



**VARIAÇÃO DO TEPE EM ONSET COMPLEXO NA COMUNIDADE
DE FALA DO RIO DE JANEIRO: UM OLHAR A PARTIR DE
FALANTES EXCLUÍDOS SOCIALMENTE**

**VARIATION OF FLAP IN COMPLEX ONSET IN THE SPEECH
COMMUNITY OF RIO DE JANEIRO: A LOOK FROM SOCIALLY
EXCLUDED SPEAKERS**

Marcelo Alexandre Silva Lopes de Melo¹

Livia Fernandes Silva²

Bruna Oliveira Ranquine³

RESUMO:

Este trabalho apresenta os resultados de um estudo sobre a variação do tepe em onset complexo – como em $p[r]incesa \sim p[\emptyset]incesa$ –, na comunidade de fala do Rio de Janeiro, a partir da análise de dados de produção espontânea de 08 falantes de uma amostra de fala composta por adolescentes excluídos socialmente (Amostra EJLA/PEUL). O objetivo deste estudo é verificar a atuação de condicionamentos estruturais e sociais para a realização do tepe em onset complexo, bem como investigar a possível atuação de condicionamentos lexicais. Partindo dos pressupostos teóricos da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]) e dos Modelos Baseados no Uso (BYBEE, 2001, 2010; PIERREHUMBERT, 2003, 2016), foram analisadas todas as ocorrências de itens com a variável em oito entrevistas selecionadas, sendo os dados levantados submetidos ao programa Rbrul (modelo de efeitos mistos). Os percentuais de ausência do tepe revelam um comportamento semelhante entre os indivíduos da Amostra EJLA e indivíduos de outro grupo e com maior inserção social (Amostra Censo 1980, cf. MOLLICA; PAIVA, 1991), o que aponta para a possibilidade de a variável não ser influenciada por fatores sociais ou de os condicionamentos estruturais atuarem da mesma forma em toda a comunidade de fala. Não foram observados condicionamentos fonéticos para a realização da variável, sendo apenas a variável linguística “tonicidade” selecionada pelo programa. A partir da análise dos itens mais frequentes na amostra, aponta-se para a possibilidade de haver diferentes organizações dos itens que contêm o tepe: itens lexicais com alto índice de ausência de tepe podem indicar uma centralidade nos exemplares sem tepe, ao passo que os com baixa probabilidade e ocorrência categórica do tepe terão a representação dominante ou central com o tepe.

PALAVRAS-CHAVE: Tepe em onset complexo; Variação; Item lexical; Modelo de exemplares.

1 Professor Adjunto, Departamento de Linguística e Filologia, UFRJ, malmelo.lopes@letras.ufrj.br.

2 Graduada em Letras (Português-Inglês), UFRJ, liviafernandes@letras.ufrj.br.

3 Graduada em Letras (Português-Inglês), UFRJ, bruna.o.ranquine@gmail.com.



ABSTRACT:

This work presents the results of a study on the variation of flap in complex onset – as in p[r]incesa ~ p[ø]incesa (princess) –, in the speech community of Rio de Janeiro, from the analysis of spontaneous production data of 08 speakers of a speech sample constituted by socially excluded adolescents (EJLA/PEUL Sample). The aim of this study is to verify the actuation of structural and social conditioning for the realization of the flap in complex onset, as well as to investigate the possible actuation of lexical conditioning. Based on the theoretical assumptions of Variationist Sociolinguistics (WEINREICH; LABOV; HERZOG, 2006 [1968]) and Usage-Based Models (BYBEE, 2001, 2010; PIERREHUMBERT, 2003, 2016), all occurrences of items with the variable in the interviews were analyzed, and the data collected were submitted to the Rbrul program (mixed effects model). The overall percentages of the absence of the flap reveal a similar behavior between individuals in the EJLA Sample and individuals belonging to another social group and with a greater social insertion (Censo 1980 Sample, according to MOLLICA; PAIVA, 1991), which points to the possibility that the variable is not influenced by social factors or that the structural conditioning acts in the same way throughout the speech community. No phonetic conditioning was observed for the realization of the variable, and “stress syllable” was the only linguistic variable selected by the program. From the analysis of the most frequent items in the sample, we point to the possibility of different organizations of the items containing the flap: lexical items with a high rate of the absence of the flap can indicate a centrality in the exemplars without flap, whereas items with low probability and categorical occurrence of the flap will have the dominant or central representation with the flap.

KEYWORDS: Flap in complex onset; Variation; Lexical item; Exemplar Models.

Introdução

O presente trabalho analisa a realização do tepe em onset complexo – como em p[r]ofessora ~ p[ø]ofessora – a partir de dados de produção de uma amostra de fala espontânea composta por adolescentes excluídos socialmente⁴ e que cumpriam medida socioeducativa em uma unidade de internação do estado do Rio de Janeiro (Amostra EJLA/PEUL). Normalmente, o comportamento de indivíduos com esse perfil social não costuma ser objeto de estudos linguísticos. Além disso, poucos são os estudos sobre variação do tepe em onset complexo no português brasileiro (doravante PB), sendo a grande maioria dos estudos voltada para alternância de líquidas em encontros consonantais (rotacismo), como em exp[l]icar ~ exp[r]icar. Portanto, neste trabalho, observa-se exclusivamente a variação da vibrante simples (tepe) em onset complexo, isto é, a realização ou ausência do tepe em encontros consonantais.

O suporte teórico do trabalho conjuga os pressupostos teóricos da Sociolinguística

4 O grau de inserção/exclusão social, neste trabalho, está sendo tomado com base na origem e situação dos indivíduos em diferentes esferas da vida pública: estabilidade de vínculos familiares e afetivos, participação no mercado de trabalho (formal ou informal), processo de escolarização, condições e local de moradia, acesso aos serviços públicos (saúde, transporte entre outros). Em outras palavras, o termo *exclusão social* designa “pessoas definitivamente fora do mercado de trabalho e de todos os direitos que ele assegura” (FONTE, 1995: 108).

Variacionista, no que diz respeito à definição do conhecimento linguístico como contendo a heterogeneidade estruturada e da relação entre língua e sociedade, postulados por Weinreich, Labov e Herzog (1968), e os pressupostos dos Modelos baseados no Uso (BYBEE, 2001, 2010; PIERREHUMBERT, 2003, 2016), no que diz respeito à relação entre conhecimento abstraído e uso, à organização do conhecimento linguístico do falante e status da variação na gramática.

Relativamente à variável em análise, o presente estudo partiu dos resultados do estudo realizado por Mollica e Paiva (1991) com dados de produção obtidos junto à Amostra Censo 1980, isto é, com dados de produção de falantes da comunidade de fala do Rio de Janeiro, inseridos socialmente e que haviam cursado o Ensino Fundamental e/ou Médio. As referidas autoras analisaram a variação de grupos consonantais formados por obstruintes seguidas de líquidas e encontraram condicionamentos fonéticos atuando para a realização do tepe (consoante do grupo e presença de outra líquida na palavra). Cristóforo-Silva (2002) também estudou o cancelamento de líquidas em grupos consonantais no PB e sugeriu que os casos de ausência do tepe em grupo consonantal são casos de difusão lexical, uma vez que, segundo a autora, não há qualquer motivação fonética para a realização ou não da variável.

Neste trabalho, dados de oito falantes da Amostra EJLA foram submetidos ao programa Rbrul, que permite tratar variáveis que são de fato independentes entre si (*fixed effect*) e variáveis que estão aninhadas em outras (*random effect*). Assim, em relação à variável em análise no presente estudo, era esperado encontrar condicionamento(s) fonético(s) semelhante(s) àqueles encontrados por Mollica e Paiva (1991), já que os falantes da Amostra EJLA pertence à mesma comunidade de fala analisada pelas autoras. Por outro lado, esperava-se, também, observar um comportamento dos indivíduos da Amostra EJLA diferente daquele descrito em estudos anteriores sobre a variável, com um percentual maior de ausência do tepe, o que poderia apontar para um condicionamento social da variável. Isto porque, em estudos sobre a variação das coda (s) e coda interna (r), Melo (2012, 2017) observou um comportamento dos indivíduos da Amostra EJLA – adolescentes excluídos socialmente – bastante distinto em relação a estudos anteriores sobre essas mesmas variáveis com indivíduos de outro grupos sociais, sendo tal diferença de comportamento compreendida em razão dos diferentes graus de inserção social dos sujeitos. Em outras palavras, apesar de ser esperado observar os mesmos condicionamentos estruturais de estudos anteriores sobre a variável, esperava-se, em razão de condicionamentos sociais (grau de inserção social dos indivíduos), observar um comportamento diferente entre os indivíduos da Amostra EJLA daquele descrito para indivíduos de outros grupos sociais da mesma comunidade de fala (Amostra Censo 1980, cf Mollica e Paiva, op.cit.). Por fim, tomando como hipótese que itens de alta frequência e que apresentam contexto fonético favorável à mudança sonora tendem a ser afetados mais rapidamente (BYBEE, 2016), buscou-se analisar

o comportamento dos itens lexicais mais frequentes na amostra, a fim de oferecer evidências sobre o papel desses itens na propagação de uma possível mudança sonora.

Serão apresentados a seguir, na seção 2, os conceitos principais que sustentam a abordagem adotada neste estudo: o status da variação na gramática, partindo os pressupostos da Sociolinguística Variacionista e incorporando a hipótese de representação postulado pelos Modelos baseados no Uso. A seção 3 contém revisão bibliográfica dos estudos sobre a variável em análise. Na seção 4, serão apresentadas a metodologia e uma breve descrição da amostra em estudo (Amostra EJLA/PEUL), de onde foram os levantados os dados analisados nesse estudo. As seções 4 e 5, respectivamente, tratam da análise dos dados obtidos e considerações finais.

Variação e gramática

Cristóforo-Silva e Gomes (2004) observam que “[o] primado da categoricidade do sistema ou da invariância da gramática data do estruturalismo linguístico” e teve continuidade com o advento da teoria gerativa. A hipótese propagada pela teoria gerativa ajudou a moldar a concepção de língua para os estudos linguísticos, os quais passaram a considerar o desempenho do falante, a exemplo da dicotomia proposta pelo estruturalismo da primeira metade do século XX, como externo à língua e, portanto, como algo que não deveria integrar as análises. A Sociolinguística Variacionista, desenvolvida no final da década de 1960 (WEINREICH, LABOV e HERZOG, 2006 [1968]), concebe um modelo de descrição e interpretação da variação linguística, a fim de superar a ideia de a língua ser um sistema homogêneo e autônomo, trazendo a possibilidade de se compreender a língua como um sistema heterogêneo que comporta regras categóricas e variáveis. Entender, por sua vez, a língua como um sistema heterogêneo pressupõe admitir que a variação não só é inerente a esse sistema, mas também é influenciada tanto por fatores internos ao sistema linguístico como por fatores de natureza social. A variação não seria, portanto, aleatória, mas sistemática e previsível tanto estruturalmente quanto socialmente. Ao assumirem que “a mudança linguística é um processo contínuo e o subproduto inevitável da interação linguística”, Weinreich, Labov e Herzog (2006 [1968]) rompem com as dicotomias entre ‘indivíduo e sociedade’, ‘sincronia e diacronia’, ‘homogeneidade e heterogeneidade’, e propõem que “uma explicação razoável da mudança dependerá da possibilidade de descrever a diferenciação ordenada dentro da língua” (p. 87-88).

Entretanto, o modelo de organização do conhecimento linguístico tradicionalmente adotado pelos estudos variacionistas se vale de uma concepção de gramática cujo núcleo é invariante e, a partir do qual, as formas variáveis são produzidas. Assim, a adoção desse modelo de arquitetura do conhecimento linguístico – isto é, invariante – traz uma questão importante: como conceber uma gramática para o falante capaz de acomodar a variação se o próprio

conceito de gramática supõe um núcleo invariante (CRISTÓFARO-SILVA e GOMES, 2004, p. 32)? Além disso, conceber gramática como um núcleo invariante implica assumir que a variação é o resultado de um processo que ocorre mediante a aplicação de uma determinada regra a uma forma única, abstrata. Pierrehumbert (1994) argumenta que a variação não pode ser concebida como produto de um processo, mas sim como representação, devendo, para tanto, ser acomodada em um modelo que incorpore a gradiência e distribuições probabilísticas. Ainda de acordo com a autora, por ocupar um lugar central no conhecimento linguístico do falante, a variação não poderia ser representada apenas por meio de processos fonológicos.

Weinreich, Labov e Herzog (2006 [1968], p. 126) mencionaram a importância de contribuições da Linguística teórica para o entendimento da variação e da mudança linguística. Assim, o tratamento representacional dado à variabilidade observada na fala pelos Modelos Baseados no Uso (doravante MBU) pode trazer novas contribuições para a compreensão da natureza do conhecimento linguístico do falante. Nesta abordagem, a representação detalhada é captada pela Teoria de Exemplos.

Primeiramente introduzida na Psicologia como um modelo de percepção e categorização, a Teoria de Exemplos foi estendida para análise dos sons da fala por Johnson (1997) e postula que a estrutura do conhecimento linguístico é construída pela representação na memória da totalidade de experiências linguísticas vivenciadas pelo indivíduo. De acordo com a Teoria de Exemplos, o conhecimento sonoro abstrato não comporta apenas informações de contraste categórico entre segmento (fonemas), mas sim o registro detalhado de toda informação que um indivíduo pode encontrar em um item lexical (FOULKES e DOCHERTY, 2006). Neste sentido, adotar uma representação em exemplos implica assumir que a estocagem dos itens lexicais é redundante, isto é, todas as ocorrências concretas de uso percebidas são armazenadas no léxico, sendo o armazenamento dos itens lexicais realizado de maneira a comportar informação fonética detalhada. Cristóvão Silva e Gomes (2020, p. 19) argumentam que “todas as variantes relacionadas a um item lexical fazem parte de sua representação no léxico”, incluindo as representações detalhadas: informações neurofisiológicas, acústicas e da indexação social.

Uma modelagem do conhecimento linguístico baseada em exemplos não é um modelo que captura somente o detalhe fonético. Para Pierrehumbert (2003, 2016), representações detalhadas no léxico serviriam de base para a emergência de representações abstratas, gerando assim uma escala de diferentes graus de abstração, que se retroalimentariam, sendo que cada nível tem seu próprio aparato representacional. As informações relacionadas às unidades mais abstratas, que mapeiam realizações acústicas diferentes, relações fonotáticas e relações morfofonológicas, constituem diferentes níveis de abstração. Esses níveis de abstração emergem das representações das palavras no léxico, as quais se encontram armazenadas em uma rede de conexões de similaridade sonora e semântica entre as palavras (BYBEE, 2001;

PIERREHUMBERT, 2003). A gramática fonológica é, portanto, emergente das formas das palavras representadas no léxico, que, por sua vez, constituem também generalizações sobre a fala (BYBEE, 2001; PIERREHUMBERT, 1994, 2001, 2003, 2016).

Nesse sentido, Cristófaros-Silva e Gomes (2017, p. 158) salientam que, para a abordagem baseada em exemplares, “o lócus da representação mental é a palavra e sua ocorrência em construções específicas”. As autoras postulam também que os exemplares mais frequentes estão representados por mais ocorrências e são compreendidos como sendo mais robustos do que os exemplares pouco frequentes, os quais estarão representados por menos ocorrências. Os exemplares semelhantes estão próximos e os dissemelhantes, distantes. Sendo assim, “a densidade dos exemplares tem impacto na organização lexical” (PISONI et al., 1985). A hipótese segundo a qual o léxico estaria organizado em uma rede de conexões sonoras e semânticas entre os itens permite que existam representações estáveis e, ao mesmo tempo dinâmicas, haja vista que a estrutura emerge da própria experiência de ouvir e falar.

A representação em exemplares permite, ainda, capturar efeitos de frequência observados em estudos de mudança linguística. Bybee (2001, p. 10) menciona dois tipos de frequência: frequência de ocorrência (*token frequency*) e frequência de tipo (*type frequency*). A frequência de ocorrência é determinada pela quantidade de vezes em que uma unidade ocorre em um determinado corpus, ao passo que frequência de tipo é determinada pela quantidade de itens lexicais que compartilham determinado padrão. Os dois tipos de frequência, ainda de acordo com Bybee (2010, 2012), atuam como mecanismos de propagação da mudança. Segundo Bybee (2001, 2010), a frequência de ocorrência dos itens lexicais impacta a representação lexical e efeitos de frequência de ocorrência podem ter papel promotor ou conservador a depender do tipo de mudança, respectivamente, se sonora ou por analogia. Assim, uma modelagem em exemplares permite que sejam capturados, simultaneamente, efeitos da frequência de uso - por meio da representação em exemplares - e efeitos da frequência de tipo - por meio da organização do léxico em rede.

O tepe em onset complexo

A realização do onset complexo no Português Brasileiro pode ser analisada levando-se em consideração dois fenômenos variáveis distintos: a alternância de líquidas em onset complexo (púb[l]ica ~ púb[r]ica; ing[l]ês ~ ing[r]ês), amplamente conhecida como rotacismo; ou a ausência/presença do tepe (p[r]ofessora ~ professora; t[r]anquilo ~ tranquilo). Neste trabalho, o objeto de análise será a realização/ausência do tepe em onset complexo, isto é, a variação da vibrante simples (tepe) em onset complexo. Portanto, nesta seção, serão apresentados alguns trabalhos que analisaram exclusivamente a variação da vibrante simples (tepe) em onset complexo.

Mollica e Paiva (1991) analisaram a variação de grupos consonantais formados por obstruintes seguidas de líquidas na comunidade de fala do Rio de Janeiro, a partir de dados de fala da Amostra Censo 1980 (PEUL/UFRJ). Essa amostra, constituída no início dos anos 1980, foi estratificada por faixa etária (07 a 14, 15 a 25, 26 a 49 e acima de 50 anos), escolaridade (1º e 2º ciclos do Ensino Fundamental e Ensino Médio) e sexo (homens e mulheres) e os falantes que compõem essa amostra foram distribuídos aleatoriamente por diferentes bairros da cidade do Rio de Janeiro. O estudo de Mollica e Paiva (*op. cit.*) observava tanto os casos de rotacismo quanto a realização/ausência do tepe em onset complexo, uma vez que um dos objetivos das autoras era verificar se os dois casos constituíam fenômenos distintos ou não. No entanto, assim como indicado anteriormente, este trabalho se detém às análises de ocorrência/ausência do tepe em onset complexo, motivo pelo qual serão discutidos apenas os resultados para esse caso.

Assim, relativamente à variação do tepe em onset complexo, as autoras partiam da hipótese segundo a qual, em razão da similaridade fonética na cadeia, os grupos consonantais constituídos por ‘r’ tendem a se simplificar, passando de ‘r’ para ‘0’, sempre que existir, na mesma palavra, um outro ‘r’ (p. 182). Segundo as autoras, essa hipótese de dissimilação pode ser fundamentada pela história do Português, em que, na literatura diacrônica, palavras constituídas por dois segmentos de natureza equivalente, um deles tende a não ser realizado, como em *fratre* > *frade*.

Nesse estudo, as autoras testaram as seguintes variáveis: presença de outro segmento líquido no item lexical e ausência de outro segmento líquido no item lexical. Os resultados confirmaram a hipótese inicial, isto é, que, nos casos em que ocorre a presença de outra consoante líquida na palavra, há uma maior probabilidade de ‘r’ se transformar em ‘0’, pois num contexto em que dois segmentos iguais aparecem em um mesmo item lexical, um deles está propenso a desaparecer. A tabela 01, a seguir, traz os resultados encontrados pelas autoras para a influência da presença de outra líquida na palavra:

Tabela 01 - Influência da presença de outra líquida (Mollica e Paiva, 1991, p. 183)

	Apl/N	%	prob
presença de outra líquida	50/306	16,34	.60
ausência de outra líquida	82/747	10,98	.39

A partir da leitura dos resultados, em relação à realização/ausência do tepe, as autoras sugerem que o fenômeno segue uma restrição diacrônica que continua a operar sincronicamente: “dois segmentos líquidos na mesma palavra, ou seja, cadeias similares levam à diferenciação dos segmentos, fazendo com que um deles se transforme ou desapareça”. Além disso, para as autoras, a passagem do tepe para zero acarreta “um total destravamento silábico, criando-se

uma sílaba simples”, o que corroboraria para “uma tendência mais geral das línguas de evitar sílabas complexas” (p. 183).

Ademais, Mollica e Paiva (1991) consideraram também a cadeia do abrandamento na passagem de ‘r’ para ‘0’ nos grupos consonantais. Para isso, as autoras averiguaram a convergência entre a consoante base do grupo consonantal e as probabilidades de supressão do ‘r’. Tal consoante foi avaliada sob duas perspectivas: modo de articulação e sonoridade. Na tabela 02, a seguir, é apresentada a influência da consoante base do grupo consonantal:

Tabela 02 - Influência da consoante base do grupo consonantal (Mollica e Paiva, 1991, p. 184)

	Apl/N	%	prob
oclusiva	745/2430	31	.75
fricativa	31/116	03	.25
surda	707/2303	31	.64
sonora	41/243	17	.36

Com base nos resultados, as autoras lançaram mão da escala de força proposta por Hooper (1976) para sugerir que a não-realização – ou ausência – do tepe em onset complexo é favorecida quando a consoante do grupo consonantal é uma oclusiva surda – [p], [t], [k], como em *semp[r]e ~ sempe* – e desfavorecida quando a consoante do grupo consonantal é uma fricativa sonora – [v], [z], [ʒ], como em *liv[r]o ~ livo*.

Cristófaros-Silva (2000) analisou o cancelamento de líquidas em grupo consonantal no PB. A autora chama a atenção para o fato de que os encontros consonantais são estruturas instáveis, devido a sílaba CV (consoante + vogal) ser universal e não marcada. No trabalho em referência, a autora examina algumas hipóteses que tentam explicar a ausência do tepe em onset complexo. A primeira hipótese analisada pela autora é a que sustenta ser a redução da consoante nos encontros consonantais explicada para que se chegue à sílaba universal CV. Porém, a autora considera que essa proposta não é adequada por alguns motivos, entre eles o fato de que não é aconselhável buscar um padrão de sílaba universal para categorias diferentes (encontros hétero e tautossilábicos) (p. 524). Além disso, a autora mostra que em estudos atuais a quebra dos encontros consonantais tautossilábicos depende da tonicidade da vogal seguinte. Isso indica que nos encontros consonantais tautossilábicos, se a vogal for pretônica > tônica > postônica, há uma forte tendência ao cancelamento da consoante (p. 525).

Dessa forma, por não ser possível capturar qualquer condicionamento estrutural capaz de explicar a ausência do tepe em grupos consonantais complexos no PB, Cristófaros-Silva propõe que a ausência do tepe nesses encontros deve ser compreendida como um caso de difusão lexical, ou seja, a autora sugere que o fato de um falante realizar ou não uma consoante no encontro consonantal não pode ser previsto, mas, sim, guiado pelo léxico do mesmo. Sendo

assim, ocorre uma “lexicalização em que falantes diferentes têm listas diferentes de palavras que têm formas opcionais (com e sem cancelamento da consoante do encontro consonantal)” (CRISTÓFARO-SILVA, 2000, p. 525).

Para confirmar a sua hipótese, a autora apresenta alguns argumentos, entre eles o de que os itens lexicais com maior frequência na fala apresentam o cancelamento da consoante em encontro consonantal, isto é, quanto mais frequente for um padrão, mais chances ele terá de se aplicar a novos itens no léxico. Ademais, a hipótese de difusão lexical postula que as mudanças são implementadas na língua gradualmente, através de determinados itens do léxico. Sendo assim, ocorreria uma mudança foneticamente abrupta – por haver o cancelamento da consoante – que é implementada de forma gradual. Por fim, tem-se o caso de convergência morfossemântica, em que palavras que inicialmente apresentam-se com estruturas distintas passam a combinar em sua forma e sentido, como em *sob ~ sobre*, *este ~ esse* etc. Portanto, os fatos mencionados anteriormente apontam para um caso de difusão lexical, em que o léxico é o responsável por implementar a mudança sonora. Desse modo, entende-se que, possivelmente, a mudança sonora não atingirá a regularidade plena no cancelamento das consoantes em encontro consonantal.

Cristóforo-Silva (2002) volta a tratar do cancelamento de líquidas em grupos consonantais no Português Brasileiro como sendo resultado de um processo de difusão lexical, sugerindo que a separação de um item lexical em duas formas diferentes, como consequência de uma lexicalização, é um indicativo para esta difusão. Ainda discorrendo sobre esta variável, a autora explora as motivações que levam à redução do tepe em onset complexo. Sendo assim, ela retoma a hipótese investigada no estudo anterior de que esta redução ocorre para que seja alcançado o padrão silábico mais simples e universal: CV. Entretanto, essa hipótese é novamente refutada pela autora, haja vista que uma sequência de duas sílabas CV (CVCV) pode se tornar CCV, como por exemplo em *xí[kar]a → xí[kr]a*. Este tipo de mudança pode ser explicado pelo fato de, no PB, ser rara a acentuação de palavras proparoxítonas e, dessa forma, o falante tende a simplificar a sílaba para que ela se encaixe no padrão comum de acentuação, no caso, a paroxítona.

A referida autora ainda sugeriu que a perda das líquidas em onsets complexos teria relação com o lugar que as líquidas ocupam na escala de sonoridade, segundo a qual às vogais é atribuída a sonoridade máxima e às obstruintes, a sonoridade mínima. Desse modo, pelo fato de os processos de lenição geralmente envolverem mudança consonantal em direção ao extremo mais sonoro da escala, para a autora, a não-realização de líquidas em onset complexo não está relacionada à estrutura silábica ou à simplificação do sistema de

tonicidade: para Cristófaró-Silva, a perda das líquidas em onsets complexos está ligada ao fato de as líquidas ocuparem um lugar elevado na escala de sonoridade e, portanto, estarem sujeitas à redução. Assim, ainda de acordo com a autora, a motivação para a redução das líquidas em onset complexo está na própria estrutura segmental destas.

Portanto, a autora reafirma a sua hipótese sobre a ausência do tepe em onset complexo refletir um caso de difusão lexical ao sugerir que “o léxico é interpretado como um sistema plástico que se adapta às necessidades dos falantes. Eu também afirmei que as mudanças sonoras sempre apresentam formas lexicalizadas em competição ou usadas de forma intercambiável pelo mesmo falante”⁵ (CRISTÓFARO-SILVA, 2002, p. 104).

Metodologia

A metodologia de trabalho é aquela utilizada nos estudos sociolinguísticos no que diz respeito à coleta de dados de fala espontânea, verificação de condicionamentos linguísticos e sociais da variação, além da análise dos dados. Os dados foram coletados a partir de uma amostra de fala espontânea da comunidade de fala do Rio de Janeiro: Amostra EJLA. Essa amostra, elaborada a partir da metodologia da sociolinguística variacionista para obtenção de fala espontânea, integra o acervo do PEUL/UFRJ e representa um grupo social com características socioeconômicas particulares e que quase nunca é objeto de estudos linguísticos.

A Amostra EJLA foi elaborada entre 2008 e 2009 e é constituída por quatorze indivíduos, todos estes do sexo masculino, com idades entre 14 e 20 anos e que, à época das entrevistas, cumpriam medida socioeducativa de internação em uma instituição pública do estado do Rio de Janeiro. Além de serem moradores de favelas e de não terem tido acesso a uma educação formal em uma base regular, os falantes da EJLA apresentam vínculos familiares bastante fragilizados. Em decorrência da situação de vulnerabilidade social em que se encontram, os indivíduos que compõem essa amostra não atuam no mercado formal de trabalho e são aliados das cadeias produtivas. Desta forma, como forma de subsistência, esses adolescentes não só recorrem a atividades relacionadas ao mercado informal de trabalho, mas também à associação a atividades ilegais, tais como o tráfico de drogas. É importante ter em mente que esses adolescentes se encontravam internados, justamente, em razão da prática – em alguns casos, reiterada – de atos infracionais análogos a crimes (tráfico de drogas, roubos, furtos entre outros)⁶. Assim, o

5 “the lexicon be interpreted as a plastic system which adapts to speakers’ necessities. I have also claimed that sound changes always present lexicalized forms either in competition or used interchangeably by the same speaker”

6 Não há, aqui, qualquer valoração de juízo moral acerca dos motivos que levaram esses adolescentes à prática dos atos infracionais e, conseqüentemente, à internação. O que se pretende é, simplesmente, chamar atenção para o processo de vulnerabilidade e exclusão social (baixo grau de inserção social).

baixo grau de inserção social coloca os adolescentes da Amostra EJLA em situação de exclusão social, a qual deve ser entendida nos termos de Fontes (1995, p. 108):

A exclusão social contemporânea é diferente das formas precedentes de discriminação ou segregação, uma vez que tende a criar, nacional e internacionalmente, indivíduos e/ou sociedades inteiramente desnecessários ao universo da produção econômica. Para eles, aparentemente, não há mais possibilidade de integração ou reintegração no mundo do trabalho e da alta tecnologia. Neste sentido, os novos excluídos parecem seres economicamente descartáveis.

Conforme já apontado em estudos anteriores com falantes da Amostra EJLA (MELO, 2012; 2017), o baixo grau de inserção dos indivíduos pode levar a um comportamento linguístico – sob o ponto de vista da produção e da percepção – bastante distinto daquele observado para indivíduos da mesma comunidade de fala, mas pertencentes a outros grupos sociais. Em outras palavras, o grau de inserção social dos indivíduos de uma mesma comunidade pode conduzir a diferentes encaminhamentos e formações, os quais, por sua vez, impactam não só na relação dos indivíduos com o mundo que os cerca, mas também na forma como esses mesmos indivíduos organizam o seu conhecimento linguístico.

As entrevistas que constituem a amostra EJLA se basearam nos princípios metodológicos de Labov (1972) com a finalidade de tornar a amostra mais próxima possível do uso dos falantes, fazendo eles se sentirem à vontade, e até esquecerem que estavam sendo gravados. As gravações ocorreram na própria unidade de internação onde os indivíduos estavam e foram realizadas durante 30 e 60 minutos com o auxílio de um gravador digital. Diante das particularidades dos falantes e por se tratar de jovens que cumpriam medida socioeducativa de internação em unidade fechada, o acesso a este grupo tornou-se difícil não apenas por questões burocráticas, mas também pela desconfiança que um desconhecido poderia gerar. Portanto, faz-se necessário reiterar o fato de que, ainda de acordo com os princípios metodológicos de Labov (1972), as entrevistas foram feitas por uma pessoa próxima aos falantes, o que as tornou mais produtivas.

Foram consideradas todas as ocorrências de itens com a variável na entrevista, com o objetivo de melhor capturar a variação no item lexical. Para o presente estudo, foi realizado o levantamento de 08 (oito) falantes de Amostra EJLA. A variável dependente comporta duas variantes: realização do tepe (at[r]asar; ag[r]adecer; cump[r]ir) e ausência do tepe (at[ø]asar; ag[ø]adecer; cump[ø]ir). Como todos os indivíduos da Amostra EJLA possuem as mesmas características sociais (indivíduos do sexo masculino, mesma faixa etária, mesmo grau de escolaridade), apenas a variável extralinguística ‘indivíduo’ foi testada, sendo todas as demais

variáveis linguísticas. As variáveis independentes testadas foram: ‘primeira consoante do grupo consonantal’, ‘vogal do grupo consonantal’, ‘tonicidade da sílaba em que o onset complexo se realiza’, ‘tamanho do item lexical’, ‘presença de outra líquida na palavra’ e ‘item lexical’. O envelope da variação, conforme já mencionado, considerou apenas a realização ou não do tepe, sem considerar a alternância de líquidas no grupo consonantal. Era esperado encontrar efeitos da consoante do grupo consonantal, bem como da presença de outra líquida na palavra, já que tais efeitos foram observados por Mollica e Paiva (1991) sobre a mesma variável e com dados da mesma comunidade de fala.

Foi utilizado o pacote Rbrul para o tratamento estatístico dos dados por meio da regressão logística. A escolha do pacote Rbrul para realização da análise estatística dos dados se deve à possibilidade de esta ferramenta realizar análise de efeitos mistos, isto é, a análise estatística inclui variáveis aleatórias, como item lexical e indivíduo, juntamente com variáveis independentes. Variáveis aleatórias são aquelas cujos fatores são abertos e podem ter características mapeadas por outras variáveis. No caso das variáveis mencionadas, propriedades dos itens lexicais como, por exemplo, tonicidade da sílaba com coda e tamanho da palavra foram analisadas conjuntamente com a variável item lexical.

Análise dos dados

Esta seção apresenta os resultados obtidos para a variação do tepe em onset complexo. A tabela 03 a seguir apresenta a distribuição das variantes na Amostra EJLA:

Tabela 03 - Resultados tepe em onset complexo na amostra EJLA

	Apl/N	%
realização do tepe	810/1128	71,8
ausência do tepe	318/1128	28,2

A distribuição das variantes para a amostra EJLA contrariou a expectativa inicial que era a de encontrar diferenças significativas entre o comportamento de indivíduos da Amostra EJLA e os resultados encontrados por Mollica e Paiva (1991) para a amostra Censo 1980 (29,3% para a ausência do tepe). Apesar das diferenças – tanto sociais quanto temporal – muito marcadas entre os indivíduos das duas amostras, praticamente não houve diferença quanto à distribuição das variantes, sendo a ausência do tepe ligeiramente menor entre os indivíduos da EJLA (28,2%). Dessa forma, é possível considerar que a variável analisada pareça não ser influenciada por fatores sociais ou que os condicionamentos estruturais atuem da mesma forma em toda a comunidade de fala.

O programa selecionou como significativas⁷ para a realização do tepe em onset complexo as variáveis ‘tonicidade’ (p-valor = 5.71e-09) e ‘indivíduo’ (p-valor = 2.64e-06). A variável ‘indivíduo’ foi rodada como sendo de efeito fixo, uma vez que não há outra variável social, isto é, não há diferença entre os indivíduos da amostra: todos eram do sexo masculino, tinham entre 16 e 20 anos, não concluíram o Ensino Fundamental e pertenciam à mesma classe social (classe baixa) e eram moradores de favelas. Assim, não é possível tratar, neste caso específico, a variável ‘indivíduo’ como aleatória, tendo em vista que não há fatores em aberto e, conseqüentemente, não é possível que outras características sociais sejam mapeadas por outras variáveis. Para essa análise, portanto, a variável ‘indivíduo’ só pode ser tratada como uma variável de efeito fixo. Por outro lado, variável ‘item lexical’ foi tratada como aleatória (*random effects*), tendo em vista que características presentes nessa variável são também mapeadas por outras variáveis relativas a questões linguísticas (primeira consoante do onset complexo, tonicidade da sílaba na palavra, vogal da sílaba em que o onset complexo ocorre, presença de outra líquida na palavra, tamanho do item). Assim, em razão de ter sido tratada como uma variável aleatória, não foi atribuído p-valor à variável ‘item lexical’.

Apesar de diferentes agrupamentos para a variável ‘consoante do grupo’ e diferentemente do que sugeriram Mollica e Paiva (1991), não foram encontrados efeitos significativos desta variável. Essa diferença pode ser explicada em razão do modelo estatístico assumido para a presente análise (efeitos mistos). A fim de buscar evidências para essa hipótese, os dados foram submetidos ao programa estatístico sem que o item lexical fosse incluído como efeito aleatório. Dessa forma, a ‘consoante do grupo’ foi selecionada (p-valor = 0.001), assim como em Mollica e Paiva (op. cit.), ao lado de ‘tonicidade’ (p-valor = 1.14e-66), ‘indivíduo’ (p-valor = 2.1e-07) e ‘vogal da sílaba’ (p-valor = 4.88e-06) e, mesmo assim, com o p-valor mais elevado entre as variáveis selecionadas. Isso parece apontar que, possivelmente em decorrência do modelo estatístico disponível à época do estudo de Mollica e Paiva (1991), efeitos de condicionamento lexical possam ter sido encobertos, o que reforça a necessidade de uma análise que considere o item lexical como uma variável independente.

Em relação às variáveis selecionadas, a ‘tonicidade’ revelou que a ausência do tepe é muito favorecida em sílabas postônicas finais e muito desfavorecida em sílabas tônicas, conforme se depreende da tabela 04 a seguir:

⁷ A relevância das variáveis selecionadas pelo Rbrul é observada em função do p-valor: quanto menor for o p-valor, maior será a significância da variável. Importante esclarecer ainda que “e-(número)” indica que a quantidade de casas decimais atribuídas a um determinado p-valor, ou seja, “e-10” indica que este p-valor é composto por 10 casas decimais que são 0 e a partir daí aparecem os demais algarismos (algo como, por exemplo, 8.38e-10 corresponde a 0,000000000838). Assim, quanto maior o “e-(número)”, menor será o p-valor e, portanto, maior será a significância de determinada variável.

Tabela 04 - Resultados: TONICIDADE para tepe em onset complexo na amostra EJLA

	Apl/N	%	peso
postônicas finais	188/269	68,9	0.849
pretônicas	81/363	20,8	0.421
tônicas	49/496	10,3	0.197

Esse resultado espelha o que geralmente acontece em processos de enfraquecimento, os quais tendem a atingir primeiramente segmentos em sílabas átonas finais e preservar segmentos em sílabas tônicas. É possível observar ainda que, mesmo em sílabas pretônicas, o tepe tende a ser preservado.

Relativamente à variável ‘indivíduo’, é possível observar uma grande variabilidade entre os indivíduos da amostra, o que revela, por sua vez, que é possível capturar, em uma mesma comunidade de fala, diferentes comportamentos entre os indivíduos, os quais contribuem diferentemente para a variação do tepe em onset complexo. A Tabela 05 a seguir traz os resultados para os indivíduos da amostra analisada:

Tabela 05 - Resultados: INDIVÍDUO para tepe em onset complexo na amostra EJLA

	Apl/N	%	peso
M104	77/187	41,2	0.734
U108	40/110	36,4	0.631
R125	47/131	35,9	0.555
L106	29/94	30,9	0.527
R110	15/66	22,7	0.403
J150	31/138	22,5	0.398
R162	43/178	24,2	0.393
R143	36/224	16,1	0.344

O indivíduo R143 mostra um percentual de ausência do tepe bem abaixo da média obtida para a amostra. Observando os dados para esse indivíduo, temos que, dos 224 dados produzidos por ele, 73 foram itens que, na amostra, foram realizados majoritariamente produzidos com o tepe: *tranquilo* (16), *três* (14), *crack* (13), *crime* (12), *patrão* (10), *frente* (8). Isso pode apontar para uma interação entre ‘item lexical’ e ‘indivíduo’ a ser analisada futuramente.

Relativamente à variável ‘item lexical’, foram observados os itens mais frequentes na Amostra EJLA, a fim de analisar o comportamento de tais itens. Após um levantamento junto banco de dados do Projeto ASPA/UFMG para saber a frequência de tais itens na língua, vemos que todos os itens são de alta frequência na língua e revelam comportamentos bastante distintos entre eles:

Tabela 06 - Resultados: ITEM para tepe em onset complexo na amostra EJLA

	Apl/N	%	peso	frequência ASPA
sinistro	4/4	10,0%	0.770	782
dentro	27/31	87,1%	0.796	58707
outro	62/75	82,7%	0.744	124867
sempre	09/11	81,8%	0.681	105307
outra	26/32	81,2%	0.689	89126
quatro	21/27	77,8%	0.650	119244
problema	10/13	76,9%	0.917	72767
professora	07/14	50,0%	0.716	11805
primeira	07/14	50,0%	0.777	136785
monstro	1/3	33,3%	0.422	1947
frente	11/39	28,2%	0.835	57755
comprar	04/16	25,0%	0.802	26994
primeiro	04/16	25,0%	0.584	168229
entra	3/15	20,0%	0.240	18049
padrão	01/11	9,1%	0.629	1827
crime	02/24	8,3%	0.570	38353
trabalho	1/14	7,1%	0.347	129023
tranquilo	3/59	5,1%	0.244	3883
tráfico	0/17	0,0%	0.253	12754
crack	0/40	0,0%	0.195	2691
três	0/48	0,0%	-	203629

As inversões de pesos relativos em relação à frequência observadas na Tabela 06 acontecem, provavelmente, por conta da distribuição de ocorrência dos itens por indivíduo: itens com menos de 08 ocorrências (por exemplo, os itens *sinistro* e *monstro*) não foram realizados por todos os indivíduos da amostra ou, ainda, um mesmo indivíduo pode ter realizado todas as ocorrências desses itens. Além disso, pode ser que mesmo aqueles itens que ocorreram mais de 08 vezes (número de indivíduos da amostra que foram analisados) não tenham sido realizados por todos os indivíduos da amostra.

De uma maneira geral, a análise dos itens mais frequentes na amostra permite observar que os itens que foram mais realizados sem o tepe são aqueles em que, com exceção de *problema*⁸,

⁸ apesar de a sílaba em que o grupo consonantal analisado acontece não ser átona final, em *problema*, a sílaba é uma pretônica – não é uma sílaba tônica – e também se verifica a presença de uma outra líquida na palavra, um condicionamento importante analisado por Mollica e Paiva (1991).

o onset complexo se realiza em sílaba átona final: *sinistro, dentro, outro, sempre, outra e quatro*. No entanto, é possível verificar a existência de itens em que o grupo consonantal está na sílaba átona final e que, mesmo assim, há um baixo percentual de ausência do tepe (*monstro e entra*). Além disso, é possível observar que itens em que o grupo consonantal está na sílaba tônica - contexto altamente desfavorecedor - tendem a preservar o tepe (*crack, tráfico e três* são categoricamente realizados com o tepe). Ocorre que, mesmo em sílaba tônica, é possível observar a ausência do tepe em grupo consonantal, como em *comprar, patrão e crime*.

Considerações finais

O presente artigo apresentou resultados de uma pesquisa que observou a realização do tepe em onset complexo, a partir de dados de produção de um grupo de adolescentes excluídos socialmente, com baixo grau de inserção social e com pouco acesso às instituições sociais responsáveis pela transmissão de valores socialmente prestigiados: todos eram moradores de favelas da cidade do Rio de Janeiro, tinham baixa escolarização, apresentavam vínculos familiares e relacionais bastante esgarçados e, à época das entrevistas, cumpriam medida socioeducativa de internação em uma instituição do estado (Amostras EJLA).

Por meio dos resultados encontrados, é possível considerar que a variável analisada parece não ser influenciada por fatores sociais ou que os condicionamentos estruturais atuem da mesma forma em toda a comunidade de fala, uma vez que a distribuição das variantes não mostrou diferença entre os indivíduos das amostras EJLA e Censo 1980 (cf. MOLLICA e PAIVA, 1991). Em outras palavras, apesar das diferenças temporais entre a constituição das amostras (mais de 25 anos) e as diferenças sociais entre os falantes que compõem as duas amostras, não foi possível observar diferenças no comportamento dos dois grupos – Amostra Censo 1980 e Amostra EJLA – em relação à ausência do tepe em onset complexo.

Até o presente momento e diferentemente do que foi proposto por Mollica e Paiva (1991), não foram encontrados condicionamentos fonéticos para a realização do tepe em onset complexo. Atribui-se essa diferença ao modelo estatístico assumido para o presente estudo: modelo de efeitos mistos. A análise estatística disponível à época em que Mollica e Paiva (op. cit.) realizaram o estudo sobre a variável pode ter encoberto possíveis efeitos do item lexical. O único condicionamento linguístico observado que se mostrou significativo foi a tonicidade da sílaba em que o tepe em grupo consonantal ocorre, sendo o contexto favorecedor para a ausência do tepe aquele que é geralmente observado em fenômenos de enfraquecimento de segmentos sonora: ocorrência do grupo consonantal em sílaba átona final.

Ao serem analisados os itens mais frequentes na amostra e tomando a representação em exemplares, podemos sugerir que há diferentes organizações dos itens lexicais que apresentam

o tepe em onset complexo: itens lexicais com alto índice de ausência de tepe podem indicar uma centralidade nos exemplares sem tepe, ao passo que os com baixa probabilidade e ocorrência categórica do tepe terão a representação dominante (cf CONNINE et al, 2008) ou central (cf MELO, 2012) com o tepe. Isto porque as diferentes representações mentais incorporam as mesmas possibilidades fonéticas disponíveis para os falantes da comunidade de fala – no caso “ausência” e “realização” do tepe –, mas com distribuições probabilísticas diferentes, em função da experiência (social) dos indivíduos. Em outras palavras, as representações abstratas dos itens lexicais contêm todas as variantes encontradas na fala, sendo a frequência das variantes dos itens lexicais a responsável pela robustez das representações de cada item. Assim, todas as possibilidades fonéticas de um item lexical se encontram representadas no léxico dos falantes, sendo a frequência das variantes com que esses itens são representados um parâmetro que atua no processamento. No caso, o comportamento bastante diferenciado entre os itens mais frequentes na Amostra EJLA pode apontar para, como já dito, diferentes centralidades nas representações desses mesmos itens. Assim, itens realizados predominantemente sem o tepe – tais como *sinistro*, *dentro*, *outro*, *sempre*, *outra* e *quatro* têm a sua representação central ou dominante sem o tepe, ao passo que itens realizados predominantemente com o tepe – tais como *patrão*, *crime*, *trabalho*, *tranquilo*, *crack*, *tráfico*, *três* – têm a sua representação central ou dominante com o tepe.

Assumindo a hipótese de representações múltiplas, segundo as quais haveria diferentes organizações dos itens lexicais, é possível melhor acomodar, ainda, o fato de os itens *monstro* e *entra*, em que o onset complexo ocorre em contexto que favorece a ausência do tepe (sílabas átonas finais), terem sido realizados predominantemente com o tepe. Entende-se que, para esses dois itens, apesar de o contexto favorecer a ausência do tepe, a representação central ou dominante inclui o tepe, sendo a representação periférica sem o tepe. Assim, é possível que outras propriedades dos itens, tais como frequência de ocorrência e grau de especialização (OLIVEIRA, 1991), possam estar atuando ao lado de tonicidade da sílaba em que o onset complexo ocorre.

Desta forma, é preciso analisar, ao lado de condicionamentos estruturais, a atuação de possíveis condicionamentos lexicais para realização das variantes, sem que um tipo de condicionamento exclua o outro. Um modelo que concebe a representação múltipla de itens lexicais pode melhor acomodar – e conciliar – tanto efeitos estruturais, como efeitos lexicais, uma vez que tal modelo propõe a organização do léxico em uma rede de conexões sonoras e semânticas entre os itens, permitindo, assim, capturar efeitos de ambas as naturezas, além de efeitos de frequência.

Por fim, em estudos futuros, pretende-se incluir falantes de outros grupos sociais da mesma comunidade de fala, a fim de confirmar (ou não) a ausência de condicionamento fonético, bem como melhor observar o papel do item lexical na propagação da mudança sonora. Além disso,

pretende-se complementar o estudo com dados de produção eliciada de palavras do PB, bem como com uma análise acústica que permita capturar a gradiência da realização/ausência do tepe.

REFERÊNCIAS

BYBEE, J. *Phonology and language use*. Cambridge: Cambridge University Press, 2001.

BYBEE, J. *Language, usage, and cognition*. Cambridge: Cambridge University Press, 2010.

BYBEE, J. Usage-based theory and exemplar representation. In: HOFFMAN, T.; TROUSDALE, G. (eds.) *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press, 2013, p. 49-69.

CONNINE, C. M.; RANBOM, L. J.; PATTERSON, D. J. Processing Variant Forms in Spoken Word Recognition: The Role of Variant Frequency. *Perception & Psychophysics*, v. 70, n.3, 2008, p. 403-411.

CRISTÓFARO-SILVA, T. Sobre a quebra de encontros consonantais no português brasileiro. *Anais de Seminários do GEL*. São Paulo, 2000.

CRISTÓFARO-SILVA, T. Branching onsets in Brazilian Portuguese. *Rev. Est. Ling.*, Belo Horizonte, v. 10, n.1, p. 91-107, jan./jun. 2002.

CRISTÓFARO SILVA, T.; GOMES, C. A. Fonologia na perspectiva dos Modelos de Exemplares. In: GOMES, C. A. (org.) *Fonologia na perspectiva dos Modelos de Exemplares*. São Paulo: Contexto, 2020, p. 13-16.

_____. Teoria de Exemplares. In: HORA, D.; MATZENAUER, C. (org.) *Fonologia, Fonologias: uma introdução*. São Paulo: Contexto, 2017, p. 157-168.

DOCHERTY, G. J.; FOULKES, P. An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua*, vol. 142, p. 42-56, 2014.

FONTES, V. Sobre a exclusão: alguns desafios contemporâneos. *Caderno CRH*. Salvador, n.23. p.98-119, jul/dez.1995

FOULKES, P.; DOCHERTY, G. J. The social life of phonetics and phonology. *Journal of Phonetics*, vol. 34, 2006, p. 409-438.

GOMES, C. A.; CRISTÓFARO SILVA, T. Variação linguística: questão antiga e novas perspectivas. *Linguagem*, ILAPEC/Macapá, v. 1, n.2, 2004, p. 31-41.

MELO, M. A. S. L. de. *Desenvolvendo novos padrões na comunidade de fala: um estudo sobre a fricativa em coda na comunidade de fala do Rio de Janeiro*. Dissertação (Mestrado) do

Programa de Pós-Graduação em Linguística, Faculdade de Letras, UFRJ, Rio de Janeiro, 2012.

MELO, M. A. S. L. de. *Direcionalidade da mudança sonora: o papel do item lexical e da avaliação social*. Tese (Doutorado) do Programa de Pós-Graduação em Linguística, Faculdade de Letras, UFRJ, Rio de Janeiro, 2017.

MOLLICA, M. C.; PAIVA, M. da C. de. Restrições estruturais atuando na relação entre [L] -> [R] e [R] -> 0 em grupos consonantais em Português. *Boletim da Associação Brasileira de Linguística*, nº 11, 1991.

OLIVEIRA, M. A. de. A controvérsia neogramática reconsiderada. Artigo originariamente escrito e publicado em inglês: The neogrammarian controversy revisited. *International Journal of the Sociology of Language*, n. 89, pp. 93-105, 1991.

PIERREHUMBERT, J. B. Probabilistic Phonology: Discrimination and Robustness. In: BOD, R.; HAY, J.; JANNEDY, S. (eds.) *Probability Theory in Linguistics*. Cambridge MA: The MIT Press, p.177-228, 2003.

HAY, J.; JANNEDY, S. Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics* v. 2, p. 33-52, 2016.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*; tradução Marcos Bagno; revisão técnica Carlos Alberto Faraco. São Paulo: Parábola Editorial, 2006.