

O DESENVOLVIMENTO DA CONSCIÊNCIA SOCIOLINGÜÍSTICA E O SUCESSO NO DESEMPENHO EM LEITURA

Raquel Meister Ko FREITAG*

- **RESUMO:** Avaliações nacionais e internacionais mostram que o Brasil falha no ensino de leitura em todo o sistema escolar. Neste artigo defende-se que a explicação para o sucesso no aprendizado inicial da leitura de uma parte destes estudantes pode estar relacionada ao desenvolvimento da consciência sociolingüística. As taxas de variação de oito variáveis fonológicas (monotongação de -ow e de -aj, -ej, ditongação, apagamento de -R em coda em nomes e em verbos, apagamento de -d em -ndo, simplificação de cluster e vocalização da lateral palatal) em três diferentes contextos estilísticos (do menos para o mais monitorado: entrevista curta - tarefa de nomeação de figuras - tarefa de leitura em voz alta) coletados em uma sala de aula da 3ª série mostram uma relação gradual entre o aumento do monitoramento estilístico e o uso da variante padrão e efeito positivo no desempenho em um teste de compreensão em leitura. Esses resultados revelam que o desenvolvimento da consciência sociolingüística pode ser medido pela transposição da variante da fala para a leitura em voz alta, sugerindo a automatização no processo de decodificação por via lexical.
- **PALAVRAS-CHAVE:** consciência sociolingüística; leitura em voz alta; compreensão em leitura.

Introdução¹

Os resultados de avaliações nacionais, como a Avaliação Nacional da Alfabetização (ANA) e internacionais, como o Programa Internacional de Avaliação dos Alunos (PISA) mostram que o Brasil falha no ensino de leitura em todo o sistema escolar. Apenas um exemplo, em Sergipe, os resultados da ANA de 2015 apontam que apenas 25% das crianças atingem a proficiência em leitura adequada para o 2º ano do ensino fundamental. Este não é um problema novo, nem é exclusivo da sociedade brasileira (LABOV, 1972a), mas que aqui parece estar estagnado, sem conseguir avanços.

* Universidade Federal de Sergipe (UFS). São Cristóvão – SE - Brasil. rkofreitag@uol.com.br. ORCID: 0000-0002-4972-4320. rkofreitag@academico.ufs.br.

¹ Versão deste texto foi apresentada no 48 *New Ways of Analyzing Variation* (NWAV), 2019, em Eugene, Oregon, Estados Unidos.

O problema brasileiro da leitura pode ser em parte devido às diferenças entre a variedade padrão e a dialetal, que ainda são pouco exploradas nos programas nacionais para o ensino de leitura. Diferentes métodos e técnicas têm sido apresentados e propostos como paliativos; e a Sociolinguística pode contribuir para a redução desta assimetria não como paliativo, mas como peça essencial, com o desenvolvimento da consciência sociolinguística. Neste artigo, assume-se a que a explicação para o sucesso na leitura de uma parte dos estudantes pode estar relacionada ao desenvolvimento da consciência sociolinguística: a consciência da existência de formas variáveis na língua e de seu significado social e/ou estilístico em relação aos contextos de uso, relações sociais e identidades pessoais (VAN COMPERNOLLE; WILLIAMS, 2013). Este artigo tem duplo objetivo:

- i) Comparação entre o comportamento de traços linguísticos variáveis e o desempenho em um teste de compreensão leitora, de modo a identificar o poder preditivo dos traços sociolinguísticos, para contribuir em programas de ensino de leitura que considerem o desenvolvimento da consciência sociolinguística.
- ii) Aprimoramento da escala de avaliação social adotada na pesquisa sociolinguística (indicadores, marcadores e estereótipos), com a testagem do comportamento de oito variáveis fonético-fonológicas do português brasileiro em diferentes situações estilísticas.

Por adentrar na relação entre consciência e valores estilísticos, a observação da emergência da consciência sociolinguística pode retroalimentar a investigação das escalas de avaliação social de traços linguísticos, que, no português brasileiro, são, via de regra, inferidas a partir da distribuição de frequências quanto aos perfis sociais (FREITAG, 2016).

Consciência linguística e sociolinguística

A avaliação da língua é determinante para a constituição da identidade linguística dos falantes, em termos de apreciação social e consciência. Diferentes abordagens tentam explicar essa relação, como as de Labov (1972b), com indicadores, marcadores e estereótipos; Bell (1984), com *audience design*; Silverstein (2003), com ordens indexicais; Eckert (2008), com campos indexicais. Em comum, estas propostas assumem que os falantes fazem escolhas durante a fala, mas essas escolhas não necessariamente estão no nível da consciência. A consciência sociolinguística é um conhecimento explícito que emerge de experiências agregadas de reconhecer as diferenças linguísticas no momento da interação, e chega à compreensão como linguística e socialmente significativas (SQUIRES, 2016).

Todos os falantes de uma língua desenvolvem uma consciência de língua (*language awareness*), uma gama de conhecimentos específicos sobre a língua: no primeiro nível,

estrutura e gramática; no segundo, conhecimento sobre os aspectos sociais de uma língua, incluindo variação linguística, variabilidade no ajuste entre falante e audiência, intenções do falante e *code-switching* (consciência sociolinguística); no terceiro nível, envolve o conhecimento sobre como a língua pode ser deliberadamente manipulada para efeitos persuasivos; e no quarto nível, o nível da consciência crítica da língua envolve o reconhecimento de como padrões sociais e discursivos são mutuamente constitutivos, e como os falantes estão amplamente imersos e condicionados por estes padrões linguísticos da comunidade (PIPER, 2003).

A escola explora a consciência linguística de modo particular no aprendizado inicial da leitura, com a consciência fonológica e morfossintática, que contribuem para a automaticidade na decodificação e compreensão leitora. A consciência sociolinguística, no entanto, é ainda pouco explorada, mesmo nos estudos sociolinguísticos.

Todos os falantes, em algum momento, mobilizam a consciência sociolinguística ao tentar explicar por que uma pessoa fala diferente. O que as pessoas comuns, não linguistas, pensam sobre a língua? Por um lado, existe uma consciência sociolinguística popular: as explicações, explanações e as crenças sobre as relações entre língua e contexto social feitas por não especialistas. Por outro, age a força do prescritivismo, conjunto de práticas metalinguísticas normativas, com foco no valor de correção, no uso “correto”, de acordo com a norma codificada na gramática. Enquanto o prescritivismo tenta se amparar em modelos científicos, a consciência sociolinguística popular é ignorada pela ciência. McKenzie (2015) destaca a importância dada ao modo como a população em geral percebe, processa e codifica a variação linguística na fala, incluindo a identificação de variedades linguísticas como marcas social ou regionalmente localizadas. Mobilizando a consciência sociolinguística, estudos de percepção podem, por exemplo, prover informação científica para os rótulos atribuídos a variantes sociais e regionais específicas por não linguistas, assim como ajudar a identificar os traços (sócio)linguísticos que determinam a sua classificação, e auxiliar o pesquisador na sua tarefa de mapeamento e classificação de variedades e traços.

Como dito anteriormente, os traços linguísticos variáveis podem ser escalonados em função do nível de apreciação social (LABOV, 1972b): estereótipos - traços linguísticos socialmente marcados de forma consciente pelos falantes; marcadores - traços linguísticos com padrões sociais e estilísticos que têm efeitos consistentes sobre o julgamento consciente ou inconsciente do ouvinte sobre o falante; e indicadores - traços socialmente estratificados, no entanto, não são sujeitos à variação estilística. E quanto à consciência, os fenômenos variáveis podem estar ao nível da consciência ou abaixo. Falantes são mais conscientes das variáveis com variantes que são foneticamente diferentes ou com variantes envolvidas na manutenção do contraste fonológico (TRUDGILL, 1986). O conhecimento da variação decorre de um processo semiótico durante o qual os falantes são expostos a diferenças linguísticas, observam-nas e sistematizam padrões de seu uso em conexão com fatos sociais (SQUIRES, 2016). Enquanto muitas pesquisas investigaram o estado do conhecimento dos falantes sobre fatos da variação, em estudos descritivos da produção linguística, o processo pelo qual

os falantes constroem esse conhecimento (consciência sociolinguística) é ainda pouco explorado. Diferenças entre percepção, reconhecimento e compreensão da variação, nos termos explicitados por Squires (2016), têm sido amplamente discutidas no campo da aquisição de segunda língua, mas ainda são pouco exploradas na língua materna.

Medidas de consciência sociolinguística podem ajudar a responder questões como: quais traços (sociais, psicológicos) de falantes individuais são ativados por traços de sua fala? Como os falantes de diferentes variedades classificam e rotulam os mesmos sons? Como explicar a habilidade de falantes perceberem certos traços variáveis da língua mas outros não? Como sistematizar a habilidade dos falantes para identificar e categorizar variedades de uma língua falada ou a origem social ou geográfica do falante?

Seguindo a proposta de apreciação social de Labov (1972b), resultados de estudos sociolinguísticos de produção linguística são interpretados quanto à indexação a grupos sociais de dadas variantes; as variantes são assumidas como marcadoras de identidade social, cujo uso distancia ou aproxima de outros.

A consciência sociolinguística popular tem bem clara a construção de estereótipos sociais: o conjunto decorrente do arranjo de certos traços linguísticos variáveis evoca atitudes para com os seus utilizadores. Assume-se que as variantes linguísticas podem marcar identidades sociais e evocar atitudes, mas não se tem ainda uma explicação para as ligações entre identidade social, estereótipos sociais e língua (KRISTIANSEN, 2001).

Squires (2016) sugere que o nível de consciência das categorias de estereótipos, marcadores ou indicadores pode ser expresso como o nível de conhecimento que os falantes têm sobre a variação. A proposta de Squires (2016) é que a consciência sociolinguística vai além do conhecimento implícito vs. explícito: com estereótipos, os falantes sabem que o traço variável está relacionado a um tipo de falante; com indicadores, os falantes não sabem a que tipo de falante o traço está relacionado.

A escola propicia espaço para reflexão sobre o uso da língua; a consciência sociolinguística é medida em tarefas que levem o aluno a reconhecer e identificar diferentes variedades de uma língua, ouvir exemplos de fala com distinção entre formal e informal em diferentes situações comunicativas, como casamento, reunião, aula de culinária, etc. (PIPER, 2003; GÖRSKI; FREITAG, 2013).

No escopo de L2, Van Compernelle e Williams (2013) definem consciência sociolinguística como a consciência ou conhecimento de formas variáveis (por exemplo, variáveis lexicais, gramaticais, fonológicas) e do seu significado social e/ou estilístico em relação a contextos de uso, relações sociais e identidades pessoais. Os autores ilustram com o exemplo da variação entre a presença e a ausência de *ne* na negação francesa, variável gramatical que carrega significado social e estilístico: a presença de *ne* é geralmente exigida em francês padrão, formal (especialmente escrito), enquanto a ausência é característica de quase todos os falantes europeus e norte-americanos de francês em francês mais informal, cotidiano (especialmente falado).

O desenvolvimento da consciência sociolinguística pressupõe primeiro ter consciência de quais traços da língua são variáveis (e quais são as variantes); e, segundo, ter consciência dos valores socioestilísticos e significados sociais das variantes.

A consciência sociolinguística implica simultaneamente conhecimento linguístico e conhecimento sociocultural: o conhecimento consciente sobre a variação socioestilística pode levar ao uso controlado de construções linguísticas nos diferentes contextos.

Consciência sociolinguística e o desempenho em leitura

A leitura de um sistema de escrita alfabético pode ocorrer por meio de um processo de reconhecimento visual direto (rota lexical) ou por um processo envolvendo mediação fonológica (rota fonológica). A abordagem de leitura que reconhece duas rotas possíveis é conhecida como Modelo de Leitura de Dupla Rota (COLTHEART; RASTLE, 1994). Ambas as rotas de leitura iniciam com o sistema de análise visual, que tem a função de identificar as letras, e sua posição na palavra e agrupá-las. Na rota fonológica, primeiro ocorre a decodificação das letras, depois a procura de uma possível pronúncia e depois o sentido. Esta é a única rota possível para ler palavras novas ou raras na ortografia regular, e neologismos. Já a rota lexical se vale de mapas ortográficos e da representação de palavras visuais que são ativadas na memória (EHRI, 2014). Esta é a rota utilizada para as palavras frequentes e indispensável para as palavras irregulares. Mapas ortográficos formam conexões entre as palavras lidas e padrões de sons, ligando na memória a grafia, a pronúncia e o significado das palavras; palavras visuais são as representações de todas as palavras lidas na memória. Ao ver a palavra, a sua pronúncia e significado são ativados na memória imediatamente, sem qualquer som ou combinação necessária. As palavras visuais são lidas como unidades inteiras sem pausas entre os sons (EHRI, 2005). É na representação visual da palavra que a consciência sociolinguística interage com a leitura: a variedade linguística da comunidade é tomada como referência para a conversão grafofonêmica. Nesse sentido, a presença da variação linguística na leitura não pode ser considerada erro, mas sim pista de acesso à rota lexical e, por conseguinte, da automaticidade na decodificação, essencial para a compreensão da leitura.

No entanto, a relação entre variação linguística e desempenho tem sido explorada no cenário nacional como negativa, nos seguintes termos: a variação linguística na fala do aluno é inversamente proporcional ao desempenho nas tarefas de leitura e escrita. É o que apontam os trabalhos de Barrera e Maluf (2004) e Guimarães (2005). Reconhecendo os efeitos da variação linguística nos processos de aprendizagem de leitura e escrita, Barrera e Maluf (2004) se propõem a “identificar e descrever os diferentes tipos de variação linguística utilizada, no início da alfabetização escolar, por um grupo de alunos provenientes das classes populares; verificar se os alunos que apresentam maior variação linguística no início do ano são os mesmos que enfrentam maior dificuldade na aquisição da linguagem escrita no fim do ano” (BARRERA; MALUF, 2004, p.40).

Assumindo que variação linguística utilizada significa não adotar padrões linguísticos tidos como de prestígio, significa dizer que alunos que fazem uso de variantes estigmatizadas são os que terão maiores dificuldades nas tarefas de leitura e escrita. Para verificar esta hipótese, as autoras acompanharam um grupo de alunos

de uma escola pública municipal, “na periferia da cidade de São Paulo [que] atende predominantemente crianças provenientes de famílias de baixo nível socioeconômico” (BARRERA; MALUF, 2004, p.40).

Os dados para análise da variação linguística foram produzidos por estratégia de narrativa reportada a partir de um estímulo visual, documentados em áudio (procedimento que no presente estudo é denominado de tarefa de nomeação, como descrito na seção de método). Posteriormente, foram identificadas as “formas linguísticas não padrão” (BARRERA; MALUF, 2004, p.40). A soma dos “tipos de variação utilizados” dividida pelo tempo da tarefa foi computada como um índice de variação linguística para cada criança (Figura 1).²

Quadro 1 – Formas linguísticas não padrão

Descrição	Exemplo	Processo
A vogal [o] é pronunciada como [u], no início das palavras	<i>buné, tumati, cumeu.</i>	Alçamento de pretônicas
A terminação “am” [ãu] dos verbos na 3ª pessoa do plural é pronunciada como [o]	<i>eles saíro, fizero, brigaro, ficaro, foro.</i>	Desnasalização/concordância verbal
A terminação nasal [em] é pronunciada como [i]	<i>homi, onti</i>	Desnasalização
Substituição de [lh] por [i]	<i>trabaia, paiaço, oio, oreia, muié, agüia, moiadu, vermeio, veia, véio</i>	Vocalização da aproximante lateral palatal
Substituição de [l] por [r]	<i>armoço, pranta, bicireta, compreto</i>	Rotacismo
Substituição do [r] por [l]	<i>tlem, tliste, tlês</i>	Lambdacismo
Omissão do [d] no grupo consonantal [nd] das formas do gerúndio	<i>correno, falano, chorano, pulano, fazeno, brincano, conversano, pegano</i>	Supressão –d em –ndo
Omissão do [r] em alguns grupos consonantais	<i>tabalha (trabalha) ôto/ôta (outro/outra), quato (quatro)</i>	Simplificação de cluster de ataque complexo

² Além destes traços, as autoras reportam ter controlado os seguintes aspectos: variações lexicais (catá (=pegar); tacá (=jogar, atirar); buchuda (=grávida); tropicá (=tropear); caçá (=procurar); abano (=ventilador); pexera (=faca); coisá (= fazer alguma coisa); mor (= para) e mais (=com)), prócope de um [a] em algumas palavras (gulha/guia, bacaxi, rancá, sustô, marelo, ranhá, cabô, inda), prótese de um [a] em algumas palavras (atropeçô, alembrandu, avoar), construções de frases tidas como atípicas (mor deli tropcá i caí; eu venhu mais o meu amigo; a bruxa tava durminu mais u gatu; mandô eli num mais robá; não se lembriu; aguá as pranta; você ponhava; ela se acordô; meu pai vai demorá de ir; ela vai dá um castigu neli), contrações, aglutinações, abreviações e outras variações relacionadas ao contexto fonético e/ou aos aspectos prosódicos do ritmo geral da fala (ca/cas (com a/as), cu/cus (com o/os), cê (você), ni (em)), e outras variações frequentes (árvi/árvi, tamém, ingual)

Descrição	Exemplo	Processo
Vogais [e] e [o] pronunciadas como [i] e [u] quando ocorrem em posição final não acentuada das palavras	<i>alegri, sorveti trabalhu, cachorru, denti, passarinhu, árvuri</i>	Alçamento em sílaba átona final
Omissão generalizada do [r] nos infinitivos verbais	<i>brincá, passeá, fazê, mordê, durmi, istudá, iscová</i>	Queda do –R em coda (verbo)
Vogais [e] e [o] pronunciadas como [i] e [u] mesmo não estando em posição final nas palavras	<i>mintira, pudia, durmia, cumida, iscola, istuda</i>	Alçamento de pretônica
Redução sistemática dos ditongos [ei], [ai] e [ou]	<i>caxa, dexe, pipoquero, dinhero/u, poquinho, tornera, ôtro, feção,</i>	Monotongação
Expansão de monotongos em ditongos	<i>mais (em lugar de mas), feiz, deiz, faiz, tambeim, ningueim</i>	Ditongação
Ausência de concordância nominal (singular/plural)	<i>as criança, minhas amiga, meus irmão, as menina, muitas letra, as pranta, às vez, duas gota, ficamos duente, as coisa;</i>	Concordância nominal
Ausência de concordância verbal	<i>elis volta, as pessoa gosta, as criança foi, é eu, elis pega, é essas, nós brinca, elis si separô, elis vai, elis fica, os dois é.</i>	Concordância verbal

(Em cinza: traços variáveis observados, mas não somados ao índice de variação linguística)

Fonte: Barrera e Maluf (2004, p.43-44).

Na relação de comportamentos linguísticos controlados, alguns considerados como não padrão não diferem significativamente dos considerados padrão; é o caso da supressão do gerúndio, que, em São Paulo (FERREIRA; TENANI; GONÇALVES, 2012), tem taxa de 76% de aplicação da variante com supressão; também não há característica linguística que justifique distinção entre os alçamentos de “cumeu”, considerado não padrão, para “cumida”, considerado como padrão (Figura 1). O lambdacismo, articulação de /l/ por /r/ em ataque complexo, de aquisição tardia, não pode ser considerada como variação (NICOLOSI; HARRYMAN; KRESHECK, 2004).

Além disso, o índice considera apenas a frequência dos itens não padrão, e não a sua proporção em relação à variável (frequência *type/token*), como é o procedimento da sociolinguística. Ainda assim, as autoras encontraram associação negativa entre o índice de variação linguística e o desempenho nas tarefas de leitura e escrita ($r = -0,27$; $p < 0,05$).

Partindo da mesma hipótese, Guimarães (2005) observa a influência da variação linguística e da consciência morfossintática nas diferenças de desempenho em leitura

e escrita de alunos da 2^a, 3^a e 4^a séries do ensino fundamental, divididos em dois grupos: Grupo 1, 18 alunos (quatro da 3^a série e quatorze da 4^a série) que apresentavam dificuldades em leitura e escrita e frequentavam, além das aulas regulares, salas de recursos; e Grupo 2, 17 alunos (três de 2^a série e quatorze da 3^a série) sem dificuldades em leitura e escrita, que frequentavam apenas as aulas regulares.

A medida de variação linguística adotada no estudo segue o que propuseram Barrera e Maluf (2004); e, dentre as medidas de consciência morfossintática, o que efetivamente está sendo medido é o uso ou não de um traço variável, como na tarefa de uso gerativo de morfemas, que consiste na flexão de formas verbais de acordo com o contexto dado (estímulo), mas o que está sendo aferido é a realização fonética variável de um morfema:

- pronúncia ou omissão do “r” nos infinitivos verbais (p. ex.: pronunciar /passear/ ou /passiá/, /cantar/ ou /cantá/, /comprar/ ou /comprá/) [queda do -R em coda em verbos];
- pronúncia ou omissão do “u” final em verbos no pretérito perfeito, terceira pessoa do singular (p. ex.: pronunciar /dançou/ ou /dançô/, /falou/ ou /falô/, /mostrou/ ou /mostrô/) [monotongação de -ow];
- pronúncia ou omissão do “s” final em verbos no pretérito perfeito, primeira pessoa do plural (p. ex.: pronunciar /comparamos/ ou /compramo/, /perdemos/ ou /perdemo/, /brincamos/ ou /brincamo/);
- pronúncia ou omissão da nasalização do ditongo nasal átono em verbos no pretérito, terceira pessoa do plural (p. ex.: pronunciar /cantaram/ ou /cantaro/, /deram/ ou /deru/). (GUIMARÃES, 2005, p.265).

Ou, ainda na tarefa de correção das violações gramaticais, em que frases supostamente incorretas deveriam ser corrigidas quanto a concordância verbal, emprego de pronomes, flexões de substantivos e adjetivos, nas quais, segundo a autora, há “anomalias morfêmicas” como, por exemplo, “Nós vai ao circo no próximo domingo”, “Marta veste sua casaco”, “Juliana tem lindas olhos”. Enquanto nas duas últimas há efetivamente violação gramatical quanto à concordância do gênero, na primeira frase há variação. Estas inconsistências, por um lado, enfraquecem o valor do resultado do teste quanto ao reconhecimento morfossintático, mas, por outro, podem fornecer pistas da consciência sociolinguística.

Quadro 2 – Formas linguísticas não padrão e taxas de aplicação

Categorias	Tipos de variação linguística	Variações linguísticas identificadas no G1		Variações linguísticas identificadas no G2	
		F	%	F	%
Pronúncia das vogais [e] e [o] como [i] e [u]	<i>alegri, mininu etc.</i>	476	40,64%	226	33,38%
Omissão do [r] nos infinitivos verbais	<i>brincá, passeá etc.</i>	94	8,02%	65	9,60%
Omissão do [u] final (verbos no pretérito perfeito, 3ª pessoa singular)	<i>falô, mostrô etc.</i>	198	16,91%	171	25,26%
Omissão da nasalização do ditongo nasal átono (verbos no pretérito, 3ª pessoa plural)	<i>sairo, fizero, brigaro, ficaro/u, pegaro etc.</i>	5	0,43%	-	-
Redução do [n] do grupo consonantal [nd] no gerúndio	<i>pensano, chorano, sorrinu, correnu etc.</i>	62	5,29%	14	2,07%
Redução dos ditongos [ei], [ai], e [ou]	<i>caxa, dexta, ôtro etc.</i>	20	1,71%	21	3,10%
Expansão de monotongos em ditongos	<i>feiz, deiz, tambeim etc.</i>	57	4,87%	59	8,71%
Contrações, aglutinações e outras variações relacionadas aos aspectos prosódicos do ritmo geral da fala	<i>ca/cas (com a/as), cê (você), ni (em) etc.</i>	152	12,98%	85	12,56%
Prótese de um [a] em formas verbais	<i>atropeço, avoar etc.</i>	5	0,43%	7	1,03%
Prócope de um [a] em algumas palavras	<i>dimirado, sustô etc.</i>	3	0,26%	3	0,44%
Substituição de [lh] por [i]	<i>óia, trabaia, muiê etc.</i>	4	0,34%	2	0,30%
Substituição de [l] por [r]	<i>craru, bicicleta etc.</i>	-	-	1	0,15%
Terminação nasal [em] pronunciada como [i]	<i>homi, ontí etc.</i>	24	2,05%	-	-
Omissão de [r] em grupos consonantais	<i>pá (pra), ôto (outro) etc.</i>	24	2,05%	9	1,33%
Ausência de concordância nominal	<i>dois gol, os aluno etc.</i>	8	0,68%	2	0,30%
Ausência de concordância verbal	<i>as pessoa gosta etc.</i>	3	0,26%	3	0,44%
Variações lexicais (de vocabulário)	<i>tacá (jogar), coisá etc.</i>	11	0,94%	3	0,44%
Variação envolvendo flexão de adjetivo	<i>mais melhor etc.</i>	1	0,09%	-	-
Outras variações	<i>ingual, tamém etc.</i>	24	2,05%	6	0,89%
TOTAL		1171	100,00%	677	100,00%

Fonte: Guimarães (2005, p.267).

Os alunos do Grupo 1 apresentaram um índice de variação linguística superior aos do Grupo 2 (quadro 2), no entanto, a hipótese de que os alunos com menores índices de variação linguística apresentariam, inversamente, escores mais elevados nas tarefas que avaliavam a consciência morfossintática não foi confirmada ($r = -0,29$; $p = 0,089$). Guimarães (2005) detalha os tipos de variação por frequência, e conclui que três quartos do total de ocorrências de variante não padrão são relativas a fenômenos não

estigmatizados e recorrentes na fala de todos os brasileiros, como o alçamento vocálico, apagamento de -R em coda externa (verbos), desnasalização em verbos, apagamento de -d no gerúndio, monotongação e ditongação. Parece haver efeito da opacidade do sistema ortográfico, mas que, diferentemente do que explica Guimarães (2005), não seria sentida apenas pelos falantes de variantes dialetais afastadas da norma padrão, mas sim por todos os falantes.

O exame dos estudos que consideram a relação entre variação linguística e o desempenho em tarefas de leitura e escrita no português brasileiro evidencia a necessidade de refinamento do controle do índice de variação linguística quanto à:

- i) recorrência das variantes (o que é padrão e o que não é padrão);
- ii) avaliação social das variantes (estigmatizadas vs. não estigmatizadas, percebidas vs. não percebidas, etc.) e;
- iii) sensibilidade à frequência das variantes (a ocorrência única de um traço variável vs. uso sistemático).

Considerando estes aspectos, a relação entre a consciência sociolinguística, medida pela sensibilidade da taxa de frequência de uma variante quanto ao contexto estilístico, e o desempenho em tarefas de leitura deve ser positiva: quanto maior a consciência sociolinguística, melhor o desempenho nos testes. A mudança da direção da hipótese em relação aos estudos anteriores se deve à mudança de foco: enquanto os estudos de Barrera e Maluf (2004) e Guimarães (2005) consideraram a taxa de variação linguística em uma atividade controlada (narrativas a partir de figuras) e a compararam com o desempenho em testes de leitura, este estudo considera a taxa de variação linguística na fala, em atividade oral controlada e na leitura em voz alta e o desempenho em teste de compreensão em leitura. A realização da variante produtiva na comunidade na leitura em voz alta é uma pista de acesso à rota lexical, e, para a avaliação da proficiência em leitura, sinaliza um perfil de leitor proficiente, que tem memória visual da palavra (EHRI, 2005). A transposição dos traços variáveis da fala para a leitura não afeta negativamente a leitura, ao contrário; pode dar pistas acerca da proficiência em leitura (MACHADO, 2018; MACHADO; FREITAG, 2019; FREITAG, 2020).

Pistas da consciência sociolinguística

Os efeitos da consciência da variação linguística podem ser observados mais diretamente na leitura, especialmente no que diz respeito ao contraste entre as relações grafonômicas não transparentes, ou seja, aquelas que não têm correspondência biunívoca. Embora o português brasileiro seja considerada uma língua semi-transparente do ponto de vista grafonômico, em contextos de variação a assimetria se acentua e fomenta questões relacionadas à correção da pronúncia, como: “*cenoura*” ou “*cenora*”?

O prescritivismo dos instrumentos normativos (como as gramáticas) abarca realizações variáveis na ortoepia, mas para itens específicos e não previsíveis por regras ou princípios do sistema alfabético da língua. Embora se assuma que a conversão grafofonêmica selecione a variedade da comunidade, orientações prescritivistas direcionam a escolha de variantes, em alguns momentos selecionando uma variedade dialetal e convertendo-a em um traço generalizado, e, em outros momentos, evidenciando um padrão colado na tradição escrita. É o que faz Bechara em relação ao processo de ditongação:

Quanto aos ditongos, cumpre notar: *ai*, *ei*, e *ou*, na pronúncia, devem guardar, na pronúncia cultivada, sua integridade, não se exagerando o valor do *i* ou do *u*, nem os eliminando, como o faz a modalidade distensa: caixa, queijo, ouro. [...] Normalmente ditongamos, pelo acréscimo de um *i*, as vogais tônicas finais seguidas de um *-z* ou *-s*. Assim não fazemos a diferença entre *pás*, *pais* e *paz*; mas e mais” (BECHARA, 2009, p.60).

Para os ditongos crescentes, o processo de conversão grafofonêmica abonado na prescrição normativa é transparente em relação ao sistema ortográfico (e a palavra “cenoura” deve ser pronunciada com ditongo, embora na fala vernacular do Brasil todo a monotongação seja o padrão); já para contextos de sílaba tônica final terminada por *-s/-z*, a conversão grafofonêmica não é transparente, há a inserção de um ditongo que é produtivo e categórico na fala do Rio de Janeiro, mas não na fala de Santa Catarina ou Rio Grande do Sul, por exemplo (LEIRIA, 2000; CALLOU; LEITE; MORAES, 2012).

Por conta da variabilidade, as pistas de consciência sociolinguística a serem controladas neste estudo foram selecionadas considerando os traços do sistema linguístico do português brasileiro que já estão amplamente documentados do ponto de vista sociolinguístico, de modo a permitir estimar seu comportamento padrão, sua recorrência dialetal e social, e as restrições contextuais³, como as tendências identificadas no Rio de Janeiro, com dados do PEUL (PAIVA; SCHERRE, 1999; PAIVA; PAREDES-SILVA, 2012): monotongação dos ditongos decrescentes *-ej* e *-ow*; estigmatização da alternância entre as líquidas ou a queda de *-R* nos grupos consonantais; estabilidade da supressão de */d/* na sequência *-ndo*; implementação da queda de *-R* pós-vocálico, em especial nos infinitivos.

O processo de monotongação consiste no apagamento do glide palatal [j], como em *caixa*, ou velar [w] em ditongo decrescente, como em *cenoura*. No português brasileiro, este fenômeno apresenta condicionamentos distintos a depender da natureza do glide e do contexto fonológico seguinte: o apagamento do glide velar tende a ser uma regra categórica em situações informais, em todos os contextos linguísticos, inclusive na escrita (ARAUJO; BORGES, 2019; GONÇALVES; AMARAL, 2014; TOLEDO,

³ O processo de monotongação, por exemplo, ocorre de modo distinto em *-ow* e em *-aj*, *-ej*, assim como a queda do *-R* final é diferente em nomes e verbos, dentre outros aspectos (MOLLICA, 2001).

2013; HAUPT; SEARA, 2012; CRISTOFOLINI, 2011). Variantes não salientes, como a realização monotongada, não apresentam diferenças de julgamento que interferem no teste de lexicalidade (FREITAG; SOUZA, 2019). O glide palatal apresenta restrições de natureza interna, decorrente do contexto fonológico seguinte, é regra semicategórica, independentemente da formalidade, para contextos em que a sílaba seguinte apresenta o traço palatal, como em *caixa*, *beijo*, e é restringido por contextos em que a sílaba seguinte é iniciada por oclusiva, como em *leito*, *caibo*. O comportamento é estável em todas as regiões, sem sensibilidade social ou dialetal.

Em direção contrária, ocorre a ditongação palatal em sílaba tônica travada por fricativas alveolares/pós-alveolares (NEVINS, 2012; CALLOU; LEITE; MORAES, 2012). O processo, abordado inclusive na prescrição normativa, tem restrições dialetais, mas não sociais.

O apagamento do -R em coda externa em verbos é uma regra categórica na fala do português brasileiro, mesmo em situações de maior formalidade e monitoramento, sem estar associado a diferenças sociais. Já o apagamento em contextos monomorfêmicos e em posição medial na palavra apresenta restrições dialetais e sociais: a região Nordeste se diferencia do Sudeste e do Sul por conta do apagamento do -R em não verbos, mas, nessa região, o apagamento de -R em posição medial apresenta diferenças sociais (CALLOU; MORAES; LEITE, 1998; MONARETTO, 2000; CALLOU; SERRA; CUNHA, 2015).

O apagamento do -d em segmentos -ndo tem recorrência dialetal e social: enquanto em algumas regiões, como o interior de São Paulo, é restrito aos verbos (FERREIRA; TENANI; GONÇALVES, 2012), em outras, como Rio de Janeiro e o Nordeste, atinge o sistema como um todo (MOLLICA; MATTOS, 1992; NASCIMENTO; ARAÚJO; CARVALHO, 2013; FREITAG; CARDOSO; PINHEIRO, 2018) e, ainda, em outras regiões, sequer é noticiado, como no sul do Brasil. O comportamento é associado à baixa escolarização ou à menor formalidade; e é restringido em situações de maior formalidade, como na leitura em voz alta (CARDOSO; PINHEIRO; SILVA, 2019).










A simplificação de clusters é uma etapa de um processo de leniência que pode resultar na eliminação do segundo segmento em ataques complexos CCV formados por uma obstruinte e uma líquida (CRISTÓFARO-SILVA, 2002). No português brasileiro, ocorre com o tepe, como em *pró/pr/ío* > *pró/p_ío*, *li/vr/o* > *li/v_ó*. O contexto tautossilábico tem aquisição tardia no português; mas, neste caso, a não produção de sílabas CCV é compensada com o alongamento (MIRANDA; CRISTÓFARO-SILVA, 2011). Ao ingressar no ensino fundamental, o aluno em situação típica já adquiriu o traço. A persistência da realização simplificada do cluster é um processo variável com recorrência associada a perfis sociais específicos, e alvo de estigma social (MOLLICA; PAIVA, 2011).

O contexto ortográfico de “lh” tem realização variável: pode ser realizado por uma aproximante lateral palatal /ʎ/, lateral alveolar palatalizada /ʎ̥/ ou glide vocalizado /y/. A realização vocalizada é social e dialetalmente delimitada, associada à baixa

escolarização e região rural (BRANDÃO, 2007). Está presente no espaço escolar, transposto para a escrita (COAN; FREITAG, 2010; FREITAG, 2011).

Junto com R retroflexo, o rotacismo é, possivelmente, o traço sociolinguístico mais estereotipado no português brasileiro: o cancelamento da distinção entre a lateral e o tepe em ataques complexos, como em *f/lanela ~ f/r/anela*, ou coda silábica, como em *a/l/moço ~ a/r/moço*, é objeto de manipulação consciente de caricaturas de “caipira”, como os personagens Jeca Tatu e Chico Bento (COAN; FREITAG, 2010; COSTA, 2013; FREITAG *et al.*, 2020). Este traço é um marcador sociodialeto do português brasileiro, recorrente na fala de regiões distantes dos grandes centros urbanos, de pessoas menos escolarizadas e mais velhas.

Quadro 3 – Distribuição dos traços controlados quanto à produtividade e recorrência

Processo	Variantes	Variante produtiva na comunidade	Escopo da recorrência		Variante padrão
			dialeto	social	
Monotongação de -ow		cen/ow/ra			
		cen/o_/ra	√		√
Apagamento de -R em coda externa		cant/aR/			
		cant/a_/	√		√
Ditongação (sílabas tônicas final -S)		arr/oz/		√	
		arr/ojz/	√		√
Monotongação de -aj, -ej		c/aj/xa			
		c/a_/xa	√		√
Apagamento de -R em coda interna		so/R/vete	√	√	√
		so/_/vete			
Apagamento de d- em -ndo		canta/ndo/		√	√
		canta/_no/	√		
Simplificação (ataque complexo)		pe/dra/	√		√
		pe/d_a/		√	
Vocalização de lateral palatal		pa/la/ço	√	√	√
		pa/ya/ço			
Rotacismo		f/l/or	√	√	√
		f/r/or			

Fonte: Elaboração própria.

A observação da leitura em voz alta pode permitir a identificação do juízo de valor atribuídos aos traços linguísticos variáveis, contribuindo para a construção científica da consciência sociolinguística (Figura 3). Por outro lado, a presença de traços variáveis na leitura é uma pista do acesso à rota lexical e, por conseguinte, da automaticidade na decodificação, essencial para a compreensão, que pode ser verificada com a taxa de sucesso em uma avaliação de leitura.

Método

Os procedimentos de observação e constituição da amostra estão detalhadamente descritos em Freitag e Sá (2019), cujo objetivo foi identificar perfil de leitores de alunos de uma turma de 3º ano do ensino fundamental de uma escola pública de Sergipe; aqui são sumarizados os procedimentos necessários para a descrição sociolinguística. Participaram da pesquisa 21 estudantes, com idade entre 7 e 11 anos, todos alunos regulares de uma turma de 3º ano em 2017, em quatro tipos de coletas de dados: entrevistas curtas; tarefas de nomeação; leitura em voz alta; e teste de compreensão em leitura. Os primeiros três tipos de coleta correspondem a um contexto estilístico de maior ou menor monitoramento da fala: entrevista curta (menos formal) → atividade oral (nomeação de figuras, intermediário) → leitura em voz alta (mais formal). Tanto os textos para a leitura em voz alta como as figuras para a nomeação foram escolhidos em função das palavras que também poderiam possibilitar a ocorrência dos mesmos fenômenos variáveis. O teste de compreensão em leitura (TCL) semelhante à ANA de língua portuguesa, disponibilizado pela Editora Moderna (2016), ao contrário das coletas anteriores, teve aplicação coletiva a 18 estudantes⁴. Toda a documentação linguística das coletas orais foi gravada individualmente, em ambiente escolar, e submetida a transcrição ortográfica com áudio alinhado no software Elan (WITTENBURG *et al.*, 2006).

Os contextos possíveis de ocorrências dos fenômenos variáveis nas entrevistas curtas, tarefas de nomeação e na leitura em voz alta foram identificados e controlados quanto à taxa de frequência da variante padrão na comunidade por aluno e análise univariada por fenômeno variável. Os acertos no TCL foram computados individualmente e analisados em um modelo de regressão linear quanto à taxa de frequência da variante padrão na comunidade nos extremos do contínuo de monitoramento estilístico (leitura em voz alta e entrevistas curtas), e posterior meta-análise. A visualização gráfica dos resultados foi desenvolvida com o pacote *ggstatsplot* (PATIL; POWELL, 2018) para a plataforma R.

Resultados e discussão

Para explicar os resultados, foram considerados três aspectos: há associação entre a ocorrência da variante e o contínuo estilístico? (observação do percentual e o resultado

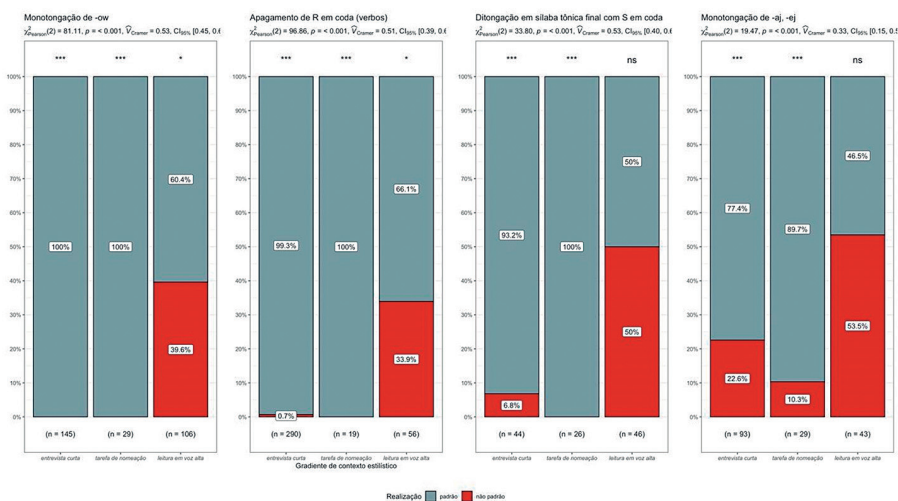
⁴ A Avaliação Nacional de Alfabetização (ANA) é uma avaliação em larga escala censitária aplicada a cada dois anos pelo INEP/MEC a estudantes de escolas da rede pública que estejam cursando o 2º ano do ensino fundamental. A matriz de referência de língua portuguesa é composta por dois eixos estruturantes (leitura e escrita) e 12 habilidades. As provas não são disponibilizadas, nem há itens comentados no portal Devolutivas (onde podem ser encontrados itens da Prova Brasil, por exemplo). Testes similares à ANA, considerando as habilidades da matriz de referência, são desenvolvidos com o intuito de preparar e pré-avaliar os estudantes. O TCL adotado neste estudo é composto por 17 itens de múltipla escolha e 3 de produção escrita. Todas as habilidades do eixo de leitura da matriz de referência estão contempladas nas questões do teste. Com exceção da habilidade 1, avaliada somente no item 1, as demais habilidades são avaliadas em dois itens. As questões foram pontuadas com escores atribuídos conforme a matriz de referência. Detalhes do estudo completo podem ser conferidos em Freitag e Sá (2019).

do teste de qui-quadrado de Pearson χ^2); o quão forte é essa associação? (coeficiente V de Cramer) e qual é a tendência de comportamento? (direção dos percentuais). Considerando estes aspectos, as taxas de ocorrência da variante assumida como padrão na comunidade apresentam quatro comportamentos distintos:

- ocorrência categórica da variante padrão nos contextos estilísticos menos formais e ocorrência variável da variante não padrão na situação mais formal (Figura 1);
- baixa recorrência da variante padrão nos contextos estilísticos menos formais e alta recorrência na situação mais formal (Figura 2);
- alta recorrência da variante padrão nos contextos estilísticos menos formais e baixa recorrência na situação mais formal (Figura 3);
- estabilidade em todas as situações estilísticas (Figura 4).

Nas situações menos formais, a realização monotongada de -ow, o apagamento do -R em coda silábica externa (verbos) e a ditongação em sílaba tônica final travada por -S se manifestam como regra categórica para os dois primeiros e semi-categórica para o último fenômeno. A variante assumida como não padrão na comunidade ocorre somente na leitura em voz alta, o que sugere que não seja necessariamente um efeito do monitoramento estilístico, mas um reflexo da transparência na conversão grafofonêmica que sinaliza o acionamento da rota fonológica e não da rota lexical. Por não serem marcados socialmente, os traços não seriam filtrados pela consciência sociolinguística na situação de maior formalidade; por conseguinte, espera-se que o leitor proficiente já tenha na memória visual das palavras em que ocorre este traço a informação de que a relação grafofonêmica não é transparente: escreve-se “cenoura” mas se lê “cenora”.

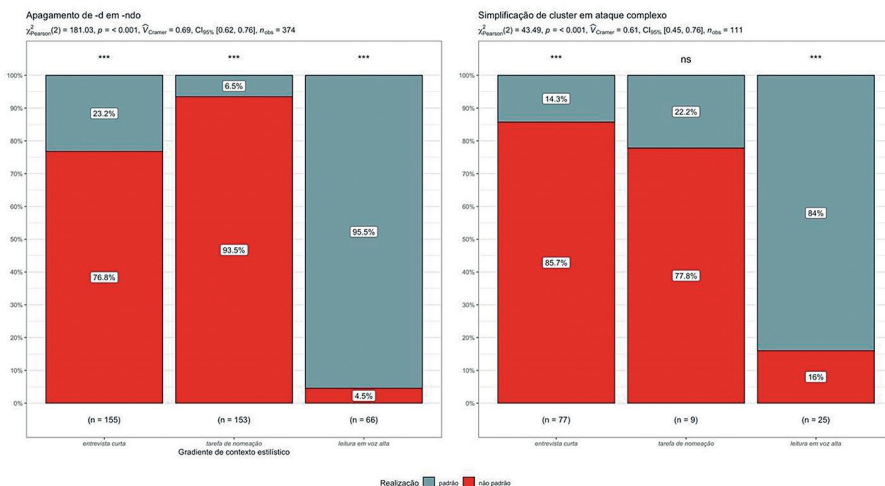
Figura 1 – Distribuição dos traços com predomínio da variante padrão na fala



Fonte: Elaboração própria.

Na leitura em voz alta, estes traços têm associação fraca (caso da monotongação de -ow) ou não significativa (-R em verbos e ditongação). O comportamento destas variáveis poderia sugerir uma distribuição de marcadores, por conta da sua distribuição no contínuo de monitoramento estilístico. No entanto, por ser uma amostra constituída por leitores em aprendizado inicial da leitura, talvez a consciência sociolinguística ainda não tenha sido despertada. A conservação da variante assumida como não padrão, a partir dos efeitos de frequência, pode sinalizar para a convergência de efeitos de orientações prescritivistas e a transparência do sistema, o que explicaria a não monotongação e a conservação do -R em coda nos verbos, mas não a ditongação, que tem uma orientação prescritivista que vai ao encontro da realização assumida como padrão (cf. BECHARA, 2009). Por outro lado, a consciência morfológica poderia explicar a conservação do -R nos verbos, como sugerem os resultados de Guimarães (2005). Uma investigação em leitores maduros (em níveis de escolarização mais avançado) poderia dar pistas mais consistentes: se o comportamento obtido se mantém, há efeito de monitoramento estilístico (e, neste caso, os traços teriam comportamento de marcadores), se o comportamento na leitura se iguala ao de entrevistas e tarefas de nomeação, há efeito do desenvolvimento da consciência sociolinguística (e, neste caso, os traços teriam comportamento categórico, não mais uma regra variável).

Figura 2 – Distribuição dos traços com predomínio da variante não padrão na fala



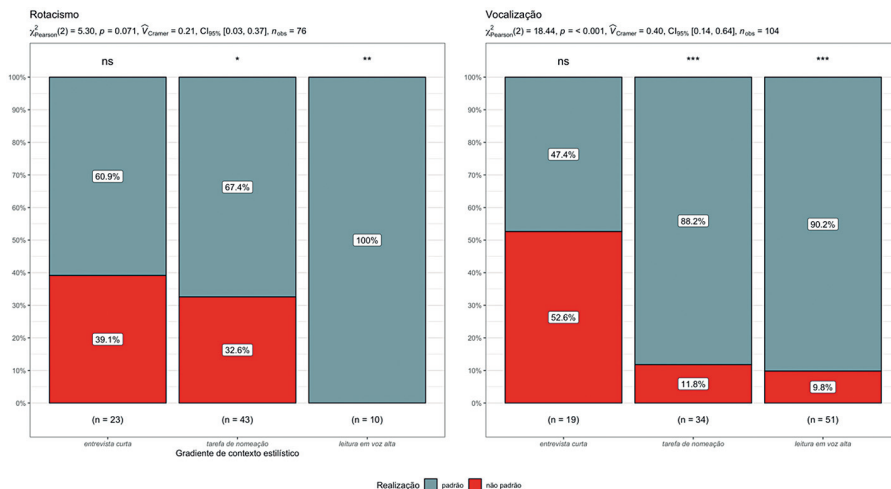
Fonte: Elaboração própria.

O apagamento do -d em segmentos -ndo e a simplificação de clusters com ataque complexo são fenômenos com alta recorrência nas situações menos formais e baixa recorrência na situação mais formal. A distribuição sugere efeito da transparência do sistema ortográfico, mas em direção diferente daquela das variáveis anteriores: a

memória visual das palavras com estes segmentos contempla duas realizações, a da fala, em que predomina a variante com apagamento (-d ou -R), e a da leitura, com a realização do segmento. Diferentemente das variáveis anteriores, a associação entre as variantes é forte e significativa em todos os contextos de monitoramento estilístico. Este resultado reforça as pistas de estudos sociolinguísticos de produção de que, quanto à avaliação social, os traços apresentam comportamento de marcadores: sensíveis ao monitoramento estilístico.

A vocalização da lateral palatal é um estereótipo no português brasileiro (MADUREIRA, 1999; CYRANKA; RONCARATI, 2010) e esta percepção de estereótipo parece ser captada pela consciência sociolinguística dos estudantes da amostra. Embora seja o traço sociolinguístico com associação mais fraca dentre os oito controlados, o comportamento das frequências da variante padrão é sensível ao contexto estilístico, com recorrência decrescente em função do aumento da formalidade, mas ainda com realizações na leitura. Isso significa que a memória visual das palavras com este segmento contempla duas realizações, a padrão e a não padrão, e a realização não padrão sofre efeitos de monitoramento social (evidenciada pela recorrência reduzida nas situações de oralidade, seja na entrevista curta, seja na tarefa de nomeação, e pela restrição na leitura em voz alta). O comportamento deste traço é diferente do comportamento dos traços reconhecidos como marcadores, porque, além da sensibilidade estilística, que está presente nos marcadores, há uma restrição à ocorrência na fala. Por ser manipulado conscientemente, alvo de metacomentários avaliativos, cujo exemplo mais emblemático é o poema “Pronominais”, de Oswald de Andrade, este traço é considerado um estereótipo (COAN; FREITAG, 2010). Mas, muito mais forte do que a vocalização, é o rotacismo: na amostra, não foram encontradas realizações da variante não padrão na situação de leitura em voz alta. Este resultado, ao mesmo tempo que corrobora os estudos de produção, aponta para o desenvolvimento da consciência sociolinguística denotada pela sensibilidade ao contexto estilístico e à seleção/barramento de traços.

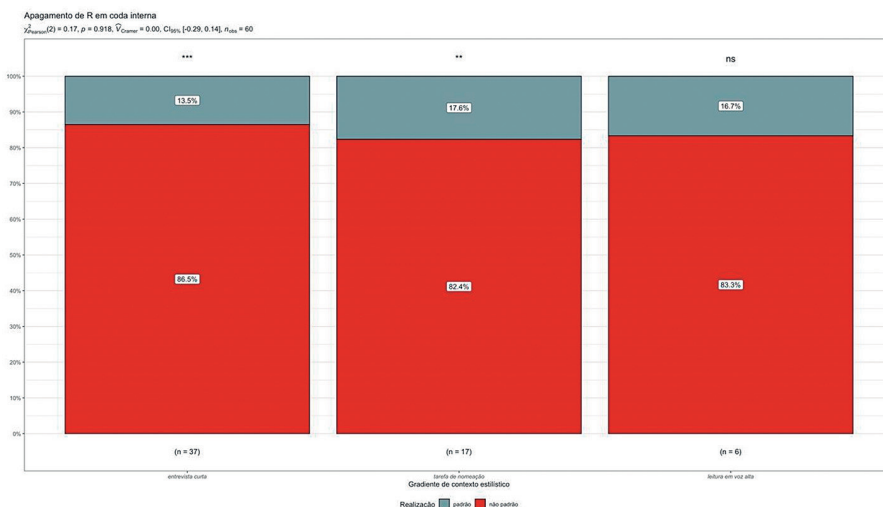
Figura 3 – Distribuição dos traços com comportamento de estereótipo



Fonte: Elaboração própria.

O comportamento do apagamento do -R em coda como tendência geral aponta para a não sensibilidade às situações estilísticas, com taxa de realização relativamente constante no caso do -R apagamento em coda interna, um comportamento de indicador, o que requer uma análise com outros perfis de leitor para dirimir o efeito do monitoramento ou do desenvolvimento da consciência sociolinguística.

Figura 4 – Distribuição de traços com comportamento de indicador












Fonte: Elaboração própria.

Considerando o comportamento dos traços controlados na leitura em voz alta, os resultados apontam que a consciência sociolinguística contribui para o monitoramento da ocorrência da variante não padrão quanto ao apagamento de d- em -ndo e simplificação do ataque complexo, que são traços sensíveis ao contexto estilístico (marcadores); o apagamento de -R em nomes não parece ser sensível ao monitoramento estilístico (indicador); e a vocalização da lateral palatal é altamente sensível, com distribuição restrita nas situações de menor monitoramento e forte restrição na leitura.

Em relação aos demais contextos, por conta de sua proximidade com o prescritivismo atrelado à transparência do sistema ortográfico, resultados com outros perfis de escolarização são necessários para mais segurança no delineamento de uma avaliação social.

Figura 5 – Distribuição dos traços controlados quanto à produtividade e recorrência

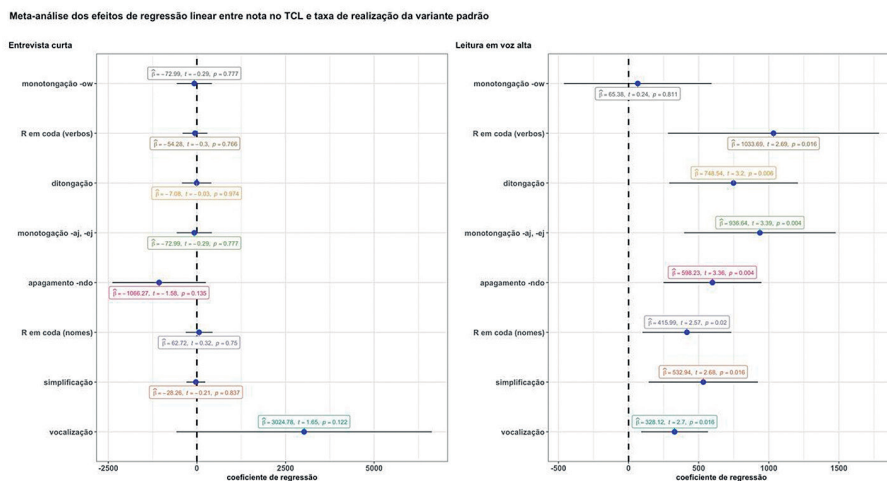
Processo	Variantes		Variante produtiva na comunidade		Escopo da recorrência		Avaliação social
			fala	leitura	dialetal	social	
Monotongação -ow		cen/ow/ra		√			marcador
		cen/o_/ra	√				
Apagamento de -R em coda externa		cant/ar/		√			
		cant/a_/	√				
Ditongação		arr/oz/		√	√		
		arr/ojz/	√				
Monotongação -aj, -ej		c/aj/xa		√			
		c/a_/xa	√				
Apagamento de d- em -ndo		canta/ndo/		√	√	√	
		canta/_no/	√				
Simplificação de ataque complexo		pe/dra/	√	√		√	
		pe/d_a/					
Apagamento de -R em coda interna		so/r/vete	√	√	√		
		so/_/vete					
Vocalização de lateral palatal		pa/ka/ço	√	√	√	√	
		pa/ya/ço					
Rotacismo		f/l/or		√	√	√	
		f/r/or					

Fonte: Elaboração própria.

A ocorrência dos traços sociolinguísticos fornece pistas da consciência sociolinguística, que têm reflexos na leitura por conta da ativação das rotas: a ocorrência do traço tido como padrão na leitura em voz alta sinaliza o acionamento de rota lexical, com maior precisão e rapidez do que o leitor que faz uso da rota fonológica. O leitor com maior precisão e rapidez na decodificação pode despende maior esforço cognitivo no processo de compreensão, e assim, ter melhor desempenho em testes de compreensão

em leitura, como são as avaliações oficiais e o TCL aplicado neste estudo. Assim, a ocorrência dos traços padrão na leitura em voz alta pode ser um fator preditivo do desempenho em um teste de compreensão em leitura. Para testar esta hipótese, para cada traço sociolinguístico foi computada a taxa de frequência da variante padrão por estudante e por situação estilística menos formal (entrevistas curtas) e mais formal (tarefa de leitura em voz alta).

Figura 6 – Modelo de regressão dos efeitos da ocorrência dos traços sociolinguísticos e o desempenho no teste de compreensão em leitura



Fonte: Elaboração própria.

Os modelos lineares para prever o desempenho no teste de compreensão em leitura e a taxa de aplicação da variante padrão de cada um dos traços controlados nas entrevistas curtas não foram estatisticamente significativos (Figura 8, à esquerda). Já os modelos lineares para prever o desempenho no teste de compreensão em leitura e a taxa de aplicação da variante padrão de cada um dos traços controlados na leitura em voz alta (Figura 5, à direita), à exceção da monotongação de -ow, se mostraram significativos, com grande efeito positivo do uso da variante padrão na pontuação do TCL, como o apagamento de -d em -ndo ($\beta = 598.23$, $p < .01$), substancial, como a ditongação, ($\beta = 748.54$, $p < .01$) e a monotongação de -aj, -ej ($\beta = 936.64$, $p < .01$), e médio, como a vocalização de lateral palatal ($\beta = 328.12$, $p < .05$) e o apagamento do -R em coda interna ($\beta = 1033.69$, $p < .05$), em verbos ($\beta = 415.99$, $p < .05$), e em clusters de ataque complexo ($\beta = 532.94$, $p < .05$).

A metanálise é uma forte evidência de que a consciência sociolinguística, medida pela taxa de ocorrência da variante padrão na comunidade na situação estilística de tarefa de leitura em voz alta, tem um forte efeito no sucesso na leitura, medida pelo escore obtido no teste de compreensão em leitura aplicado ao mesmo conjunto de alunos.

Em termos gerais, este resultado quer dizer que ouvir o aluno ler em voz alta pode ajudar a identificar a capacidade de compreensão em leitura. Parece uma medida simples, que não requer recursos específicos, basta uma formação inicial para o professor das séries iniciais ser sensível à variação sociolinguística. A recomendação de formação inicial sociolinguisticamente sensível é proposta também por Barrera e Maluf (2004) e Guimarães (2005), no entanto, a realidade pedagógica brasileira ainda não incorpora a dimensão do tratamento variável da língua nas séries iniciais (FREITAG; SÁ, 2019), talvez por falta de detalhamento mais voltado para a prática pedagógica dos estudos sociolinguísticos descritivos. As experiências em L2 podem ser um ponto de partida: Van Compernelle e Williams (2013) defendem que a consciência sociolinguística pode ser desenvolvida por instrução; o uso controlado da variação pode se desenvolver à medida que os alunos internalizam seu conhecimento sociolinguístico como uma ferramenta cognitiva. As recomendações pedagógicas para L2 (mas podemos estendê-las às séries iniciais de L1) envolvem a exposição dos aprendizes a variantes linguísticas socioestilisticamente diferenciadas, a explicação dos seus valores socioestilísticos e a oportunização de situações aos alunos para utilizarem as variantes em tarefas comunicativas.

Schwalm (1985) argumenta que o desenvolvimento da consciência sociolinguística é muito mais um problema para falantes de uma L2 do que para falantes nativos da língua, uma vez que o primeiro grupo demanda muito mais ajustes socioculturais em relação ao uso da língua. No entanto, o processo de transição para o espaço escolar, especialmente em situações em que a variedade linguística da comunidade é mais distante do que se assume como variedade de prestígio ou padrão, e pode ser uma experiência que demande a sensibilização dos estudantes, afinal, a escola é uma cultura diferente do cotidiano dos alunos, e para terem sucesso, precisam aprender os valores e os traços linguísticos variáveis que são característicos dessa cultura. Pendergast e Kaplan (2015) reforçam a importância de maior preparo dos professores para observar efeitos culturais e linguísticos nas práticas de leitura em voz alta nas séries iniciais. A adoção destas medidas pode contribuir para a redução da assimetria nos resultados das avaliações de leitura.

Conclusão

A investigação do comportamento de um mesmo traço em diferentes situações de monitoramento estilístico pelos mesmos falantes permite tecer considerações para o refinamento da escala de avaliação social das variáveis sociolinguísticas, e, por ter sido desenvolvida no espaço escolar e envolver como parâmetro um teste de compreensão leitora, também permite tecer considerações acerca do desenvolvimento da consciência sociolinguística e o desempenho em leitura.

Esteretótipos negativos, como a vocalização e o rotacismo, ainda que presentes na fala, não transpassam para as situações de maior monitoramento estilístico, como

a tarefa de nomeação e a leitura. Variáveis com comportamento de marcador, como o apagamento de -d em -ndo e a simplificação de ataques complexos, também foram barradas na leitura em voz alta.

A monotongação, a ditongação e o apagamento do -R em coda externa (especialmente verbos), traços que são tradicionalmente considerados como indicadores, neste estudo, apresentaram comportamento de marcadores, por serem sensíveis ao contexto de monitoramento estilístico, mas em uma direção contrária ao esperado: a variante não padrão é mais recorrente no contexto de maior monitoramento, a leitura em voz alta. Neste caso, o efeito pode ser atribuído à falta de transparência entre a forma ortográfica e a realização padrão na comunidade; por estarem em situação de aprendizado inicial da leitura, os estudantes ainda acessam a rota fonológica, ao invés da rota lexical.

Os resultados do estudo mostram que os alunos que apresentaram as maiores taxas de realização de variáveis padrão na tarefa de leitura em voz alta foram os que tiveram melhor desempenho em um teste de compreensão em leitura. A transposição da variante padrão da fala espontânea para a leitura em voz alta sugere a automatização no processo de decodificação por via lexical, e consequente disponibilidade cognitiva para efeitos de processamento.

Pode parecer um resultado óbvio, mas uma observação empírica, como o desempenho nas avaliações oficiais, mostra que a realidade é muito distante. Reverter essa situação requer, por um lado, que resultados descritivos de estudos sociolinguísticos sejam mobilizados para a formação de professores de língua materna, de modo a contribuir para a identificação de traços variáveis da fala que passam para a leitura e seu respectivo valor social; e, por outro, a observação dos traços variáveis da fala que passam para a leitura pode contribuir para o refinamento das descrições sociolinguísticas no que diz respeito ao comportamento de variáveis do tipo marcadores, estereótipos e indicadores: enquanto estereótipos negativos têm padrão de comportamento estável e demarcado, marcadores parecem atuar em duas direções: aquele em que a variante padrão do traço é recorrente na fala, mas não passa para a leitura (emergindo uma variante não padrão), e aquele em que a variante não padrão do traço é recorrente na fala, e também não passa para a leitura (este o comportamento mais prototípico associado aos estereótipos). A diversificação dos contextos estilísticos para a observação da variação linguística tem ganhado espaço na agenda de pesquisa atual (cf. VIEIRA; LIMA, 2019), com resultados que particularizam os padrões de comportamento em diferentes gêneros e tipos textuais. A inclusão de tarefas de leitura e a documentação da produção linguística no espaço escolar para identificar traços variáveis da fala que passam para a leitura e seu respectivo valor social pode contribuir para o desenvolvimento de estratégias de diagnóstico de leitura em situação de sala de aula.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, selecionado pelo edital POSGRAP/COPGD/UFS 06/2020.

FREITAG, R. The development of sociolinguistic awareness and success in reading performance. *Alfa*, São Paulo, v.65, 2021.

- *ABSTRACT: National and international assessments show that Brazil fails to teach reading in its entire school system. This paper argues that the explanation for part of these students' success in early learning of reading might be related to the development of sociolinguistic awareness. The rates of variation in eight phonological variables (-ow monophthongization, -aj, -ej monophthongization, diphthongization, coda -r deletion in nouns and in verbs, -d deletion in -ndo, cluster simplification, and vocalization of palatal lateral) in three different stylistic contexts (from - to + monitored speech: short interview -> naming picture task -> reading aloud task) collected in a 3rd grade classroom show a gradual relation between the increase in stylistic monitoring and the use of standard variant and a positive effect in the performance in a reading comprehension test. These results reveal that the development of sociolinguistic awareness can be measured by the transposition of the variant into spontaneous speech for the reading aloud, suggesting automaticity in the decoding process by lexical route.*
- *KEYWORDS: reading aloud; Sociolinguistic awareness; social evaluation; reading comprehension.*

REFERÊNCIAS

ARAÚJO, A. S.; BORGES, D. K. V. Atitudes linguísticas de estudantes universitários: o fenômeno da monotongação em foco. **Tabuleiro de Letras**, Salvador, v.12, p.97-113, 2019.

BARRERA, S. D.; MALUF, M. R. Variação lingüística e alfabetização: um estudo com crianças da primeira série do ensino fundamental. **Psicologia Escolar e Educacional**, Campinas, v.8, n.1, p.35-46, 2004.

BECHARA, E. **Moderna gramática portuguesa**. Rio de Janeiro: Nova Fronteira, 2009.

BELL, A. Language style as audience design. **Language in society**, Cambridge, v.13, n.2, p.145-204, 1984.

BRANDÃO, S. F. Um estudo variacionista sobre a lateral palatal. **Letras de Hoje**, Porto Alegre, v.42, n.3, p.89-99, 2007.

- CALLOU, D.; LEITE, Y.; MORAES, J. A ditongação no português do Brasil: estudo de dois casos. **Vivacité et diversité de la variation linguistique**, [s.l.], v.3, p.95-102, 2012.
- CALLOU, D.; MORAES, J.; LEITE, Y. Apagamento do R final no dialeto carioca: um estudo em tempo aparente e em tempo real. **DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada**, São Paulo, v.14, nesp., 1998. Disponível em: <https://doi.org/10.1590/S0102-44501998000300006>. Acesso em: 05 ago. 2021
- CALLOU, D.; SERRA, C.; CUNHA, C. Mudança em curso no português brasileiro: o apagamento do R no dialeto nordestino. **Revista da ABRALIN**, [s.l.], v.14, n.1, p.195-219, 2015.
- CARDOSO, P. B.; PINHEIRO, B. F. M.; SILVA, L. S. Variação do segmento /d/ no contexto -ndo: efeitos prosódicos e de leitura. **Leitura**, [s.l.], v.2, n.63, p.174-191, 2019.
- COAN, M.; FREITAG, R. M. K. Sociolinguística variacionista: pressupostos teórico-metodológicos e propostas de ensino. **Domínios de Linguagem**, Uberlândia, v.4, n.2, p.173-194, 2010.
- COLTHEART, M.; RASTLE, K. Serial processing in reading aloud: evidence for dual-route models of reading. **Journal of Experimental Psychology: human perception and performance**, Arlington, v.20, n.6, p.1197-1211, 1994.
- COSTA, L. T. da. Fenômenos variáveis e variantes líquidas produzidas no ataque complexo. **Acta Scientiarum, Language and Culture**, Maringá, v.35, n.2, p.179-186, 2013.
- CRISTÓFARO-SILVA, T. Branching onsets in Brazilian Portuguese. **Revista de Estudos da Linguagem**, Belo Horizonte, v.10, n.1, p.91-107, 2002.
- CRISTOFOLINI, C. Estudo da monotongação de [ow] no falar florianopolitano: perspectiva acústica e sociolinguística. **Revista da ABRALIN**, [s.l.], v.10, n.1, p.205-229, 2011.
- CYRANKA, L. F. de M.; RONCARATI, C. Atitudes lingüísticas: uma pesquisa em escolas públicas de Juiz de Fora (MG-BRASIL). **Diacrítica**, Braga, v.1, n.24, p.211-231, 2010.
- ECKERT, P. Variation and the indexical field. **Journal of Sociolinguistics**, Chichester, v.12, n.4, p.453-476, 2008.
- EHRI, L. C. Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. **Scientific Studies of Reading**, Philadelphia, v.18, n.1, p.5-21, 2014.
- EHRI, L. C. Learning to read words: theory, findings, and issues. **Scientific Studies of reading**, Philadelphia, v.9, n.2, p.167-188, 2005.

FERREIRA, J. S.; TENANI, L. E.; GONÇALVES, S. C. L. O morfema de gerúndio 'ndo' no Português Brasileiro: análise fonológica e sociolinguística. **Letras & Letras**, Uberlândia, p.167-188, 2012.

FREITAG, R. M. K. Reparos na leitura em voz alta como pistas de consciência sociolinguística. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, São Paulo, v.36, n.2, 2020. Disponível em: <https://doi.org/10.1590/1678-460X2020360206>. Acesso em: 05 ago. 2021.

FREITAG, R. M. K. Uso, crença e atitudes na variação na primeira pessoa do plural no Português Brasileiro. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, São Paulo, v.32, n.4, p.889-917, 2016.

FREITAG, R. M. K. Implicações da variação na alfabetização: a lateral palatal e seus correspondentes grafêmicos. **Leitura**, [s.l.], v.2, n.46, p.37-56, 2011.

FREITAG, R. M. K.; CARDOSO, P. B.; PINHEIRO, B. F. M. Saliência na conservação de /d/ no segmento -ndo: efeitos sociais e estilísticos. **Gragoatá**, Niterói, v.23, n.46, p.654-678, 2018.

FREITAG, R. M. K.; SÁ, J. J. de S. Leitura em voz alta, variação linguística e o sucesso na aprendizagem inicial da leitura. **Ilha do Desterro**, Florianópolis, v.72, n.3, p.41-62, 2019.

FREITAG, R. M. K.; SOUZA, V. R. A. 2019. Discriminação de palavras e efeitos da variação linguística. *In*: SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY AND COLLOCATES EVENTS, 12., 2019, Salvador. **Proceedings**. Salvador: SBC, 2019. p.345-353

FREITAG, R. M. K. *et al.* O uso da língua para a discriminação. **A Cor das Letras**, Feira de Santana, v.21, n.1, p.185-207, 2020.

GONÇALVES, G. F.; AMARAL, V. S. Produções orais e escritas dos ditongos [aj],[ej] e [ow]: dados de São José do Norte/RS. **Diadorim: revista de estudos linguísticos e literários**, Rio de Janeiro, v.14, p.127-154, 2014.

GÖRSKI, E. M.; FREITAG, R. M. K. O papel da sociolinguística na formação dos professores de língua portuguesa como língua materna. *In*: MARTINS, M. A.; TAVARES, M. A. (org.). **Contribuições da Sociolinguística e da Linguística histórica para o ensino de língua portuguesa**. Natal: EdUFRN, 2013. p.11-52.

GUIMARÃES, S. R. K. Influência da variação linguística e da consciência morfossintática no desempenho em leitura e escrita. **Interação em Psicologia**, Curitiba, v.9, n.2, p.261-271, 2005.

HAUPT, C.; SEARA, I. C. Caracterização acústica do fenômeno de monotongação dos ditongos [aj, ej, oj] no falar florianopolitano. **Revista Linguagem & Ensino**, Pelotas, v.15, n.1, p.263-290, 2012.

- KRISTIANSEN, T. Two standards: one for the media and one for the school. **Language Awareness**, Clevedon, v.10, n.1, p.9-24, 2001.
- LABOV, W. **Language in the inner city: studies in the black English vernacular**. Pennsylvania: University of Pennsylvania Press, 1972a.
- LABOV, W. **Sociolinguistic patterns**. Pennsylvania: University of Pennsylvania Press, 1972b.
- LEIRIA, L. L. A ditongação variável em sílabas tônicas finais travadas por -S. **Organon**, Porto Alegre, v.14, n.28-29, p.133-141, 2000.
- MACHADO, A. P. G.; FREITAG, R. M. K. Pistas dos processos de decodificação que levam à compreensão da leitura. **Letras de hoje**, Porto Alegre, v.54, n.2, p.132-145, 2019.
- MACHADO, A. P. G. Variação linguística e leitura: fenômenos variáveis da fala na leitura em voz alta. **A Cor das Letras**, Feira de Santana, v.19, n.4, p.196-218, 2018.
- MADUREIRA, E. D. Reanálise de alguns aspectos da vocalização da lateral palatal no português. **Revista de Estudos da Linguagem**, Belo Horizonte, v.8, n.1, p.125-145, 1999.
- MCKENZIE, R. M. The sociolinguistics of variety identification and categorisation: free classification of varieties of spoken English amongst non-linguist listeners. **Language Awareness**, Clevedon, v.24, n.2, p.150-168, 2015.
- MIRANDA, I. C. C.; CRISTÓFARO-SILVA, T. Aquisição de encontros consonantais tautossilábicos: uma abordagem multirrepresentacional. **Revista Linguística**, Rio de Janeiro, v.7, n.1, p.14-31, 2011.
- MOLLICA, M. C. Enfoques de pesquisa sobre a relação língua e sociedade. **Veredas-Revista de Estudos Linguísticos**, Juiz de Fora, v.5, n.1, p.7-19, 2001.
- MOLLICA, M. C.; MATTOS, P. B. Pela conjugação das abordagens variacionista e difusionista. **Revista de estudos da linguagem**, Belo Horizonte, v.1, n.1, p.53-64, 1992.
- MOLLICA, M. C.; PAIVA, M. da C. Restrições estruturais atuando na relação entre [l] ~ [r] e [r] ~ [0] em grupos consonantais em português. **Boletim da Abralin**, Campinas, n.11, p.181-190, 1991.
- MONARETTO, V. N. de O. O apagamento da vibrante posvocálica nas capitais do sul do Brasil. **Letras de hoje**, Porto Alegre, v.1, n.35, p.275-284, 2000.
- NASCIMENTO, K. R. S.; ARAÚJO, A. A.; CARVALHO, W. J. A. A redução do gerúndio no falar popular de Fortaleza: um olhar variacionista. **Veredas-Revista de Estudos Linguísticos**, Juiz de Fora, v.17, n.2, p.398-413, 2013.
- NEVINS, A. Vowel lenition and fortition in Brazilian Portuguese. **Letras de Hoje**, Porto Alegre, v.47, n.3, p.228-233, 2012.

NICOLOSI, L.; HARRYMAN, E.; KRESHECK, J. **Terminology of communication disorders: speech-language-hearing**. Philadelphia: Lippincott Williams & Wilkins, 2004.

PAIVA, M. da C.; SCHERRE, M. M. P. Retrospectiva sociolingüística: contribuições do PEUL. **DELTA: Documentação e Estudos em Linguística Teórica e Aplicada**, São Paulo, v.15, n.3, p.201-232, 1999.

PAIVA, M. da C.; SILVA, V. L. P. Cumprindo uma pauta de trabalho: contribuições recentes do Peul. **Alfa: Revista de Lingüística**, São Paulo, v.56, n.3, p.739-771, 2012.

PATIL, I.; POWELL, C. **ggstatsplot: “ggplot2” Based Plots with Statistical Details**. [S.l.]: ggstatsplot, 2018.

PENDERGAST, L. L.; KAPLAN, A. Instructional context and student motivation, learning, and development: commentary and implications for school psychologists. **School Psychology International**, London, v.36, n.6, p.638-647, 2015.

PIPER, D. Language awareness in Nova Scotia schools: an exploratory study. **Brock Education Journal**, St. Catharines, v.13, n.1, p.61-71, 2003.

SCHWALM, D. E. Degree of difficulty in basic writing courses: insights from the oral proficiency interview testing program. **College English**, Urbana, v.47, n.6, p.629-640, 1985.

SILVERSTEIN, M. Indexical order and the dialectics of sociolinguistic life. **Language & communication**, Oxford, v.23, n.3-4, p.193-229, 2003.

SQUIRES, L. Processing grammatical differences: perceiving versus noticing. *In*: BABEL, Anna (ed.). **Awareness and control in sociolinguistic research**. Cambridge: Cambridge University Press, 2016. p.80-103.

TOLEDO, E. E. Estudo em tempo real da monotongação do ditongo decrescente/ej/em amostra de Porto Alegre. **Letrônica**, Porto Alegre, v.6, n.1, p.94-107, 2013.

TRUDGILL, P. **Dialects in contact**. Oxford: Blackwell, 1986.

VAN COMPERNOLLE, R. A.; WILLIAMS, L. The effect of instruction on language learners' sociolinguistic awareness: an empirical study with pedagogical implications. **System**, [s.l.], v.41, n.2, p.298-306, 2013.

VIEIRA, S. R.; LIMA, M. D. A. de O. **Variação, gêneros textuais e ensino de Português: da norma culta à norma-padrão**. Rio de Janeiro: Letras UFRJ, 2019.

WITTENBURG, P. *et al.* Elan: a professional framework for multimodality research. **Proceedings of LREC**, p. 1556-1559, 2006.

Recebido em 6 de outubro de 2019

Aprovado em 10 de maio de 2020