



Estudo de vogais suarabácticas na fala de paranaenses e paulistas sob uma perspectiva estatística e sociodialetológica

Study of suarabacti vowels in the speech of Paraná and São Paulo under a statistical and socio-dialectology perspective

Valter Pereira ROMANO
(Universidade Federal de Lavras)

Rodrigo Duarte SEABRA
(Universidade Federal de Itajubá)

RESUMO

*O estudo analisa o comportamento da epêntese vocálica em encontros consonantais na fala de brasileiros naturais e residentes nos estados de São Paulo e Paraná. Utiliza-se como corpus os dados coletados pela equipe do Projeto ALiB em 55 municípios sumarizando 220 informantes. Selecionaram-se três questões do Questionário Fonético-Fonológico: 54, 72 e 101, cujo caput são os vocábulos *afiosa*, *pneu* e *advogado*, respectivamente. O estudo considera aspectos teóricos e metodológicos da Dialectologia e da Sociolinguística para analisar o fenômeno fonético-fonológico compreendido sob uma perspectiva pancrônica. Por meio de tratamentos estatísticos são confirmadas hipóteses sobre a influência de variáveis linguísticas e extralinguísticas na realização do fenômeno.*

Palavras-chave: *epêntese vocálica; encontro consonantal; Projeto ALiB; estados de São Paulo e Paraná.*

ABSTRACT

*The study analyzes the behavior of vowel epenthesis in consonant clusters in the speech of natural and resident Brazilians in the states of São Paulo and Paraná. We used as corpus the data collected by the ALiB project team in 55 municipalities totaling 220 informants. Three questions from the Phonetic-Phonological Questionnaire were selected: 54, 72 and 101, the caput of which are the words *aftosa*, *pneu* and *advogado*, respectively. The study considers theoretical and methodological aspects of Dialectology and Sociolinguistics to analyze the phonetic-phonological phenomenon understood under a panchronic perspective. The hypotheses about the influence of linguistic and extralinguistic variables in the occurrence of phenomenon in question are confirmed by statistical treatments.*

Key-words: *vowel epenthesis; consonant cluster; ALiB Project; states of São Paulo and Paraná.*

Introdução

Os metaplasmos são modificações fonéticas que sofrem as palavras ao longo do tempo, podendo, segundo Coutinho (1976), ser classificados em quatro categorias: (i) metaplasmos por permuta; (ii) por aumento; (iii) por subtração e (iv) por transposição.

Dentre os metaplasmos por aumento (ii), encontra-se o processo fonético-fonológico conhecido como epêntese, “acréscimo de fonema no interior de palavras, exs.: pign(o)ra > pendra (arc.) (>prenda), um(e)ru > ombro, area (<arena) > areia [...]” (Coutinho 1976: 146), cuja peculiaridade se manifesta também na “intercalação de uma vogal para desfazer um grupo de consoantes” (Carvalho; Nascimento 1984: 36), como se observa na passagem do latim para o português em: planu > prão > porão; blatta > brata > barata; grupa (kruppa-germânico) > garupa. Segundo Coutinho (1976: 147), a este tipo de epêntese especial dá-se o nome de anaptixe ou suarabácti, encontrados já em vocábulos do português arcaico como *carônica* (arc.) por *crônica*, cuja forma atual não evidencia mais o fenômeno.

Este artigo discute o comportamento da vogal suarabáctica em três vocábulos do português brasileiro: *aftosa*, *pneu* e *advogado*.

Mais especificamente, o estudo se atém à qualidade da vogal com base no *corpus* coletado pela equipe do Projeto Atlas Linguístico do Brasil em 17 cidades paranaenses e 38 municípios paulistas junto a indivíduos naturais das localidades de no máximo nível fundamental de escolaridade.

O estudo utiliza como método de análise e descrição ferramentas computacionais que visam a facilitar o tratamento dos dados, são elas: o [SGVCLin] – *Software para geração e visualização de cartas linguísticas* (Romano; Seabra; Oliveira 2014), o programa SPSS - *Statistical Package for the Social Sciences* (SPSS 2008) e o GoldVarb 2001 (Robinson; Lawrence; Tagliamonte 2001).

Uma breve reflexão sobre os dados permite afirmar que há um comportamento diferenciado quanto ao uso da vogal epentética, que, mesmo sendo investigada por estudiosos como Lee (1993), Massini-Cagliari (2000) e, sobretudo, Collischonn (1996, 2000, 2002, 2003, 2004), carece de abordagens que detalhem a qualidade desta vogal. Nesse sentido, este trabalho discute os resultados não somente a partir da estatística descritiva, mas também, da inferencial por meio de testes de 24 hipóteses, seguindo os procedimentos teórico-metodológicos da Dialetoлогия e da Sociolinguística.

Nas próximas seções, apresenta-se a revisão da literatura acerca da epêntese no português, seguida da exposição dos materiais e métodos empreendidos no estudo. Na seção 3, discorre-se sobre a descrição e análise do *corpus* a partir de três perspectivas: (i) estatística descritiva: resultados dos relatórios obtidos a partir do [SGVCLin] que subsidiou a elaboração de hipóteses; (ii) estatística inferencial, cujo conteúdo versa sobre os resultados obtidos para o tratamento das hipóteses formuladas; e (iii) tratamento dispensado pelo GoldVarb 2001 na relação de variáveis linguísticas e extralinguísticas aplicadas à regra do abaixamento da vogal suarabática [*i*] > [*e*]. Seguem, ao final, algumas conclusões e as referências bibliográficas.

1. A epêntese no português

Ao tratar as estruturas das sílabas do português, Camara Jr. (1977: 57) reconhece a intercalação de uma vogal nos encontros consonantais

de vocábulos como *compacto*, *apto*, *ritmo*, *afta*, insistindo que a epêntese, nestes contextos, não pode ser fonemicamente desprezada, “apesar da tendência a reduzir a sua emissão no registro formal da língua culta. Ela é /i/ na área do Rio de Janeiro e /e/ ([a] neutro em Portugal)”.

Para Cagliari (1981: 107), algumas palavras no português brasileiro variam foneticamente, podendo ter uma sílaba a mais ou a menos dependendo da ocorrência de uma vogal breve e átona, em geral [i], “entre uma oclusiva, uma nasal bilabial e uma fricativa alveolar por um lado, e outra consoante por outro lado”. O autor ilustra as ocorrências com exemplos e apresenta os segmentos consonantais em que pode ocorrer a inserção da vogal (Tabela 1).

Cagliari (1981: 109) afirma que a vogal “pode se realizar com uma qualidade mais baixa e mais central [ə] sempre que ocorrer uma oclusiva velar precedendo-a sendo seguida por uma oclusiva alveolar surda ou por uma nasal alveodental”, como em: [fa-ki-tu-aw] ~ [fa-kə-tu-aw]; [a-ki-ni] ~ [a-kə-ni], concluindo que a formação do tipo de encontros consonantais descritos mostra a variabilidade das formas fonéticas de muitas palavras refletidas nas pronúncias dos falantes. “Foneticamente, o caso é simples; ou ocorre a vogal ou não ocorre” (Cagliari 1981: 109).

Tabela 1 – Segmentos consonantais que podem apresentar vogal epentética [i]

consoante		Consoante	exemplos
b	+	p, t, d, k, m, n, s, z, x, ʒ, v, l	subproduto, obter, súbdito, subconsciente, submarino, abnegado, absoleto, obséquio, sub-reptício, objeto, óbvio, sub-locação
p	+	t, s	captou, psicose
d	+	d, m, v, ʒ	admirar, advogado, adjetivo
t	+	m	ritmo
k	+	t, s, n	compacto, fixe, técnica
g	+	g, m, n	pigmeu, ignorância
m	+	m, n	amnésia
f	+	f, t	afta

Fonte: Cagliari (1981: 108).

Um estudo clássico sobre a epêntese no português brasileiro é o trabalho de Lee (1993), no qual o estudioso aborda o fenômeno sob a perspectiva da teoria da sílaba e da Fonologia Lexical. O autor considera que a epêntese no português é sempre realizada fonologicamente como /e/ e que, dependendo do contexto, pode ocorrer o alçamento para /i/. Para discussão dos dados e defesa de sua tese, o autor apresenta uma série de exemplos sobre a estrutura silábica no português, e, no que se refere à inserção de vogal epentética, em vocábulos que possuem *onset* /pn/, afirma que a vogal /e/ é inserida entre duas consoantes a fim de evitar a colisão na hierarquia da sonoridade, de modo a obter a eufonia.

Como observamos, /e/ sempre é inserida, embora, na forma ortográfica, não seja realizada: em alguns casos, evitando as estruturas silábicas não-possíveis; no português falado, em caso de eufonia, parece que há uma tendência a se guardar a forma CV, para evitar dificuldades na pronúncia. (Lee 1993: 849)

Outro trabalho sobre a epêntese vocálica é o de Collischonn (1996) no qual a autora discute, também à luz da teoria da sílaba, três casos do fenômeno em português, em meio de palavra com consoantes heterossilábicas (rpto), depois de consoante final (VARIG) e diante de grupo consonantal (spa). A autora não se detém na qualidade desta vogal, se alta, média-alta ou baixa central, restringindo-se a discutir a aplicabilidade da teoria nos padrões silábicos encontrados. Novos estudos sobre a epêntese vocálica foram realizados por Collischonn (2000; 2002; 2003; 2004), sob a perspectiva da sociolinguística variacionista, com base nos dados do Projeto VARSUL, considerando também teorias fonológicas. Entretanto, conforme a autora,

[...] consideramos como variável dependente a presença ou não de um elemento epentético vocálico; isto é, consideramos aplicação quando houve epêntese e não-aplicação quando não houve. Não consideramos como variável a qualidade da vogal realizada. Ela pode realizar-se como vogal reduzida, especialmente nas posições postônicas, ou como semivogal (a PU[kja]visa). Além destas, há também a realização como schwa (a[kə]ne). (Collischonn 2004: 64)

Massini-Cagliari (2000) também fez estudo sobre epêntese vocálica, porém com vistas a comparar e distinguir a epêntese e a paragoge¹ no português arcaico com base em cantigas trovadorescas. Em seu estudo, conclui que os dois processos se diferenciam, uma vez que a epêntese destina-se a ‘corrigir’ estruturas silábicas no início ou no meio do vocábulo, enquanto a paragoge altera estrutura que já apresenta boa formação em posição final. Ainda sobre a epêntese vocálica, com o foco no estudo da paragoge, encontram-se os trabalhos de Massini-Cagliari (2006), Migliorini (2010) e Migliorini; Massini-Cagliari (2011).

Ou seja, embora Camara Jr (1977), Cagliari (1981) e Lee (1993) já apontassem a existência de uma vogal anaptítica variável quanto à sua qualidade (alta, média ou central), os trabalhos mais recentes que visam a descrever o fenômeno com base em *corpora* não se detêm na qualidade dessa vogal. Todavia, no cenário nacional, trabalhos gerais sobre o português, bem como os primeiros trabalhos de cunho dialetológico, que remontam o início do século XX, documentam a presença da vogal e suas diferentes qualidades.

Teyssier (1994: 82), ao tratar das características do português, considera o fenômeno na categoria dos aspectos inovadores da fonética brasileira, afirmando que “os grupos consonantais que ocorrem em certas palavras de origem erudita (ex.: admirar, advogado, observar, psicologia, ritmo) são eliminados pelo aparecimento de um *i*, mais raramente de um *e*: adimirar, adivogado ou adevogado, obisservar, pissicologia, rítimo”. Já Boléo (1943), em sua clássica obra, *Brasileirismos: problemas de método*, afirma que

a eliminação de alguns grupos consonânticos por meio de um *e* anaptítico também não é facto peculiar ao Brasil. É possível (não posso asseverar) que a introdução de um *i* para desfazer os grupos seja desconhecida em Portugal: é o caso de palavras como *abissolutamente*; mas já o mesmo se não verifica com respeito ao *e* anaptítico. [...] Não é raro ouvirem-se na linguagem familiar e, sobretudo, na popular, pronúncias como *felor*, *belusa*, *guelória*, *adevogada nossa*. A única diferença está em que na fala brasileira este *e* é mais fechado e mais demorado, podendo passar a *i*: Quelemência (Clemência), quilaridade (claridade). (Boleo 1943: 24)

1. “Paragoge ou epítese é a adição de fonema no fim do vocábulo exs.: ante > antes” (Coutinho 1976: 147).

Em contraponto a esse posicionamento, mas em consonância às ideias de Teyssier (1994), Cunha (1986) discute, além de outros aspectos, traços de inovações do português brasileiro observados na fonética e na morfossintaxe. Dentre os inúmeros fatos apresentados, encontra-se a “intercalação de um *i* ou de um *e* para desfazer certos grupos consonânticos: advogado por advogado, peneu por pneu” (Cunha 1986: 209), embora o autor reconheça que “o fenômeno da anaptixe ocorre em alguns falares e em certos tipos de elocução do português europeu, mas tudo faz crer que seja uma solução paralela à brasileira” (Cunha 1986: 223).

Noll (2008), em estudo contrastivo do português brasileiro face ao português europeu afirma que “a tendência ampliada, na língua coloquial brasileira, de desfazer *clusters* consonantais que se encontram em limite silábico, por meio de uma vogal epentética, é indício para o desenvolvimento linguístico determinado, em primeiro plano, pelo uso oral” (Noll 2008: 228). Assim, a epêntese vocálica corresponde aos princípios da hierarquia sonora e da lei de contato entre sílabas uma vez que “certos encontros consonantais em construções eruditas da tradição escrita ou em estrangeirismos são sentidas como alheias ao sistema do português” (Noll 2008: 228).

Argumenta o estudioso que a dissolução de encontros consonantais já é conhecida nos primórdios da história do português e que a epêntese também pode ser reconhecida no português europeu, não se tratando, portanto, de influência primária de línguas africanas ou indígenas nesse caso. O autor constata a epêntese vocálica em textos setecentistas onde já se observara o vocábulo *adeministra* por *administra*. Nas palavras do pesquisador, “Esse testemunho é interessante porque demonstra que epêntese também fazia parte da fala de pessoas instruídas” (Noll 2008: 228) e, inclusive, é encontrada em textos literários.

Elia (1963) assevera que a epêntese nos grupos consonantais impróprios é uma tendência geral do português do Brasil, revelando a repugnância pelas sílabas fechadas por consoantes. De acordo com o estudioso, “No período românico, tal repugnância se traduzia pelo embebedimento da primeira consoante na segunda (assimilação regressiva), tipo *septe* – *sette* – *sete* ; *ipse* – *esse* – *êsse*; ou então, pela vocalização da primeira consoante, como em *octo* – *oito*” (Elia 1963: 273).

Para Elia, depois de constituída a língua portuguesa, duas tendências diversas passaram a atuar: uma em Portugal, com a valorização da consoante e outra, no Brasil, com a valorização da vogal. Ou seja, tendência ao apagamento da vogal no português europeu ao passo que, no português do Brasil, o segmento vocálico tornou-se mais distinto, excetuando-se, as tônicas igualmente distintas em ambos os países. Segundo o autor:

valoriza-se a consoante em Portugal, nos grupos consonantais, mediante o acréscimo, por assim dizer, de um /e/ brevíssimo, que não chega a permitir a constituição de uma sílaba. No Brasil, ao contrário, a palavra ganha facilmente uma sílaba. Tal fato se dá inclusive entre pessoas cultas. Entre nós [brasileiros], p. ex. o normal é ouvir-se a pronúncia ‘pissicologia’ em vez de psicologia. (Elia 1963: 273)

Elia (1963) prossegue sua argumentação utilizando-se de observações feitas por Souza da Silveira acerca da tonicidade dos vocábulos no português. Assim, segundo a pauta prosódica do português brasileiro, os vocábulos em que ocorrem encontros consonantais impróprios ganham mais uma sílaba e o acento secundário recai sobre a vogal epentética.

O autor conclui que não era preciso apelar para influxos afro-índios na realização deste fenômeno fonético, uma vez que “a pronúncia distinta, clara, das vogais é a tradição mais antiga da língua portuguesa” (Elia 1963: 274), reconhecida por eminentes filólogos como Gonçalves Viana, e constatada como fenômeno românico que ocorreu inclusive no próprio latim. Desse modo,

A debilidade das consoantes pós-vocálicas no português oral do Brasil faz com que, mesmo na pronúncia das pessoas cultas, sejam desmanchados os grupos cuja segunda consoante não seja uma líquida (l ou r). Por isso a tendência é dizer adevogado, obijeto, iguinarar, adimirar, dóguima, rítimo, pissicose. (Elia 1963: 293)

O autor salienta, no entanto, que em grupos próprios, as classes cultas não fazem a epêntese. O mesmo não se dá na fala das classes populares, em que se ouve: *gulória, quilaridade, reculuta (recruta), quelemente (Clemente)* etc., formas consideradas por Silva Neto (1950) como fruto da pronúncia relaxada.

Quanto aos trabalhos dialetológicos, a obra inaugural da dialetologia brasileira, *O Dialeto Caipira* (1982 [1920]), de Amadeu Amaral, apresenta, mesmo que brevemente, casos de epêntese vocálica na pronúncia do caipira do interior de São Paulo, em vocábulos como “*rec-u-luta, Ing-a-laterra, g-a-rampo*” (Amaral 1982: 54), desfazendo-se, desse modo, encontros consonantais com cl-, gl-, gr-.

Antenor Nascentes (1953), em *O Linguajar Carioca*, cuja primeira edição remonta ao ano de 1922, por sua vez, traz um panorama mais abrangente sobre o fenômeno, definindo o segmento vocálico inserido para desfazer encontros consonantais como ‘vogal parasitária’. De acordo com Nascentes (1953: 57): “Em Portugal a realização chega quase à perfeição, porque a língua possui o *e* brevíssimo que falta ao Brasil”.

Para este estudioso, as vogais parasitárias são três: *e*, *i* ou *u* cujo uso está associado à classe inculta da sociedade, uma vez que esta “não nutre preocupações de boa pronúncia e intercala francamente uma vogal sem o caráter de parasitaria e sim com o de uma vogal normal (Nascentes 1953: 57), por outro lado, “a classe semiculta deixa ouvir a vogal parasitaria com menos intensidade do que a classe inculta” (Nascentes 1953: 57). O autor arrola os seguintes encontros consonantais, muitos dos quais não aparecem na linguagem da classe inculta:

bc – obcecado-obicecado, bd – obdurar-obidurar, bj – obfirmar-obifirmar, bg – Abgar-Abegar, bj – abjurar-abijurar, bn – abnegado-abenegado, bs – abisóluto-abesóluto, bt – obter-obeter, bv – obvio-óbivio, cç – convicção-convição, cm – acme-áquime, cn – tenico-téquinico, ct – provecto-provéquito, cz – czar-quezar, dg – Edgar-Edegar, dj – adjacente-adejacente, dm – admirar-adimirar, dn – adnato-adenato, dq – adquirir-adquirir, dv – advogado-adevogado, gm – pragmática-praguimática, gn – indignar-se-indiguar-se, pç – recepção, recepção, mn – amnesia-amnesia, pn – pneumático-pneumático, ps – psicologia-pissicologia, pt – optar-opitar, tm – ritmar-ritimar, zl – guzla-guzila, x (c e s) – fixar-fiquisar. Exemplos com *u* parasitário: Ludgero-Ludugero, Edvirges-Eduvirge (Nascentes 1953: 58).

Pautando-se em Camara Jr., Nascentes afirma que a persistência do /e/ em vez do /i/ ou do /o/ em vez do /u/, é determinada pela natureza da vogal tônica com que a vogal átona tende a harmonizar-se em abrimento bucal. “É por este motivo que encontramos as formas

populares *peneu*, *adevogado* com suarabacti de um /e/ fechado e por outro lado *abissoluto*, onde o /u/ tônico da penúltima sílaba regula o timbre da pretônica imediata e determina o timbre de /i/ para a vogal anaptítica” (Nascentes 1953: 59-60).

Clóvis Monteiro (1933), em a *Linguagem dos Cantadores*, cujo *corpus* se refere à língua falada no Ceará, declara que, no meio de palavras, “evita-se o encontro de consoantes pertencentes a sílabas diferentes mediante a interposição de uma vogal: objeto-obijeto, admira-adimira, advogado-adèvogado” (Monteiro 1933: 61).

Marroquim (1934), em *A Língua do Nordeste*, baseando-se na linguagem de Alagoas e Pernambuco, aponta que “um fenômeno geral de dialeção popular é o desdobramento de grupos consonantais pelo acréscimo de uma vogal entre dois fonemas” (Marroquim 1934: 82), acrescentando que

As classes cultas em certas palavras também dividem os grupos consonantais acrescentando um i: obicecação, abidicar, subijugar (o povo pronuncia sujigar), obinubilação, abistenção, obiturar, óbivio [...] adivogado ou adevogado, dóguima, aguinóstico, minemônica, aquicepção, pineumonia piscicóse, apitidão, néquiso, rítimo. (Marroquim 1934: 85)

Teixeira (1938), acerca do falar mineiro, também arrola uma série de encontros consonantais entre os quais se intercalam vogais, fenômeno denominado pelo estudioso como abrandamento. Neste abrandamento dos encontros consonantais, ocorrem também a vogal média [e] e a alta [u], conforme se observa: “bc (obiceno), bj (obijeto), bn (abinegado), bs (obiservar), bt (obiter), cç (infequição), cl (quilemente), cr (recluta), dq (adquirir), dj (adijacente), dm (adimirar), dv (adivogado), gn (iguinorante), pç (percepção), pn (pineumonia), pt (opitado), tm (atmosfera)”. (Teixeira 1938: 21-22)

Conforme anotado nos trabalhos, casos de suarabácti são comuns tanto no português brasileiro quanto no português europeu, com a tendência à redução da vogal em Portugal e presença de uma vogal alta no Brasil. Por outro lado, observa-se, sobretudo, nos trabalhos de Nascentes e Marroquim que a vogal suarabáctica tende a se realizar como /e/, principalmente, em vocábulos como *advogado* e *pneu* cujas marcas de uso estão associadas a fatores sociais do usuário. O foco

deste trabalho recai sobre o comportamento desta vogal epentética que, no *corpus* analisado, ou ocorre como vogal média alta anterior [e] ou como alta anterior [i] associando-se ao perfil dos informantes, ou seja, nesta oportunidade, o fenômeno é tratado sob uma perspectiva sociodialetoológica.

No tocante a isso, cabe salientar brevemente que os trabalhos atuais em Dialectologia têm adentrado os “veios sociolinguísticos” (Cardoso 2010), agregando à dimensão diatópica outras variáveis extralinguísticas, fazendo-se, portanto, uma Dialectologia Pludimensional (Thun 1998), ou uma abordagem sociodialetoológica da variação linguística (Guy 2012) a partir da combinação espaço geográfico e o espaço social, buscando ora dialetologizar a Sociolinguística ora socializar a Dialectologia (Thun 1998). Desse modo, nas próximas seções, passa-se à descrição dos materiais e métodos empregados no estudo, seguidos da análise do *corpus*.

2. Materiais e Métodos

Para este trabalho, consideram-se dados referentes a dois estados da Federação, Paraná (PR) e São Paulo (SP), coletados pela equipe do Projeto ALiB em 55 municípios. De acordo com os procedimentos metodológicos estabelecidos para o ALiB, no Paraná, a rede de pontos é composta por 17 cidades e, em São Paulo, por 38 localidades.² Em cada um dos municípios, foram entrevistados quatro informantes naturais da região linguística, tendo como nível máximo de escolaridade o ensino

2. A rede de pontos do ALiB é identificada por números. No estado de São Paulo, foram selecionados os seguintes municípios: pontos 150 (Jales), 151 (Votuporanga), 152 (São José do Rio Preto), 153 (Barretos), 154 (Franca), 155 (Andradina), 156 (Araçatuba), 157 (Ribeirão Preto), 158 (Lins), 159 (Ibitinga), 160 (Mococa), 161 (Presidente Epitácio), 162 (Adamantina), 163 (Araraquara), 164 (Teodoro Sampaio), 165 (Presidente Prudente), 166 (Marília), 167 (Bauru), 168 (Mogi Mirim), 169 (Assis), 170 (Bernardino de Campos), 171 (Botucatu), 172 (Piracicaba), 173 (Campinas), 174 (Bragança Paulista), 175 (Taubaté), 176 (Guaratinguetá), 177 (Itapetininga), 178 (Sorocaba), 179 (São Paulo), 180 (Caraguatatuba), 181 (Itararé), 182 (Capão Bonito), 183 (Itanhaém), 184 (Santos), 185 (Ribeira), 186 (Registro), 187 (Cananeia). No Paraná, a rede de pontos é composta pelas seguintes cidades: 207 (Nova Londrina), 208 (Londrina), 209 (Terra Boa), 210 (Umuarama), 211 (Tomazina), 212 (Campo Mourão), 213 (Cândido de Abreu), 214 (Piraí do Sul), 215 (Tolledo), 216 (Adrianópolis), 217 (São Miguel do Iguaçu), 218 (Imbituva), 219 (Guarapuava), 220 (Curitiba), 221 (Morretes), 222 (Lapa), 223 (Barracão).

fundamental. Assim, em cada ponto linguístico, tem-se um homem e uma mulher de 18 a 30 anos (faixa 1) e um homem e uma mulher de 50 a 65 anos (faixa 2), totalizando 220 informantes³.

Nesta oportunidade, foram selecionadas três questões do Questionário Fonético-Fonológico (Comitê Nacional do Projeto ALiB 2001):

- **Questão 54 (AFTOSA):** ... *uma doença que dá no gado, em geral na boca? Dá uma febre. Se não separar o gado doente, ela pega nos outros. É preciso vacinar o gado para ele não ter essa doença.*
- **Questão 72 (PNEU):** ... *aquilo que o carro tem; preto, redondo, se passar por um prego, fura e se esvazia?*
- **Questão 101 (ADVOGADO):** ... *que profissional se pode contratar para defender os interesses na Justiça?*

Estas perguntas buscam documentar a presença ou não de uma vogal suarabáctica para desfazer os encontros consonantais impróprios: *ft, pn e dv*.

Os dados foram transcritos e revisados pela equipe do projeto ALiB – regional Paraná – e, de acordo com o material disponível no banco de dados ainda inédito, constata-se que houve a ocorrência da vogal epentética em todos os inquéritos, seja com a realização da vogal alta anterior [i] ou da vogal média anterior [e], tratando-se, portanto, de uma regra categórica no *corpus* analisado. Focaliza-se, desse modo, como regra variável, o abaixamento da vogal em função de fatores linguísticos e extralinguísticos.

Os dados levantados foram armazenados no software [SGVCLin], que possibilita o acesso a relatórios diversos, bem como a geração de cartas linguísticas de produtividade e de isofônicas (Romano; Seabra; Oliveira 2014). A análise foi realizada sob três perspectivas:

- Tratamento quantitativo considerando-se a distribuição diatópica por meio de cartas linguísticas, bem como sua distribuição em

3. Nas capitais de cada estado foram entrevistados quatro informantes de nível superior, cujos dados não são utilizados neste trabalho.

dados percentuais e relativos segundo variáveis extralinguísticas (sexo, idade e estado);

- Tratamento estatístico no software *SPSS* versão 17.0 (SPSS, 2008) com vistas a realizar testes de hipóteses, discriminadas no final desta seção;
- Tratamento estatístico no software *GoldVarb 2001*, considerando-se variáveis linguísticas e extralinguísticas que podem ou não influenciar na aplicação da regra [i] > [e]. Considera-se como variáveis independentes os seguintes grupos de fatores:
 - a) **Extralinguísticos:** (i) sexo; (ii) faixa etária; (iii) estado;
 - b) **Linguístico:** (iv) item lexical (aftosa, pneu, advogado)⁴.

2.1. Contextualização da análise estatística

O objetivo de uma análise estatística consiste em extrapolar seus resultados para a população estudada. Neste estudo, entende-se por população os informantes representativos do falar de cada estado. Apresentam-se testes de hipóteses com o objetivo de inferir conclusões acerca dessa população. Testes de hipóteses são utilizados para realizar inferências estatísticas por meio do intervalo de confiança, buscando-se delimitar o parâmetro populacional desconhecido. Finalmente, uma hipótese é formulada quanto ao valor do parâmetro e, com base nos elementos amostrais, aplica-se um ou mais testes que indicarão a aceitação ou rejeição da hipótese. Para maiores detalhes sobre o uso de estatística inferencial em estudos linguísticos, confira Oliveira (2009) e Romano; Seabra (2014).

O tratamento estatístico realizado a partir do SPSS considerou os testes de 24 hipóteses nulas, a saber:

- *Hipótese nula 1:* há diferença significativa na distribuição de produtividade da vogal média alta [e] entre as faixas 1 e 2;

4. Estes são os únicos vocábulos do QFF que podem suscitar a realização da suarabácti, desse modo, compreende-se que não se faz necessário controlar o contexto antecedente e subsequente, e também a posição do encontro consonantal no vocábulo.

- *Hipótese nula 2*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre as faixas 1 e 2;
- *Hipótese nula 3*: há diferença significativa na distribuição de produtividade da vogal média alta [e] entre homens e mulheres;
- *Hipótese nula 4*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre homens e mulheres;
- *Hipótese nula 5*: há diferença significativa na distribuição de produtividade da vogal média alta [e] nos estados de São Paulo e Paraná;
- *Hipótese nula 6*: há diferença significativa na distribuição de produtividade da vogal alta [i] nos estados de São Paulo e Paraná;
- *Hipótese nula 7*: há diferença significativa na distribuição de produtividade da vogal média alta [e] entre as faixas 1 e 2 para AFTOSA;
- *Hipótese nula 8*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre as faixas 1 e 2 para AFTOSA;
- *Hipótese nula 9*: há diferença significativa na distribuição de produtividade da vogal média alta [e] entre homens e mulheres para AFTOSA;
- *Hipótese nula 10*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre homens e mulheres para AFTOSA;
- *Hipótese nula 11*: há diferença significativa na distribuição de produtividade da vogal média alta [e] nos estados de São Paulo e Paraná para AFTOSA;
- *Hipótese nula 12*: há diferença significativa na distribuição de produtividade da vogal alta [i] nos estados de São Paulo e Paraná para AFTOSA;
- *Hipótese nula 13*: há diferença significativa na distribuição de produtividade da vogal média alta [e] entre as faixas 1 e 2 para PNEU;
- *Hipótese nula 14*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre as faixas 1 e 2 para PNEU;

- *Hipótese nula 15*: há diferença significativa na distribuição de produtividade da vogal média alta [e] entre homens e mulheres para PNEU;
- *Hipótese nula 16*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre homens e mulheres para PNEU;
- *Hipótese nula 17*: há diferença significativa na distribuição de produtividade da vogal média alta [e] nos estados de São Paulo e Paraná para PNEU;
- *Hipótese nula 18*: há diferença significativa na distribuição de produtividade da vogal alta [i] nos estados de São Paulo e Paraná para PNEU;
- *Hipótese nula 19*: há diferença significativa na distribuição de produtividade da vogal média alta [e] entre as faixas 1 e 2 para ADVOGADO;
- *Hipótese nula 20*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre as faixas 1 e 2 para ADVOGADO;
- *Hipótese nula 21*: há diferença significativa na distribuição de produtividade da vogal média alta [e] entre homens e mulheres para ADVOGADO;
- *Hipótese nula 22*: há diferença significativa na distribuição de produtividade da vogal alta [i] entre homens e mulheres para ADVOGADO;
- *Hipótese nula 23*: há diferença significativa na distribuição de produtividade da vogal média alta [e] nos estados de São Paulo e Paraná para ADVOGADO;
- *Hipótese nula 24*: há diferença significativa na distribuição de produtividade da vogal alta [i] nos estados de São Paulo e Paraná para ADVOGADO.

3. Descrição e análise dos dados

Os relatórios obtidos a partir do [SGVCLin] possibilitam a visualização da distribuição de produtividade em números absolutos e relativos de acordo com as variáveis envolvidas. No *corpus*, observa-se baixo índice de abstenção de respostas, sendo registradas 612 ocorrên-

cias distribuídas entre as duas variantes: 333 registros de [e] e 309 de [i]. Ou seja, ocorre, predominantemente, a vogal média como vogal suarabáctica. No entanto, verifica-se a presença de variantes fonéticas e morfofonêmicas para cada um dos três vocábulos.

Para a questão 54, documentaram-se 11 variantes, perfazendo o total de 204 ocorrências: [afi'tɔzə] (98 ocorrências), [afe'tɔzə] (38), [fe'tɔzə] (20), [afi'tɔzi] (13), [fi'tɔzi] (12), [afe'tɔzi] (6), [fe'tɔzi] (6), [efi'tɔzə] (2), [ve'tɔzi] (1) e [vi'tɔzi] (1).

Para a questão 72, documentaram-se 5 variantes: [pe'new] (98 ocorrências), [pi'new] (40), [pe'news] (5), [pi'news] (3), [pi'newʃ] (1), que totalizam 219 registros.

Para a questão 142, registraram-se 8 variantes: [adzivo'gadʊ] (105 ocorrências), [adevo'gadʊ] (71), [dʒivo'gadʊ] (14), [devo'gadʊ] (14), [adivo'gadʊ] (11), [divo'gadʊ] (2), [adzivo'gadʊs] (1), [dʒivo'gadʊ] (1), totalizando 219 ocorrências.

As diferentes realizações morfofonêmicas dos itens não foram consideradas, pois o objetivo principal é discutir a qualidade da vogal suarabáctica nos encontros consonantais impróprios *ft*, *pn* e *dv*. A Figura 1 ilustra a distribuição percentual dessas vogais de acordo com os vocábulos.

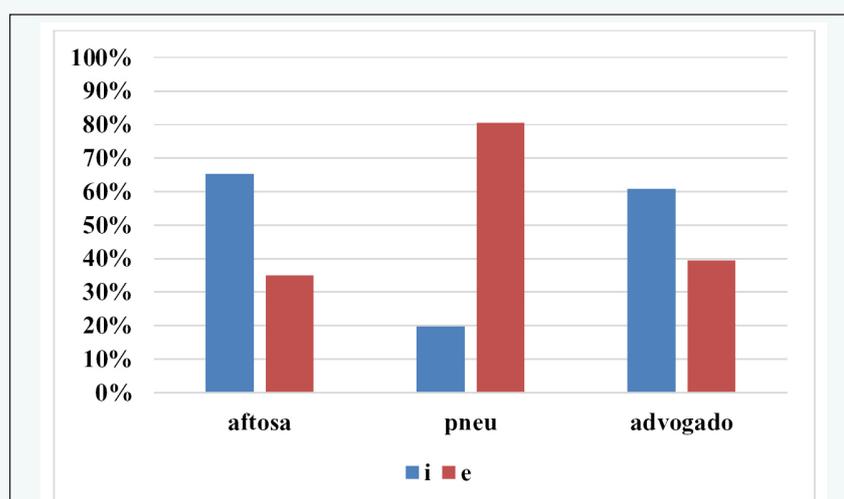


Figura 1 – Distribuição percentual das vogais suarabásticas de acordo com os vocábulos. Fonte: [SGVCLin].

Observa-se que em *aftosa* e *advogado* há a predominância da vogal alta anterior [i] (65,20% e 60,73%, respectivamente). Este comportamento, porém, se inverte em *pneu*, no qual se verifica a predominância da vogal média alta anterior [e] em 80,37% das respostas. O segundo item lexical que mais apresenta a realização da vogal [e] é *advogado* (39,27%). Em *aftosa*, a vogal média figura em, aproximadamente, 35% das realizações.

3.1. Comportamento sociolinguístico: faixa etária e sexo

Esta distribuição percentual, por sua vez, também se associa ao perfil dos informantes. A Tabela 2 apresenta a produtividade dessas vogais considerando as faixas etárias e os itens lexicais.

Tabela 2 – Distribuição de produtividade das vogais suarabáticas segundo a faixa etária e os vocábulos

		[i]	[e]	Total
<i>Aftosa</i>	Faixa 1	85,87% (79)	14,13% (13)	92
	Faixa 2	51,79% (58)	48,21% (54)	112
Total		137	67	204
<i>Pneu</i>	Faixa 1	22,73% (25)	77,27% (85)	110
	Faixa 2	16,51% (18)	83,49% (91)	109
Total		43	176	219
<i>Advogado</i>	Faixa 1	81,65% (89)	18,35% (20)	109
	Faixa 2	40% (44)	60% (66)	110
Total		133	86	219

Fonte: [SGVCLin].

A Tabela 2 mostra que, para *aftosa*, os informantes da faixa 1 utilizam com mais frequência a vogal [i] (85,87%), ao passo que entre os informantes da segunda faixa, há maior equilíbrio entre a realização da vogal alta (51,79%) e média (48,21%). Nesta faixa, a diferença reside apenas em quatro registros a mais de [i] em relação ao [e], conforme se observa pelo número de ocorrências expressos entre parênteses. Para *pneu*, em ambas as faixas, predomina o uso da vogal [e]. O vocábulo *advogado*, por sua vez, apresenta comportamento análogo à *aftosa*, predominando o uso da vogal alta entre os informantes da primeira

faixa (81,65%) em face a 18,35% dos informantes dessa faixa que utilizam a vogal média. No entanto, entre os informantes da faixa 2, nota-se uma inversão na distribuição de ambas as variantes, uma vez que mais da metade dos informantes com este perfil prefere o uso de [e] a [i] (60% e 40%, respectivamente).

A Tabela 3 apresenta a produtividade das variantes considerando-se os sexos e os itens lexicais.

Tabela 3 – Distribuição de produtividade das vogais suarabáticas [i] e [e] segundo o sexo e os vocábulos

		[i]	[e]	Total
<i>Aftosa</i>	<i>Masculino</i>	61,68% (66)	38,32% (41)	107
	<i>Feminino</i>	69,07% (67)	30,93% (30)	97
Total		133	71	204
<i>Pneu</i>	<i>Masculino</i>	21,82% (24)	78,18% (86)	110
	<i>Feminino</i>	17,43% (19)	82,57% (90)	109
Total		43	176	219
<i>Advogado</i>	<i>Masculino</i>	57,80% (63)	42,20% (46)	109
	<i>Feminino</i>	63,64% (70)	36,36% (40)	110
Total		133	86	219

Fonte: [SGVCLin]

Observa-se que para *aftosa* predomina o uso da vogal [i] tanto entre os homens (61,68%) quanto entre as mulheres (69,07%). Verifica-se também que a vogal média, proporcionalmente à vogal alta, ocorre com maior produtividade entre os homens (38,32%), diminuindo sua ocorrência entre as mulheres.

Para *pneu*, há a predominância da vogal média, sendo, contudo, mais produtiva entre as mulheres (82,57%). Proporcionalmente, há menor incidência da vogal alta entre as informantes femininas, apresentando-se mais produtiva entre os informantes masculinos (21,82%).

Quanto ao vocábulo *advogado*, predomina o uso da vogal alta [i], sobretudo, entre as mulheres (63,64%). Neste vocábulo, a vogal média é mais produtiva entre os homens (42,20%).

3.2. Comportamento geolinguístico: estados

Em números absolutos e relativos, a Tabela 4 apresenta a distribuição das ocorrências de acordo com o estado e os vocábulos.

Tabela 4 – Distribuição de produtividade das vogais suarabáticas segundo o estado e os vocábulos

		[i]	[e]	Total
<i>Aftosa</i>	<i>SP</i>	72,03% (103)	27,97% (40)	143
	<i>PR</i>	49,18% (30)	50,82% (31)	61
Total		133	71	204
<i>Pneu</i>	<i>SP</i>	21,05% (32)	78,95% (120)	152
	<i>PR</i>	16,42% (11)	83,58% (56)	67
Total		43	176	219
<i>Advogado</i>	<i>SP</i>	67,76% (103)	32,24% (49)	152
	<i>PR</i>	44,78% (30)	55,22 (37)	67
Total		133	86	219

Fonte: [SGVCLin]

Observa-se que os estados apresentam comportamento diferenciado quanto à produtividade das vogais alta e média. Para *aftosa*, no PR, há equilíbrio na distribuição percentual de ambas as variantes: 49,18% de [i] face a 50,82% de [e]. Já em SP, predomina em 72,03% das respostas a ocorrência da vogal alta em relação aos 27,97% da vogal média.

No que se refere a *pneu*, ambos os estados apresentam maior incidência da vogal média, predominando, entretanto, maior produtividade no PR (83,58%), embora, em números absolutos, haja menor quantidade de ocorrências. Em SP, a vogal média figura, aproximadamente, em 79% e a vogal alta ocorre em 21,05%. Para *advogado*, observa-se, em SP, a predominância da vogal [i] (67,76%), ao passo que, no PR, a maior incidência se registra para a vogal média (55,22%).

A partir do software [SGVCLin], foram geradas cartas linguísticas para cada um dos vocábulos⁵. Trata-se de *cartas diatópicas* com a produtividade das variantes por ponto linguístico e *cartas de isófonas*, que

5. As cartas apresentadas neste estudo são experimentais, elaboradas *ad hoc* a partir da ferramenta [SGVCLin].

apresentam a área de ocorrência das variantes, bem com o percentual de produtividade.

A Figura 2 ilustra a distribuição de ambas as variantes para o vocábulo *aftosa* no PR e em SP. A partir da leitura da carta, verifica-se que, no PR, ocorre a alternância das vogais suarabácticas nas 17 localidades que compõem a rede de pontos do estado. Ademais, a vogal média atinge produtividade máxima em dois pontos linguísticos: 214 (Piraí do Sul) e 223 (Barracão), cuja ocorrência representa 75% das respostas nessas localidades. Por outro lado, a vogal [i] apresenta maior produtividade em duas localidades, ponto 209 (Terra Boa) – 75% e ponto 207 (Nova Londrina) – 87,5%.

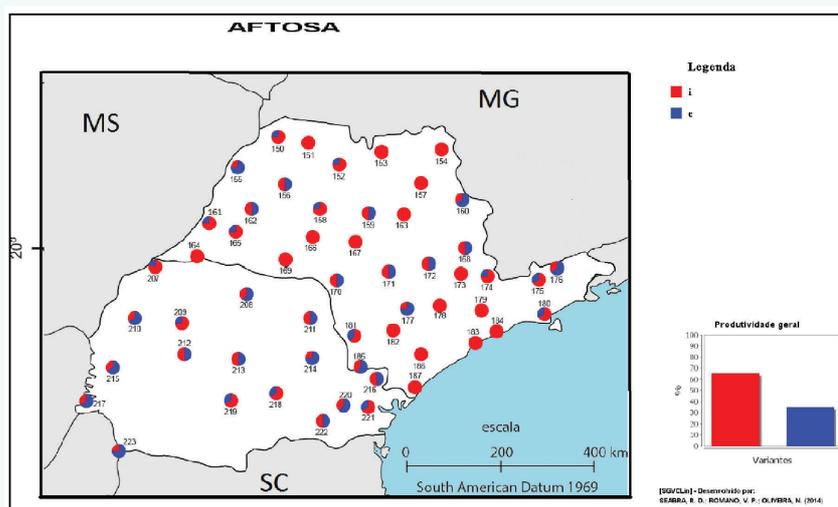


Figura 2 – Distribuição diatópica das vogais suarabácticas [i] e [e] nos estados de SP e PR para *aftosa*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

No estado de SP, dos 37 municípios que compõem a rede de pontos, em 17 deles, não há a alternância no uso das vogais em *aftosa*, sendo registrada apenas a vogal alta. A vogal média apresenta maior índice de ocorrência apenas em duas localidades – pontos 155 (Andradina) e 177 (Itapetininga), cujo percentual figura em 75%.

As Figuras 3 e 4 apresentam, respectivamente, as isófonas da vogal alta [i] e vogal média [e] evidenciando as regiões que apresentam maior e menor produtividade das vogais suarabácticas para *aftosa*.

Na Figura 3, observa-se que a vogal alta ocorre em todas as localidades, apresentando-se com 100% de ocorrência em uma localidade do litoral paulista (ponto 183- (Itanhaém). A vogal alta também se apresenta mais produtiva (80%) em três cidades localizadas no noroeste paulista (pontos 153 – Barretos, 154 – Franca, 157 – Ribeirão Preto), em duas localidades do interior (ponto 163 – Araraquara e 178 – Sorocaba), em uma localidade do litoral sul do estado de SP (187 – Itanhaém), no oeste deste estado (ponto 164 – Teodoro Sampaio) e noroeste paranaense (ponto 207 – Nova Londrina).

Encontram-se também dois grandes feixes de isófonas da vogal [i] que perpassam o estado de SP em sentido transversal, com produtividade de 60%: (i) o primeiro feixe contempla o norte de SP, partindo do ponto 150 (Jales) em direção ao noroeste paranaense (ponto 209 – Terra Boa) e (ii) o segundo parte do Vale do Paraíba (174 – Bragança Paulista), adentra a região metropolitana da capital, atingindo localidades do litoral sul paulista, expandindo-se até uma localidade do interior do estado (ponto 181 – Itararé).

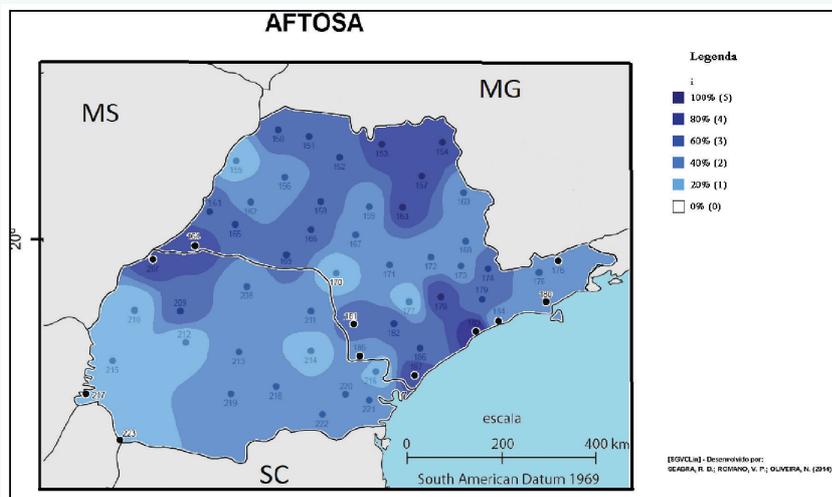


Figura 3 – Áreas de isófonas da vogal suarabática [i] nos estados de SP e PR para *aftosa*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

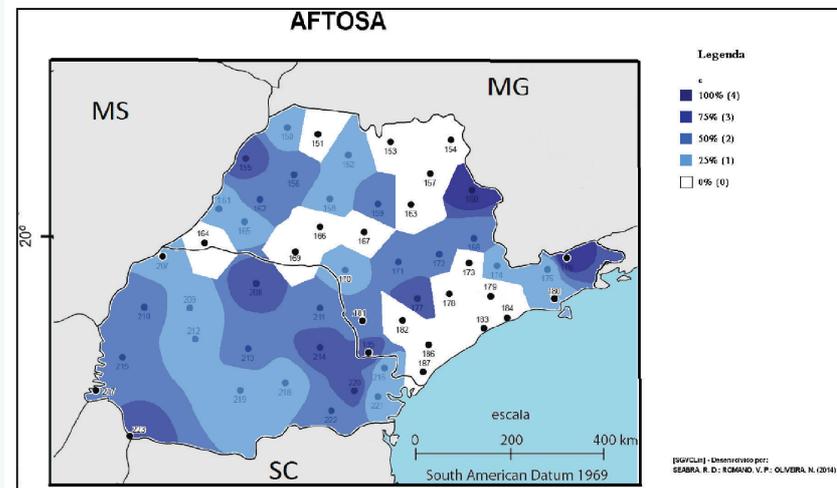


Figura 4 – Áreas de isófonas da vogal suarabática [e] nos estados de SP e PR para *aftosa*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

Quanto às áreas de isófonas da vogal média, a Figura 4 apresenta maior distribuição no território paranaense, contemplando todas as localidades com índice de produtividade no intervalo de 25% a 75%. No Paraná, cidades como Curitiba (ponto 220), Piraí do Sul (ponto 214) e Barracão (ponto 223) apresentam o maior índice da vogal [e] para *aftosa* (75%). Dois feixes de isófonas ocorrem em 50% dos registros: (i) o primeiro localiza-se na porção oeste do PR (pontos 210 – Umuarama, 215 – Toledo e 217 – São Miguel do Iguçu) e (ii) o segundo feixe parte do sul do PR (ponto 222 – Lapa) em direção ao centro e norte do estado (ponto 211 – Tomazina). Esses dois feixes estão separados por uma faixa intermediária que perpassa o estado no sentido transversal (do noroeste ao sul) cuja ocorrência de [e] figura em 25%. O mesmo índice encontra-se no leste do PR, presente em duas localidades (216 – Adrianópolis e 221 – Morretes).

No estado de SP, no entanto, a área de ocorrência da vogal [e] para *aftosa*, apresenta distribuição diatópica irregular. Verifica-se que há cidades paulistas em que não está registrado tal segmento. Todavia, observa-se que o maior índice de ocorrência da vogal [e] localiza-se em duas cidades paulistas (160 – Mococa, 176 – Guaratinguetá), com 100% das respostas.

Quanto a *pneu*, observa-se comportamento diferenciado em relação a *aftosa*, havendo preferência, na maioria dos pontos, pela vogal suarabática [e] (Figuras 5, 6 e 7).

Nas Figuras 5 e 6, é possível visualizar que, em muitos dos pontos linguísticos de SP e PR, a vogal suarabática [e] apresenta-se de forma categórica, atingindo 100% de ocorrência em 22 localidades. Em uma ampla área de ambos os estados, a vogal [e] varia de 50 a 75%, contemplando 32 pontos linguísticos. Apenas em uma localidade – ponto 150 (Jales), esta vogal apresenta-se em 25% das respostas. Por outro lado, a maior incidência da vogal alta ocorre neste ponto linguístico com 100% de produtividade.

Na Figura 7, observa-se que há pontos isolados em ambos os estados em que a vogal alta representa 66% das ocorrências, e uma ampla área do território paulista em que essa vogal ocorre com 33%, bem como áreas no estado de SP e, principalmente, do PR, em que não há o registro da vogal.

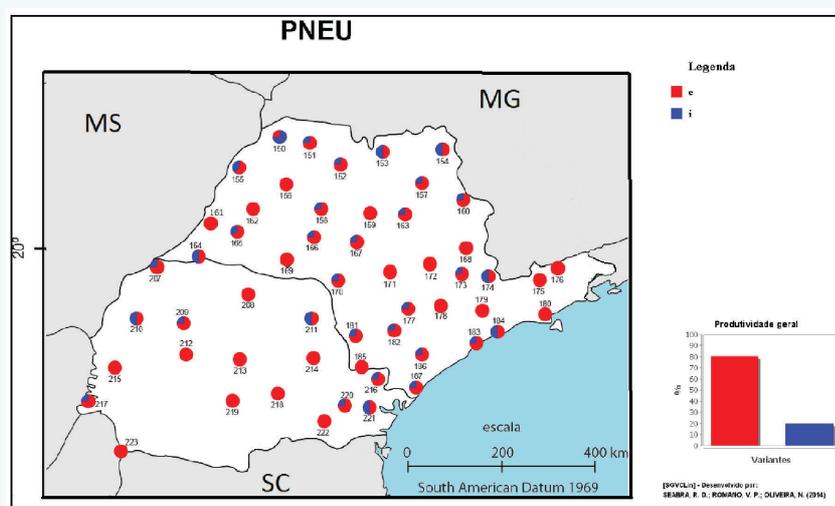


Figura 5 – Distribuição diatópica das vogais suarabáticas [i] e [e] nos estados de SP e PR para *pneu*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

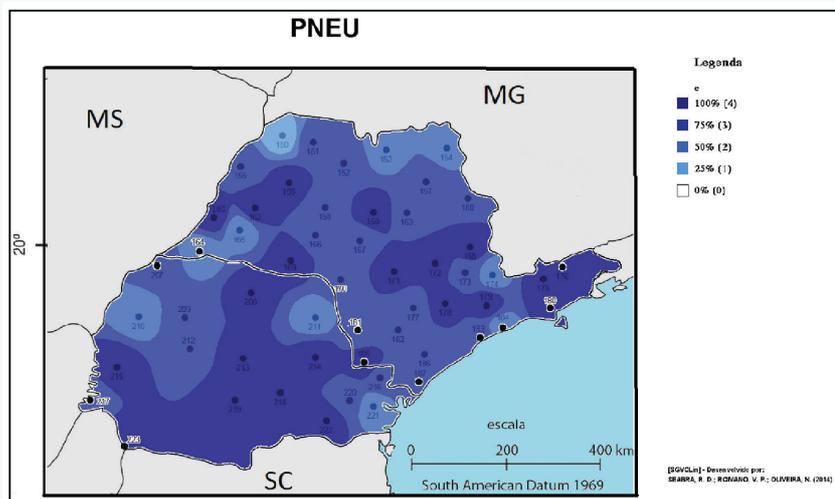


Figura 6 – Áreas de isófonas da vogal suarabática [e] nos estados de SP e PR para *pneu*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

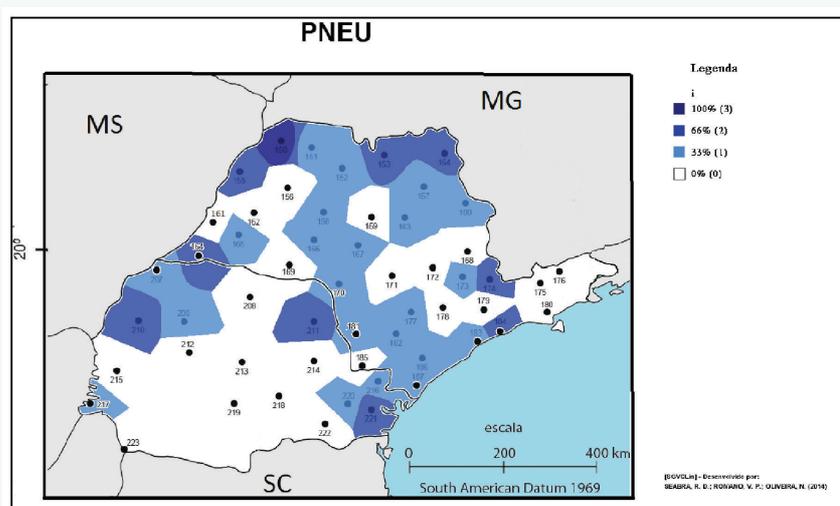


Figura 7 – Áreas de isófonas da vogal suarabática [i] nos estados de São Paulo e Paraná para *pneu*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

Para o vocábulo *advogado*, observa-se a presença da vogal alta e média em ambos os estados (Figuras 8, 9 e 10).

Estudo de vogais suarabácticas na fala de paranaenses e paulistas ...

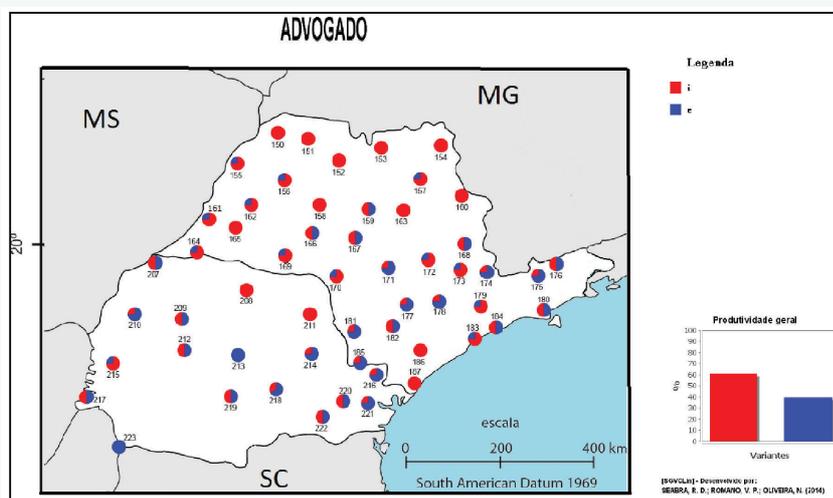


Figura 8 – Distribuição diatópica das vogais suarabácticas [i] e [e] nos estados de SP e PR para *advogado*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

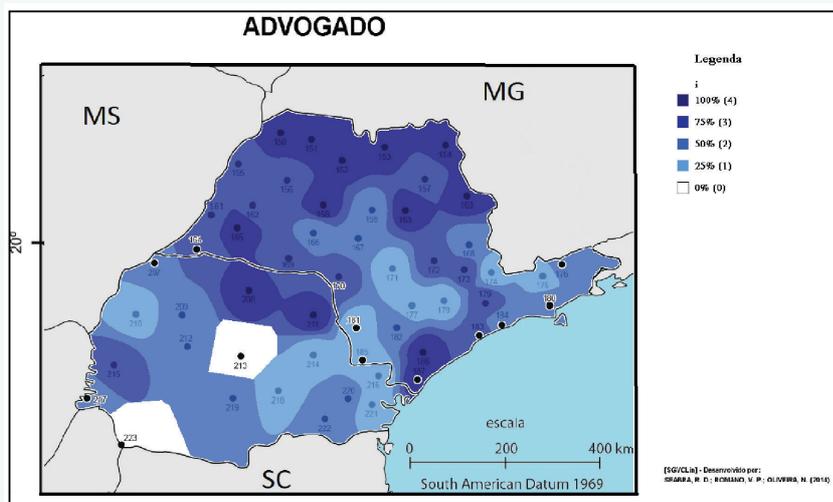


Figura 9 – Áreas de isófonas da vogal suarabáctica [i] nos estados de SP e PR para *advogado*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

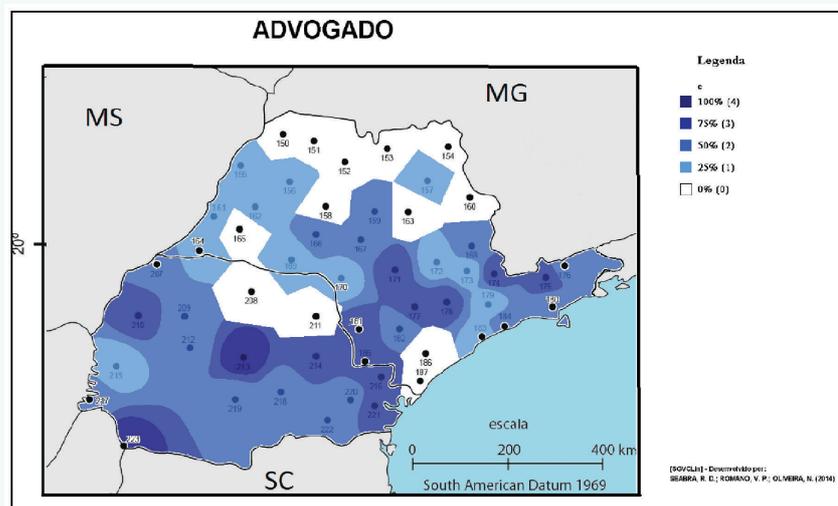


Figura 10 – Áreas de isófonas da vogal suarabática [e] nos estados de São Paulo e Paraná para o vocábulo *advogado*. Fonte: Banco de dados do ALiB (2014) – [SGVCLin].

No PR, a vogal alta apresenta-se com 100% de produtividade em duas localidades do norte do estado: pontos 208 (Londrina) e 211 (Tomazina). Nas demais localidades, predomina o índice de 25 a 50%. Em SP, por sua vez, verifica-se maior número de localidades em que este segmento vocálico ocorre, principalmente, no norte desse estado e em um ponto do interior e litoral sul, em que há registro de 100%.

Quanto à vogal [e] (Figura 10), a maior produtividade registra-se em dois pontos paranaenses (213 – Candido de Abreu e 223 – Barracão). O índice de 75% ocorre no PR, contemplando o leste desse estado, por onde atinge em feixe contínuo o interior de SP (Sorocaba - 178). Esse percentual ocorre também em duas localidades do Vale do Paraíba (pontos 174 – Bragança Paulista e 175 – Taubaté). O índice de incidência de 25 a 50% contempla a maioria das localidades. Verifica-se ainda que cidades do norte paranaense, localidades do litoral sul de SP e ampla faixa territorial do norte paulista não registram a vogal [e] para *advogado*.

3.3. Análise estatística – SPSS

As variáveis numéricas envolvidas no estudo e representadas pelo número de ocorrências de cada variante foram expressas em média e desvio padrão. Em relação às variações categóricas, representadas pelas variantes [e] e [i], foram utilizadas frequências simples e relativas, além do intervalo de confiança de 95%. A análise envolveu o uso do teste de *Kolmogorov-Smirnov (KS)* para avaliar o pressuposto de normalidade dos dados. Valores de p-valor maiores que 0,05 indicam que a variável tem adesão à distribuição normal. A comparação de médias entre grupos, necessária para os testes de hipóteses relativas às variáveis categóricas, considerou a adesão dessas variáveis à normal. Utilizou-se o teste *t de Student* para comparar dois grupos independentes – *faixa etária, sexo e estados*. Os dados referentes a cada variante são analisados segundo sua ocorrência e os resultados são considerados estatisticamente significantes se p-valor < 0,05.

A Tabela 5 apresenta os resultados obtidos para a verificação da normalidade dos dados, que revelaram aderência à distribuição normal ($p = 0,159$; $p = 0,221$).

A aplicação do teste *t de Student* (Tabela 6) confirma as hipóteses nulas 1 e 2, revelando que o teste alcançou nível significativo nos dois casos, assumindo-se que as variâncias dos grupos não são homogêneas. A partir da análise de variância, nota-se que há diferença significativa entre as médias das faixas ($p < 0,001$ nos dois casos) e as hipóteses nulas 1 e 2 podem ser aceitas.

Tabela 5 – Resumo estatístico segundo as vogais

	[e]	[i]
Respostas: média (N*)	6,05 (55)	5,62 (55)
Desvio padrão	2,28	2,24
KS** (p-valor)	1,125 (0,159)	1,049 (0,221)

Fonte: Banco de dados do ALiB (2014)

Nota: *N = Quantidade de localidades

**Teste Kolmogorov-Smirnov

Tabela 6 – Comparação das ocorrências das vogais segundo a faixa etária

	<i>Faixa</i>	<i>Média (N*)</i>	<i>dp</i>	<i>Teste**(p-valor)</i>
<i>[e]</i>	<i>Faixa 1</i>	1,07 (110)	0,751	-7,776 (p < 0,001)
	<i>Faixa 2</i>	1,95 (110)	0,923	
<i>[i]</i>	<i>Faixa 1</i>	1,75 (110)	0,815	5,913 (p < 0,001)
	<i>Faixa 2</i>	1,05 (110)	0,937	

Fonte: Banco de dados do ALiB (2014)

Nota: *N = Quantidade de localidades

** Teste t de Student

Para as hipóteses nulas 3 e 4, a aplicação do teste não alcançou nível significativo ($p = 0,357$ e $p = 0,831$, respectivamente), assumindo-se que as médias entre os grupos (homens e mulheres) para as variantes *[e]* e *[i]* não alcançaram diferença significativa, podendo-se rejeitar ambas as hipóteses.

A Tabela 7 apresenta os resultados dos testes para as hipóteses nulas 5 e 6.

Tabela 7 – Comparação das ocorrências das vogais segundo os estados

	<i>Estado</i>	<i>Média (N*)</i>	<i>dp</i>	<i>Teste**(p-valor)</i>
<i>[e]</i>	<i>SP</i>	1,38 (152)	0,934	-3,315 (0,001)
	<i>PR</i>	1,82 (68)	0,913	
<i>[i]</i>	<i>SP</i>	1,57 (152)	0,911	3,911 (p < 0,001)
	<i>PR</i>	1,04 (68)	0,921	

Fonte: Banco de dados do ALiB (2014)

Nota: *N = Quantidade de localidades

** Teste t de Student

O resultado mostra que o teste alcançou nível significativo nos dois casos. Com base na análise, nota-se que há diferença significativa entre as médias dos estados ($p = 0,001$ para *[e]*; $p < 0,001$ para *[i]*) e as hipóteses nulas 5 e 6 podem ser aceitas.

A partir dos resultados gerais acerca do comportamento das vogais suarabáticas segundo a faixa etária, sexo e estado, passa-se às análises específicas para cada um dos vocábulos. A Tabela 8 apresenta os resultados do teste para *afiosa* segundo a faixa etária.

Tabela 8 – Comparação das ocorrências das vogais segundo a faixa etária para AFTOSA

	<i>Faixa</i>	<i>Média (N*)</i>	<i>dp</i>	<i>Teste**(p-valor)</i>
[e]	<i>Faixa 1</i>	0,12 (110)	0,351	6,407 (p < 0,001)
	<i>Faixa 2</i>	0,53 (110)	0,57	
[i]	<i>Faixa 1</i>	0,72 (110)	0,472	-3,275 (0,001)
	<i>Faixa 2</i>	0,49 (110)	0,554	

Fonte: Banco de dados do ALiB (2014)

Nota: *N = Quantidade de localidades

** Teste t de Student

Novamente, o teste alcançou nível significativo nos dois casos. A partir da análise, nota-se que há diferença significativa entre as médias das faixas etárias (p < 0,001 para [e]; p = 0,001 para [i]) e as hipóteses nulas 7 e 8 podem ser aceitas.

Os resultados dos testes das hipóteses 9 e 10 não alcançaram níveis significativos. Ambas as variantes [e] e [i], considerando o sexo, não apresentam diferença estatisticamente significativa quanto à realização das vogais suarabáticas em *aftosa* (p = 0,150 e p = 0,898, respectivamente).

Quanto à distribuição estatística das vogais suarabáticas nos estados, a Tabela 9 apresenta os resultados do teste.

Tabela 9 – Comparação das ocorrências das vogais segundo os estados para AFTOSA

	<i>Estado</i>	<i>Média (N*)</i>	<i>dp</i>	<i>Teste**(p-valor)</i>
[e]	<i>SP</i>	0,26 (152)	0,498	-2,598 (0,01)
	<i>PR</i>	0,46 (68)	0,531	
[i]	<i>SP</i>	0,68 (152)	0,51	3,143 (0,002)
	<i>PR</i>	0,44 (68)	0,529	

Fonte: Banco de dados do ALiB (2014)

Nota: *N = Quantidade de localidades

** Teste t de Student

O resultado mostra que o teste alcançou nível significativo nos dois casos. A análise revelou que há diferença significativa entre as médias dos estados (p = 0,01 para [e]; p = 0,002 para [i]) e as hipóteses nulas 11 e 12 podem ser aceitas.

Para *pneu* nenhuma das seis hipóteses formuladas foram aceitas, rejeitando-se as hipóteses nulas 13 a 18 (H13: $p = 0,314$, H14: $p = 0,236$, H15: $p = 0,502$, H16: $p = 0,398$, H17: $p = 0,562$, H18: $p = 0,402$).

As Tabelas 10 e 11 apresentam os resultados dos testes das hipóteses estabelecidas para *advogado*.

Tabela 10 – Comparação das ocorrências das vogais segundo a faixa etária para ADVOGADO

	<i>Faixa</i>	<i>Média (N*)</i>	<i>dp</i>	<i>Teste**(p-valor)</i>
<i>[e]</i>	<i>Faixa 1</i>	0,18 (110)	0,387	7,002 ($p < 0,001$)
	<i>Faixa 2</i>	0,6 (110)	0,492	
<i>[i]</i>	<i>Faixa 1</i>	0,81 (110)	0,395	-6,8 ($p < 0,001$)
	<i>Faixa 2</i>	0,4 (110)	0,492	

Fonte: Banco de dados do ALiB (2014)

Nota: *N = Quantidade de localidades

** Teste t de Student

A Tabela 10 mostra que o teste alcançou nível significativo nos dois casos. Observa-se que há diferença significativa entre as médias das faixas etárias ($p < 0,001$ nos dois casos) e as hipóteses nulas 19 e 20 podem ser aceitas. Os testes das hipóteses 21 e 22 não apresentaram diferença estatística significativa ($p = 0,409$ e $p = 0,337$, respectivamente).

A Tabela 11 apresenta os resultados dos testes para as hipóteses nulas 23 e 24.

Tabela 11 – Comparação das ocorrências das vogais segundo os estados para ADVOGADO

	<i>Estado</i>	<i>Média (N*)</i>	<i>dp</i>	<i>Teste**(p-valor)</i>
<i>[e]</i>	<i>SP</i>	0,32 (152)	0,469	-3,171 (0,002)
	<i>PR</i>	0,54 (68)	0,502	
<i>[i]</i>	<i>SP</i>	0,68 (152)	0,469	3,385 (0,001)
	<i>PR</i>	0,44 (68)	0,5	

Fonte: Banco de dados do ALiB (2014)

Nota: *N = Quantidade de localidades

** Teste t de Student

O teste alcançou nível significativo nos dois casos. Nota-se que há diferença significativa entre as médias dos estados ($p = 0,002$ para *[e]*; $p = 0,001$ para *[i]*) e as hipóteses nulas 23 e 24 podem ser aceitas.

Do exposto, conclui-se que das 24 hipóteses testadas, 12 foram aceitas. Duas observações gerais podem ser auferidas: (i) os testes que envolviam a variável extralinguística sexo não confirmam diferenças estatísticas significativas na distribuição de produtividade, e as hipóteses 3, 4, 9, 10, 15, 16, 21 e 22 foram rejeitadas. Ademais, (ii) *pneu*, comportou-se diferentemente dos outros vocábulos, uma vez que *aftosa* e *advogado* apresentam-se, predominantemente, com a vogal alta, ao passo que *pneu* figura, na maioria dos casos, com vogal média. Entretanto, a distribuição entre vogal média e alta neste vocábulo, considerando-se as variáveis faixa etária e estado não são estatisticamente significativas, e as hipóteses nulas 13, 14, 17 e 18 são rejeitadas. Não obstante, as hipóteses gerais 1 e 2 são confirmadas e as que envolviam faixa etária e estado para os vocábulos *aftosa* e *advogado* também revelaram diferenças estatísticas significativas (Hipóteses: 5, 6, 7, 8, 11, 12, 19, 20, 23 e 24).

Os dados foram submetidos ainda ao tratamento estatístico a partir do software GoldVarb 2001. Os resultados revelam informações que confirmam os testes de hipóteses empreendidos.

3.4. Análise estatística – GoldVarb 2001

A primeira rodada dos dados apresentou um total de 641 registros. De acordo com o relatório da rodada binomial, a melhor rodada (*up run*) eliminou a variável sexo na aplicação da regra e considerou os seguintes grupos de fatores significativos, em ordem de relevância: (i) item lexical; (ii) faixa etária; (iii) estado.

Quanto ao item lexical, a Tabela 12 sintetiza os seguintes resultados:

Tabela 12 – Aplicação da regra [i] > [e] segundo o item lexical

Fatores	Aplicação / Total	%	Peso relativo (Pr)
<i>Aftosa</i>	69/202	34	0,28
<i>Pneu</i>	176/220	80	0,8
<i>Advogado</i>	86/229	39	0,35
Total	331/641	51	

Input: 0,531
Significância: 0,000

Os vocábulos *aftosa* e *advogado* não favorecem a aplicação da regra (Prs: 0,28 e 0,35, respectivamente). Por outro lado, *pneu* comporta-se como um item que favorece a ocorrência da vogal média, ou seja, observa-se que há maior tendência à aplicação da regra neste vocábulo.

Considerando a faixa etária, a Tabela 13 apresenta o seguinte panorama:

Tabela 13 – Aplicação da regra [i] > [e] segundo a faixa etária do informante

Fatores	Aplicação / Total	%	Peso relativo (Pr)
<i>Faixa 1 (18 a 30 anos)</i>	118/312	37	0,32
<i>Faixa 2 (50 a 65 anos)</i>	213/329	64	0,67
Total	331/641	51	

Input: 0,531
Significância: 0,000

O grupo indica que, entre os informantes da faixa 1, a regra não tende a se aplicar (Pr: 0,32), mostrando que a maioria desses informantes utiliza a vogal [i]. Por outro lado, entre os informantes da faixa 2, a regra tende a se aplicar (Pr: 0,67), predominando a vogal média em 64% das respostas.

O grupo de fatores (iii) – estado – também foi considerado relevante para a aplicação da regra (Tabela 14).

Tabela 14 – Aplicação da regra [i] > [e] segundo o estado

Fatores	Aplicação / Total	%	Peso relativo (Pr)
<i>PR</i>	124/196	63	0,65
<i>SP</i>	207/466	46	0,43
Total	331/642	51	

Input: 0,48
Significância: 0,001

Conforme se observa, no PR, há maior probabilidade de a regra acontecer (Pr: 0,65), pois, das 196 ocorrências, 124 são com vogal suarabáctica [e]. No estado de SP, predomina o uso da vogal [i], desfa-

vorecendo a aplicação da regra (Pr: 0,43), embora a diferença do peso relativo e dos dados percentuais não seja tão discrepante em relação ao uso da vogal majoritária [i].

Os resultados do GoldVarb 2001 não indicam que a variável sexo interfere na aplicação da regra, portanto, o grupo (i) foi desconsiderado nos resultados da binomial, pois esta variável extralinguística não favoreceu nem desfavoreceu a aplicação da regra, uma vez que os pesos relativos mantiveram-se próximos a 0,5 (Pr: 0,51 e 0,48), para os homens e mulheres, respectivamente, bem como os dados percentuais e o número de ocorrências.

Considerações Finais

A epêntese vocálica do português é um fenômeno fonético que já fora observado pelos primeiros dialetólogos brasileiros: Amaral (1920), Nascentes (1922/1953), Monteiro (1933), Marroquim (1934) e Teixeira (1938). Ademais, a qualidade da vogal para desfazer encontros consonantais impróprios é variável, podendo se realizar como vogal alta, média ou central, conforme também atestam Câmara Junior (1977), Cagliari (1981) e Lee (1993). Os estudiosos apontam que a presença da vogal suarabática não é uma peculiaridade do português brasileiro, uma vez que é atestada na passagem do latim ao português. Porém, observa-se uma tendência ao apagamento desta vogal no português europeu, ao passo que, no português brasileiro, a presença da vogal ocorre, inclusive, entre as pessoas cultas, como uma forma de evitar estruturas silábicas não-possíveis, priorizando-se a eufonia (Lee 1993). Nascentes (1953: 57), todavia, associa a presença da ‘vogal parasitária’ às pessoas incultas “que não nutrem preocupação com a boa pronúncia” e atribui a qualidade desta vogal à natureza do segmento vocálico tônico com que a vogal átona tende a harmonizar-se.

Os métodos de análise empreendidos no estudo possibilitaram verificar o comportamento das variáveis envolvidas na realização das variantes [e] e [i] nos vocábulos selecionados.

Quanto à faixa etária, verifica-se que: (i) em *aftosa*, há o predomínio da vogal alta [i] entre os informantes da primeira faixa e, entre os informantes da segunda, há equilíbrio no uso da vogal alta ou média;

(ii) *pneu* apresenta alta produtividade da vogal média em ambas as faixas; e (iii) *advogado* figura com vogal alta, predominantemente, entre os informantes da faixa 1, ao passo que, na faixa 2, predomina o uso da vogal média. Tanto o SPSS quanto o GoldVarb ratificam estatisticamente a importância desta variável na realização da regra.

No que tange ao estado, nas localidades paranaenses há maior uso da vogal média, principalmente, em *aftosa* e *advogado* em relação às cidades paulistas que, para estes mesmos vocábulos, apresentam maior produtividade da vogal alta. Por outro lado, para *pneu*, verifica-se maior incidência da vogal média tanto em SP quanto no PR, este último tendendo ao equilíbrio entre as variantes.

Diante do exposto, os resultados do GoldVarb permitem afirmar que o item lexical é uma variável importante para a realização da regra, sendo o vocábulo *pneu* o que tende a ocorrer com vogal epentética [e], ao passo que *aftosa* e *advogado* apresentaram pesos relativos desfavoráveis à regra variável.

No que se refere à variável sexo, podem-se salientar duas conclusões: (i) homens e mulheres preferem vogal alta em *pneu* e *advogado*, predominando, entretanto, maior produtividade entre as mulheres no uso dessa vogal; e (ii) há inversão na produtividade das vogais suarabáticas na realização de *pneu*, ocorrendo entre as mulheres o uso da vogal média. Contudo, a análise estatística empreendida no GoldVarb não considerou esta variável como um fator que interfere na realização da regra, pois os pesos relativos são neutros (próximos a 0,5) e o tratamento empregado pelo SPSS confirma que a diferença entre os grupos (homem e mulher) não é estatisticamente significativa.

O fenômeno da suarabácti no português, portanto, pode ser compreendido sob uma perspectiva pancrônica, uma vez que está atestado na história da língua portuguesa, caracterizando-se como um típico fenômeno em variação. A qualidade da vogal está condicionada, no *corpus* analisado, pelo item lexical e por variáveis extralinguísticas como faixa etária e estado, conforme comprovado estatisticamente.

Recebido em junho de 2014

Aprovado em dezembro de 2016

E-mails: valter.romano@hotmail.com

rdsea@hotmail.com

Referências bibliográficas

- AMARAL, Amadeu. 1982 [1920]. *O Dialeto Caipira: gramática e vocabulário*. 4.ed. São Paulo: Hucitec; INL.
- BOLÉO, Manuel de Paiva. 1943. *Brasileirismos: problemas de método*. Separata da Revista Brasília, vol. III. Coimbra: Coimbra Ed. Ltda.
- CARDOSO, Suzana Alice Marcelino Silva. 2010. *Geolinguística: tradição e modernidade*. São Paulo: Parábola Editorial.
- CAMARA JR., Joaquim, Mattoso. 1977. *Estrutura da língua portuguesa*. 8. Ed. Petrópolis: Vozes.
- CARVALHO, Dolores Garcia & NASCIMENTO, Manoel. 1984. *Gramática história: segundo grau e vestibulares*. 14. Ed. São Paulo: Ática.
- COLLISCHONN, Gisela. 1996. Um estudo da epêntese à luz da Teoria da Sílabas de Junko Itô (1986). *Letras de Hoje*, Porto Alegre, v. 31, n. 2. p. 149-158.
- _____. 2000. A epêntese vocálica no português do Sul do Brasil: análise variacionista e tratamento pela Teoria da Otimalidade. *Letras de Hoje*, Porto Alegre, v. 35, n.1, p.285-318.
- _____. 2002. A epêntese vocálica no português do sul do Brasil. In: BISOL, Leda; BRESCANCINI, Cláudia (Orgs.). *Fonologia e variação: recortes do português*. Porto Alegre: EDIPUCRS, p. 205-230.
- _____. 2003. A epêntese vocálica no português do Sul do Brasil: variáveis extralinguísticas. *Revista Letras*, Curitiba, n. 61 (especial), p. 285-297.
- _____. 2004. Epêntese vocálica e restrições de acento no português do Sul do Brasil. *Signum: Estudos da Linguagem*, Londrina, v. 7, n. 1, p. 61-78.
- COMITÊ NACIONAL DO PROJETO ALiB. 2001. *Questionários 2001*. Londrina: EDUEL.
- COUTINHO, Ismael de Lima. 1976. *Pontos de gramática histórica*. 7. Ed. Rio de Janeiro: Ao Livro Técnico.
- CUNHA, Celso. 1986. Conservação e inovação no português do Brasil. *O eixo e a roda*. n.5, Belo Horizonte, p. 199-230. Disponível em: <http://www.lettras.ufmg.br/poslit/08_publicacoes_txt/CELSOCUNHA.pdf>. Acesso em: 27 mai. 2014.
- ELIA, Sílvio. 1963. *Ensaio de filologia*. Rio de Janeiro: Livraria Acadêmica.
- GUY, Gregory R. 2012. Rumos da sociodialetoлогия da América Latina. In: CONGRESSO INTERNACIONAL DE DIALETOLOGIA E SOCIOLINGUÍSTICA (2.: 2012: Belém, PA). Diversidade linguística e políticas de ensino: *Anais...* p. 44-60. São Luís: EDUFMA.

- LEE, Seung-Hwa. 1993. Epêntese no português. *Estudos Lingüísticos XXII* – Anais de Seminários do GEL, Ribeirão Preto, Instituição Moura Lacerda, v. II, p. 847-854.
- MARROQUIM, Mário. 1996 [1934]. *A Língua do Nordeste*. 3. ed. Curitiba: HD Livros.
- MASSINI-CAGLIARI, Gladis. 2000. Epêntese e paragoge: processos fonológicos distintos. In: II CONGRESSO INTERNACIONAL DA ABRALIN, Florianópolis. *Anais...* Florianópolis: Associação Brasileira de Lingüística, p. 400-410. CD-ROM.
- MIGLIORINI, Livia. 2010. Revistando a paragoge em português arcaico sob a ótica da Fonologia Lexical. *Revista Eutomia: Revista online de literatura e lingüística*, Pernambuco, ano III, v. 2, p. 1-17.
- MIGLIORINI, Livia & MASSINI-CAGLIARI, Gladis. 2011. A epêntese vocálica no português brasileiro: regra lexical ou pós-lexical? *Todas as Letras: Revista de Língua e Literatura*, São Paulo, v.13, n.1, p.72-84.
- MONTEIRO, Clóvis. 1933. *A Linguagem dos Cantadores*. Rio de Janeiro: Borsoi.
- NASCENTES, A. 1953 [1922]. *O Linguajar Carioca*. 2. ed. Rio de Janeiro: Organizações Simões.
- NOLL, Voker. 2008. *O português brasileiro: formação e contrastes*. Trad. Mário Eduardo Viário. São Paulo: Globo.
- OLIVEIRA, Alan Jardel de. 2009. Análise quantitativa no estudo da variação lingüística: noções de estatística e análise comparativa entre Varbrul e SPSS. *RELin: Revista de Estudos da Linguagem*. Belo Horizonte, UFMG, v.17, n.2, p. 93-119.
- ROBINSON, John S.; LAWRENCE, Helen R. & TAGLIAMONTE, Sali A. 2001. *GOLDVARB 2001: a multivariate analysis application for Windows*. Department of Language and Linguistic Science, York, Canada: University of York.
- ROMANO, Valter Pereira & SEABRA, Rodrigo Duarte. 2014. Dados geolinguísticos sob uma perspectiva estatística: a variação lexical no Centro-Oeste, Sudeste e Sul do Brasil. *RELin: Revista de Estudos da Linguagem*. Belo Horizonte, UFMG, v. 22, n.2, p.59-92.
- ROMANO, Valter Pereira; SEABRA, Rodrigo Duarte & OLIVEIRA, Nathan. 2014. [SGVCLin] – Software para geração e visualização de cartas lingüísticas. *RELin: Revista de Estudos da Linguagem*. Belo Horizonte, UFMG, v. 22, n.1, p. 119-151.
- SEABRA, Rodrigo Duarte; ROMANO, Valter Pereira & OLIVEIRA, Nathan. 2014. A Brazilian Contribution for Automated Linguistic Cartography. *American Journal of Linguistics*, Oxford, v.3, n. 2, p 27-40.

- SILVA NETO, Serafim da. 1950. *Introdução ao estudo da língua portuguesa no Brasil*. Rio de Janeiro: INL.
- SPSS for Windows. 2008. Version 17.0. [S.l]: IBM, 1 CD-ROM.
- TEIXEIRA, José Aparecido. 1938. *O falar mineiro*. Separata da Revista do Arquivo Municipal de São Paulo, São Paulo: Prefeitura Municipal de São Paulo. Departamento de Cultura.
- TEYSSIER, Paul. 1994. *História da língua portuguesa*. Trad. Celso Cunha, 6. Ed. Lisboa: Sá da Costa.
- THUN, Harald. 1998. La geolingüística como lingüística variacional general. *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*, v. 5, p. 701-729, Tübingen: Niemeyer.