

“SE VOCÊ ACHA QUE CAPIXABA NÃO TEM SOTAQUE, FALA ARROZ AÍ RAPIDINHO”:

A DITONGAÇÃO NA VARIEDADE CAPIXABA¹

“IF YOU THINK CAPIXABAS DON’T HAVE AN ACCENT, SAY ‘ARROZ’ REAL QUICK”:

DIPHTHONGIZATION IN THE CAPIXABA ACCENT

Ana Carolina Carnieli Duarte²

Raquel Gomes Chaves³

RESUMO

Este trabalho investigou o processo de ditongação com glide /j/ em sílabas tônicas terminadas em /S/ em dados de fala da região Metropolitana da Grande Vitória (RMGV), a partir da perspectiva teórico-metodológica da Teoria da Variação e Mudança Linguística (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972], 1994) e com o suporte da fonética acústica. Como *corpus* de análise, utilizamos a Amostra Duarte (2025), composta por sete informantes estratificados por *sexo* (homens e mulheres), *faixa etária* (Faixa A: 22 a 35 anos e Faixa C: acima de 60 anos) e *nível de escolarização* (ensino médio completo e ensino superior completo). As variáveis controladas foram: *vogal-núcleo do segmento-alvo*, *posição do acento*, *estilo*, *sexo*, *faixa etária* e *escolarização* (variáveis de efeito fixo) e *item lexical* e *informante* (variáveis de efeito aleatório). Os resultados indicaram que o fenômeno apresentou taxa de aplicação de 35,5%, índice próximo ao registrado pela literatura para os dados de fala carioca. As variáveis linguísticas e extralinguísticas estatisticamente significativas foram *faixa etária*, a ditongação é mais frequente na fala dos informantes mais jovens; *posição do acento*, a ditongação é prevalente em monossílabos tônicos e oxítonas; *vogal núcleo do segmento-alvo*, a vogal média-alta anterior colabora com o fenômeno em oposição à vogal média-baixa posterior; e *estilo*, a leitura desfavorece a incidência de ditongação. Os dados sugerem, portanto, que a ditongação na RMGV é um processo influenciado pelo monitoramento estilístico, pela posição do acento na palavra, pela qualidade da vogal do item lexical e pela faixa etária dos falantes.

PALAVRAS-CHAVE: Ditongação. Variedade capixaba. Variação fonética-fonológica.

ABSTRACT

This study investigated the process of diphthongization with the glide /j/ in stressed syllables ending in /S/ in speech data from the Metropolitan Region of Vitória (RMGV), based on the theoretical-methodological framework of Variationist Theory (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972], 1994) and supported by acoustic phonetics. The *corpus* consists of the Duarte Sample (2025), comprising seven informants stratified by *sex* (men and women), *age group* (Group A: 22 to 35 years old; Group C: over 60 years old), and *education level* (high school and university graduates). The controlled variables included:

¹ Na etapa de redação deste manuscrito, utilizamos a ferramenta ChatGPT (OpenAI, 2026, versão GPT-4) com o objetivo de aprimorar a fluidez do texto. Após o uso da ferramenta, revisamos criticamente todo o conteúdo, assumindo integral responsabilidade pelo manuscrito publicado, incluindo eventuais imprecisões ou incorreções.

² Universidade Federal do Espírito Santo (UFES), Edital Piic 2024/2025 (Bolsa FAP), carnieli.carol@gmail.com, <https://orcid.org/0009-0003-1835-9505>.

³ Universidade Federal do Espírito Santo (UFES), raquel.chaves@ufes.br, <https://orcid.org/0000-0002-6310-0194>.

target segment nucleus vowel, stress position, style, sex, age group, and education (fixed effect variables), as well as *lexical item* and *informant* (random effect variables). Results indicated a 35.5% application rate, close to the rates reported in the literature for the Rio de Janeiro dialect. Statistically significant variables included: *age group*, with diphthongization being more frequent among younger speakers; *stress position*, prevalent in stressed monosyllables and oxytone words; *nucleus vowel*, where the high-mid front vowel favors the phenomenon compared to the low-mid back vowel; and *style*, as reading discourages diphthongization. The data suggest that diphthongization in the RMGV is a process influenced by stylistic monitoring, word stress position, the quality of the lexical item's vowel, and the speakers' age group.

KEYWORDS: Diphthongization. Capixaba variety. Phonetic-phonological variation.

1. Introdução

A discussão mais recorrente acerca da variedade capixaba gira em torno de quais seriam os traços linguísticos indexadores do falar espírito-santense. A dificuldade em determinar esses aspectos resulta na concepção geral de que o capixaba não teria sotaque (Tesch, 2022; Lacerda; Leppaus, 2023; Yacovenco, 2012). Segundo Yacovenco *et al.* (2012), o isolamento histórico do Espírito Santo e os fluxos migratórios posteriores contribuíram para uma variedade linguística pouco marcada, cujos traços passam despercebidos, inclusive pelos próprios falantes.

Nesse cenário, a ditongação variável em sílabas tônicas travadas por /S/ configura-se como um dos poucos traços salientes na identificação da variedade em questão. Tesch (2022), ao aplicar um questionário de percepção a 1400 respondentes, relativo às marcas identificadas como específicas do sotaque capixaba, apontou que alguns traços fonéticos/fonológicos foram mais recorrentes nas respostas, a saber: a realização do /S/ capixaba, identificado como diferente do /S/ carioca, e o fenômeno da ditongação (*arroz ~ arroiz*). Para os habitantes locais, a impressão de que a ditongação seria um traço identitário do capixaba também pode ser ilustrada por comentários, como os encontrados na rede social *X* (antigo *Twitter*): “Capixaba é uma raça engraçada. Acrescentam o i em todas as palavras: *arroiz, treis*; aí, quando a palavra tem *i*, ignoram e falam *pexe, feção*.” e “Se você acha que capixaba não tem sotaque, fala *arroz* aí rapidinho. Você fala *arroiz*.”.

Partindo dessa discussão, neste trabalho, investigamos a variação da ditongação em sílabas tônicas travadas por /S/ (*três ~ treis, mas ~ mais, Jesus ~ Jesuis*) em dados de espírito-santenses que residem na Região Metropolitana da Grande Vitória (RMGV). Nos pautamos na perspectiva teórico-metodológica da Teoria da Variação e Mudança Linguística (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972], 1994), com o suporte da Fonética Experimental.

O processo variável referido, do ponto de vista fonológico, é caracterizado pela inserção de uma semivogal na mesma sílaba que o monotongo, criando um ditongo decrescente, isto é, uma sequência de dois segmentos vocálicos consecutivos (na qual o primeiro é uma vogal e o segundo, uma semivogal) ocupando uma mesma sílaba (Bisol, 1999). Do ponto de vista fonético, os ditongos decrescentes são abordados como eventos articulatórios gradientes que envolvem uma configuração vocálica estável, a qual se altera para uma segunda configuração mais fechada. Acusticamente, essa trajetória se manifesta nas transições formânticas, principalmente do primeiro e segundo formantes (Barbosa; Madureira, 2015).

As perguntas de pesquisa que guiam esta análise são as seguintes: (i) qual a frequência da ditongação variável em sílabas tônicas travadas por /S/ na Grande Vitória?; (ii) há evidências acústicas que corroboram a análise de oitiva na identificação das formas ditongadas *versus* monotongadas?; (iii) quais fatores linguísticos e extralinguísticos se correlacionam com a variante ditongada?

Adicionalmente, com base nos estudos revisados sobre o fenômeno na Região Sudeste (Silva, 2014; Aragão, 2014; Silveira, 2023; Callou, 2009), formulamos as seguintes hipóteses: (i) o índice de ditongação é alto na comunidade investigada e difundido entre as categorias sociais, com base na identificação desse traço pelos falantes (marca dialetal da comunidade como um todo); (ii) há diferenças significativas entre ditongos e monotongos, levando-se em conta que ditongos apresentariam transições de F1 e F2; (iii) os fatores favorecedores do fenômeno são: as vogais anteriores e a central, em comparação às posteriores; os monossílabos tônicos e oxítonas em relação às paroxítonas e proparoxítonas; o estilo de fala registra maior ocorrência do fenômeno do que o estilo leitura; os informantes mais jovens ditongam mais, assim como os menos escolarizados; mulheres tendem a ditongar mais do que homens.

No contexto capixaba não há registro de uma descrição abrangente e detalhada da atuação do fenômeno de ditongação, à exceção do trabalho de Silva (2014), que analisa a ditongação nas capitais brasileiras. Com vistas a preencher essa lacuna, este trabalho tem como objetivo geral investigar o processo de ditongação com glide /j/ em sílabas tônicas na variedade da Grande Vitória. Os objetivos específicos são: (i) compreender a influência dos fatores linguísticos e sociais nesse aspecto da fala; (ii) contribuir para a descrição fonético-fonológica da fala capixaba; e (iii) enquadrar essa discussão no quadro mais amplo da Teoria da Variação e Mudança Linguística (Weinreich; Labov; Herzog, 1968; Labov, 2008 [1972], 1994).

O presente trabalho organiza-se da seguinte forma: na seção 2, apresentamos a revisão de literatura acerca do fenômeno investigado, com foco em trabalhos sobre a Região Sudeste; na seção 3, descrevemos a metodologia empregada, com ênfase nos procedimentos de constituição da Amostra Duarte (2025), no tratamento dos dados, bem como na inspeção acústica dos segmentos analisados e na delimitação do envelope de variação; na seção 4, apresentamos e discutimos os resultados obtidos, contemplando tanto a análise quantitativa descritiva quanto a análise estatística inferencial, com base em modelos de regressão de efeitos mistos; e, por fim, na seção 5, tecemos as considerações finais, nas quais retomamos os principais achados do estudo e apontamos suas contribuições para a descrição sociolinguística da variedade capixaba.

2. Estudos sobre o fenômeno variável da ditongação na Região Sudeste

A ditongação por inserção de glide /j/ em sílabas tônicas terminadas em /S/ é um fenômeno recorrente em diversas variedades do PB, com ocorrências registradas em diferentes regiões do país. Silva (2014, p. 144), ao mapear o comportamento do fenômeno em 25 capitais brasileiras, por meio dos questionários lexicais que constituem o *corpus* do Projeto Atlas Linguístico do Brasil (ALiB),

conclui que a distribuição geográfica constitui fator importante para a aplicação do fenômeno, sendo o país dividido entre áreas fortemente favorecedoras da ditongação, como o Nordeste, e áreas de manutenção do monotongo, como o Sul.

Especificamente sobre a Região Sudeste, de particular interesse para a presente pesquisa, Silva (2014, p. 245) identifica um comportamento homogêneo nas frequências de ditongação entre as capitais: Rio de Janeiro apresentou índices de 19%, Vitória, frequência de 15%, Belo Horizonte, 12% e São Paulo registrou 13%. As taxas inferiores a 20%, no entanto, podem ser justificadas pela seleção de itens lexicais constantes no protocolo de coleta do Projeto ALIB, uma vez que Silva (2014) analisa as respostas ao Questionário Fonético-Fonológico (QFF), o qual é composto por 159 questões previamente elaboradas e organizadas e que incentiva respostas específicas dos participantes (com perguntas como: “a carne do porco não é magra porque tem...?”), ao que o informante deveria responder “colesterol”).⁴ O questionário, porém, foi desenvolvido objetivando o estudo de diversos fenômenos, e não especificamente a ditongação diante de /S/ e, por isso, os contextos são, muitas vezes, pouco favorecedores. Além disso, a autora não analisou apenas a ditongação em sílabas tônicas, como foi feito no presente trabalho e em demais pesquisas sobre o fenômeno, considerando todos os contextos em que o /S/ ocupasse a posição de coda silábica. É importante destacar ainda que a análise da distribuição da ditongação em faixas etárias revelou indícios de mudança em curso, com favorecimento para a realização entre os mais jovens.

Outro estudo sobre a ditongação nas capitais brasileiras é o de Aragão (2014), que utiliza do mesmo *corpus* de Silva (2014). A autora também analisa itens lexicais isolados do QFF para testar se os processos da ditongação e da monotongação constituem variantes diatópicas, diastráticas ou puramente fonéticas (compreendendo a influência das variáveis linguísticas controladas, sendo elas: *tipo da vogal, contexto posterior, tonicidade, extensão da palavra e registro*). Os resultados globais apontaram para um alto percentual de ditongação no Brasil, com 72% de aplicação. No que tange à distribuição geográfica, incluindo as capitais da Região Sudeste, Aragão (2014, p. 12) nega a hipótese de que esses processos funcionem como marcadores regionais. A autora argumenta que o fenômeno ocorre em todo o território nacional, não sendo possível isolar o sudeste ou qualquer outra região como área exclusiva ou preferencial de ocorrência com base nos itens analisados. No entanto, apesar dos dados encontrados no estudo apontarem para esse resultado (já que a ditongação acontece em diversas regiões do país), se faz necessário compreender se o fenômeno ocorre da mesma maneira em todas as áreas do Brasil, identificando a influência de outros aspectos dos sotaques (como a realização do /S/), as trajetórias formânticas dos dados e a avaliação da população sobre esse aspecto de fala.

⁴ As palavras analisadas por Silva (2014) a partir das respostas dos informantes ao QFF foram as seguintes: *alastrou, aliás, arroz, as, asfalto, aspas, atrás, através, vasculhante, bastante, besta, bosque, busca, capataz, capaz, casca, casas, caspa, chaminés, churrasqueira, colegas, colesterol, coroas, costas, das, desvio, dez, dias, duas, elas, encosta, escola, esquerdo, fósforo, frutas, gasta, gosto, inglês, japonês, juntas, labaredas, lâmpadas, mas, mazela, mesma, muitas, nas, nos, os, palafitazinhas, pescoço, pesquisa, placas, português, prateleiras, produz, quiosques, rapaz, rasgando, rasgar, respostas, rodas, ruas, rústicas, talvez, três, vocês.*

Os estudos que analisaram o comportamento variável da ditongação na fala carioca (Callou, 2009; Silveira, 2023) apontam essa variedade como de alta manifestação do fenômeno. No que se refere à frequência de aplicação da variante ditongada, Callou (2009, p. 47), ao analisar comparativamente dados do Projeto da Norma Urbana Linguística Culta (NURC) coletados em dois períodos (1970 e 1990), identifica taxas de ditongação entre 41% e 65% (na década de 1970) e de 19% a 62% (na década de 1990), a depender do contexto linguístico.

Silveira (2023, p. 30), por sua vez, ao desenvolver uma análise em tempo real comparando dados do Programa de Estudos sobre o Uso da Língua (PEUL) e contrapondo os resultados das décadas de 1980 e 2000, registra frequência de ditongação de 57% em monossílabos na década de 1980 e de 72% na de 2000. Ambos os estudos apontam para um condicionamento linguístico semelhante: a ditongação é verificada em maior escala quando a vogal núcleo é baixa ou média anterior, co-ocorre com a variante palatal em coda silábica e é fortemente condicionada pela tonicidade da sílaba, com os monossílabos apresentando frequência significativamente superior.

Quanto às variáveis extralinguísticas, embora o fenômeno seja analisado exclusivamente em dados de fala culta (o que indica que a ditongação não se trata de um marcador de baixa escolarização), observam-se padrões de variação relacionados à faixa etária, com distribuição curvilínea que sugere um processo de mudança linguística em curso, e ao sexo, com leve preferência pela ditongação entre falantes do sexo feminino.

3. Metodologia

Para a realização deste estudo, optamos pela constituição de uma amostra de dados de fala coletados segundo os moldes labovianos (Amostra Duarte, 2025). A referida amostra é composta por sete entrevistas sociolinguísticas realizadas com informantes residentes da Região Metropolitana da Grande Vitória⁵, que, segundo dados do IBGE (2022), possui aproximadamente 1.880.000 habitantes distribuídos entre as sete cidades: Cariacica, Fundão, Guarapari, Serra, Viana, Vila Velha e a capital Vitória. Os sete informantes que fazem parte deste estudo residem em Cariacica, Serra, Vila Velha e Vitória. Na imagem 1, apresentamos um mapa da RMGV.

⁵ A amostra encontra-se, ainda, em fase de constituição, sendo as sete entrevistas analisadas as disponíveis para estudo até o presente momento.

Imagem 1: Mapa da Região Metropolitana da Grande Vitória

Fonte: Instituto de Defesa Agropecuária e Florestal do Espírito Santo (IDAF-ES).

Disponível em: <https://idaf.es.gov.br/mapas-do-espírito-santo>. Acesso em: 05 jan. 2026

Cabe destacar que não delimitamos como pré-requisito que os pais dos participantes da pesquisa fossem, também, nascidos na localidade, como se observa na constituição de outras amostras sociolinguísticas. No entanto, analisamos os dados de fala de cada um dos participantes, observando se atendiam aos traços fonéticos capixabas, nesse caso, considerando o /S/ alveolar e o /R/ aspirado como variantes características da comunidade em estudo⁶ – o que foi aferido a partir da observação pessoal da entrevistadora no contato prévio com os participantes. Os informantes foram estratificados por *sexo* (homens e mulheres), *faixa etária* (*Faixa A*, de 22 a 35 anos, e *Faixa C*, acima de 60 anos) e *nível de escolarização* (ensino médio completo e ensino superior completo). A estratificação da amostra encontra-se disposta no quadro 1.

⁶ Schmidt (2024) e Ruela (2025) realizaram pesquisas de iniciação científica analisando o /S/ e o /R/ em posição de coda final na variedade capixaba, respectivamente. Schmidt (2024, 2025) identificou a variante fricativa alveolar como praticamente categórica para a realização do /S/ em coda e Ruela (2024, 2025) constatou a variante fricativa glotal para a realização do /R/ em posição final.

Quadro 1: Estratificação social e dados gerais dos informantes da Amostra Duarte (2025)

Informante	Sexo	Idade	Local de nascimento	Cidade que reside atualmente	Local de nascimento dos pais
Informante 1	Masculino	23 anos	Vargem Alta (ES)	Vitória (ES)	Vargem Alta (ES)
Informante 2	Masculino	27 anos	Vitória (ES)	Vila Velha (ES)	Mãe: Vitória (ES); Pai: Vila Velha (ES)
Informante 3	Masculino	71 anos	Vitória (ES)	Vitória (ES)	Baixo Guandu (ES)
Informante 4	Masculino	67 anos	Cariacica (ES)	Cariacica (ES)	Cariacica (ES)
Informante 5	Feminino	30 anos	Vitória (ES)	Vitória (ES)	Mãe: Vitória (ES); Pai: Rio de Janeiro (RJ)
Informante 6	Feminino	60 anos	São José do Calçado (ES)	Serra (ES)	São José do Calçado (ES)
Informante 7	Feminino	22 anos	Vitória (ES)	Vitória (ES)	Vitória (ES)

Fonte: Elaboração das autoras.

A coleta foi realizada em três momentos. Primeiramente, realizamos uma entrevista de experiência pessoal nos moldes variacionistas. Após a realização das entrevistas, procedemos à leitura de uma lista de itens lexicais, preparada utilizando itens previamente selecionados (sendo palavras paroxítonas e oxítonas com a sílaba tônica terminada em /S/, propiciando a realização da ditongação, além de itens distratores). Por fim, gravamos a leitura de cinco textos que incluíam os itens lexicais passíveis de ditongação⁷. Com vistas a garantir a qualidade da gravação, utilizamos o gravador digital ZOOM H4N e o microfone unidirecional AKG C520L. Os áudios foram gravados em taxa de amostragem de 44100Hz.

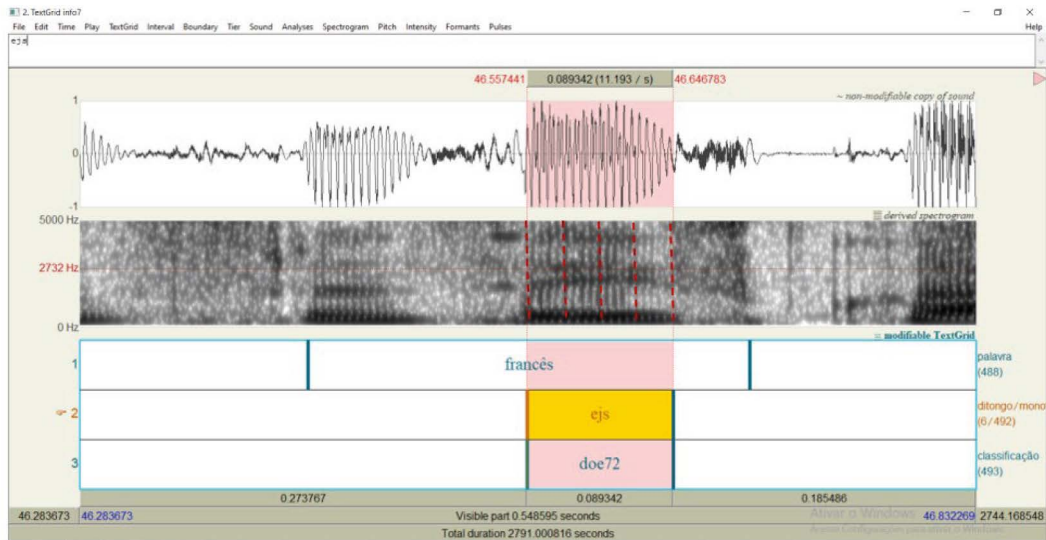
Posteriormente à coleta, realizamos as seguintes etapas: (i) transcrição, etiquetagem manual e classificação dos dados; (ii) levantamento das ocorrências que apresentam sílabas tônicas travadas por /S/; (iii) análise da trajetória formântica dos dados (valores de F1 e F2) no *software* Praat (versão 6.4.47); (iv) organização dos dados em planilhas de forma a possibilitar o estudo da influência das variáveis preditoras de efeito fixo – *posição do acento, estilo, vogal, sexo, nível de escolarização, faixa etária* – e das variáveis de efeito aleatório – *item lexical e informante*; (v) submissão dos dados à análise estatística na Plataforma R (RStudio); (vi) análise e interpretação dos resultados.

Para o exame das trajetórias formânticas das ocorrências, extraímos as informações dos primeiro e segundo formantes (F1 e F2)⁸ de cinco pontos equidistantes no intervalo correspondente à etiquetagem do segmento-alvo, visando delimitar a trajetória dos valores. Esse procedimento pode ser verificado na imagem 2.

⁷ Projeto aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal do Espírito Santo (CEPE/UFES), sendo 86924924.6.0000.5542 o número do processo.

⁸ F1 e F2 são medidas em Hz que indicam, respectivamente, o grau de elevação da língua no momento de produção do dado e do segundo formante e o grau de anteriorização da língua durante a realização do segmento-alvo.

Imagem 2: Representação da coleta dos dados das trajetórias formânticas dos segmentos-alvo analisados em 5 pontos equidistantes da etiquetagem (linhas tracejadas em vermelho)



Fonte: Elaboração das autoras.

Com base na análise de oitiva atrelada à inspeção acústica dos dados, buscamos delimitar um critério que guiasse a categorização da variável-resposta em *forma ditongada versus forma não ditongada*.

No processo de definição das variáveis preditoras linguísticas e extralinguísticas, bem como os níveis que compõem cada uma delas, tomamos como base os resultados de estudos acerca do fenômeno da ditongação realizados em outras localidades do Brasil (*cf.* seção 2). O quadro 2 exibe as variáveis consideradas.

Quadro 2: Envelope de variação

Variável resposta	
ditongação	formas não ditongadas (<i>mas, três, cuscuz</i>) formas ditongadas (<i>mais, três, cuscuiz</i>)
Variáveis preditoras de efeito fixo	
vogal-núcleo do segmento-alvo	/a/ (<i>mas ~ mais</i>) /e/ (<i>mês ~ mês</i>) /ɛ/ (<i>dez ~ deiz</i>) /o/ (<i>arroz ~ arroiz</i>) /ɔ/ (<i>voz ~ voiz</i>) /u/ (<i>luz ~ luiz</i>)
posição do acento na palavra	monossílabo tônico (<i>três ~ três</i>) oxítone (<i>capuz ~ capuiz</i>) paroxítone (<i>mesmo ~ mesmo</i>) proparoxítone (<i>ginástica ~ ginaística</i>)

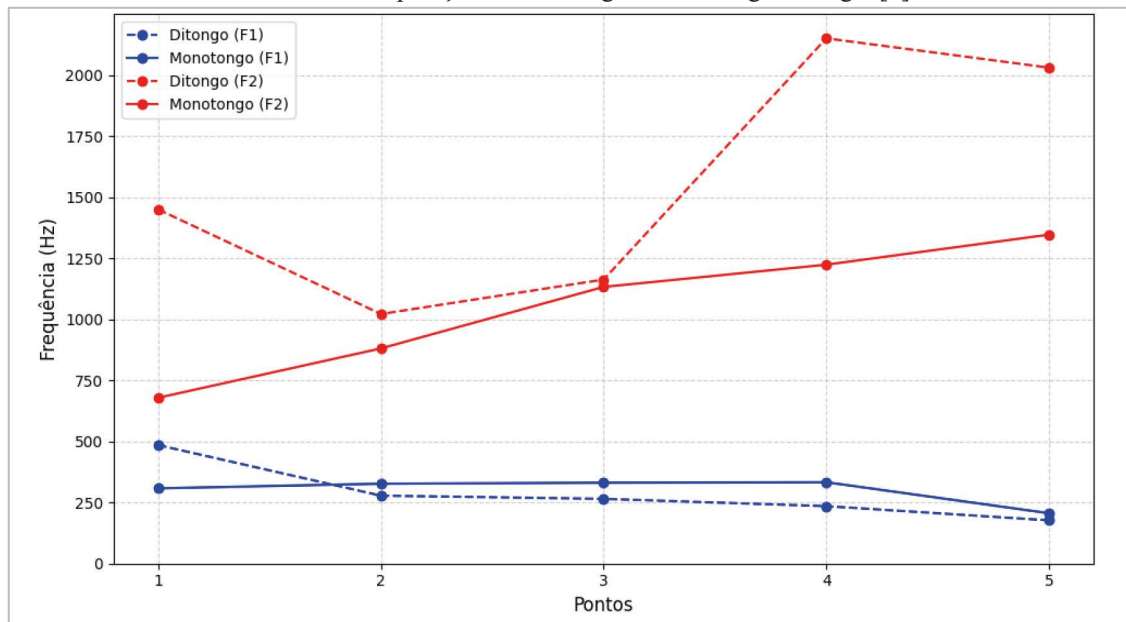
estilo	fala leitura
sexo	feminino masculino
faixa etária	faixa A (de 22 a 35 anos) faixa C (Mais de 60 anos)
escolarização	ensino médio ensino superior
Variáveis preditoras – efeito aleatório	
item lexical	
informante	

Fonte: Elaboração das autoras.

Na próxima seção, apresentamos os resultados, contemplando tanto a análise descritiva, com a exposição das distribuições gerais e das tendências observadas, quanto a análise inferencial, na qual se discutem os efeitos estatísticos das variáveis linguísticas e extralinguísticas sobre a ocorrência da ditongação.

4. Resultados e discussão

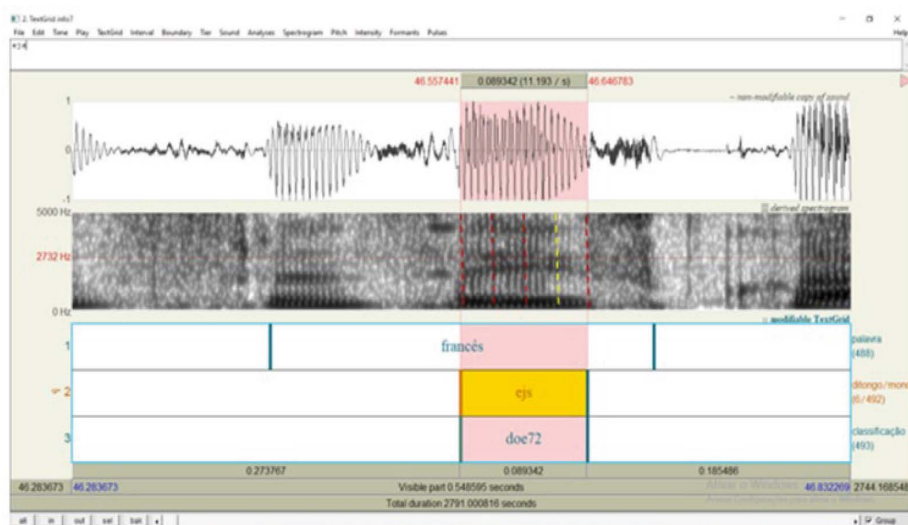
O primeiro passo da análise quantitativa da ditongação consistiu na definição de um critério acústico adequado para distinguir, de forma alinhada à percepção auditiva, vogais ditongadas de monotongos. A análise preliminar dos correlatos acústicos indicou que o segundo formante (F2), associado à anterioridade/posterioridade da língua, mostrou-se um parâmetro eficiente na identificação da presença/ausência do glide /j/ nos dados em exame. Em contraste, o primeiro formante (F1), relacionado à altura vocálica, permaneceu praticamente inalterado, independentemente da ocorrência de ditongação. O gráfico 1 ilustra essa diferença para a vogal /o/. Observamos que a trajetória de F1 se mantém estável tanto no monotongo [o] quanto no ditongo [oj]. Já a trajetória de F2 apresenta um aumento acentuado no final da produção do ditongo, refletindo o deslocamento da língua em direção à posição anterior, característico do glide /j/. Esse aumento ocorre no momento final da produção vocálica, imediatamente antes da transição para a sibilante.

Gráfico 1: Comparação entre ditongo e monotongo da vogal [o]

Fonte: Elaboração das autoras.

A partir dessa constatação, o valor de F2 foi definido como o parâmetro para a classificação das ocorrências em monotongos ou ditongos. Para garantir a precisão desse parâmetro, foi necessário determinar um ponto específico na trajetória formântica que capturasse a articulação do glide amenizando a interferência dos efeitos de coarticulatórios do contexto seguinte. Como os segmentos-alvo da pesquisa são seguidos pela sibilante /S/ (realizada categoricamente como alveolar em nossa amostra), o ponto final da vogal (último tracejado) já apresenta influências da produção dessa consoante. Por essa razão, optamos por extrair o valor de F2 no penúltimo ponto de medição (linha tracejada amarela), conforme ilustrado na imagem 3.

Imagem 3: Representação da coleta dos dados das trajetórias formânticas dos segmentos-alvo analisados com enfoque no penúltimo ponto (linha tracejada amarela)



Fonte: Elaboração das autoras.

Na análise quantitativa dos dados, o fenômeno da ditongação em sílaba tônica ocorreu em 35,5% do *corpus* analisado, correspondendo a 428 de um total de 1.171 ocorrências. O índice geral de ditongação observado na amostra (35,5%) mostra-se superior ao valor de 15% reportado por Silva (2014) para Vitória – que analisa a ditongação em contextos tônicos e átonos, diferentemente do que realizamos neste estudo e nos aqui analisados sobre a fala carioca, que levam em consideração a realização da ditongação diante de /S/ em sílabas tônicas -, e próximo ao das taxas registradas para a variedade carioca, superior a 40% (Callou, 2009; Silveira, 2023).

O percentual geral é inferior à nossa expectativa inicial, de que o valor seria superior a 50%, hipótese conjecturada com base no fato de a ditongação ser frequentemente apontada como um traço característico da variedade capixaba. Acreditamos que, embora a ditongação em sílabas tônicas travadas por /S/ seja frequentemente apontada, tanto na literatura quanto no senso comum, como um marcador saliente do falar espírito-santense, os dados empíricos revelam um quadro mais complexo.

Em uma análise inicial, o valor que encontramos poderia ser interpretado como contraditório. No entanto, acreditamos que essa aparente discrepância (entre os valores esperados e os valores observados) seria reflexo da saliência sociolinguística do fenômeno. Em outras palavras, acreditamos que a ditongação mesmo não apresentando frequência elevada seria muito perceptível. Um exemplo clássico de fenômeno saliente no português é a não marcação da concordância verbal de terceira pessoa do plural (*eles come* vs. *eles comem*), que, conforme indica a literatura (Scherre; Naro, 1997; Monguilhott, 2009; Chaves, 2017; Naro, 1981; Vieira, 1997), nas comunidades urbanas, é verificada em cerca de 20 a 30% dos dados. Apesar da relativa baixa frequência, o fenômeno é altamente perceptível e estigmatizado (caracterizando-se, nos termos labovianos, como um estereótipo negativo).

Além disso, conforme Freitag (2018), há indícios de que os significados sociais da variação fonética (no caso a ditongação) seriam avaliados pelos ouvintes com maior facilidade do que variáveis de nível gramatical (no caso da concordância verbal). Essa percepção seria mediada pelo que Labov *et al.* (2006) denominaram *monitor sociolinguístico*, mecanismo cognitivo que atuaria no processamento de informações socialmente marcadas. Em estudo de percepção conduzido sobre a variação na pronúncia de –ing (*working* ~ *workin*), Labov *et al.* demonstraram que os sujeitos são sensíveis a oscilações de frequência a partir de 10%.

Dessa forma, defendemos que a taxa de 35,5% é suficiente para que o fenômeno da ditongação tenha se cristalizado como um índice identitário da variedade capixaba ainda que não seja, em termos absolutos, um traço linguístico altamente frequente. Diferentemente de variáveis estigmatizadas que operam “abaixo da consciência” (indicadores), a ditongação na RMGV parece comportar-se como um marcador ou mesmo como um estereótipo, situando-se acima do nível de consciência social. Assim, acreditamos que a percepção de que a ditongação ocorre com maior frequência do que os dados quantitativos sugerem pode ser atribuída à sua elevada saliência. Estudos nesse sentido, como o de Tesch (2022), dão suporte a essa afirmação.

A seguir, na tabela 1, apresentamos os resultados referentes à distribuição dos dados para cada variável social e linguística considerada.

Tabela 1: Valores brutos e percentuais referentes a cada variável linguística e extralinguística controlada (análise descritiva)

variável		n (ditongação)	total	%
sexo	feminino	275	604	45,53
	masculino	153	567	26,98
faixa etária	A (mais jovens)	285	635	44,88
	C (mais velhos)	143	536	26,68
escolarização	EM	190	529	35,92
	ES	238	642	37,07
posição do acento	monossílabo	297	447	66,44
	oxítona	122	250	48,8
	paroxítona	8	467	1,71
	proparoxítona	1	7	14,29
vogal-núcleo	ε	3	28	17,3
	ɔ	29	216	8,8
	e	123	309	39,7
	o	6	38	15,8
	u	7	53	13,2
	a	245	393	62,3
estilo	fala	277	490	56,53
	leitura	151	681	22,17
		428	1.171	35,5%

Fonte: Elaboração das autoras.

A tabela 1 apresenta uma distribuição da ditongação em relação às variáveis sociais e linguísticas controladas em nosso *corpus*. No que tange aos fatores sociais, destacam-se as variáveis *sexo* e *faixa etária*. As informantes do sexo feminino apresentaram taxa de ditongação de 45,5%, consideravelmente superior aos 27% registrados para os sujeitos do sexo masculino. Em relação à *faixa etária*, a *faixa A* (mais jovens) ditongou em 44,9% dos casos, quase o dobro da *faixa C* (mais velhos), a qual apresentou uma taxa de 26,7%. Para a variável *escolarização*, por sua vez, registramos valores aproximados, 35,9% (Ensino Médio) e 37,1% (Ensino Superior).

No que toca aos fatores linguísticos, a variável *posição do acento* apresentou prevalência do fenômeno em monossílabos tônicos (66,4%), como em *mês ~ mêis*, e oxítonas (48,8%), como em *você ~ vocêis*, e aplicação reduzida em palavras paroxítonas (1,7%), como em *mesmo ~ meismo* e proparoxítonas (14,9%), como em *ginástica ~ gináistica*. A variável *vogal-núcleo do segmento-alvo* também se mostrou associada a diferentes taxas de ocorrência: o fenômeno foi mais recorrente nos dados com a vogal baixa (62,3%), como em *paz ~ paiz*, e com a vogal média-alta (39,8%), como em *você ~ vocêis*. Para as demais vogais, as proporções de ditongação observadas foram inferiores, situando-se entre 8,8% e 17,3%. Em reação à variável estilística, a *fala espontânea* apresentou mais ocorrência do fenômeno (56,5%) em detrimento do estilo de *leitura* (22,2%).

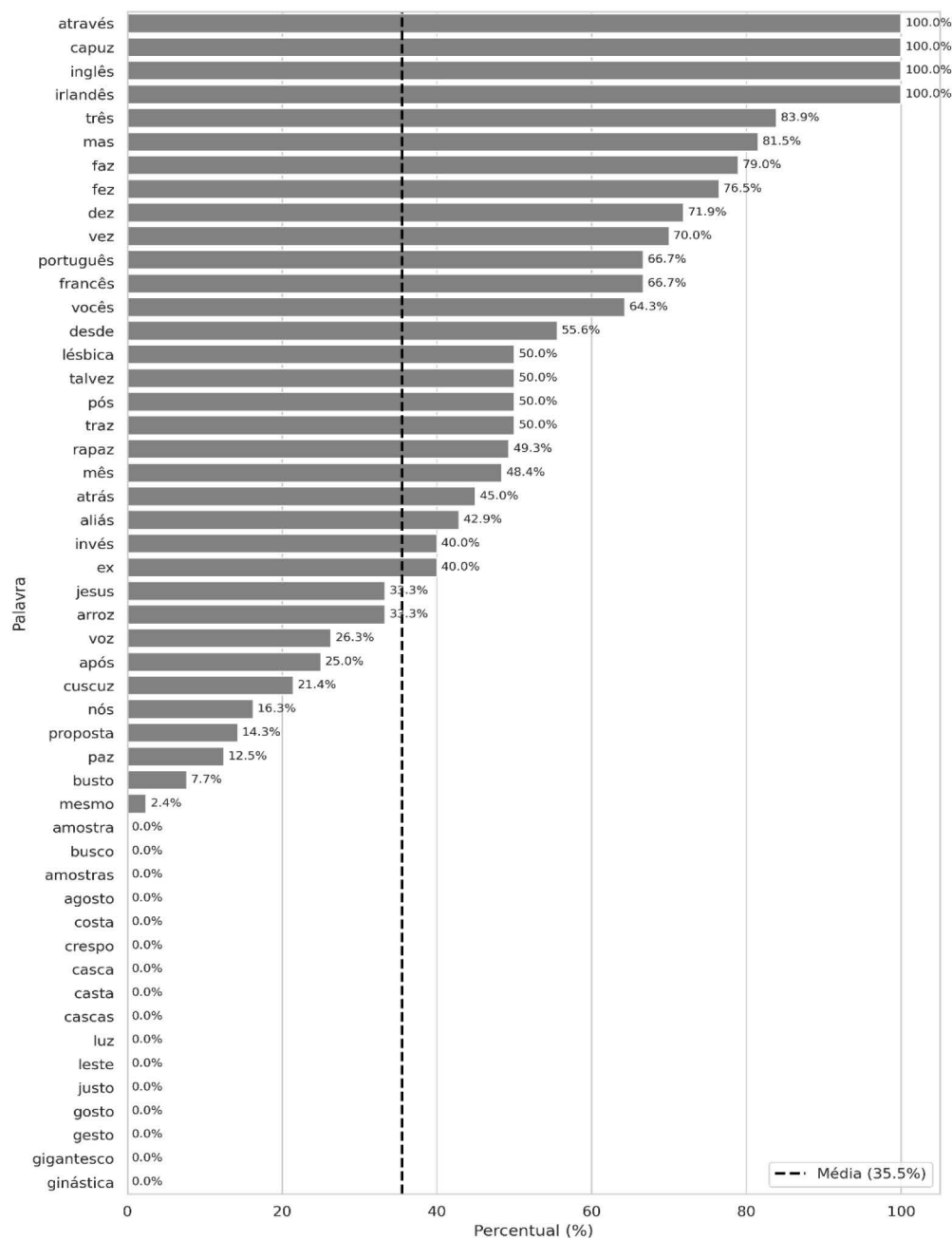
Para além das tendências gerais, a análise da variabilidade individual, exibida na tabela 2, revela a heterogeneidade do fenômeno entre os falantes.

Tabela 2: Valores brutos e percentuais do fenômeno de ditongação por informante

Categoria	n (Ditongação)	Total	Percentual (%)
INFORMANTE 1 (23 anos, homem, Vitória)	54	129	41,86
INFORMANTE 2 (27 anos, homem, Vila Velha)	34	115	29,57
INFORMANTE 3 (71 anos, homem, Vitória)	33	154	21,43
INFORMANTE 4 (67 anos, homem, Cariacica)	32	169	18,93
INFORMANTE 5 (30 anos, mulher, Vitória)	94	145	64,83
INFORMANTE 6 (60 anos, mulher, Serra)	78	213	36,62
INFORMANTE 7 (22 anos, mulher, Vitória)	103	246	41,87

Fonte: Elaboração das autoras.

A *Informante 5*, jovem de 30 anos natural de Vitória, destaca-se como a participante que mais ditongou, com uma taxa de 64,8%. Cabe, no entanto, destacar que o pai da participante é carioca, o que pode ter impacto no índice verificado, já que no Rio de Janeiro a ditongação apresenta taxas próximas a essa, de acima de 40%. Seguindo padrão semelhante, o *Informante 1* (23 anos, Vargem Alta, residente em Vitória) e a *Informante 7* (22 anos, natural de Vitória) também apresentaram taxas elevadas, aproximadamente 42%. Em contrapartida, os informantes mais velhos, como o *Informante 3* (71 anos, Vitória) e o *Informante 4* (67 anos, Cariacica), exibiram um comportamento mais conservador, com as taxas inferiores de ditongação, 21,4% e 18,9%, respectivamente. Esse contraste entre falantes mais jovens e mais velhos reforça os padrões observados na tabela 1 e sugere que a ditongação pode estar passando por um processo de mudança liderado por mulheres jovens na variedade capixaba.

Gráfico 2: Proporção de ocorrência de ditongação por item lexical (50 das 73 palavras registradas em nosso *corpus*)

Fonte: Elaboração das autoras.

Tabela 3: Valores brutos e percentuais da ditongação por item

palavra	ditongação/ total de dados	% de ditongação
através	2/2	100
capuz	1/1	100
inglês	12/12	100
irlandês	1/1	100
três	26/31	83,87
mas	172/211	81,52
faz	15/19	78,95
fez	13/17	76,47
dez	23/32	71,88
vez	14/20	70
português	8/12	66,67
francês	2/3	66,67
vocês	18/28	64,29
desde	5/9	55,56
lésbica	1/2	50
talvez	3/6	50
pós	1/2	50
traz	2/4	50
rapaz	36/73	49,32
mês	15/31	48,39
atrás	9/20	45
aliás	9/21	42,86
invés	2/5	40
ex	4/10	40
Jesus	3/9	33,33
arroz	6/18	33,33
voz	5/19	26,32
após	6/24	25
cuscuz	3/14	21,43
nós	7/43	16,28
proposta	1/7	14,29
paz	1/8	12,5
busto	1/13	7,69
mesmo	1/41	2,44
amostra	0/5	0

Fonte: Elaboração das autoras.

Conforme podemos observar no gráfico 2 e tabela 3, a ditongação atua com maior ou menor frequência a depender do item lexical. Foi registrada aplicação categórica de 100% do fenômeno, por exemplo, na palavra *inglês*, a qual contabilizou 12 ocorrências, e de 0% em palavras como *texto*, computada 42 vezes. As palavras que apresentaram taxas acima da média encontrada de ditongação e que foram produzidas ao menos cinco vezes em nosso *corpus* foram: *inglês* (100%), *três* (83,9%),

mas (81,5%), *faz* (78,9%), *fez* (76,5%), *dez* (71,9%), *vez* (70%), *português* (66,7%), *vocês* (64,3%), *desde* (55,6%), *talvez* (50%), *rapaz* (49,3%), *mês* (48,4%), *atrás* (45%), *aliás* (42,9%), *invés* (40%). As palavras com maiores índices de ditongação são monossílabas tônicas, oxítonas e com duas sílabas (exceto *português*, com três sílabas). Uma única paroxítona figura entre as palavras com altos percentuais de ditongação, o item *desde* (*deisde*). Interpretamos que, nesse caso, os falantes poderiam estar analisando o item como *des de*. Nessa reanálise, o primeiro elemento estaria sendo interpretado como um monossílaboônico e o segundo elemento, como uma preposição/clítico. Assim, esse não seria um exemplar paroxítono, mas sim, de um monossílaboônico (*des ~ deis*). Por outro lado, as palavras em que não se registrou ocorrência do fenômeno em estudo foram majoritariamente paroxítonas. Dentre os itens desse grupo acentual, verificamos ditongação apenas em *proposta* (14,3%), *busto* (7,7%) e *mesmo* (4%).

Tal constatação indica a necessidade de uma investigação mais aprofundada, especialmente no que se refere à possível interação entre forças neogramáticas *versus* difusionistas (Labov, 1994; Phillips, 2006). Nesse sentido, acreditamos que o léxico, atrelado à saliência, possa desempenhar papel relevante na percepção do fenômeno como um traço reconhecível como típico do falar capixaba.

Para investigar os fatores linguísticos e sociais que se correlacionam ao fenômeno variável da ditongação, utilizamos um modelo de regressão logística de efeitos mistos. O modelo final incluiu as variáveis *faixa etária*, *estilo*, *posição do acento* e *vogal núcleo do segmento-alvo* (variáveis de efeito fixo), além das variáveis para *informante* e *palavra* (variáveis de efeito aleatório). Os resultados deste modelo encontram-se dispostos na tabela 4.

Tabela 4: Resultados da análise de regressão de efeitos mistos

Variável (efeito fixo)	Estimativa	Erro-padrão	z-valor	p-valor
Intercepto: Faixa A - mais velhos, proparoxítona, vogal média-alta posterior, estilo:fala	-3,4163	0,73920	-4,622	< 0,001***
FAIXA ETÁRIA: Faixa A - mais jovens	1,8223	0,35788	5,092	< 0,001***
POSIÇÃO DO ACENTO: monossílabo	3,4814	0,66845	5,208	< 0,001***
POSIÇÃO DO ACENTO: oxítona	3,5651	0,61490	5,798	< 0,001***
POSIÇÃO DO ACENTO: paroxítona	-1,4178	0,46290	-3,063	0,002**
VOGAL: média-baixa anterior	0,9903	0,52389	1,890	0,059
VOGAL: média-baixa posterior	-0,9286	0,40216	-2,309	0,021*
VOGAL: baixa	0,6397	0,46722	1,369	0,171
VOGAL: média-alta anterior	0,9586	0,34515	2,777	0,005**
VOGAL: alta posterior	-0,6321	0,70428	-0,897	0,369
ESTILO: leitura	-1,9202	0,43523	-4,412	< 0,001***

Modelo (DITONCAÇÃO ~ FAIXA.ETARIA + POSIÇÃO DO ACENTO + VOGAL + ESTILO+(1|ITEM>LEXICAL) + (1|INFORMANTE), data = dados, family = binomial)

Fonte: Elaboração das autoras.

Os resultados do modelo, apresentados na tabela 4, indicam efeitos significativos das variáveis *faixa etária*, *posição do acento*, *vogal núcleo do segmento-alvo* e *estilo*. Observamos forte efeito da faixa etária, com a *Faixa A* (mais jovens) apresentando coeficiente positivo e estatisticamente significativo ($\beta = 1,8223$; $p < 0,001$), o que indica aumento na probabilidade de ocorrência da ditongação em comparação à faixa de referência (*Faixa C* – mais velhos).

A distribuição da ditongação nas faixas etárias investigadas sinaliza uma mudança em curso. Essa afirmação é feita com cautela, já que estamos diante de um estudo em tempo aparente. A maior probabilidade de aplicação da variante ditongada entre falantes mais jovens, aliada ao comportamento mais conservador dos falantes mais velhos, reproduz um padrão clássico de mudança linguística (Labov, 2008 [1972]; 1994). Esse resultado é particularmente relevante à luz da literatura sobre o sudeste, uma vez que Silva (2014) já havia identificado indícios de mudança nas capitais.

Entre os fatores linguísticos, o efeito mais expressivo foi o da *posição do acento*. As categorias *monossílabo* ($\beta = 3,4814$; $p < 0,001$) e *oxítona* ($\beta = 3,5651$; $p < 0,001$) elevaram significativamente a chance de ocorrência da ditongação em relação ao nível de referência (proparoxítonas). Esses resultados indicam que a posição do acento exerce papel central na realização da variante analisada.

No que se refere à *vogal núcleo do segmento-alvo*, apresentaram efeito significativos a vogal média-alta anterior ($\beta = 0,9586$; $p < 0,05$) e a vogal média-baixa posterior ($\beta = -0,9286$; $p < 0,05$). A vogal média-alta anterior (*três ~ treis*, *você ~ vocêis*) mostrou aumento da probabilidade de ditongação, ao passo que a vogal média baixa-posterior (*após ~ apósis*, *nós ~ nósis*) diminui a probabilidade de ocorrência do fenômeno ambas comparadas à vogal central. A tendência observada na análise descritiva de que a vogal baixa /a/ e as vogais médias anteriores (tanto a média-alta /e/ quanto a média-baixa /ɛ/) colaboram com a ocorrência do fenômeno, em consonância com os resultados reportados para a variedade carioca por Callou (2009) e Silveira (2023), foi confirmada para a vogal /e/. Tal favorecimento da vogal média-alta anterior pode ser explicado por fatores articulatórios e acústicos: vogais anteriores apresentam maior propensão à emergência do glide /j/ devido à maior amplitude do movimento articulatório da língua e à maior saliência das transições formânticas. Em contraste, vogais posteriores e mais fechadas impõem maiores restrições articulatórias a esse movimento, reduzindo a probabilidade de ditongação.

Por fim, o *estilo* também se mostrou relevante: o *estilo* leitura reduziu significativamente a probabilidade de ocorrência do fenômeno ($\beta = -1,9202$; $p < 0,001$). A redução da ditongação no estilo de leitura indica uma tentativa dos informantes de reprodução fiel da escrita. Esse comportamento estilístico aproxima a ditongação do estatuto de marcador sociolinguístico, definido por Labov (2008 [1972]) como uma variável sensível à atenção do falante.

5. Considerações finais

Este estudo teve como objetivo descrever o comportamento variável da ditongação com glide /j/ em sílabas tônicas travadas por /S/ na fala da RMGV, sob a perspectiva da Teoria da Variação e

Mudança Linguística. A análise de 1.171 ocorrências revelou um índice geral de 35,5% de ditongação, valor superior ao previamente registrado para Vitória. Esse resultado indica que a ditongação constitui um traço perceptível e socialmente reconhecido da fala capixaba, embora não se configure como um traço com altas taxas de aplicação.

No que se refere às perguntas de pesquisa formuladas na introdução, os resultados permitem responder que: (i) o índice de ditongação na Grande Vitória é intermediário no panorama do sudeste, sendo superior ao valor anteriormente registrado para Vitória (ii) as trajetórias formânticas, em especial os valores de F2, corroboram de forma consistente à análise de oitiva, validando a distinção entre monotongos e ditongos a partir de critérios acústicos objetivos; e (iii) os principais fatores correlacionados à variante ditongada foram a *posição do acento*, a *faixa etária*, o *estilo de fala* e a *vogal núcleo do segmento-alvo*.

Por fim, este trabalho contribui com a descrição de um aspecto fonético-fonológico do sotaque capixaba. Os resultados reforçam que a fala da Grande Vitória constitui dados que podem ser proveitosos para a discussão sobre saliência estrutural, especialmente em uma variedade em que poucos traços são apontados como marcas dialetais. Nesse sentido, a análise vai ao encontro a premissa de Labov (2006) de que traços identitários podem apresentar frequências de aplicação inferiores a 50%, por vezes em torno de 10%, e ainda assim serem socialmente salientes. Portanto, a ditongação na RMGV é um elemento relevante para o estudo do *monitor sociolinguístico*, sugerindo que pesquisas futuras focadas na percepção da própria população em relação a este traço da fala possam expandir as discussões aqui apresentadas, mapeando como esses índices de frequência se convertem em avaliações subjetivas e estigmas ou prestígios na comunidade de fala.

Referências

- ARAGÃO, Maria do Socorro Silva de. Ditongação e Monotongação nas Capitais Brasileiras. In: CONGRESSO INTERNACIONAL DA ASSOCIAÇÃO DE LINGÜÍSTICA E FILOGIA DA AMÉRICA LATINA, 17., 2014, João Pessoa. *Anais* [...]. João Pessoa: Ideia, 2014. pp. 2089-2101.
- BARBOSA, Plínio; MADUREIRA, Sandra. *Manual de fonética acústica experimental: aplicações a dados do português*. São Paulo: Cortez, 2015.
- BISOL, Leda. O ditongo na perspectiva da fonologia atual. D.E.L.T.A., Rio Grande do Sul, v. 5, n. 2, pp. 185-224, 1989.
- CALLOU, Dinah. Variação e mudança no âmbito do consonantismo. In: PAIVA, Maria da Conceição de; D'ANGELIS, Wilmar da Rocha (org.). *Estudos de Sociolinguística: variações linguísticas no português brasileiro*. Rio de Janeiro: Editora UFRJ, 2009.
- CHAVES, R. G. *A redução/desnasalização de ditongos nasais átonos finais e a marcação explícita de CVP6: um estudo de correlação*. 2017. 359 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2017.

DUARTE, Ana Carolina Carnieli. *A ditongação em sílabas tônicas travadas por /S/ no dialeto capixaba*. Vitória, 2025. Relatório Técnico-Científico.

FREITAG, Raquel Meister Ko. Saliência estrutural, distribucional e sociocognitiva. *Acta Scientiarum. Language and Culture*, Maringá, v. 40, n. 2, e41173, pp. 1-11, 2018. Disponível em: <https://edueojs.uem.br/ojs/index.php/ActaSciLangCult/article/view/41173>. Acesso em: 13 jun. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Vitória (ES): Cidades e Estados*. Rio de Janeiro: IBGE, 2022. Disponível em: <https://www.ibge.gov.br/cidades-e-estados/es/vitoria.html>. Acesso em: 20 abr. 2026.

LABOV, William. *Principles of linguistic change: linguistic factors*. Oxford: Blackwell, 1994.

LABOV, William *et al.* Listeners' sensitivity to the frequency of sociolinguistic variables. In: FRIESNER, M. L.; RAVINDRANATH, M. (eds.). *Penn Working Papers in Linguistics: selected papers from NWAV 34*. Philadelphia: University of Pennsylvania; Penn Linguistics Club, 2006. v. 12, n. 2, pp. 105-129.

LABOV, William. *Padrões Sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].

LACERDA, Marcela Langa; LEPPAUS, Ana Clara S. Capixaba tem sotaque? Quem pode dar essa resposta? Questões para a sociolinguística da sociedade. *Confluência*, n. 65, pp. 73-94, 2023.

OPENAI. *ChatGPT* [software]. Versão GPT-4. [S.l.]: OpenAI, 2026. Disponível em: <https://chat.openai.com>. Acesso em: 31 jul. 2026.

PHILLIPS, Betty. Word frequency and the actuation of sound change. *Language*, v. 60, n. 2, pp. 320-342, 1984.

MONGUILHOTT, Isabel de Oliveira e Silva. *Estudo sincrônico e diacrônico da concordância verbal de terceira pessoa do plural no PB e no PE*. 2009. 229 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2009.

NARO, Anthony Julius. The social and structural dimensions of a syntactic change. *Language*, Baltimore, v. 57, n. 1, pp. 63-98, 1981.

RUELA, Emilene. */R/ em coda medial na fala capixaba*. Vitória, 2025. Relatório Técnico-Científico.

SCHERRE, Maria Marta Pereira; NARO, Anthony Julius. A concordância de número no português do Brasil: um caso típico de variação inerente. In: HORA, Dermeval da (org.). *Diversidade linguística no Brasil*. João Pessoa: Ideia, 1997. pp. 93-114.

SCHMIDT, Sophia. *A variação de /S/ em coda: análise longitudinal da fala de Viviane Mosé e Elisa Lucinda*. Vitória, 2024. Relatório Técnico-Científico.

SILVA, Amanda dos Reis. *A ditongação em sílabas fechadas por /s/ no corpus do Projeto ALIB: nas trilhas das capitais brasileiras*. 2014. Dissertação (Mestrado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2014.

SILVEIRA, Anna Carolina Lopes da. A trajetória da ditongação de monossílabos na variedade carioca: um estudo em tempo real. 2023. Trabalho de Conclusão de curso. Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2023.

TESCH, Leila. O sotaque capixaba: um estudo de percepção. *(Con)Textos Linguísticos*, Vitória, v. 16, n. 34, pp. 225-242, 2022.

VIEIRA, S. A não-concordância em dialetos populares: uma regra variável. *Graphos*, v. 2, n. 1, pp. 115-133, 1997.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006 [1968].

YACOVENCO, Lilian Coutinho *et al.* Projeto PortVix: a fala de Vitória/ES em cena. *Alfa: Revista de Linguística*, São José do Rio Preto, v. 56, n. 2, pp. 771-792, 2012.