujeitos apresentaram produção variável do *tap*, sendo que ST2 e ST3 atingiram índices superiores a 80% de produção correta em posição medial. Para a produção do *tap*, verificou-se uma maior variabilidade nas produções entre os sujeitos, fato esse sinalizado por maiores valores de desvio padrão para a ocorrência do segmento. Em linhas gerais, o recurso mais utilizado, para ambas posições silábicas, foi a omissão. Já a semivocalização com [j], apresentou-se como um recurso importante para a posição medial. Todavia, as características idiossincráticas das crianças precisam ser levadas em consideração. Sendo assim, serão expostos os resultados apresentados individualmente pelos informantes.

O Gráfico 10 revela os índices das produções da informante ST1, uma menina de 3 anos e 02 meses.

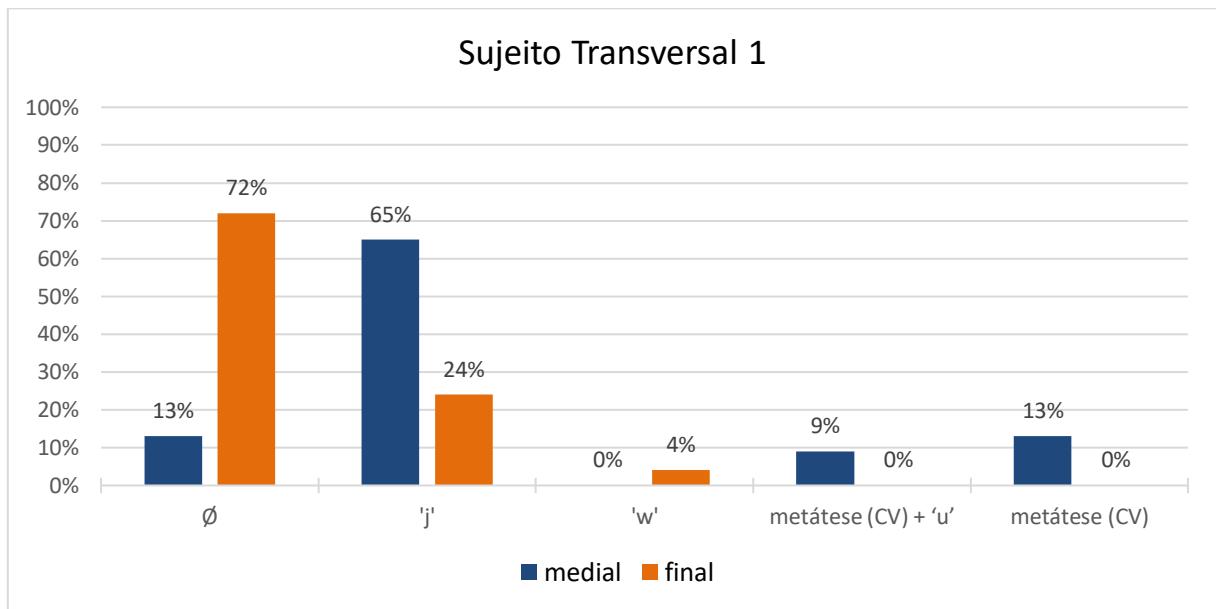


Gráfico 10 - Índices de produção *do tap* e recursos utilizados por ST1 para posição medial (azul) e para posição final (laranja)

A informante utiliza cinco recursos quando não consegue alcançar a produção alvo, três em posição final e quatro em posição medial. De acordo com o Gráfico 10, é possível visualizar que, para a posição final, o recurso mais empregado, a omissão, apresenta um índice de 72%, seguido pela semivocalização [j], 24%. Quando o *tap* deveria figurar em posição medial, o principal recurso utilizado foi a semivocalização, que apresentou um índice de 65% de ocorrência. As vogais [a] e [u] formam o ambiente que favoreceu a semivocalização. Os outros três recursos utilizados nessa posição apresentaram índices entre 9% e 13%.

Verificou-se que nenhuma ocorrência do *tap* foi percebida auditivamente. A produção de metátese (CV) foi observada na palavra *círculo*, produzida como ['si.ku.ro], dado que o alvo [r] foi deslocado para a direita, alterando a consoante da última sílaba. O fenômeno intitulado metátese (CV) mais acréscimo da vogal [u] ocorreu nas palavras *circo* e *turco*, que foram produzidas como ['si.ku.ru] e ['tu.ku.ru]. As palavras que eram dissílabas tornaram-se trissílabas devido ao deslocamento para a direita, criando, assim, uma nova sílaba CV composta pelo [r] acompanhada da vogal [u]. Destaca-se, portanto, que, embora não produza o *tap* em coda silábica, já o realiza em sílaba CV.

O Gráfico 11 exibe os índices das produções do informante ST2, um menino de 3 anos e 04 meses.

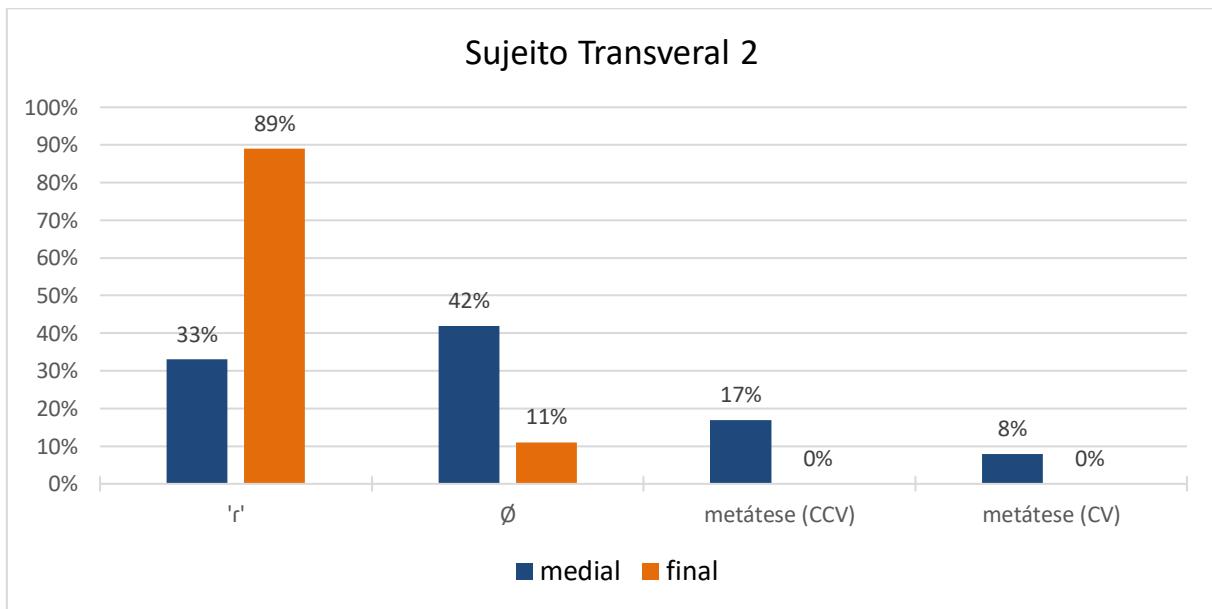


Gráfico 11 - Índices de produção do *tap* e recursos utilizados por ST2 para posição medial (azul) e para posição final (laranja)

Os índices do informante apontam uma maior facilidade de produção do segmento alvo em final de sílaba, posição essa que apresenta 89% de ocorrência e 11% de omissões. No entanto, a posição medial demonstra ter um maior uso de metáteses se comparada à posição final. É preciso atentar para o fato que as vogais [a] e [i] foram o ambiente que propiciaram a omissão do *tap*. Para a palavra *carta*, apenas uma produção de [r] foi percebida. Já a palavra *círculo* foi percebida como ['si.ku.ru] em duas das suas três repetições, ou seja, na sílaba alvo o segmento foi omitido. A palavra *farpa* foi percebida auditivamente como ['fa.pra]; já para a palavra *turco*, o informante utilizou, em todas as repetições, o recurso de metátese CCV e a palavra alvo foi produzida como ['tru.ku]. Desse modo, uma clara tentativa de produção do alvo é evidenciada pela utilização desse recurso.

O Gráfico 12 apresenta os índices das produções do informante ST3, um menino de 3 anos e 11 meses.

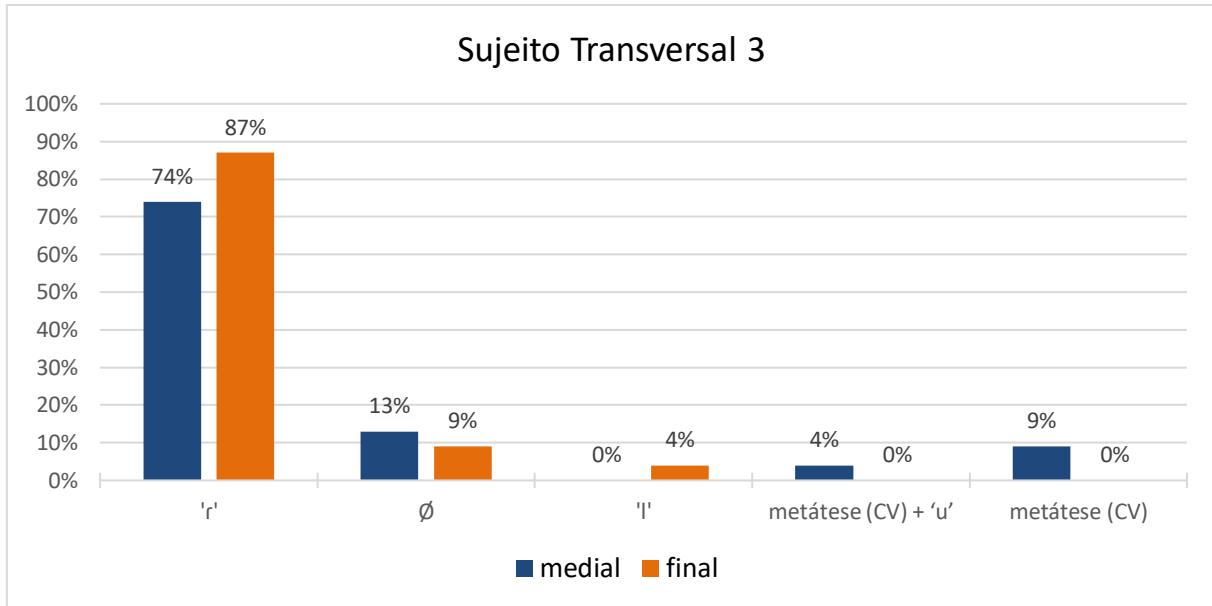


Gráfico 12 - Índices das produções do *tap* e recursos utilizados por ST3 para posição medial (azul) e para posição final (laranja)

Para a posição medial, ST3 apresenta um índice de 74% de ocorrência do segmento e utiliza quatro recursos quando não há a produção da consoante: omissão, substituição por [l], metátese (CV) mais acréscimo da vogal [u] e metátese (CV). Por outro lado, em final de sílaba, foi detectado auditivamente 87% de produção do *tap*, sendo 20 ocorrências para 23 possibilidades. Importante também evidenciar algumas especificidades dos dados. A primeira é a segunda repetição de *turco*, em que a criança produziu como ['tur.ku.ru]; a segunda é a produção de [ar.ba.ʒur] para a primeira repetição de *abajur*. Observa-se que, nos dois casos, o menino reforça a produção do *tap*, evidenciando um equívoco na coordenação dos gestos. É provável que um erro de intrusão gestual foi cometido devido a adição de um gesto de forma antecipada.

O Gráfico 13 expõe os índices das produções da informante ST4, uma menina de 4 anos e 05 meses.

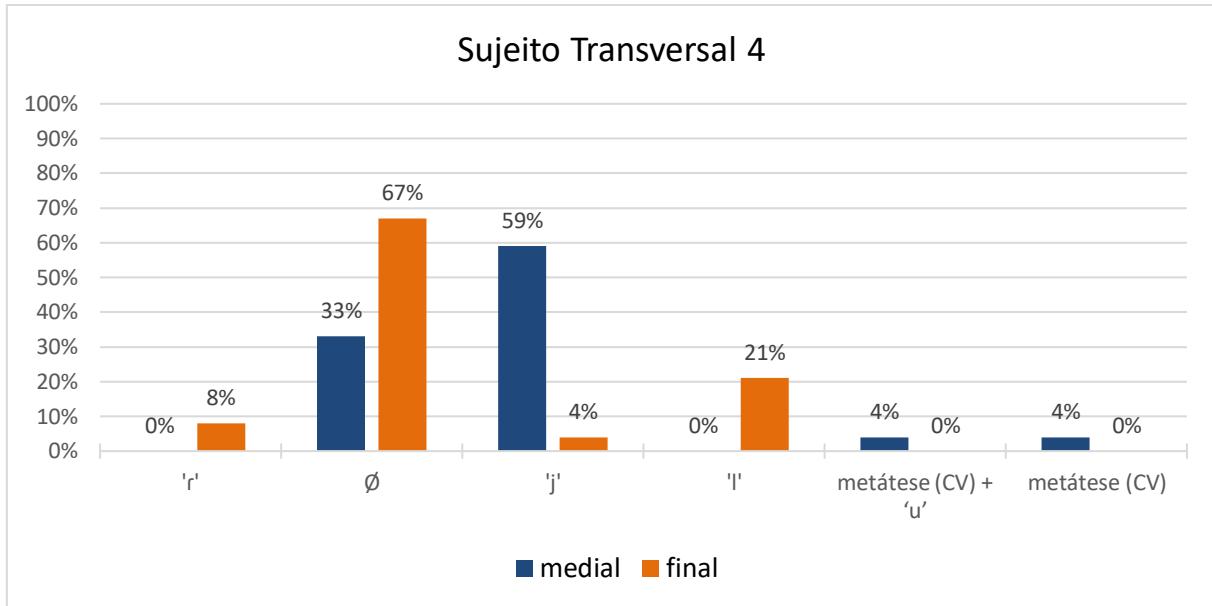


Gráfico 13 - Índices de produção *do tap* e recursos utilizados por ST4 para posição medial (azul) e para posição final (laranja)

Os resultados indicam que o *tap* não foi produzido em nenhuma das possibilidades, em posição medial, e quatro recursos foram utilizados para a consoante não produzida: omissão, semivocalização [j], metátese (CV) mais acréscimo da vogal [u] e, por fim, metátese (CV). Em posição final, o segmento emergiu em apenas 8% dos casos. Sendo assim, foi necessário o emprego de três recursos nas tentativas de realização das palavras com a consoante ainda não articulada. A semivocalização e a omissão demonstraram ser recursos importantes quando o *tap* figura no meio da palavra, enquanto que a omissão e a substituição pela lateral apresentaram preferência no momento em que o segmento não era produzido em final da palavra. Importante reportar que alguns alvos foram produzidos de forma inesperada, como os casos de *Artur*, realizado como [la.'tu] e *Zair*, articulado como [sa.'iej] e [za'je']. Fica claro que a criança produz, de forma recorrente, um movimento de metátese com substituições ou semivocalizações, pois ela ainda não possui estabilizada a coordenação gestual necessária para a realização do som de acordo com a forma alvo.

De modo geral, ao analisar os gráficos dos informantes individualmente, foi possível constatar que quanto maior o índice de produção do *tap*, menor foi o número de recursos utilizados pela criança. Esse fato evidencia que o segmento está próximo de sua aquisição, visto que recursos para a substituição do som já não se fazem necessários.

Após a análise dos resultados individuais dos informantes, serão expostos, no Gráfico 14, os resultados gerais para os sujeitos transversais. Dessa maneira, será possível observar o padrão geral de aquisição do *tap* pelos informantes.

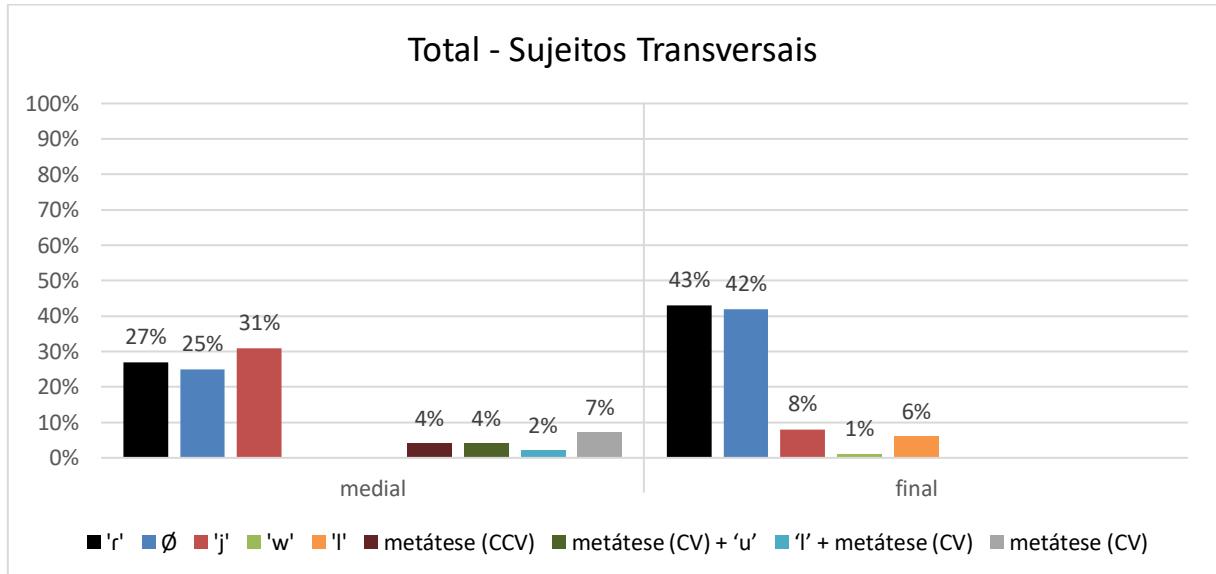


Gráfico 14 - Médias das produções do *tap* e recursos utilizados pelos informantes transversais para posição medial (à esquerda) e para posição final (à direita)

Com base nos percentuais reportados no Gráfico 14, verifica-se que o *tap* é produzido, em posição medial, em 27% dos casos das possibilidades de sua ocorrência. Quando a articulação do segmento não é concretizada, os recursos mais empregados foram o apagamento e a semivocalização com a vogal alta anterior, com índices de 25% e 31%, respectivamente. Nessa posição silábica, os casos de metáteses somam 17%. Para o contexto silábico final, o [r] ocorre em 43% dos dados e o recurso de apagamento é o principal recurso que opera nessa posição. Portanto, nota-se que os informantes transversais adotaram modos diferentes para manejar a ausência do segmento, tendo por base as posições diferentes desse dentro da palavra.

#### 4.1.3 Comparação entre dados longitudinais e transversais

A realização de uma comparação entre os resultados do sujeito longitudinal e os resultados dos sujeitos transversais possibilitará o levantamento de inferências sobre o processo de aquisição dos róticos e quais os recursos dispuseram de uma preferência maior por parte das crianças que participaram deste estudo.

Nesse sentido, o Gráfico 15 concentra uma comparação dos resultados entre os sujeitos, de acordo com o recurso utilizado, para a posição medial de palavra.

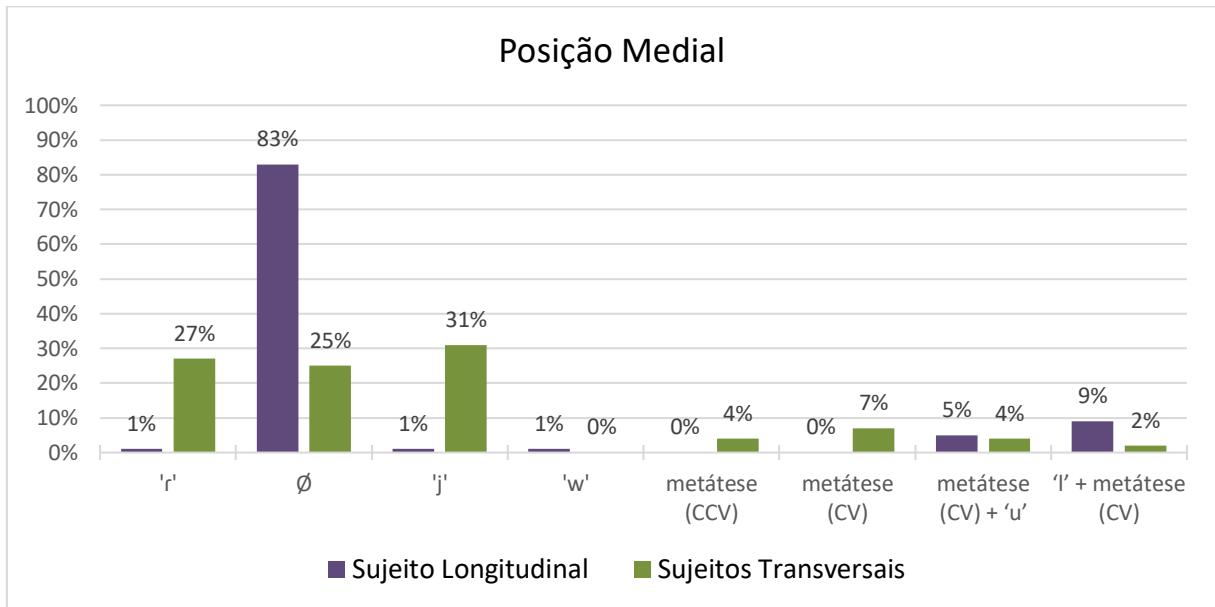


Gráfico 15 - Médias das produções do *tap* e recursos utilizados pelo informante longitudinal (roxo) e pelos informantes transversais (verde) para posição medial

Os índices apontam que a informante longitudinal não possui a constelação gestual necessária para a articulação do *tap*, visto que, ao longo das coletas, o segmento não foi praticamente produzido. Os sujeitos transversais alcançaram um índice de produção de 27%. As produções de metáteses desses sujeitos abrangem a articulação do segmento, no entanto, esse está fora de sua posição alvo. Quando o segmento não é articulado, a informante longitudinal opta por empregar o recurso da omissão em 83% dos casos. Por outro lado, os dois recursos mais empregados pelos sujeitos transversais foram a semivocalização [j], com um índice de 31%, seguido da omissão, que compreende 25% dos casos. Logo, é possível inferir que esses sujeitos, ainda que produzam o segmento em alguns casos, necessitam de um refinamento mais adequado dos gestos articulatórios do *tap*.

O Gráfico 16 apresenta os índices das produções do *tap* da informante longitudinal e dos informantes transversais, em posição final da palavra.

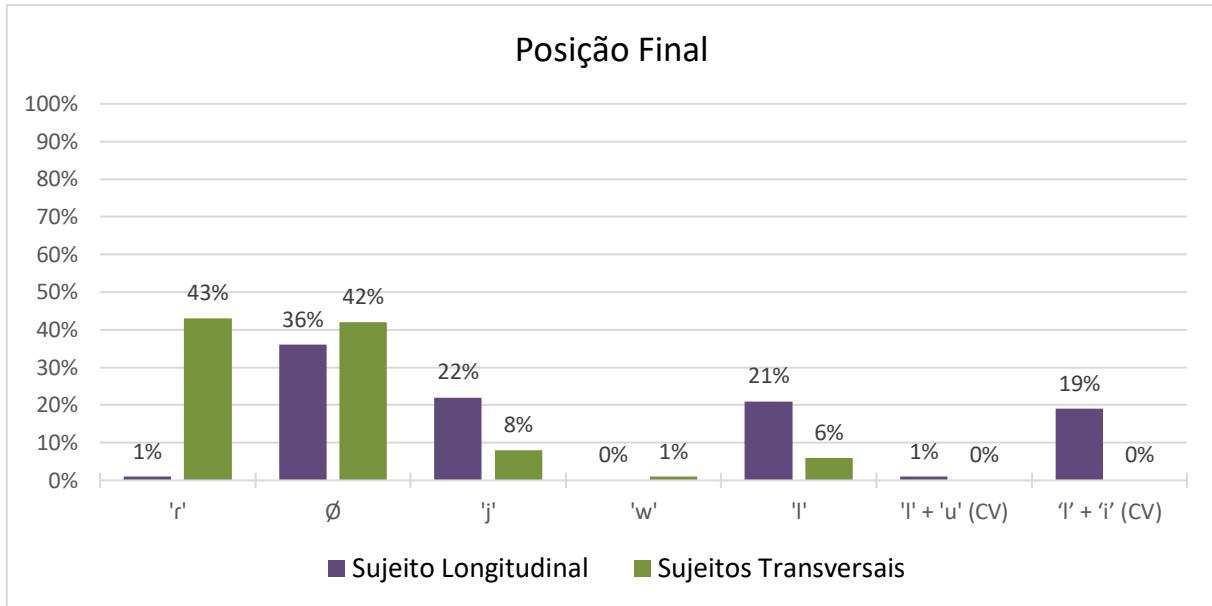


Gráfico 16 - Médias das produções do *tap* e recursos utilizados pelo informante longitudinal (roxo) e pelos informantes transversais (verde) para posição final

O segmento foi produzido, no final dos vocábulos, pelos sujeitos transversais em 43% das possíveis ocorrências, enquanto a informante longitudinal obteve um índice de produção de apenas 1%. Este corresponde a uma única produção presente na segunda coleta, a palavra *vir*. O recurso de omissão ocorreu de forma similar para ambos os grupos de sujeitos, sendo que os transversais apresentaram um índice levemente maior que a longitudinal, 42% e 36%, respectivamente.

Entre os grupos, a semivocalização [j] foi predominante nas produções da informante longitudinal, 22%. Na fala dos transversais, esse recurso foi pouco expressivo, 8%. Já a semivocalização por [w] ocorreu uma única vez na fala de ST1 – *vir* foi produzido como ['viw].

Outro ponto a ser destacado é a alta taxa de ocorrência de [l] na produção da informante longitudinal. Considerando-se os índices de substituição do *tap* por [l] e substituição por [l] mais acréscimo da vogal alta anterior, esse recurso apresenta uma taxa de 41% de utilização. Esse fato pode expressar que há um progresso nas tentativas de realização do segmento, tendo em vista que [r] e [l] são compostos por gestos de ponta de língua.

Ao observar os resultados dos Gráficos 15 e 16, é possível inferir que há uma diferença na emergência do segmento de acordo com a sua posição na palavra. A posição final, conforme previsto na literatura, parece favorecer a aquisição do *tap*. Os dois grupos de informantes apresentaram um índice maior de produção de [r] quando

esse encontrava-se em posição final. A produção de semivocalização [j] e, especialmente, de substituição por [l] apontam que há mais tentativas de articulação do som também nessa posição. Ainda, para a informante longitudinal, o índice de omissões é menor em final de palavra, 36%, quando comparado à posição medial, 83%. Esses achados vão ao encontro dos resultados presentes nas pesquisas de Hernandorena (1990), Mezzomo (2004) e Miranda (1996).

A substituição do *tap* por [l], principalmente na fala da informante longitudinal, pode estar relacionada com a emergência do segmento em final de palavra, antes do seu surgimento em posição medial. Esse fato evidencia a tentativa da criança de realizar o *tap* corretamente, no entanto, esse realiza-se de outra forma. Sendo, assim, a emergência da lateral é uma consequência da falta de regulação dos gestos articulatórios por parte do falante, visto que há similaridade nas características gestuais que unem esses dois sons. A produção de um *tap* seria “um encurtamento dos componentes coronais da lateral” (ALBANO, 2001, p. 136). Para a autora, a diferença entre os dois segmentos está na regulação da magnitude temporal do gesto coronal (ponta de língua).

A produção do fenômeno denominado *metátese* (CV), na presente pesquisa, ocorre em posição medial para ambos grupos de sujeitos, destacando-se que, para os longitudinais, a metátese (CV) se realiza com a inserção da vogal “u” na sílaba seguinte, como na produção de [tu.ku.ru], referente à palavra *turco*. As palavras alvo são modificadas em virtude do deslocamento do *tap* para a direita da palavra, gerando, assim, uma sílaba (CV) constituída pelo [r] acompanhada de uma vogal.

Em posição final, não há ocorrência de metátese, pois não há contexto disponível para a ocorrência desse fenômeno. Sendo assim, a criação de uma sílaba CV ocorre com a *substituição por [l]* mais *acrúscimo das vogais [i]* ou *[u]*, por exemplo, a produção de [aba.'ʒu.li] para a palavra *abajur* e ['ku.lu] para *cur*. A produção da lateral no dialeto pelotense, de fato, só ocorre em sílaba CV ou CCV, assim, para a produção do gesto de ponta para o alvo [r] – mesmo que ainda sem a magnitude adequada –, originando a emergência de um [l], o acrúscimo das vogais [i] ou [u] acaba sendo uma boa estratégia realizada pelo aprendiz.

Outra hipótese seria considerar que, sendo a produção do rótico, em posição CVC, acompanhada de um elemento vocálico na forma alvo, a vogal [u], produzida pelo aprendiz, poderia simplesmente decorrer da emergência do elemento vocálico, com maior duração temporal, tendo em vista a substituição do *tap* pela lateral alveolar

realizada pela criança. Das quinze produções de SL1 que incluem a aplicação do recurso *substituição por [l] mais acréscimo das vogais [i] ou [u]*, no entanto, apenas duas mantêm correspondência entre a vogal nuclear da sílaba CVC alvo e a vogal da nova sílaba CV: [za.'li] para *Zair* e ['ku.lu] para *cur*. Todas as demais produções se referem aos alvos *abajur*, *Artur*, *azar* e *mar*, produzidos com a vogal [i] em sílaba CV, como [a.ba.'ʒu.li], [a.'za.li], [a.'tu.li] e ['ma.li].

É preciso considerar que os recursos utilizados pelas crianças aparentam ser diferentes quando comparadas as posições silábicas, observando-se apenas os índices para cada estratégia adotada. No entanto, parece existir um mesmo padrão de estratégia adotada pelos informantes: a criação de uma sílaba CV mais à direita da palavra. Ocorre que há a necessidade de colocar o *tap* em posição CV, uma vez que todos os informantes já adquiriram o segmento nessa posição silábica. Esse fato é claramente expressado pelo sujeito longitudinal.

De acordo com a análise de oitiva, os sujeitos transversais demonstram estarem em uma etapa de aquisição mais avançada que a informante longitudinal. Em termos articulatórios, é possível afirmar que quando a criança não produz o *tap* e opta pelo recurso de omissão, ocorre a redução da magnitude gestual. As substituições do alvo pela lateral [l] demonstram que há um ajuste equivocado do gesto de ponta de língua. Por fim, quando há a um movimento de metátese nas produções, fica evidente que há a tentativa de articulação do alvo, porém, é necessária uma refinação gestual para a coordenação acertada de *timing* entre o gesto de ponta de língua do *tap* e de corpo envolvido na produção da vogal nuclear. A produção do alvo em sílaba CV, portanto, parece acabar sendo mais simples do que em sílaba CVC.

## 4.2 Análise acústica

Para a realização da análise acústica, foram coletados 624 dados de áudio. Contudo, para validar a análise dos dados, alguns dados foram descartados em razão de ruídos ambientais ou sobreposição de vozes. Assim, 581 apresentaram a qualidade necessária para a realização de uma análise confiável.

Esta seção divide-se em três subseções. Em 4.2.1, serão reportados os resultados referentes à análise dos dados por parte do sujeito controle, com a discussão dos achados de cada parâmetro acústico. Ainda, apenas para esse sujeito, será apresentada uma análise acerca da natureza do elemento vocálico que acompanha o *tap*. Na sequência, a subseção 4.2.2, apresenta os resultados referentes a análise dos áudios da informante longitudinal e a subseção 4.2.3 reporta os resultados oriundos das análises dos dados dos sujeitos transversais.

### 4.2.1 Sujeito controle

O sujeito controle, S0, possui uma gravação de áudio e nenhum dado precisou ser descartado, pois esses apresentavam qualidade acústica suficiente para uma análise confiável. Logo, 78 *tokens* compõem a amostra de fala da criança.

Para melhor estruturação do reporte de resultados, as próximas subseções – 4.2.1.1 a 4.2.1.4 – apresentam a análise dos dados referente à inspeção visual do espectrograma, aspectos duracionais, trajetória formântica e natureza formântica do elemento vocálico, respectivamente.

#### 4.2.1.1 Inspeção visual do espectrograma

A representação gráfica do sinal da fala comprehende o espectrograma e a forma de onda. A partir desses, é possível tecer inferências acústico-articulatórias durante a produção de determinado tipo de som. Deste modo, a inspeção visual dos dados do sujeito controle revelou a produção de cinco variantes róticas: *tap*, vibrante, aproximante retroflexa, fricativa glotal e *tap* fricatizado, como mostra a Tabela 3.

Variante rótica produzida	CVC MEDIAL				CVC FINAL			
	[a]	[i]	[u]	Total	[a]	[i]	[u]	Total
Tap	6/9	6/6	6/9	<b>18/24</b>	0/9	0/9	0/9	<b>0/27</b>
Vibrante	3/9	0/6	3/9	<b>6/24</b>	5/9	7/9	9/9	<b>21/27</b>
Aproximante Retroflexa	0/9	0/6	0/9	<b>0/24</b>	3/9	1/9	0/9	<b>4/27</b>
Fricativa Glotal Desvozeada	0/9	0/6	0/9	<b>0/24</b>	1/9	0/9	0/9	<b>1/27</b>
Tap Fricatizado	0/9	0/6	0/9	<b>0/24</b>	0/9	1/9	0/9	<b>1/27</b>

Tabela 3 - Produção das variantes róticas de acordo com o ambiente vocálico e a posição na palavra por S0

A Tabela 3 revela que a variante rótica mais empregada, em posição medial de palavra, foi o *tap*, com 18 ocorrências (75%). Além dessa variante, verificaram-se 6 produções (25%) de vibrantes, sendo 3 em contexto da vogal baixa e 3 em contexto da vogal alta posterior. Para posição final de sílaba, o *tap* não ocorreu em nenhum ambiente vocálico. No entanto, a vibrante foi a variante predominante, apresentando 21 ocorrências (78%), seguida pela aproximante retroflexa. Esta foi percebida em 4 ocorrências (15%), como os alvos *lugar* e *azar*. Outras duas variantes para o rótico foram observadas, a saber: fricativa glotal e *tap* fricatizado. Faz-se necessário destacar que a ocorrência de vibrantes, para ambas posições silábicas, foi identificada por Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2021) nas produções de locutores adultos, falantes do mesmo dialeto que o referido sujeito controle.

A fim de inferir o modo como alguns alvos foram produzidos, serão expostos, ao longo dessa seção, exemplos para cada posição na palavra da sílaba CVC contendo o rótico. Para ilustrar a inspeção visual dos dados produzidos por S0 em posição medial, seguem as Figuras 19 e 20.

A Figura 19 apresenta a forma de onda e o espectrograma da palavra *turco*, na qual o sujeito controle produz um *tap*, a variante rótica mais articulada para posição medial.

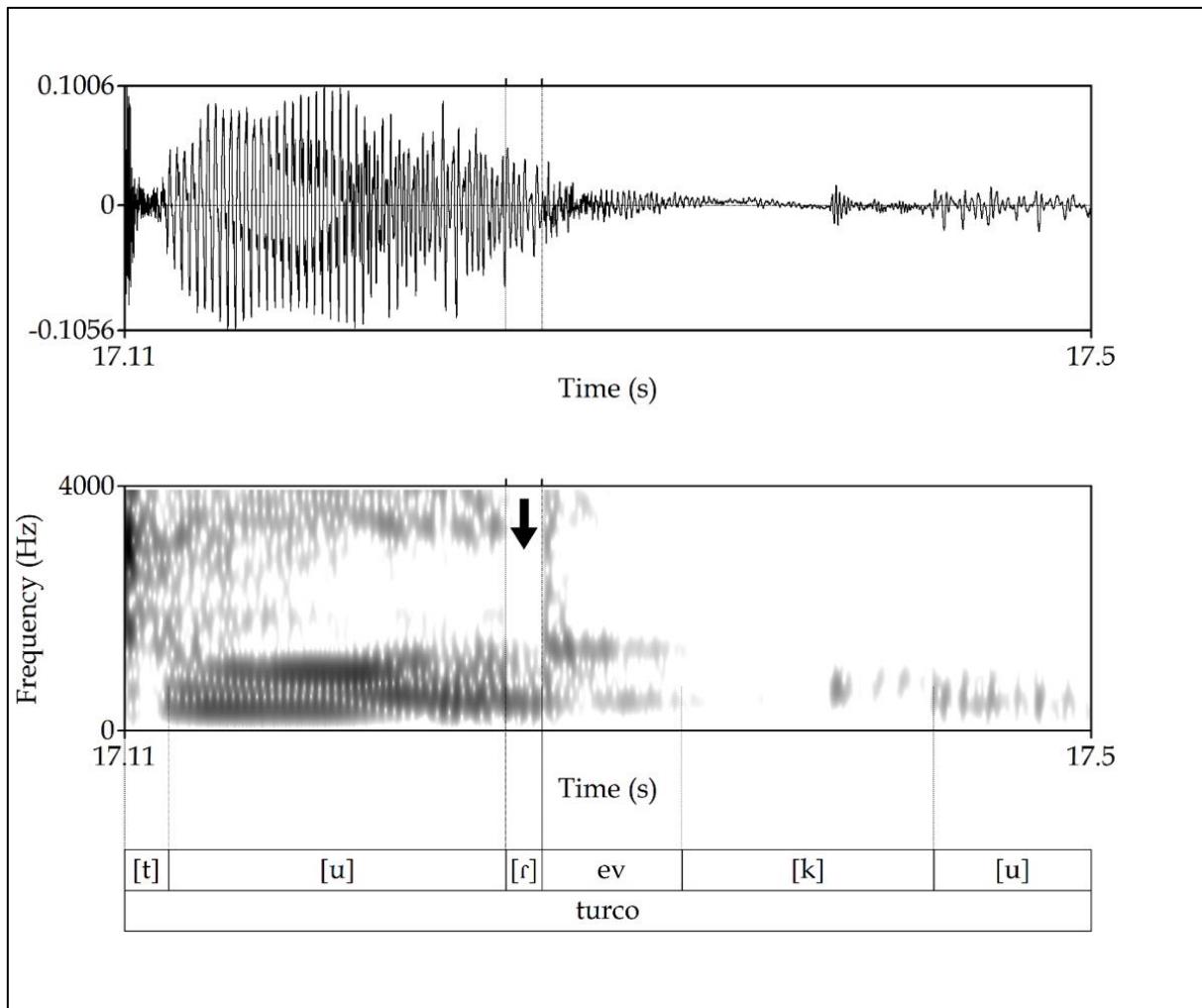


Figura 19 - Forma de onda e espectrograma de banda larga da produção de ['tur.ku] do alvo "turco" por S0. Em destaque pela seta a produção do *tap*.

O espectrograma gerado para o alvo *turco* exibe uma faixa quase em branco, seguido por uma plosão e um elemento vocálico. Na Figura 19, a seta aponta o momento de queda abrupta da amplitude<sup>20</sup>. Em seguida, é possível observar uma breve plosão, seguida por momento acústico de estrutura formântica contínua e similar a uma vogal, o elemento vocálico (ev). Essas qualidades acústicas permitem inferir que um *tap* foi articulado pelo sujeito controle. No que se refere à produção de elementos vocálicos, em todos os casos em que a produção de um *tap* foi identificada, em posição medial, o segmento estava acompanhado pelo elemento vocálico.

Ainda em posição medial, a Figura 20 exibe a representação gráfica da produção de uma vibrante alveolar.

<sup>20</sup> O espectrograma registra a amplitude das frequências do sinal acústico. Uma maior a amplitude do sinal é representada por um maior escurecimento dos traçados. Em contrapartida, a menor amplitude pode ser identificada por faixas mais claras no espectrograma (SILVA et al., 2019).

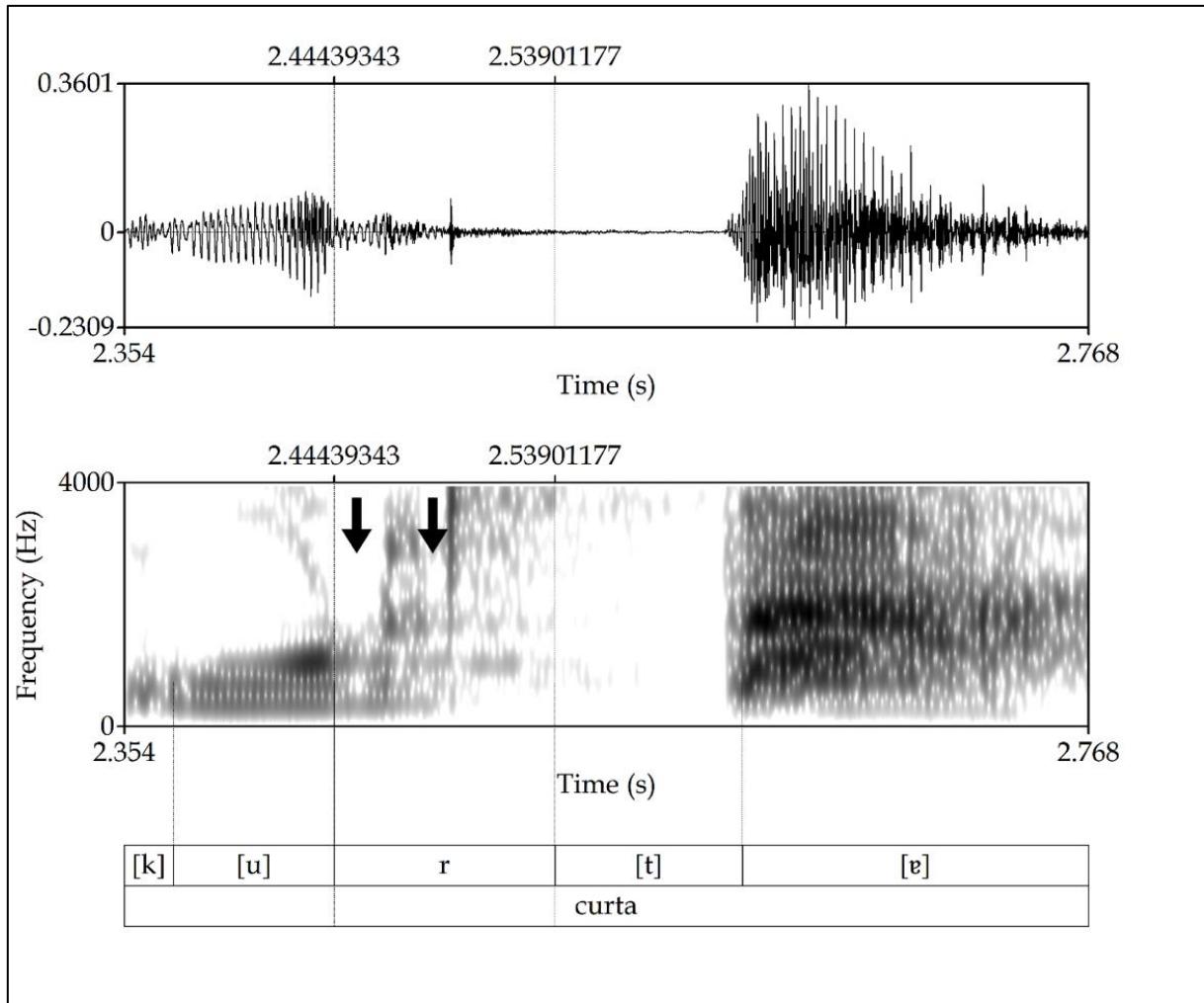


Figura 20 - Forma de onda e espectrograma de banda larga da produção de ['kur.tə] do alvo "curta" por S0. Em destaque pela setas a produção da vibrante.

Ao observar o espectrograma e a forma de onda da palavra *curta*, é possível identificar os períodos de vibração do segmento. As setas apontam dois espaçamentos quase em branco, que são os correlatos acústicos do fechamento do trato vocal. Após cada espaçamento, há um elemento vocálico que reflete o momento de abertura do trato vocal. A partir dessas qualidades acústicas, pode-se inferir que o sujeito controle produziu uma vibrante alveolar em posição medial.

Para ilustrar a inspeção visual dos dados em posição final de S0, seguem as figuras 21 e 22.

A Figura 21 ilustra o espectrograma e a forma de onda gerados para o alvo *azar*, na qual o sujeito controle articula uma vibrante, a variante rótica predominante em posição final.

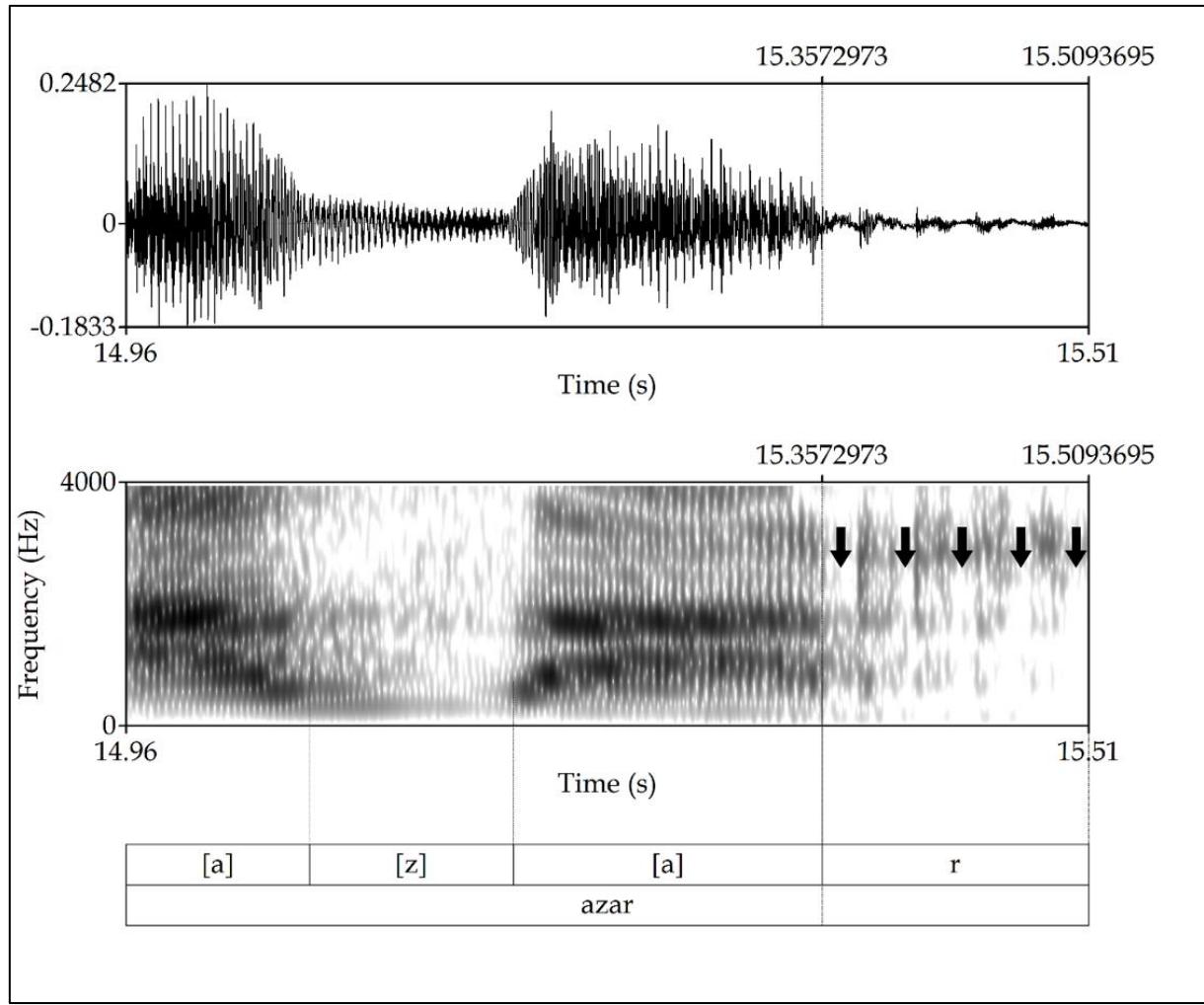


Figura 21 - Forma de onda e espectrograma de banda larga da produção de [‘a.zar] do alvo "azar" por S0. Em destaque pelas setas a produção da vibrante.

É possível identificar os períodos de vibração ao observar a forma de onda e o espectrograma do alvo *azar*. As setas apontam vários espaçamentos em branco que demonstram os momentos de baixa energia gerados pelo movimento de fechamento do trato vocal. Do mesmo modo, é possível observar a produção de um elemento vocálico que acompanha cada espaço em branco no espectrograma. Com base nessas evidências acústicas, infere-se que o sujeito controlou articulou uma vibrante alveolar.

A Figura 22 apresenta o espectrograma e a forma de onda da palavra *lugar*, na qual o sujeito controle produz uma aproximante, a segunda variante rótica mais articulada na posição final.

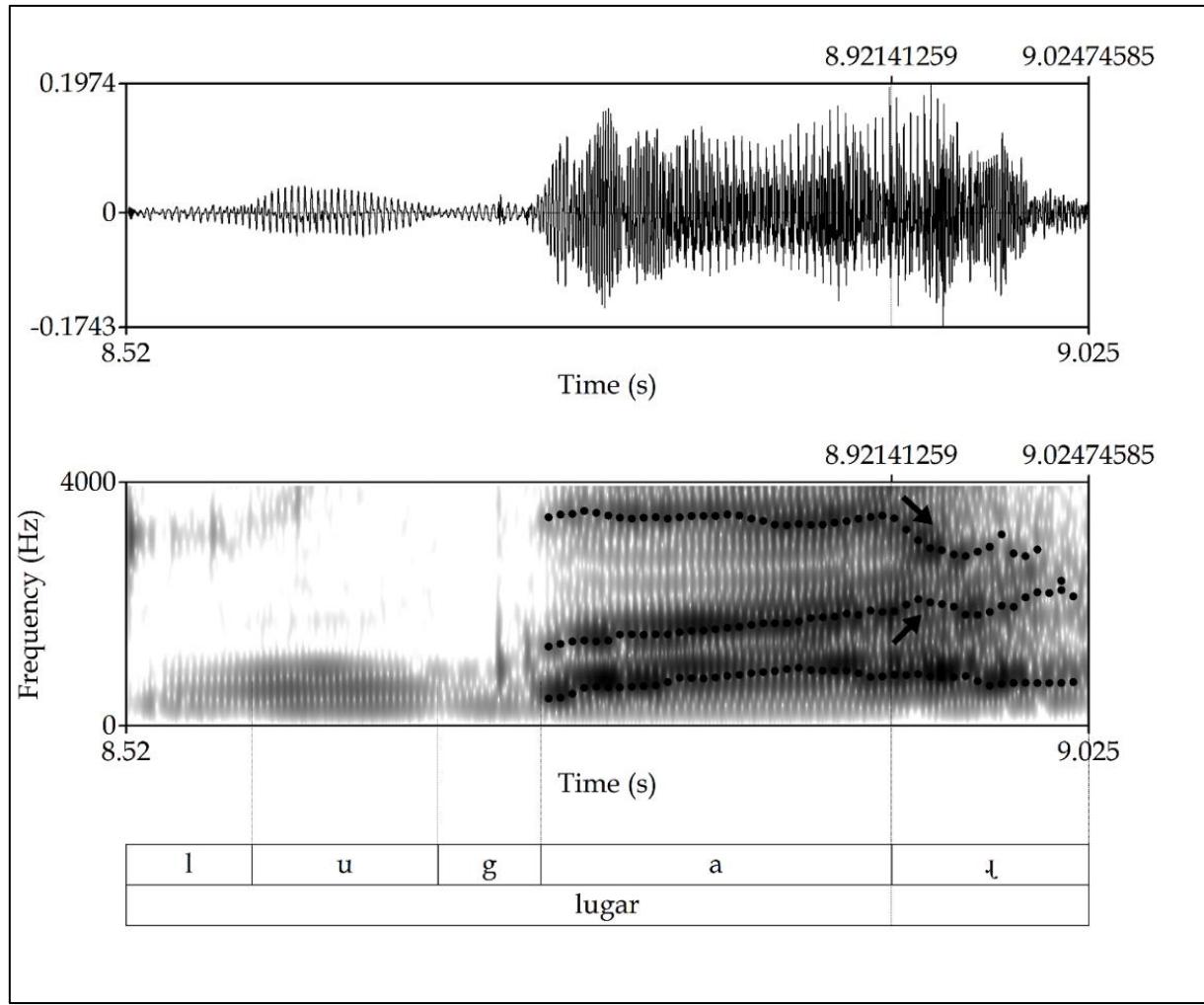


Figura 22 - Forma de onda e espectrograma de banda larga da produção de [‘lu.gau] do alvo “lugar” por S0. Em destaque pelas setas a trajetória formântica.

Após a imagem acústica da vogal baixa, no espectrograma, é possível observar um segmento de estrutura formântica regular e contínua. A primeira seta aponta a trajetória descendente de F3 e a segunda seta indica uma ascendência de F2. A aproximante retroflexa é caracterizada acusticamente por Silva *et al.* (2019) em quatro aspectos: (i) trajetória contínua dos três primeiros formantes; (ii) trajetória ascendente de F2; (iii) trajetória descendente de F3 e (iv) F3 acima de 2.000 Hz. Os valores encontrados para o dado representado pela Figura 29 são: F1 – 735 Hz, F2 – 1810 Hz, F3 – 2641 Hz. Para dados do português brasileiro, Ferraz (2005) defende que a principal característica para a aproximante retroflexa é a trajetória descendente do terceiro formante, apesar de seu valor ser alto ou baixo. Cabe ressaltar que, conforme esperado, os valores reportados para as frequências de crianças são mais altos do

que valores encontrados nas produções de adultos<sup>21</sup>. Sendo assim, a partir das qualidades acústicas apresentadas, infere-se que o sujeito controle produziu uma aproximante retroflexa.

A inspeção visual das formas de ondas e dos espectrogramas, gerados para as produções do sujeito controle, permitem inferir que houve a produção de cinco variantes róticas, a saber: *tap*, vibrante, aproximante retroflexa, fricativa glotal e *tap* fricatizado. Em todas as palavras que continham o *r* como alvo, esse não foi omitido ou substituído por outro som. As formas predominantes, no entanto, foram o *tap* em posição medial e a vibrante em posição final.

#### 4.2.1.2 Aspectos duracionais

A análise dos valores de duração auxilia na identificação do tipo de som articulado, como no caso de *taps* e vibrantes, bem como fornece pistas, por meio de inferências, acerca do aprimoramento dos gestos articulatórios produzidos pela criança. Assim, serão apresentados os valores de duração absoluta e relativa do segmento rótico produzido pelo sujeito controle. Mesmo que o foco principal da presente pesquisa seja analisar a aquisição do *tap*, é preciso considerar o alto número de produções de vibrantes articuladas pelo sujeito controle, especialmente em posição final de palavra. Faz-se necessária, portanto, a descrição dos valores para ambos os sons. Ainda, esta seção busca descrever os dados do sujeito controle que serão utilizados para estabelecer comparações com as produções da informante longitudinal e dos sujeitos transversais.

A partir das referidas considerações, a Tabela 4 exibe as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição medial, abarcavam o rótico como alvo – carpa, farpa, carpa, circo, círculo, curta, curto, turco. Na mesma tabela, são exibidos os valores de duração relativa do rótico em relação à sílaba. Os valores reportados a seguir são referentes às produções nas quais o sujeito controle articula um *tap*.

---

<sup>21</sup> Conforme a teoria fonte-filtro, o comprimento do trato vocal influencia nas frequências de ressonância. O trato de uma criança é menor do que o trato vocal de um adulto. Logo, as frequências de ressonância para as cavidades orais de crianças são mais altas. Para saber mais, ver Kent e Read (2015), capítulo 2.

CODA MEDIAL					
AR		Duração absoluta (ms)			Duração relativa (%)
		Sílaba	Rótico	Vogal	(rótico - sílaba)
	MÉDIA	246	19	120	8,29
IR	DP	77,59	1,02	3,78	2,25
	MÉDIA	307	17	99	5,65
UR	DP	7,31	4,24	1,41	1,39
	MÉDIA	202	21	121	10,52
	DP	21,62	11,30	24,15	2,34

Tabela 4 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como *tap*) produzidos pelo sujeito controle em posição medial, conforme contexto vocálico

Com base nos valores reportados pela Tabela 4, é possível observar que o rótico apresenta valores de duração absoluta entre 17 a 21 ms, ocupando de 5,65% a 10,52% da sílaba. A vogal alta posterior, além de exibir um valor maior de duração absoluta, também é o segmento vocálico que apresenta maior variabilidade em seus valores. Essa afirmação apoia-se nos maiores valores de desvio padrão reportados pelos dados da vogal [u].

As medidas de duração relativa, no entanto, não são consideradas relevantes aqui, uma vez que as sílabas CVC analisadas incluem diferentes modos de articulação na primeira consoante, o que responde pelo baixo percentual de duração relativa no contexto da vogal [i] – as sílabas CVC de *circo* e *círculo* apresentam um segmento fricativo de maior duração – e pelo alto percentual no contexto da vogal [u] – as sílabas CVC de *curta*, *curto* e *turco* apresentam consoantes plosivas de menor duração.

O Gráfico 17 apresenta as médias de duração absoluta da sílaba, da vogal e do rótico das produções das palavras realizadas por S0. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa; (IR) circo, círculo; (UR) curto, curta, turco. Os valores reportados para o rótico são referentes às produções nas quais a criança articula um *tap*.

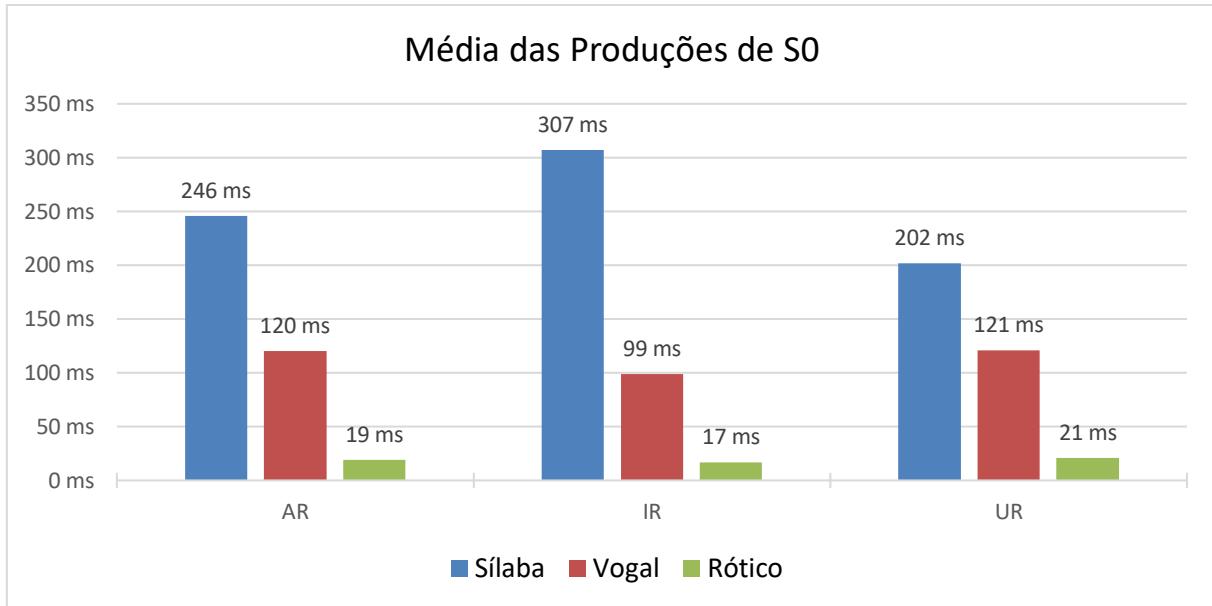


Gráfico 17 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico articulado como *tap* (verde) das produções de S0, para a posição medial

Ao analisar do Gráfico 17, pode-se observar que há diferença de duração da sílaba composta pela vogal [i] em comparação com as outras sílabas. Essa diferença, conforme já reportado, está relacionada com a consoante que integra as palavras com essa vogal. Quanto à duração do rótico, pode-se afirmar que esse foi produzido da forma convencional da variante *tap*, pois seu valor de duração está próximo ao reportado pela literatura do português brasileiro.

Após a análise dos dados, nos quais o sujeito controle produz o *tap*, a Tabela 5 apresenta os valores de duração referentes às produções nas quais a criança articula uma vibrante. A referida tabela exibe as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição medial, abarcavam o rótico como alvo – carpa, farpa, carpa, circo, círculo, curta, curto, turco. Do mesmo modo, são apresentados os valores de duração relativa do rótico em relação à sílaba.

CODA MEDIAL					
AR		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(rótico - sílaba)
	MÉDIA	208	70	110	33,48
IR	DP	11,67	25,46	15,20	10,28
	MÉDIA	—	—	—	—
UR	DP	—	—	—	—
	MÉDIA	208	68	101	33,12
	DP	19,45	3,18	1,06	2,25

Tabela 5 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como vibrante) em posição medial do sujeito controle, conforme contexto vocálico

Na Tabela 5, é possível observar que o rótico vibrante foi produzido apenas em contexto de [a] e [u] para a posição medial, apresentando duração de 70 ms e 68 ms, respectivamente. Ambos os segmentos ocuparam a mesma proporção da sílaba, cerca de 33%. Ao examinar os valores médios de duração da vogal, nota-se que a vogal alta posterior apresenta 101 ms de duração absoluta e a vogal baixa, 110 ms.

O Gráfico 18 exibe as médias de duração absoluta da sílaba, da vogal e do rótico das produções realizadas por S0. Os valores apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa; (UR) curto, curta, turco. Os valores reportados para o rótico são referentes às produções nas o sujeito controle articula uma vibrante.

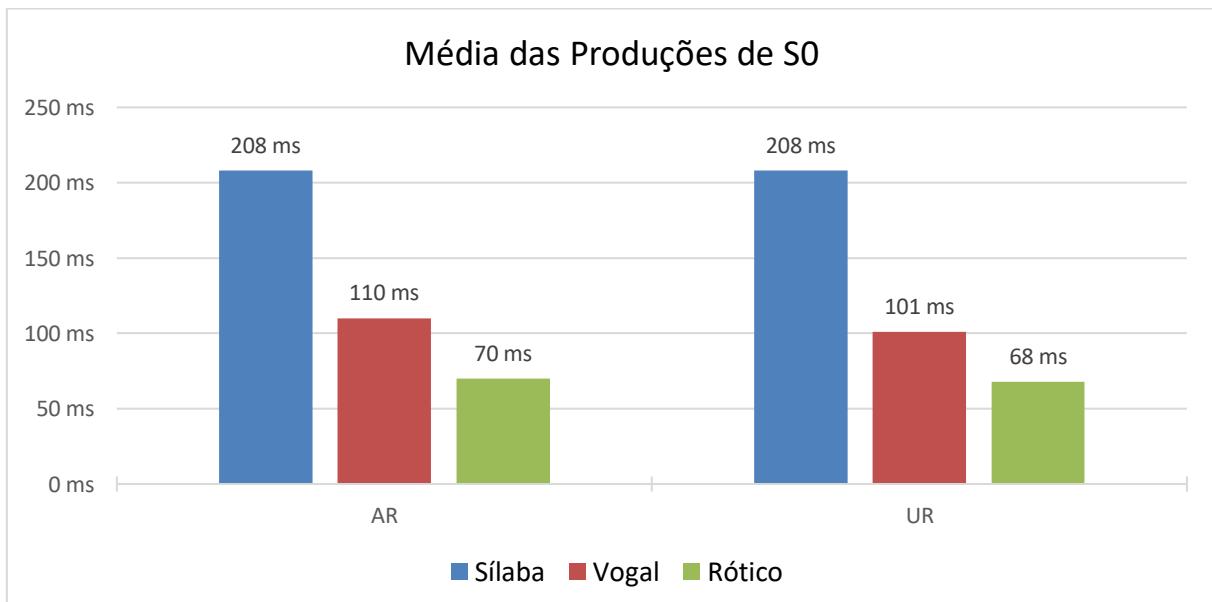


Gráfico 18 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico articulado como vibrante (verde) das produções de S0, para a posição medial

A partir da análise do Gráfico 18, percebe-se que não há uma diferença expressiva entre as médias de duração absoluta da sílaba, vogal e rótico das palavras em ambiente vocálico [a] e [u]. A média de duração da vibrante que está acompanhada pela vogal baixa produzida pela criança, 70 ms, aproxima-se do valor reportado por Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2021) para uma informante adulta. A informante S4, do referido estudo, apresentou uma média de duração de 68 ms durante a produção da vibrante em contexto de [a], em posição medial. Sendo assim, pode-se afirmar que as produções de vibrantes em posição medial, articuladas pelo sujeito controle, assemelham-se a produções já encontradas em dados de adultos para a mesma posição silábica.

Na sequência, serão reportados os resultados referentes às médias de durações absoluta e relativa das produções em posição final de sílaba.

Conforme reportado pela análise visual dos dados do sujeito controle, não foi constatada a produção do *tap* em posição final. Outras três variáveis do rótico foram identificadas, no entanto, para o caráter descritivo dos dados, a partir do parâmetro de duração, faz-se necessário aqui apenas para a identificação de *taps* e vibrantes. Da mesma forma, o intuito desta seção é fornecer os resultados relacionados às produções do sujeito controle para que se possa estabelecer comparações com os dados dos demais informantes da pesquisa.

Com base nas considerações tecidas, a Tabela 6 apresenta as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição final, abarcavam o rótico como alvo – mar, lugar, azar, vir, Nadir, Zair, cur, abajur, Artur. Além disso, são exibidos os valores de duração relativa do rótico em relação à sílaba. Os valores expostos a seguir são referentes às produções nas quais a criança produz uma vibrante.

CODA FINAL					
AR	MÉDIA	Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(rótico - sílaba)
	DP	465	201	174	41,73
IR	MÉDIA	397	157	150	39,01
	DP	23,41	20,04	67,18	3,35
UR	MÉDIA	372	155	140	41,65
	DP	14,11	6,83	3,75	3,60

Tabela 6 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como vibrante) em posição final produzidos pelo sujeito controle, conforme contexto vocálico

Ao observar os resultados exibidos na Tabela 6, percebe-se que a vibrante em contexto de [a] apresenta o maior valor de duração absoluta, 201 ms. Para o contexto de [i], a variante rótica apresenta uma média de 157 ms de duração absoluta. Por fim, a vibrante produzida acompanhada de [u], apresentou 155 ms de duração, sendo o menor valor apresentado. Ao considerar a porção que o rótico ocupava no interior da sílaba, nota-se que estes ocupam entre 39,01% e 41,73% da mesma.

O Gráfico 19 apresenta os resultados das médias de duração absoluta da sílaba, da vogal e do rótico – articulado como vibrante – das produções dos itens lexicais contendo coda final realizados por S0. A disposição dos dados se dá conforme os seguintes contextos: (AR) mar, lugar, azar; (IR) vir, Nadir, Zair; (UR) cur, abajur, Artur.

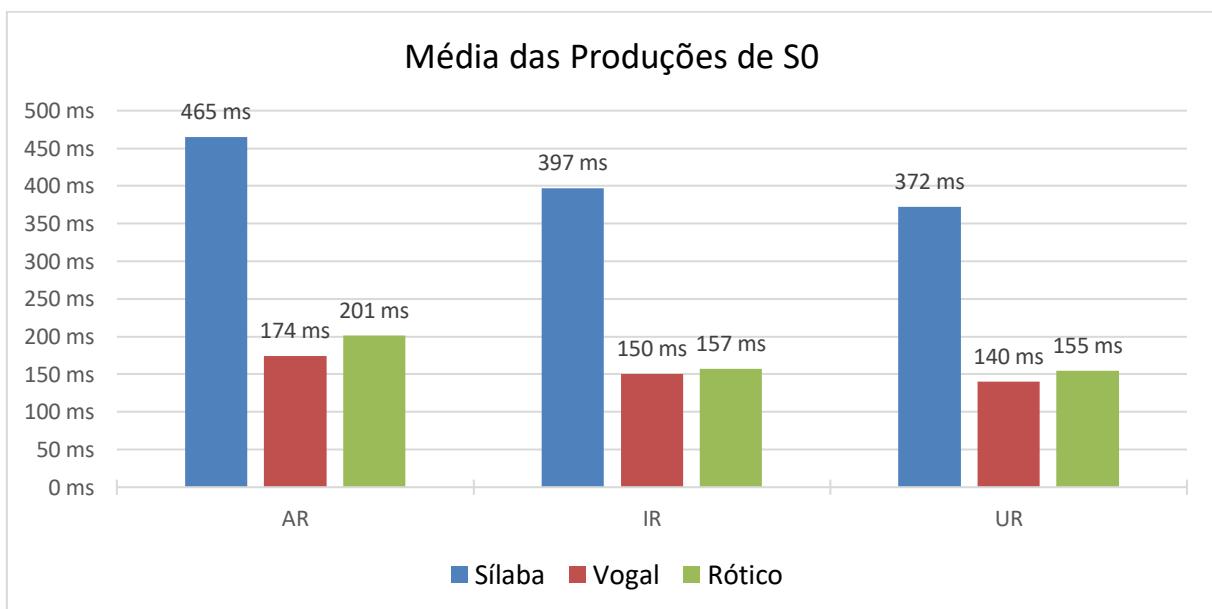


Gráfico 19 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico articulado como vibrante (verde) das produções de S0, para a posição final

O Gráfico 19 revela que a duração da vibrante foi maior do que a duração da vogal nuclear em todos os ambientes vocálicos. Os valores de duração absoluta do rótico giram em torno de 155 ms e 201 ms. Posto isto, nota-se que o segmento rótico foi produzido como uma vibrante, tendo em vista que seus valores de duração se assemelham aos que foram descritos por outros estudos do português brasileiro.

Ao observar os resultados oriundos da análise de duração das produções do sujeito controle, percebe-se que, em posição medial, o *tap* apresentou valores de duração entre 17 ms e 21 ms, ocupando de 5,65% a 10,52% da sílaba. Para a mesma posição, a vibrante foi produzida apenas em contexto de [a] e [u], com uma duração de aproximadamente 70 ms, o que representou cerca de 33% da sílaba. Em posição final, não ocorreu a articulação de *tap*, no entanto, a vibrante foi a variante rótica predominante. O segmento apresentou valores de duração entre 155 ms e 201 ms, ocupando de 39,01% a 41,73% da sílaba. Ao comparar os valores dos segmentos vibrantes que se apresentam em ambas posições silábicas, nota-se que a vibrante produzida em final de palavra dura cerca de 2,2 vezes mais do que o segmento articulado em posição medial.

Como já mencionado, os resultados de duração do sujeito controle aproximam-se dos reportados em trabalhos que analisaram dados de adultos, falantes do mesmo dialeto. Fato esse que sinaliza para uma aquisição e estabilização do som no inventário fonético-fonológico da criança. Logo, infere-se que ao sujeito controle consegue coordenar o *timing* dos gestos articulatórios para realizar o segmento rótico em coda silábica.

#### 4.2.1.3 Trajetória formântica

A análise da trajetória formântica nos dados do sujeito controle auxilia a compreender como desenrola-se a transição dos formantes na presença do rótico nos dados de uma criança com a fase de aquisição segmental concluída. Assim, é possível, nas próximas seções, verificar a presença de pistas de produções gradientes nos dados dos demais sujeitos, pois, uma vez que se sabe o que pode ser encontrado na trajetória formântica dos dados do sujeito controle, torna-se possível tecer algumas comparações e inferências acerca de produções outras.

A análise da trajetória formântica dos dados do sujeito controle foi feita, para a posição medial, observando o comportamento da vogal quando essa estava

acompanhada pelo *tap*, quando estava em contexto CV ou quando estava acompanhada pelo glide em sílabas CVV.

Para tanto, serão apresentados gráficos e tabelas com os dados presentes no *corpus* de tríades. Os gráficos são oriundos de uma produção de cada alvo, ao passo que os quadros foram construídos a partir da média geral das produções.

Os Gráficos 20, 21 e 22 apresentam as trajetórias formânticas de F1, F2 e F3 de produções realizadas por S0, em contexto da vogal baixa: carta X cata X caixa.

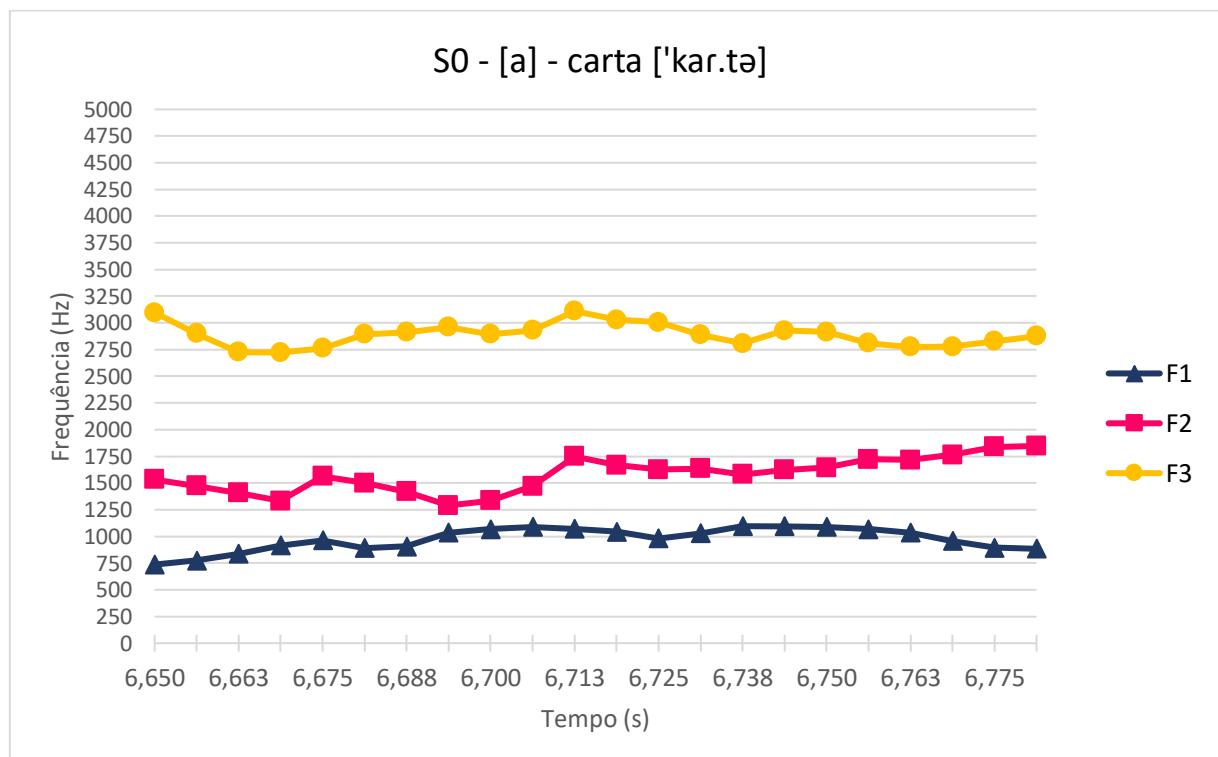
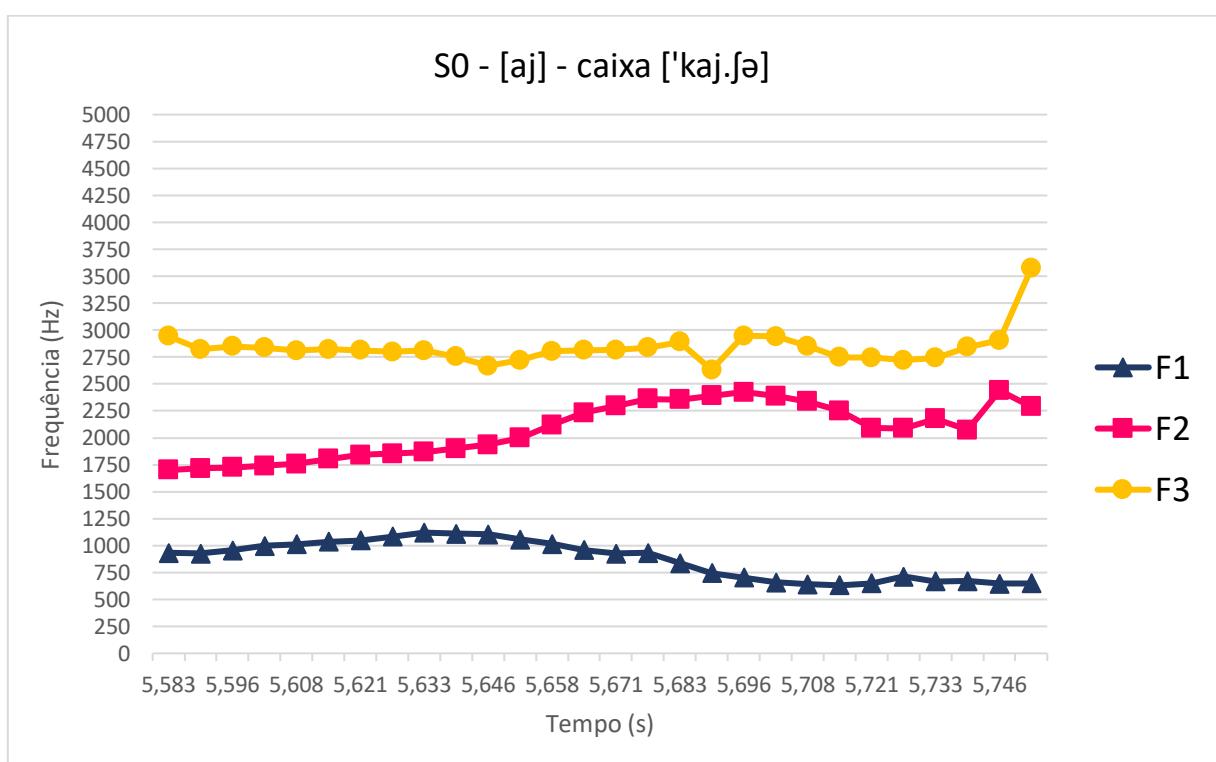
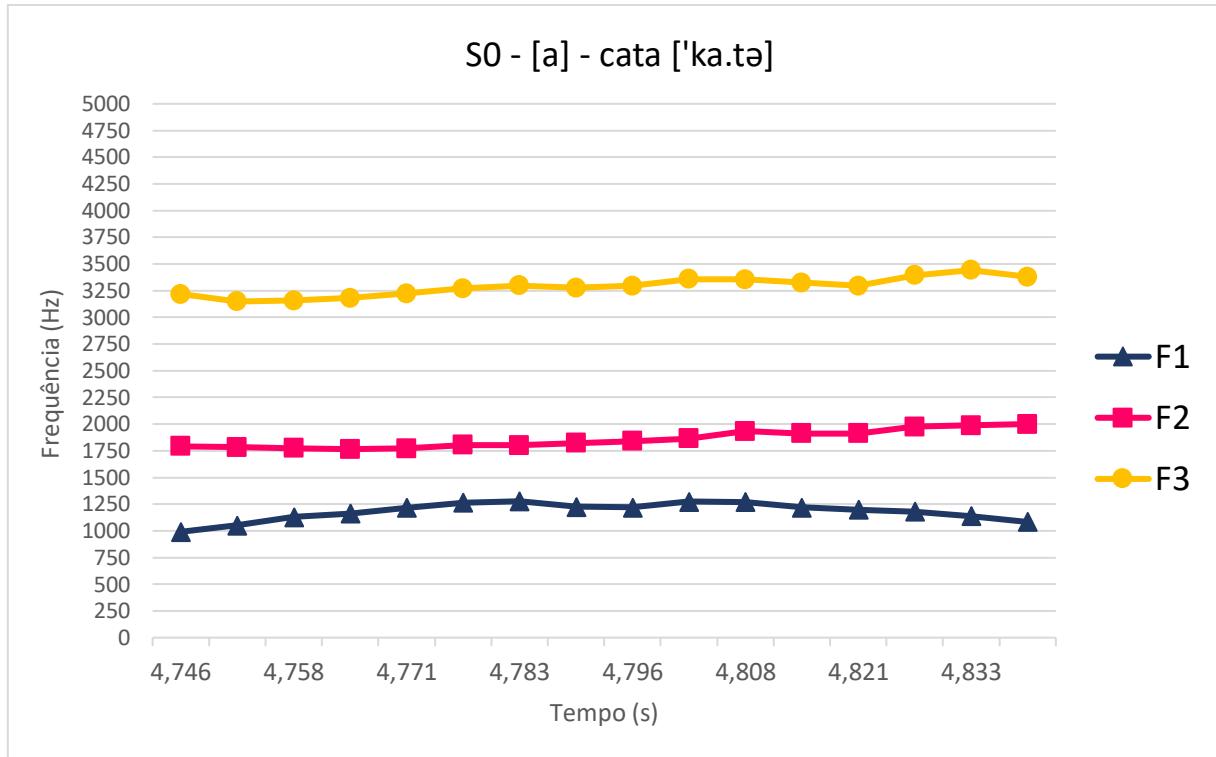


Gráfico 20 – Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [a] que antecede o rótico na produção de ['kar.tə] por S0



Ao comparar a tríade de palavras, percebe-se que, a partir do Gráfico 20, há sensíveis curvas de subida e decida durante a trajetória dos três formantes. Ainda, é possível notar uma queda da frequência para F1, um aumento no segundo formante e um suave aumento de F3 na porção final da produção da vogal [a] que antecede o *tap*, ou seja, na chamada zona de transição formântica.

No Gráfico 21, nota-se que há uma certa estabilidade durante a trajetória dos três primeiros formantes, o que indica que não há alterações articulatórias expressivas na produção da vogal em contexto de [a]. É possível notar um distanciamento maior entre os valores das frequências do segundo e terceiro formante, especialmente pela ausência de elevação de F2.

A trajetória formântica para o alvo em contexto de [aj], presente no Gráfico 22, evidencia a transição gradual do primeiro e segundo formante entre os dois segmentos vocálicos que compõem o ditongo. Em outras palavras, o distanciamento entre os valores das duas primeiras frequências que são típicos da articulação de ditongos, segundo Silva *et al.*, (2019).

A Tabela 7 revela as médias formânticas do início, meio e fim da vogal/ditongo nas palavras em contexto [ar] (carta, farpa, carpa), [a] (cata) e [aj] (caixa) realizadas por S0.

[a] carta, carpa, farpa									
INÍCIO			MEIO			FIM			
	F1	F2	F3	F1	F2	F3	F1	F2	F3
MÉDIA (Hz)	708	1595	2933	992	1681	3009	847	1772	2863
DP	141,01	479,95	424,75	226,95	395,51	523,39	118,50	340,25	706,19
[a] cata									
INÍCIO			MEIO			FIM			
	F1	F2	F3	F1	F2	F3	F1	F2	F3
MÉDIA (Hz)	880	1831	3139	1234	1841	3492	1104	1900	3622
DP	80,12	167,62	309,36	52,21	117,46	324,24	168,59	187,73	322,13
[aj] caixa									
INÍCIO			MEIO			FIM			
	F1	F2	F3	F1	F2	F3	F1	F2	F3
MÉDIA (Hz)	852	1737	2816	886	2116	3052	572	2434	3364
DP	91,14	64,90	126,05	79,76	161,55	412,01	54,03	539,97	218,92

Tabela 7 - Médias e desvio-padrão para os valores de F1, F2 e F3 nos trechos de início, meio e fim da trajetória formântica de [a], em contexto [ar], [a] e de [aj], nas produções de S0

Com base na Tabela 7, para o contexto [ar], é possível constatar que, para os valores de F1, F2 e F3, há oscilações com elevações e quedas nos valores das frequências. Esse comportamento foi reportado no Gráfico 20 como *sensíveis curvas de subida e decida*. É possível observar, no entanto, do meio para o fim, uma queda de F1 e elevação de F2. Para F3, observa-se uma elevação do início para o meio da vogal. Quanto aos dados em contexto de [a], nota-se que o primeiro formante apresenta uma sensível curva de subida e descida, ao passo que os valores se elevam sensivelmente para F2 e de forma um pouco mais acentuada para F3. No que concerne ao comportamento dos formantes para as produções em contexto de [aj], observa-se uma considerável queda na média reportada para F1, à medida que as frequências do segundo e do terceiro formante apresentam elevação.

Os Gráficos 23 e 24 apresentam as trajetórias formânticas de F1, F2 e F3 de produções realizadas por S0, em contexto da vogal alta anterior: círculo X chico.

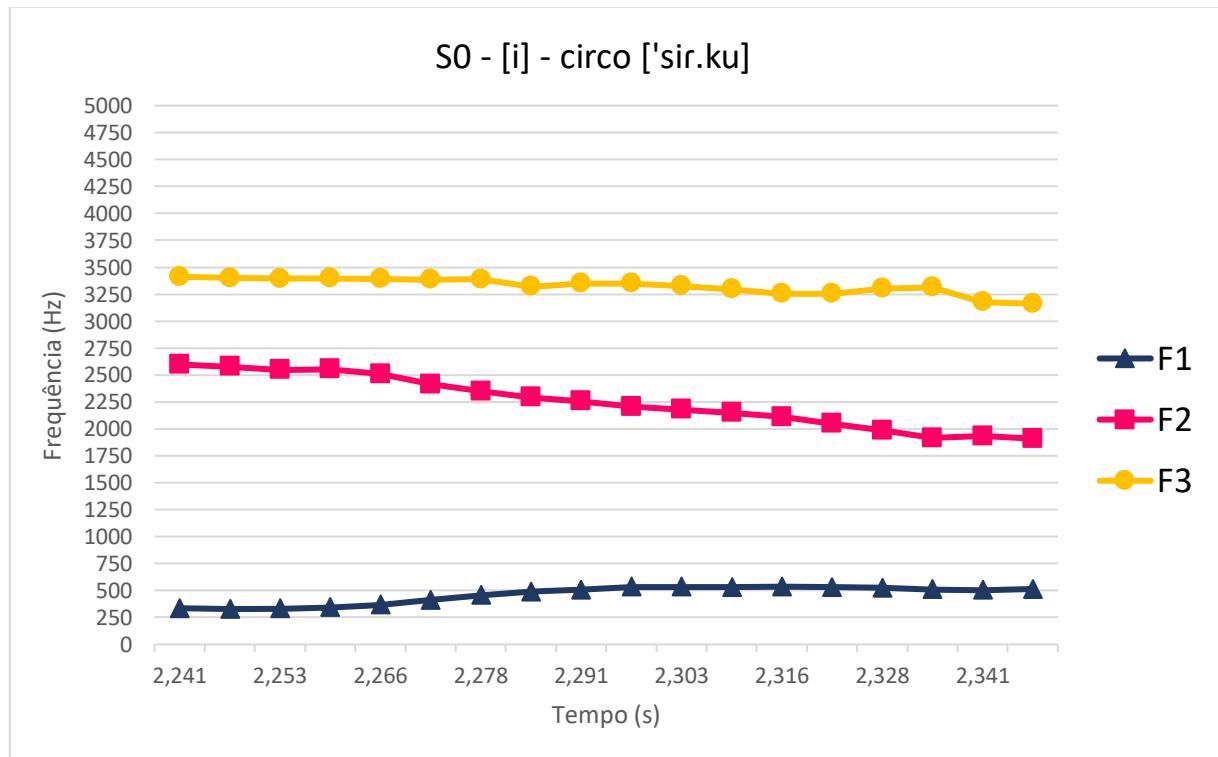


Gráfico 23 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] que antecede o rótico na produção de ['sir.ku'] por S0

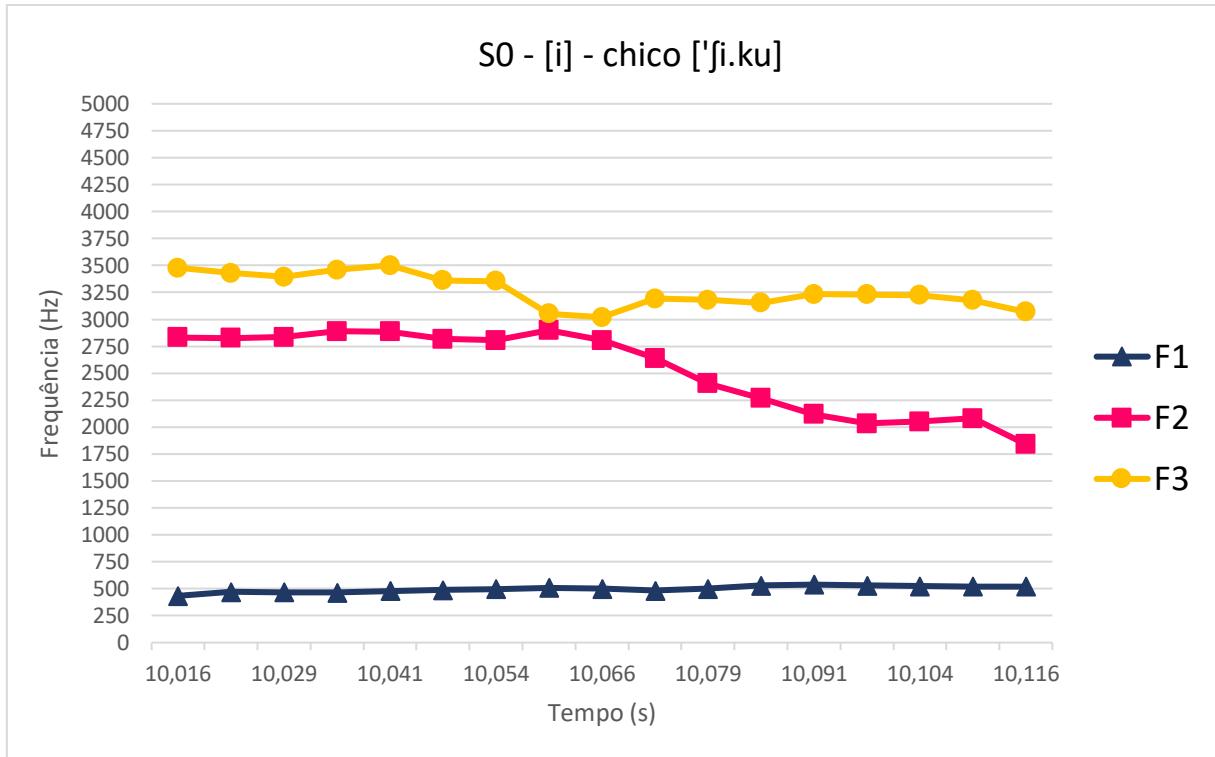


Gráfico 24 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] de ['ʃi.ku] por S0

A inspeção visual das produções de *circo* e *chico* permitiu concluir que, para o contexto IR, a partir do Gráfico 23, o primeiro formante apresenta uma breve elevação, ao passo que as frequências de F2 e F3 diminuem, sendo de modo mais acentuado para o segundo formante. Quanto aos dados em contexto de I, percebe-se que, no Gráfico 24, o primeiro formante está horizontalmente alinhado. Já o segundo e terceiro formantes apresentam uma considerável queda, sendo de forma mais intensa para o F2.

A Tabela 8 revela as médias formânticas do início, meio e fim da vogal nas palavras em contexto IR (circo, circulo) e I (chico) produzidas por S0.

[i] circo, circulo									
	INÍCIO			MEIO			FIM		
	F1	F2	F3	F1	F2	F3	F1	F2	F3
MÉDIA (Hz)	383	2409	3361	499	2359	3230	528	2092	3047
DP	33,24	26,10	21,03	120,84	61,23	88,24	19,65	72,51	1,62
[i] chico									
	INÍCIO			MEIO			FIM		
	F1	F2	F3	F1	F2	F3	F1	F2	F3
MÉDIA (Hz)	551	2699	3252	467	2405	3011	492	2003	2985
DP	281,13	131,97	94,97	58,92	79,50	31,84	28,15	86,05	89,46

Tabela 8 - Médias e desvio-padrão para os valores de F1, F2 e F3 nos trechos de início, meio e fim da trajetória formântica de [i], em contexto [ir] e [i], nas produções de S0

Com base nos valores reportados pela Tabela 8, observa-se que, para as produções em contexto de IR, há uma elevação do F1, ao passo que F2 e F3 apresentam trajetórias decrescentes. Para as produções, em contexto de I, o primeiro formante manteve-se próximo de 500 Hz. Em relação à média de F2, essa apresenta uma considerável queda, já a queda de F3 exibe-se de forma menos acentuada. Desse modo, corrobora-se o que já tinha sido apontado pela inspeção visual para os dois contextos.

Os Gráficos 25, 26 e 27 apresentam as trajetórias formânticas de F1, F2 e F3 de produções realizadas por S0, em contexto da vogal alta posterior: curto X cuca X cuida.

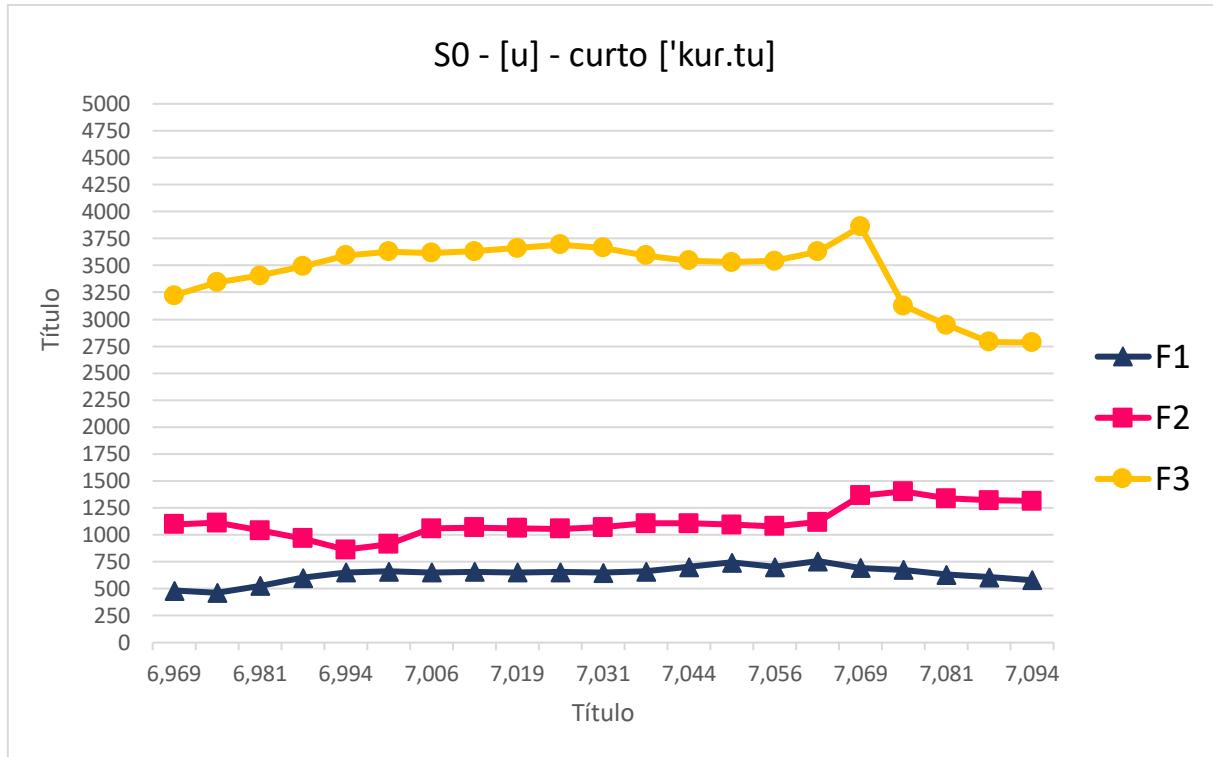


Gráfico 25 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u] que antecede o rótico na produção de ['kur.tu] por S0

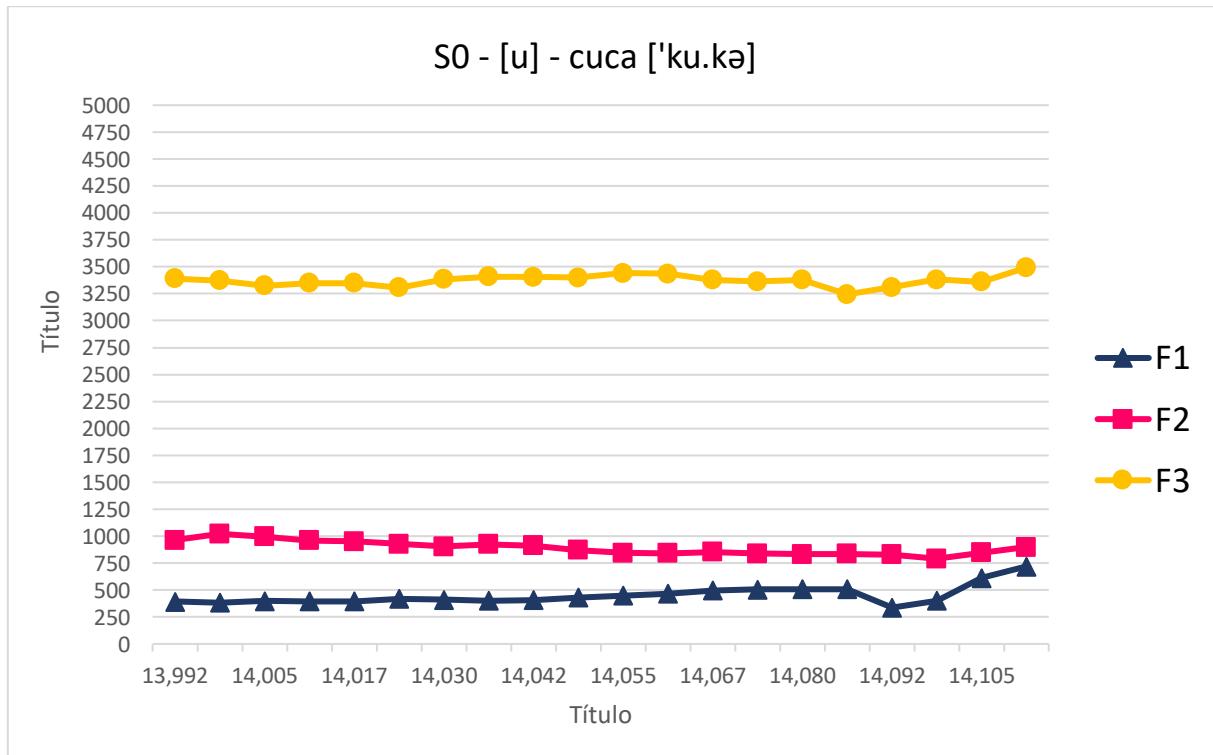


Gráfico 26 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u] na produção de ['ku.kə] por S0

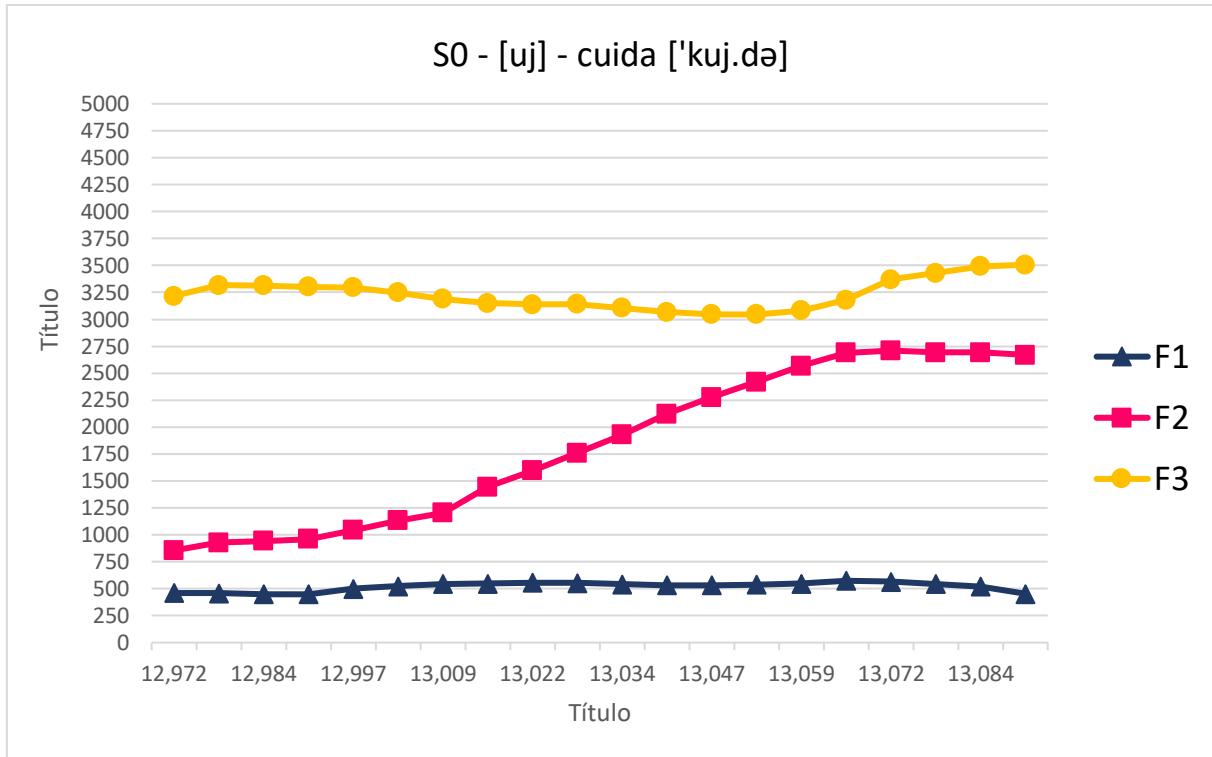


Gráfico 27 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) do ditongo [uj] na produção de ['kuj.də] por S0

Ao comparar as palavras em contexto [ur] x [u] x [uj], é possível observar, a partir do Gráfico 25, que os valores das frequências para o terceiro formante decrescem na parte final da produção da vogal quando essa é seguida pelo *tap*. Do mesmo modo, o segundo formante parece se distanciar de forma suave do primeiro na porção final da vogal. Com base na trajetória formântica para o alvo em contexto de [u], presente no Gráfico 26, percebe-se que os três formantes estão horizontalmente alinhados, apresentando um distanciamento expressivo entre F2 e F3. Há uma elevação, na porção final da vogal, em F1 e F3. Os valores reportados para as frequências do segundo formante permanecem em torno de 1000 Hz, ao passo que para F3 as frequências mantêm-se próximo de 3400 Hz. Para [uj], no Gráfico 27, nota-se que uma sensível elevação de F3 na porção final da vogal. Os valores de F1 e F2 evidenciaram a transição gradual entre as vogais que constituem o ditongo, o primeiro cai suavemente enquanto o segundo aumenta de forma gradual e intensa.

A Tabela 9 revela as médias formânticas do início, meio e fim da vogal/ditongo nas palavras em contexto [ur], [u] e [uj] realizadas por S0.

[u] curta, curto, turco												
	INÍCIO			MEIO			FIM			F1	F2	F3
	F1	F2	F3	F1	F2	F3	F1	F2	F3			
MÉDIA (Hz)	506	1186	3051	552	1073	3302	497	1337	2874			
DP	72,28	284,80	482,67	64,00	182,00	446,68	94,25	311,50	563,82			
[u] cuca												
	INÍCIO			MEIO			FIM			F1	F2	F3
	F1	F2	F3	F1	F2	F3	F1	F2	F3			
MÉDIA (Hz)	448	806	3413	432	847	3528	567	842	3347			
DP	50,41	164,94	239,43	42,26	13,42	134,27	155,76	99,70	681,17			
[uj] cuida												
	INÍCIO			MEIO			FIM			F1	F2	F3
	F1	F2	F3	F1	F2	F3	F1	F2	F3			
MÉDIA (Hz)	548	1629	3110	519	1486	3134	462	2546	3248			
DP	85,89	960,78	22,79	35,25	380,46	124,84	64,41	95,94	383,77			

Tabela 9 - Médias e desvio-padrão para os valores de F1, F2 e F3 nos trechos de início, meio e fim da trajetória formântica de [u], em contexto [ur], [u] e de [uj], nas produções de S0

De acordo com os resultados presentes na Tabela 9, é possível constatar que, para os alvos em contexto [ur], o terceiro formante apresenta uma queda, ao passo que os valores de F2 sobem, distanciando-se de F1. Quanto às médias em contexto de [u], nota-se que o primeiro formante exibiu uma suave subida, enquanto o segundo manteve-se na faixa dos 800 Hz. O terceiro formante apresenta uma sensível curva de subida e descida. No que tange o comportamento dos formantes para os alvos em contexto de UI, observa-se uma pequena elevação de F3 e queda de F1, ao passo que F2 exibe uma abrupta elevação. Assim, a análise do referido quadro corrobora o que havia sido apontado pela inspeção visual dos gráficos.

Ao observar os resultados oriundos da análise da trajetória formântica das produções do sujeito controle para a posição medial, percebe-se que, para a produção da vogal [a] que antecede o rótico, há uma pequena queda de F1 e um aumento nos valores de F2 e F3, sendo mais acentuado para o primeiro. Para a vogal [a], em contexto CV, os três formantes apresentam-se horizontalmente alinhados. A produção do ditongo [aj] exibe o distanciamento de F1 e F2, de acordo com o comportamento típico para esse tipo de segmento.

A trajetória formântica dos dados em contexto da vogal alta anterior revela que, quando o rótico é articulado, o valor de F1 sofre uma leve elevação, ao passo que F2

e F3 caem. Para a vogal [i], em contexto CV, F1 apresenta-se horizontalmente alinhado, à medida que F2 e F3 decrescem.

Por fim, a comparação das trajetórias formânticas em contexto da vogal alta posterior evidenciou que F2 se eleva, distanciando-se de F1, enquanto F3 exibe uma queda. No entanto, o comportamento dos formantes para a vogal [u], em contexto CV, exibe um padrão diferente. As três frequências apresentam-se horizontalmente alinhadas, à medida que há uma suave elevação na porção final da vogal nos valores de F1 e F3.

#### 4.2.1.4 Natureza do elemento vocálico

A análise da natureza do elemento vocálico adjacente à produção do *tap* permite investigar se esse elemento se comporta como uma vogal neutra ou se o segmento entrecorta a vogal nuclear. Verifica-se, então, se o elemento vocálico apresenta uma natureza formântica diferente da vogal nuclear da sílaba ou se o mesmo exibe as mesmas características da vogal que acompanha o *tap*.

Os alvos analisados para o ambiente vocálico de [a] foram *carpa* e *carta*, para o contexto da vogal [i], examinaram-se as palavras *circo* e *círculo*, e, para [u], foram considerados os itens lexicais *curta* e *curto*. Sendo assim, os elementos vocálicos presentes nas palavras *farpa* e *turco* não foram analisados, pois buscou-se manter o mesmo ponto de articulação das consoantes, para minimizar efeitos de coarticulações distintas. Dos 18 *tokens* disponíveis para análise, 5 foram articulados como vibrantes e 4 foram descartados por conta da falta de correspondência entre a posição do formante e o traçado LPC. Assim, a mostra contou com 9 *tokens*.

Dois tipos de análises foram feitos para examinar a natureza do elemento vocálico. O primeiro com base em Silva, Clemente e Nishida (2006), buscou extrair os valores de F1 e F2 da vogal núcleo e do elemento vocálico em seus respectivos pontos médios. O segundo, apresenta uma comparação entre a parte final de uma vogal nuclear em CV e o elemento vocálico, conforme já descrito no capítulo Metodologia da presente dissertação.

A fim de realizar a análise conforme Silva, Clemente e Nishida (2006), serão apresentadas as Tabelas 10 e 11 e o Gráfico 28. A Tabela 10 exibe os valores de F1 e F2 das vogais nucleares (VN) e dos elementos vocálicos (EV) presentes nos dados do sujeito controle, em posição medial.

VOGAL NUCLEAR				ELEMENTO VOCÁLICO			
		F1	F2			F1	F2
[a]	MÉDIA (Hz)	1116	1890	[a]	MÉDIA (Hz)	966	2096
	DP	94,40	226,63		DP	94,40	11,31
[i]	MÉDIA (Hz)	518	2358	[i]	MÉDIA (Hz)	519	2032
	DP	127,63	42,43		DP	46,32	4,95
[u]	MÉDIA (Hz)	553	931	[u]	MÉDIA (Hz)	541	1738
	DP	66,47	15,56		DP	41,72	199,40

Tabela 10 - Médias gerais de F1 e F2 do ponto médio da vogal nuclear e do elemento vocálico, em contexto de [a], [i] e [u], produzidos por S0

Os dados apresentados na Tabela 10 foram utilizados para gerar o Gráfico 28. A partir da plotagem dos valores de F1 e F2, pode-se compreender como a vogal nuclear e o elemento vocálico se comportam no espaço acústico.

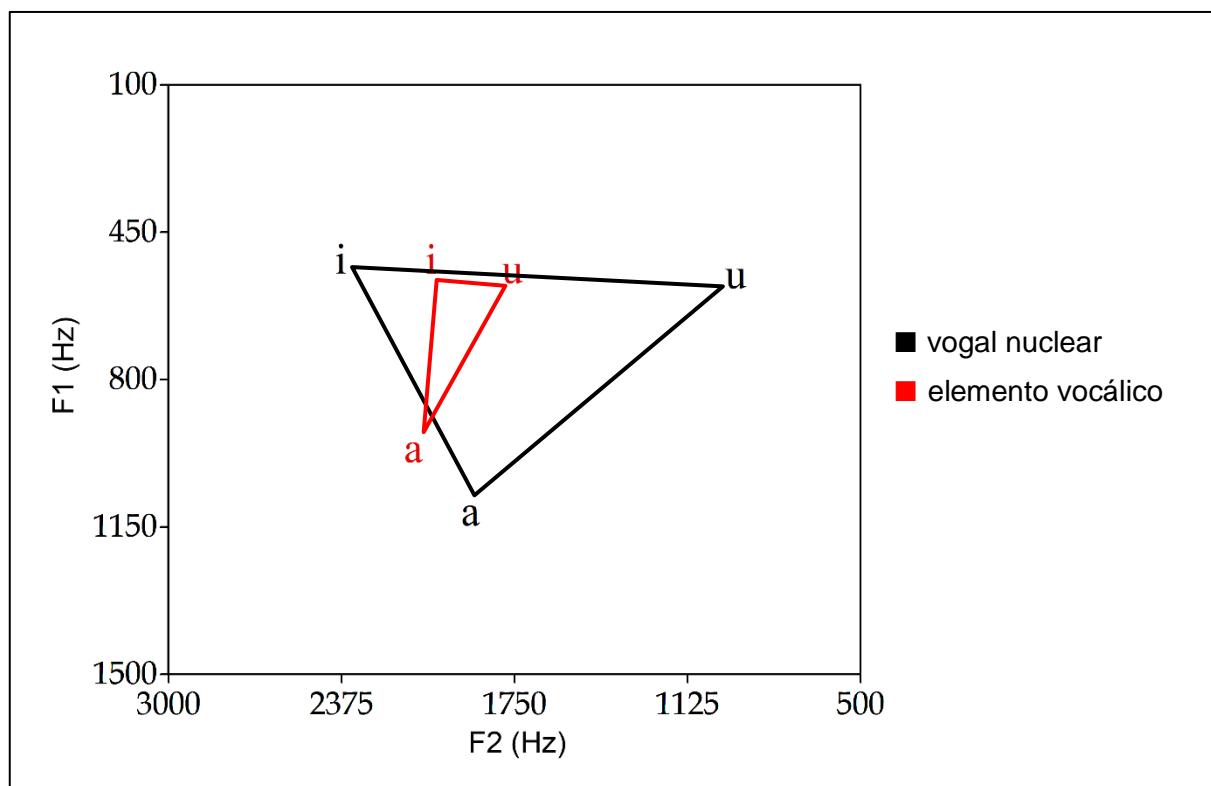


Gráfico 28 - Espaço acústico de [a], [i] e [u] para a vogal nuclear (preto) e o elemento vocálico (vermelho) produzidos por S0

Ao analisar a distribuição dos elementos vocálicos no espaço acústico, a partir do Gráfico 28, percebe-se que esses apresentam-se de forma mais centralizada, compondo um triângulo vocálico mais fechado, com configuração menos regular. Os elementos vocálicos que ocorrem em contexto de vogais altas exibem valores de F2

mais próximos a uma vogal central, apresentando pouca distinção entre si no que concerne aos movimentos de anterioridade e posterioridade da língua<sup>22</sup>. A posição da vogal alta posterior é a mais distante da vogal núcleo. Salienta-se, no entanto, que os elementos vocálicos, apesar de pouco se distinguirem em relação a F2, comportam-se como suas respectivas vogais nucleares em relação a F1.

Silva, Clemente e Nishida (2006) adotaram um procedimento para auxiliar a comparação entre a vogal núcleo da sílaba e o elemento vocálico. Esse consiste em subtrair as médias das frequências de EV dos valores das médias de VN. Considerem-se, pois, os resultados da aplicação do referido procedimento na Tabela 11:

DIFERENÇA ENTRE A VOGAL NUCLEAR E O ELEMENTO VOCÁLICO						
	F1 VN (Hz)	F1 EV (Hz)	Diferença (Hz)	F2 VN (Hz)	F2 EV (Hz)	Diferença (Hz)
[a]	1116	966	150	1890	2096	-206
[i]	518	519	-1	2358	2032	326
[u]	553	541	13	931	1738	-807

Tabela 11 - Diferenças dos valores médios de F1 e F2 entre as vogais nucleares (VN) e os elementos vocálicos (EV), produzidos por S0

A partir dos resultados apresentados na Tabela 11, é possível verificar que as maiores diferenças se concentram no âmbito de F2. Nesse caso, percebe-se que os elementos vocálicos apresentam maiores diferenças em F2, pois esses encaminham-se para a direção central do espaço acústico. O elemento vocálico em contexto de [u] é o que apresenta maior diferença, apresentando maior deslocamento em relação à vogal núcleo.

O segundo modo de análise compara os valores formânticos do ponto médio da porção final de uma vogal nuclear em sílaba CV e do elemento vocálico. Para tanto, serão apresentadas as Tabelas 12 e 13, e os Gráficos 29 e 30.

Nesse sentido, foram escolhidas três vogais [a, i, u], produzidas em sílaba CV, para serem comparadas, em termos de duração e valores de frequências, com cada elemento vocálico analisado. Desse modo, com base no valor de duração de cada EV, foi calculado quanto, em termos de proporção, ele ocupava da vogal CV, conforme

<sup>22</sup> De acordo com Kent e Read (2015), os valores de frequências do segundo formante inferem o movimento do eixo horizontal da língua.

já reportado na Metodologia. A Tabela 12 apresenta os dados para a extração do ponto médio da vogal nuclear.

PROPORÇÃO DO ELEMENTO VOCÁLICO				
	Duração CV (ms)	Duração EV (ms)	Proporção de EV (%)	Porção final para extração do PM de VN (ms)
CARPA 1	115	23	20	23
CARTA 1	115	38	33	38
CARTA 2	115	25	22	25
CIRCO 1	100	28	28	28
CIRCO 2	100	41	41	41
CIRCULO 1	100	31	31	31
CIRCULO 2	100	38	38	38
CURTA 3	125	16	13	16
CURTO 3	125	19	15	19

Tabela 12 – Proporção do elemento vocálico em relação à vogal nuclear CV, por S0

Além da proporção de cada elemento vocálico, é possível observar, a partir da Tabela 12, que a duração absoluta varia entre 16 ms e 41 ms. O EV produzido em contexto de [u] apresenta os menores índices de proporção, 13% e 15%, ao passo que o ambiente vocálico de [i] proporcionou índices maiores, entre 28% e 41%.

A Tabela 13 exibe os valores de F1 e F2, extraídos proporcionalmente, do ponto médio das vogais nucleares (VN) e dos elementos vocálicos (EV) presentes nos dados do sujeito controle, em posição medial.

ELEMENTO VOCÁLICO						
		F1 (Hz)	F2 (Hz)		F1 (Hz)	F2 (Hz)
a <sub>1</sub>	CATA	1118	1728	CARPA 1	1033	2104
a <sub>2</sub>	CATA	1164	1768	CARTA 1	1000	2259
a <sub>3</sub>	CATA	1126	1732	CARTA 2	799	1917
i <sub>1</sub>	CHICO	505	1867	CIRCO 1	299	2007
i <sub>2</sub>	CHICO	514	1912	CIRCO 2	804	2050
i <sub>3</sub>	CHICO	516	1868	CIRCULO 1	343	1934
i <sub>4</sub>	CHICO	515	1899	CIRCULO 2	629	2137
u <sub>1</sub>	CUCA	676	871	CURTA 3	511	1879
u <sub>2</sub>	CUCA	652	850	CURTO 3	570	1597

Tabela 13 – Ponto médio de F1 e F2 das vogais protótipos (extraídos da porção final) e dos elementos vocálicos produzidos por S0

Os dados exibidos na Tabela 13 foram utilizados para gerar os Gráficos 29 e 30. A partir da plotagem dos valores de F1 e F2, pode-se compreender como os elementos vocálicos se comportam no espaço acústico, tendo em vista sua comparação com a parte final da vogal CV.

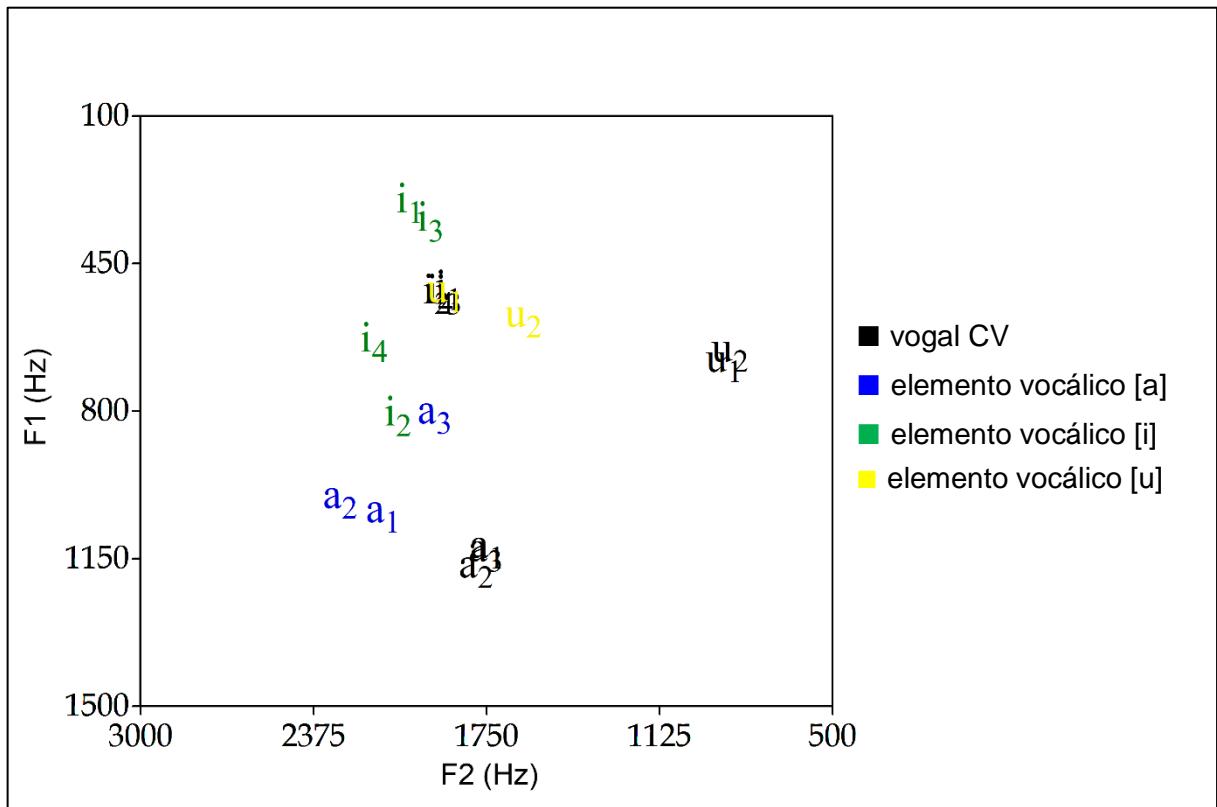


Gráfico 29 - Espaço acústico de [a], [i] e [u] para o ponto médio da porção final vogal CV (preto) e elemento vocálico produzidos por S0 – contexto de [a] (azul), [i] (verde) e [u] (amarelo)

Os resultados presentes no Gráfico 29 evidenciam que os elementos vocálicos apresentam certa variação em sua distribuição, especialmente os EVs de [i], no que concerne a F1. Em relação à comparação entre os formantes dos elementos vocálicos e os formantes da porção final da sílaba CV, observa-se, para o contexto de [a], uma anteriorização em relação à VN. Já o contexto de [i] apresenta uma maior diferença em relação à altura, com similaridade entre EV e VN no que se refere a F2. O contexto de [u] também indica um elemento vocálico um pouco mais anterior e outro mais centralizado.

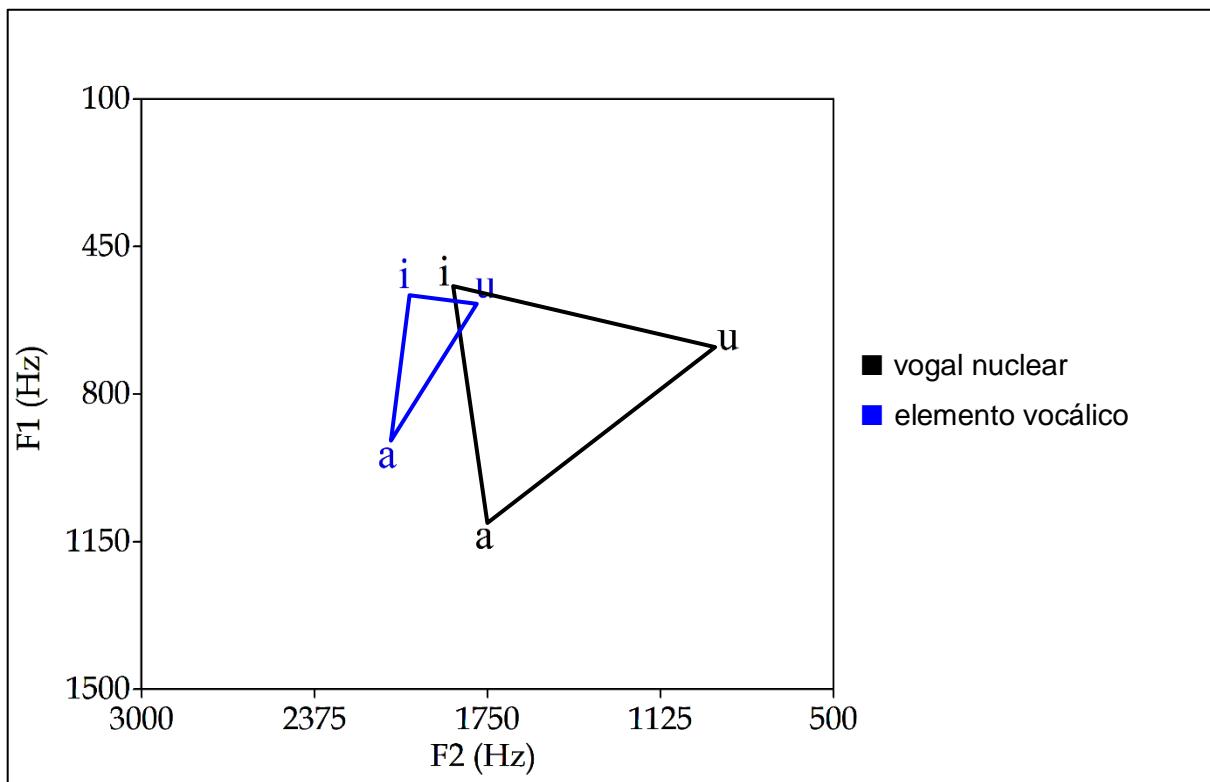


Gráfico 30 - Espaço acústico de [a], [i] e [u] para o ponto médio da porção final vogal CV (preto) e para a média do ponto médio do elemento vocálico produzidos por S0 – contexto de [a] (azul), [i] (verde) e [u] (amarelo)

O Gráfico 30, que apresenta uma comparação entre as médias do ponto médio de EV e do ponto médio da porção final de VN, deixa mais evidente a anteriorização de [a] e a centralização de [u]. Sendo assim, a produção de EV como um elemento independente da vogal nuclear parece pouco provável. Observe-se, mais uma vez, que ficam preservadas as diferenças de altura entre os três tipos de EV produzidos, em acordo com a natureza vocálica das vogais nucleares.

Retoma-se, assim, as duas possibilidades de natureza do elemento vocálico produzido adjacente ao *tap* em final de sílaba, ou seja, (i) a de uma vogal neutra, com comportamento distinto da vogal que antecede o *tap*, e (ii) a de um elemento vocálico que entrecorta o rótico, apresentando as mesmas características da vogal nuclear da sílaba. A primeira análise feita neste estudo, cujos resultados estão expostos no Gráfico 28, converge, apenas em relação a F2, com os resultados de Silva, Clemente e Nishida (2006), segundo o qual afirma que EV, quando produzido em final de sílaba, exibe características acústicas mais próximas de uma vogal central. Destaca-se que os elementos vocálicos convergem com a vogal nuclear da sílaba no que concerne aos padrões de F1. Na segunda análise, com base nos Gráficos 29 e 30, a centralização parece ocorrer apenas em ambiente vocálico de [u]. Na verdade, todos

os elementos vocálicos produzidos são um pouco mais anteriores do que suas respectivas vogais centrais, sendo que, a anteriorização para [u], acaba posicionando-o em uma região mais central do espaço acústico. Sendo assim, os resultados parecem indicar a existência de uma relação mais estreita entre o elemento vocálico e a vogal nuclear da sílaba CVC.

#### 4.2.2 Sujeito longitudinal

A informante longitudinal dispõe de três gravações de áudio. Após o descarte dos dados que não apresentavam qualidade suficiente para uma análise acústica confiável, 216 produções compõem a amostra de fala da criança acompanhada longitudinalmente.

Para melhor organização do texto, as próximas subseções – 4.2.1.1 a 4.2.1.3 – discutem a análise dos dados referente à inspeção visual do espectrograma, aspectos duracionais e trajetória formântica.

##### 4.2.2.1 Inspeção visual do espectrograma

A análise visual do espectrograma permite a observação de detalhes fonéticos que não são percebidos por meio de uma análise de oitiva. A partir das três características elencadas por Silva *et al* (2019), (i) ausência de energia; (ii) ruído transiente ou soltura da oclusão e (iii) presença de elemento vocálico, foi gerada a Tabela 14, contendo as possibilidades e ocorrências de características do *tap* na forma de onda e no espectrograma. A presença de tais características fornece, assim, pistas acerca da existência de prováveis contrastes encobertos observados na fala da informante longitudinal em CVC medial.

POSSIBILIDADES E OCORRÊNCIAS DE CONTRASTES ENCOBERTOS								
1 <sup>a</sup> COLETA								
	FORMA DE ONDA			ESPECTROGRAMA			Contraste	
	(i)	(ii)	(iii)	(i)	(ii)	(iii)		
	A	1/8	0/8	0/8	2/8	0/8	0/8	2/8
	I	3/6	0/6	4/6	4/6	3/6	4/6	4/6
	U	2/6	0/6	1/6	3/6	2/6	1/6	3/6
	TOTAL	6/20	0/20	5/20	9/20	5/20	5/20	9/20
2 <sup>a</sup> COLETA								
	A	3/8	0/8	2/8	3/8	0/8	2/8	3/8
	I	4/5	2/5	4/5	5/5	4/5	3/5	5/5
	U	5/7	2/7	3/7	5/7	2/7	3/7	5/7
	TOTAL	12/20	4/20	9/20	13/20	6/20	8/20	13/20
	3 <sup>a</sup> COLETA							
	A	0/9	0/9	0/9	0/9	0/9	0/9	0/9
	I	4/6	0/6	1/6	4/6	0/6	1/6	4/6
	U	8/9	1/9	4/9	8/9	1/9	4/9	8/9
	TOTAL	12/24	1/24	5/24	12/24	1/24	5/24	12/24

Tabela 14 – Possibilidades e ocorrências de contrastes encobertos a partir das características acústicas do *tap*, sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala da informante longitudinal, em CVC medial

Ao analisar a Tabela 14, nota-se que a ausência de energia foi a característica do *tap* que mais evidenciou a presença de possíveis contrastes encobertos em posição medial, sendo o espectrograma a forma que facilitou a sua observação. A presença de elementos vocálicos foi a segunda característica que mais corroborou para a observação de tais contrastes. Observa-se, também, que, da primeira para a segunda coleta, houve um aumento da presença de contrastes para todos os contextos vocálicos.

A Tabela 15, apresenta as possibilidades e ocorrências de características do *tap* na forma de onda e no espectrograma para a posição final. Na referida tabela, constam os prováveis contrastes encobertos observados na fala da informante longitudinal.

POSSIBILIDADES E OCORRÊNCIAS DE CONTRASTES ENCOBERTOS							
1ª COLETA							
	FORMA DE ONDA			ESPECTROGRAMA			Contraste
	(i)	(ii)	(iii)	(i)	(ii)	(iii)	
	A	0/9	0/9	0/9	0/9	0/9	0/9
	I	3/9	0/9	3/9	3/9	0/9	3/9
TOTAL	A	0/9	0/9	0/9	0/9	0/9	0/9
	TOTAL	0/27	0/27	0/27	0/27	0/27	3/27
2ª COLETA							
	A	1/8	1/8	1/8	1/8	1/8	1/8
	I	4/8	0/5	4/8	4/8	0/8	4/8
	U	5/9	0/9	3/9	5/9	0/9	5/9
	TOTAL	11/25	1/25	9/25	11/25	1/25	9/25
3ª COLETA							
	A	1/8	1/8	1/8	1/8	1/8	1/8
	I	2/9	0/9	2/9	2/9	0/9	2/9
	U	1/9	0/9	1/9	1/9	0/9	1/9
	TOTAL	4/26	1/26	4/26	4/26	1/26	4/26

Tabela 15 – Possibilidades e ocorrências de contrastes encobertos a partir das características acústicas do *tap*, sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala da informante longitudinal, para CVC final

A análise da Tabela 15 revela que, assim como para posição medial, a ausência de energia foi a característica do *tap* que mais apontou a presença de contrastes encobertos, seguida da presença de elementos vocálicos. O espectrograma e a forma de onda contribuíram igualmente para a observação de tal característica. Igualmente, constata-se que, na segunda coleta, a presença de pistas acústicas relativas aos contrastes é mais recorrente.

As características mais relevantes para a constatação de contrastes encobertos foram, portanto, a ausência de energia e a presença de elementos vocálicos, a primeira mais recorrente que a segunda. Sendo assim, infere-se que a dificuldade da informante pode estar relacionada à regulação da magnitude do gesto de ponta de língua, pois ocorre o breve momento de obstrução, caracterizado, no sinal acústico, como um espaçamento em branco no espectrograma ou pela baixa amplitude da onda sonora. O baixo índice de correlatos visuais de ruído transiente ou soltura da oclusão pode evidenciar uma imprecisão durante a execução do afastamento dos articuladores. Sendo assim, a realização imprecisa dos movimentos articulatórios tornaria o segmento inaudível.

Com o propósito de sintetizar as informações anteriores, foi gerada a Tabela 16, contendo as possibilidades e ocorrências de prováveis contrastes encobertos de acordo com o ambiente vocálico e com a posição do rótico na sílaba.

POSSIBILIDADES E OCORRÊNCIAS DE CONTRASTES ENCOBERTOS						
	MEDIAL			FINAL		
	[a]	[i]	[u]	[a]	[i]	[u]
<b>1<sup>a</sup> COLETA</b>	2/8	4/6	3/6	0/9	3/9	0/9
<b>2<sup>a</sup> COLETA</b>	3/8	5/5	5/7	1/8	4/8	5/9
<b>3<sup>a</sup> COLETA</b>	0/9	4/6	8/9	1/8	2/9	1/9
<b>TOTAL</b>	<b>5/25</b>	<b>13/17</b>	<b>16/22</b>	<b>2/25</b>	<b>9/26</b>	<b>6/27</b>
		<b>34/64</b>			<b>17/78</b>	

Tabela 16 – Possibilidades e ocorrências de prováveis contrastes encobertos na fala da informante longitudinal, de acordo com a posição silábica e o ambiente vocálico em posições medial e final

A análise da Tabela 16 permite identificar que a posição medial foi o contexto silábico que facilitou a ocorrência de prováveis contrastes encobertos. Em 64 produções, foram identificadas 34 possibilidades de produções gradientes, o que representa cerca de 53,13% das tentativas de articulação do rótico. Para a posição final, foram 17 prováveis contrastes percebidos, em um total de 78 produções, ou seja, 21,79%.

Embora a análise de oitiva tenha detectado um alto índice de omissão do *tap* em posição medial, com o registro de 83%, e uma baixa aplicação de recursos outros, como semivocalização e realização de um segmento lateral, há, justamente nessa posição, a maior presença de contrastes encobertos. Desta forma, a posição medial não parece constituir um ambiente desfavorecedor para a emergência do *tap*.

Na sequência, serão expostos alguns dados acústicos da informante longitudinal a fim de que se possa inferir o modo como alguns alvos foram articulados. Serão apresentados exemplos para cada coleta.

Em relação à primeira coleta, a análise de oitiva não constatou a produção do segmento, o que indica, portanto, a presença de possíveis contrastes encobertos. No entanto, os resultados obtidos apresentaram apenas três pistas visuais de produções gradientes do *tap* em posição final, conforme apresentado na Tabela 16. Assim, ao levar em conta o contexto silábico, é possível observar que a posição medial foi o contexto que facilitou a ocorrência do provável fenômeno para os casos de omissão.

Salienta-se que, tanto em CVC medial quanto final, foi possível visualizar a produção de outro segmento na posição em que o *tap* deveria figurar. Para ilustrar a inspeção visual dos dados da primeira coleta de SL1, segue a Figuras 23.

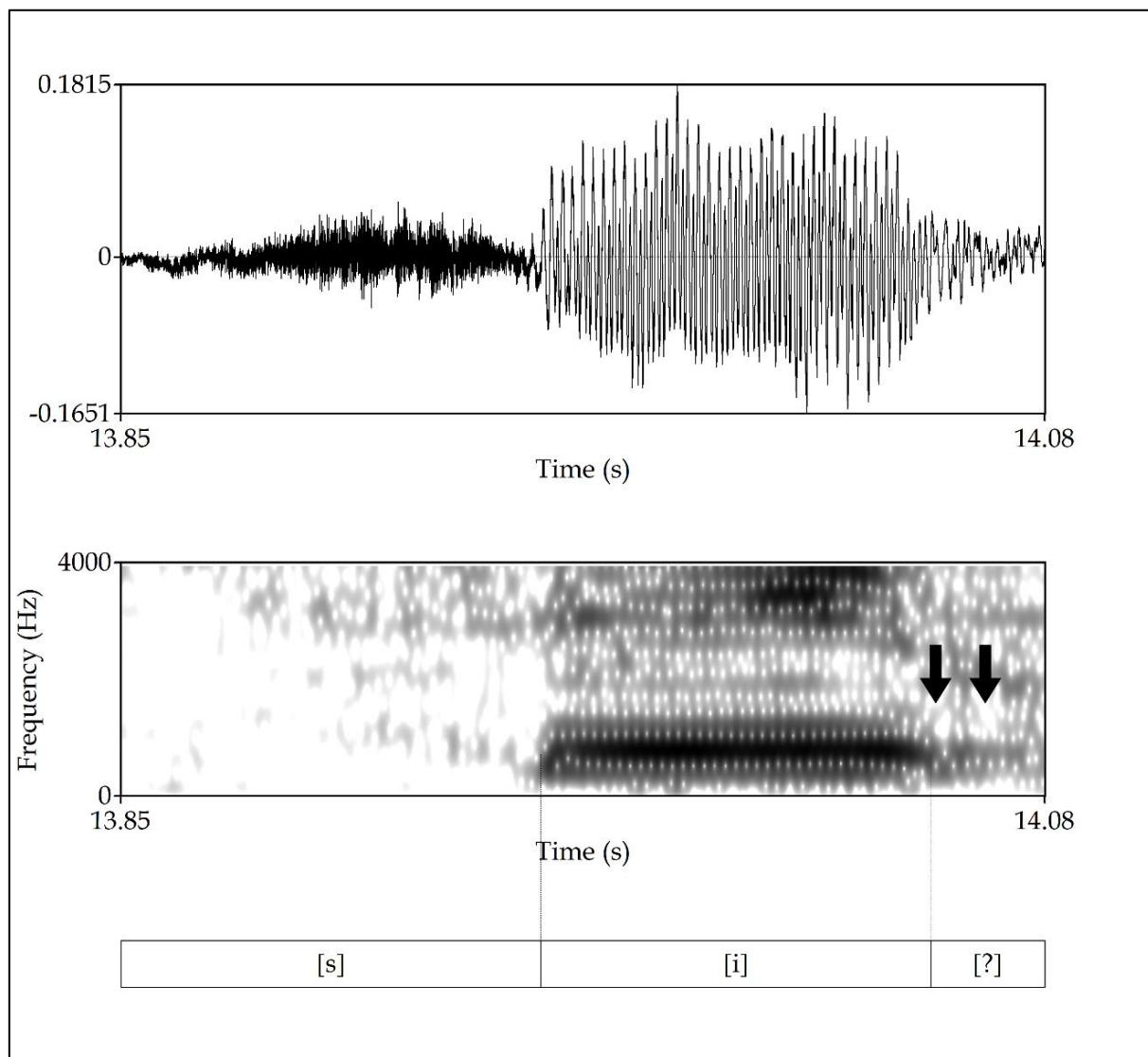


Figura 23 - Forma de onda e espectrograma de banda larga da produção de ['sɪ] do alvo "círculo" produzido na primeira coleta por SL1. As setas indicam a presença do contraste encoberto.

Embora o *tap* não seja visível no espectrograma, para a produção do alvo *círculo*, foram encontrados dois correlatos acústicos que podem evidenciar traços residuais de uma variante rótica, a vibrante<sup>23</sup>. Na Figura 23, percebe-se, na forma de onda, uma queda da amplitude ao final da vogal [i]. As setas apontam, no

<sup>23</sup> O trabalho de Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2021) observa, em posição medial de palavra, a produção de vibrantes alveolares em sujeitos adultos, falantes do mesmo dialeto que a informante em questão.

espectrograma, o momento de queda abrupta da amplitude. Embora essas qualidades acústicas indiciem a produção gradiente de uma possível vibrante, esse alvo não foi percebido com a articulação do segmento na análise de oitiva.

Já na segunda coleta, as pistas visuais apontam uma maior ocorrência de produções gradientes em posição medial de sílaba. Todas as produções que abarcavam a vogal alta anterior apresentaram possíveis contrastes encobertos no sinal acústico. Apesar disso, o julgamento de oitiva não detectou a produção do *tap*. Para exibir a inspeção visual realizada nos dados da segunda coleta de SL1, apresentam-se as Figuras 24 e 25.

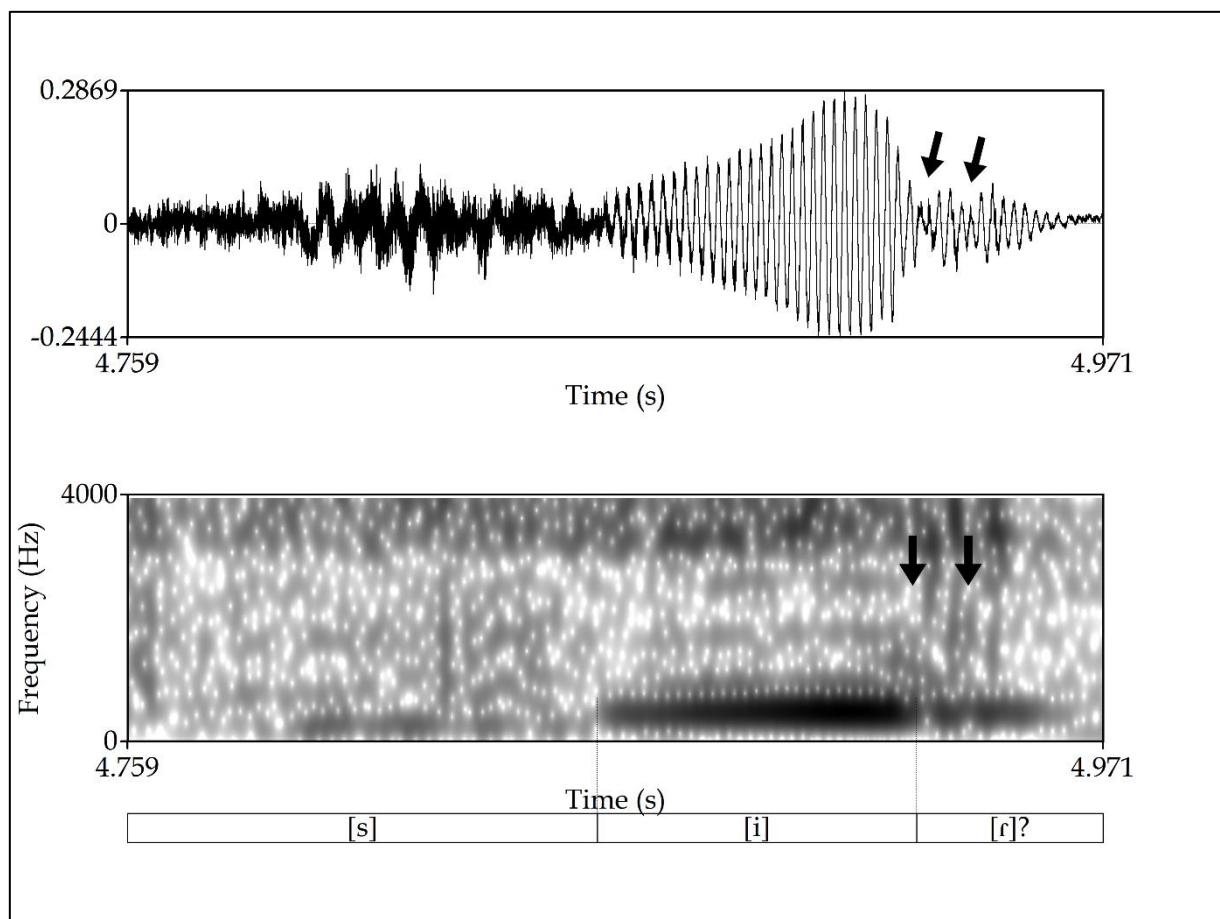


Figura 24 – Forma de onda e spectrograma de banda larga da produção de [si], do alvo "circo" produzido na segunda coleta por SL1. As setas indicam a presença do contraste encoberto.

Ao observar a forma de onda e o spectrograma do alvo *circo*, notam-se pistas visuais da produção de uma vibrante. Na Figura 24, as setas, na forma de onda, indicam a queda de amplitude e, na parte inferior da figura, percebem-se os espaçamentos em branco no spectrograma. Essas características, chamadas de períodos de vibração, revelam que a informante parece tentar articular um segmento

que aproximasse de seu alvo. No entanto, a análise de oitiva não constatou a produção do segmento.

Ainda para a segunda coleta, um outro exemplo é a produção do alvo *vir*, conforme disposto na Figura 25.

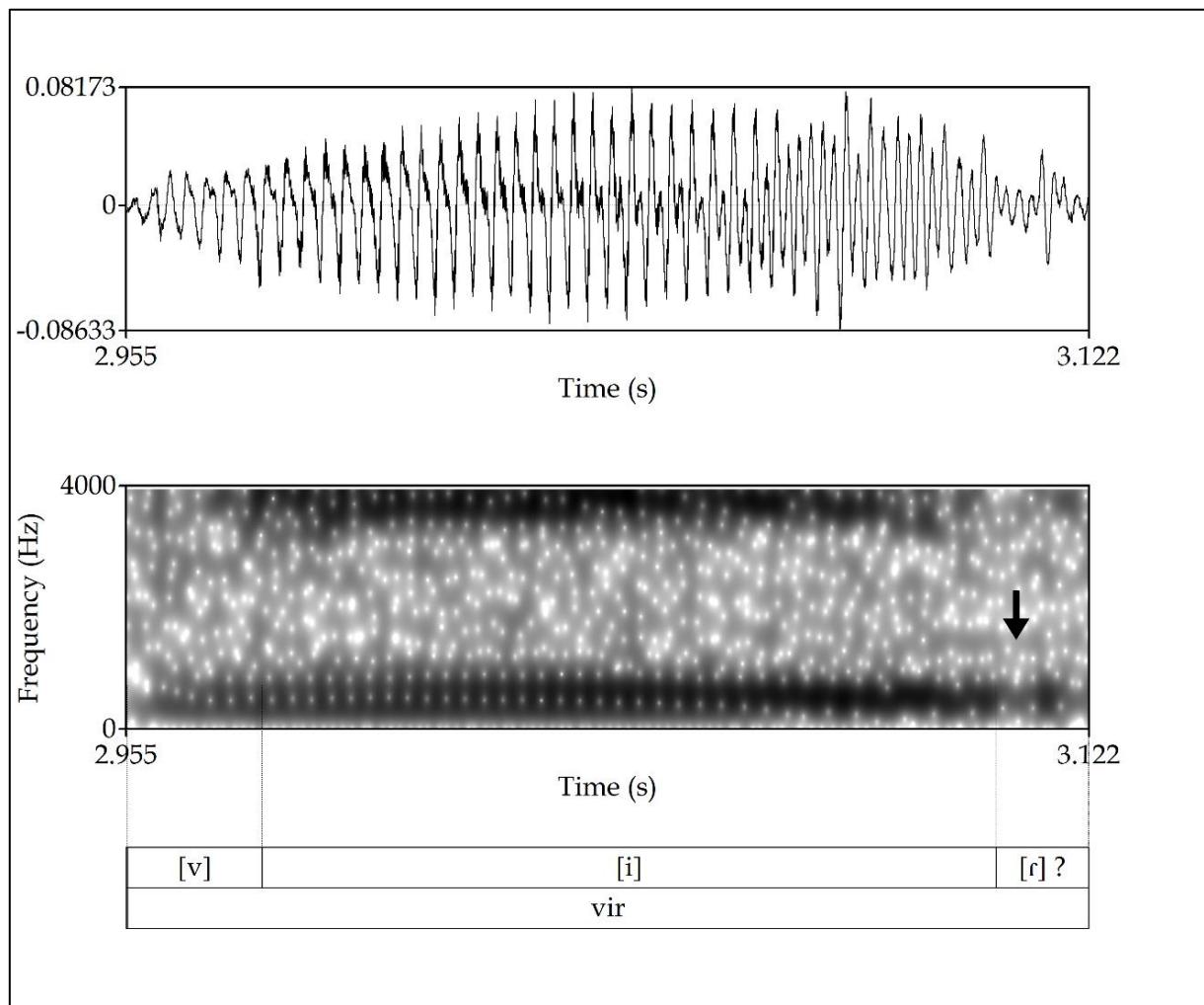


Figura 25 - Forma de onda e espectrograma de banda larga da produção de [vi], do alvo "vir" produzido na segunda coleta por SL1. A seta indica a presença do contraste encoberto.

A análise de oitiva para o alvo *vir* não constatou a presença do *tap*, decisão unanime entre os três juízes. Apesar disso, o espectrograma apresentou uma variação do sinal acústico ao final da vogal [i]. Na Figura 25, é possível observar uma forte queda de amplitude, sinalizado pela seta, e, após, algo como se fosse a porção final da vogal. Essas seriam duas características acústicas da produção do segmento. Seria possível inferir que a informante tenta fazer o movimento articulatório de ponta de língua, contudo, esse não apresentaria magnitude suficiente para ser percebido auditivamente.

A inspeção visual realizada para a terceira coleta de SL1 indica que as produções gradientes foram predominantes na posição medial, sendo a vogal alta posterior o contexto facilitador para a ocorrência do fenômeno. Casos de semivocalização também foram encontrados de forma expressiva. Para ilustrar a análise visual dos dados da terceira coleta da informante longitudinal, um exemplo de semivocalização, disposto na Figura 26.

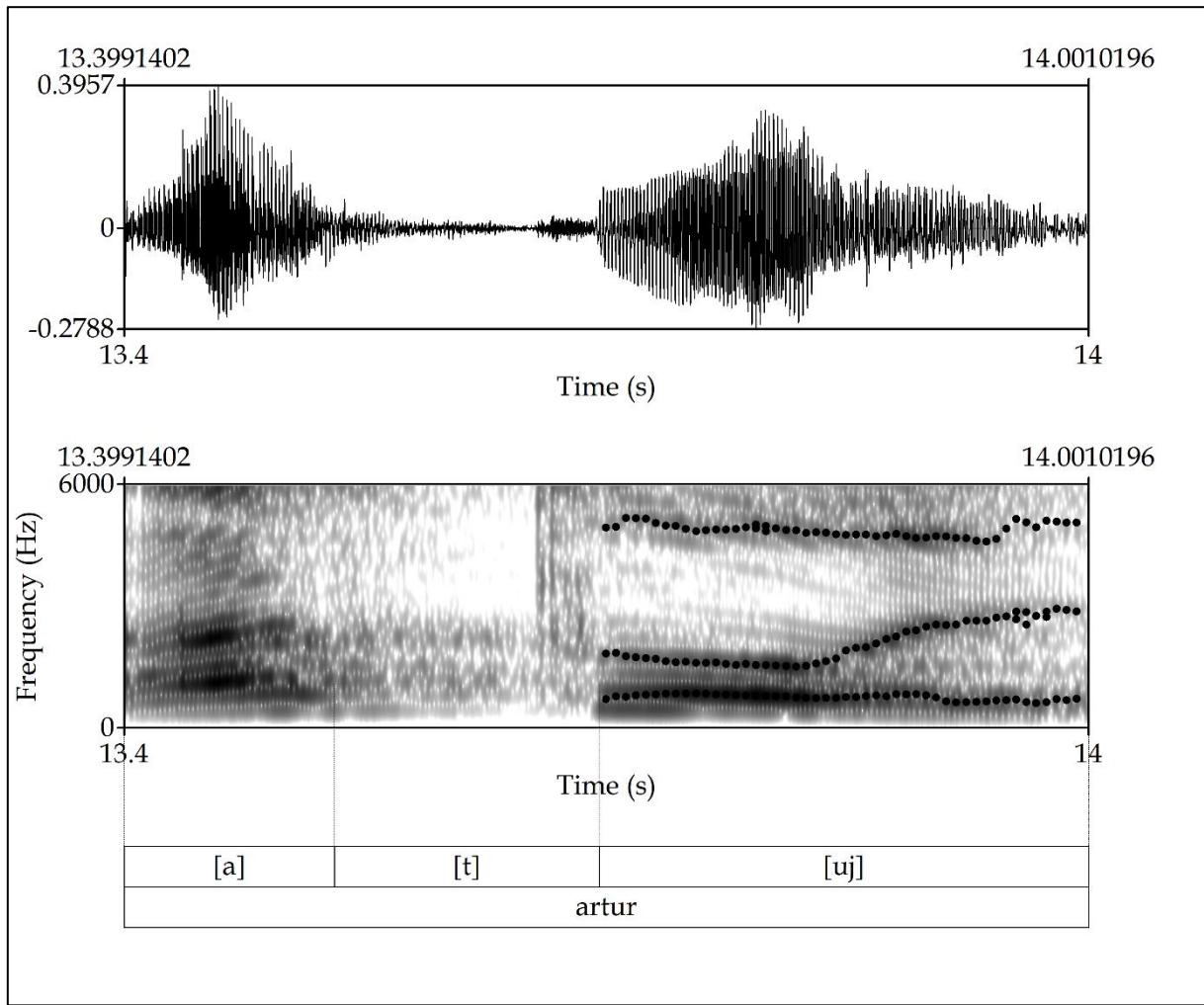


Figura 26 - Forma de onda e espectrograma de banda larga da produção de [‘a.tu] do alvo "Artur" produzido por SL1 na terceira coleta.

A análise visual do espectrograma gerado para a palavra *Artur* indica a produção da semivogal [j] no final da palavra, lugar este que deveria ser ocupado pelo [r]. Após a produção da vogal alta posterior, é possível observar a transição formântica para a ocorrência do ditongo. Há uma elevação contínua de F2 e uma leve queda de F1, configuração formântica que indica a produção do glide palatal [j]. Para a vogal [u], as frequências de F1 e F2 são de 827 Hz e 1586 Hz, respectivamente. O glide [j]

apresenta 624 Hz para F1 e 2526 Hz para F2. Desse modo, a percepção inferida por meio da análise de oitiva confirmou-se. Ocorreu a semivocalização do *tap* no alvo *Artur*.

A inspeção visual dos espectrogramas, gerados para os dados da informante longitudinal, aponta, portanto, que, apesar de o *tap* não ter sido produzido em boa parte dos dados – já que está presente em algumas produções decorrentes de metáteses, como ['tu.ku.ru] e ['si.ku.lo] –, produções gradientes foram detectadas. Assim, as pistas acústicas permitem inferir que talvez a informante produza um rótico mais enfraquecido, devido a uma redução na magnitude gestual. Os casos em que mais de um espaçamento em branco foram identificados no espectro demostram uma tentativa de articulação do som conforme o alvo, ou seja, uma vibrante, especialmente em posição final, ainda que seja necessário, por parte da informante, refinar os gestos articulatórios para que o segmento se torne audível.

#### 4.2.2.2 Aspectos duracionais

A análise dos valores de duração tem como propósito determinar possíveis diferenças e semelhanças na duração da vogal nuclear entre os pares de palavras que foram produzidas pelas crianças, como carta x cata e mar x má, tendo em vista que a informante longitudinal não produziu o segmento alvo ao longo das três coletas.

Desse modo, busca-se obter respostas sobre fenômenos de produções gradientes encontrados nas análises visuais dos espectrogramas, como o alongamento compensatório, já reportado em Mezzomo (2004). Para melhor organização textual, essa subseção inicialmente reporta os resultados obtidos para a posição medial, efetuando algumas comparações entre os pares de palavras e com os dados do sujeito controle. Em seguida, o mesmo dar-se-á para os resultados referentes à posição final.

Com base nas referidas considerações, a Tabela 17 apresenta as médias de duração absoluta da sílaba, do rótico e da vogal nuclear das palavras que, em posição medial, abarcavam o rótico como alvo – carpa, farpa, carpa, circo, círculo, curta, curto, turco. Na mesma tabela, são exibidos os valores de duração relativa da vogal em relação à sílaba. Os valores reportados a seguir são referentes às produções nas quais a informante utiliza o recurso de omissão do *tap*.

CODA MEDIAL					
		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(vogal - sílaba)
AR	Coleta 1	174	0	133	76,77
	Coleta 2	156	0	119	76,29
	Coleta 3	182	0	130	73,88
	<b>MÉDIA</b>	<b>171</b>	-	<b>128</b>	<b>75,65</b>
	DP	13,31	-	7,36	1,55
IR	Coleta 1	208	0	92	43,94
	Coleta 2	217	0	74	32,77
	Coleta 3	216	0	69	31,41
	<b>MÉDIA</b>	<b>214</b>	-	<b>78</b>	<b>36,04</b>
	DP	4,85	-	11,89	6,87
UR	Coleta 1	136	0	110	80,83
	Coleta 2	132	0	102	76,61
	Coleta 3	134	0	97	71,95
	<b>MÉDIA</b>	<b>134</b>	-	<b>103</b>	<b>76,46</b>
	DP	2,28	-	6,60	4,44

Tabela 17 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos, com o recurso de omissão, pela informante longitudinal, conforme contexto vocálico

Conforme a Tabela 17, os itens lexicais que apresentam a maior variabilidade em relação a duração absoluta da vogal são relativos à vogal [i], que apresenta valores entre 69 e 92 ms, que pode ser verificado com o maior valor de desvio padrão reportado. As vogais que mais ocuparam a sílaba foram [u] e [a], com valores médios de duração relativa de 76,46% e 75,65%. A vogal alta anterior foi a que menos ocupou a sílaba, 36,04%. Essa diferença nos valores de duração relativa deve-se, segundo Silva *et al.* (2019), à duração intrínseca da vogal, considerando suas características articulatórias. A articulação de vogais abertas exige maior abertura da mandíbula, assim, esse movimento influencia diretamente em maiores valores de duração para [a]. No entanto, vogais fechadas ou altas são articuladas com a mandíbula praticamente fechada, desse modo, valores menores de duração são apresentados para esses tipos de segmentos vocálicos. Destaca-se, no entanto, que, o fato de [i] apresentar os menores valores de duração relativa também está relacionado ao fato de que as sílabas CVC de circo e círculo serem compostas por fricativas, ao contrário de curto e turco, que apresentam sílabas CVC constituídas por plosivas.

Com o objetivo de traçar uma comparação entre a duração das vogais realizadas nas produções que continham omissões de [r] e a duração das vogais

produzidas em itens lexicais com o molde silábico CV, a Tabela 18 apresenta as médias de duração absoluta da sílaba e da vogal nuclear das palavras que, em posição medial, incluíam a vogal em alvos CV – cata, Chico, cuca. Inclui também os valores de duração relativa da vogal em relação à sílaba.

		CODA MEDIAL	
		Duração absoluta (ms)	Duração relativa (%)
		Sílaba	Vogal
A	Coleta 1	185	134
	Coleta 2	149	99
	Coleta 3	137	106
	<b>MÉDIA</b>	<b>157</b>	<b>113</b>
	<b>DP</b>	<b>8,25</b>	<b>18,65</b>
I	Coleta 1	272	112
	Coleta 2	283	102
	Coleta 3	220	82
	<b>MÉDIA</b>	<b>258</b>	<b>99</b>
	<b>DP</b>	<b>44,19</b>	<b>15,09</b>
U	Coleta 1	170	142
	Coleta 2	224	175
	Coleta 3	144	98
	<b>MÉDIA</b>	<b>179</b>	<b>138</b>
	<b>DP</b>	<b>55,98</b>	<b>38,54</b>

Tabela 18 - Médias de duração absoluta e relativa dos alvos CV em posição medial da informante longitudinal, conforme contexto vocálico

Com base nos dados presentes na Tabela 18, observa-se que a vogal alta posterior é a que compreende maior variação nos valores de duração, apresentando valores médios entre 98 ms e 175 ms ao longo das três coletas. Ao examinar os valores de duração da sílaba, foi possível constatar grande oscilação nos valores referentes ao contexto da mesma vogal, pois, para [u], a sílaba apresentou valores entre 144 ms e 224 ms. Tais fatos são sinalizados pelos maiores valores de desvio padrão apresentados por essa vogal. Os dados referentes à duração relativa da vogal em relação à sílaba apontam que a vogal [i] ocupa 39,38% da sílaba. Já as vogais [a] e [u] representam 72,78% e 75,40% da sílaba, respectivamente.

Com o propósito de investigar a possibilidade de que a informante, apesar de não produzir o [r], fosse capaz de utilizar algum recurso para marcar a distinção entre as palavras que apresentavam o *tap* como alvo e as que não continham a presença do segmento, serão apresentados os Gráficos 30, 31 e 32. Esses efetuam algumas

comparações entre as produções da informante em diferentes contextos, CVC e CV, ao passo que também realizam comparações com os dados do sujeito controle.

Tendo por base as considerações apresentadas, o Gráfico 30 reúne os resultados das médias de duração absoluta da sílaba e da vogal das produções das palavras realizadas por SL1. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, carpa; (IR) circo, círculo; (UR) curto, curta, turco; (A) cata; (I) chico; (U) cuca. Os valores reportados para as palavras compostas por [r] – AR, IR, UR – são referentes às produções nas quais a informante utiliza o recurso de omissão do *tap*.

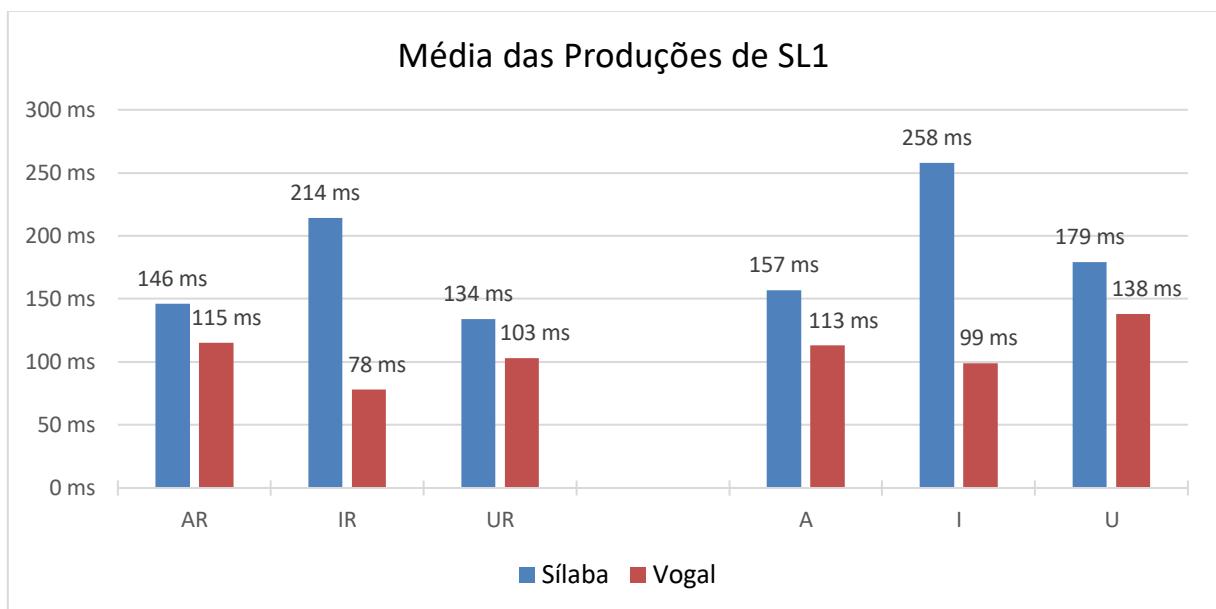


Gráfico 31 – Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1, para a posição medial

A análise do Gráfico 31 permite observar que existem algumas diferenças em relação à duração do segmento vocálico quando comparamos os alvos CVC e CV, sendo que apenas a vogal [a] apresenta valores aproximados. As vogais [i] e [u], em alvos CVC, apresentam menor duração do que seus pares em posição CV. Tal fato parece indicar que a estratégia de alongamento compensatório, constatada por Mezzomo (2004) para alguns segmentos em coda, não é utilizada por SL1.

O Gráfico 32 reúne os resultados das médias de duração relativa da vogal em relação à sílaba das produções realizadas por SL1. Os valores reportados para as palavras compostas por [r] – AR, IR, UR – igualmente são referentes às produções nas quais a informante utiliza o recurso de omissão do *tap*.

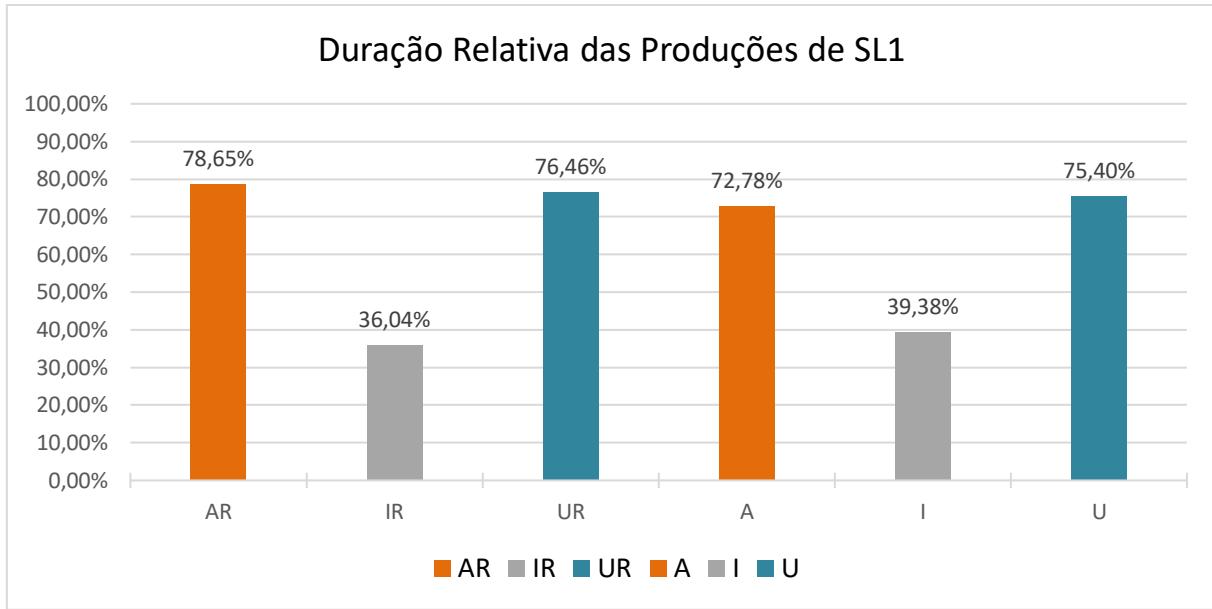


Gráfico 32 – Médias de duração relativa da vogal em contexto de [a] (laranja), [i] (cinza) e [u] (verde) em produções de SL1 para a posição medial

A partir da observação do Gráfico 32, nota-se que não há uma diferença expressiva entre os índices de porcentagens de duração relativa da vogal nas palavras que continham ou não o *tap* como alvo, como em carta x cata. As vogais em ambiente AR e UR ocupam 78,65% e 76,46% da sílaba, já as vogais presentes em contexto de A e U preenchem 72,78% e 75,40% da sílaba, ou seja, os valores percentuais são bastante aproximados. Apenas no contexto de [a], a diferença é um pouco maior, com seis pontos percentuais, o que poderia indicar algum tipo de alongamento.

O menor índice de duração para as palavras em ambiente de IR e I pode apoiar-se no fato de que a consoante inicial se trata de um segmento fricativo, como já reportado, sendo assim, as produções de IR ocupam 36,04% da sílaba e as produções de I preenchem 39,38% dessa.

Logo, não há um indicativo de que a informante estabeleça um contraste, a partir do parâmetro de duração entre as vogais [i] e [u] resultantes da omissão do *tap* e as vogais produzidas em itens lexicais os quais não tinham como alvo o segmento. Para a vogal [a], o distanciamento de seis pontos percentuais sugere a observação de outros dados.

Para fazer a comparação dos dados entre o informante longitudinal e o sujeito controle, foram analisados os valores de duração absoluta de elementos de mesma natureza. Sendo assim, o Gráfico 33 reúne os resultados das médias de duração

absoluta da sílaba e da vogal das produções realizadas por SL1 e por S0. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa; (IR) circo, círculo; (UR) curto, curta, turco. Os valores reportados para os itens lexicais constituídos por [r] pertencem as produções nas quais a informante longitudinal não produziu o *tap*.

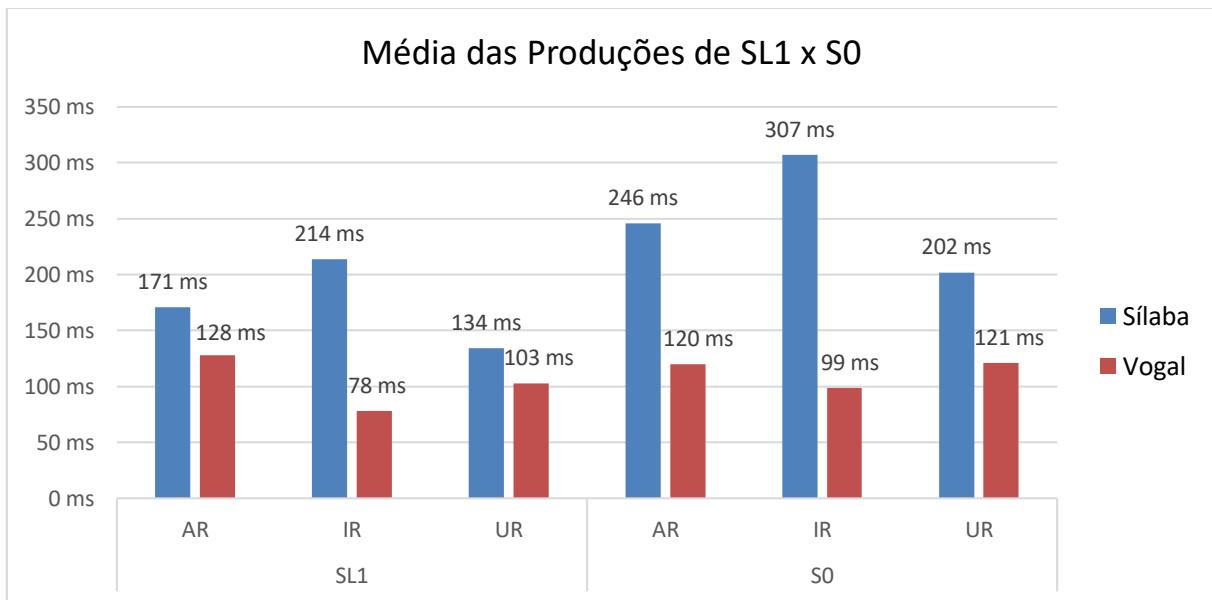


Gráfico 33 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita) em alvos CVC medial

Ao comparar as médias de duração absoluta da sílaba entre SL1 e S0, a partir do Gráfico 33, constata-se maior tempo de produção para o sujeito controle, que apresenta sílabas com duração de 202 ms à 307 ms. As produções da informante longitudinal apresentam valores menores de duração absoluta, de 134 ms à 214 ms. Essa diferença expressiva pode ser explicada pela ausência do segmento rótico nas produções de SL1, a qual pode deixar a sílaba menor.

No que se refere aos valores de duração absoluta da vogal, o Gráfico 33 sugere que, de acordo com o ambiente vocálico, as produções apresentam valores aproximados de duração entre os informantes, com produções vocálicas um pouco mais alongadas por S0. Apenas em relação à vogal [a], constata-se uma duração levemente maior por parte de SL1. Tal fato pode indicar que, para alvos CVC, em contexto de [a], a aprendiz estaria, de fato, utilizando uma estratégia de alongamento compensatório.

Após a análise dos dados, nos quais a informante longitudinal omite o *tap*, a Tabela 19 apresenta os valores referentes às produções nas quais a informante utiliza

o recurso de semivocalização do segmento. A referida tabela exibe as médias de duração absoluta da sílaba, rótico e ditongo das palavras que, em posição medial, abarcavam o rótico como alvo – carpa, farpa, carpa, curta, curto, turco. Em seguida, são exibidos os valores de duração relativa da vogal em relação à sílaba.

CODA MEDIAL					
		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Ditongo	(ditongo - sílaba)
AR	Coleta 1	272	–	261	95,96
	Coleta 2	–	–	–	–
	Coleta 3	–	–	–	–
	<b>MÉDIA</b>	<b>272</b>	–	<b>261</b>	<b>95,96</b>
UR	DP	–	–	–	–
	Coleta 1	181	–	140	77,71
	Coleta 2	–	–	–	–
	Coleta 3	–	–	–	–
	<b>MÉDIA</b>	<b>181</b>	–	<b>140</b>	<b>77,71</b>
	DP	–	–	–	–

Tabela 19 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos, com o recurso da semivocalização, pela informante longitudinal, conforme contexto vocálico

A partir da Tabela 19, observa-se que a semivocalização do rótico ocorreu apenas na primeira coleta, em ambos contextos vocálicos. Em AR, o alvo farpa foi produzido como ['paw.pə], em uma única repetição, apresentando 272 ms de duração absoluta da sílaba e 261 ms de duração do ditongo decrescente. Para UR, o alvo curta foi produzido como ['kuj.tə] e curto foi articulado como ['kuj.tu], em todas as repetições. A média de duração dessas produções é de 181 ms para a sílaba e 140 ms para o glide palatal.

Para examinar as possíveis semelhanças e/ou diferenças entre as semivocalizações de [r] e os ditongos produzidos em itens lexicais que não tinham o tap como alvo, a Tabela 20 apresenta as médias de duração absoluta da sílaba, rótico e ditongo presentes nas palavras que, em posição medial, abarcavam o ditongo como alvo – caixa, cuida. Na mesma tabela são exibidos os valores de duração relativa do ditongo em relação à sílaba.

		CODA MEDIAL		
		Duração absoluta (ms)		Duração relativa (%)
		Sílaba	Ditongo	(ditongo - sílaba)
AI	Coleta 1	185	151	81,62
	Coleta 2	144	96	66,67
	Coleta 3	167	96	57,49
	<b>MÉDIA</b>	<b>165</b>	<b>114</b>	<b>68,59</b>
UI	Coleta 1	185	153	82,52
	Coleta 2	148	120	80,95
	Coleta 3	—	—	—
	<b>MÉDIA</b>	<b>166</b>	<b>136</b>	<b>81,73</b>
	<b>DP</b>	<b>25,93</b>	<b>23,33</b>	<b>1,11</b>

Tabela 20 - Médias de duração absoluta e relativa dos alvos CVV que contêm ditongo em posição medial produzidos pela informante longitudinal, conforme contexto vocálico

Ao analisar os dados do item lexical *caixa*, notaram-se casos de monotongação em duas das três repetições de cada coleta. Segundo Câmara Jr. (1979), a monotongação ocorre quando um ditongo é transformado em uma vogal simples (um monotongo), assim, ['kaj.ʃə] é produzido como ['ka.ʃə]. Quando esse tipo de articulação foi percebido, esses dados não foram contabilizados para os cálculos das médias. O material acústico referente ao alvo *cuida*, da terceira coleta, foi produzido em contexto de frase<sup>24</sup> em todas as repetições, apresentando valores discrepantes<sup>25</sup> em relação à primeira e segunda coleta e, por essa razão, foram retirados do cálculo da média.

A partir da Tabela 20, é possível perceber que as médias de duração absoluta da sílaba, para ambos os contextos são similares. A sílaba composta pelo ditongo [aj] apresenta 165 ms de média de produção, enquanto as sílabas constituídas pelo ditongo [uj] contêm 166 ms de valor médio de produção. Para a produção dos ditongos, apresentam-se os valores de 114 ms para [aj] e 136 ms para [uj].

Conforme já observado, ao não produzir o [r], a criança semivocaliza algumas produções. Para observar a capacidade do parâmetro de duração ser capaz de marcar distinção entre as palavras que semivocalizaram o *tap* e as que não abrangiam a presença do segmento, serão apresentados, a seguir, os Gráficos 34, 35, 36 e 37. Esses estabelecem comparações entre as produções da informante em diferentes

<sup>24</sup> A imagem referente à palavra *cuida* era de uma mãe segurando um bebê, sendo assim, a criança produziu o alvo na seguinte frase “*cuida de bebê*”. A imagem está presente no Anexo B.

<sup>25</sup> Os valores apresentados para *cuida* da terceira coleta são de 290 ms para a sílaba e 243 ms para o ditongo.

contextos, CVC e CVV, ao passo que também realizam comparações com os dados do sujeito controle.

O Gráfico 34 reúne os resultados das médias de duração absoluta da sílaba e da vogal/ditongo das produções de SL1. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa e (UR) curto, curta, turco; (AI) caixa e (UI) cuida. Os valores reportados para as palavras compostas por [r] – AR e UR – são referentes às produções nas quais a informante utiliza o recurso de semivocalização do *tap*.

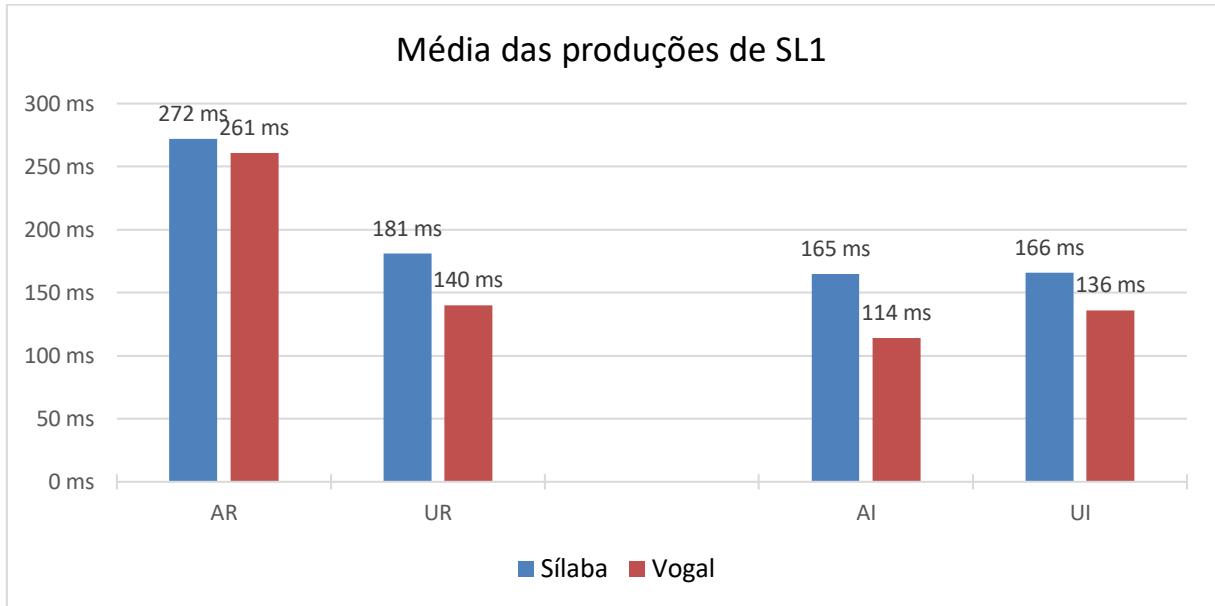


Gráfico 34 - Médias de duração absoluta da sílaba (azul) e da vogal/ditongo (vermelho) das produções de SL1, para a posição medial

Ao observar os dados de AR e AI, a partir do Gráfico 34, é necessário destacar que os valores reportados para AR se referem a um dado isolado. Assim, ainda que seja insuficiente, para fins de comparação com o alvo CVV, chama a atenção pela diferença expressiva de 114 ms. Para o contexto de UR e UI, é possível verificar que há apenas uma diferença de 4 ms entre a duração da sílaba e da vogal/ditongo entre os alvos.

O Gráfico 35 exibe os resultados das médias de duração relativa do ditongo em relação à sílaba das produções realizadas por SL1. Os valores reportados para as palavras compostas por [r] – AR, UR – são referentes às produções nas quais a criança emprega o recurso de semivocalização do *tap*.

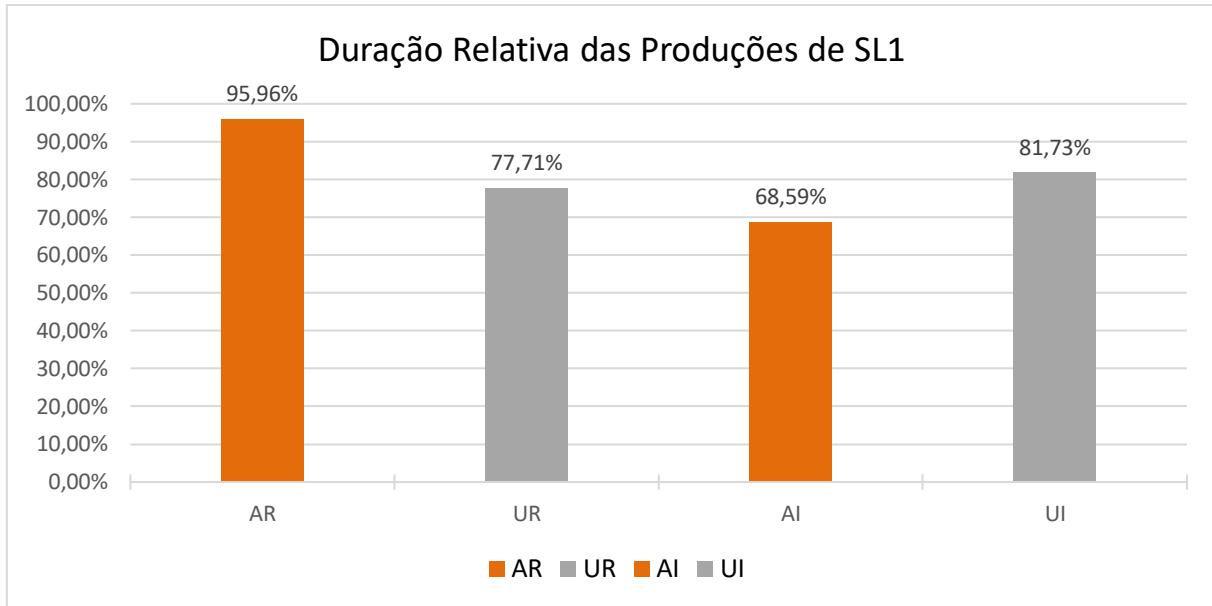


Gráfico 35 - Médias de duração relativa do ditongo em contexto de [a] (laranja) e [u] (cinza) produzido por SL1, para a posição medial

Ao observar o Gráfico 35, é possível verificar que AR apresenta uma duração relativa superior em relação às médias de duração apresentadas nas demais produções. No entanto, esse dado refere-se a uma única produção, conforme já reportado, não sendo viável estabelecer comparação com AI. Os índices de porcentagens, ao comparar UR e UI, são novamente aproximados, ainda que levemente maiores em contexto CVV. As palavras semivocalizadas compreendem 77,71% da sílaba, enquanto os alvos que continham o ditongo [uj] compreenderam 81,73% de duração dentro da sílaba. Uma menor duração do glide, realizado para alvos CVC, justamente seria esperada como pista duracional, tendo em vista que o rótico justamente ocupa uma janela temporal menor do que a da semivogal. Sendo assim, a duração dos ditongos, resultantes de semivocalizações do [r], pode ser um recurso utilizado pela informante para distinguir essas produções dos alvos CVV, constituindo um contraste encoberto.

Para observar se o mesmo comportamento era percebido nos dados do sujeito controle, o Gráfico 36 exibe os resultados das médias de duração relativa da vogal e do ditongo em relação à sílaba.

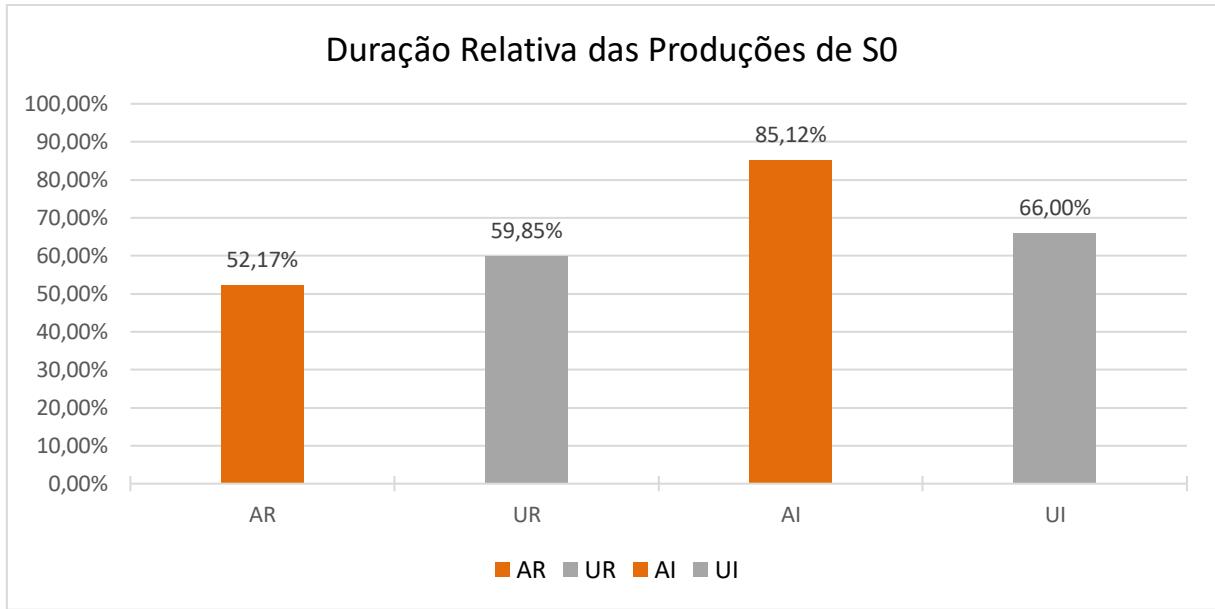


Gráfico 36 - Médias de duração relativa da vogal/ditongo em contexto de [a] (laranja) e [u] (cinza) produzidos por S0, para a posição medial

Observe-se que uma maior duração da média de duração relativa dos alvos CVV, em contexto de AI e UI, quando comparada à média de duração relativa dos alvos CVC nos mesmos contextos, pode também ser constada nos dados de S0, corroborando, assim, a menor duração esperada para os alvos CVC produzidos como CVV por SL1.

Para comparar os dados entre o informante longitudinal e o sujeito controle, o Gráfico 37 concentra os resultados das médias de duração absoluta da sílaba e da vogal das produções realizadas por SL1 e por S0. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa e (UR) curto, curta, turco. Os valores reportados para os itens lexicais constituídos por [r] pertencem às produções nas quais a informante longitudinal semivocalizou o *tap*.

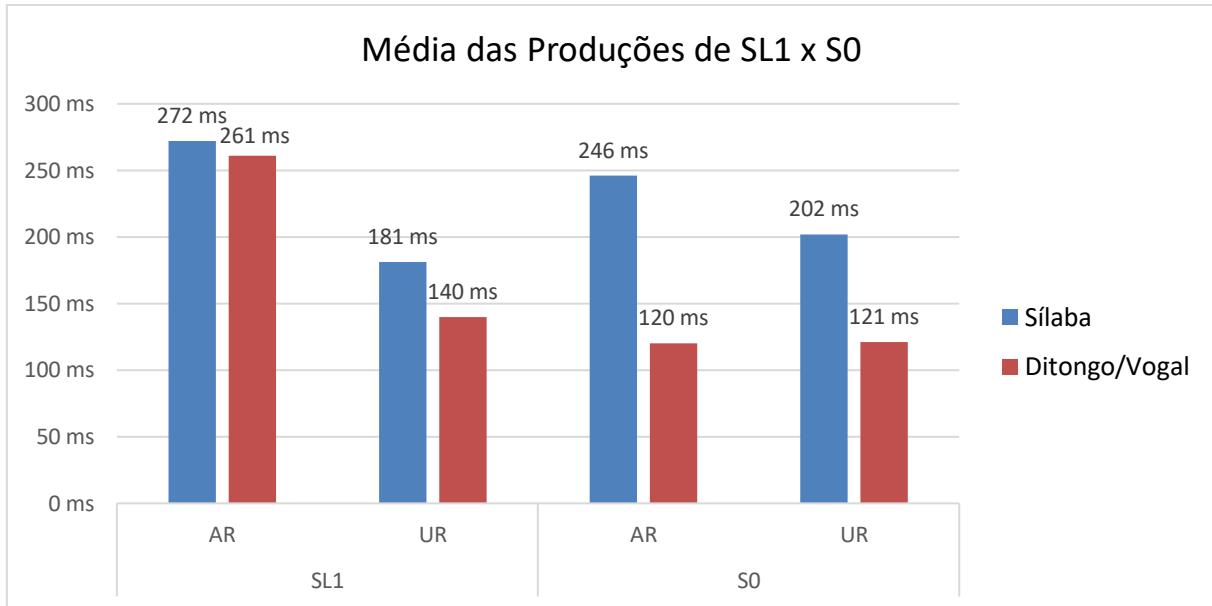


Gráfico 37 - Médias de duração absoluta da sílaba (azul) e do ditongo/vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial

Assim como ocorre com o uso do recurso da omissão, ao observar o Gráfico 37, percebe-se maior duração da sílaba nos dados do sujeito controle, 202 ms a 246 ms. A informante longitudinal exibe médias menores para a produção da sílaba em UR, quando comparado com os dados de S0. Os valores de AR de SL1 não são passíveis de comparação, pois tratam-se de uma produção esporádica.

Ao analisar os dados referentes a UR, é possível perceber que há diferença entre os dados dos dois sujeitos. Uma maior duração nos dados do sujeito longitudinal deve-se, provavelmente, à presença da semivogal nas produções dessa informante.

A partir desse momento serão apresentados os resultados referentes às médias de durações absoluta e relativa das produções em posição final de sílaba.

A Tabela 21 apresenta as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição final, abarcavam o rótico como alvo – mar, lugar, azar, vir, Nadir, Zair, cur, abajur, Artur. Em seguida, são exibidos os valores de duração relativa da vogal em relação à sílaba. Os valores reportados a seguir são referentes às produções nas quais a informante utiliza o recurso de omissão do *tap*.

CODA FINAL					
		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(vocal - sílaba)
AR	Coleta 1	228	0	123	53,95
	Coleta 2	—	—	—	—
	Coleta 3	298	0	233	78,19
	<b>MÉDIA</b>	<b>263</b>	<b>0</b>	<b>178</b>	<b>66,07</b>
IR	Coleta 1	322	0	267	83,06
	Coleta 2	204	0	118	60,58
	Coleta 3	291	0	225	80,29
	<b>MÉDIA</b>	<b>272</b>	<b>0</b>	<b>204</b>	<b>74,64</b>
UR	Coleta 1	—	—	—	—
	Coleta 2	—	—	—	—
	Coleta 3	186	0	121	65,05
	<b>MÉDIA</b>	<b>186</b>	<b>0</b>	<b>121</b>	<b>65,05</b>
	DP	—	—	—	—

Tabela 21 - Médias de duração absoluta e relativa da sílaba e da vogal dos alvos CVC em posição final, produzidos, com o recurso de omissão, pela informante longitudinal, conforme contexto vocálico

Com base nos resultados reportados pela Tabela 21, verifica-se que, para AR e UR, o recurso da omissão não é utilizado com regularidade pela criança. Apenas três produções omitiram o *tap* em sua articulação, estas foram: *azar* (coleta 1), *lugar* (coleta 3) e *abajur* (coleta 3). As demais produções foram semivocalizadas ou articuladas com a lateral [l]. Ao observar as produções de IR, foi constatado que o rótico foi omitido em todas as produções de *vir* e *Nadir*. No entanto, quando o *tap* não foi produzido para o alvo *Zair*, notou-se a presença de um ditongo ou de um hiato.

Para realizar a segmentação do referido alvo, foi observada a transição formântica entre as duas vogais. Segundo Silva *et al.* (2019), os ditongos apresentam uma mudança continua dos dois primeiros formantes, enquanto hiatos permitem a observação dos estados estacionários de cada vogal e sua transição formântica ocorre de uma forma menos contínua. Sendo assim, os valores de IR são referentes à presença do hiato no alvo *Zair* e à omissão do *tap* nas palavras *vir* e *Nadir*.

Para caracterizar as produções que não tinham o rótico como alvo, a Tabela 22 apresenta as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição final, compreendem a vogal como alvo – má, vi, tatu. Em seguida são exibidos os valores de duração relativa da vogal em relação à sílaba.

		CODA FINAL		
		Duração absoluta (ms)		Duração relativa (%)
		Sílaba	Vogal	(vogal - sílaba)
<b>A</b>	Coleta 1	250	189	76,70
	Coleta 2	351	308	86,62
	Coleta 3	319	219	66,87
	<b>MÉDIA</b>	<b>306</b>	<b>239</b>	<b>76,73</b>
	<b>DP</b>	<b>22,86</b>	<b>62,02</b>	<b>9,87</b>
<b>I</b>	Coleta 1	250	186	74,42
	Coleta 2	215	168	78,76
	Coleta 3	284	203	71,34
	<b>MÉDIA</b>	<b>250</b>	<b>185</b>	<b>74,84</b>
	<b>DP</b>	<b>48,55</b>	<b>17,51</b>	<b>3,73</b>
<b>U</b>	Coleta 1	282	151	52,83
	Coleta 2	378	352	85,18
	Coleta 3	281	169	59,80
	<b>MÉDIA</b>	<b>313</b>	<b>224</b>	<b>65,94</b>
	<b>DP</b>	<b>68,24</b>	<b>111,02</b>	<b>17,03</b>

Tabela 22 - Médias de duração absoluta e relativa da sílaba e da vogal dos alvos CV em posição final produzidos pela informante longitudinal, conforme contexto vocálico

Ao observar os valores reportados pela Tabela 22, nota-se que a vogal [u] apresenta maior variação nos seus resultados. A sílaba apresentou valores de duração absoluta entre 281 ms a 378. Ao examinar os dados referentes a vogal, constata-se uma grande oscilação nos valores de duração, 151 ms a 352 ms. Tais aspectos são apontados pelos elevados valores de desvio padrão apresentados por esse contexto vocálico.

Do mesmo modo que foi feito para a posição medial, os Gráficos 38, 39 e 40 realizam comparações para investigar se a criança utiliza algum recurso para diferenciar os pares de palavras, com e sem o rótico.

A partir das observações reportadas, o Gráfico 38 exibe os resultados das médias de duração absoluta da sílaba e da vogal das produções das palavras realizadas por SL1. A organização dos dados se dá conforme os seguintes contextos: (AR) mar, lugar, azar; (IR) vir, Nadir, Zair; (UR) cur, abajur, Artur; (A) má; (I) vi; (U) tatu. Os valores apresentados para as palavras compostas por [r] – AR, IR, UR – são referentes às produções nas quais a informante utiliza o recurso de omissão do *tap*.

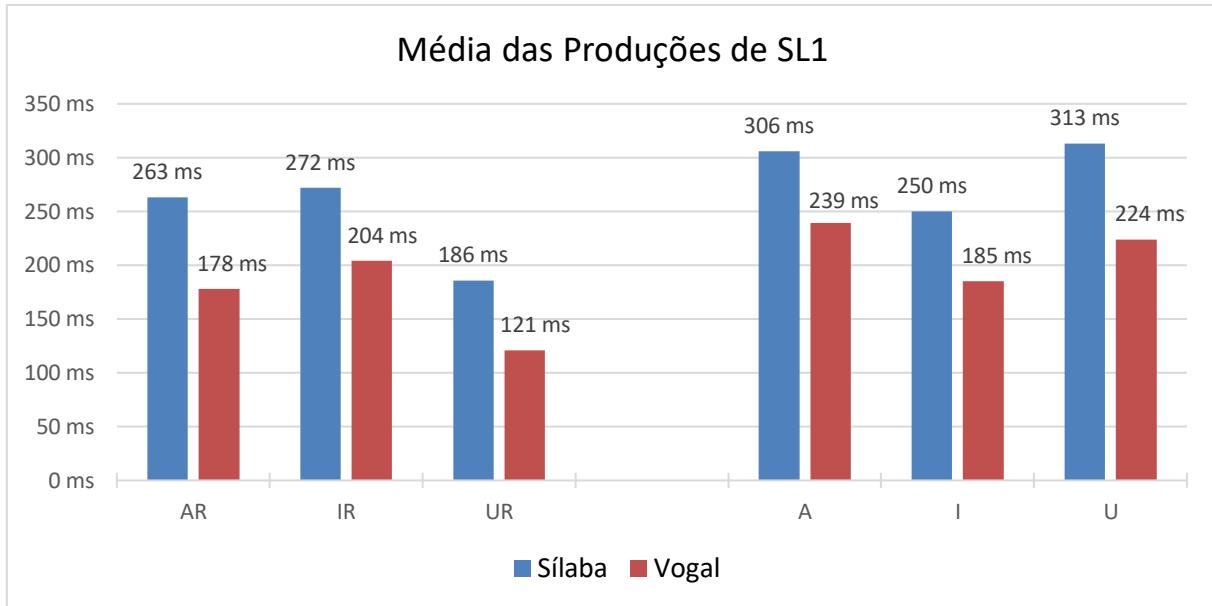


Gráfico 38 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1, para a posição final

Conforme revela o Gráfico 38, as médias de duração absoluta da sílaba e da vogal verificadas para AR e UR registram diferenças expressivas ao compará-las com os resultados de A e U. As palavras que abarcam o rótico como alvo apresentam duração menor, AR e UR contêm a sílaba com 263 ms e 186 ms, enquanto A e U apresentam duração absoluta da sílaba de 306 ms e 313 ms. O mesmo acontece com a duração da vogal desses respectivos contextos. Contudo, os resultados para IR e I exibem um comportamento diferente, os valores de duração são mais aproximados. As palavras que são constituídas pelo rótico apresentam valores de duração maiores quando comparadas aos itens lexicais que não possuem o segmento como alvo, tanto no que concerne à duração da sílaba quanto no que concerne à duração da vogal. Sendo assim, a maior duração absoluta da vogal [i] para alvos CVC, em comparação à duração da vogal [i] para alvos CV, poderia indicar a presença de alongamento compensatório na aquisição do rótico em coda.

O Gráfico 39 exibe as médias de duração relativa da vogal das produções realizadas por SL1. Os resultados reportados para os itens lexicais compostos por [r] – AR, IR, UR – concernem às produções nas quais a criança emprega o recurso de omissão do *tap*.

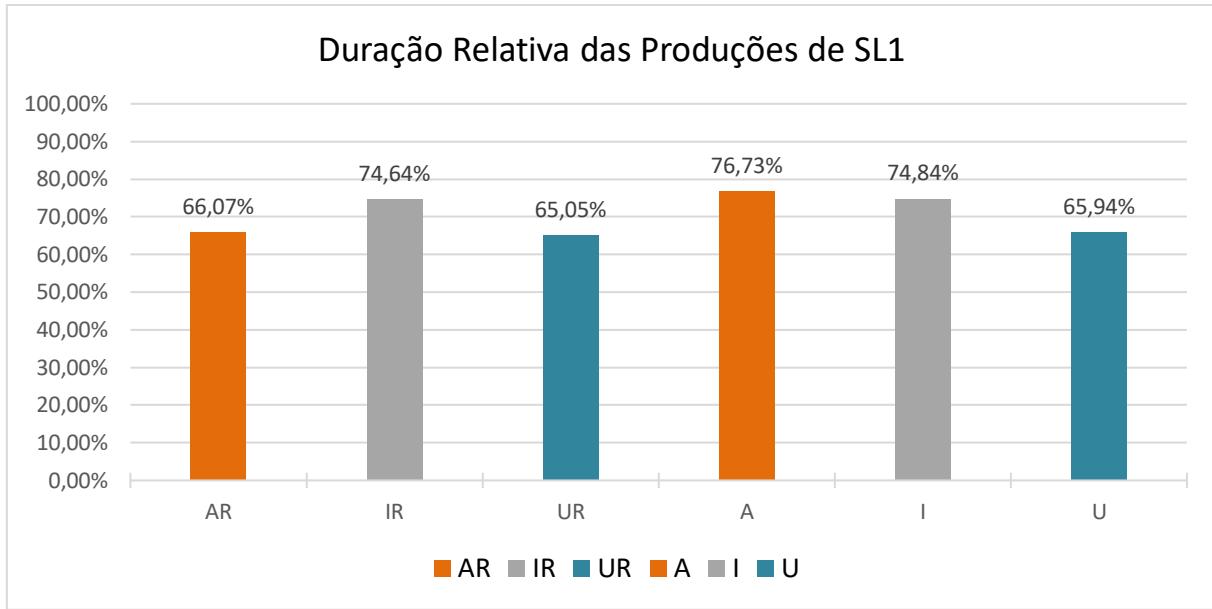


Gráfico 39 - Médias de duração relativa da vogal em contexto de [a] (laranja), [i] (cinza) e [u] (verde) produzida por SL1, para a posição final

Ao analisar o Gráfico 39, percebe-se que os índices de porcentagens de duração relativa das vogais, nas palavras que continham o *tap* como alvo e nas que não o continham, apresentam valores aproximados. Ainda assim, os itens lexicais que incluem o rótico como alvo exibem uma porcentagem um pouco menor do que as outras. O segmento vocalico que mais ocupou a sílaba, nas palavras que possuíam apenas a vogal como alvo, foi A, com 76,73%. No entanto, IR foi o contexto que exibiu uma maior porcentagem, 74,64%. Sendo assim, os dados de duração relativa sugerem que não é possível afirmar que a criança, a partir do parâmetro de duração, marque distinção entre alvos CVC – contendo o *tap* – e CV, uma vez que a diferença entre as porcentagens é pouco expressiva.

Com o objetivo de comparar as produções do sujeito controle com os dados da informante longitudinal, foram observados os valores de duração absoluta da sílaba e da vogal das palavras que abarcam o rótico como alvo. Posto isto, o Gráfico 40 exibe as médias das produções de SL1 e S0. Os resultados apresentados estão dispostos conforme os seguintes contextos: (AR) mar, lugar, azar; (IR) vir, Zair, Nadir; (UR) cur, abajur, Artur. Os valores exibidos para as palavras constituídas por [r] concernem às produções em que a criança omitiu o *tap*.

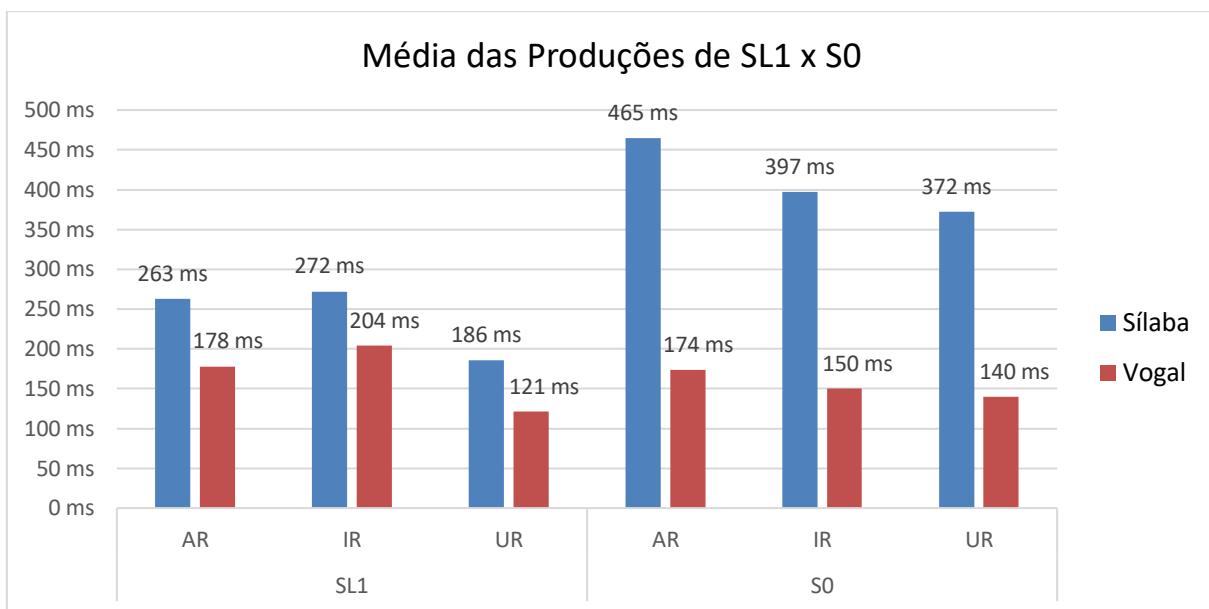


Gráfico 40 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita), para a posição final

A análise do Gráfico 40 permite constatar maior tempo de duração silábica nas produções do sujeito controle. Os valores de duração da sílaba de S0 giram em torno de 465 e 372 ms. As produções da informante longitudinal apresentam valores menores de duração absoluta, de 272 ms a 186 ms. Esse comportamento ocorreu também em posição medial, no entanto, de forma menos acentuada. Essa diferença mais expressiva nos dados de posição final parece explicar-se pelo fato de que as produções do rótico pelo sujeito controle foram de fato articuladas, e, ainda, como vibrantes. Sendo assim, a omissão do segmento rótico nas produções de SL1 diminui os valores de duração da sílaba.

Ao observar os resultados de duração da vogal, o Gráfico 40 indica que, de acordo com o ambiente vocálico, os informantes exibem valores aproximados apenas para o contexto de [a]. Contudo, uma maior duração da vogal alta anterior nas produções de SL1 chama a atenção. A informante longitudinal apresenta 204 ms de duração para [i], enquanto o sujeito controle exibe 150 ms de duração para o mesmo segmento vocálico. A maior duração da vogal anterior poderia indicar que a pista de duração vocálica é relevante para a distinção do par CVC – CV, o que parece ser corroborado pelos resultados dispostos no Gráfico 38. Os resultados do Gráfico 39, no entanto, ainda não permitem fazer essa afirmação.

Após a observação dos resultados oriundos da omissão do *tap*, a Tabela 23 exibe os dados de produções resultantes do emprego do recurso de semivocalização do segmento. A tabela mencionada apresenta as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição final, continham o rótico como alvo – mar, lugar, azar, cur, abajur, Artur. Em seguida são exibidos os valores de duração relativa da vogal em relação à sílaba.

CODA FINAL			
		Duração absoluta (ms)	Duração relativa (%)
		Sílaba	(ditongo - sílaba)
AR	Coleta 1	424	338
	Coleta 2	295	238
	Coleta 3	213	157
	<b>MÉDIA</b>	<b>310</b>	<b>244</b>
	<b>DP</b>	<b>106,27</b>	<b>4,03</b>
UR	Coleta 1	459	329
	Coleta 2	242	170
	Coleta 3	296	258
	<b>MÉDIA</b>	<b>332</b>	<b>252</b>
	<b>DP</b>	<b>112,97</b>	<b>9,21</b>

Tabela 23 - Médias de duração absoluta e relativa dos alvos CVC em posição final, produzidos, com o recurso da semivocalização, pela informante longitudinal, conforme contexto vocálico

A Tabela 23 aponta que a semivocalização do *tap* ocorreu para os dois ambientes vocálicos presentes no instrumento de coleta. Em AR, as principais palavras semivocalizadas foram *mar* e *azar*. Essas produções apresentaram valores de 310 ms para duração da sílaba e 244 ms para a duração absoluta do ditongo. Para UR, o principal item lexical que apresentou o uso do recurso de semivocalização foi *Artur*, sendo que o *tap* em *abajur* foi semivocalizado apenas uma vez na segunda coleta. Esses tipos de produções apresentaram uma média de 332 ms para a duração da sílaba e 252 ms de duração absoluta do ditongo. A semivocalização também ocorreu para o contexto de IR final na palavra *Zair*. No entanto, ao contrário dos contextos de AR e UR, nesse caso, a semivocalização pode ser simplesmente decorrente da omissão do *tap*.

A fim de investigar prováveis semelhanças e diferenças entre as palavras que possuíam o ditongo como alvo e as semivocalizações de [r], a Tabela 24 exibe as médias de duração absoluta da sílaba, rótico e ditongo presentes na palavra que, em

posição final, abarcava o ditongo como alvo – *pai*. Também exibe os valores de duração relativa do ditongo em relação à sílaba.

CODA FINAL			
AI		Duração absoluta (ms)	Duração relativa (%)
	Sílaba	Ditongo	(ditongo - sílaba)
	Coleta 1	224	208
	Coleta 2	148	130
	Coleta 3	279	262
<b>MÉDIA</b>		<b>217</b>	<b>200</b>
DP		65,90	66,45
			3,22

Tabela 24 - Médias de duração absoluta e relativa dos alvos CVV em posição final produzido pela informante longitudinal, conforme contexto vocálico

Ao elaborar a metodologia da presente pesquisa, houve uma falha na seleção do item lexical que, provavelmente, transcorreu de um erro de digitação<sup>26</sup>. Sendo assim, o ditongo [uj], em posição final, não foi inserido no instrumento de coleta, dessa forma, produções como [aba.'3uj] e [a.'tuj] não possuem um par para estabelecer uma análise comparativa.

A partir a Tabela 24, nota-se que a sílaba formada pelo ditongo [aj] apresenta 217 ms de duração absoluta. Já o ditongo ocupa 91,11% da sílaba, com uma duração de 200 ms.

De acordo com as observações anteriores, os Gráficos 41, 42 e 43 estabelecem algumas comparações entre as produções da informante longitudinal em contextos diferentes, CVC e CVV, ao passo que também realiza comparações com as produções do sujeito controle. Essa comparação novamente busca verificar se o parâmetro de duração é usado para marcar contraste entre as palavras que apresentam semivocalização do *tap* e as que contêm o alvo CVV.

Após as considerações apresentadas, o Gráfico 41 reúne os resultados das médias de duração absoluta da sílaba e da vogal/ditongo das produções das palavras realizadas por SL1. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) *mar*, *lugar*, *azar*; (UR) *cur*, *abajur*, *Artur*; (AI) *pai*. Os valores

<sup>26</sup> O item lexical que foi gravado para ser o par dos alvos semivocalizados foi *rubi*, no entanto, o correto seria a palavra *Rui*.

reportados para as palavras compostas por [r] – AR, UR – são referentes às produções nas quais a informante utiliza o recurso de semivocalização do *tap*.

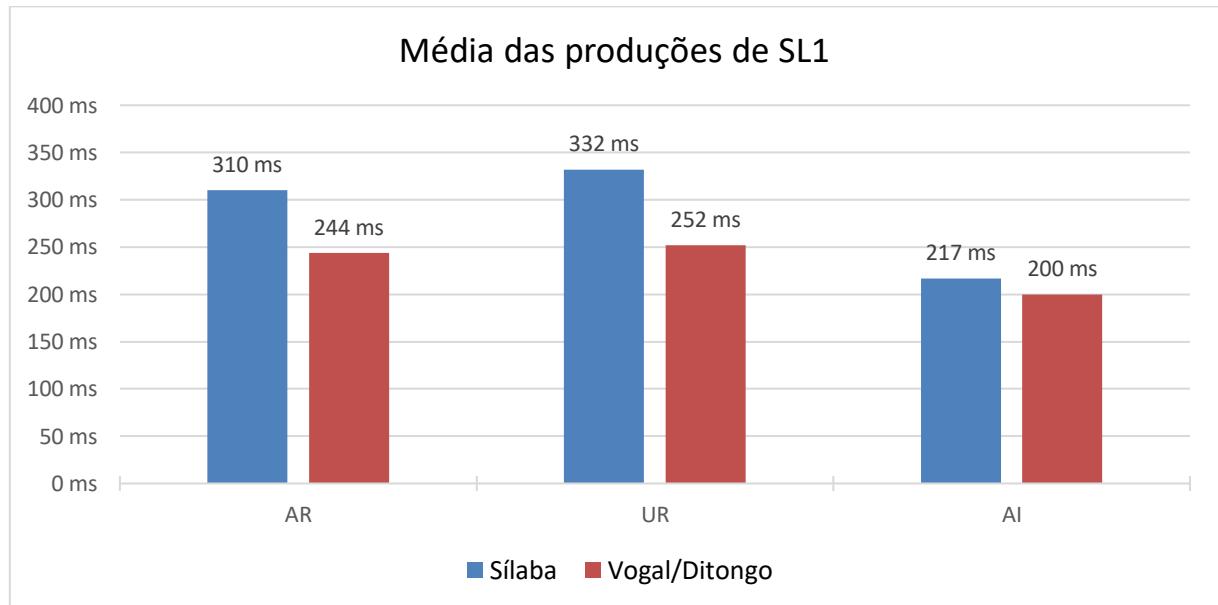


Gráfico 41 - Médias de duração absoluta da sílaba (azul) e da vogal/ditongo (vermelho) das produções de SL1, para a posição final

O Gráfico 41 permite observar que, em termos de duração absoluta, as produções semivocalizadas apresentam maiores durações do que as produções compostas por ditongo. Embora claramente se estabeleça o contraste, por meio de diferenças de duração, entre alvos CVC e CVV, esperar-se-ia uma menor duração para os alvos CVC, no entanto, não é o que se apresenta para o par AR x AI.

O Gráfico 42 apresenta as médias de duração relativa do ditongo em relação à sílaba das produções da informante longitudinal. Os resultados reportados para as palavras compostas por [r] – AR, UR – referem-se às produções nas quais criança emprega o recurso de semivocalização do *tap*.

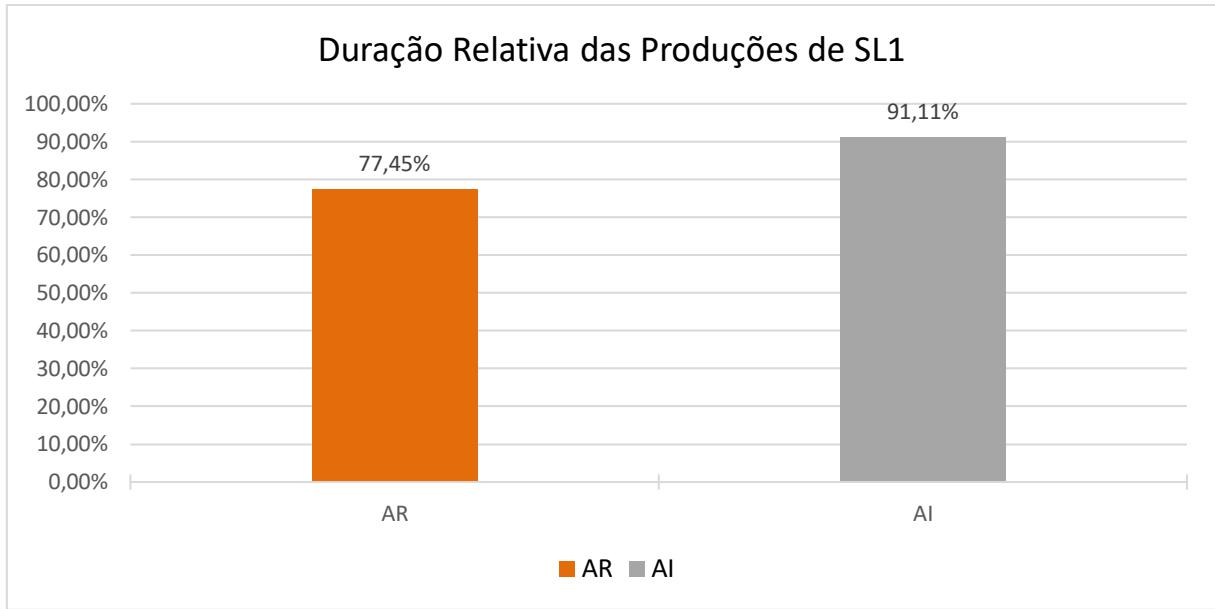


Gráfico 42 - Médias de duração relativa de produções semivocalizadas em contexto de AR e IR (laranja) e de produções do ditongo [aj] (cinza) produzido por SL1, para a posição final

A partir da observação do Gráfico 42, verifica-se que os valores de duração relativa entre as produções de AR e AI apresentam diferença expressiva, com 77,45% e 91,11%, respectivamente. Sendo assim, a duração dos ditongos, decorrentes de semivocalizações do *tap* em contexto da vogal baixa, parece ser um recurso utilizado pela criança para distinguir essas produções das quais o segmento rótico não era alvo.

A fim de comparar as produções da informante longitudinal e do sujeito controle, foram examinados dados de elementos de mesma natureza. Desse modo, o Gráfico 43 apresenta os resultados das médias de duração absoluta da sílaba e da vogal/ditongo das produções realizadas por SL1 e por S0. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) mar, lugar, azar; (UR) cur, abajur, Artur. Os valores exibidos para as palavras constituídas por [r] concernem às produções que a informante longitudinal semivocalizou o *tap*.

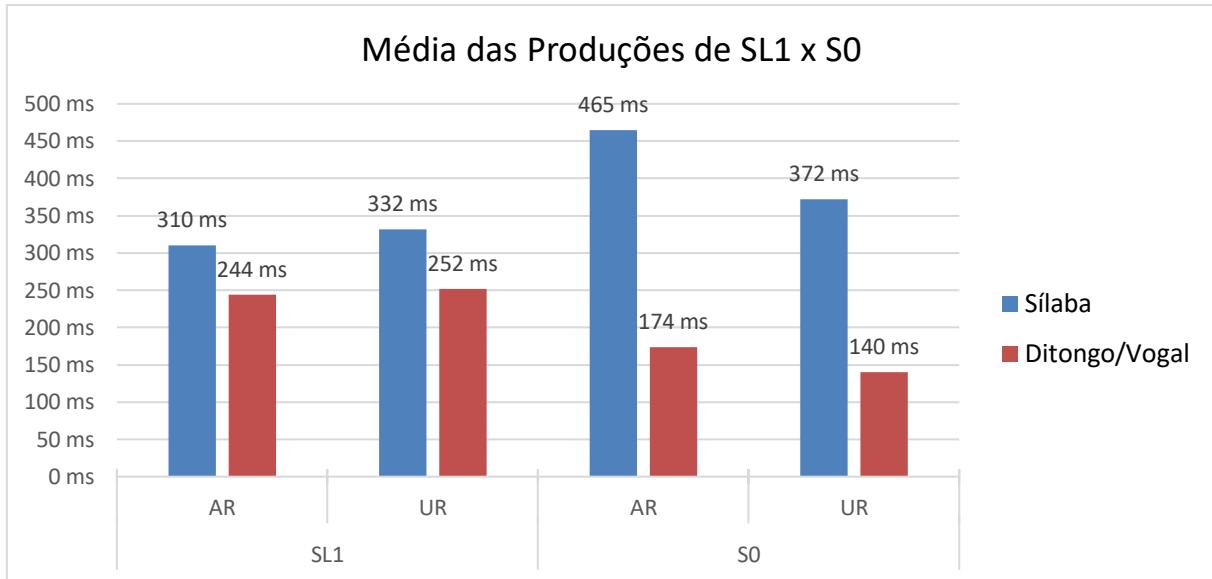


Gráfico 43 - Médias de duração absoluta da sílaba (azul) e do ditongo/vogal (vermelho) das produções de SL1 (à esquerda) em comparação às produções de S0 (à direita), para a posição final

Do mesmo modo que se verificou para o uso do recurso de omissão do rótico, o Gráfico 43 evidencia maior tempo de duração para os dados do sujeito controle. A diferença expressiva entre os sujeitos, para os valores referentes à duração da sílaba, expressa as produções de vibrantes por parte do sujeito controle. A considerável diferença entre os informantes, para o contexto vocálico, evidencia a produção de ditongos por parte de SL1 e de vogais por S0. Posto isto, a articulação de um segmento vibrante, por parte de S0 aumenta o tempo de duração do constituinte silábico. Já a semivocalização do *tap* nas produções de SL1 aumenta os valores de duração do contexto vocálico.

A partir da análise dos aspectos duracionais das produções de SL1, em posição medial, foi constatado que a informante não produziu o *tap* ao longo das 3 coletas, tampouco estabeleceu uma distinção entre a vogal oriunda da omissão e a vogal de uma sílaba CV, a partir de pistas duracionais. Do mesmo modo, não é possível afirmar que suas produções apresentem diferenças quando comparadas com os dados do sujeito controle, uma vez que as vogais de ambos os sujeitos exibem valores de duração absoluta aproximados. No entanto, os valores de duração relativa referentes ao emprego do recurso da semivocalização, em ambiente vocálico de [u], parecem evidenciar que a informante estabelece um contraste por meio de uma menor duração de suas produções em contexto CVC das produções em contexto CVV. Pois, nesse caso, é esperada uma menor duração do glide resultante da semivocalização, uma

vez que o *tap* ocuparia uma janela temporal menor do que a semivogal. Do mesmo modo, esse comportamento é encontrado nos dados, em contexto de CVC e CVV, do sujeito controle. Para os resultados em posição final, a análise dos gráficos não permite inferir que a informante estabeleça contraste entre as produções a partir da aplicação do recurso da omissão. Já para a utilização do recurso de semivocalização, é possível inferir que a criança vale-se de uma menor duração do glide, presente em contexto CVC, para estabelecer diferença entre essas produções e as produções em contexto CVV.

#### 4.2.2.3 Trajetória Formântica

A análise da trajetória formântica nos dados da informante longitudinal tem como principal propósito efetuar inferências acerca de possíveis contrastes encobertos presentes nas produções da criança em fase de aquisição do *tap*. Para tanto, foi feita a comparação entre as palavras que foram classificadas, durante a análise de oitiva, como omissão ou semivocalização e as palavras produzidas em contexto CV ou CVV. Por exemplo, a produção em que foi percebida a omissão do *tap*, como na articulação da palavra *carta*, produzida como ['ka.tə], foi comparado com a produção da palavra *cata* (contexto CV) produzida como ['ka.tə]. Do mesmo modo foi feito para a semivocalização, o ditongo oriundo da articulação do alvo *curto*, produzido como ['kuj.tu], foi comparado com o ditongo produzido em ['kuj.də], referente à palavra *cuida* (contexto CVV).

Tendo em vista que não é possível apresentar a trajetória formântica de todas as produções, serão apresentados aqui alguns dados em que foi constatado, de forma mais evidente, a presença de possíveis contrastes encobertos.

Assim, foram gerados gráficos das trajetórias formânticas que permitiram observar o movimento dos três formantes no transcorrer do tempo do segmento vocálico, a fim de examinar possíveis semelhanças e diferenças entre as produções mencionadas. Cada gráfico representa apenas uma produção do alvo.

Os Gráficos 44 e 45 apresentam as trajetórias formânticas de F1, F2 e F3 de produções realizadas por SL1, em contexto de [u] (curto X cuca). O *tap* foi omitido na produção de *curto*.

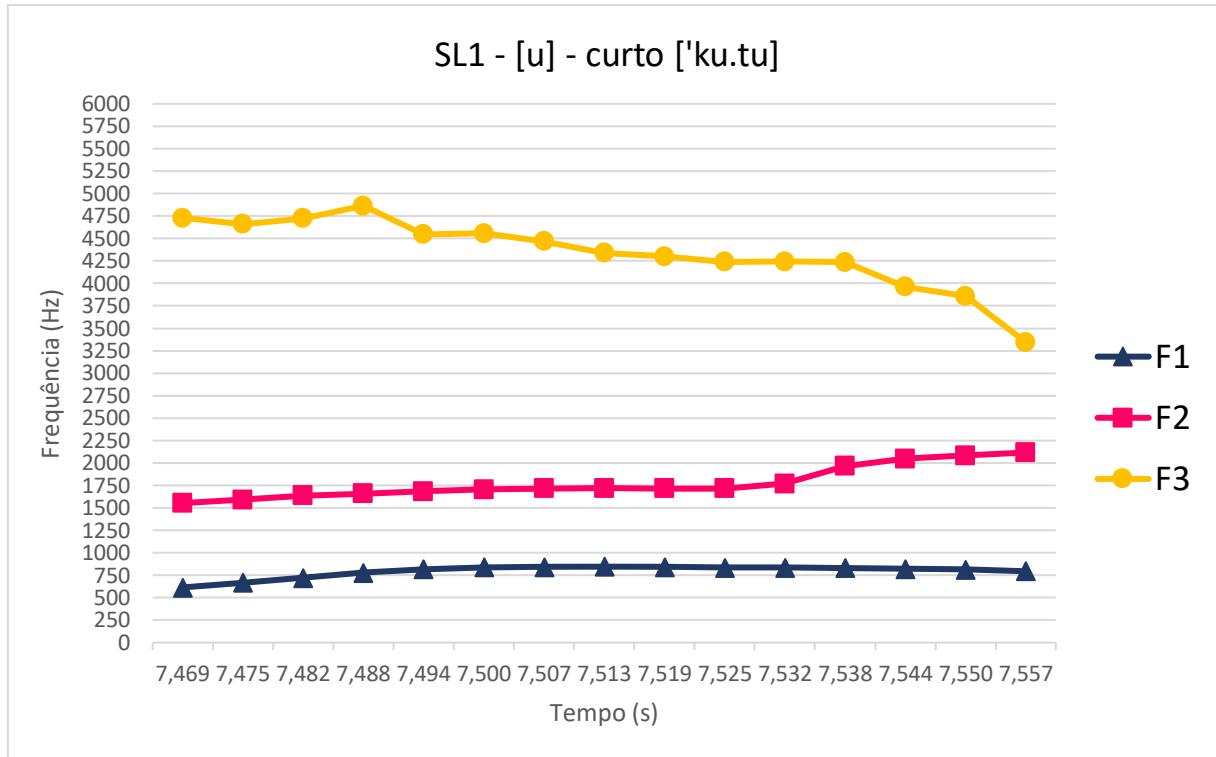


Gráfico 44 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u], resultante da omissão do rótico em ['ku.tu] da segunda coleta de SL1

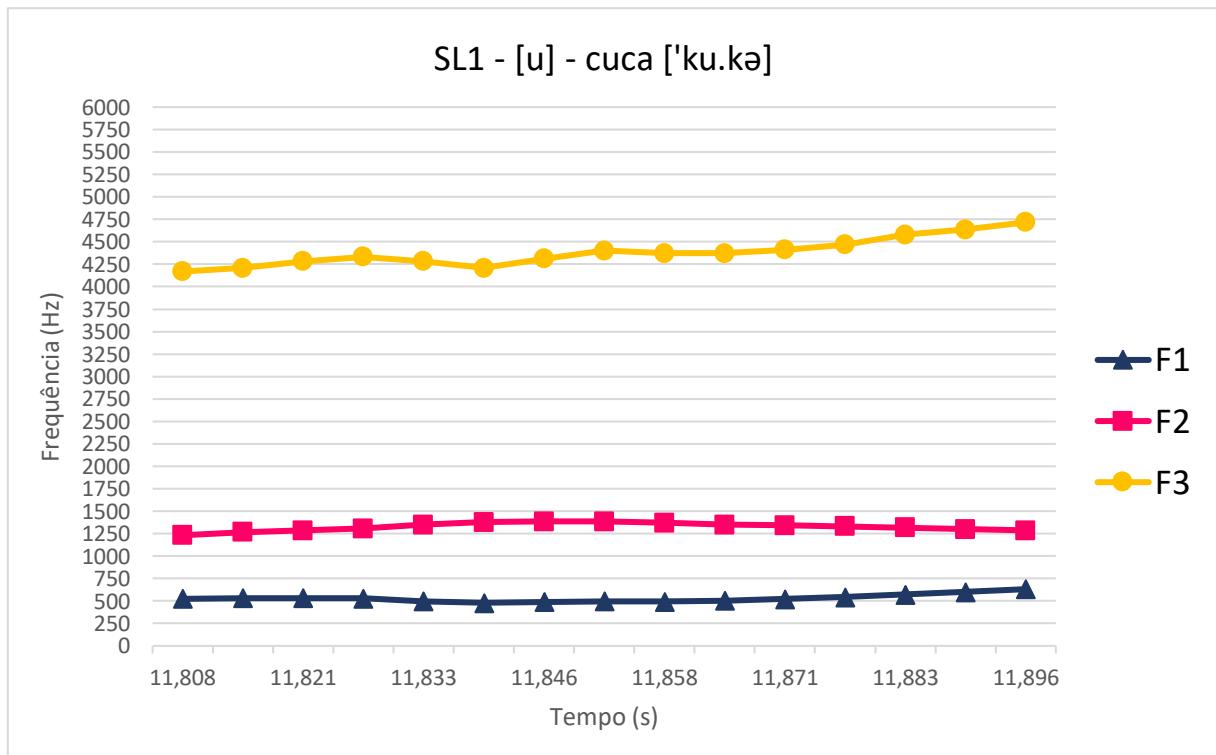


Gráfico 45 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [u], em contexto CV, da produção de ['ku.tu], da segunda coleta de SL1

Ao comparar as duas produções, a partir dos Gráficos 44 e 45, percebe-se que, embora sejam relativas à mesma vogal, a trajetória formântica para ambas apresenta um comportamento diferente. Para a produção de *cuca*, os dois primeiros formantes estão horizontalmente alinhados e o terceiro exibe uma sensível elevação. Quanto à omissão do *tap* no alvo *curto*, observa-se que os valores para F3 decrescem e os valores de F2 se elevam na parte final da vogal [u]. Esse mesmo comportamento foi evidenciado durante a produção do mesmo alvo pelo sujeito controle (Gráfico 25), no entanto, após a produção da vogal, S0 articula um *tap*. Desse modo, a trajetória formântica, na omissão do segmento pela informante longitudinal, parece revelar a presença de um contraste encoberto.

Os Gráficos 46 e 47 exibem as trajetórias formânticas de F1, F2 e F3 de produções articuladas por SL1, em contexto de [i] (circo X chico) sendo que o *tap* foi omitido na produção de *circo*.

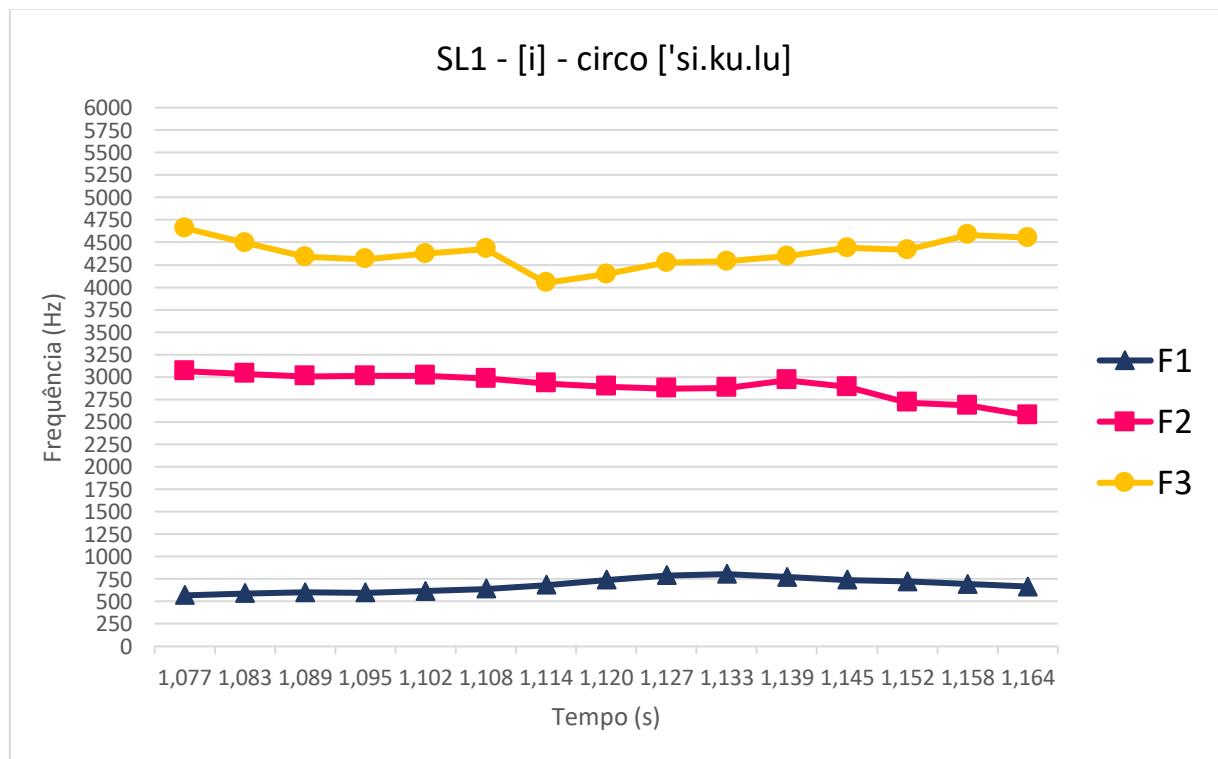


Gráfico 46 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i], resultante da omissão do rótico em ['si.ku.lu] da segunda coleta de SL1

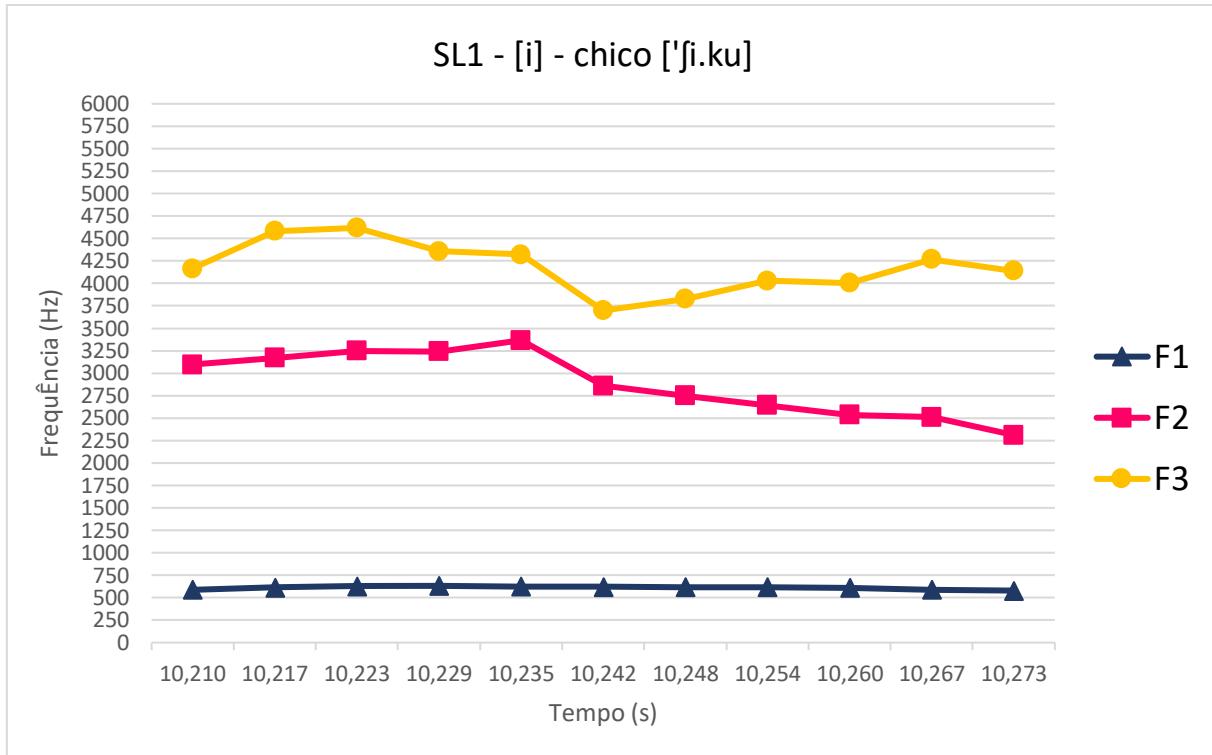


Gráfico 47 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i], em contexto CV, da produção de ['ʃi.ku], da segunda coleta de SL1

Observando os Gráficos 46 e 47, nota-se que, para a omissão do *tap* em *círcio*, o primeiro formante exibe uma sensível elevação, ao passo que os valores de F2 e F3 diminuem de forma suave. Em relação à produção da vogal CV em *Chico*, percebe-se que o primeiro formante está horizontalmente alinhado. Já o segundo formante cai de forma considerável. Ao comparar o valor de F3 das duas produções, tem-se, para o fim de [i] resultante da omissão, o valor de 4552 Hz e para o [i] produzido em CV, 4139 Hz. O comportamento da vogal durante a omissão é similar à trajetória formântica apresentada para a produção do segmento no dado do sujeito controle, apresentado no Gráfico 23. Assim, é possível que o contraste encoberto, envolvendo uma menor queda de F2 em contexto de IR, possa estar relacionado com a posição ântero-posterior da língua.

Retomando a análise dos resultados obtidos pelos gráficos de trajetória formântica, tem-se que, para a produção de [u] em contexto CV, F1 e F2 apresentam-se horizontalmente alinhados e F3 exibe uma leve elevação. Já para mesma vogal, oriunda da omissão do *tap*, o valor de F2 se eleva e F3 cai. Quanto à produção de [i] em contexto CV, F1 não sofre alterações, mantendo-se horizontalmente alinhado, e o valor de F2 cai de forma expressiva. No entanto, o F1 da vogal oriunda da omissão do *tap*, eleva-se um pouco e os valores de F2 e F3 diminuem.

Embora essas comparações sejam feitas entre vogais da mesma natureza, as trajetórias formânticas se modificam a depender do contexto silábico, evidenciando, assim, contrições em regiões diferentes no trato oral para as mesmas vogais. Esse fato é capaz de sugerir que a análise da trajetória formântica das produções de SL1 apresenta-se como uma ferramenta eficaz na revelação de produções gradientes.

#### 4.2.3 Sujeitos Transversais

Os informantes transversais dispõem de uma gravação de áudio cada. A amostra de dados era composta por 312 *tokens* e, após avaliação de qualidade acústica, 25 precisaram ser descartados. Portanto, 287 dados foram submetidos à análise, sendo que ST1 dispõe de 74, ST2 possui 68, à ST3 pertencem 72 e, por fim, ST4 apresenta 74 dados.

Para melhor organização da apresentação dos resultados, as próximas subseções – 4.2.3.1 a 4.2.3.3 – reportam a análise dos dados referentes à inspeção visual do espectrograma, aspectos duracionais e trajetória formântica, respectivamente.

##### 4.2.3.1 Inspeção Visual do Espectrograma

A inspeção visual dos dados dos sujeitos transversais revelou a produção de cinco variantes róticas: *tap*, vibrante, fricativa glotal, aproximante alveolar e *tap* fricatizado, como mostra a Tabela 25.

Variante rótica produzida		MEDIAL				FINAL			
		[a]	[i]	[u]	Total	[a]	[i]	[u]	Total
<b>ST1</b>	<i>Tap</i>	0/9	0/5	0/9	<b>0/23</b>	0/9	0/7	0/9	<b>0/25</b>
	Não produzido	9/9	5/5	9/9	<b>23/23</b>	9/9	7/7	9/9	<b>25/25</b>
<b>ST2</b>	<i>Tap</i>	1/9	1/6	5/9	<b>7/24</b>	4/6	0/6	5/7	<b>9/19</b>
	Fricativa Glotal Desvozeada	0/9	0/9	0/9	<b>0/9</b>	2/6	6/6	0/7	<b>8/19</b>
	Não produzido	8/9	5/6	4/9	<b>17/24</b>	0/6	0/6	2/7	<b>2/19</b>
<b>ST3</b>	<i>Tap</i>	5/9	0/6	7/8	<b>12/23</b>	5/6	7/8	5/8	<b>17/22</b>
	Vibrante	4/9	0/6	1/8	<b>5/23</b>	1/6	0/8	0/8	<b>1/22</b>
	Aproximante Alveolar	0/9	0/9	0/8	<b>0/23</b>	0/6	0/8	1/8	<b>1/22</b>
	<i>Tap</i> Fricatizado	0/9	0/9	0/8	<b>0/23</b>	0/6	1/8	1/8	<b>2/22</b>
<b>ST4</b>	Não produzido	0/9	6/6	0/8	<b>6/23</b>	0/6	0/8	1/8	<b>1/22</b>
	<i>Tap</i>	0/9	0/5	0/8	<b>0/22</b>	0/8	0/9	1/9	<b>1/26</b>
	Não produzido	9/9	5/5	8/8	<b>22/22</b>	8/8	9/9	8/9	<b>25/26</b>

Tabela 25 - Produção das variantes róticas de acordo com o ambiente vocálico e posição na palavra para cada informante transversal

Ao analisar a Tabela 25, percebe-se que, quando produzida uma variante rótica, o *tap* foi o mais articulado. Ao analisarem-se os dados individualmente, constata-se que ST1 não produziu nenhuma variante rótica. Para ST2, as produções de *taps* correspondem a 29% das articulações em posição medial e 53% em posição final. O sujeito ST3 foi a criança que mais apresentou produções de róticos, sendo que, em posição medial, 52% das articulações apresentam-se como *taps* e, em posição final, 77%. Além da produção do *tap*, verificou-se, para ST3, 5 produções (22%) de vibrantes em posição medial. O mesmo informante apresentou outras duas variantes de róticos, a saber: aproximante alveolar e *tap* fricatizado. O sujeito ST4 apresentou apenas uma produção do *tap* em contexto de [u].

Para identificar os possíveis contrastes encobertos presentes nos dados dos sujeitos transversais, observaram-se as três principais características do *tap*, segundo Silva *et al.* (2019): (i) ausência de energia; (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico. Sendo assim, a Tabela 26 apresenta as possibilidades e ocorrências de características do *tap* na forma de onda e no espectrograma. Na mesma tabela, constam os prováveis contrastes encobertos observados na fala dos sujeitos transversais para a posição medial.

---

**POSSIBILIDADES E OCORRÊNCIAS DE CONTRASTES ENCOBERTOS**


---

ST1							
FORMA DE ONDA			ESPECTROGRAMA				Contraste
	(i)	(ii)	(iii)	(i)	(ii)	(iii)	
<b>A</b>	0/9	0/9	0/9	0/9	0/9	0/9	<b>0/9</b>
<b>I</b>	2/5	0/5	1/5	2/5	0/5	1/5	<b>2/5</b>
<b>U</b>	5/9	1/9	5/9	5/9	5/9	5/9	<b>5/9</b>
<b>TOTAL</b>	<b>7/23</b>	<b>1/23</b>	<b>6/23</b>	<b>7/23</b>	<b>5/23</b>	<b>6/23</b>	<b>7/23</b>
ST2							
<b>A</b>	3/8	2/8	2/8	3/8	2/3	2/8	<b>3/8</b>
<b>I</b>	1/5	0/5	1/5	1/5	1/5	1/5	<b>1/5</b>
<b>U</b>	0/4	0/4	0/4	0/4	0/4	0/4	<b>0/4</b>
<b>TOTAL</b>	<b>4/17</b>	<b>2/17</b>	<b>3/17</b>	<b>4/17</b>	<b>3/17</b>	<b>3/17</b>	<b>4/17</b>
ST3							
<b>A</b>	0/0	0/0	0/0	0/0	0/0	0/0	<b>0/0</b>
<b>I</b>	4/6	3/6	3/6	4/6	3/6	3/6	<b>4/6</b>
<b>U</b>	0/0	0/0	0/0	0/0	0/0	0/0	<b>0/0</b>
<b>TOTAL</b>	<b>4/6</b>	<b>3/6</b>	<b>3/6</b>	<b>4/6</b>	<b>3/6</b>	<b>3/6</b>	<b>4/6</b>
ST4							
<b>A</b>	0/9	1/9	1/9	1/9	1/9	1/9	<b>1/9</b>
<b>I</b>	2/5	2/5	2/5	2/5	2/5	2/5	<b>2/5</b>
<b>U</b>	0/8	0/8	0/8	0/8	0/8	0/8	<b>0/8</b>
<b>TOTAL</b>	<b>2/22</b>	<b>3/22</b>	<b>3/22</b>	<b>3/22</b>	<b>3/22</b>	<b>3/22</b>	<b>3/22</b>

Tabela 26 - Possibilidades e ocorrências de contrastes encobertos a partir das características do *tap*, sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala dos sujeitos transversais, para posição medial

Ao observar a Tabela 26, nota-se que a ausência de energia foi a característica do *tap* que mais colaborou para a observação de prováveis contrastes encobertos, em posição medial. O elemento vocálico foi a segunda característica que mais evidenciou a presença de possíveis contrastes.

A Tabela 27 exibe as possibilidades e ocorrências das características do *tap* na forma de onda e no espectrograma, para a posição final. Na referida tabela, encontram-se os prováveis contrastes encobertos observados na fala dos sujeitos transversais.

---

**POSSIBILIDADES E OCORRÊNCIAS DE CONTRASTES ENCOBERTOS**


---

ST1							
FORMA DE ONDA			ESPECTROGRAMA			Contraste	
	(i)	(ii)	(iii)	(i)	(ii)	(iii)	
<b>A</b>	2/9	0/9	2/9	2/9	0/9	2/9	<b>2/9</b>
<b>I</b>	3/7	1/7	2/7	3/7	1/7	2/7	<b>3/7</b>
<b>U</b>	1/9	0/9	1/9	1/9	0/9	1/9	<b>1/9</b>
<b>TOTAL</b>	<b>6/25</b>	<b>1/25</b>	<b>5/25</b>	<b>6/25</b>	<b>1/25</b>	<b>5/25</b>	<b>6/25</b>
ST2							
<b>A</b>	0/1	0/1	0/1	0/1	0/1	0/1	<b>0/1</b>
<b>I</b>	0/6	0/6	0/6	0/6	0/6	0/6	<b>0/6</b>
<b>U</b>	0/2	0/2	0/2	0/2	0/2	0/2	<b>0/2</b>
<b>TOTAL</b>	<b>0/9</b>	<b>0/9</b>	<b>0/9</b>	<b>0/9</b>	<b>0/9</b>	<b>0/9</b>	<b>0/9</b>
ST3							
<b>A</b>	0/0	0/0	0/0	0/0	0/0	0/0	<b>0/0</b>
<b>I</b>	0/0	0/0	0/0	0/0	0/0	0/0	<b>0/0</b>
<b>U</b>	0/2	0/2	0/2	0/2	0/2	0/2	<b>0/2</b>
<b>TOTAL</b>	<b>0/2</b>	<b>0/2</b>	<b>0/2</b>	<b>0/2</b>	<b>0/2</b>	<b>0/2</b>	<b>0/2</b>
ST4							
<b>A</b>	0/8	0/8	0/8	0/8	0/8	0/8	<b>0/8</b>
<b>I</b>	3/9	2/9	0/9	3/9	2/9	0/9	<b>3/9</b>
<b>U</b>	1/8	0/8	0/8	1/8	1/8	1/8	<b>1/8</b>
<b>TOTAL</b>	<b>4/25</b>	<b>2/25</b>	<b>0/25</b>	<b>4/25</b>	<b>3/25</b>	<b>1/25</b>	<b>4/25</b>

Tabela 27 - Possibilidades e ocorrências de contrastes encobertos a partir das características do *tap*, sendo (i) ausência de energia, (ii) ruído transiente ou soltura da oclusão e (iii) elemento vocálico na fala dos sujeitos transversais, para posição final

Ao analisar a Tabela 27, percebe-se que, bem como para posição medial, a ausência de energia foi a característica do *tap* que mais corroborou para a observação de possíveis contrastes encobertos, seguida da presença de elementos vocálicos em posição final. Para a observação de tais características, o espectrograma e a forma de onda contribuíram de modo similar.

Portanto, a ausência de energia e a presença de elemento vocálico foram as características que favoreceram a observação de produções gradientes. A partir disso, é possível inferir que a maioria dos sujeitos transversais apresentem dificuldades em regular a magnitude do gesto de ponta de língua, pois é possível observar o correlato acústico, um espaçamento em branco no espectrograma, do momento de breve obstrução. A baixa recorrência dos correlatos visuais da soltura da oclusão ou de

ruídos transientes, podem apontar uma falha durante o afastamento dos articuladores. Posto isto, é possível que o segmento se torne inaudível devido a execução imprecisa dos movimentos articulatórios.

A fim de condensar as informações das Tabelas 26 e 27, foi elaborada a Tabela 28 que apresenta as possibilidades e ocorrências de prováveis contrastes encobertos de acordo com o ambiente vocálico e com a posição do rótico na sílaba.

POSSIBILIDADES E OCORRÊNCIAS DE CONTRASTES ENCOBERTOS						
	MEDIAL			FINAL		
	[a]	[i]	[u]	[a]	[i]	[u]
<b>ST1</b>	0/9	2/5	5/9	2/9	3/7	1/9
<b>ST2</b>	3/8	1/5	0/4	0/1	0/6	0/2
<b>ST3</b>	0/0	4/6	0/0	0/0	0/0	0/2
<b>ST4</b>	1/9	2/5	0/8	0/8	3/9	1/8
<b>TOTAL</b>	<b>4/26</b>	<b>9/21</b>	<b>5/21</b>	<b>2/18</b>	<b>6/22</b>	<b>2/21</b>
	<b>18/68</b>			<b>10/61</b>		

Tabela 28 - Lista de possibilidades e ocorrências de prováveis contrastes encobertos na fala dos sujeitos transversais, de acordo com a posição silábica e o contexto silábico em posições medial e final

A posição medial foi o contexto que favoreceu a ocorrência de produções gradientes. Das 68 possibilidades de produções, 18 foram identificadas como prováveis contrastes encobertos, o que representa cerca de 26,47% das articulações. A posição final apresentou um índice menor de aparições do fenômeno, 10 dados foram identificados com traços residuais de um possível *tap*, cerca de 16,39% das produções. Menos possibilidades de produções gradientes são sinalizadas para ST3, pois esse sujeito apresenta um alto índice de produções de róticos. As quatro ocorrências de contrastes encobertos estão presentes em posição medial. Provavelmente, isso ocorra, pois, para essa posição, o rótico ainda não emergiu completamente, uma vez que essa foi a posição que menos apresentou a ocorrência do segmento, de acordo com a Tabela 25. O mesmo ocorre com ST2, porém, em menor proporção. Esse informante apresenta um menor índice de produções do rótico em ambas as posições, no entanto, a posição medial parece favorecer a articulação de possíveis produções gradientes. Destaca-se que, para o sujeito longitudinal, a posição medial também foi favorecedora para a visualização de contrastes encobertos.

Ao longo dessa seção, serão expostos alguns dados acústicos dos sujeitos transversais, de forma a identificar como se apresentam as produções gradientes. Para ilustrar a inspeção visual dos dados, em posição medial, seguem as Figuras 27, 28, 29 e 30.

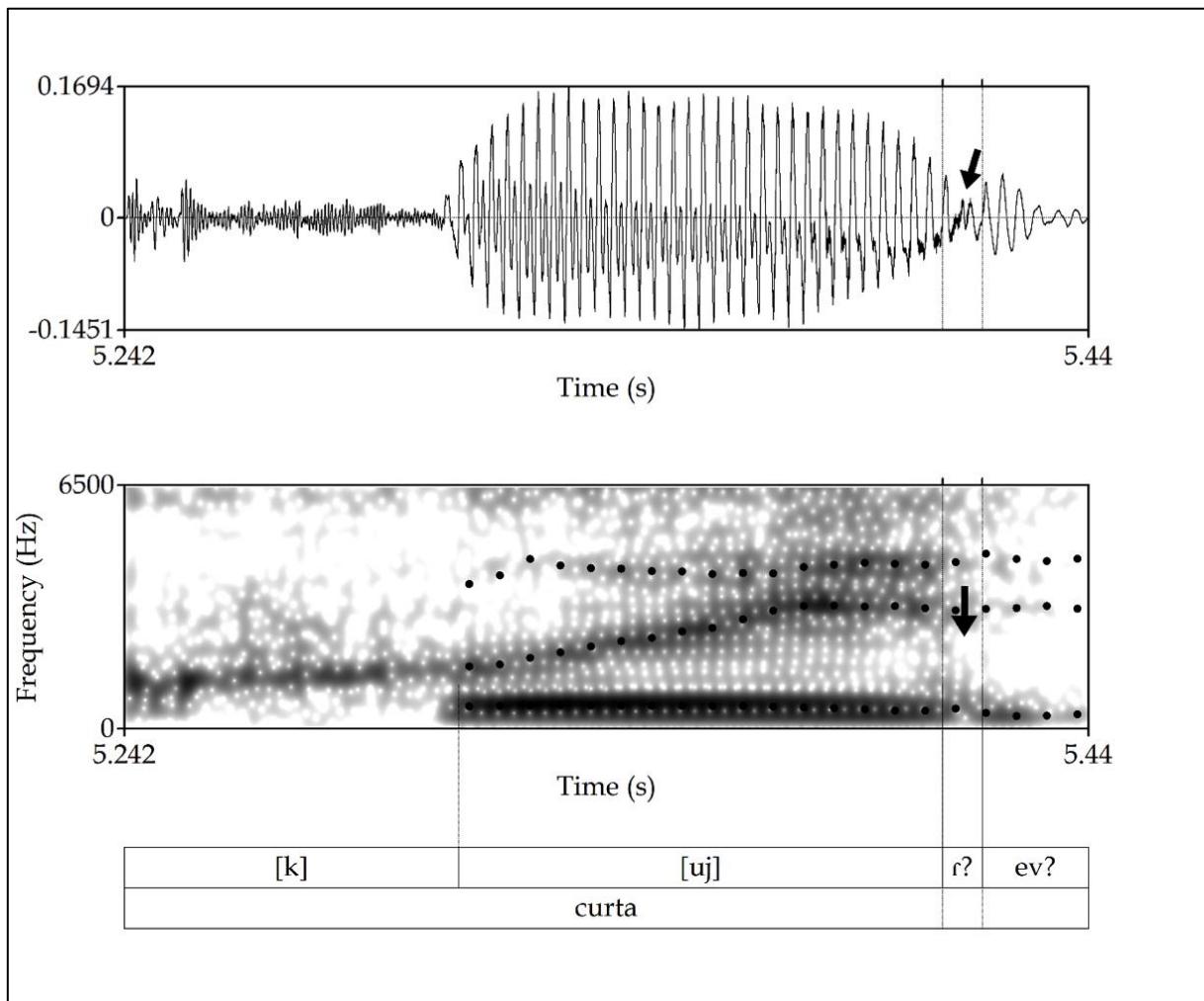


Figura 27 - Forma de onda e espectrograma de banda larga da produção de [‘kuj] do alvo "curta" por ST1. Em destaque pela seta o contraste encoberto

Todos os juízes perceberam a produção do alvo *curta* de ST1 como [‘kuj.te]. Ao observar o espectrograma e a forma de onda da Figura 27, é possível observar que esse dado apresenta a produção de um ditongo decrescente, seguido de um possível contraste encoberto. Há a transição gradual do primeiro e segundo formantes entre os dois segmentos vocálicos que compõem o ditongo. A vogal [u] apresenta F1 de 617 Hz e F2 de 2304 Hz, enquanto o glide [i] exibe F1 de 522 Hz e F2 de 3263 Hz. Embora o *tap* não tenha sido percebido auditivamente, dois correlatos acústicos permitem a inferência de uma eventual tentativa de articulação. Destaca-se, inicialmente, a queda de amplitude da onda sonora, apontada pela seta na parte

superior da figura, enquanto no espectrograma é possível perceber uma pequena faixa quase em branco, sinalizada pela seta na parte inferior da mesma figura. Além disso, nota-se um momento acústico de estrutura formântica contínua e similar a semivogal [j]. Sendo assim, a produção do ditongo decrescente compreende um movimento articulatório para a produção de [u] e, de forma contínua, uma anteriorização e elevação da ponta da língua para uma configuração próxima à vogal [i]. Do mesmo modo, a produção de um *tap* envolve a elevação, em direção aos alvéolos, da ponta da língua. A revelação do contraste encoberto indica que há uma tentativa de articulação do [r], porém, a criança ainda necessita refinar o gesto de ponta de língua para atingir a produção do segmento alvo, evitando, assim, a ocorrência de semivocalizações.

Na Figura 28 é possível observar o espectrograma gerado para o alvo *círculo*, produzido por ST2.

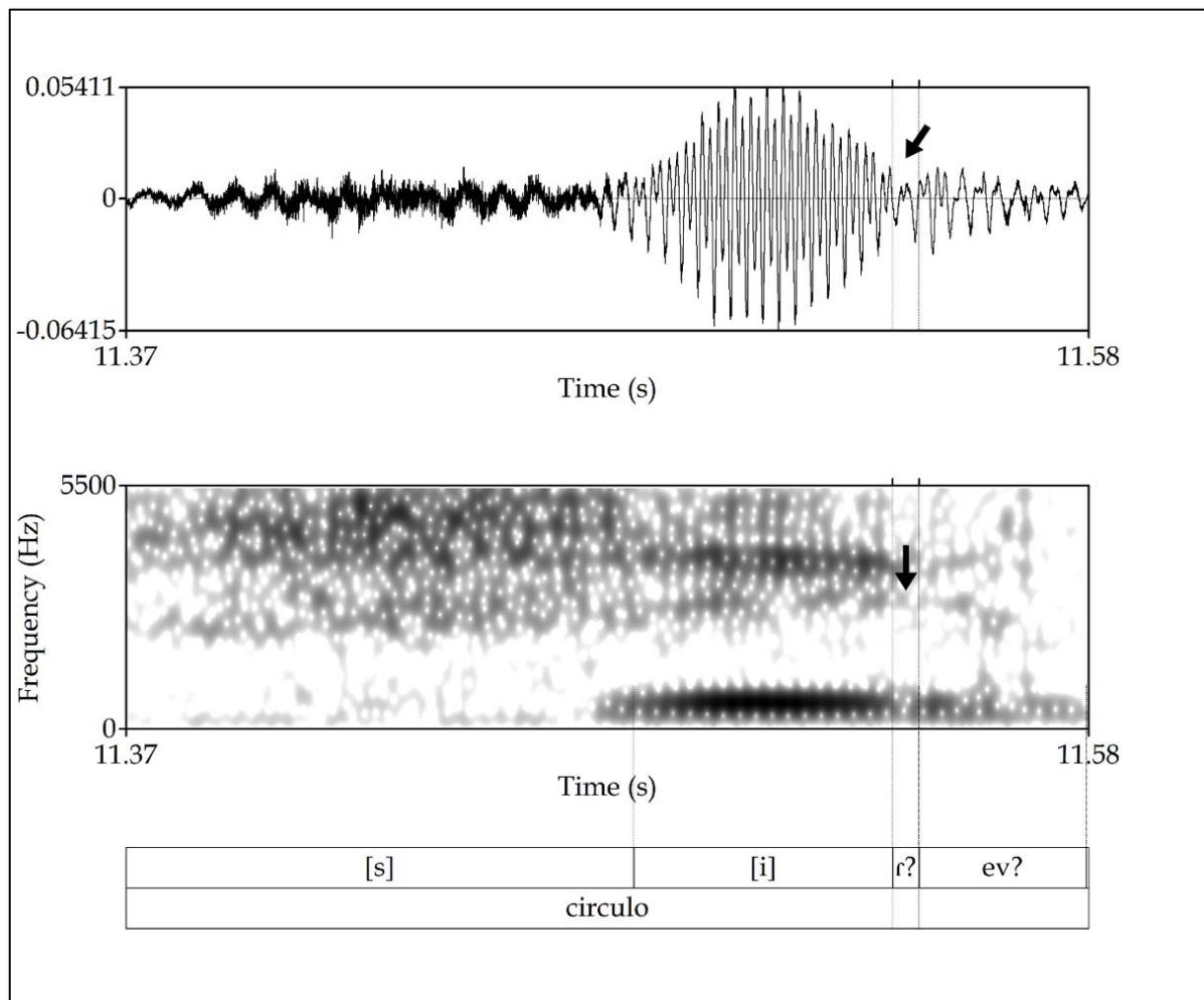


Figura 28 - Forma de onda e espectrograma de banda larga da produção de [si] do alvo "círculo" por ST2. Em destaque pela seta o contraste encoberto

Conforme a Figura 28, observa-se um espaçamento em branco, sinalizado pela seta. Destacado pelas linhas pontilhadas, o momento de pouca energia apresenta duração de apenas 5 ms. No oscilograma, a seta aponta a queda de amplitude na forma de onda, em seguida, percebe-se um momento acústico que aparenta ser a porção final da vogal. Embora essas pistas visuais indiquem a articulação de um *tap* ao final da primeira sílaba, esse alvo foi transscrito como ['si.ku.ru], evidenciando, assim, a ocorrência de uma metátese CV. Logo, infere-se que há uma tentativa de produção do segmento ainda em posição de coda, não apenas a sua realização em *onset*, na sílaba final, resgatada de oitiva.

O dado parece sugerir investigações futuras no que concerne à presença de contrastes encobertos para os casos de metáteses encontrados no transcorrer da aquisição fonológica das estruturas CVC e CCV.

Na Figura 29, ao analisar a forma de onda e o espectrograma da palavra *circo*, produzida por ST3, percebe-se uma forte queda de amplitude, sinalizada pela seta, seguida por um momento acústico de estrutura formântica contínua e similar a uma vogal.

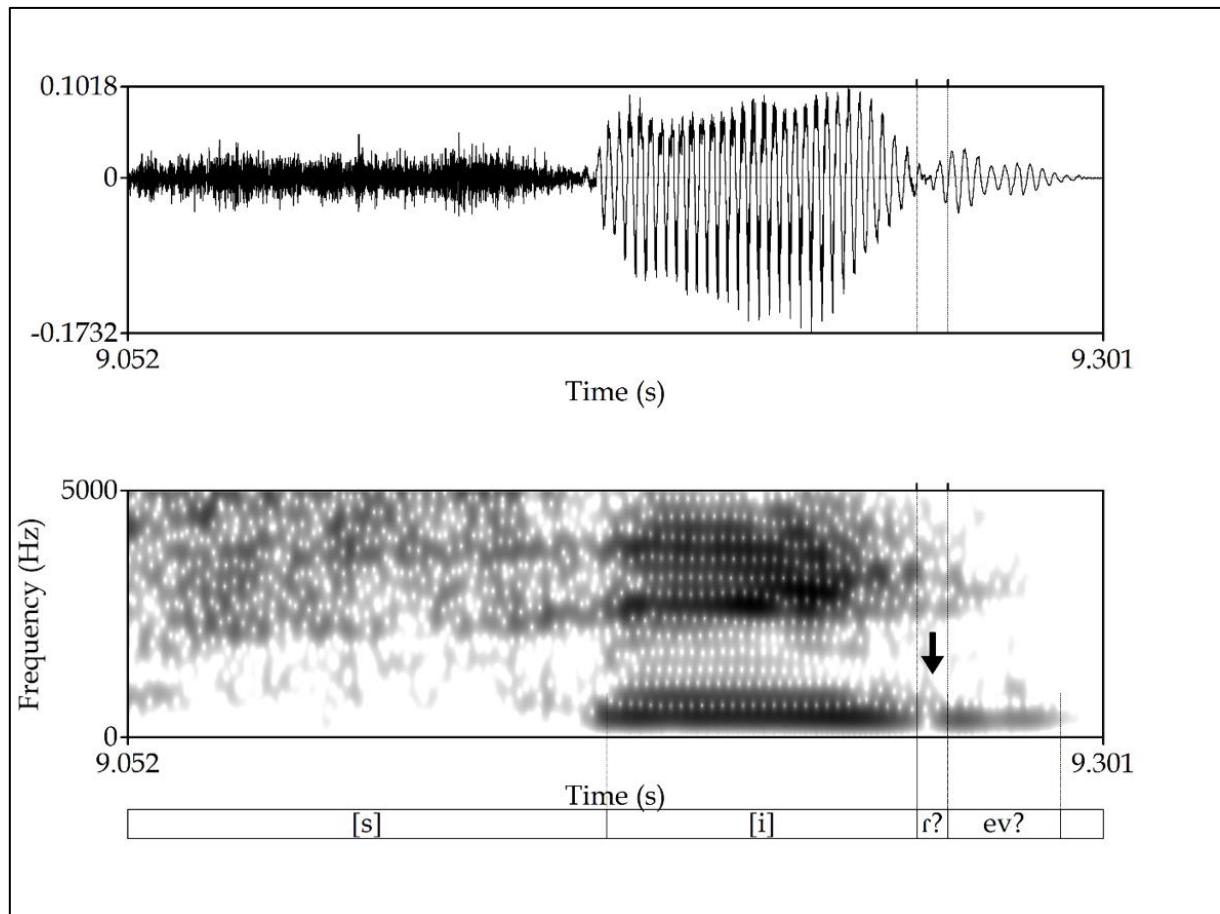


Figura 29 - Forma de onda e espectrograma de banda larga da produção de [‘si] do alvo "circo" por ST3. Em destaque pela seta o contraste encoberto

O momento de pouca energia, destacado pelas linhas pontilhadas, na Figura 29, apresenta uma duração muito curta, apenas 7 ms. Apesar dessas pistas visuais apontarem a articulação de um [r], a análise de oitiva não constatou a produção do segmento, decisão unânime entre os três juízes. Sendo assim, infere-se que a criança execute de forma imprecisa o movimento de ponta de língua, sendo necessário um ajuste na magnitude no *timing* do gesto para esse ser percebido auditivamente.

A análise visual do alvo *carpa*, produzido por ST4, indicou pistas de um contraste encoberto, como pode ser observado na Figura 30.

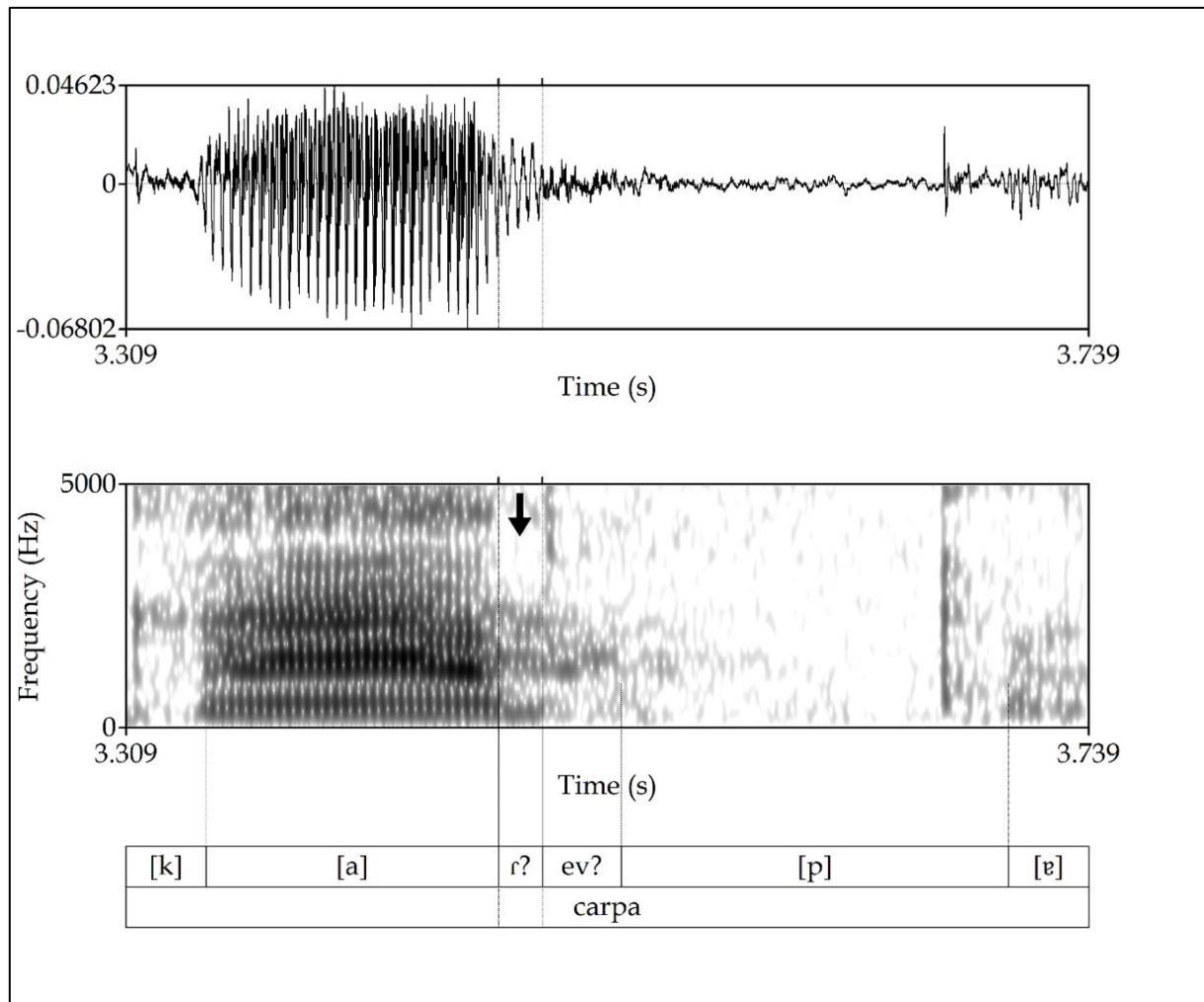


Figura 30 - Forma de onda e espectrograma de banda larga da produção de ['ka.pe] do alvo "carpa" por ST4. Em destaque pela seta o contraste encoberto

A seta aponta um espaçamento em branco no espectrograma, o momento de ausência de energia. Em seguida, nota-se um momento acústico similar à porção final da vogal. Embora essas pistas visuais apontem a produção de um *tap*, esse segmento não foi percebido durante o julgamento de oitiva, sendo transscrito como ['ka.pe].

A partir desse momento, serão apresentados os resultados referentes à inspeção visual dos dados em posição final. Para o sujeito ST2 não foram identificados possíveis contrastes encobertos em final de palavra. Sendo assim, seguem as Figuras 31 e 32.

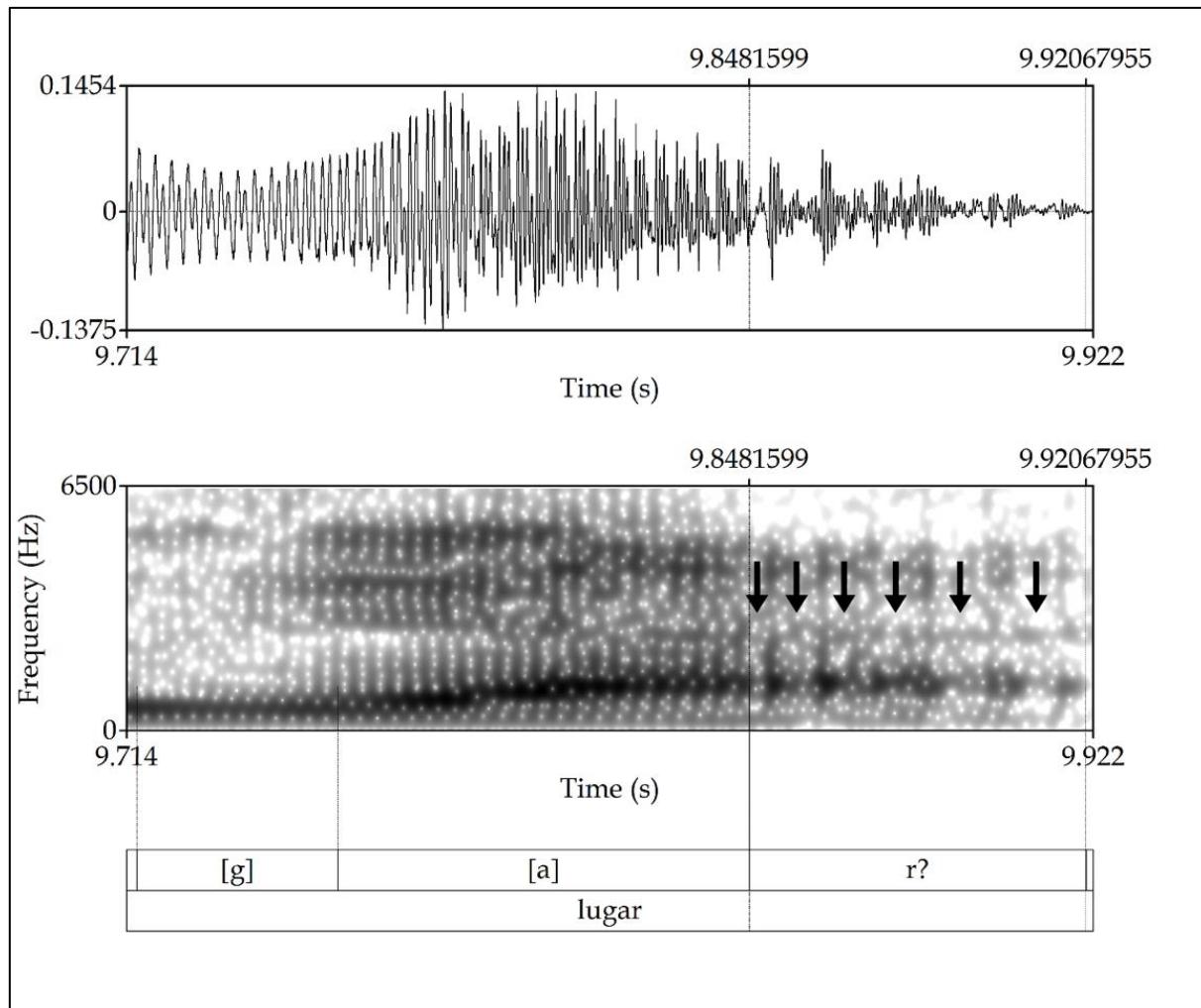


Figura 31 - Forma de onda e espectrograma de banda larga da produção de ['ga] do alvo "lugar" por ST1. Em destaque pelas setas o contraste encoberto

Ao observar a forma de onda e o espectrograma do alvo *lugar*, produzido por ST1, a partir da Figura 31, notam-se pistas visuais da produção de uma vibrante. As setas indicam as várias faixas verticais com menor energia, ainda que não estejam em branco. Esses períodos de vibração parecem evidenciar traços residuais da variante rótica. No entanto, a vibrante não foi detectada pelos três juízes. Sendo assim, é possível inferir que a criança tenta alcançar a produção do alvo, embora não tenha sucesso.

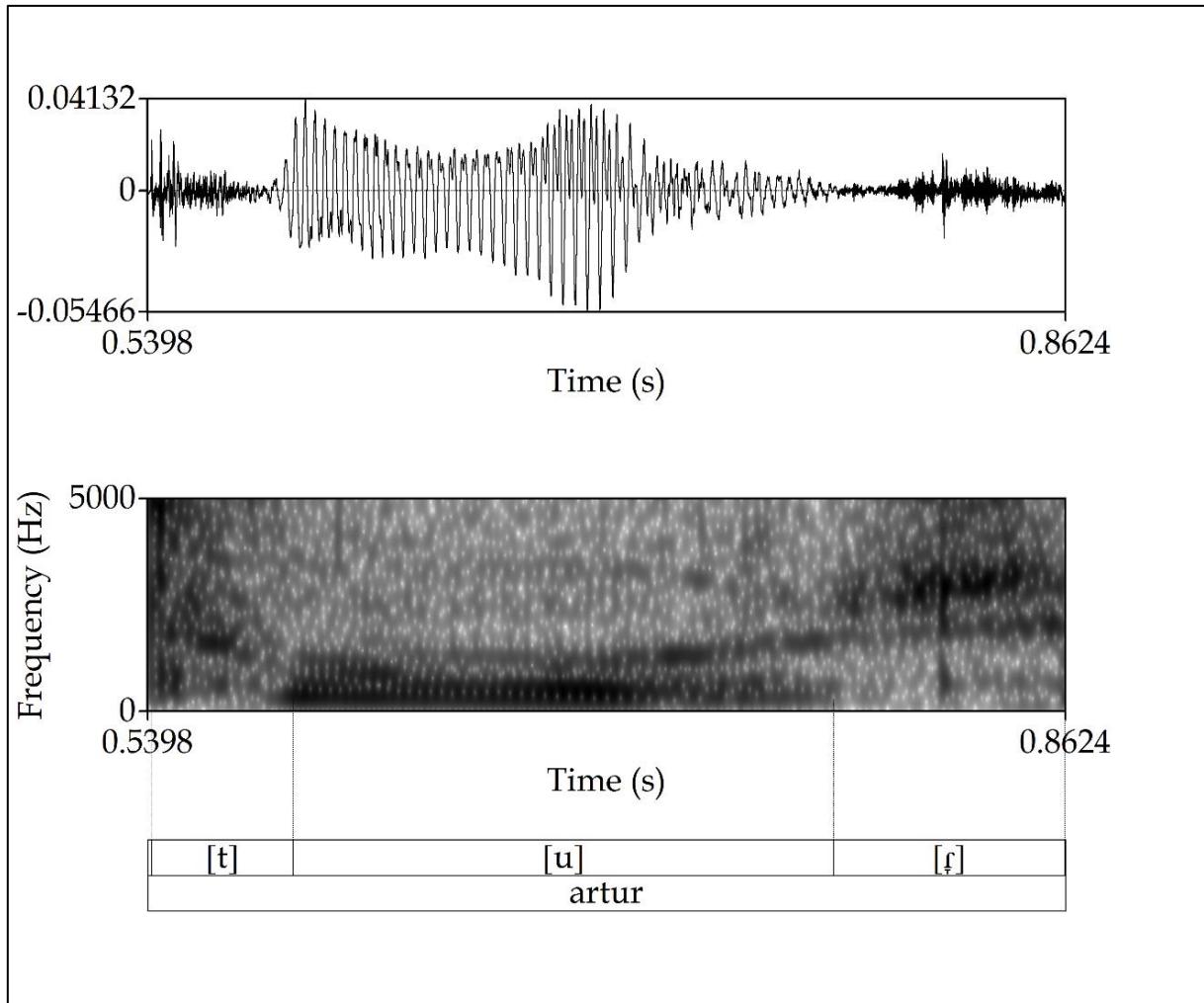


Figura 32 - Forma de onda e espectrograma de banda larga da produção de [‘tur̥] do alvo "Artur" por ST3

A Figura 32 exibe a forma de onda e o espectrograma da produção de um *tap* fricativizado. Conforme Bilharva-da-Silva (2019), é possível identificar a produção desse tipo de segmento ao verificar uma maior concentração de energia nas regiões de alta frequência no espectrograma. Essa marca acústica parece evidenciar a fricção causada durante a passagem de ar.

A inspeção visual dos dados dos sujeitos transversais revela que foram percebidos possíveis contrastes encobertos quando a produção do *tap* não foi audível na análise dos juízes, especialmente em posição medial. Por meio das pistas acústicas encontradas, é possível inferir que a coordenação dos gestos ainda precisa ser ajustada, uma vez que há algum tipo de correlato acústico da produção do segmento, mas esse não é percebido na análise de oitiva. Isso indica que o segmento está emergindo na fala das crianças.

#### 4.2.3.2 Aspectos Duracionais

A análise dos aspectos duracionais, conforme já reportado, tem como um de seus propósitos identificar o tipo de som articulado ou determinar diferenças e semelhanças entre os pares de palavras produzidos pelas crianças. Tendo em vista que alguns sujeitos não produziram o *tap*, faz-se necessário dois tipos de análises. Nos casos em que o rótico foi produzido, essas produções são comparadas com os dados do sujeito controle. Para os casos de omissão e semivocalização, é observado o comportamento da vogal núcleo da sílaba ou do ditongo e, posteriormente, comparado com dados da mesma natureza do sujeito controle. Inicialmente, essa subseção apresenta os resultados das análises feitas para a posição medial, observando as produções de róticos, omissões e semivocalizações. Em seguida, dar-se-á o mesmo para os resultados referentes à posição final.

A partir das referidas considerações, a Tabela 29 apresenta as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição medial, abarcavam o rótico como alvo – carpa, farpa, carpa, circo, círculo, curta, curto, turco. Na mesma tabela, são exibidos os valores de duração relativa do rótico em relação à sílaba. Os valores reportados a seguir são referentes às produções nas quais os sujeitos transversais articulam um *tap*.

CODA MEDIAL					
		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(rótico - sílaba)
AR	ST1	—	—	—	—
	ST2	222	20	123	9,01
	ST3	280	15	124	5,57
	ST4	—	—	—	—
	<b>MÉDIA</b>	<b>251</b>	<b>18</b>	<b>123</b>	<b>7,29</b>
IR	DP	<b>40,84</b>	<b>3,27</b>	<b>0,65</b>	<b>2,43</b>
	ST1	—	—	—	—
	*ST2	300	15	142	5,00
	ST3	—	—	—	—
	ST4	—	—	—	—
UR	<b>MÉDIA</b>	<b>300</b>	<b>15</b>	<b>142</b>	<b>5,00</b>
	DP	—	—	—	—
	ST1	—	—	—	—
	ST2	233	21	138	9,05
	ST3	192	23	102	12,43
	<b>MÉDIA</b>	<b>212</b>	<b>22</b>	<b>120</b>	<b>10,74</b>
	DP	<b>29,13</b>	<b>1,47</b>	<b>25,67</b>	<b>2,39</b>

Tabela 29 - Médias de duração absoluta e relativa dos alvos que contém rótico (articulados como *tap*) em posição medial dos sujeitos transversais, conforme contexto vocálico

\*Valores obtidos com base em uma única produção, assim, não constitui média.

Com base nos valores exibidos na Tabela 29, verifica-se que o rótico apresenta valores médios de duração absoluta entre 15 e 23 ms, ocupando de 5,00% a 12,43% da sílaba. Em contexto de AR e UR, apenas os sujeitos ST2 e ST3 produziram o segmento. Para IR, somente ST2 articulou o segmento em uma única produção, o alvo *círculo*. Ao observar os valores de duração da vogal núcleo da sílaba, nota-se que a vogal baixa apresenta 123 ms e 124 ms de duração absoluta e a vogal alta posterior, 102 ms e 138 ms.

O Gráfico 48 apresenta as médias de duração absoluta da sílaba, da vogal e do rótico das produções das palavras em contexto de AR (carta, farpa e carpa) realizadas por S0, ST2 e ST3, em posição medial. Os valores reportados para o rótico são referentes às produções nas quais as crianças articulam um *tap*.

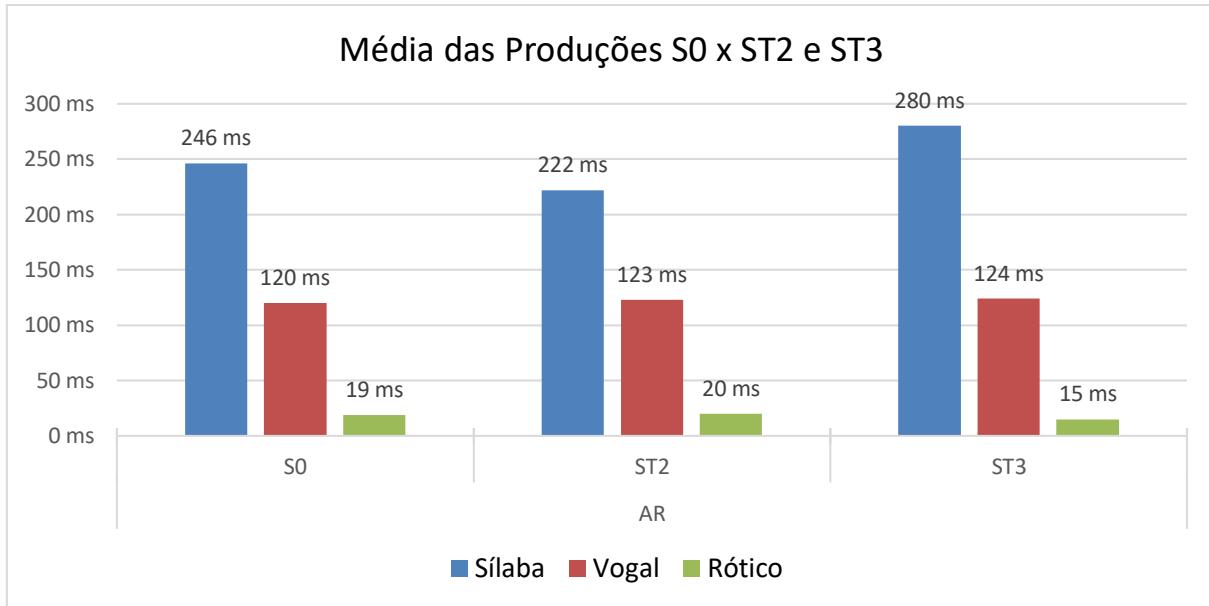


Gráfico 48 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de S0, ST2 e ST3 em contexto vocalico de AR, para a posição medial

Ao comparar as médias de duração absoluta da sílaba entre o sujeito controle e os sujeitos transversais, a partir do Gráfico 48, constata-se que ST3 apresenta a maior média de duração da sílaba, o que parece estar simplesmente associado à duração da consoante inicial, possivelmente o segmento fricativo. As médias apresentadas para as produções da vogal nuclear da sílaba são muito aproximadas, bem como às relativas ao rótico. Quanto à duração do rótico, pode-se afirmar que os sujeitos transversais produziram o *tap* de forma convencional, uma vez que as médias de duração, 15 ms e 20 ms, aproximam-se da média reportada pelo sujeito controle, 19 ms.

O Gráfico 49 exibe as médias de duração absoluta da sílaba, da vogal e do rótico das produções dos itens lexicais em contexto de UR (curto, curta, turco) realizadas por S0, ST2 e ST3, em posição medial. Os valores reportados para o rótico são referentes às produções nas quais as crianças articulam um *tap*.

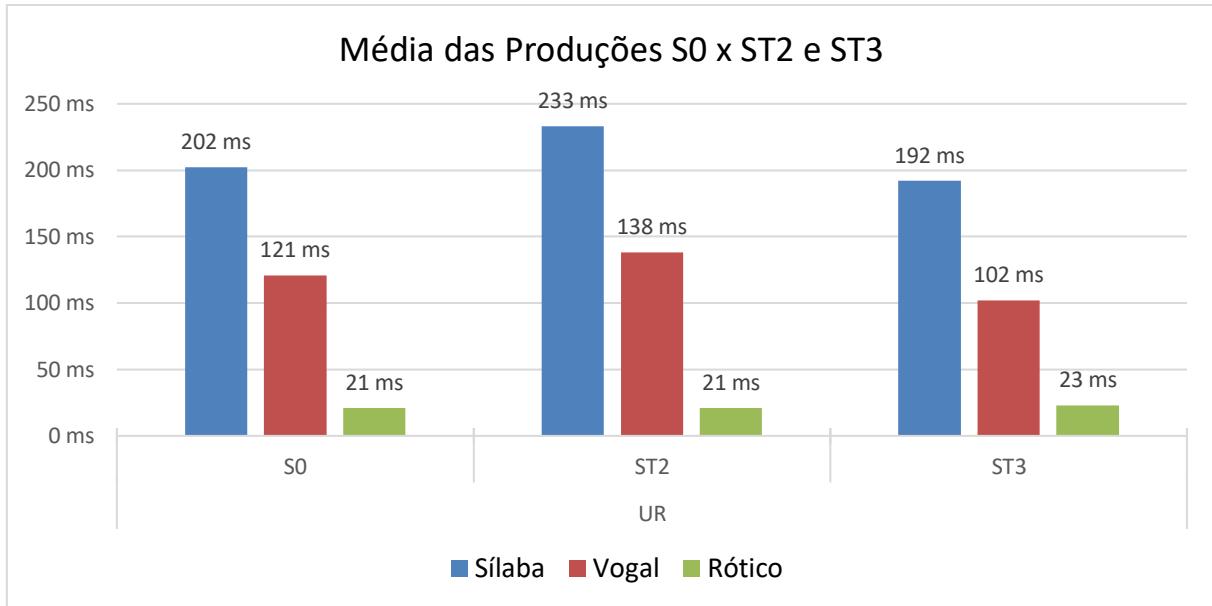


Gráfico 49 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de S0, ST2 e ST3 em contexto vocálico de UR, para a posição medial

A análise do Gráfico 49 evidencia que, ao contrário do que ocorreu para as médias de duração da sílaba em contexto de AR, as produções de ST2 são maiores do que as reportadas para o sujeito controle, o que, provavelmente, esteja associado a uma maior duração da vogal nuclear, maior para ST2. Por fim, ao observar as produções do rótico dos sujeitos transversais, pode-se afirmar que esses produziram um *tap* com duração muito aproximada, entre 21 ms e 23 ms.

O Gráfico 50 reúne os resultados das médias de duração relativa do *tap* em relação à sílaba das produções, em contexto AR e UR, realizadas por S0, ST2 e ST3.

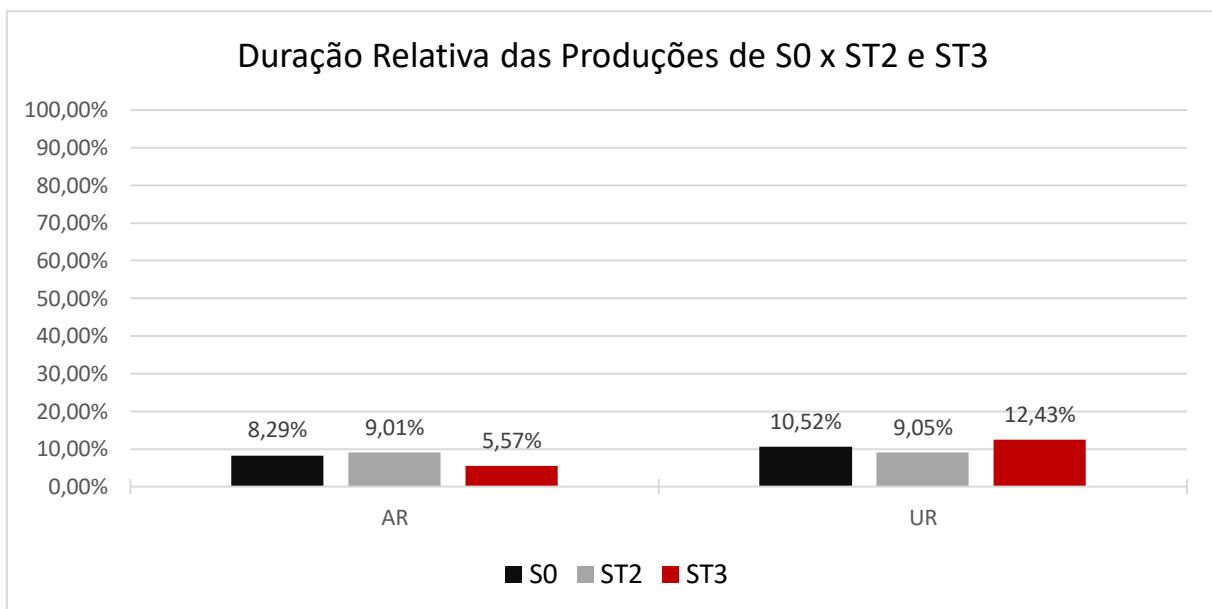


Gráfico 50 - Médias de duração relativa do rótico de S0 (preto), ST2 (cinza) e ST3 (vermelho), em contexto de AR e UR, para posição medial

Ao observar os dados de ST2 e ST3, presentes no Gráfico 50, é possível verificar que os índices de porcentagens reportados para as produções de ST2 se aproximam dos valores apresentados por S0. O sujeito ST3, no entanto, apresenta maiores diferenças na duração relativa do rótico, especialmente no contexto de AR. Tal fato parece estar associado a uma maior duração do segmento fricativo, e da sílaba, consequentemente, não a uma dificuldade em si na realização do *tap*.

Os róticos presentes nos itens lexicais compostos por [a], produzidos pelos sujeitos transversais, ocupam 9,01% e 5,57% da sílaba, enquanto o rótico articulado pelo sujeito controle ocupa 8,29% da sílaba. As palavras compostas por [u] exibem percentuais ligeiramente maiores, em relação ao contexto de [a]. O *tap* produzido por ST2 ocupa 9,05% da sílaba e o segmento articulado por ST3 preenche 12,43% da sílaba, ao passo que o [r] produzido por S0 apresenta um índice de 10,52% de duração relativa.

Assim, com base nos resultados reportados, é possível observar que, quando os sujeitos ST2 e ST3 articularam um *tap*, em contexto de AR e UR, suas produções apresentam-se conforme o esperado para a produção do segmento, uma vez que os valores de duração se aproximam das médias reportadas para o sujeito controle.

Após a análise dos dados, nos quais os sujeitos transversais produzem o *tap*, a Tabela 30 apresenta os valores referentes às produções nas quais as crianças

omitem o segmento. A fim de observar o comportamento da vogal núcleo da sílaba, a referida tabela exibe as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição medial, abarcavam o rótico como alvo – *carpa*, *farpa*, *carpa*, *circo*, *círculo*, *curta*, *curto*, *turco*. Em seguida são exibidos os valores de duração relativa da vogal em relação à sílaba.

CODA MEDIAL					
		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(vogal - sílaba)
AR	*ST1	473	–	264	55,81
	ST2	241	–	183	76,61
	ST3	–	–	–	–
	ST4	264	–	172	68,58
	<b>MÉDIA</b>	<b>326</b>	–	<b>206</b>	<b>67,00</b>
IR	DP	<b>127,90</b>	–	<b>50,36</b>	<b>10,49</b>
	ST1	247	–	101	45,19
	ST2	280	–	140	47,26
	ST3	233	–	78	34,67
	ST4	292	–	132	45,44
UR	<b>MÉDIA</b>	<b>263</b>	–	<b>113</b>	<b>43,14</b>
	DP	<b>27,53</b>	–	<b>28,52</b>	<b>5,72</b>
	ST1	183	–	162	88,74
	*ST2	126	–	112	88,89
	ST3	–	–	–	–
	ST4	185	–	164	88,52
	<b>MÉDIA</b>	<b>165</b>	–	<b>146</b>	<b>88,72</b>
	DP	<b>33,35</b>	–	<b>29,40</b>	<b>0,19</b>

Tabela 30 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos com o recurso de omissão pelos sujeitos transversais, conforme contexto vocálico

\* Valores obtidos com base em uma única produção, assim, não constitui média

Na Tabela 32, é possível perceber que ocorreram omissões em todos os contextos vocálicos para ST1, ST2 e ST4. Para ST3, não foram observadas omissões em AR e UR, uma vez que esse sujeito produziu o *tap* em todas as possibilidades de ocorrência nesses ambientes vocálicos, mas quando o rótico estava acompanhado da vogal alta anterior, todas as produções foram articuladas sem o segmento. O sujeito ST1 omitiu o *tap* uma única vez para o alvo *farpa*. Da mesma forma, ST2 apresenta uma única omissão para o alvo *curto*. Ainda, o sujeito ST4 omitiu apenas uma produção para os alvos *carpa* e *farpa*.

Os itens lexicais compostos pela vogal [a] apresentam valores médios de duração da vogal entre 172 ms e 264 ms. A vogal [i], oriunda da omissão do *tap*, apresenta uma média de 78 ms a 140 ms. Para a vogal [u], foram reportados valores médios de 112 ms a 164 ms de duração absoluta. Quanto à duração relativa da vogal em relação à sílaba, a vogal alta anterior apresenta os menores percentuais de duração, de 34,67% a 47,26%.

A fim de comparar a duração das vogais realizadas nas produções que continham omissões de [r] e a duração das vogais produzidas em itens lexicais com o molde silábico CV, a Tabela 31 apresenta as médias de duração absoluta da sílaba e da vogal nuclear das palavras que, em posição medial, incluíam a vogal em alvos CV – cata, Chico, cuca. A mesma tabela exibe os valores de duração relativa da vogal em relação à sílaba.

CODA MEDIAL				
		Duração absoluta (ms)		Duração relativa (%) (vogal - sílaba)
		Sílaba	Vogal	
A	ST1	187	148	75,70
	ST2	228	137	58,65
	ST3	139	86	62,31
	ST4	127	87	68,78
	<b>MÉDIA</b>	<b>170</b>	<b>114</b>	<b>66,36</b>
I	DP	<b>46,72</b>	<b>32,79</b>	<b>7,50</b>
	ST1	240	79	32,02
	ST2	221	85	38,22
	ST3	272	91	33,00
	ST4	323	133	41,22
U	<b>MÉDIA</b>	<b>264</b>	<b>97</b>	<b>36,11</b>
	DP	<b>44,52</b>	<b>24,44</b>	<b>4,36</b>
	ST1	115	69	61,29
	ST2	100	60	71,40
	ST3	129	88	67,93
	ST4	154	119	76,42
	<b>MÉDIA</b>	<b>125</b>	<b>84</b>	<b>69,33</b>
	DP	<b>23,05</b>	<b>25,92</b>	<b>6,39</b>

Tabela 31- Médias de duração absoluta e relativa da vogal dos alvos CV em posição medial produzidos pelos sujeitos transversais, conforme contexto vocálico

Com base nos valores apresentados na Tabela 31, percebe-se, em contexto de vogais altas, todas as médias reportadas por ST4 são maiores do que as médias

dos demais sujeitos. A vogal baixa exibe uma média de 86 ms a 148 ms, a vogal alta anterior apresenta valores médios de 79 ms a 133 ms, ao passo que, para a vogal alta posterior, é reportado de 60 ms a 119 ms de duração absoluta. Os dados referentes à duração relativa da vogal em relação à sílaba apontam que a vogal [i] ocupa a menor duração, preenchendo a sílaba em 32,02% a 38,22%. Tal fato está relacionado à presença da consoante fricativa em *Chico*, a qual ocupa proporção bem maior do que as plosivas dos contextos com A e U.

A fim de investigar a possibilidade de que os informantes, apesar de não produzirem o [r], fossem capazes de utilizar algum recurso para marcar a distinção entre as palavras que apresentavam o *tap* como alvo e as que não continham a presença do segmento, serão apresentados os Gráficos 51, 52, 53, 54, 55 e 56. Nesses, constam algumas comparações entre as produções dos sujeitos transversais em contextos CVC e CV, por ambiente vocálico. Em seguida, nos gráficos 57, 58, 59 e 60, comparações entre os dados de cada sujeito transversal e os dados do sujeito controle.

Tendo por base as considerações apresentadas, o Gráfico 51 reúne os resultados das médias de duração absoluta da sílaba e da vogal das produções realizadas por ST1, ST2 e ST4. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa e (A) cata. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais as crianças utilizaram o recurso de omissão do *tap*.

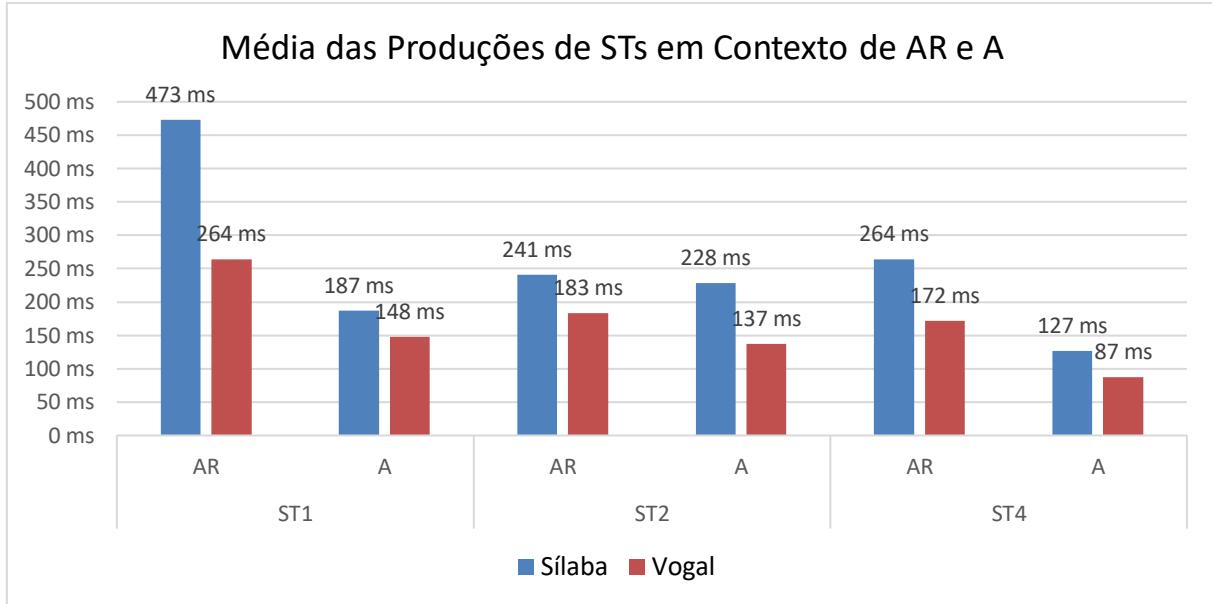


Gráfico 51 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2 e ST4, em contexto de AR (com omissão do *tap*) e A, para a posição medial

Ao observar os valores apresentados pelo Gráfico 51, é necessário ressaltar que os valores reportados para ST1 e ST4 em AR são oriundos de uma ou duas produções, como anteriormente reportado. Assim, não é possível estabelecer efetivamente uma comparação entre a duração da vogal em alvos CVC e CV. Os resultados, no entanto, demonstram que as vogais para alvos CVC são mais longas do que as do alvo CV. Os dados de ST2 revelam valores aproximados para as médias de duração da sílaba, no entanto, o mesmo não ocorre para a vogal nuclear da sílaba. A vogal baixa, em alvo CVC, apresenta uma duração de 183 ms, enquanto a vogal do alvo CV apresenta 137 ms de duração. Os resultados de ST2 poderiam, portanto, estar indicando a presença de um alongamento compensatório.

O Gráfico 52 exibe os resultados das médias de duração relativa da vogal nuclear em relação à sílaba das produções realizadas por ST1, ST2 e ST4, em contexto AR e A. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais a informante utiliza o recurso de omissão do *tap*.

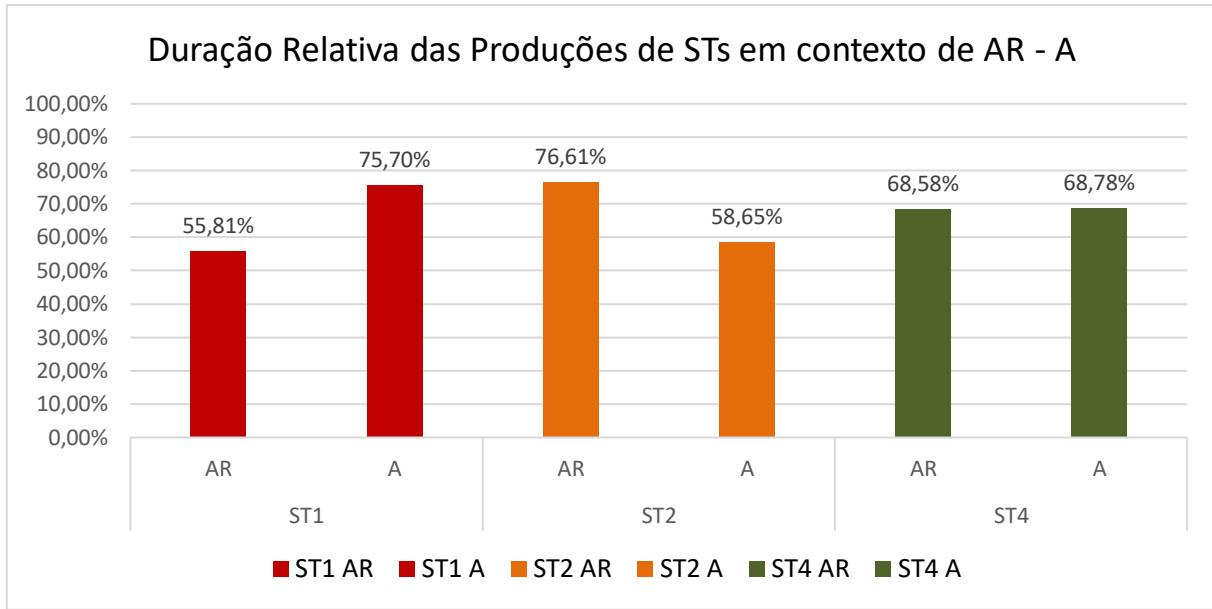


Gráfico 52 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja) e ST4 (verde), em contexto de AR e A, para posição medial

Uma vez que os dados de ST1 e ST4 não são efetivamente suscetíveis a análises comparativas, observa-se, no Gráfico 52, que há uma diferença expressiva entre os índices de porcentagem de duração relativa da vogal em relação à sílaba nos dados de ST2. As vogais resultantes das omissões do *tap*, em AR, ocupam 76,61% da sílaba, ao passo que um menor índice é reportado para as vogais em CV, 58,65%. Logo, há um indicativo de que, a partir do parâmetro de duração, essa informante estabeleça um contraste entre as vogais resultantes da omissão do segmento e as vogais produzidas em palavras sem o rótico como alvo.

O Gráfico 53 apresenta os resultados das médias de duração absoluta da sílaba e da vogal das produções das produções realizadas por ST1, ST2, ST3 e ST4. Os dados apresentados estão organizados conforme os seguintes contextos: (IR) circo, círculo e (I) Chico. Os valores referentes às palavras compostas por [r] são de produções nas quais as crianças utilizaram o recurso de omissão do *tap*.

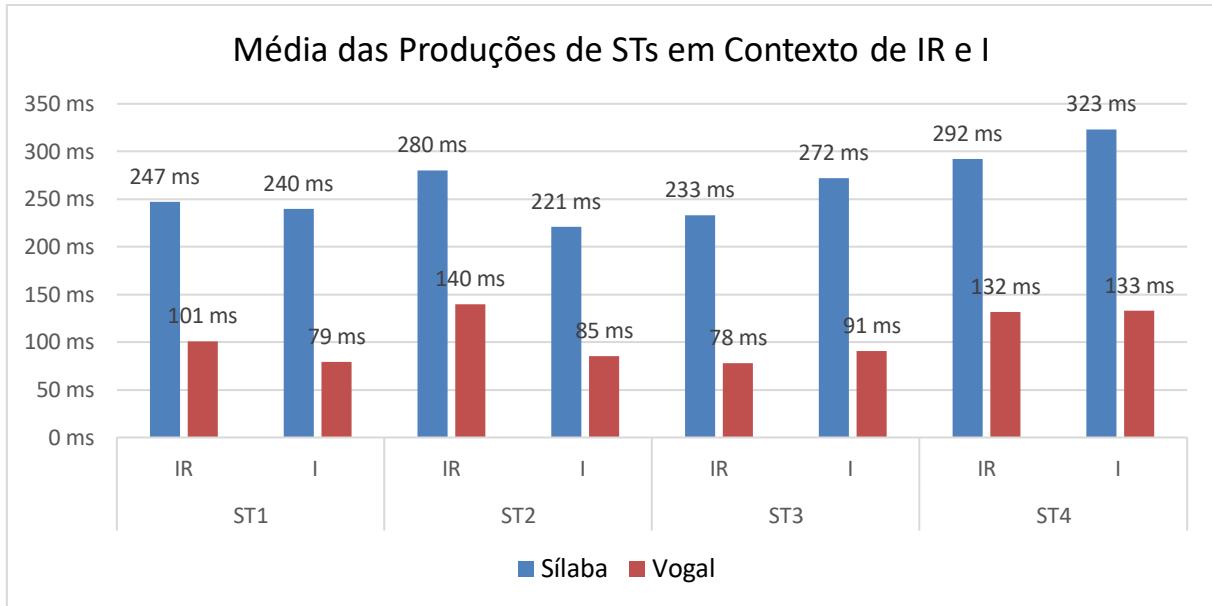


Gráfico 53 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2, ST3 e ST4, em contexto de IR (com a omissão do *tap*) e I, para a posição medial

Ao examinar os dados do Gráfico 53, identifica-se uma maior duração da vogal em contexto de IR do que em contexto de I nas produções de ST1 e ST2. Assim como reporta Mezzomo, Keske-Soares e Mota (2004), parece haver um alongamento da vogal e da consoante fricativa, tendo em vista que as palavras são compostas de consoante fricativa inicial. No que tange à duração das vogais, os itens lexicais articulados por ST1 e ST2 apresentam uma diferença considerável entre a duração da vogal nuclear de IR e I. Para os demais sujeitos, a diferença de valores é pequena, sendo quase nula para ST4.

O Gráfico 54 reúne os resultados das médias de duração relativa da vogal nuclear em relação à sílaba das produções, em contexto IR e I, realizadas por ST1, ST2, ST3 e ST4. Os valores apresentados para os itens lexicais compostos por [r] são referentes às produções nas quais a informante utiliza o recurso de omissão do *tap*.

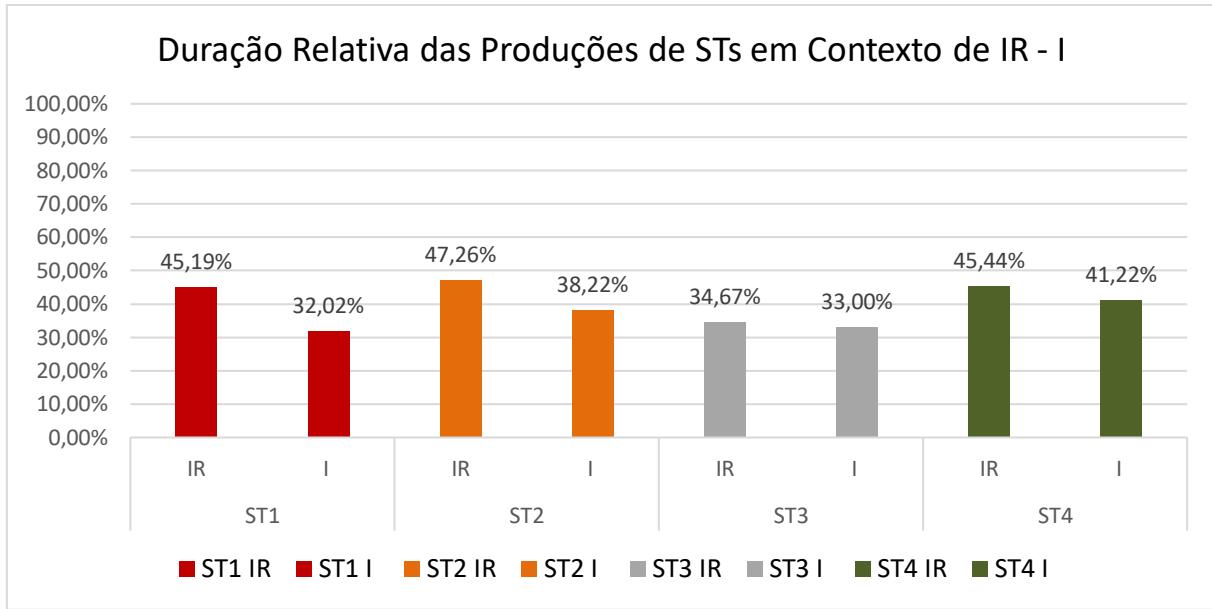


Gráfico 54 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja), ST3 (cinza) e ST4 (verde), em contexto de IR (com omissão do *tap*) e I, para posição medial

A análise do Gráfico 54 revela que os índices de porcentagens de duração relativa das vogais nas palavras em contexto de IR e I apresentam valores aproximados para ST3 e ST4. Contudo, as médias apresentadas por ST1 e ST2, para as vogais resultantes de omissão, apresentam índices de porcentagens elevados em relação às vogais de sílabas CV. As vogais dos itens lexicais compostos por CV, produzidos por ST1 e ST2, ocupam 32,02% e 38,22%, da sílaba, respectivamente. Já as vogais oriundas de omissão compreendem 45,19% e 47,26% da sílaba. Sendo assim, parece que essas crianças se valem do alongamento da vogal proveniente de omissão para estabelecer distinção entre palavras de sílabas CVC e CV.

O Gráfico 55 exibe os resultados das médias de duração absoluta da sílaba e da vogal das produções das produções realizadas por ST1, ST2 e ST4. Os dados apresentados estão organizados conforme o seguinte contexto: (UR) curta, curto, turco e (U) cuca. Os valores referentes às palavras compostas por [r] são referentes às produções nas quais os sujeitos transversais utilizaram o recurso de omissão do *tap*.

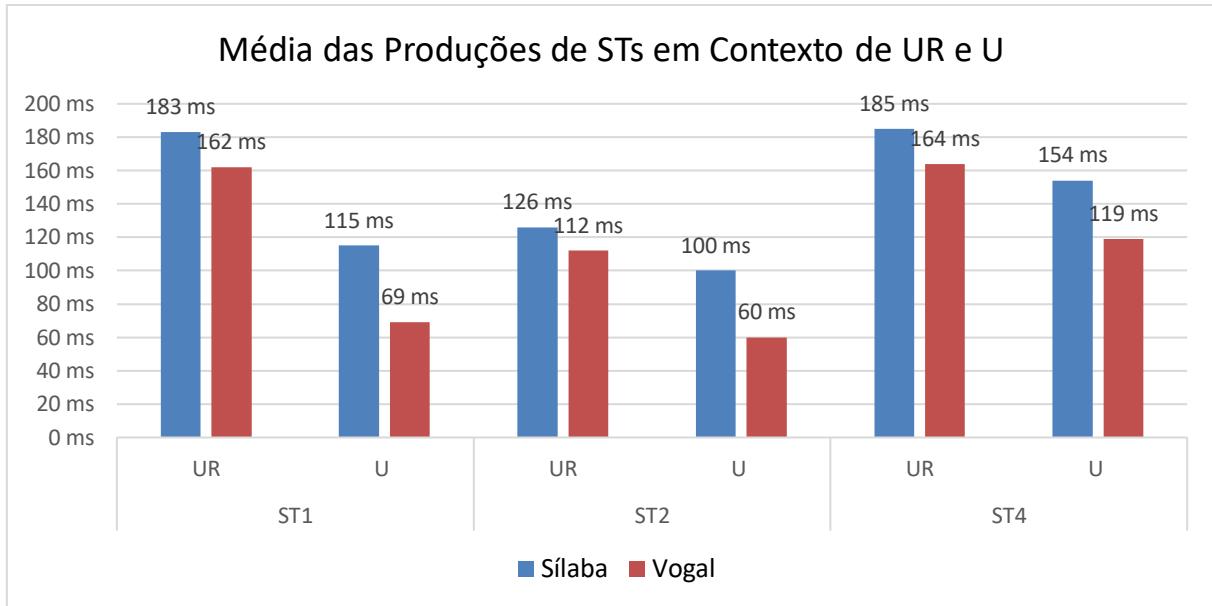


Gráfico 55 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2 e ST4, em contexto de UR (com omissão do *tap*) e U, para a posição medial

Ao observar os dados de ST2, a partir do Gráfico 55, é necessário destacar que os valores reportados para UR se referem a um dado isolado, não passível de análise comparativa. As médias de duração absoluta da sílaba e da vogal verificadas para ST4 apresentam diferença entre os alvos CVC e CV. No entanto, ao examinar os dados de ST1, é possível observar um comportamento diferente, as médias de duração da vogal e da sílaba registram uma diferença expressiva ao comparar UR e U. As palavras que deveriam ser constituídas pelo rótico apresentam duração maior, 183 ms para duração da sílaba e 162 ms para duração da vogal, enquanto os itens lexicais com alvo CV apresentam duração absoluta de 115 ms para sílaba e 69 ms para vogal.

O Gráfico 56 apresenta os resultados das médias de duração relativa da vogal nuclear em relação à sílaba das produções realizadas por ST1, ST2 e ST4, em contexto UR e U. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais as crianças utilizaram o recurso de omissão do *tap*.

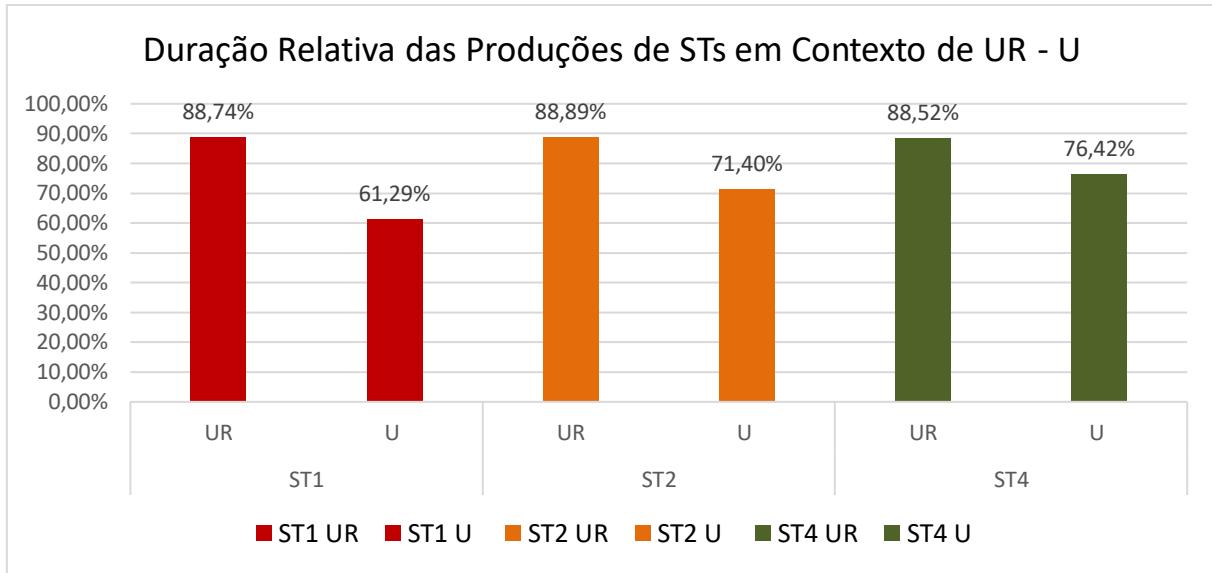


Gráfico 56 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja) e ST4 (verde), em contexto de UR (com omissão do *tap*)e U, para posição medial

Conforme revela o Gráfico 56, percebe-se que os índices de porcentagens de duração relativa das vogais em CVC apresentam valores maiores do que as vogais presentes em CV, para todos os sujeitos. No entanto, os dados de UR de ST2 referem-se a uma única produção, conforme já reportado, não sendo viável estabelecer comparação com U. Ao analisar as produções de ST4, nota-se uma pequena diferença entre os índices de duração relativa. As palavras que abrangiam o rótico apresentam uma média de 88,52% de duração relativa, ao passo que a vogal nuclear dos itens lexicais que possuíam apenas a vogal alta posterior como alvo ocupavam 76,42% da sílaba. Entretanto, os dados de ST1 exibem um comportamento distinto, pois há uma diferença expressiva entre os índices de porcentagem. As vogais dos alvos em contexto U ocupavam 61,29% da sílaba, à medida que as vogais resultantes da omissão em contexto UR ocupavam 88,74% da sílaba.

Posto isto, é possível inferir que o parâmetro de duração marca contraste entre os alvos que contém o *tap*, em CVC, e os que não o contêm, em CV. No entanto, o papel que esse parâmetro parece exercer entre os sujeitos tende a ser mais forte para ST1 do que para ST4, uma vez que a diferença entre as porcentagens é mais expressiva para o primeiro. Desse modo, corrobora-se o que já tinha sido apontado pelas médias de duração absoluta.

Para fazer a comparação dos dados entre os sujeitos transversais e o sujeito controle, foram analisados os valores de duração absoluta de elementos de mesma natureza. Sendo assim, o Gráfico 57 apresenta os resultados das médias de duração

absoluta da sílaba e da vogal das produções realizadas por ST1 e por S0. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa; (IR) circo, círculo; (UR) curto, curta, turco. Os valores reportados para os itens lexicais constituídos por [r] pertencem às produções nas quais o sujeito transversal não produziu o *tap*.

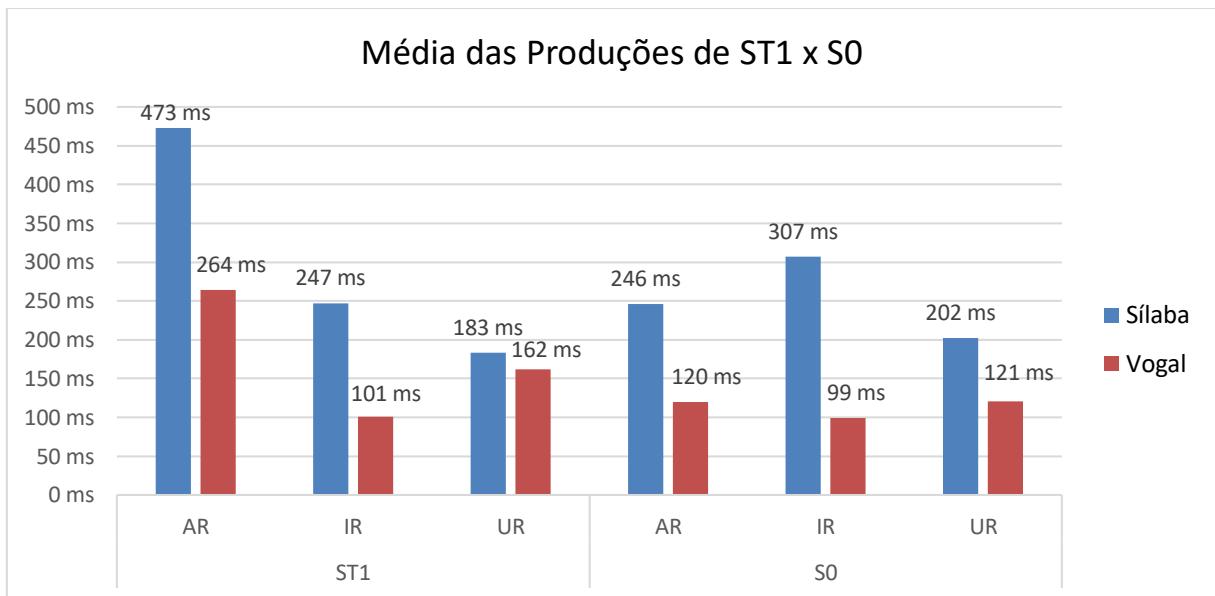


Gráfico 57 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de AR, IR e UR (com omissão do *tap*), das produções de ST1 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial

Ao observar os dados de AR do sujeito transversal, a partir do Gráfico 57, é necessário destacar que esses valores se tratam de uma produção esporádica. Quanto aos dados dos demais contextos vocálicos, percebe-se maior média de duração para as produções do sujeito controle. Esse fato pode estar associado à produção do rótico nos dados do sujeito controle, podendo, assim, elevar a duração absoluta da sílaba.

Em relação às médias de duração da vogal, é possível observar que os valores de IR para ambos informantes são similares. O sujeito transversal apresenta 101 ms de duração segmental da vogal e o sujeito controle a produz com uma média de 99 ms de duração absoluta. No entanto, para UR, há uma diferença entre os valores de duração entre os dois sujeitos. A duração absoluta da vogal para o sujeito controle é de 121 ms, ao passo que o sujeito transversal apresenta 162 ms de média. Logo, há indícios de que, quando o *tap* deveria figurar precedido pela vogal alta posterior, ocorre um alongamento da vogal nas produções do sujeito transversal.

O Gráfico 58 exibe os valores das médias de duração absoluta da sílaba e da vogal das produções realizadas por ST2 e por S0. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa; (IR) circo, círculo; (UR) curto, curta, turco. Os valores reportados para os itens lexicais constituídos por [r], produzidos pelo sujeito transversal, pertencem às produções nas quais o *tap* foi omitido.

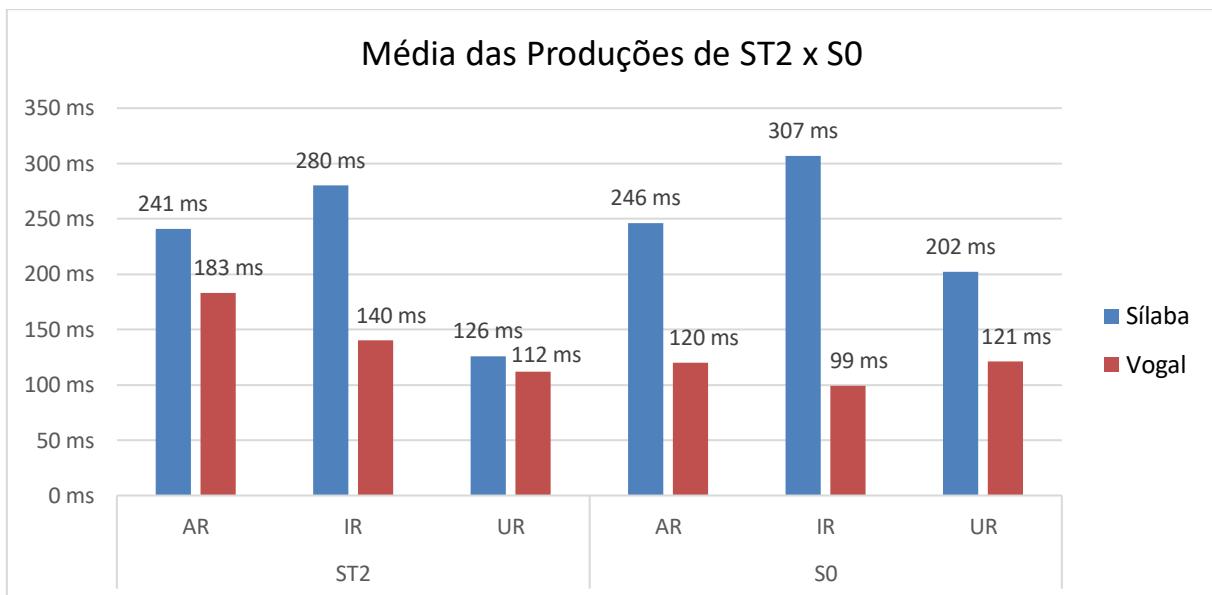


Gráfico 58 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de AR, IR e UR, (com omissão do *tap*) das produções de ST2 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial

Os valores de UR de ST2 não são passíveis de comparação, pois tratam-se de uma produção esporádica. Posto isto, ao compararem-se as médias de duração da sílaba, a partir do Gráfico 58, constatam-se maiores valores para as médias do sujeito controle.

Quanto às produções de vogais, percebe-se que o sujeito transversal apresenta maiores médias de duração em comparação com o sujeito controle. Para IR, há diferença expressiva entre as médias, pois ST2 apresenta 140 ms de duração para a vogal, ao passo que, para S0, é reportado o valor de 99 ms. No que se refere aos valores da vogal em contexto AR, a diferença é ainda maior. Enquanto ST2 exibe um valor médio de 183 ms para a vogal, S0 apresenta uma média de 120 ms de duração absoluta da vogal nuclear da sílaba. Desse modo, há indícios de que o sujeito transversal emprega o alongamento da vogal para marcar uma posição que deveria

ser do rótico. Essa marcação é evidente tanto para a vogal baixa quanto para a vogal alta anterior.

O Gráfico 59 reúne os resultados das médias de duração absoluta da sílaba e da vogal das produções realizadas por ST3 e por S0. Os dados apresentados são referentes ao contexto de (IR) circo, círculo, uma vez que esse informante não apresentou omissões para os demais contextos vocálicos. Os valores reportados para os itens lexicais constituídos por [r], produzidos pelo sujeito transversal, pertencem as produções nas quais o *tap* foi omitido.

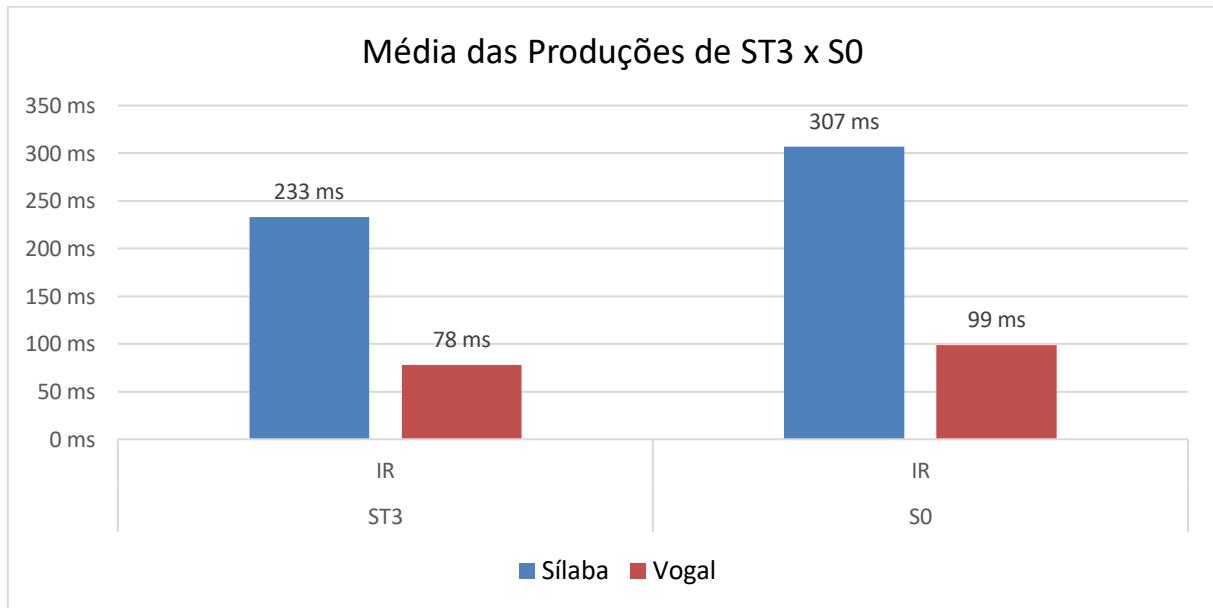


Gráfico 59 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de IR, das produções de ST3 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial

Ao comparar as médias de duração absoluta da sílaba entre ST3 e S0, observa-se menor tempo de duração para o sujeito transversal, que apresenta uma média de 233 ms de duração da sílaba. Para o sujeito controle, essa média é de 307 ms para a duração da sílaba. Ao retomar os dados do Gráfico 53, nota-se que as médias de duração absoluta de ST3, tanto para a sílaba quanto para vogal em contexto CVC, são as menores dentre todos os sujeitos. O mesmo ocorre quando a média de duração absoluta da vogal nuclear do alvo CVC, no contexto de IR, de ST3 é comparada com a de S0, ou seja, ST3 apresenta uma menor duração vocálica em suas produções, não indicando a presença de alongamento compensatório.

O Gráfico 60 apresenta os resultados das médias de duração absoluta da sílaba e da vogal das produções realizadas por ST4 e por S0. Os dados apresentados

estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa; (IR) circo, círculo; (UR) curto, curta, turco. Os valores reportados para os itens lexicais constituídos por [r] pertencem as produções nas quais ST4 não produziu o *tap*.

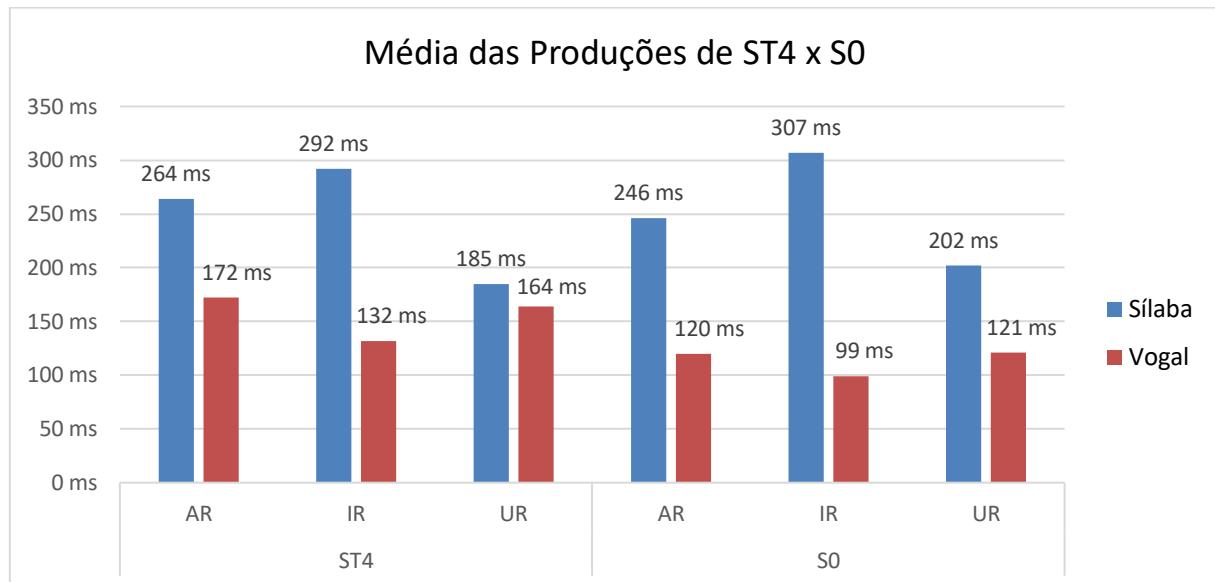


Gráfico 60 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho), em contexto de AR, IR e UR (com omissão do *tap*), das produções de ST4 (à esquerda) em comparação às produções de S0 (à direita), para a posição medial

Antes de examinar os valores apresentados pelo Gráfico 60, é necessário relembrar que as médias reportadas para ST4, em AR, são procedentes de duas produções, como já mencionado. Desse modo, essas médias não são passíveis de comparação com os dados de S0.

Ao compararem-se as médias entre os sujeitos, em contexto de IR, percebe-se que a duração da sílaba apresenta valores aproximados. Quanto à duração da vogal alta anterior, há certa diferença, pois ST4 exibe uma média de 132 ms, enquanto que, para S0, é reportado 99 ms de duração da vogal.

No que se refere aos dados de UR, observa-se que a sílaba e a vogal, produzidas por ST4, exibem médias aproximadas. A duração segmental da sílaba é de 185 ms, ao passo que a vogal apresenta 164 ms de duração absoluta. Para o sujeito controle, há uma diferença expressiva entre os valores de duração da sílaba e da vogal, pois há a produção da consoante no início da sílaba e a articulação do *tap* no final da sílaba. Sendo assim, parece que, para o contexto da vogal alta posterior, o sujeito transversal preenche o espaço do rótico não produzido com o alongamento da vogal.

Em suma, tanto a vogal alta anterior quanto a posterior apresentam produções mais longas nos dados de ST4, quando comparados aos de S0, indicando, portanto, a presença de alongamento compensatório para marcar a não realização do rótico.

Após a análise das produções, nas quais os sujeitos transversais não produzem o *tap*, a Tabela 32 apresenta os valores referentes aos dados nos quais as crianças semivocalizam o segmento. A fim de observar o comportamento do ditongo oriundo da semivocalização do rótico, a referida tabela exibe as médias de duração absoluta da sílaba, rótico e ditongo das palavras que, em posição medial, abarcavam o *tap* como alvo – carpa, farpa, carpa, curta, curto, turco. Do mesmo modo, são exibidos os valores de duração relativa da vogal em relação à sílaba.

CODA MEDIAL					
AR		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Ditongo	(ditongo - sílaba)
		ST1	347	–	242
		ST2	–	–	–
		ST3	–	–	–
	ST4	300	–	217	75,14
<b>MÉDIA</b>		<b>323</b>	–	<b>230</b>	<b>73,47</b>
DP		33,23	–	17,52	2,35
UR		ST1	188	–	137
		ST2	–	–	–
		ST3	–	–	–
		ST4	244	–	207
		<b>MÉDIA</b>	<b>216</b>	–	<b>172</b>
	DP	39,72	–	49,73	9,10

Tabela 32 - Médias de duração absoluta e relativa dos alvos CVC em posição medial, produzidos com o recurso da semivocalização pelos sujeitos transversais, conforme contexto vocálico

A partir da Tabela 32, percebe-se que a semivocalização do *tap* ocorreu apenas nas produções de S1 e ST4, em ambos contextos vocálicos. Esse recurso ocorreu de forma predominante nos dados dessas duas crianças. Em AR, apenas um alvo não foi semivocalizados por ST1, ao passo que ST4 não semivocalizou apenas dois alvos. Para UR, ambos informantes semivocalizaram todas as produções de *curta* e *curto*, enquanto o rótico na palavra *turco* foi omitido em todas as suas ocorrências. As médias de duração absoluta do ditongo [aj] giram em torno de 217 e 242 ms. Para [uj], as produções de ST1 exibem 137 ms de média e as de ST4, 207 ms de duração absoluta.

Para examinar as possíveis semelhanças e/ou diferenças entre as semivocalizações de [r] e os ditongos produzidos em palavras com alvo CVV, a Tabela 33 apresenta as médias de duração absoluta da sílaba e do ditongo, e a duração relativa do ditongo presentes nos itens lexicais que, em posição medial, abarcavam o ditongo como alvo – caixa, cuida.

CODA MEDIAL				
		Duração absoluta (ms)	Duração relativa (%)	
		Sílaba	Ditongo	(ditongo - sílaba)
AI	ST1	234	207	88,47
	ST2	233	193	82,52
	*ST3	195	144	73,85
	ST4	180	155	86,06
	<b>MÉDIA</b>	<b>211</b>	<b>175</b>	<b>82,73</b>
UI	<b>DP</b>	<b>27,35</b>	<b>30,24</b>	<b>6,40</b>
	ST1	290	222	77,51
	ST2	199	128	63,61
	ST3	208	173	83,24
	ST4	194	151	77,27
	<b>MÉDIA</b>	<b>223</b>	<b>169</b>	<b>75,41</b>
	<b>DP</b>	<b>45,24</b>	<b>40,00</b>	<b>8,34</b>

Tabela 33 - Médias de duração absoluta e relativa os alvos CVV que contêm ditongo em posição medial produzidos pelos sujeitos transversais, conforme contexto vocálico

\* Valores obtidos com base em uma única produção, assim, não constitui média

Ao examinar os dados da palavra *caixa*, constatou-se que ocorreu monotongação em duas produções de ST3. Sendo assim, o valor reportado pelo quadro é oriundo de apenas uma ocorrência do alvo na forma esperada. No entanto, esse dado não será utilizado para uma análise comparativa, tendo em vista que esse informante não semivocalizou suas produções em contexto de vogal baixa.

A partir da Tabela 33, observa-se que as médias de duração da sílaba, para AI, são próximas entre os sujeitos transversais. A sílaba composta por [aj] apresenta uma duração em torno de 180 a 234 ms. No entanto, para os dados de UI, nota-se que a média reportada por ST1 se distancia dos valores apresentados pelos demais informantes. Para ST2, ST3 e ST4, as médias apresentadas são 199, 208 e 194 ms de duração absoluta da sílaba, respectivamente. Entretanto, ST1 exibe uma média de 290 ms de duração da sílaba.

Quanto à produção dos ditongos, os sujeitos transversais apresentam uma média de 144 ms a 207 ms para [aj] e 128 ms a 222 ms para [uj]. Assim como ocorre para a duração da sílaba, as médias de ST1 para a produção dos ditongos são maiores em relação às médias dos demais informantes, especialmente em [uj].

De acordo com as observações aqui tecidas, ao não produzir o *tap*, a criança semivocaliza algumas produções. A fim de observar a capacidade do parâmetro de duração ser capaz de marcar distinção entre as palavras que apresentam semivocalização do rótico e os alvos CVV, serão apresentados os Gráficos 61, 62, 63 e 64.

O Gráfico 61 apresenta os valores das médias de duração absoluta da sílaba e do ditongo das produções de ST1 e ST4. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa e (AI) caixa. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais a informante utiliza o recurso de semivocalização do *tap*.

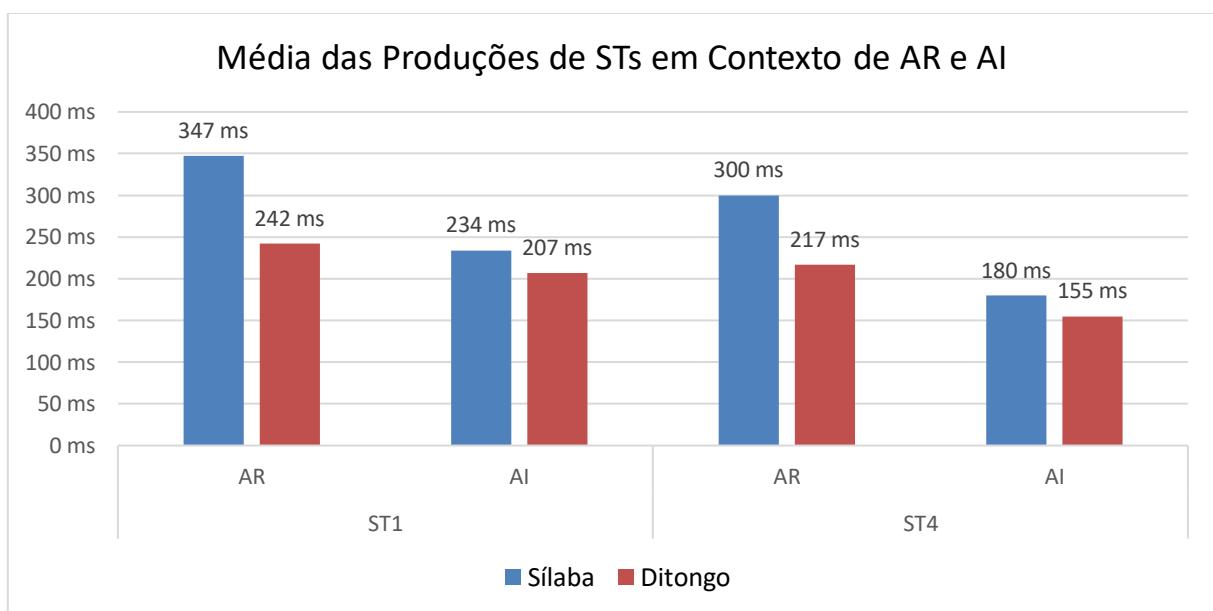


Gráfico 61 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de AR e AI, para a posição medial

Ao observar os dados de ST1 e ST4, presentes no Gráfico 61, é possível verificar que os valores reportados para AR são maiores em relação às médias de AI. Nos dados de ST1, as médias de duração da sílaba apresentam uma grande diferença entre AR e AI. Esse contraste entre essas médias pode estar relacionado com a presença de uma consoante fricativa no início de um dos itens lexicais que compõem

o contexto de AR. Quanto à produção do ditongo por ST1, em contexto de AI, exibe uma média de 207 ms, enquanto o ditongo produzido em AR apresenta 242 ms de duração absoluta. Para ST4, há uma diferença expressiva entre os valores de duração da sílaba e do ditongo entre as palavras em contexto de CVC e CVV. A duração do ditongo, presente em CVV, é de 155 ms, ao passo que o ditongo oriundo da semivocalização do rótico apresenta uma média de 217 ms de duração absoluta.

O Gráfico 62 apresenta os resultados das médias de duração relativa do ditongo em relação à sílaba, em contexto AR e AI, realizadas por ST1 e ST4. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais as crianças utilizaram o recurso de semivocalização do *tap*.

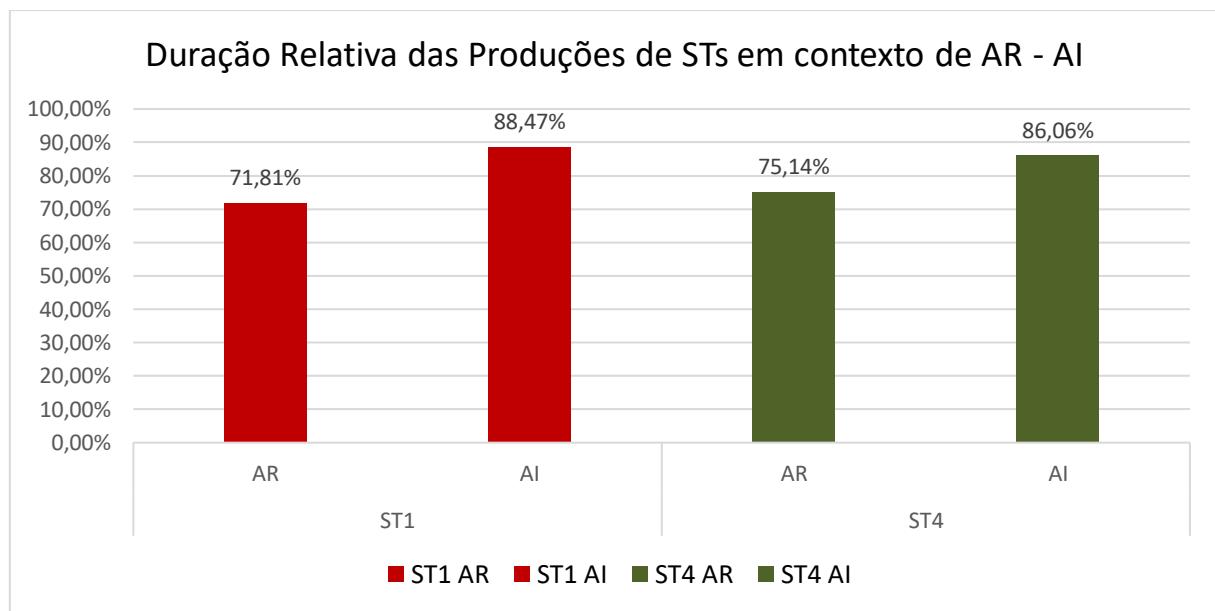


Gráfico 62 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de AR e AI, para posição medial

A partir da observação do Gráfico 62, nota-se que as produções das vogais semivocalizadas apresentam índices de duração menor do que as vogais em contexto CVV. Em contexto de [a], ST1 exibe 72,81% de duração relativa da vogal em relação à sílaba e os dados de ST4 apresentam uma média de 75,14%. As palavras que continham o ditongo [aj] como alvo continham produções com médias de 88,47% e 86,06% para ST1 e ST4, respectivamente. Considerando que o *tap* preenche um espaço pequeno na sílaba, menor do que a produção de uma semivogal, seria esperada, das produções semivocalizadas, uma menor duração para tentar aproximar

à do segmento alvo. Sendo assim, é possível que a criança tente marcar um contraste entre as produções por meio de uma menor duração do segmento semivocalizado.

O Gráfico 63 apresenta os valores das médias de duração absoluta da sílaba e do ditongo das produções de ST1 e ST4 para os seguintes contextos: (UR) curto, curta, turco e (UI) cuida. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais a informante utiliza o recurso de semivocalização do *tap*.

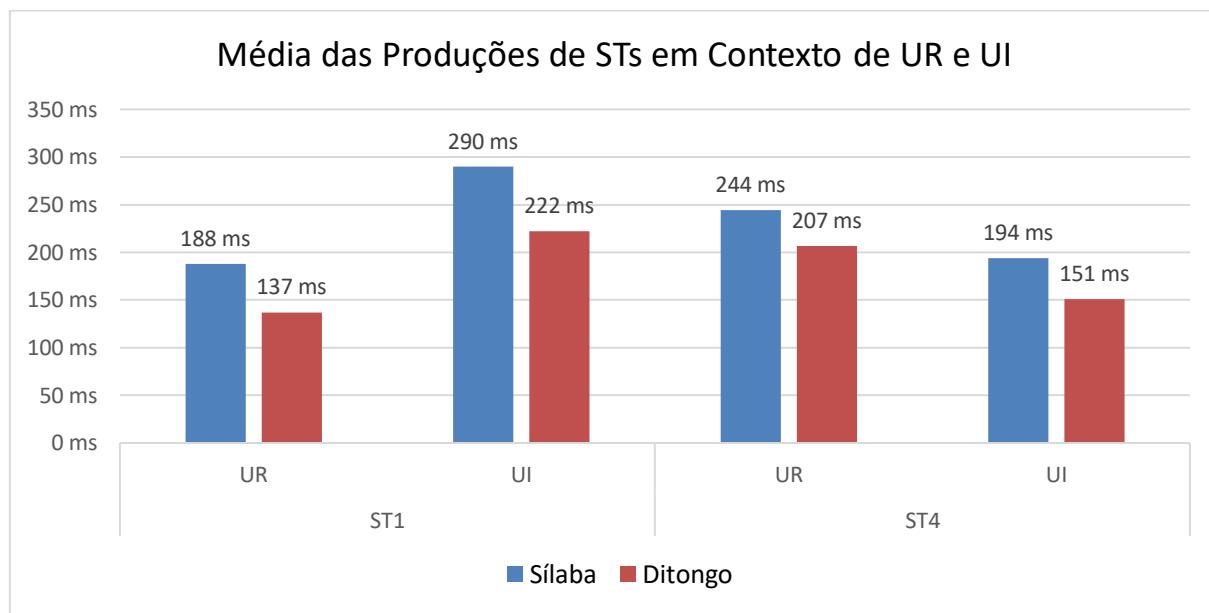


Gráfico 63 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de UR e UI, para a posição medial

As produções de ST1 e ST4 comportam-se de modo diferente, de acordo com o Gráfico 63. É possível observar que as médias de UR, para o primeiro, são menores do que as médias para UI, uma vez que o ditongo oriundo da semivocalização apresenta 137 ms de duração absoluta e o ditongo de CVV exibe 222 ms de média. Já as articulações de UR de ST4 são maiores em valores absolutos em relação às produções em contexto de UI, apresentando uma diferença expressiva entre ambas.

O Gráfico 64 exibe as médias de duração relativa do ditongo em relação à sílaba das produções realizadas por ST1 e ST4, em contexto de UR e UI. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais as crianças utilizaram o recurso de semivocalização do *tap*.

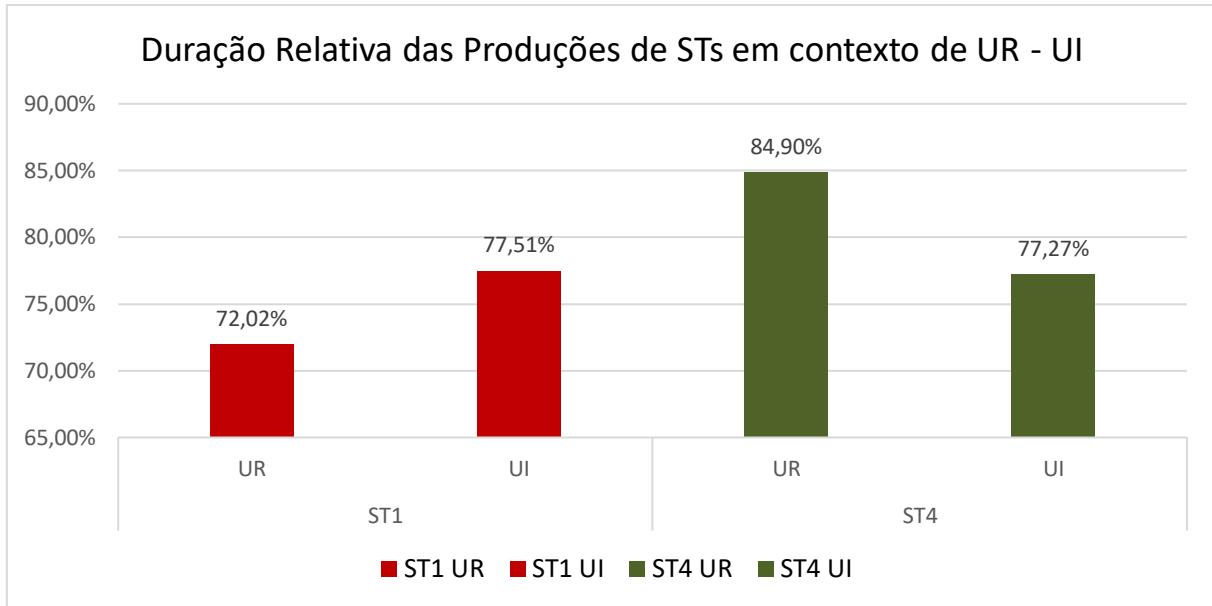


Gráfico 64 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de UR e UI, para posição medial

Ao observar os resultados do Gráfico 64, nota-se que os índices de porcentagens dos informantes apresentam comportamentos diferentes. O ditongo oriundo da semivocalização o *tap*, produzido por ST1, ocupa 72,02% da sílaba. Já a duração segmental do ditongo [uj], presente em CVV, apresenta um índice de 77,51% de duração relativa. O ditongo produzido por ST4, em contexto de UR, ocupa 84,90% da sílaba, enquanto, em contexto de UI, exibe uma maior porcentagem, 77,27%.

Desse modo, corrobora-se o que já havia sido indiciado pelos valores absolutos de duração, ou seja, ST1, a partir do parâmetro de duração, parece marcar distinção entre alvos CVC – contendo o *tap* – e CVV. A diferença de duração relativa entre as duas médias é menor, pois, destaca-se o fato de, justamente para alvos CVC, apresentar uma menor duração, o que parece mais coadunado com um ditongo produzido em detrimento de um rótico.

A partir desse momento serão apresentados os resultados referentes às médias de durações absoluta e relativa das produções em posição final de sílaba.

A Tabela 34 apresenta as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição final, abarcavam o rótico como alvo – mar, lugar, azar, vir, Nadir, Zair, cur, abajur, Artur. Na mesma tabela, são exibidos os valores de duração relativa da vogal em relação à sílaba. Os valores reportados a seguir são referentes às produções nas quais os sujeitos transversais articulam um *tap*.

CODA FINAL					
		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(rótico - sílaba)
AR	ST1	—	—	—	—
	ST2	424	35	183	8,38
	ST3	294	31	157	10,62
	ST4	—	—	—	—
	<b>MÉDIA</b>	<b>359</b>	<b>33</b>	<b>170</b>	<b>9,50</b>
IR	DP	<b>91,39</b>	<b>2,83</b>	<b>18,09</b>	<b>1,58</b>
	ST1	—	—	—	—
	ST2	—	—	—	—
	ST3	351	45	199	13,02
	ST4	—	—	—	—
UR	<b>MÉDIA</b>	<b>351</b>	<b>45</b>	<b>199</b>	<b>13,02</b>
	DP	—	—	—	—
	ST1	—	—	—	—
	ST2	380	24	167	7,29
	ST3	325	47	138	14,17
UR	*ST4	526	16	281	3,04
	<b>MÉDIA</b>	<b>415</b>	<b>29</b>	<b>197</b>	<b>8,74</b>
	DP	<b>98,21</b>	<b>15,79</b>	<b>74,16</b>	<b>6,55</b>

Tabela 34 - Médias de duração absoluta e relativa dos alvos que contêm rótico (articulados como *tap*) em posição final produzidos pelos sujeitos transversais, conforme contexto vocálico

\* Valores obtidos com base em uma única produção, assim, não constitui média

Com base nos valores exibidos na Tabela 34, verifica-se que, em contexto de AR, apenas os sujeitos ST2 e ST3 produziram o segmento. Para IR, somente ST2 articulou o rótico como a variante *tap*. Quanto à UR, os sujeitos ST2, ST3, e ST4 produziram o segmento, no entanto, o valor reportado para ST4 corresponde a uma única produção, o alvo *abajur*. Sendo assim, observa-se que o rótico apresenta valores de duração absoluta entre 24 e 47 ms, ocupando 7,29% e 14,17% da sílaba nos dados de ST2 e ST3, respectivamente. Ao observar os valores médios de duração da vogal núcleo da sílaba, nota-se que a vogal baixa apresenta 157 ms e 183 ms de duração absoluta, a vogal alta anterior exibe 199 ms de duração e a vogal alta posterior, 138 ms e 167 ms.

Outras formas de róticos foram encontradas para dois sujeitos: ST2 e ST3. A única variante rótica articulada por ST2, à exceção do *tap*, foi a fricativa glotal desvozeada. Essa apresentou-se nas produções da palavra *mar*, com uma média de 436 ms de duração da sílaba, 161 ms de duração da fricativa e 235 ms de duração da

vocal. Esse tipo de segmento também foi articulado em todas as palavras que abarcam a vogal alta anterior como ambiente vocálico, sendo elas: *vir*, *Nadir* e *Zair*. Para esses itens lexicais, a média geral foi de 519 ms de duração da sílaba, 155 ms de duração da fricativa glotal desvozeada e 229 ms de duração da vogal [i].

Quanto às produções de ST3, esse produziu uma vibrante para palavra *azar*, com 379 ms de duração da sílaba, 143 ms de duração do rótico e 131 ms de duração da vogal. Uma produção de *Zair* foi articulada com um *tap* fricativo, apresentando 329 ms de duração da sílaba, 116 ms de duração da variante rótica e 242 ms de duração da vogal. O rótico do item lexical *cur* foi produzido uma vez como uma aproximante alveolar desvozeada, exibindo uma duração da sílaba de 225 ms, 61 ms de duração da aproximante e 145 ms de duração da vogal. Por fim, uma produção de *Arthur* foi articulada como um *tap* fricativo, apresentando 434 ms de duração da sílaba, 139 ms de duração do rótico e 135 ms de duração da vogal.

Após a descrição das produções articuladas como *tap*, por parte dos sujeitos transversais, esperar-se-ia que fosse feita a comparação desses dados com as produções do sujeito controle. No entanto, essa análise comparativa não poderá ser realizada para os dados em posição final, uma vez que S0 não produziu *taps*, pois articulou apenas vibrantes nesse contexto. Do mesmo modo, não é possível comparar as produções de vibrantes entre os informantes, uma vez que os sujeitos transversais não possuem dados suficientes para que uma análise confiável seja realizada. Sendo assim, os Gráficos 65, 66 e 67 reúnem as produções de *taps* de ST2 e ST3, de acordo com o ambiente vocálico.

O Gráfico 65 reúne os resultados das médias de duração absoluta da sílaba e da vogal das produções das palavras realizadas por ST2. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa e (UR) curto, curta, turco. Os valores reportados para o rótico são oriundos das produções nas quais a criança produz um *tap*, em posição final.

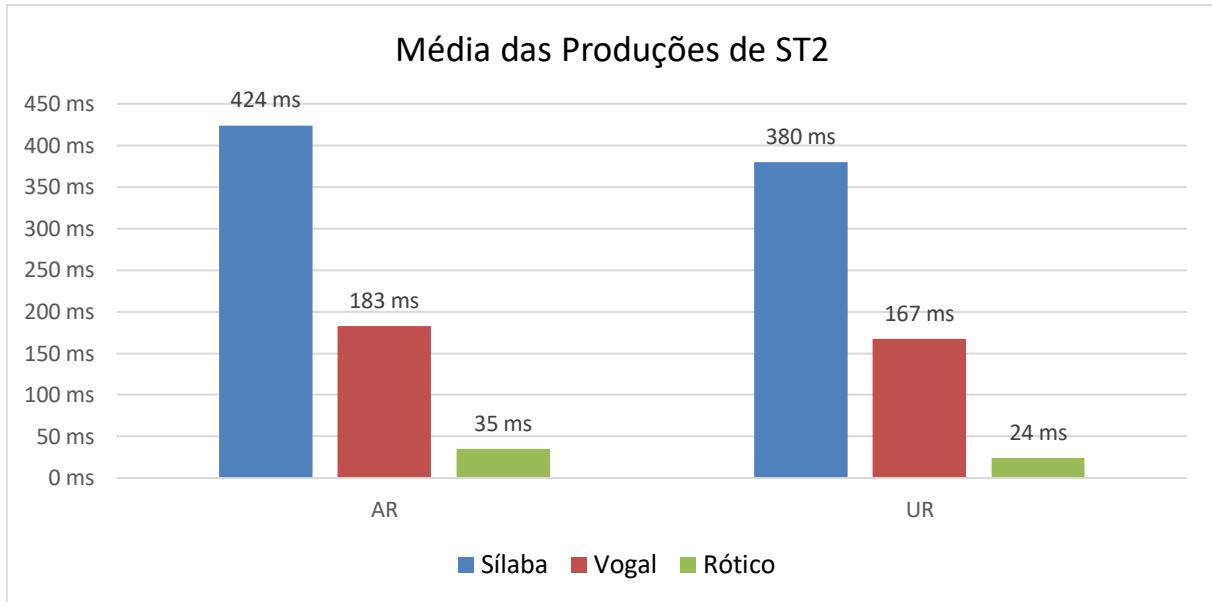


Gráfico 65 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de ST2, para a posição final

Ao observar os valores de AR e UR, a partir do Gráfico 65, nota-se que a duração segmental dos dados em contexto de vogal baixa é maior em relação aos dados em contexto da vogal alta posterior. A vogal [a] apresenta 183 ms de média de duração absoluta, enquanto [u] exibe 167 ms. Esse comportamento das vogais pode estar relacionado à duração intrínseca dos segmentos vocálicos. Quanto à duração do rótico, em contexto de AR, o segmento é produzido com uma média de 35 ms. No entanto, quando o *tap* encontra-se em contexto de UR, é reportada uma média um pouco mais baixa, 24 ms de duração absoluta.

O Gráfico 66 apresenta os resultados das médias de duração absoluta da sílaba e da vogal das produções realizadas por ST3. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) carta, farpa, carpa; (IR) circo, círculo; (UR) curto, curta, turco. Os valores reportados para o rótico são referentes às produções nas quais a criança articula um *tap*, em posição final.

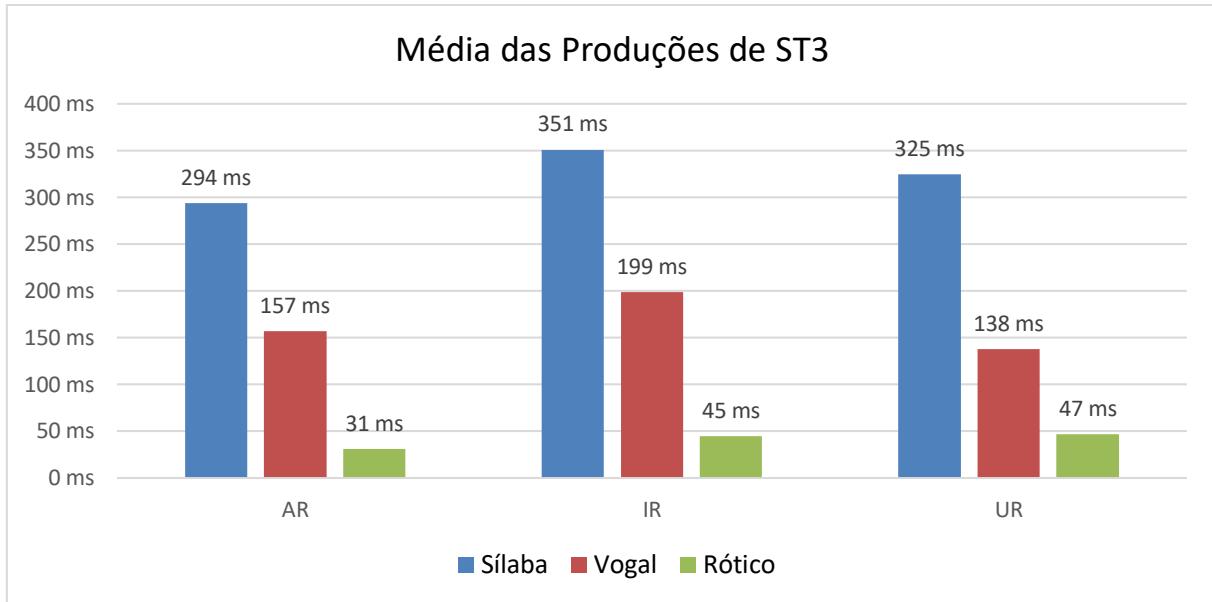


Gráfico 66 - Médias de duração absoluta da sílaba (azul), da vogal (vermelho) e do rótico (verde) das produções de ST3, para a posição final

A partir da observação do Gráfico 66, nota-se que as palavras produzidas em contexto de IR apresentam maior duração da sílaba e da vogal. Quanto à produção do *tap*, ao contrário de ST2, percebe-se que o ambiente vocálico de AR exibe a menor duração do segmento, 31 ms. Para a produção do rótico em contexto de vogais altas, não há diferença expressiva de duração, pois, para IR, é reportado 45 ms de duração do *tap* e, para UR, é apresentado 47 ms de duração absoluta do segmento.

O Gráfico 67 reúne os resultados das médias de duração relativa do *tap* em relação à sílaba das produções, em contexto AR, IR e UR, realizadas por ST2 e ST3.

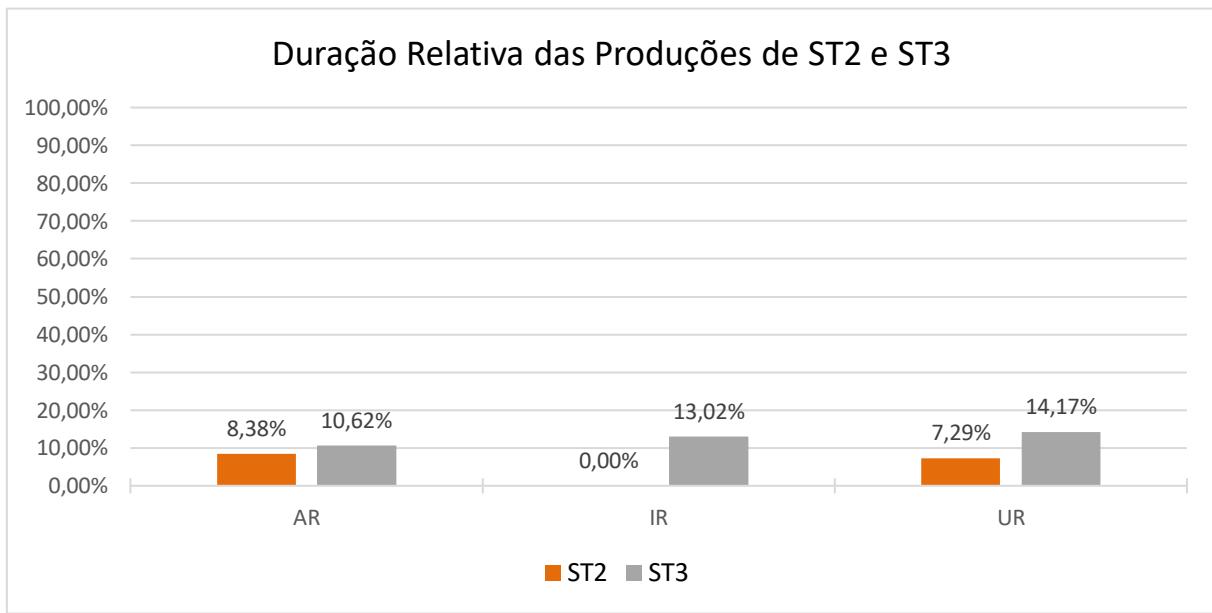


Gráfico 67 - Médias de duração relativa do rótico de ST2 (laranja) e ST3 (cinza), em contexto de AR, IR e UR, para posição final

A análise do Gráfico 67 revela que os índices de porcentagens observados para as produções de ST2 e ST3, em contexto AR, são relativamente aproximados. O róticos produzidos por ST3 ocupam 10,62% da sílaba, enquanto o rótico articulado por ST2 ocupa 8,38% da sílaba. Quanto à duração relativa das palavras em contexto de [i], ST3 é o único informante que articula um *tap*, exibindo 13,02% de índice de duração. As palavras compostas por [u], produzidas por ST3, exibem percentuais maiores, em relação às produções de ST2. O *tap* produzido por ST3 ocupa 14,17% da sílaba e o segmento articulado por ST2 preenche apenas 7,29% da sílaba.

Com base nos resultados aqui apresentados, é possível observar que, quando os sujeitos ST2 e ST3 articularam um *tap*, em contexto de AR, IR e UR, suas produções comportam-se conforme o esperado para a produção do segmento, uma vez que seus valores de produção se aproximam das médias reportadas pela literatura.

Após a análise dos dados, nos quais os sujeitos transversais produziram o *tap*, a Tabela 35 apresenta os valores referentes às produções nas quais as crianças omitem o segmento. A fim de observar o comportamento da vogal núcleo da sílaba, a referida tabela exibe as médias de duração absoluta da sílaba, rótico e vogal nuclear das palavras que, em posição final, abarcavam o rótico como alvo – mar, lugar, azar, vir, Nadir, Zair, cur, abajur, Artur. Em seguida, são exibidos os valores de duração relativa da vogal em relação à sílaba.

CODA FINAL					
		Duração absoluta (ms)		Duração relativa (%)	
		Sílaba	Rótico	Vogal	(vocal - sílaba)
AR	ST1	373	—	264	69,89
	ST2	—	—	—	—
	ST3	—	—	—	—
	ST4	302	—	238	78,88
	<b>MÉDIA</b>	<b>338</b>		<b>251</b>	<b>74,38</b>
IR	DP	<b>50,70</b>		<b>18,58</b>	<b>6,36</b>
	ST1	443	—	234	53,10
	ST2	—	—	—	—
	ST3	—	—	—	—
	ST4	289	—	198	68,92
UR	<b>MÉDIA</b>	<b>366</b>	—	<b>216</b>	<b>61,01</b>
	DP	<b>108,64</b>	—	<b>25,67</b>	<b>11,18</b>
	ST1	337	—	221	66,83
	ST2	271	—	126	45,07
	*ST3	225	—	162	72,00
	ST4	426	—	213	53,75
	<b>MÉDIA</b>	<b>315</b>	—	<b>181</b>	<b>59,41</b>
	DP	<b>87,21</b>	—	<b>44,89</b>	<b>12,26</b>

Tabela 35 - Médias de duração absoluta e relativa dos alvos CVC em posição final, produzidos, com o recurso da omissão, pelos sujeitos transversais, conforme contexto vocálico

\* Valores obtidos com base em uma única produção, assim, não constitui média

A Tabela 35 permite observar que os sujeitos ST2 e ST3 não empregaram o recurso de omissão nas palavras em contexto de AR e IR. Em relação aos itens lexicais em contexto de UR, todos os sujeitos transversais não produziram o *tap* em algum momento. Tal fato pode indicar o contexto da vogal alta posterior como desfavorecedor para a emergência do rótico. No entanto, é necessário ressaltar que uma única omissão, oriunda do alvo *abajur*, foi observada para ST3. Já o sujeito ST2 omitiu o *tap* em duas produções, nos alvos *abajur* e *Artur*.

Ao observar o comportamento da vogal, percebe-se que a vogal [a], resultante da omissão do *tap*, apresenta 238 ms e 264 ms de duração. A vogal [i] exibe uma média de 198 ms e 234 ms de duração. Por fim, para a vogal [u], é reportado uma média de duração absoluta de 221 ms para ST1 e 213 ms para ST4. Quanto à duração relativa da vogal em relação à sílaba, os maiores índices foram reportados para a vogal [a], pois essa ocupa 69,89% e 78,88% da sílaba.

A fim de comparar a duração das vogais realizadas nas produções que continham omissões de [r] e a duração das vogais produzidas em itens lexicais com o molde silábico CV, a Tabela 36 apresenta as médias de duração absoluta da sílaba e da vogal nuclear das palavras que, em posição final, incluíam a vogal em alvos CV – cata, Chico, cuca. Essa inclui também os valores de duração relativa da vogal em relação à sílaba.

CODA FINAL				
		Duração absoluta (ms)		Duração relativa (%)
		Sílaba	Vogal	(vogal - sílaba)
A	ST1	292	187	64,08
	ST2	427	332	76,90
	ST3	176	101	56,73
	ST4	243	144	61,19
	<b>MÉDIA</b>	<b>285</b>	<b>191</b>	<b>64,72</b>
I	DP	<b>106,07</b>	<b>100,18</b>	<b>8,66</b>
	ST1	394	189	50,35
	ST2	252	180	69,15
	ST3	250	147	58,91
	ST4	297	176	62,27
U	<b>MÉDIA</b>	<b>298</b>	<b>173</b>	<b>60,17</b>
	DP	<b>67,65</b>	<b>18,19</b>	<b>7,81</b>
	ST1	217	190	86,26
	ST2	276	159	57,38
	ST3	266	125	47,89
	ST4	288	131	44,78
	<b>MÉDIA</b>	<b>262</b>	<b>151</b>	<b>59,08</b>
	DP	<b>31,08</b>	<b>29,90</b>	<b>18,90</b>

Tabela 36 - Médias de duração absoluta e relativa dos alvos CV em posição final dos sujeitos transversais, conforme contexto vocálico

Com base nos valores apresentados na Tabela 36, é possível observar que ST2 apresenta maior média para as palavras em contexto de vogal baixa. Ao analisar os dados referentes à duração segmental das vogais, observa-se que, para a vogal [a], foi reportado médias entre 101 ms a 187 ms, à exceção de ST2 que apresenta 332 ms de duração para esse ambiente vocálico. A vogal [i] apresenta médias entre 147 ms a 189 ms e, por fim, a vogal [u] exibe médias entre 125 ms e 190 ms de duração absoluta. Quanto aos valores de duração relativa, esses apontam que a vogal baixa ocupa de 56,73% a 64,08% da sílaba, à exceção do segmento produzido por

ST2, que ocupa 76,90% da sílaba. A vogal alta anterior compreende de 50,35% a 69,15% do constituinte silábico. Para a vogal alta posterior, foi reportada um índice de 44,78% a 86,26% de duração relativa.

Da mesma forma que foi realizado para as produções em posição medial, os Gráficos 68, 69, 70, 71, 72 e 73 efetuam comparações para investigar a possibilidade de as crianças valerem-se de aspectos duracionais para diferenciar as palavras em contexto CVC e CV.

A partir das referidas considerações, o Gráfico 68 apresenta os resultados das médias de duração absoluta da sílaba e da vogal das produções realizadas por ST1 e ST4. Os dados apresentados estão organizados conforme os seguintes contextos: (AR) mar, lugar, azar e (A) má. Os valores exibidos para os itens lexicais compostos por [r] compreendem as produções nas quais as crianças omitiram o *tap*.

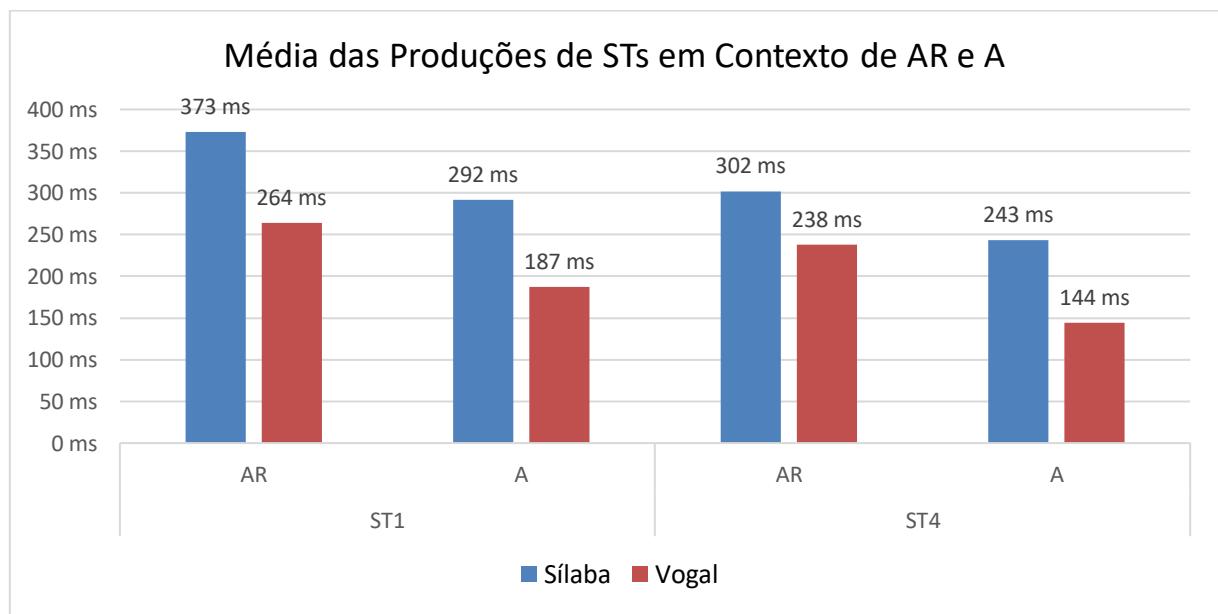


Gráfico 68 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de AR e A, para a posição final

A partir da observação do Gráfico 68, verifica-se que ambos sujeitos registram uma diferença expressiva entre suas produções de AR e A. Quanto aos dados de ST1, percebe-se o mesmo comportamento das médias de duração absoluta. Em ambiente AR, é reportado para sílaba uma média de 373 ms e para a vogal 264 ms de duração absoluta, ao passo que, em contexto de A, a sílaba exibe uma duração de 292 ms e a vogal é articulada com uma média de 187 ms de duração absoluta. Já para ST4, os itens lexicais que deveriam ser constituídos pelo rótico exibem duração segmental

maior, 302 ms de duração da sílaba e 238 ms para a vogal oriunda da omissão do segmento. As produções referentes à estrutura CV apresentam 243 ms para duração da sílaba e 144 ms de duração da vogal nuclear.

O Gráfico 69 reporta os valores das médias de duração relativa da vogal nuclear em relação à sílaba das produções realizadas por ST1 e ST4, em contexto AR e A. Os valores exibidos para as palavras compostas por [r] são referentes às produções nas quais as crianças não produziram o *tap*.

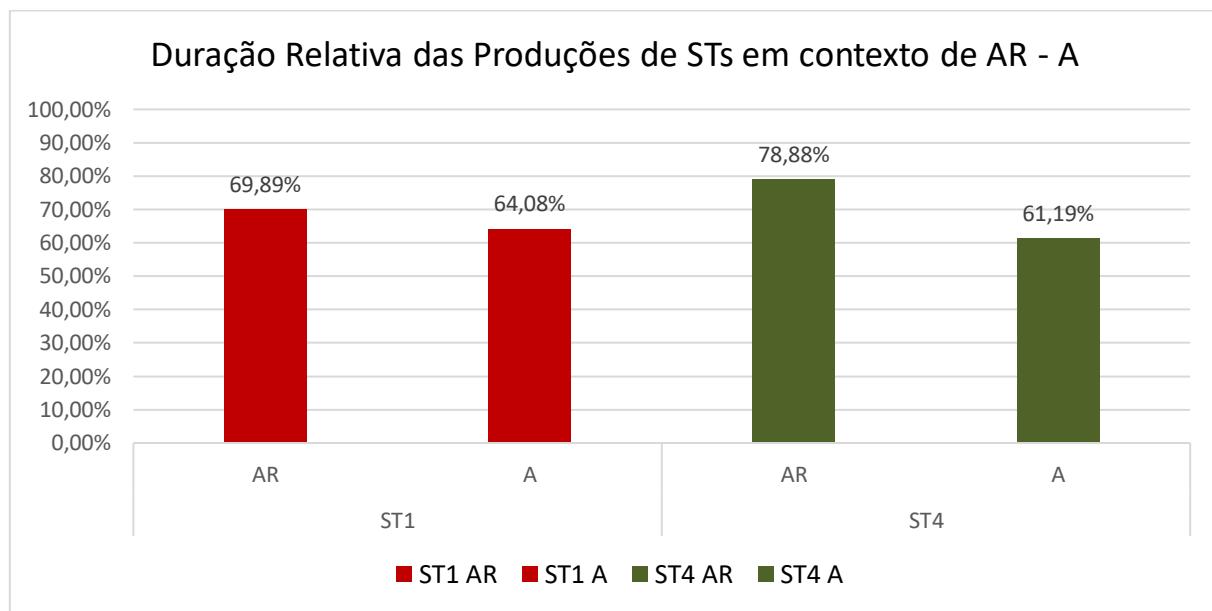


Gráfico 69 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de AR e A, para posição final

Conforme revela o Gráfico 69, para as duas crianças, os índices de porcentagens de duração relativa das vogais oriundas da omissão do rótico, em CVC, são maiores do que as vogais em contexto silábico CV. Os dados de ST4 apresentam uma diferença expressiva entre seus valores. As vogais em ambiente AR ocupam 78,88% da sílaba, à medida que as vogais presentes em A preenchem 61,19% do constituinte silábico. Para ST1, o mesmo comportamento é registrado, porém, não com a mesma intensidade, pois a diferença entre os pares de palavras é menos expressiva. As vogais referentes à omissão do *tap* compreendem 69,89% da sílaba, enquanto as vogais presentes em CV ocupam 64,08% da sílaba.

Assim, pode-se afirmar que o parâmetro de duração diferencia as produções entre os alvos CVC e CV. Apesar disso, há indícios de que esse parâmetro tenha um papel mais forte nas produções de ST4, pois esse apresenta uma diferença mais expressiva entre seus índices de duração relativa.

O Gráfico 70 exibe os resultados das médias de duração absoluta da sílaba e da vogal das produções realizadas por ST1 e ST4 nos seguintes contextos: (IR) vir, Nadir, Zair e (I) vi. Os valores reportados às palavras compostas por [r] são de produções nas quais as crianças utilizaram o recurso de omissão do *tap*.

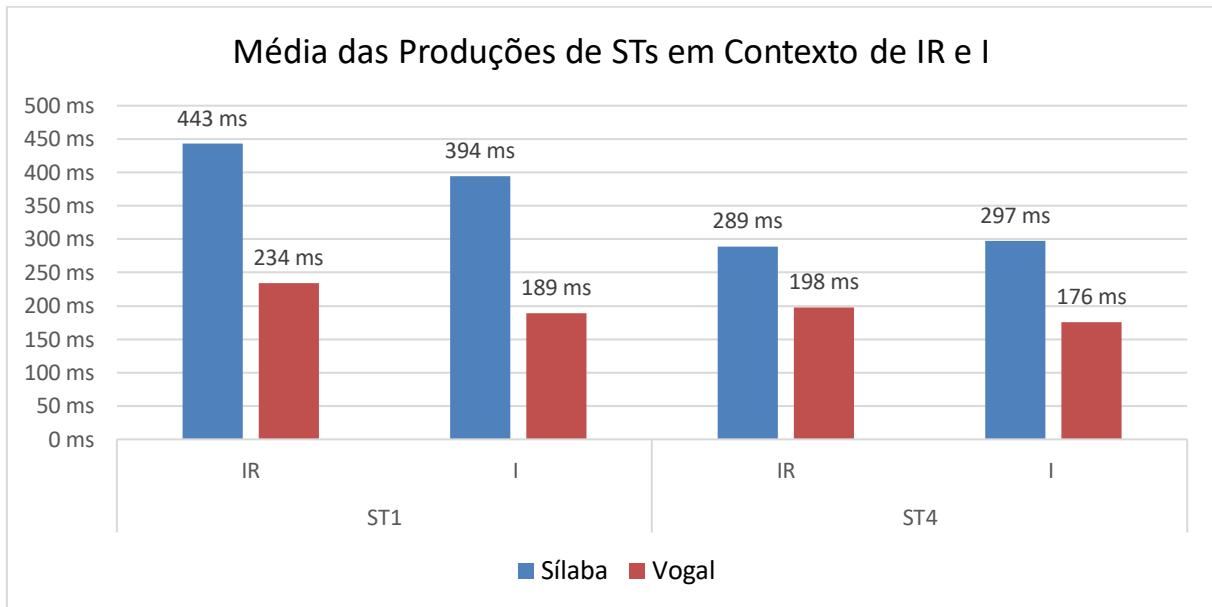


Gráfico 70 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1 e ST4, em contexto de IR e I, para a posição final

A análise do Gráfico 70 permite identificar uma maior duração segmental nas produções de ST1 em relação ao sujeito ST4. No entanto, essa diferença não é expressiva, pois para ST4, a sílaba, produzida em IR, apresenta duração de 289 ms e a vogal, 198 ms, à medida que a sílaba articulada em contexto de I exibe uma média de 297 ms e a vogal abarca 176 ms de duração absoluta. Para ST1, o mesmo desempenho é encontrado. A média de duração da sílaba, em IR, registra 443 ms de duração da sílaba e 234 ms de duração para a vogal resultante da omissão, enquanto a sílaba produzida em contexto de I registra 394 ms de duração e a vogal nuclear da sílaba apresenta 189 ms de duração absoluta.

O Gráfico 71 apresenta os valores das médias de duração relativa da vogal nuclear em relação à sílaba das produções, em contexto IR e I, realizadas por ST1 e ST4. Os dados reportados para os itens lexicais que abarcam o [r] são referentes às produções nas quais os sujeitos não produziram o *tap*.

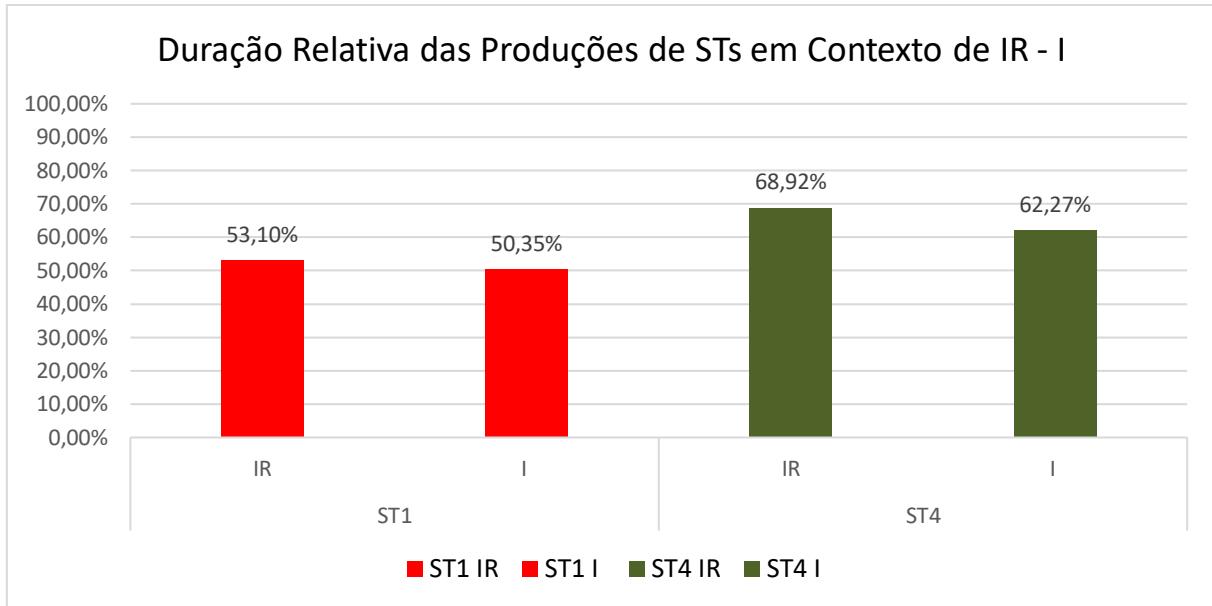


Gráfico 71 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho) e ST4 (verde), em contexto de IR e I, para posição final

Ao examinar os dados do Gráfico 71, percebe-se que os índices de porcentagens de duração relativa da vogal, oriunda da omissão do *tap*, apresentam valores superiores aos das vogais presentes nas palavras CV, ainda que aproximados. Para ST1, as palavras que abarcavam o rótico exibiam uma média de 53,10% de duração relativa, ao passo que a vogal nuclear dos itens lexicais que possuíam apenas a vogal alta anterior como alvo compreendiam 50,35% da sílaba. Analisando as produções de ST4, é possível observar que os índices para IR são de 68,92%, ao passo que I apresenta 62,27% de duração relativa da vogal em relação à sílaba.

Considerando as diferenças de duração absoluta e de duração relativa para a vogal nuclear [i], entre CVC e CV, é possível inferir que ambos os sujeitos parecem marcar distinção entre os alvos.

O Gráfico 72 exibe os resultados das médias de duração absoluta da sílaba e da vogal das produções das produções realizadas por ST1, ST2, ST3 e ST4 nos seguintes contextos: (UR) cur, abajur, Artur e (U) tatu. Os valores referentes às palavras compostas por [r] são referentes às produções nas quais os sujeitos transversais utilizam o recurso de omissão do *tap*.

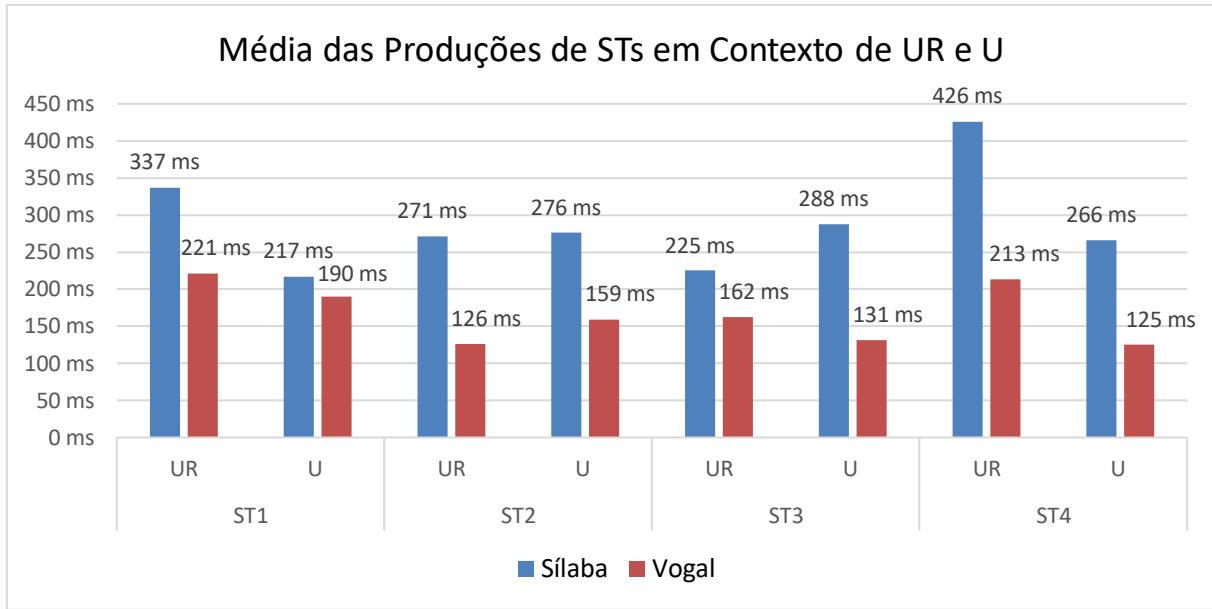


Gráfico 72 - Médias de duração absoluta da sílaba (azul) e da vogal (vermelho) das produções de ST1, ST2, ST3 e ST4, em contexto de UR e U, para a posição final

Tendo em vista que as produções de ST2 e ST3 não são passíveis de comparação, pois correspondem a ocorrências esporádicas, é possível observar que a duração segmental de UR, presente nos dados de ST4, exibe uma diferença expressiva em relação aos dados em contexto de U. A vogal produzida por esse sujeito, oriunda da omissão do *tap*, apresenta 213 ms de duração, à medida que a vogal presente no alvo *cuca* exibe 125 ms de duração absoluta. Nos dados de ST1, essa diferença é, aparentemente, pequena, visto que a vogal articulada em UR possui 221 ms de duração e a vogal presente nas palavras em contexto de U detém 190 ms de duração absoluta.

Um aspecto interessante a ser destacado é que, mesmo com base em apenas uma produção, chama a atenção que a produção de ST2 apresenta uma menor duração da vogal para o alvo CVC quando comparado a CV. A não realização do alongamento compensatório poderia, justamente, ser esperada aqui, uma vez que ST2 não realizou omissões para os alvos AR e IR, o que indicia estar em estágio mais avançado da aquisição do rótico em coda final, quando comparado a sujeitos que apresentaram o alongamento compensatório, como ST4.

O Gráfico 73 exibe os resultados das médias de duração relativa da vogal nuclear em relação à sílaba das produções realizadas por ST1, ST2, ST3 e ST4, em contexto UR e U. Os valores reportados para as palavras compostas por [r] são referentes às produções nas quais as crianças utilizaram o recurso de omissão do *tap*.

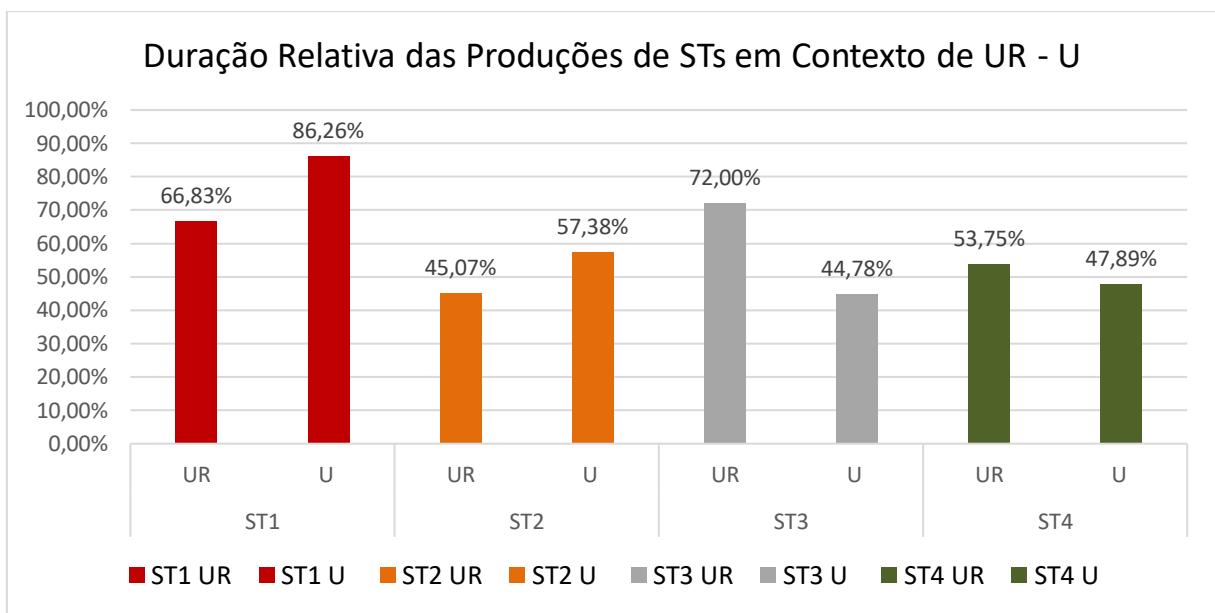


Gráfico 73 - Médias de duração relativa da vogal núcleo da sílaba de ST1 (vermelho), ST2 (laranja), ST3 (cinza) e ST4 (verde), em contexto de UR e U, para posição final

Como já foi mencionado, os dados de ST2 e ST3 não são suscetíveis à análises comparativas. Sendo assim, ao examinar os dados de ST1 e ST4, percebe-se as vogais produzidas por esses sujeitos demostram desempenho diferente no que tange à duração relativa em relação à sílaba. O índice de porcentagem apresentado por ST4, em UR, é maior do que o valor reportado para U. O primeiro é 53,75%, enquanto o último valor exibido é de 47,89%. Ao observar os valores apresentados para ST1, nota-se que o comportamento das vogais é diferente do que recém fora descrito. Para esse informante, a vogal nuclear do alvo CV ocupa 86,26% da sílaba, ao passo que a vogal resultante da omissão do *tap* preenche 66,83% da sílaba.

Os resultados da duração relativa confirmam o já reportado para a duração absoluta, ou seja, a presença de alongamento compensatório nas produções de ST4. O mesmo foi constatado para ST3, ainda que com base em apenas uma produção.

Após a exploração dos dados, nos quais os sujeitos transversais não produzem o *tap*, a Tabela 37 exibe os resultados oriundos das produções nas quais as crianças semivocalizam o segmento. Com o intuito de observar o comportamento do ditongo proveniente da semivocalização do rótico, a referida tabela exibe as médias de duração absoluta da sílaba e do ditongo das palavras que, em posição final, abarcavam o *tap* como alvo – mar, lugar, azar, vir, Nadir, Zair, cur, abajur, Artur. Do

mesmo modo, são exibidos os valores de duração relativa da vogal em relação à sílaba.

FINAL			
		Duração absoluta (ms)	Duração relativa (%)
		Sílaba	(ditongo - sílaba)
AR	ST1	609	466
	ST2	—	—
	ST3	—	—
	ST4	—	—
	<b>MÉDIA</b>	<b>609</b>	<b>466</b>
IR	<b>DP</b>	—	—
	ST1	659	294
	ST2	—	—
	ST3	—	—
	<b>MÉDIA</b>	<b>659</b>	<b>294</b>
UR	<b>DP</b>	—	—
	ST1	370	263
	ST2	—	—
	ST3	—	—
	*ST4	431	373
	<b>MÉDIA</b>	<b>400</b>	<b>318</b>
	<b>DP</b>	<b>43,49</b>	<b>77,60</b>

Tabela 37- Médias de duração absoluta e relativa dos alvos CVC em posição final, produzidos com o recurso da semivocalização, pelos sujeitos transversais, conforme contexto vocálico

\* Valores obtidos com base em uma única produção, assim, não constitui média

De acordo com a Tabela 37, é possível observar que o *tap* foi semivocalizado em todos os ambientes vocálicos para ST1 e apenas em contexto da vogal alta posterior para ST4. No entanto, faz-se necessário tecer algumas observações sobre os dados produzidos pelos referidos informantes.

Para ST1, apenas cinco semivocalizações foram percebidas nas suas produções e são oriundas das palavras *mar*, *Zair*, *cur* e *Artur*, sendo esse último produzido em duas repetições. O sujeito ST4 semivocalizou apenas uma produção para o alvo *cur*. Desse modo, não é possível tecer nenhuma inferência acerca de tais produções, uma vez que essa amostra não é suficiente para estimar o comportamento real do fenômeno.

A partir da análise dos aspectos duracionais das produções dos sujeitos transversais, percebe-se que apenas ST2 e ST3 articulam o *tap* em posição medial.

Quando o segmento está acompanhado pela vogal baixa, a média de duração do rótico é de 20 ms e 15 ms, respectivamente. Em contexto da vogal alta posterior, esse valor é de 21 ms para ST2 e 23 ms para ST3. Sendo assim, é possível afirmar que a duração nas produções do *tap* se encontra dentro do esperado.

Quando a criança não consegue articular o *tap* e emprega o recurso da omissão, é possível inferir que ST1 exibe produções de vogais mais alongadas, especialmente em contexto de vogais altas. Já para ST2, esse comportamento é percebido nos dados em contexto de vogais baixas e altas anteriores. Do mesmo modo, as produções de ST4 exibem maiores valores de duração para a vogal alta posterior em contexto CVC do que em contexto CV. Quanto às produções de ST3, não é possível afirmar que a criança marque algum contraste entre as produções a partir do parâmetro de duração das vogais. Ao comparar as produções dos sujeitos transversais com as produções do sujeito controle, corroboram-se os resultados de ST2 em contexto de [a] e [i]. Para ST1, apenas a vogal alta posterior apresenta duração maior do que as vogais articuladas por S0. Quanto às produções de ST4, constatou-se que [u] exibe maior duração em contexto CVC do que CV. Além disso, foi contatado que as produções em contexto de [i] apresentam maior duração do que as produções do sujeito controle.

Ainda sobre a posição medial, o recurso de semivocalização do *tap* exibe um padrão contrário à omissão. Menores valores de duração relativa das produções semivocalizadas parecem estabelecer contraste entre as produções em contexto CVC e CVV. Essa característica é evidenciada nos dados de ST1 para os dois contextos investigados, AR e UR. Enquanto ST4 apresentou valores menores de duração para as semivocalizações em contexto de AR.

Ao analisar os dados referentes à posição final, é possível identificar que ST2, ST3 e ST4 articulam o *tap* a depender do ambiente vocálico. O informante ST2 produz o segmento em contexto de vogal baixa e vogal alta posterior. Já ST3 produz a consoante em todos os contextos vocálicos analisados. Por fim, ST4 articula o *tap* apenas quando esse está precedido da vogal alta posterior.

Durante o emprego do recurso de omissão do rótico, percebeu-se a possível presença de alongamento compensatório nas produções de ST1 em contexto de vogal baixa e da vogal alta anterior. O informante ST4 parece alongar todas as vogais analisadas. Já o recurso de semivocalização foi utilizado na posição final para produções esporádicas de ST1 e ST4.

#### 4.2.3.3 Trajetória formântica

A análise da trajetória formântica nos dados dos sujeitos transversais busca investigar a existência de produções gradientes presentes na fala das crianças. O método de análise feito na presente seção segue a comparação feita nos dados da informante longitudinal. A trajetória formântica das produções classificadas como omissão ou semivocalização pela análise de oitiva foi comparada com as palavras em contexto CV e CVV.

Sendo assim, foram gerados gráficos das trajetórias formânticas dos três primeiros formantes a fim de observar o movimento dessas frequências ao longo do tempo. Tais gráficos são provenientes de uma produção do alvo e serão apresentados alguns contrastes encobertos detectados, uma vez que não é possível a exposição de todos os dados.

Os Gráficos 74 e 75 apresentam as trajetórias formânticas de F1, F2 e F3 de produções realizadas por ST1, em contexto de IR e I (circo X chico). O *tap* foi omitido na produção de *circo*.

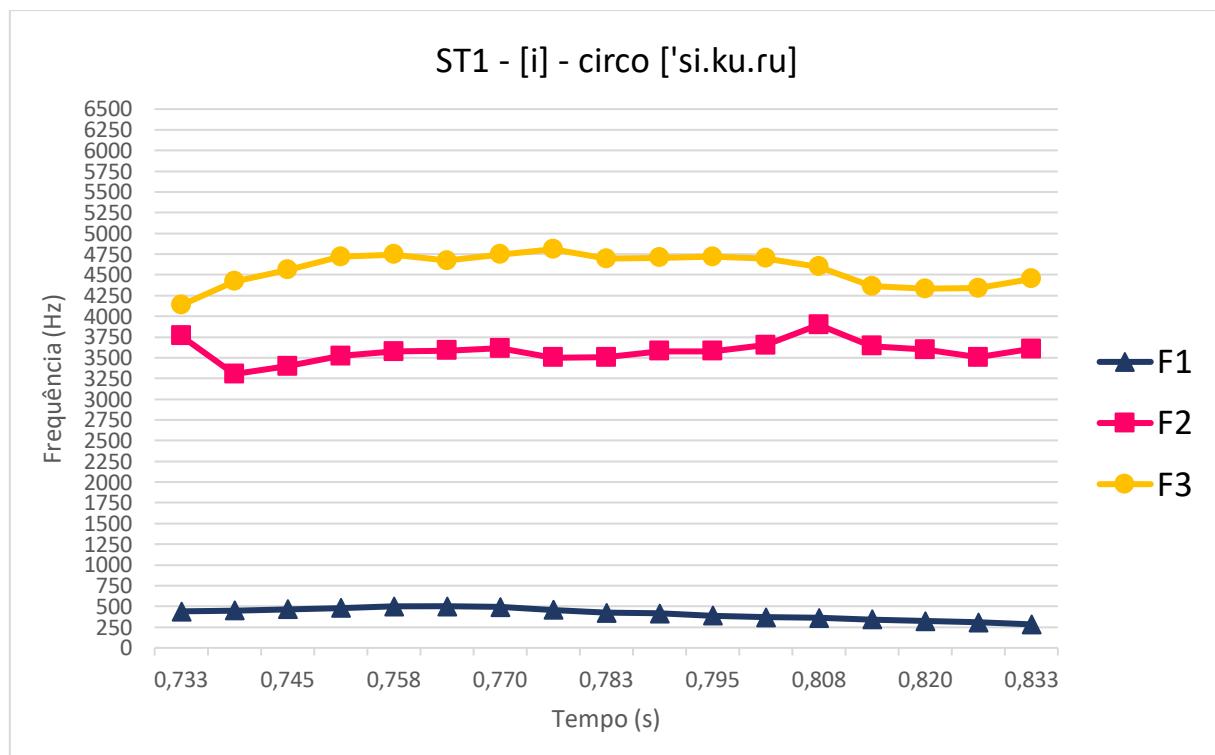


Gráfico 74 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] resultante da omissão do rótico em ['si.ku], produzida por ST1

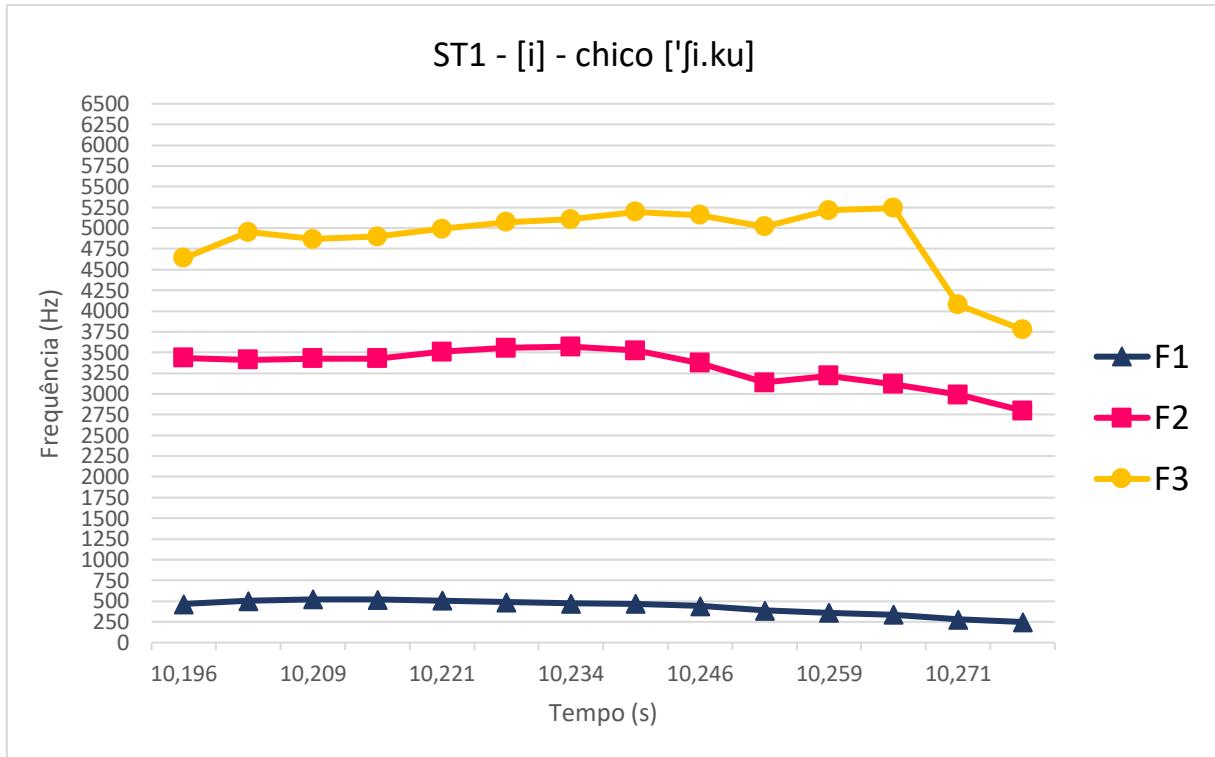


Gráfico 75 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] de ['ʃi.ku], produzida por ST1

Ao analisar a trajetória formântica dos alvos *círculo* e *Chico*, concluiu-se que, para a omissão do *tap*, a partir do Gráfico 74, as frequências de F1, F2 e F3 exibem-se relativamente alinhadas, com suaves quedas em F1 e F3 na porção final. Para a produção do alvo em CV, no Gráfico 75, nota-se que o segundo e o terceiro formantes apresentam queda, sendo de forma mais intensa para F3. Ao comparar o valor de F2 e F3 na porção final da vogal das duas produções, percebe-se um comportamento distinto. A vogal resultante da omissão apresenta F2 de 3607 Hz e F3 de 4453 Hz, à medida que a vogal em contexto CV exibe F2 de 2799 Hz e F3 de 3775 Hz. Rodrigues (2007) observou um comportamento parecido para a omissão do *tap* em contexto de [i]. O valor de F2 na vogal resultante da omissão permaneceu maior do que para a vogal em contexto CV. Desse modo, percebe-se que ST1 marca uma distinção entre o [i] resultante da omissão e o [i] da estrutura CV. Assim, revela-se um possível contraste encoberto presente da produção do alvo *círculo*, envolvendo os valores do segundo e terceiro formantes.

Os Gráficos 76, 77 e 78 apresentam as trajetórias formânticas de F1, F2 e F3 de produções realizadas por ST2, em contexto de AR e A (carta X cata).

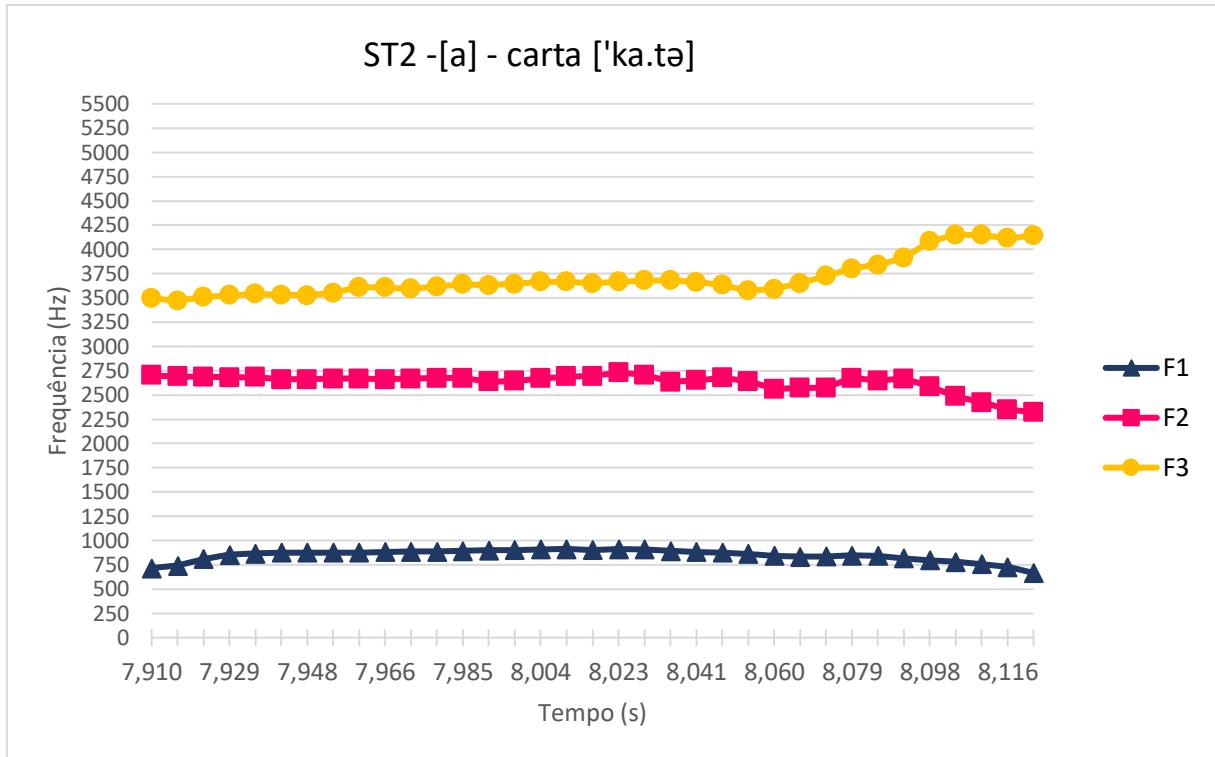


Gráfico 76 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [a] resultante da omissão do rótico em ['ka.tə] produzida por ST2

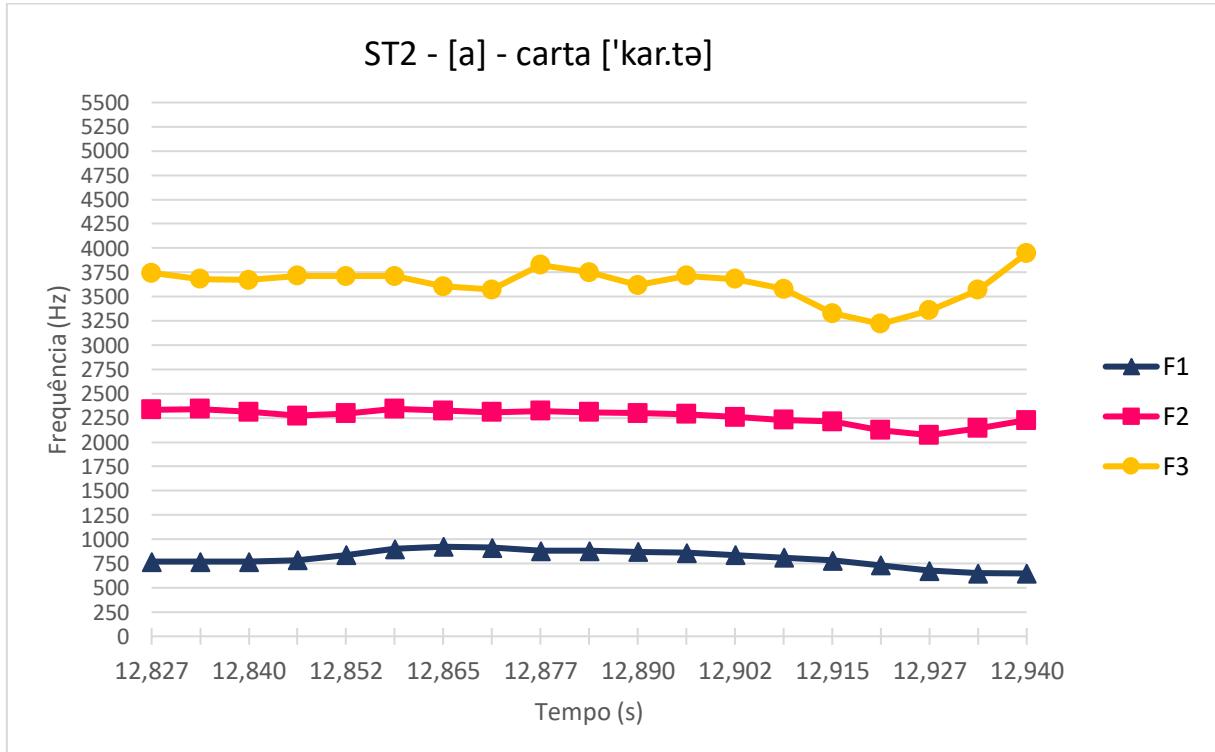


Gráfico 77 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [a] que antecede o rótico na produção de ['kar.tə'] por ST2

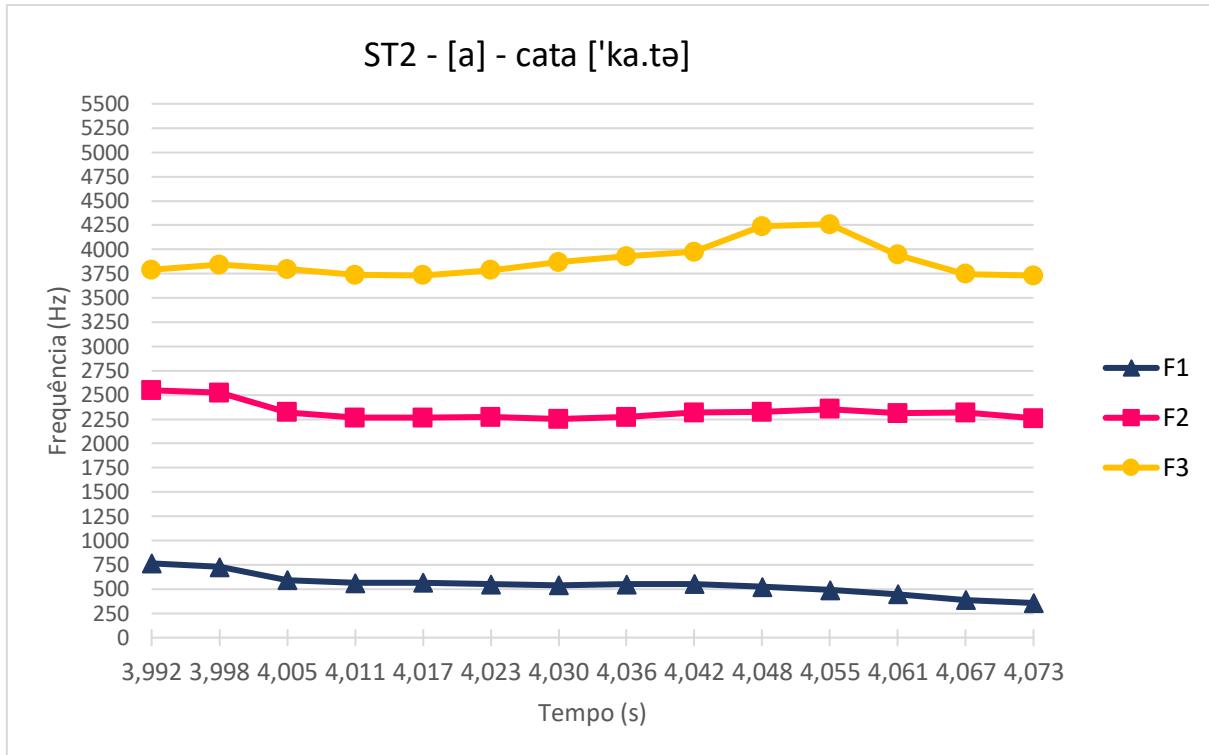


Gráfico 78 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [a] de ['ka.tə] produzida por ST2

Com base no Gráfico 76, observa-se que a produção da vogal resultante da omissão do *tap* exibe o primeiro formante horizontalmente alinhado, ao passo que as frequências do segundo formante caem e do terceiro formante se elevam. Para a vogal que antecede a produção do rótico, percebe-se uma queda de F1 e elevação de F3. Em relação à vogal produzida em contexto CV, nota-se que o terceiro formante apresenta uma sensível curva de subida e descida. Ao observar a trajetória formântica presente nos Gráficos 76 e 77, percebe-se que os formantes mantêm um comportamento parecido, além de frequências em faixas próximas. No entanto, ao comparar os dados presentes nos Gráficos 76 e 78, nota-se, a partir do ponto médio, para a vogal resultante da omissão, F1 de 878 Hz e F3 de 3759 Hz, à medida que a vogal CV exibe F1 de 548 Hz e F3 de 3912 Hz.

Os Gráficos 79 e 80 exibem as trajetórias formânticas de F1, F2 e F3 em dados produzidos por ST3, em contexto de IR e I (circo X Chico). O *tap* foi omitido na articulação de *circo*.

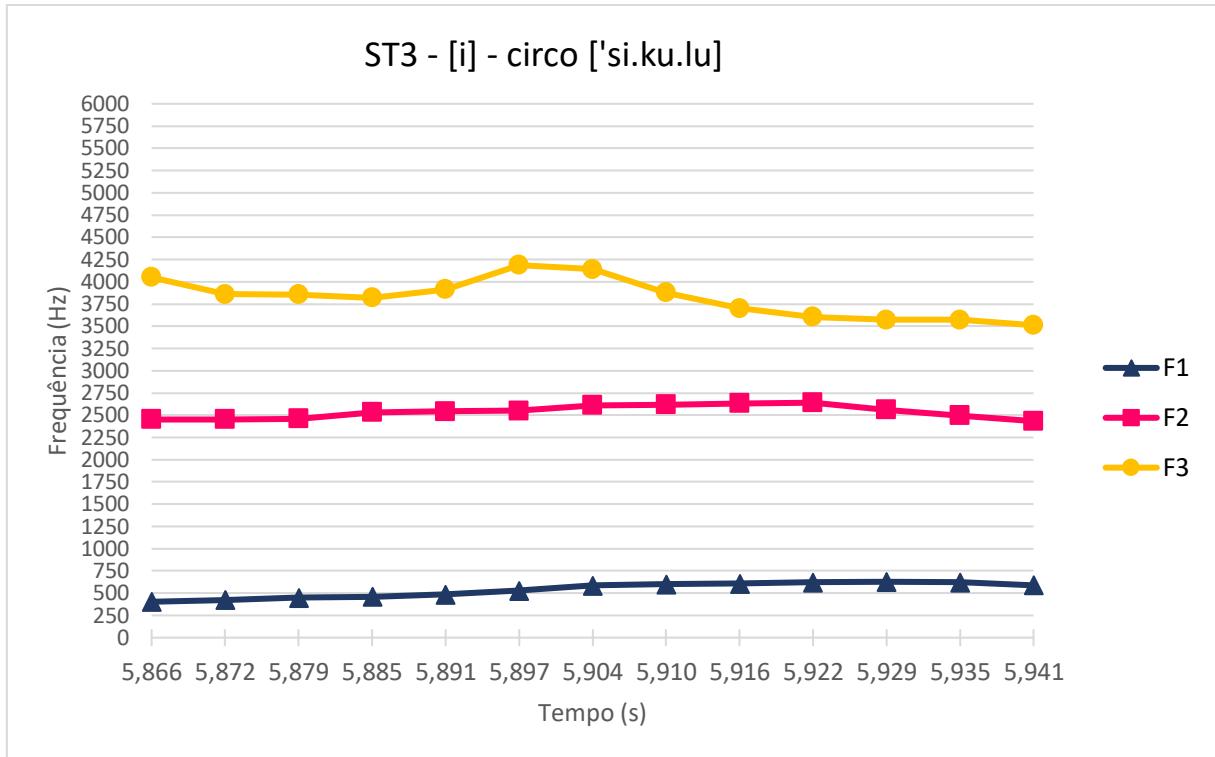
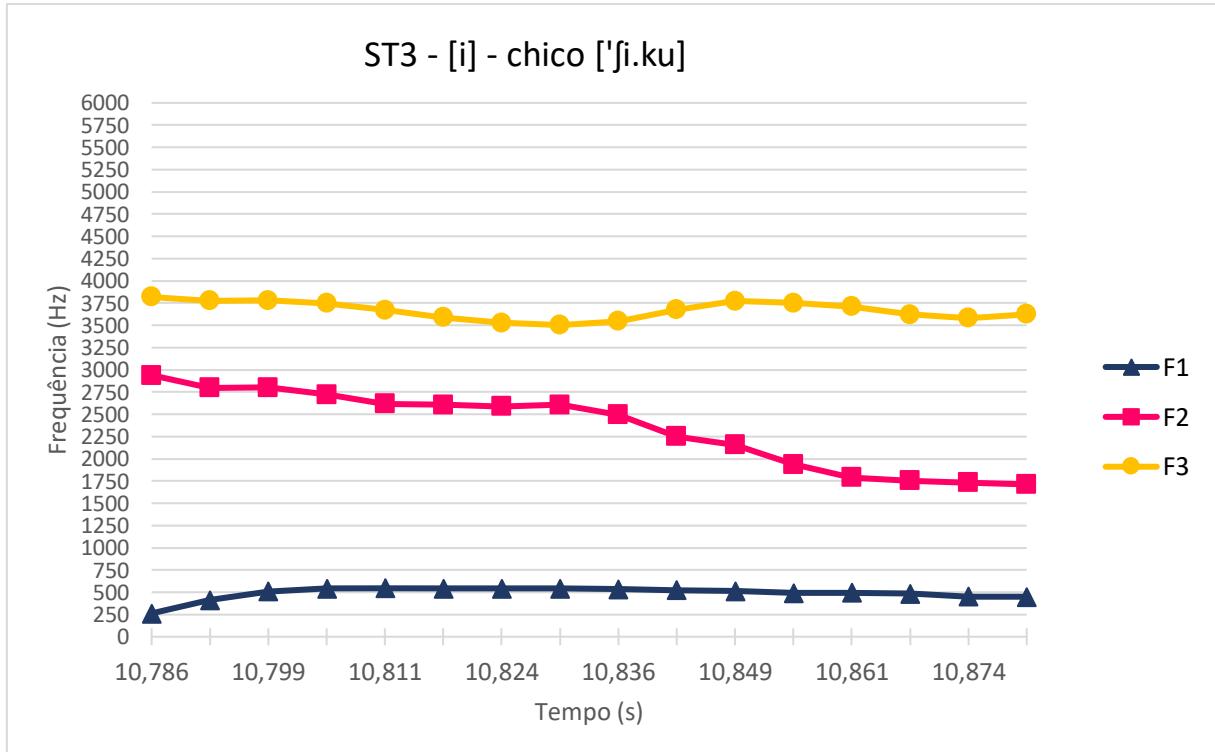


Gráfico 79 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] resultante da omissão do rótico em ['si.ku.lu] por ST3



Observando os Gráficos 79 e 80, observa-se que, para a vogal resultante da omissão, os dois primeiros formantes estão horizontalmente alinhados, enquanto o terceiro apresenta uma sensível queda. No entanto, para vogal presente na estrutura CV, apenas o primeiro formante encontra-se horizontalmente alinhado, à medida que as frequências do segundo e terceiro formante se afastam. Ao comparar o valor de F2 na porção final das duas produções, tem-se, para [i], resultante da omissão, o valor de 2434 Hz e, para o [i], produzido em CV, 1713 Hz. Ainda, o comportamento dos formantes ao longo do tempo, presentes no Gráfico 79, assemelham-se à trajetória formântica exibida durante a produção de [i] no alvo *circo*, pelo sujeito controle, conforme Gráfico 23.

Após essas observações, percebe-se que, mais uma vez, a produção gradiente, resultante da omissão do rótico, foi identificada por meio do valor exibido pelo segundo formante. Assim, pode-se inferir que, articulatoriamente, ST3 realizou uma contrição em uma parte mais anterior do trato ao produzir a vogal resultante da omissão, devido ao seu valor mais alto para F2.

Os Gráficos 81 e 82 apresentam as trajetórias formânticas de F1, F2 e F3 de produções realizadas por ST1, em contexto de IR e I (circo X Chico). O *tap* foi omitido na produção de *circo*.

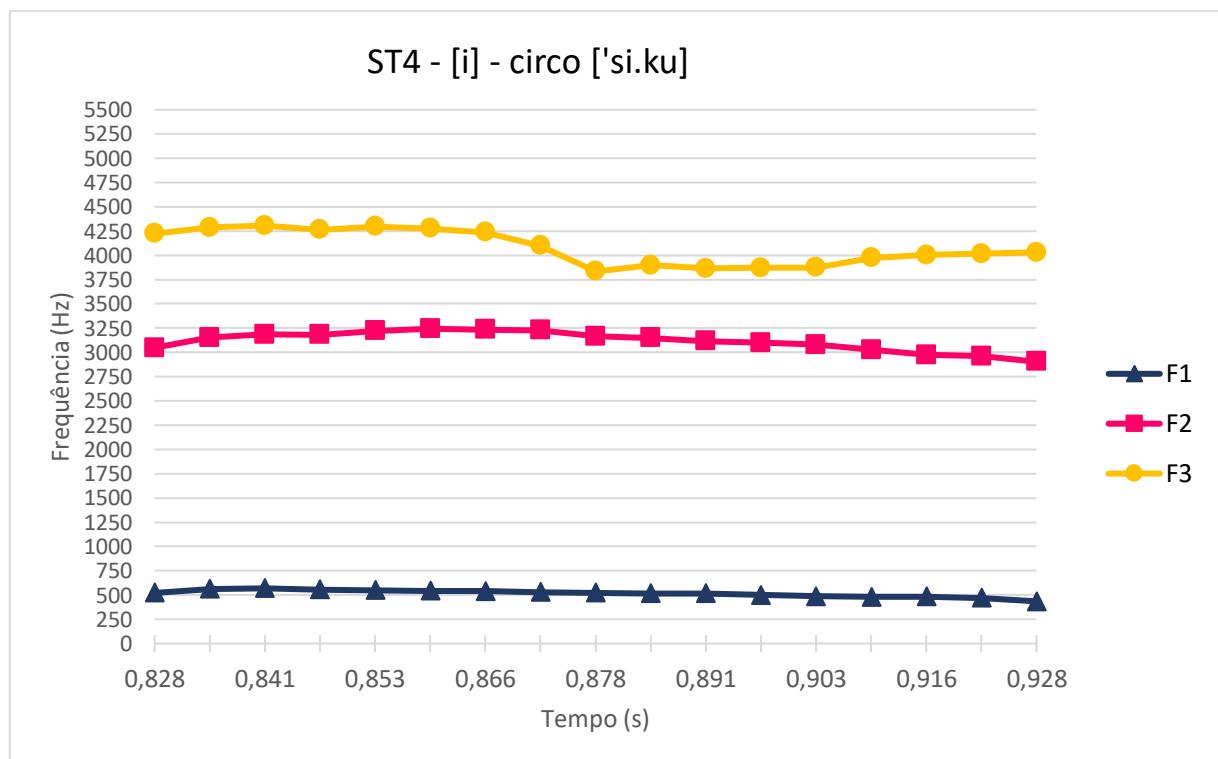


Gráfico 81 - Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] resultante da omissão do rótico em ['si.ku'] por ST4

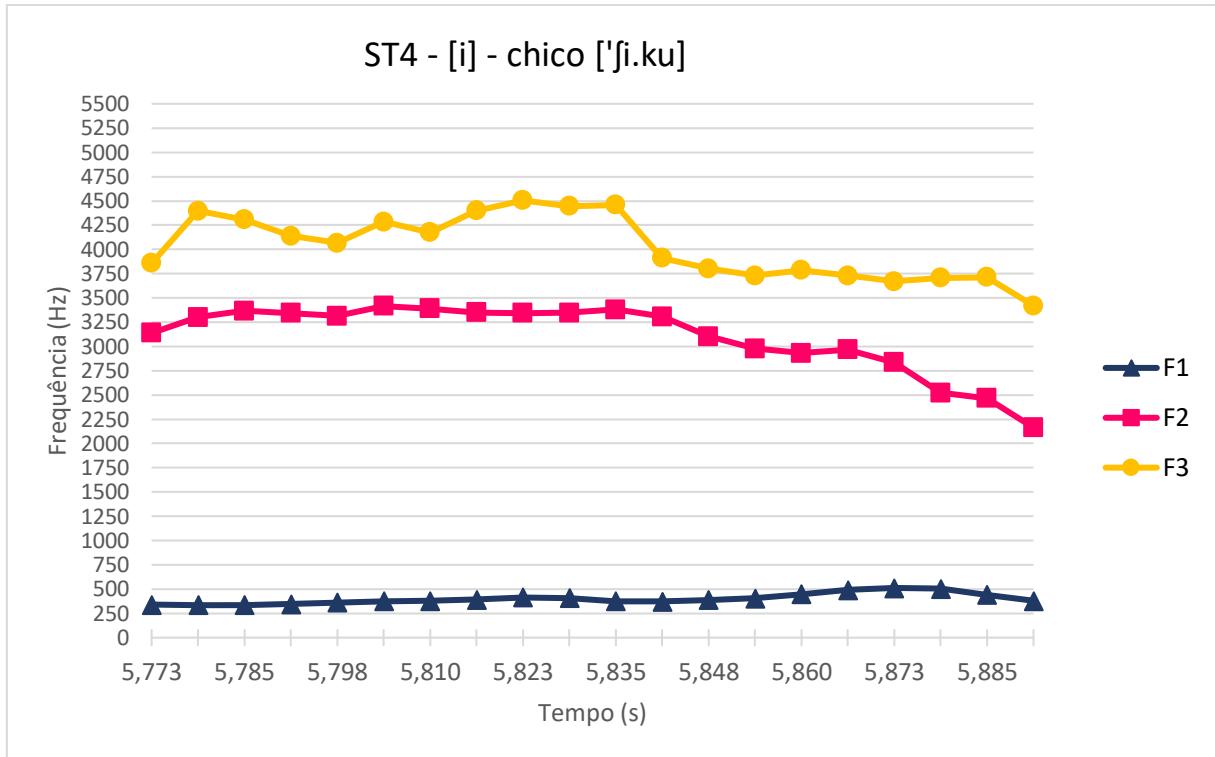


Gráfico - 82 Trajetórias formânticas relacionadas a F1 (azul), F2 (rosa) e F3 (amarelo) da vogal [i] de ['ʃi.ku] por ST4

Ao examinar a trajetória formântica das palavras *circo* e *Chico*, percebe-se que, para a omissão do *tap*, a partir do Gráfico 81, as frequências de F1 e F2 permanecem horizontalmente alinhadas, enquanto F3 exibe uma pequena curva de descida e subida. Para a produção da vogal em CV, no Gráfico 82, nota-se que o primeiro formante está alinhado na horizontal, ao passo que o segundo e o terceiro apresentam queda, sendo de forma mais intensa para F2. Ao aferir o valor de F2 e F3 no final da vogal das duas produções, observa-se um comportamento distinto. A vogal oriunda da omissão apresenta F2 de 2903 Hz e F3 de 4028 Hz; já a vogal em contexto CV exibe F2 de 2165 Hz e F3 de 3421 Hz. Com isso, o valor do segundo formante na vogal referente à omissão é maior do que para a vogal em estrutura CV, o que pode indicar uma constrição na região mais anterior do trato. Portanto, um possível contraste encoberto presente da produção do alvo *circo* é percebido por meio da trajetória formântica de F2 mais elevada, inferindo, assim, uma constrição mais anterior no trato vocal daquela apresentada pela vogal [i] na palavra *chico*.

Ao retomar a análise dos resultados obtidos pelos gráficos de trajetória formântica, é possível perceber que, para a produção de [i] resultante da omissão do rótico, os valores de F1, F2 e F3 apresentam-se horizontalmente alinhados, à medida

que a vogal presente em silva CV apresenta modificações no que concerne aos valores de F2 e F3. Quanto à produção de [a] em contexto CVC, é possível verificar um comportamento similar para a sua produção e sua omissão, a elevação de F3. Para a vogal em contexto CV, percebe-se um movimento de subida e descida para F3. Embora a análise da trajetória formântica tenha se dado de forma visual, é possível perceber que há contrastes encobertos em andamento entre a produção do *tap* e sua omissão. Todos os informantes transversais parecem marcar de alguma forma uma distinção entre as palavras por meio de sutis alterações nas constrições no trato vocal, apesar de essas serem imperceptíveis ao ouvido humano.

## 5 Conclusão

Esta pesquisa buscou descrever e analisar a emergência dos gestos articulatórios do *tap* em sílaba CVC, em posição medial e final de palavra a partir de aspectos acústicos. Para tanto, foram investigadas produções de cinco crianças em fase de aquisição do português e uma criança com a fase de aquisição fonético-fonológica já concluída.

Desse modo, considerando os resultados obtidos ao longo das análises propostas neste trabalho, serão retomadas as questões de pesquisa e suas hipóteses apresentadas na Introdução.

A primeira questão de pesquisa (Q1) aborda se a produção do rótico apresentará o padrão de *tap*, embora algumas variações possam ser identificadas. A hipótese de investigação postula que sim, pois, apesar de a maioria dos róticos produzidos nessas posições apresentar um padrão de *tap* (MEZZOMO, 1999, 2004; MIRANDA, 1996), algumas variações do segmento serão observadas. É necessário destacar que o trabalho de Melcheque, Ferreira-Gonçalves e Brum-de-Paula (2021) encontrou a produção de vibrantes alveolares em sujeitos adultos, especialmente em coda final. Sendo assim, esse tipo de segmento pode também aparecer na fala das crianças.

Essa hipótese foi confirmada, uma vez que a análise dos dados do sujeito controle revelou que a variante rótica mais empregada, em posição medial de palavra, foi o *tap*. Além dessa variante, foram observadas produções de vibrantes acompanhadas pela vogal baixa e pela vogal alta posterior. Para posição final de sílaba, o *tap* não ocorreu em nenhum ambiente vocálico. No entanto, a vibrante foi a variante predominante. Do mesmo modo, outras duas variantes para o rótico foram observadas, como a aproximante retroflexa, fricativa glotal e *tap* fricatizado. A informante longitudinal não produziu nenhuma variante rótica. Em relação aos dados dos sujeitos transversais, a análise acústica revelou que, quando o rótico era articulado, a maioria das produções deu-se na forma de *tap*. Além disso, foram observadas a produção das seguintes variantes: vibrante, fricativa glotal, aproximante alveolar e *tap* fricatizado.

A segunda questão de pesquisa (Q2) investiga se a criança usará principalmente a estratégia de redução gestual a fim de alcançar a produção de um

*tap*. A hipótese de investigação afirma que sim, pois espera-se que as crianças utilizem, na maior parte dos casos, a estratégia de redução da magnitude dos gestos, visto que, no trabalho de Vassoler (2016), as crianças típicas valeram-se dessa estratégia a fim de produzir o *tap* presente em sílaba CCV. Já nas produções das crianças atípicas, foi constatada uma maior realização de gestos intrusivos, apresentando dificuldades em identificar os gestos de ponta e de dorso de língua. A autora infere que esse fato ocorre em virtude dos desvios e atrasos de controle dessas regiões do corpo da língua.

Essa hipótese foi confirmada parcialmente, uma vez que, ao serem analisadas as estratégias de produção, quando as crianças não conseguiam produzir o segmento alvo, notou-se que, fundamentalmente, esteve presente a estratégia de redução gestual, o que ocasionou a presença recorrente de contrastes encobertos, no entanto, outras estratégias, como substituições e intrusão de gestos também foram observadas, a depender da posição na palavra, medial ou final.

Infere-se que a estratégia de redução gestual foi amplamente empregada, em posição medial, para a informante longitudinal ao longo das três coletas, pois foram percebidos muitos casos de omissão do *tap*. No entanto, para posição final, foi observada uma alta taxa de emprego da lateral [l], sozinha ou acompanhada da vogal [i] ou [u], o que possibilita inferir que ocorreu a substituição dos gestos de [r] por [l].

No que tange às produções dos sujeitos transversais, é possível inferir que as estratégias de redução e intrusão gestual apresentaram um papel semelhante para a posição medial, uma vez que, acusticamente, não houve o predomínio apenas de omissão ou semivocalização. Já para a posição final, foi observada uma maior utilização do recurso de omissão do rótico, o que permite inferir que a estratégia de redução gestual apresentou uma maior prevalência. Em alguns casos, pode-se inferir que houve uma redução parcial do gesto, pois a trajetória formântica evidenciou que a maioria das vogais resultantes de omissão comportaram-se de modo parecido às vogais produzidas pelo sujeito controle acompanhadas do *tap*. Sendo assim, a redução parcial da magnitude gestual modifica a constrição, mas não há ponto de essa ser perceptível ao ouvido humano.

Para a terceira questão de pesquisa (Q3), é verificado se a análise acústica será capaz de revelar contrastes encobertos que não foram percebidos na análise de oitiva. A hipótese de investigação afirma que sim, pois, conforme constatou Rodrigues (2007), a trajetória formântica foi um dos principais parâmetros que revelou a presença

desses contrastes no processo de aquisição do rótico pela criança. Barbieri (2019) também encontrou contrastes encobertos na produção do *tap* em sílaba CCV na fala infantil.

Essa hipótese foi confirmada, uma vez que as análises subsidiadas pela acústica evidenciaram algum tipo de contraste encoberto. A inspeção visual dos espectrogramas apresentou evidências de que, embora o *tap* não tenha sido articulado em sua totalidade, produções gradientes foram detectadas. Visto que, por meio das pistas acústicas encontradas, foi possível inferir que houve equívocos na coordenação dos gestos, à medida que há algum tipo de correlato acústico da produção do segmento, mas esse não é percebido na análise de oitiva.

Assim, os casos em que mais de um espaçamento em branco foi identificado no espectro demostram uma clara tentativa de articulação do som conforme o alvo, ou seja, uma vibrante, especialmente em posição final. A característica mais relevante para a constatação de possíveis contrastes encobertos foi a ausência de energia, seguida pela presença de elementos vocálicos. A posição medial foi o contexto que mais favoreceu a identificação de traços residuais do segmento. A partir disso, é possível inferir que a maioria das crianças apresentam dificuldades em regular a magnitude do gesto de ponta de língua, pois é possível observar o correlato acústico, um espaçamento em branco no espectrograma, do momento de breve obstrução. A baixa recorrência dos correlatos visuais da soltura da oclusão ou de ruídos transientes pode apontar uma falha durante o afastamento dos articuladores. Posto isto, é possível que o segmento se torne inaudível devido à execução imprecisa dos movimentos articulatórios.

A análise da trajetória formântica apresentada para os dados da informante longitudinal e dos sujeitos transversais foi fundamental para evidenciar possíveis contrastes em andamento na fala das crianças, entre a produção do *tap* e sua omissão. Acusticamente, as distinções aparecem apoiarem-se nos valores de F2 e F3. Todos os informantes pareceram marcar de alguma forma uma distinção entre as palavras por meio de sutis alterações nas constrições no trato vocal. No entanto, ainda falta regular a magnitude dos gestos articulatórios para que esses sejam perceptíveis auditivamente. Assim, a trajetória formântica apresenta-se como uma ferramenta eficaz na revelação de produções gradientes.

Por fim, a quarta questão de pesquisa (Q4) investiga se existe relação entre a natureza formântica do elemento vocálico e a vogal núcleo da sílaba. A hipótese de

investigação postula que sim, ainda que o “elemento vocálico que sucede o *tap* em final de palavra parece tender à centralização, assemelhando-se a um schwa” (SILVA; CLEMENTE; NISHIDA, 2006, p. 22), Bilharva-da-Silva (2019) não encontra indícios fortes para afirmar que a qualidade do elemento vocálico seja de uma vogal neutra.

Essa hipótese foi parcialmente confirmada por meio do método experimental de análise proposto por esta pesquisa. Ao usar o método de Silva, Clemente e Nishida (2006), ocorre que os elementos vocálicos produzidos em final de sílaba apresentaram características semelhantes a uma vogal central, no que concerne ao eixo F2. No entanto, os resultados evidenciam que as vogais [a], [i] e [u] mantêm diferenças em relação ao eixo F1, configurando a formação de um triângulo vocálico, ainda que mais central.

Ao observar o comportamento dos elementos vocálicos a partir da comparação entre esses e a porção final de uma vogal nuclear, observa-se, no entanto, que apenas os elementos vocálicos produzidos em contexto de [u] tendem a centralização. Assim, os resultados indiciam uma relação entre o elemento vocálico e a vogal nuclear em sílaba CVC.

Além das questões de pesquisa aqui discutidas, foi possível perceber, por meio da análise de oitiva, que a informante longitudinal não produziu o *tap* em nenhuma posição silábica, no entanto, adotou oito diferentes recursos para se aproximar da produção do segmento. Em posição medial, a principal estratégia empregada foi a omissão do rótico. Para posição final de sílaba, os percentuais de empregos dos recursos estão mais distribuídos entre semivocalização e substituições pela lateral, isolada ou acompanhada por uma vogal. Os sujeitos transversais apresentaram produções do *tap* em ambas posições, sendo esse mais produzido em posição final. Para manejar a ausência do segmento, esses informantes adotaram modos diferentes. A omissão e a semivocalização foram recursos importantes para a posição medial, à medida que, em posição final, a omissão foi a estratégia majoritariamente adotada.

O índice expressivo de ocorrências de [l] nos dados da informante longitudinal parece evidenciar que há um progresso nas tentativas de realização do *tap*, uma vez que [r] e [l] são compostos por gestos de ponta de língua. Desse modo, a emergência da lateral pode ser fruto da falta de regulação da coordenação dos gestos articulatórios, tendo em vista que existem características articulatórias gestuais similares entre os dois sons.

Outro ponto a ser destacado é a criação de uma sílaba CV mais à direita da palavra. Dado que os informantes já adquiriram o segmento em *onset*, há a necessidade de transpor o *tap* para essa posição, por meio do emprego de metátese. Logo, a produção do alvo em sílaba CV parece ser mais simples do que em sílaba CVC.

Por meio dos aspectos duracionais, percebe-se que, quando os sujeitos ST2, ST3 e ST4 articulam um *tap*, suas produções encontram-se dentro do esperado, uma vez que seus valores de duração se aproximam das médias reportadas por S0. Quando os sujeitos ST1, ST2 e ST4 não conseguem produzir o segmento, eles parecem utilizar o alongamento compensatório da vogal proveniente de omissão do *tap*, em alguns ambientes vocálicos, a fim de estabelecer distinção entre palavras de sílabas CVC e CV. Para Miranda e Silva (2011) a produção de uma vogal mais longa conceberia um ambiente para a articulação do seguimento, assim, a partir do aprimoramento das rotinas motoras, a criança pode viabilizar as condições físicas necessárias para produção de um *tap*. O emprego do recurso da semivocalização, nos dados de SL1, ST1 e ST4, parece evidenciar que as crianças estabelecem uma distinção por meio de uma menor duração de suas produções em contexto CVC quando comparadas ao contexto CVV. Nesse caso, justamente, é esperada uma menor duração do glide resultante da semivocalização, uma vez que o *tap* ocupa uma janela temporal menor do que a semivogal.

Sendo assim, a criança parece saber que há um alvo a ser alcançado e lança mão de estratégias a fim de alcançá-lo, no entanto, é necessário um longo caminho para o êxito de sua produção. Além de perceber o segmento e tentar produzi-lo, o aprendiz precisa ajustar o *timing* dos gestos articulatórios. O estudo de Studdert-Kenned e Goodell (1995), ao estudar a aquisição da linguagem com base em produções de uma menina de 2 anos, por um período de 15 semanas, compreenderam que os erros de produção podem ser considerados como execução e *timing* imprecisos dos gestos que constituem a palavra alvo.

Espera-se, assim, que este trabalho tenha se somado a outros, como Rodrigues (2007), que contribuem para evidenciar que produções gradientes e variações de produções, previstas pela Fonologia Gestual, são reveladoras de importantes avanços no percurso de aquisição fonológica. Ainda, que os dados em aquisição da fonologia são reveladores de pistas relevantes para a discussão de

questões teóricas, como a coordenação gestual envolvida na produção de sílabas CVC.

## Referências Bibliográficas

- ALBANO, Eleonora Cavalcante. **O gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro.** Campinas: Mercado das Letras, 2001.
- ALBANO, Eleonora Cavalcante. Representações dinâmicas e distribuídas indícios do português brasileiro adulto e infantil. **Letras de Hoje**, Porto Alegre, v. 42, n. 1, p. 131–150, 2007.
- ALBANO, Eleonora Cavalcante. Uma introdução à dinâmica em fonologia, com foco nos trabalhos desta coletânea. **Revista da ABRALIN**, v. 11, n. 1, 2012.
- ALVARENGA, Daniel; OLIVEIRA, Marco Antônio de. Canonicidade silábica e a aprendizagem da escrita. **Revista de Estudos Linguísticos**, Belo Horizonte, v. 1, n. 5, p. 127–158, 1997.
- BARBERENA, Luciana da Silva. **Caracterização de Aspectos da Produção Articulatória do [r] tap por Análise Instrumental e Resultados de Intervenção.** 2016. Tese (Doutorado em Distúrbios da Comunicação Humana) - Centro de Ciências da Saúde, Universidade Federal de Santa Maria, Santa Maria, 2016.
- BARBERENA, Luciana da Silva; KESKE-SOARES, Márcia; BERTI, Larissa Cristina. Descrição dos gestos articulatórios envolvidos na produção dos sons /r/ e /l/. **Audiology - Communication Research**, São Paulo, v. 19, n. 4, p. 338–344, 2014.
- BARBIERI, Thais Telles. **Aquisição de encontros consonantais com tap no português brasileiro: análises acústica e articulatória.** 2019. Dissertação (Mestrado em Letras) - Centro de Letras e Comunicação, Universidade Federal de Pelotas, Pelotas, Pelotas, 2019.
- BARBOSA, Plínio Almeida; MADUREIRA, Sandra. **Manual de fonética acústica experimental: aplicações a dados do português.** São Paulo: Cortez, 2015.

BATES, Elizabeth; DALE, Philip S.; THAL, Donna. Diferenças individuais e suas implicações para as teorias do desenvolvimento da linguagem. *In: FLETCHER, Paul; MACWHINNEY, Brian (org.). Compêndio da Linguagem da Criança*. Tradução Marcos A. G. Domingues. Porto Alegre: Artes Médicas, 1997. p. 87–130.

BILHARVA-DA-SILVA, Felipe. **O contato português-pomerano na produção dos grupos [cr] e [rc]: o caso das vogais suarabácticas**. 2019. Tese (Doutorado em Linguística) - Escola de Humanidades, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2019.

BISOL, Leda. A sílaba e seus constituintes. *In: Gramática do português falado. Vol. VIII*. Campinas: Editora da Unicamp, 1999.

BOERSMA, Paul; WEENINK, David. **Praat: Doing phonetics by computer**. Amsterdam: Phonetic Sciences, 2019. Version: 6.0.56.

BONILHA, Giovana Ferreira Gonçalves. **Aquisição fonológica do português brasileiro: uma abordagem conexionista da Teoria da Otimidade**. 2005. Tese (Doutorado em Letras) - Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2005.

BOTASSINI, Jacqueline Ortelan Maia. A variação no uso dos róticos em Porto Alegre. **Estudos Linguísticos**, São Paulo, v. 40, n. 2, p. 1060–1072, 2011.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Towards an articulatory phonology. **Phonology Yearbook**, v. 3, p. 219–252, 1986.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Tiers in articulatory phonology, with some implications for casual speech. **Haskins Laboratories Status Report on Speech Research**, v. 92, p. 1987, 1987.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Some notes on syllable structure in articulatory phonology. **Haskins Laboratories Status Report on Speech Research**, p. 85–102, 1988.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Articulatory gestures as phonological units. **Phonology**, v. 6, n. 2, p. 201–251, 1989 a.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Gestural structures and phonological patterns. **Haskins Laboratories Status Report on Speech Research**, v. 97/98, p. 1–23, 1989 b.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Gestural specification using dynamically-defined articulatory structures. **Haskins Laboratories Status Report on Speech Research**, v. 103/104, p. 95–110, 1990 a.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Representation and reality: physical systems and phonological structure. **Haskins Laboratories Status Report on Speech Research**, v. 105/106, p. 83–92, 1990 b.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Articulatory Phonology: an overview. **Phonetica**, v. 49, p. 155–180, 1992.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Dynamics Articulatory Phonology. **Haskins Laboratories Status Report on Speech Research**, v. 113, p. 51–62, 1993.

CAGLIARI, Luiz Carlos. Análise Fonológica. *In: Introdução à teoria e à prática com especial destaque para o modelo fonêmico*. Campinas: Mercado das Letras, 2002.

CALLOU, Dinah; MORAES, João A.; LEITE, Yonne. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. *In: KOCH, Ingedore G. Villaça (org.). Gramática do português falado. Vol. VI: desenvolvimentos*. Campinas: Editora da Unicamp, 1996. p. 465–493.

CÂMARA JR., Joaquim Mattoso. **Para um estudo da fonêmica portuguesa**. Rio de Janeiro: Simões, 1953.

CÂMARA JR., Joaquim Mattoso. **Estrutura da língua portuguesa**. Rio de Janeiro: Vozes, 1977.

CÂMARA JR., Joaquim Mattoso. **História e estrutura da língua portuguesa**. Rio de Janeiro: Padrão, 1979.

CHOMSKY, Noam. *Lectures on government and binding*. **Foris**, Dodricht, 1981.

CHOMSKY, Noam. *Knowledge of language: its nature, origin and use*. **Praeger**, Londres, 1986.

CIRIGLIANO, R. J. R. *et al.* Um Conjunto de 1000 Frases Foneticamente Balanceadas para o Português Brasileiro Obtido Utilizando a Abordagem de Algoritmos Genéticos. *In: XXII Simpósio Brasileiro de Telecomunicações (SbrT 2005)*. Campinas, 2005. p. 110–114.

COLLISCHONN, Gisela. **Análise prosódica da sílaba em Português**. 1997. Tese (Doutorado em Linguística e Letras) - Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1997.

CRISTÓFARO-SILVA, Thaïs; YEHIA, Hani Camille. **Sonoridade em Artes, Saúde e Tecnologia**. 2009. Disponível em: [http://www.fonologia.org/fonetica\\_consoantes.php](http://www.fonologia.org/fonetica_consoantes.php). Acesso em: 2 maio. 2017.

FERRAZ, Irineu da Silva. **Características fonético-acústicas do /r/ retroflexo do português brasileiro: dados de informantes de Pato Branco (PR)**. 2005. Dissertação (Mestrado em Linguística) - Universidade Federal do Paraná, Curitiba, 2005.

FOWLER, Carol A. Coarticulation and theories of extrinsic timing. **Journal of Phonetics**, v. 8, p. 113–133, 1980.

FREITAS, Maria Cláudia C. de; ALBANO, Eleonora Cavalcante. O gesto fônico na aquisição “desviante”: movimentos entre a produção e a percepção. **Revista da ABRALIN**, v. 11, n. 1, p. 197–220, 2012.

FUDGE, E. Syllables. **Journal of Linguistics**, Cambridge, UK, v. 5, p. 254–287, 1969.

GICK, Bryan *et al.* A motor differentiation model for liquid substitutions in children’s speech. **The Journal of the Acoustical Society of America**, v. 1, p. 1–7, 2007.

GOLDSTEIN, Brian; CINTRÓN, Patricia. An investigation of phonological skills in Puerto Rican Spanish-speaking 2-year-olds. **Clinical Linguistics and Phonetics**, v. 15, n. 5, p. 343–361, 2001.

GOLDSTEIN, Louis *et al.* Dynamic action units slip in speech production errors. **Cognition**, v. 103, p. 386–412, 2007.

HERNANDORENA, Carmen Lúcia Matzenauer. **A aquisição da fonologia do português: estabelecimento de padrões com base em traços distintivos**. 1990. Tese (Doutorado em Linguística e Letras) - Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1990.

HERNANDORENA, Carmen Lúcia Matzenauer; LAMPRECHT, Regina Ritter. A aquisição das consoantes líquidas do português. **Letras de Hoje**, Porto Alegre, v. 32, n. 4, p. 7–22, 1997.

ILHA CERON, Marizete *et al.* Normative Features of Phoneme Acquisition in Brazilian Portuguese. **Folia Phoniatrica et Logopaedica**, v. 72, n. 3, p. 228–241, 2019.

JAKOBSON, Roman. **Child Language Aphasia Phonolog**. The Hague: Mouton, 1968.

KEHOE, Margaret M.; STOEL-GAMMON, Carol. Development of syllable structure in English-speaking children with particular reference to rhymes. **Journal of Child Language**, v. 28, n. 2, p. 393–432, 2001.

KENT, Ray; READ, Charles. **Análise acústica da fala**. Tradução Alexsandro Meireles. 1. ed. São Paulo: Cortez, 2015.

LADEFOGED, Petter; MADDIESON, Ian. Rhotics. *In: The sounds of the world's languages*. Oxford: Blackwell, 1996. p. 215–245.

LAMPRECHT, Regina Ritter. **Perfil da aquisição normal da fonologia do Português. Descrição longitudinal de 12 crianças: 2:9 a 5:5**. 1990. Instituto de Letras e Artes, Tese (Doutorado em Letras) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1990.

LAMPRECHT, Regina Ritter. **Aquisição fonológica do português: perfil de desenvolvimento e subsídios para terapia**. Porto Alegre: Artmed, 2004.

LINDAU, Mona. The story of /r/, in Fromkin, V. **Phonetic Linguistics: Essays in Honor of Peter Ladefoged**, p. 157–168, 1985.

MADDIESON, Ian. **Patterns of Sounds**. Cambridge, UK: Cambridge University Press, 1984.

MATEUS, M. H.; D'ANDRADE, E. **The Phonology of Portuguese**. Oxford: Oxford University Press, 2000.

MELCHEQUE, Patrícia Pereira; FERREIRA-GONÇALVES, Giovana; BRUM-DE-PAULA, Mirian Rose. The gestural configuration of the rhotic sounds in the pelotense dialect. *In: Trabalho apresentado no Workshop em Estudos Ultrassonográficos de Dados de Fala*. Pontifícia Universidade Católica de São Paulo, São Paulo: 10 dez, 2018.

MELCHEQUE, Patrícia Pereira; FERREIRA-GONÇALVES, Giovana; BRUM-DE-PAULA, Mirian Rose. Configuração gestual do *tap* na fala de pelotenses. *In: FERREIRA-GONÇALVES, Giovana; BRUM-DE-PAULA, Mirian Rose (org.). A ultrassonografia e os gestos da fala*. São Carlos: Pedro & João Editores, 2021. p. 233–252.

MEZZOMO, Carolina Lisbôa. **Aquisição dos fonemas na posição de coda medial do português brasileiro, em crianças com desenvolvimento fonológico normal**. 1999. Dissertação (Mestrado em Letras) - Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1999.

MEZZOMO, Carolina Lisbôa. **Aquisição da coda no português brasileiro: uma análise via Teoria de Princípios e Parâmetros**. 2004a. Tese (Doutorado em Letras) - Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2004.

MEZZOMO, Carolina Lisbôa. Sobre a Aquisição da Coda. *In: LAMPRECHT, Regina Ritter (org.). Aquisição fonológica do Português: perfil de desenvolvimento para terapia*. Porto Alegre: Artmed, 2004 b. p. 129–150.

MEZZOMO, Carolina Lisbôa; KESKE-SOARES, Márcia; MOTA, Helena Bolli. Análise acústica como instrumento de auxílio na descrição do sistema fonológico infantil.

**Anais do XII Congresso Brasileiro de Fonoaudiologia e II Congresso Sulbrasileiro de Fonoaudiologia**, Foz do Iguaçu, 2004.

MIRANDA, Ana Ruth. **A aquisição do “r”: uma contribuição à discussão sobre seu status fonológico**. 1996. Dissertação (Mestrado em Letras) - Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1996.

MIRANDA, Izabel Cristina Campolina; SILVA, Thais Cristófaro. Aquisição de encontros consonantais tautossilábicos: uma abordagem multirrepresentacional. **Revista Linguística**, Rio de Janeiro, v. 7, n. 1, p. 14–30, 2011.

MONARETTO, Valéria Neto de Oliveira. **Um reestudo da vibrante: análise variacionista e fonológica**. 1996. Tese (Doutorado em Letras) - Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Porto Alegre, 1996.

MONARETTO, Valéria Neto de Oliveira. Análise sociolinguística da vibrante no sul do país. **Revista Graphos**, João Pessoa, v. 1, n. 2, p. 25–34, 1997.

MONARETTO, Valéria Neto de Oliveira. O apagamento da vibrante pós-vocálica nas capitais do sul do Brasil. **Letras de Hoje**, Porto Alegre, v. 35, n. 5, p. 275–284, 2000.

NISHIDA, Gustavo. **A natureza intervocálica do tap em PB**. 2009. Dissertação (Mestrado em Letras) - Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, Curitiba, 2009.

PIKE, K.; PIKE, E. Immediate constituents of Mazateco syllables. **International Journal of Applied Linguistics**, v. 13, p. 78–91, 1947.

PROCTOR, Michael. Towards a gestural characterization of liquids: Evidence from Spanish and Russian. **Laboratory Phonology**, v. 2, n. 2, p. 451–485, 2011.

RECASENS, Daniel. What is and what is not an articulatory gesture in speech production: the case of lateral, rhotic and (alveolo)palatal consonants. **Gradus**, v. 1, n. 1, p. 24–42, 2016.

RODRIGUES, Luciana Lessa. **Aquisição dos róticos em crianças com queixa fonoaudiológica**. 2007. Dissertação (Mestrado em Linguística) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2007.

SALIDIS, Joanna; JOHNSON, Jacqueline S. The Production of Minimal Words : A Longitudinal Case Study of Phonological Development. **Language Acquisition**, v. 6, n. 1, p. 1–36, 1997.

SALTZMAN, Elliot; KELSO, J. Skilled actions: a task-dynamic approach. **Psychological Review**, v. 94, n. 1, 1987.

SELKIRK, Elisabeth. The syllable. In: VAN DER HULST, Harry; SMITH, N. (org.). **The structure of phonological representation. Part 2**. Holanda: Foris, 1982.

SILVA, Adelaide Hercília Pescatori. **Para a descrição fonético-acústica das líquidas no português brasileiro : dados de um informante paulistano**. 1996. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Campinas, 1996.

SILVA, Adelaide Hercília Pescatori. **As fronteiras entre Fonética e Fonologia e a alopofonia dos róticos iniciais em PB: dados de dois informantes do sul do país**. 2002. Tese (Doutorado em Linguística) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2002.

SILVA, Adelaide Hercília Pescatori; CLEMENTE, Felipe Costa; NISHIDA, Gustavo. Para a representação dinâmica do *tap* em grupos e cudas: evidências acústicas. **Revista Virtual de Estudos da Linguagem – ReVEL**, v. 4, n. 7, p. 1–26, 2006.

SILVA, Thaís Cristófaro *et al.* **Fonética Acústica: os sons do Português Brasileiro**. São Paulo: Contexto, 2019.

SILVEIRA, Francine; SEARA, Izabel Christine. Vogal de apoio em grupos consonantais CCV no Português Brasileiro. **Revista da ABRALIN**, v. 7, n. 1, p. 27–47, 2008.

STAMPE, David. **A dissertation on Natural Phonology**. 1973. Tese de Doutorado - Department of Linguistics, University of Chicago, Chicago, 1973.

- STERIADE, Donca. Gestures and autosegments. *In: KINGSTON, John; BECKMAN, Mary E. (org.). Papers in Laboratory Phonology*. Cambridge: Cambridge University Press, 1990. p. 382–397.
- STRAUSS, Sidney. **U-shaped Behavioral Growth**. New York: Academic Press, 1982.
- STUDDERT-KENNEDY, Michael; GOODELL, Elizabeth Whitney. Gestures, Features and Segments in Early Child Speech. *In: GELDER, Beatrice de; MORAIS, José (org.). Speech and Reading: a comparative approach*. Hove: Taylor & Francis Routledge, 1995. p. 65–88.
- THOMAS, Erik. Rhotics. *In: Sociophonetics: an introduction*. Basingstoke: Palgrave Macmillan, 2011. p. 129–135.
- VASSOLER, Aline Mara de Oliveira. **Coordenação gestual na produção de encontros consonantais em crianças com desenvolvimento típico e atípico**. 2016. Dissertação (Mestrado em Estudos Linguísticos) - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto 2016, 2016.
- WERTZNER, Haydée Fiszbein; PAPP, Ana Carolina Camargo Salvatti; GALEA, Daniela Evaristo dos Santos. Picture naming and imitation tests as tools for the diagnosis of phonological disorder. **Pró-fono: revista de atualização científica**, Barueri (SP), v. 18, n. 3, p. 303–312, 2006.
- WIESE, Richard. The phonology of /r/. *In: HALL, Tracy Alan (org.). Distinctive Feature Theory*. Berlim: Mounton de Gruyter, 2001. p. 335–368.
- YAVAS, Mehmet; HERNANDORENA, Carmen Lúcia Matzenauer; LAMPRECHT, Regina Ritter. **Avaliação fonológica da criança**. Porto Alegre: Artes Médicas, 1991.

## **Apêndices**

## Apêndice A – Documento apresentado para o responsável pela criança



UNIVERSIDADE FEDERAL DE PELOTAS  
CENTRO DE LETRAS E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS  
LABORATÓRIO EMERGÊNCIA DA LINGUAGEM ORAL



### EXPLICAÇÃO AOS PAIS E RESPONSÁVEIS SOBRE A PESQUISA

Estamos desenvolvendo uma pesquisa sobre aquisição fonológica (aquisição dos sons da língua) com crianças com idades entre 3 e 5 anos. Realizaremos uma entrevista com a finalidade de investigar em qual etapa da aquisição da linguagem a criança se encontra.

Somente realizaremos esta entrevista com as crianças cujos pais ou responsáveis autorizarem. Essa autorização será dada pela sua assinatura no **Termo de Consentimento** que será entregue junto com este documento.

Essa entrevista é composta por um instrumento de nomeação de imagens, no qual a criança visualizará figuras e será condicionada a produzir as palavras relacionadas com as imagens. Concomitantemente, será feita uma gravação em áudio, por meio de um gravador digital, para análise acústica posterior.

Ao término da entrevista, os responsáveis da criança serão informados das próximas etapas da pesquisa caso ocorra a continuidade de coletas de dados com esse informante.

As próximas coletas serão realizadas no Laboratório Emergência da Linguagem Oral, localizado na Universidade Federal de Pelotas, Campus Anglo. Esse é equipado com sistemas de aquisição de dados dinâmicos capazes de captar dados acústicos e articulatórios em tempo real e de alta definição. Um exemplo, é o emprego da tecnologia do ultrassom em trabalhos vinculados à linguística. No Brasil, o seu uso é recente e ainda é restrito a poucos laboratórios de pesquisa. As pesquisas realizadas no LELO visam a contribuir para que tecnologias atuais sejam empregadas nos estudos sobre a linguagem e o seu funcionamento.

Certas de contar com a colaboração dos senhores Pais ou Responsáveis, agradecemos antecipadamente e nos colocamos à disposição para maiores esclarecimentos.

Atenciosamente,

Patrícia Pereira Melcheque  
Mestranda em Letras

Giovana Ferreira Gonçalves  
Doutora em Linguística e Letras

## Apêndice B – Autorização para a gravação da entrevista



UNIVERSIDADE FEDERAL DE PELOTAS  
CENTRO DE LETRAS E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS  
LABORATÓRIO EMERGÊNCIA DA LINGUAGEM ORAL



### AUTORIZAÇÃO PARA A REALIZAÇÃO DA GRAVAÇÃO

Eu, \_\_\_\_\_, RG \_\_\_\_\_,  
responsável por \_\_\_\_\_,  
autorizo a realização da gravação de dados de fala da criança acima citada.

\_\_\_\_\_  
Assinatura do representante legal da criança

Pelotas, \_\_\_\_\_ de \_\_\_\_\_ de 2020.

## Apêndice C – Ficha de Cadastro



UNIVERSIDADE FEDERAL DE PELOTAS  
CENTRO DE LETRAS E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS  
LABORATÓRIO EMERGÊNCIA DA LINGUAGEM ORAL



### FICHA DE CADASTRO

<b>Dados da Criança</b>	
Nome Completo:	
Idade:	Data de Nascimento: / /
Nacionalidade:	Naturalidade:
Residente em Pelotas?	Já residiu em outra cidade?
Se sim, por quanto tempo? Qual cidade?	
Frequenta Escola?	Se sim, há quanto tempo?
Se não frequenta a escola, com quem passa o dia?	
Possui algum conhecimento de uma língua estrangeira?	
Se sim, qual?	
Tratamento fonoaudiológico:	Se sim, por qual motivo?
Língua presa (problema no freio lingual):	

<b>Dados do Responsável</b>	
Nome Completo:	
Idade:	Data de Nascimento: / /
Telefone:	e-mail:
Nacionalidade:	Naturalidade:
Residente em Pelotas?	Já residiu em outra cidade?
Se sim, por quanto tempo? Qual cidade?	
Nível de Escolaridade:	Situação:
Se você está fazendo ou já concluiu o ensino superior, qual curso?	
Possui algum conhecimento de uma língua estrangeira?	
Se sim, qual?	
Esse conhecimento é de disciplinas cursadas no ensino médio?	
Você fez curso de língua estrangeira?	Se sim, por quanto tempo?
Qual nível de conhecimento de L.E você considera ter?	
Você consegue manter uma conversa de nível básico nessa língua estrangeira?	

\_\_\_\_\_  
Assinatura do representante legal da criança

Pelotas, \_\_\_\_\_ de \_\_\_\_\_ de 2020.

**Apêndice D – Termo de Consentimento Livre e Esclarecido assinado  
pelo responsável pela criança participante**



UNIVERSIDADE FEDERAL DE PELOTAS  
CENTRO DE LETRAS E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS  
LABORATÓRIO EMERGÊNCIA DA LINGUAGEM ORAL



**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

Senhores pais ou responsáveis,

convidamos o menor de sua responsabilidade a participar da pesquisa **Emergência dos gestos articulatórios na aquisição da linguagem**, sob a responsabilidade da Prof.<sup>a</sup> Dr.<sup>a</sup> Giovana Ferreira Gonçalves e da mestrandona Patrícia Pereira Melcheque.

Por meio dessa pesquisa, pretende-se investigar o processo de aquisição e a produção de sons na fala infantil. Se você autorizar a participação da criança, estará contribuindo para o avanço dos estudos da emergência de sistemas linguísticos em processo de aquisição e desenvolvimento.

Para que a participação da criança seja possível, você deverá estar ciente das seguintes informações:

- 1) A participação consiste em uma coleta de dados, a qual ocorrerá no Laboratório Emergência da Linguagem Oral (LELO). Esse laboratório se situa nas dependências do Centro de Letras e Comunicação (CLC) da Universidade Federal de Pelotas (UFPel), campus Porto (situado à rua Gomes Carneiro, nº 1, em Pelotas/RS). A coleta consiste em gravação, em áudio, de interação guiada por figuras. O informante identificará figuras e passará por uma familiarização com palavras novas;
- 2) A participação da criança nesta pesquisa é voluntária, podendo ser interrompida a qualquer momento sem qualquer prejuízo aos envolvidos. Pelo mesmo motivo, não haverá qualquer tipo de despesa ou remuneração para o participante ou para o seu responsável.
- 3) Os dados fornecidos pelos informantes desta pesquisa serão analisados e poderão ser utilizados em eventos científicos, publicações e estudos futuros. Nesses casos, será preservada a identidade dos envolvidos – informantes e seus responsáveis.

Eu, \_\_\_\_\_, RG \_\_\_\_\_, responsável por \_\_\_\_\_, fui informado sobre as intenções desta pesquisa e entendo por que a participação da criança é necessária. Estou ciente de que não terei despesas e não serei remunerado por essa participação, bem como posso interrompê-la quando julgar necessário. Assino este documento por livre e espontânea vontade, certificando que estou de acordo com a realização desta pesquisa e autorizando a participação da criança.

---

Assinatura do representante legal da criança

Eu, Patrícia Pereira Melcheque, RG 3098448834, confirmo que entreguei este documento ao responsável nele nomeado. Confirmo, também, que esclareci eventuais dúvidas acerca da pesquisa e do conteúdo deste documento ao responsável nomeado.

---

Patrícia Pereira Melcheque  
Mestranda em Letras

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Giovana Ferreira Gonçalves  
Orientadora

---

Pelotas, \_\_\_\_\_ de \_\_\_\_\_ de 2020.

---

## **Anexos**

**Anexo A – Imagens apresentadas na seleção de informantes**



### Anexo B – Imagens apresentadas na coleta acústica



## Anexo C – Script extract and save intervals

```

# extract_and_save_intervals
#
# This script extracts and saves intervals in an independent sound file, the name of files will be
# the label the intervals have.
#
# Wendy Elvira-GarcÃ-a (2015). Extract_and_save_intervals, v.1 [Praat script]
# wendyelviragarcia@ g m a i l . c o m
# Laboratori de FonÃ-tica. Universitat de Barcelona.
# This script is free software under GNU Public License v.3 http://www.gnu.org/copyleft/gpl.html
#####
#####
# INSTRUCTIONS
#
# 0. Before you start: You need a sound and its matching TextGrid with the intervals you want to
# extract.
#           SELECT THE SOUND YOU WANT TO ANALISE
# 1. Open the script (Open/Read from file...), click Run in the upper menu and Run again.
# 2. Tell the script if you have any label whose intervals you don't want to extract. For example,
#    don't extract empty intervals, or don't extract intervals where there is a "x" written or a "no".
#    The files will be saved with the same name the interval has, but you can choose a prefix and
# a suffix that will apply to all files.
#    For exemple, the code of the speaker as a prefix.
# 3. A folder selector will open select, where do you want to save the new files.
#
#
#####
#####
form Extract_and_save_intervals
    comment Select the Sound you want to extract the intervals from
    comment
    comment Do you want to extract all interval?
    comment If you don't write here the label of the intervals you don't want to extract.
    word do_not_extract x
    comment If you do not want to extract empty intervals write ""
    word prefix wl07
    word suffix 1
endform

marca_de_silencio$ = do_not_extract$
informante$ = prefix$
repeticion$ = suffix$
folder$ = chooseDirectory$ ("Choose the folder where the sounds will be saved:")
#falta que lo haga del sonido seleccionado
@extract_and_save_intervals: folder$


procedure extract_and_save_intervals: .directoryOutput$
    .directoryOutput$ = chooseDirectory$: "Choose a directory to save all the new files in"
    if .directoryOutput$ <> ""
        .numberOfSelectedSounds = numberOfSelected ("Sound")
        if .numberOfSelectedSounds = 0
            pause Select the Sound you want to extract from and click OK
        endif
    endif

```

```

.base$ = selected$ ("Sound")
selectObject: "TextGrid " + .base$

#compruebo que el grid sea de ese Sound
.soundDur = Get total duration
.gridDur = Get total duration
if .soundDur<>.gridDur
    El sonido y el TextGrid tienen duraciones diferentes. ¿Seguro que son del
mismo archivo?
endif

#me apetecia que el margen de los archivos fuera silencio puro, pero para limpiarlos
no irÃ¡ bien...
#selectObject: "Sound " + .base$
# sampl_frequ = Get sampling frequency
# silence = Create Sound from formula: "silence", 1, 0, 0.2, sampl_frequ, "0"
# silence2 = Create Sound from formula: "silence", 1, 0, 0.2, sampl_frequ, "0"
.numberOflntervals = Get number of intervals: 1
for .interval from 1 to .numberOflntervals

    selectObject: "TextGrid " + .base$
    .codigo$ = Get label of interval: 1, .interval
    if .codigo$ <> marca_de_silencio$
        echo Extrayendo intervalo '.interval' '.codigo$'
            .int_start = Get start point: 1, .interval
            .int_end = Get end point: 1, .interval
            selectObject: "Sound " +.base$
            #extraigo los intervalos con un margen porque sino van muy justos,
            en el caso de que sea el primer sonido o el Ãºltimo no hay posibilidad de silencio
            if .interval = 1
                Extract part: .int_start, .int_end+0.2, "rectangular", 1, 0
            elif .interval = .numberOflntervals
                Extract part: .int_start-0.2, .int_end, "rectangular", 1, 0
            else
                Extract part: .int_start-0.2, .int_end+0.2, "rectangular", 1, 0
            endif
            Save as WAV file: .directoryOutput$ + "/" + informante$ + .codigo$ +
            repeticion$ + ".wav"
            Remove
        endif
    endfor
endif
endproc

```

## Anexo D – *Script calculate segment durations*

```

# This script will calculate the durations of all labeled segments in a TextGrid object.
# The results will be save in a text file, each line containing the label text and the
# duration of the corresponding segment..
# A TextGrid object needs to be selected in the Object list.
#
# This script is distributed under the GNU General Public License.
# Copyright 12.3.2002 Mietta Lennes

# ask the user for the tier number
form Calculate durations of labeled segments
    comment Which tier of the TextGrid object would you like to analyse?
    integer Tier 1
    comment Where do you want to save the results?
    text textfile durations.txt
endform

# check how many intervals there are in the selected tier:
numberOfIntervals = Get number of intervals... tier

# loop through all the intervals
for interval from 1 to numberOfIntervals
    label$ = Get label of interval... tier interval
    # if the interval has some text as a label, then calculate the duration.
    if label$ <> ""
        start = Get starting point... tier interval
        end = Get end point... tier interval
        duration = end - start
        # append the label and the duration to the end of the text file, separated with a tab:

        resultline$ = "label$"      'duration"newline$"
        fileappend "textfile$"" 'resultline$'
    endif
endfor

```

### Anexo E – *Script vowel plot*

```

# Ricardo Bion
# January 2006

form PLOT VOWELS
integer f1 250
integer f2 1100
optionmenu vowel_symbol: 1
    option i
    option y
    option u
    option \ic
    option \hs
    option e
    option \o
    option \ef
    option a
    option \as
    option \ab
    option \sw
    option \vt
    option \ct
optionmenu Color_of_the_vowel: 5
    option Black
    option Blue
    option Grey
    option Yellow
    option Red
    option Navy
    option Purple
    option Lime
    option Sylver
    option Green
    option Maroon
boolean erase_vowels: 0
choice size: 2
    button big
    button small
endform

if erase_vowels
Erase all
endif

max_F2 = 3000
min_F2 = 500
max_F1 = 1500
min_F1 = 100
label$ = vowel_symbol$

Select outer viewport... 0 5 0 4

if size = 1
Select outer viewport... 0 10 0 8
endif

Black
Line width... 1

```

Plain line

Font size... 12

if size = 1

Font size... 18

endif

Axes... max\_F2 min\_F2 max\_F1 min\_F1

Marks bottom... 5 yes yes no

Marks left... 5 yes yes no

Draw inner box

Text left... yes F1 (Hz)

Text bottom... yes F2 (Hz)

'color\_of\_the\_vowel\$'

if size = 1

Text special... 'f2:0' Centre 'f1:0' Half Times 24 0 'label\$'

else

Text special... 'f2:0' Centre 'f1:0' Half Times 16 0 'label\$'

endif

Plain line

Line width... 1