

UM OLHAR DINÂMICO PARA QUESTÕES DE PRODUÇÃO DE FALA EM L2: ESTUDO PRELIMINAR SOBRE A DESSONORIZAÇÃO TERMINAL (DT)

Jeniffer Imaregna Alcantara de ALBUQUERQUE
Universidade Federal do Paraná
jeniffer.albuquerque@gmail.com

Adelaide Hercília Pescatori SILVA
Universidade Federal do Paraná
adelaidehpsilva@gmail.com

RESUMO: Este trabalho analisa a produção de oclusivas finais do inglês, produzidas por aprendizes paranaenses, a fim de discutir o fenômeno da dessonorização terminal (DT). A discussão sobre a (DT), que teve início com Eckman (1987), descrevendo-a como sendo a perda do traço [sonoro] em algumas obstruintes em posição final, ganha, posteriormente, outro olhar a partir de uma perspectiva dinâmica. A partir dos resultados de um experimento preliminar, observou-se a existência de uma gradiência no fenômeno da DT. Este estudo tem como foco analisar pesquisas que se debruçaram sobre a DT, propondo desdobramentos do experimento para o processo de ensino-aprendizagem.

PALAVRAS-CHAVE: dessonorização terminal; processos de ensino-aprendizagem; sistemas dinâmicos.

ABSTRACT: This work aims to analyze the production of English stop consonants which were produced by language learners from Paraná. Eckman (1987) described the phenomenon of terminal devoicing as the loss of a [sonorous] feature in some final positions. This phenomenon will be reviewed through another paradigm which leads to a dynamic perspective. Based on a preliminary study, it was observed the presence of a gradience in the DT. Therefore, this study focuses on the description of this other view towards DT and the teaching-learning process.

KEYWORDS: terminal devoicing; teaching-learning process; dynamic systems.

0. Introdução

Acompanhamos ao longo dos anos uma mudança significativa no cenário das pesquisas relacionadas à aquisição de segunda língua e à prática docente. Com o crescimento da incorporação de exercícios

fonéticos nos materiais didáticos, os professores têm se debruçado sobre tais práticas por vezes discordando delas, uma vez que existe a queixa de que, apesar da quantidade de horas de treino, o aluno acaba por não incorporar os sons apresentados à sua produção. Além disso, mesmo quando os docentes não discordam de todo e acabam por utilizar tais exercícios em sala de aula, notam que parece haver uma discrepância entre a produção do aluno-aprendiz e a de um nativo. Tal discrepância faz com que, muitas vezes, o aluno seja descrito como alguém que não é capaz de produzir os sons da L2 alvo de modo correto, e que sua produção é desviante.

Diversos estudos que lidam com fenômenos ligados à aquisição de primeira e segunda línguas têm focalizado as semelhanças e as distinções de ambos os processos de aquisição. Trabalhos como o de Brown (1977) acabam por discutir a existência de diferenças entre os processos de aquisição de L1 e L2, mas, ao mesmo tempo, assumem uma postura de pensar que a semelhança entre as línguas tem influência na aprendizagem do aluno, o que facilitaria o processo de aquisição. Igualmente, em Krashen (1982) encontramos uma preocupação não só linguística, mas também de ordem pedagógica, sobre os processos de aquisição de L2. Nesse estudo e, principalmente em outros trabalhos de Krashen, a questão gira em torno do fato de a aquisição de uma língua estar conectada com o que o autor chama de “filtro afetivo”, ou seja, mesmo que um aluno compreenda um enunciado, se este tiver barreiras com relação ao aprendizado da L2 em questão, não conseguirá utilizar, de fato, a informação ouvida. De acordo com Stevick (1976), relativamente ao “filtro afetivo”, o aluno não conseguiria fazer uso da informação, pois o *input* recebido não atingiria a parte do cérebro responsável pela aquisição de uma segunda língua. Ainda que o foco deste estudo não seja questionar os conceitos de aquisição e aprendizagem, é importante ressaltar que, para Krashen, existem diferenças entre ambos os conceitos. O primeiro está relacionado a um processo natural de aprendizagem e o segundo é resultado de instrução formal. Nesse sentido, para o autor, a pronúncia, por exemplo, é algo adquirido e não aprendido.

Consequentemente, podemos dizer que estudos como os de Krashen guiam-nos a uma noção de aquisição e aprendizagem de L2 pautada em acertos e erros. Estudos como esses adotam estão pautados na crença de que é necessário que exista um estímulo ao aprendiz para que ele desenvolva uma resposta. A motivação dos trabalhos parece ser a descoberta de como fazer com que o aprendiz de uma L2 seja bem sucedido no processo de aquisição, de maneira a descrever suas limitações com relação aos mecanismos da língua-alvo por ele desconhecidos.

Essa visão behaviorista aplicada ao ensino de línguas estrangeiras é suplantada, em parte, pelo surgimento de novos modelos dentro da psicologia, que passam a ser empregados na relação de ensino-aprendizagem, como, por exemplo, o modelo interacionista ou o modelo cognitivista, e, em parte, dentro da linguística, pelo surgimento de ferramentas automáticas para análise de dados de fala, que permitem verificar, muitas vezes, que a produção inicialmente percebida como “erro” traduz uma tentativa do aprendiz em aproximar sua produção à produção de um falante nativo. Dessa maneira, tais ferramentas contribuem para o ensino de pronúncia de línguas estrangeiras, já que podem mostrar ao professor quais aspectos, de fato, devem ser trabalhados com seus alunos. Cabe mencionar que, no que concerne à utilização de ferramentas para análise automática do sinal de fala, a literatura já menciona, desde a década de 90, estudos que utilizam essas ferramentas no ensino de línguas estrangeiras. Podemos citar como exemplo, os estudos de Bernstein (1998), Devlieger (1998), Davies e Poesio (1998), Townshend et al (1998). Ainda que exploratórios, tais estudos começam a difundir os resultados preliminares sobre a aplicabilidade de tais ferramentas no campo de aprendizagem de segunda língua. Contudo, como menciona Rocca (2003:21):

O ensino de línguas baseado na tecnologia é uma atividade que requer um trabalho interdisciplinar, muitas vezes dificultado pela necessidade de reunir profissionais de formação bastante diversa, tais como engenheiros, programadores e professores. No entanto, uma abordagem multidisciplinar, que envolva tecnólogos de fala, linguistas e especialistas em ensino de línguas, pode indicar caminhos promissores para o uso da tecnologia [...].

É importante enfatizar, com base no excerto acima, que programas de análise de dados de fala oferecem percentuais que são extraídos pelos pesquisadores, a partir do que se mostra relevante para uma dada pesquisa. Como menciona Price (1998), ferramentas tecnológicas nada fazem sem a instrução do pesquisador e, sem o olhar do pesquisador, “a tecnologia para dados da fala em nada contribui para o aprendizado da linguagem”¹.

Nesse sentido, este trabalho pretende discutir um experimento que conta com a produção de plosivas finais do inglês, produzidas por falantes paranaenses, utilizando gravações que foram analisadas posteriormente através do programa Praat (Boersma e Weenick, 2006).

¹ Trecho original de Price (1998:81) “Without people who know how to use them, books do nothing for literacy, computers do nothing for computation, and speech technology does nothing for language learning”.

Temos como objetivo, também, fazer uma ponte entre os resultados do experimento e o processo de ensino-aprendizagem de segunda língua. Não iremos, no entanto, traçar estratégias de ensino, mostrar uma proposta de como seria um plano de trabalho que incorporasse a utilização desses resultados. Contudo, esperamos que o trabalho possa suscitar uma reflexão por parte de docentes e discentes no tocante à produção de sons na língua alvo, nesse caso particular, de plosivas finais do inglês.

1. Dessonorização Terminal (DT): revisão bibliográfica

Como foi dito na introdução deste trabalho, este estudo tem como objetivo analisar a produção de oclusivas finais do inglês produzidas por aprendizes paranaenses. Vários estudos discorreram ao longo dos anos sobre a dificuldade na distinção entre plosivas sonoras e surdas, no nível da representação, e sobre a existência de um processo de neutralização do fenômeno de dessonorização terminal por parte dos aprendizes de uma L2.

Nessa direção, temos os primeiros estudos realizados sobre a DT: tal fenômeno não é recente, este vem sendo objeto de estudo desde a década de 80, com o trabalho de Eckman (1981). Nesse estudo e em um posterior (1987), o autor caracteriza a DT como sendo um processo que envolve a perda do traço [sonoro] em algumas posições finais e, ainda, como um processo que ocorre em língua materna e estrangeira. Outros autores, como Major (1987), também descreveram a DT como se tratando da perda do traço, mas, para este, o processo da DT se dá na interlíngua. A questão aqui é que, a partir desses resultados, os pesquisadores preveem que o mesmo deve ocorrer para a produção de oclusivas em todas as línguas. Contudo, esses resultados não parecem se confirmar em pesquisas com outras línguas.

Ambos os trabalhos acima citados trazem a DT como um processo que se divide em duas categorias: a presença ou não de sonorização. Entretanto, pesquisas recentes apontam para outra direção. Zimmer e Alves (2007) e trabalhos posteriores apontam para a necessidade de se observar os contextos adjacentes à obstruente e os correlatos acústicos envolvidos na caracterização da DT.

No inglês, segundo Flege (1987 e outros trabalhos), as oclusivas que sofrem desvozeamento não compartilham as mesmas propriedades das oclusivas surdas. Os falantes nativos distinguem, com facilidade, pares surdo-sonoro, tais como as oclusivas que serão observadas.

Ao analisar a DT em outras línguas, autores como Ernestus e Baayen (2006) verificaram que, no holandês, o fenômeno se dá de maneira incompleta, pois algumas distinções se mantêm, como o [t] e [d]. Como discorrem Piroth e Janker (2004), o mesmo ocorre no alemão,

em que há uma neutralização total de alguns correlatos acústicos ligados à DT, como duração da vogal e vozeamento na oclusão, mas não em outros, como duração da plosiva, como citado abaixo, por Oostendorp (2007:1):

1. Alemão:

i. blin[t] 'blind(PRED.)'- blind[ə] 'angry(ATT)'

ii. bunt 'colourful(PRED.)'- bunt[ə] 'colourful(ATT)'

O que podemos verificar nessa citação é que em ambos os exemplos – (i) e (ii) – temos as primeiras palavras terminando com a plosiva bilabial alveolar. No entanto, quando da presença de contexto vocálico adjacente em (i), temos a mudança da surda para sua contraparte sonora. Contudo, em (ii) temos o mesmo contexto adjacente e não temos a mudança da consoante surda em sonora. Segundo Oostendorp, para línguas que apresentam casos como do alemão em (1), a literatura acaba por explicar que o desvozeamento final é foneticamente incompleto: “[...] a neutralização não é completamente perdida, mas é passível de ser recuperada através do detalhe fonético fino” (Oostendorp, 2007:2)². De acordo com o autor, esse tipo de visão sobre o fenômeno faz com que se assumam que teorias formalistas não são capazes de oferecer uma análise que dê conta da neutralização. Oostendorp continua a argumentação pontuando que trabalhos como os de Port e Leary (2005, *apud* Oostendorp 2007) acabam por atribuir a não adequação de modelos formalistas para tratar fenômenos como a DT ao modelo proposto por Chomsky e Halle (1968), cujas representações, segundo Oostendorp, são demasiadamente simples.

No entanto, apesar de algumas mudanças nos paradigmas teóricos, ainda é possível notar uma distinção entre os níveis fonético e fonológico.

Além da abordagem citada acima, contamos com a tentativa de analisar o fenômeno da DT como dinâmico, que procura descrever a DT identificando a gradiência presente no grau de sonorização entre as plosivas. Van Gelder e Port (1995) mencionam que os sistemas dinâmicos, advindos da matemática, são sistemas que se configuram pela presença de estados numéricos que se desenvolvem através do *cotínuum* temporal, mas não de modo aleatório, e sim de acordo com regras que são internas e externas ao sistema. Uma característica interessante desses sistemas é que são passíveis de mudanças, não sendo estáticos, portanto. Quando analisamos seu funcionamento, observamos que um “estado jamais é separado de forma estanque do

² Trecho original “[...] the neutralisation is not completely lost, but recoverable in fine-grained phonetic detail”.

próximo, já que há uma ideia de continuidade entre os estados, que deve ser preservada na relação temporal” (Elman, 1998).

Ainda que não se tenha uma discussão em aberto sobre mecanismos formais que deem conta da complexidade desses sistemas, o sistema linguístico de um aprendiz parece configurar-se como um sistema dinâmico, uma vez que a mudança em um dos subsistemas, se assim podemos chamar, seja ela sintática, fonético-fonológica, semântica, morfológica, etc., causa uma mudança em outro subsistema. No entanto, os sistemas dinâmicos também possuem restrições, como dissemos anteriormente, ou seja, regras que estabelecem uma ordem para a complexidade do sistema.

Além do exposto acima, é importante ressaltar que esses sistemas funcionam a partir das suas propriedades que podem variar em função do tempo e, a partir das alterações que são feitas no sistema, se configuram como sistemas não-lineares, mas que tendem à linearidade em alguns momentos.

Dentro da perspectiva de ensino-aprendizado, torna-se extremamente importante considerar uma abordagem dinâmica, uma vez que a variável tempo mostra-se relevante na análise dos dados de produção e percepção em sala de aula. Ainda que estudos perceptuais não constituam o foco deste trabalho, é relevante mencionar o trabalho de Sancier e Fowler (1997), com a realização de um experimento que suscita duas importantes questões: 1) a percepção de falantes nativos quanto à presença do que eles chamaram de “sotaque”, quando expostos às produções de uma informante falante nativa de português e de inglês como L2; e 2) a importância de verificar os valores de VOT³, como fator relevante para a diferença na percepção da produção da informante. Um fato interessante sobre essa informante é que ela, periodicamente, realizava viagens para seu país de origem e, posteriormente, retornava aos Estados Unidos. O seu “sotaque”, como chamaram os autores, era percebido em ambas as línguas, indicando que havia influência da língua materna na sua L2 e vice-versa. Embora uma investigação mais aprofundada sobre o processo de interferências de uma língua em outra não seja aqui realizada, experimentos como o de Sancier e Fowler iluminam as discussões da área.

No início dos estudos fonológicos muitos trabalhos debruçaram-se sobre o desempenho dos aprendizes de L2. Como menciona Escudero

³ VOT é sigla para *voice onset time*, que é o tempo decorrido entre a soltura da oclusão de uma consoante plosiva até o início do vozeamento do som seguinte. Diferentes durações de VOT são associadas à sonoridade na literatura fonética. A diferença mais marcante é a seguinte: plosivas surdas têm valores positivos para o VOT e plosivas sonoras, valores negativos (cf. Kent e Read, 1992). Mas há também diferenças na duração do VOT que se refletem em consoantes aspiradas e não aspiradas, por exemplo, de modo que o VOT destas é menor do que o daquelas.

(2005), Polivanov (1931) relatou, em diversos estudos, a manifestação de fonemas da L1 na L2. De acordo com Escudero (2005), esses achados poderiam indicar que algumas das dificuldades encontradas na produção da L2 seriam devidas à influência da percepção em L1. A autora também faz referência aos trabalhos de Trubetzkoy (1939/1969), os quais indicaram a possibilidade de que “erros” de produção na L2 tenham influência da percepção, uma vez que, para Trubetzkoy, o sistema fonológico da L1 desempenha o papel de “filtro” e, assim, os sons da L2 são percebidos através desse filtro.

O que podemos entender, de acordo com esse paradigma exposto por Escudero, é que as inadequações da produção de sons na L2 podem ter como base não conseguir perceber esses sons de modo correto. Em decorrência desses achados, observamos que, como nos estudos de Llisterri (1995) e Leather (1999) (apud Escudero, 2005), a percepção se daria antes da produção, pois, assim como foi dito acima, uma vez que não se percebe, com acuidade, o som no sistema da L2, a produção acaba por ficar comprometida.

Outro modelo que procura enxergar uma intersecção entre a produção e a percepção, mas lançando mão de outros primitivos de análise (não mais os traços, mas gestos articulatórios) é a Fonologia Articulatória. Modelos dinâmicos de produção da fala, como o realismo direto e a Fonologia Articulatória (FAR) de Browman e Goldstein (1992) tomam como primitivos de análise os “gestos articulatórios”. Estes são, ao mesmo tempo, unidades fonológicas e unidades articulatórias (físicas, portanto). De acordo com esse modelo, cada som se caracteriza pela presença de um gesto articulatório, o qual pode fazer parte de um dos subsistemas articulatórios: oral, laríngeo ou vélico. Os gestos recebem especificações, em cada subsistema, para grau e para local de constricção. A partir de tais especificações, os gestos acabam por representar variações como as de rigidez ou amortecimento, as quais “geram variações de magnitude gestual e de ativação temporal do gesto” (Silva, 2003). Nesse sentido, podemos entender a DT, além da análise feita pela OT, a partir de uma abordagem dinâmica, na qual existe faseamento de gestos, do movimento de pregas, que, para os aprendizes de uma L2 não se configura, necessariamente, como realizando vibração ou não, mas, como veremos a partir do experimento exposto, em um nível intermediário.

A partir da exposição desses paradigmas teóricos e da solução apresentada por cada um para dar conta da DT, mostraremos um experimento, de natureza preliminar, que procura rediscutir a DT.

2. A pesquisa

Antes de mencionar a pesquisa em si, é necessário mencionar que esta foi realizada em parceria com a Prof^ª Dr^ª Márcia Zimmer, da UCPel.

Como relatamos anteriormente, já foram realizados vários experimentos, em várias línguas, os quais procuram analisar o processo de neutralização total e parcial que se encontram associados à DT.

Como já foi comentado anteriormente, nos experimentos de Major e Eckman, os autores analisaram um único fator: a presença ou não de sonoridade nas oclusivas sonoras. Também, é preciso mencionar que os autores utilizaram a inspeção auditiva para realizar a análise dos dados. Zimmer, até seu trabalho de 2007, fazia uso da análise de outiva, mas, neste último trabalho aqui citado (ano), a autora apontou vários fatores que se mostram relevantes para a descrição da DT. Foram observados três deles: percentual de vozeamento, duração da oclusão e a duração da vogal que precede a oclusiva. A questão é: por que observar esses fatores? O nosso ouvido não consegue fazer uma análise acurada de algumas nuances como, por exemplo, uma sonorização que se dá pela metade ou quando essa parece se dar a partir de um segmento sonoro que antecede ou precede o segmento plosivo alvo, mas que quando analisado no espectrograma, não conseguimos enxergar a barra de sonoridade deste. Com a análise acústica dos dados, percebeu-se também, como dissemos acima, que não bastava olhar para a presença da barra de sonoridade nos segmentos, mas que outros fatores, como a duração da vogal e a duração da oclusão podiam ser fatores importantes para a análise, uma vez que a sonoridade do segmento oclusivo também dependia dos contextos adjacentes. Nesse sentido, mostra-se de extrema importância que, ao falar da metodologia de coleta e análise de dados, seja levado em conta um *corpus* que contemple ambientes adjacentes distintos, para averiguar se há ou não uma distinção de produção a depender do contexto, e, também, para observar se os resultados encontrados são os mesmos para todas as oclusivas.

Com relação à metodologia, o *corpus* é composto por 72 sentenças, das quais 48 eram sentenças-alvo e 24 eram sentenças distratoras. As palavras-alvo foram escolhidas de modo que tivéssemos dois exemplos de todas as oclusivas. Essas palavras possuem as oclusivas em posição final. Escolhemos analisá-las, pois, no Português Brasileiro (PB), não temos consoantes oclusivas nessa posição. Abaixo, podemos conferir a lista das palavras utilizadas nas sentenças, as quais foram dispostas em fichas para serem lidas pelos informantes:

1. Seis itens terminados em obstruintes sonoras, escolhidos de forma a neutralizar a variável transferência grafo-fônico-fonológica: *bad, pad, bob, cab, rag, bag*;
2. Dois itens terminados em fricativas sonoras, escolhidos de forma a ensejar a transferência grafo-fônico-fonológica: *falls, bobs*;

3. Oito palavras correlatas terminadas em obstruintes e fricativas surdas das palavras do itens 1 e 2 , para comparação das medidas acústicas: *bat, pat, bop, cap, rack, back, false, pops*.

Além de o *corpus* contemplar sentenças que possuem palavras com obstruintes de diferentes pontos de articulação, decidiu-se elaborar sentenças que possuíssem 3 contextos adjacentes diferentes, para observar, como Zimmer e Alves (2007), se havia a influência dos contextos na produção das oclusivas: inexistente, vocálico e consonantal (plosivo). As sentenças-veículo utilizadas encontram-se abaixo:

- 1.Say _____.
- 2.Say _____ again.
- 3.Say _____ promptly.

Para este experimento, foi realizada uma coleta de 5 repetições de cada conjunto de sentenças, com todos os 6 informantes (5 paranaenses, entre 19 e 24 anos e 1 nativo, de 27 anos). Foram gravadas 5 repetições de cada uma das sentenças, para que não houvesse nenhum viés idiossincrático por parte dos informantes. Os dados foram colhidos em cabine com tratamento acústico no LEFON⁴, a uma taxa de amostragem de 44.100Hz. Em seguida, procedeu-se à sua análise através do *software* Praat (Boersma e Weenick, 2006).

Todas as sentenças foram recortadas no programa Praat. Após recortadas, extraímos a duração dos três parâmetros já aqui elencados. Com as medidas extraídas, foi calculada a duração relativa dos dados, para observar qual era a duração real dos segmentos através de uma média, pois, por vezes, tendo somente a extração bruta dos dados, podemos pensar que alguns dos resultados encontrados sejam, talvez, mais relevantes do que realmente são.

Um importante aspecto da análise dos segmentos é seguir um único padrão de extração de medida. Por exemplo, se consideramos que a continuação da barra de sonoridade não mais pertence à vogal, devemos medir esta a partir do término da vogal. É preciso tomar cuidado para que o dado acústico não seja observado como correspondente a uma análise prévia, mas sim como um dado acurado sobre os dados, que pode, muitas vezes, ir contra as inspeções de outiva.

Essa análise poderá ser aqui visualizada nos parágrafos seguintes. Optamos por não realizar uma análise quantitativa dos dados, pois não tínhamos um número suficiente de informantes para obter uma análise estatística, além de que este estudo preliminar teve como objetivo obter

⁴ O Lefon é o Laboratório de Estudos Fônicos da UFPR. O laboratório, localizado 9º andar da universidade Federal do Paraná, é acusticamente tratado, e é o espaço no qual são desenvolvidas pesquisas concernentes ao nível fonético-fonológico das línguas.

algumas pistas do que estava ocorrendo no sinal acústico com a produção dos falantes de inglês como L2.

3. Hipóteses

Um dos objetivos de nossa investigação, conforme relatado anteriormente, era o de realizar uma análise acústica de dados relativos à dessonorização terminal e para os quais, até então, só havia análises de outiva, como no caso dos estudos da década de 80. A análise acústica mostrou-se necessária a partir das evidências de neutralização parcial em línguas como o alemão e o holandês e de pesquisas como a de Zimmer e Alves (2007), que observou a importância de se analisar os contextos adjacentes à plosiva alvo. Diante desse quadro, elaboramos as seguintes hipóteses: 1) uma vez que o padrão silábico do português não possui oclusivas em posição final, além dos informantes realizarem epêntese, fenômeno há muito descrito na literatura, os aprendizes ensurdeceriam as plosivas sonoras, transformando-as em surdas; 2) os informantes ensurdeceriam as consoantes sonoras diante de outra plosiva surda e sonorizariam diante de um contexto vocálico, sonoro; 3) haveria uma transferência grafo-fônico-fonológica em dados como [falls] e [false] uma vez que a lateral final em inglês é pronunciada e no PB (português brasileiro) a lateral palatalizada sofreu um processo de vocalização, que leva à produção da aproximante lábio-velar [w], em posição final de sílaba ou palavra; 4) com relação ao ensino-aprendizagem, e de acordo com a abordagem dinâmica, podemos pensar que os alunos realizam produções bem próximas às de um falante nativo.

4. Análise dos dados

Uma das primeiras questões a serem analisadas é a diferença de produção que pode ser observada nos dados de um mesmo falante. Essa distinção na produção retoma o fato de que a produção do falante não é categórica. Não podemos afirmar que um dado informante produzirá sempre um mesmo som para um mesmo conjunto de dados. Na Figura 1, abaixo, temos a primeira repetição da sentença *say bag promptly*. Nessa produção, podemos notar que, na parte em destaque do espectrograma, o falante não produziu a oclusiva velar sonora [g], pois não há barra de sonoridade. Contudo, na segunda repetição dessa mesma sentença, Figura 2, o informante produz a oclusiva. Isso não só reforça a necessidade de se fazer uso da análise acústica, pois, em uma análise de outiva, talvez, um pesquisador igualasse as duas produções, de modo ou a dizer que o falante sonoriza ou não. Além disso, observamos uma oscilação no que diz respeito às produções do informante GU, mas isso não indica que ele não entenda a articulação

de sons plosivos finais, mesmo que não tenha essa produção em sua língua materna. Esse é só um exemplo. Contudo, todos os informantes, inclusive a informante nativa, realizaram variações a depender da repetição e não somente com um determinado segmento. Além dessas observações, podemos notar, também, que o contexto adjacente é uma plosiva surda, o que, de acordo com nossas hipóteses prévias e estudos anteriores, faz com que o informante produza algo diferente do esperado. Se observarmos, novamente, a Figura 2, notamos que o informante GU sonoriza mesmo diante de um segmento surdo. Com isso, começamos a endossar, aos poucos, a hipótese de que os alunos estão produzindo os segmentos-alvo de uma maneira que não se distancia das produções de um nativo como veremos em outro espectrograma.

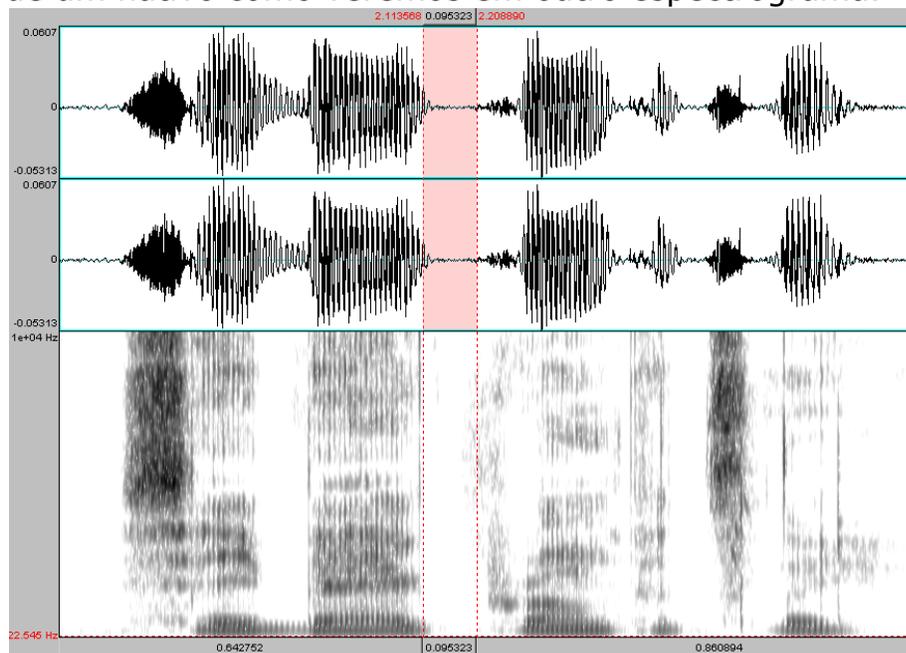


FIGURA 1 – No espectrograma, entre linhas verticais pontilhadas, a não presença de sonoridade no segmento [g] na primeira repetição da sequência *bag* na sentença “say bag promptly”, produzida pelo informante GU.

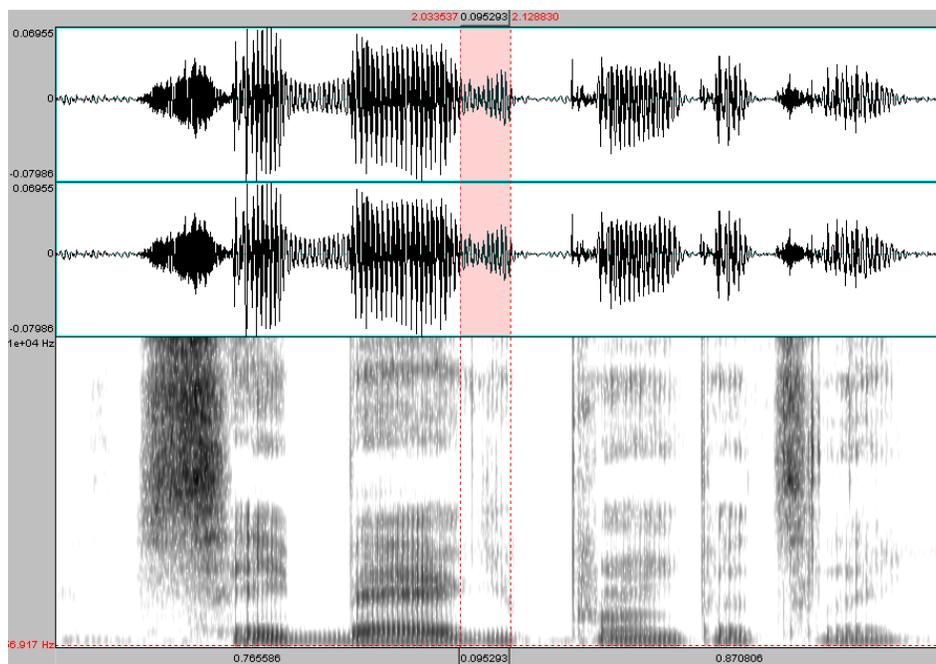


FIGURA 2: No espectrograma, entre linhas verticais pontilhadas, a presença de sonoridade no segmento [g] na segunda repetição da sequência *bag* na sentença "say bag promptly", produzida pelo informante GU.

Outro aspecto interessante é que nem sempre encontraremos a produção da oclusiva final, seja ela surda ou sonora. Em uma inspeção de outiva, um falante nativo poderia dizer que o aprendiz produziu ou [t] ou [d], mas, como apresentado nos espectrogramas 3 e 4, abaixo, há uma variação entre a visualização da barra de sonoridade, característica de sons sonoros e do *burst*⁵, parâmetro acústico das oclusivas. Na figura 3, não observamos a presença da barra de sonoridade, nem do *burst*, o que indica que não temos a produção de uma oclusiva. Já na Figura 4, observamos o *burst* e um pouco da barra de sonoridade, mas, dada a duração desse parâmetro, podemos interpretar tal barra como a continuação da vogal. Mais uma vez, se não tivéssemos utilizado a análise acústica, por conta da sonoridade do contexto adjacente, uma vogal, poderíamos dizer que o informante teria sonorizado em 3, quando, de fato, ele nem chegou a produzir a consoante. Além disso, temos, também, que lançar um olhar crítico para o percentual de sonorização nesta e em outras pesquisas. Quando obtemos um percentual X para a não sonorização, não quer dizer, necessariamente, que o informante não sonorizou. Pode ser que ele não tenha realizado o segmento. Cabe aí investigar o que o falante está entendendo como sonorização da consoante oclusiva. Como

⁵ O *burst* é a "explosão" de ar promovido pela soltura dos articuladores, durante a realização de uma oclusiva. No espectrograma, ele é caracterizado por uma barra vertical.

mencionamos anteriormente, não vamos entrar, aqui, em uma discussão aprofundada sobre aspectos perceptuais ou a percepção de sons em uma L2, mas, se pensarmos um pouco em como esse dado se apresenta para um ouvinte qualquer, este poderia dizer que ambas as produções, das Figuras 3 e 4, são iguais ou pensar no que mencionam Port e Leary (2005) sobre o nosso conhecimento estar pautado em um conjunto de símbolos alfabéticos. Nesse sentido, quando ouvimos algo como [rag] ou [ra], mas sabemos que [ra] não se configura como uma palavra possível em uma dada língua alvo, em nosso caso, o inglês, o falante preenche tal som com o que deveria ser, com o que sabe que graficamente é. É preciso notar, também, que a produção e a percepção, como menciona Port (1996), podem estar conectadas através da ideia de expectativa vs. realização. De acordo com o autor, é o que a literatura costuma chamar de “*responses bias*”. Port relata que, à medida que os informantes são apresentados a um determinado conjunto de dados e, se esses parecem ter uma complexidade alta, os informantes ativarão outro processo para fazer a distinção ou equiparação. Uma das questões extremamente interessantes que surgem a partir disso é que a decisão dos sujeitos por uma das produções pode ser afetada por diversos fatores, inclusive a expectativa destes antes mesmo de escutarem o conjunto de dados.

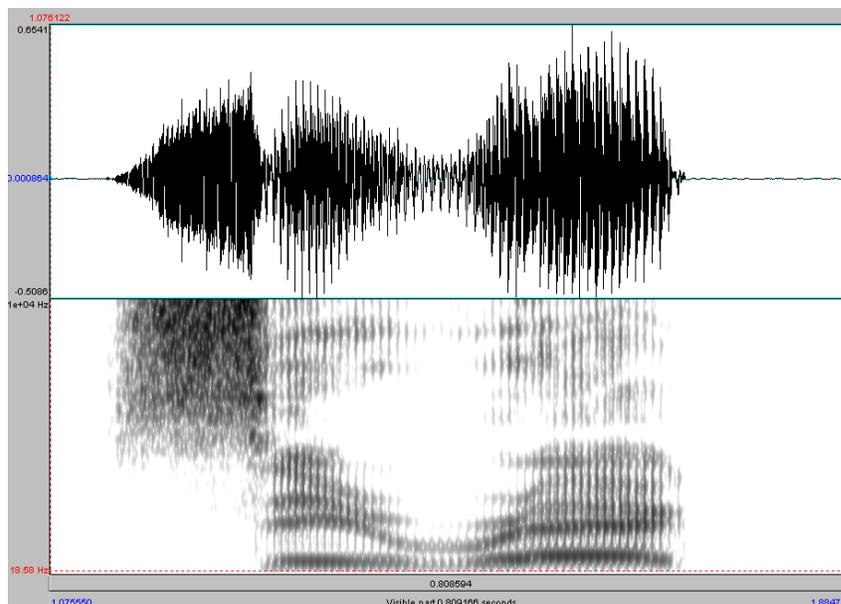


FIGURA 3 – O espectrograma da quinta repetição do segmento *rag*, sem a produção do *burst* final da sentença “say rag”, produzida pelo informante JU.

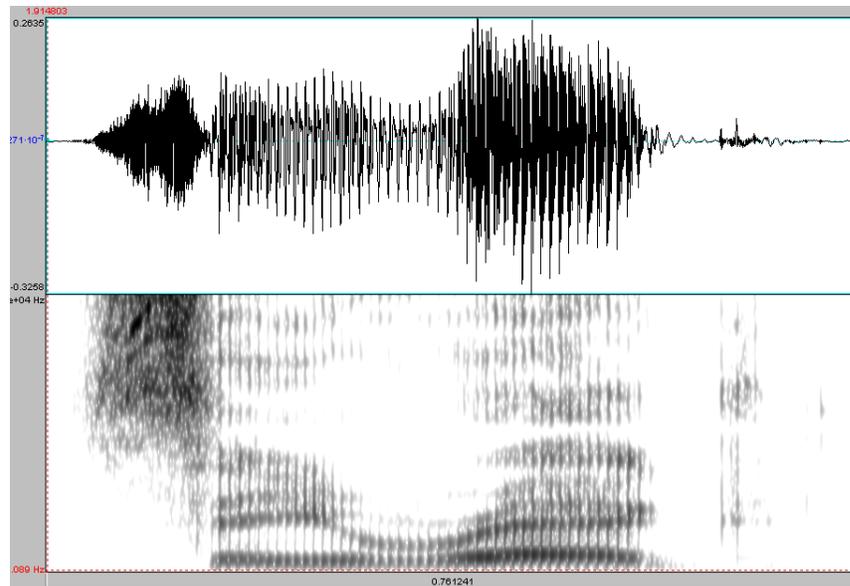


FIGURA 4 – O espectrograma da primeira repetição da palavra *rag*, na sentença “say rag”, produzida pelo informante JU.

Ainda mais interessante é notar que na produção da informante nativa PM (Figura 5) não observamos a presença do *burst*. A informante produziu um [b]; sabíamos que a sentença tinha um segmento *bag*, porém, como não há presença do *burst*, podemos caracterizar esse segmento como uma oclusiva? Sim, podemos. A *conditio sine qua non* para a caracterização de uma plosiva é a oclusão. O *burst* pode, muitas vezes, ser realizado pela ação tão tênue dos articuladores, que o som produzido é muito pouco intenso e, por isso, não captado pelo microfone durante a gravação. Apesar de notarmos a barra de sonoridade, sem a presença do *burst*, temos um som sonoro, que não é a vogal adjacente, pois não notamos a presença de formantes. A pista para a visualização da oclusiva está na transição da vogal que a precede. É, inclusive, através desse aspecto, que é possível recuperar a informação sobre o ponto de articulação da consoante. O parâmetro acústico obrigatório para caracterizar um segmento como oclusivo é o espaço em branco, correspondente à oclusão. Contudo, de acordo com Kent e Read (1992), em final de palavra, especialmente, é comum que o *burst* não seja realizado. Nesse sentido, podemos aproximar a produção de PM do informante paranaense JU.

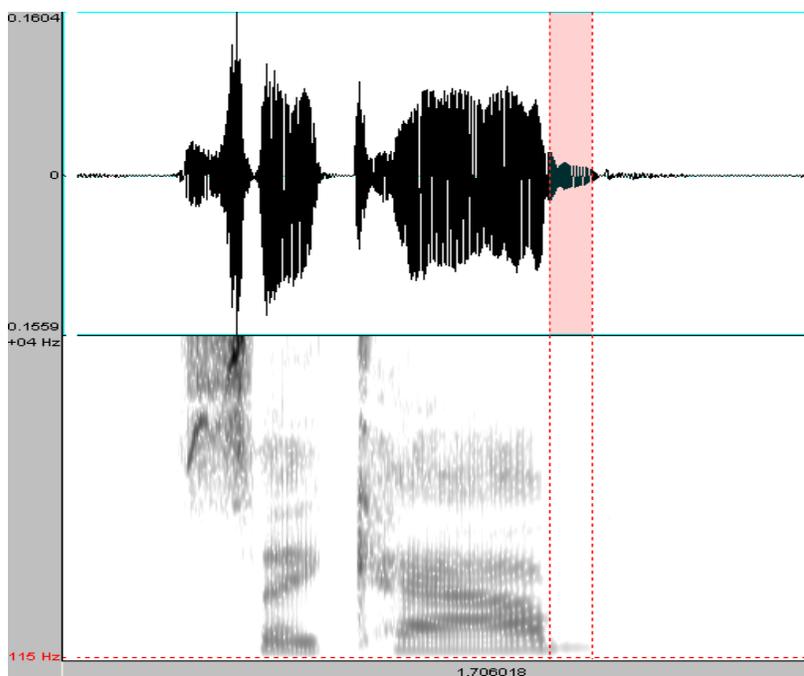


FIGURA 5 – No espectrograma, entre linhas verticais pontilhadas, a barra de sonoridade e a não produção do *burst*, no segmento [g], na primeira repetição da sentença “say cab”, produzida pelo informante PM (nativa).

No que diz respeito à transferência grafo-fônico-fonológica, ela ocorreu com todos os informantes aprendizes. As palavras [falls] e [false] foram produzidas da mesma maneira no que diz respeito à produção da fricativa final como surda, em ambas as produções. Contudo, foi interessante observar que a maioria dos informantes tentou deslocar a distinção final para a vogal e as vogais foram produzidas com aberturas diferentes. Para a produção de [falls], os alunos realizaram uma vogal aberta e para [false], fechada. Mesmo que notemos, aqui, a presença da influência dos sons da L1 na L2, observamos, também, que o aprendiz marca as diferenças utilizando-se de outros parâmetros acústicos.

Passemos, então, a algumas informações, reunidas em gráficos, sobre os parâmetros acústicos. Novamente, gostaríamos que deixar claro que os gráficos a seguir servem, apenas, como uma inspeção visual de alguns parâmetros.

Os primeiros resultados relativos à duração da vogal diante de oclusivas sonoras e surdas apontam para um aspecto interessante: a duração da vogal é maior diante de oclusivas anteriores. Em todos os gráficos, para surdas e sonoras observamos que quando a consoante é velar – [k] e [g] – a vogal é mais longa. Temos aí um fator que influencia a duração da vogal: ponto de articulação (ver gráficos abaixo). Também temos uma diferença de percentual de duração para os três ambientes. A duração é maior para o primeiro ambiente, no qual

não há contexto adjacente à direita. Isso pode indicar que, quando não há a presença de um contexto adjacente, os informantes tendem a alongar a vogal para ressaltar a produção. É interessante, também, destacar que, apesar de possuímos contextos adjacentes distintos, (Figura 7 e 8), nos quais se tem uma vogal e uma oclusiva, o percentual de duração de vogal não é discrepante.

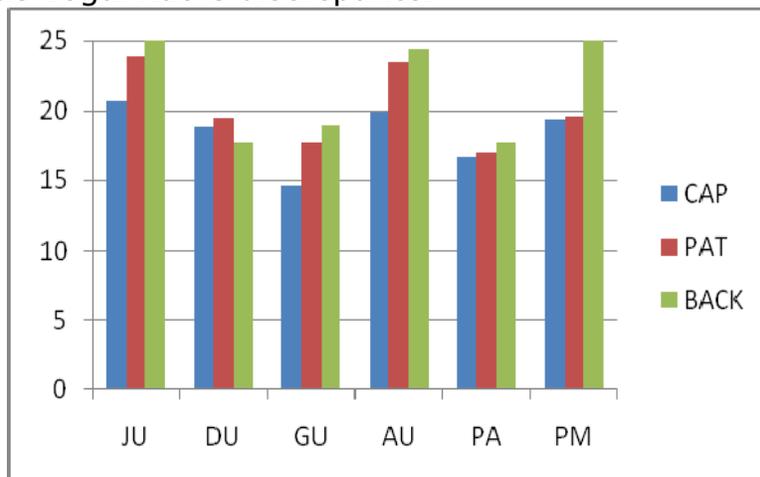


FIGURA 6 – Gráfico correspondente ao percentual de duração de vogal para o contexto 1, consoantes oclusivas finais surdas, com contexto adjacente inexistente.

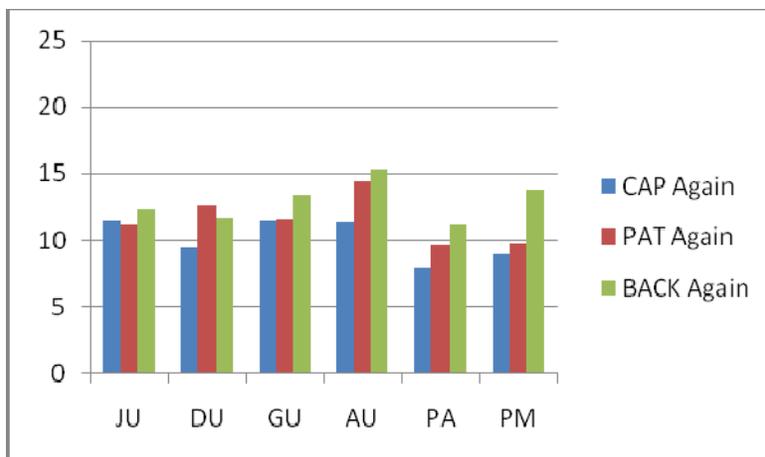


FIGURA 7 – Gráfico correspondente ao percentual de duração de vogal para o contexto 2, consoantes oclusivas finais surdas, com contexto vocálico adjacente.

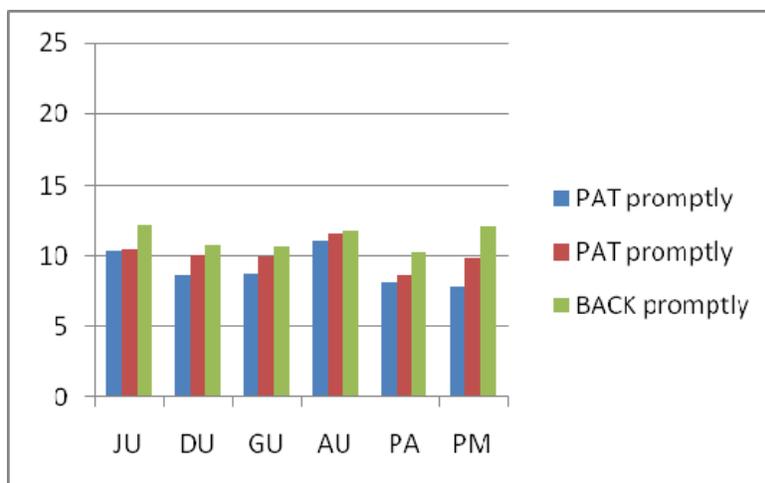


FIGURA 8 – Gráfico correspondente ao percentual de duração de vogal para o contexto 3, consoantes oclusivas finais surdas, com contexto consonantal adjacente.

Para o parâmetro vozeamento, Figuras 9, 10 e 11, observamos que o ponto de articulação também parece exercer influência. Embora a questão não pareça tão resolvida para consoantes sonoras alveolares e velares, pois para alguns informantes, o percentual de vozeamento é o mesmo ou a diferença é muito pequena, para a bilabial sonora [b] observa-se um maior índice de vozeamento.

Em alguns pontos do gráfico, como, por exemplo, em sentenças "say bad again", os informantes paranaenses, de modo mais abrangente em relação à nativa, realizam um *flap*, movimento que resulta na produção de um rótico e não de uma oclusiva, o que visualmente é perceptível por um decréscimo na taxa de vozeamento. Visualmente, quando temos um *flap*, não há uma interrupção no sinal acústico. Esse tipo de fenômeno – denominado "flapping" na literatura relativa à fonologia do inglês (cf., e.g., Fox e Terbeek, 1977) – ocorre porque o rótico e as oclusivas [t] e [d] são articulados no mesmo ponto. A ocorrência do flapping suscita uma questão relativa à maneira como tal fato se reflete nos resultados finais de uma análise. Dizemos que o informante realizou uma oclusiva sonora, pois há a presença de barra de sonoridade, mesmo sem haver a oclusão? Existe uma oclusão na ocorrência do *flap*, mas esta é distinta da oclusão da plosiva. Mais uma vez, temos que estar atentos para os experimentos, pois várias informações não são perceptíveis com a leitura de um gráfico. Nos gráficos, a presença do *flap* entra como se o informante não tivesse produzido a oclusão, mas sim como se tivesse sonorizado.

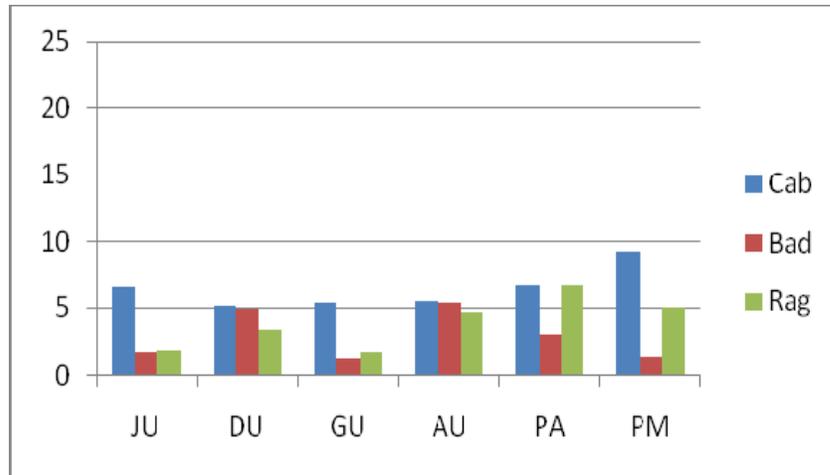


FIGURA 9 – Gráfico correspondente ao percentual de vozeamento nas consoantes oclusivas finais sonoras, com contexto adjacente inexistente.

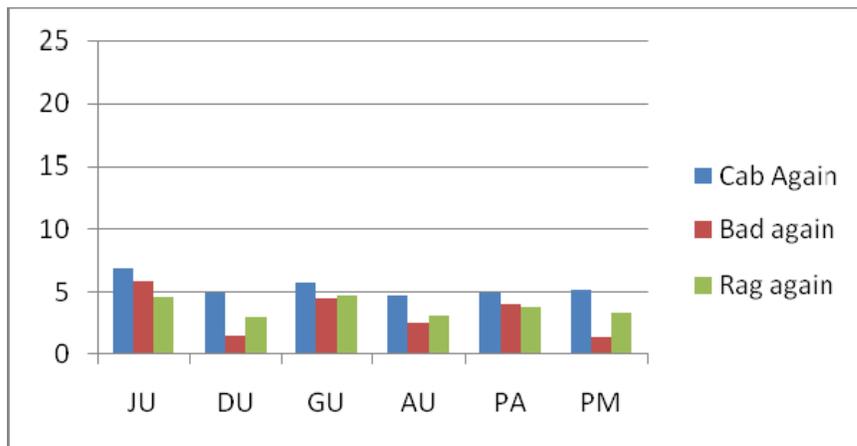


FIGURA 10 – Gráfico correspondente ao percentual de vozeamento nas consoantes oclusivas finais sonoras, com contexto vocálico adjacente.

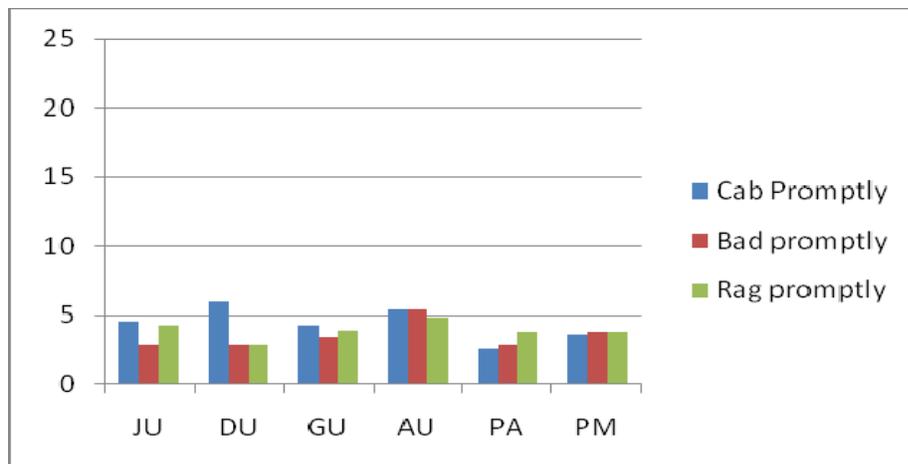


FIGURA 11 – Gráfico correspondente ao percentual de vozeamento nas consoantes oclusivas finais sonoras, com contexto consonantal adjacente.

Além do que foi mencionado no parágrafo anterior, os contextos adjacentes influenciam no vozeamento: segundo relatos da literatura, quando o contexto adjacente é uma vogal, o segmento anterior tende a sonorizar. Além de se observar tal fato, verificou-se que em outros contextos, mesmo diante de oclusiva (como na sentença 3 da página 5), os informantes continuam sonorizando.

5. Conclusões

Ainda que este seja um estudo preliminar, acreditamos que pudemos lançar algumas pistas para os estudos sobre DT, bem como para o processo de ensino-aprendizagem.

Na seção 4, lançamos algumas hipóteses, as quais foram parcialmente corroboradas e refutadas. No que diz respeito ao fato de que os alunos iriam dessonorizar segmentos sonoros diante de contextos que favorecessem o processo (ex: diante de oclusiva surda) não verificamos tal hipótese. As análises visuais, baseadas nos espectrogramas, apontam para o fato de que, mesmo em contexto favorecedor de dessonorização, os falantes continuam sonorizando. Verificamos, porém, a existência de uma transferência grafo-fônico-fonológica em dados como [falls] e [false]. No entanto, essa transferência deixou marcas nas produções. Como mencionamos anteriormente, talvez uma maneira de os informantes fazerem a distinção perceptual das duas palavras tenha sido transferir a diferença para a vogal e não para a fricativa final. Além disso, pôde ser observado que as produções dos alunos não se mostram desviantes quando comparadas com os dados de um falante nativo. As produções possuíam percentuais diferentes e como não elaboramos nenhuma análise estatística, não é possível assegurar se houve uma diferença significativa. Os gráficos, no entanto, dão-nos uma mostra de que há distinção na produção de estrangeiros e nativos. Essa produção não está relacionada a um determinado segmento, mas, muito provavelmente, à magnitude dos gestos articulatórios envolvidos no processo de sonorização de um som.

No tocante às questões de ensino-aprendizagem, é importante destacar que além dos fatores linguísticos relacionados à L1 e L2, os aprendizes têm de lidar com fatores extra-linguísticos e, também, metalinguísticos. E, talvez, nesse sentido, pensar nesse processo como um sistema dinâmico faça a diferença, uma vez que esses fatores extra-linguísticos são difíceis de serem controlados pelo experimentador, mas

são relevantes, pois são reveladores de informações sobre as produções dos informantes.

Em sala de aula, o professor não faz uso de ferramentas como a análise acústica para visualizar a produção dos alunos e, assim, sinalizar-lhes o quanto essas produções aproximam-se ou distanciam-se das produções dos falantes nativos. Aliás, não faz sentido dar à análise acústica um uso semelhante a esse, como se os espectrogramas fossem a fotografia da fala do sujeito e como se houvesse padrões acústicos "certos" ou "errados" para a produção de um som ou outro. A análise acústica deve ser ponto de partida para as reflexões dos professores acerca da produção de seus alunos – o que produzem e por quê? Dessa forma, o professor poderá focalizar sua prática didática nas reais dificuldades dos alunos, o que certamente poderá promover uma mudança na produção de alguns sons realizados por tais alunos. No entanto, uma vez ciente das tentativas dos alunos em produzir sons de uma L2, os professores podem trazer, para a aula, reflexões sobre estas de maneira que o aluno possa perceber, aos poucos, uma mudança na produção de alguns sons.

REFERÊNCIAS BIBLIOGRÁFICAS

- BERNSTEIN, J. New Uses for Speech Technology in Language Education. *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Marholmen Conference Center, Sweden, 1998.
- BOERSMA, Paul; WEENICK, David. Praat: Doing phonetics by computer (Versão 4.4.03), 2006. Disponível em: <<http://www.praat.org/>>. Acesso em: 20.dez.2011.
- BROWMAN, C.; GOLDSTEIN, L. Articulatory Phonology: an overview. *Phonetica*, 49:155-180, 1992.
- BROWN, H.D. Cognitive and affective characteristics of good language learners. Paper presented at *Los Angeles Second Language Acquisition Research Forum*, UCLA, February, 1977.
- CHOMSKY, N.; HALLE, M. *The Sound Pattern of English*. Cambridge, MA: The MIT Press, 1968.
- DAVIES, S. e POESIO, M. A CSLUrp-Based Spoken Dialogue System for Teaching English as a Foreign Language. In: *Proceedings of the European Speech Community Association on Speech Technology in Language Learning*. Marholmen Conference Center, Suécia, 1998.
- DEVLIEGER, M. The Applicability of Speech Recognition in the Context of Task-Based Language Learning for Young Children. In: *Proceedings of the European Speech Community Association on Speech Technology in Language Learning*. Marholmen Conference Center, Suécia, 1998.
- ECKMAN, F.R. On the naturalness of interlanguage phonological rules. *Language Learning*, 31:195-216, 1991.

- _____.; FRED, R. Markedness and the contrastive analysis hypothesis. In: IOUP, Georgette; WEINBERGER, Steven H. *Interlanguage phonology: the acquisition of a second language sound system*. Newbury, 1987.
- ELMAN, J. Connectionism, artificial life, and dynamic systems. In: BECHTEL, W.; GRAHAM, G. (eds.) *A Companion to Cognitive Science*. Oxford: Basil Blackwood, 1998.
- ERNESTUS, Miryam; BAAYEN, Harald. The functionality of incomplete neutralization in Dutch: the case of past-tense formation. In: GOLDSTEIN, Louis et al. (eds) *Laboratory Phonology 8*. Berlin: De Gruyter, 2006.
- ESCUADERO, P. *Linguistic Perception and Second Language Acquisition*, Vol. 113. Utrecht: LOT, 2005.
- FLEGE, J.E. The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15:47-65, 1987.
- FOX, Robert A.; TERBEEK, Dale. Dental flaps, vowel duration, and rule ordering in American English. *Journal of Phonetics*, 5:27-34, 1977.
- KENT, Ray; READ, Charles. *The Acoustic Analysis of Speech*. San Diego: Singular Publishing Group Inc., 1992.
- KRASHEN, S. *Principles and Practice in Second Language Acquisition*, Pergamon, 1982.
- MAJOR, Roy. A model for interlanguage phonology. In: IOUP, Georgette. WEINBERGER, Steven H. *Interlanguage phonology: the acquisition of a second language sound system*. Newbury, 1987, p. 101-124.
- PIROTH, Hans Georg; JANKER, Peter. Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, v. 32:81-109, 2004.
- PORT, R.F. The discreteness of phonetic elements and formal linguistics: response to A. Manaster Ramer. *Journal of Phonetics*, 24(4):491-511, 1996.
- PRICE, P. How can speech technology replicate and complement good language teachers to help people learn language? Proceedings of *ESCA Workshop on Speech Technology in Language Learning*. Sweden: Marholmen Conference Center, 1998.
- ROCCA, Paulina D. Artimonte. A Tecnologia de fala aplicada ao ensino de entoação da língua inglesa para falantes nativos de língua Portuguesa. Tese (Doutorado em Linguística Aplicada e Estudos da Linguagem). São Paulo: Pontifícia Universidade Católica de São Paulo, 2003.
- SANCIER, Michele L.; FOWLER, Carol A. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. In: *Journal of Phonetics*, v. 25:421-436, 1997.

Albuquerque, Jeniffer Imaregna Alcântara de; Silva, Adelaide Hercília Pescatori. Um olhar dinâmico para questões de produção de fala em L2: Estudo preliminar sobre a dessonorização terminal (DT). *Revista Intercâmbio*, v. XXIII: 106-127, 2011. São Paulo: LAEL/PUCSP. ISSN 2237-759X.

SILVA, A.H.P. Proposta de representação dinâmica para os róticos iniciais do PB. *Língua Viva* (Guajará-Mirim), EDUFRO - Guajará-Mirim, v. 02:25-30, 2003.

STEVICK, E.W. *Memory, Meaning and Method*. Rowley, MA: Newbury House, 1976.

TOWNSHEND, B., BERNESTEIN, J., TODIC, O. & E. WARREN. Estimation of Spoken Language Proficiency. *Proceedings of ESCA Workshop on Speech Technology in Language Learning - Marholmen Conference Center*. Sweeden, 1998.

Van GELDER, T.; PORT, R. It's about time. In: PORT, R.; Van GELDER, T. (eds.). *Mind as motion*. Cambridge, MA: MIT, 1995.

Van OOSTENDORP, Marc. Incomplete devoicing in formal phonology. *Lingua*, 118(9): 1362-1374, doi:10.1016/j.lingua, 2007.

ZIMMER, M.C.; ALVES, U.K. A dessonorização terminal na aprendizagem da L2: evidências do continuum fonética-fonologia. *Letras de Hoje*. Porto Alegre, v. 42, n. 3:56-68, set/2007.