

ANÁLISE DE FENÔMENOS COARTICULATÓRIOS EM INGLÊS COM BASE NA FONOLOGIA ARTICULATÓRIA

Coarticulatory phenomena analysis in English based on the articulatory phonology

Amaury Flávio SILVA (Faculdade de Tecnologia de Jacareí, Jacareí, Brasil)

RESUMO: *Durante o processo de produção de fala, os falantes nativos de inglês tendem a coarticular os segmentos fônicos que compõem as palavras, dando origem a uma redução no movimento dos articuladores. Essa redução altera a produção dos segmentos fônicos gerando dificuldades de percepção por aprendizes brasileiros de inglês. Além disso, tais aprendizes hiperarticulam segmentos que geralmente são reduzidos por falantes nativos, o que acarreta um maior esforço articulatório e dificulta a produção dos segmentos. Devido a esses fatores, produzimos um estudo para investigar os fenômenos coarticulatórios presentes na produção de fala em língua inglesa com base na fonologia articulatória. O estudo revelou que o falante nativo de inglês investigado utilizou uma estratégia caracterizada por não incluir pausa entre as palavras nos contextos analisados e por fazer uso das estratégias denominadas de “blending” e “hiding.”*

PALAVRAS-CHAVE: Coarticulação; Percepção; Produção; Fonologia articulatória

ABSTRACT: *During the speech production process, native speakers of English tend to coarticulate the phonemic segments that comprise the words so that a reduction in the movement of the articulators takes place. Such reduction modifies the phonemic segment production causing perception difficulties by Brazilian learners of English. Furthermore, these learners hyperarticulate segments that are usually reduced by native speakers, which causes a greater articulatory effort and hampers the segment production. Due to these factors, we have carried out a study to investigate the coarticulatory effects that exist in the English language speech production based on the articulatory phonology. The study revealed that the native speaker of English used a strategy characterized by not including a pause between the words in the contexts which were analyzed and by using the strategies called blending and hiding.*

KEYWORDS: Coarticulation; Perception; Production; Articulatory phonology

INTRODUÇÃO

A produção de fala concatenada possui características que fazem com que a pronúncia canônica dos segmentos fônicos que compõem as palavras seja alterada pelo contexto fonético. Nesse sentido, uma palavra como “can’t”, /kænt/, quando produzida em um contexto como “I can’t take it” pode sofrer um processo de redução (SILVA, 2016) de modo que a soltura da articulação da constrição do articulador da consoante /t/ poderá não ocorrer por ela ser coproduzida com a consoante homorgânica /t/ da palavra “take.”

A alteração dos segmentos fônicos, como a descrita acima, facilita o fluxo da fala devido à economia de movimento dos articuladores. Entretanto, isso pode acarretar dificuldades de percepção por aprendizes brasileiros de língua inglesa (ABLI). Além disso, ABLI tendem a hiperarticular os segmentos que compõem as palavras. Sendo assim, na produção de “it’s important that”, as consoantes adjacentes /t/ da palavra “important” e /ð/ da palavra “that”, por compartilharem o mesmo articulador e possuírem locais de constrição distintos, seriam produzidas com maior esforço articulatorio, prejudicando o fluxo da fala.

Os ABLI, de acordo com uma pesquisa realizada por Silva (2009), também apresentam um hábito articulatorio que consiste na inserção de uma vogal, tal como /i/, em contextos como “bed as”, /bed əz/, ou seja, tais falantes produzem [bedi əz]. Outro hábito articulatorio identificado na pesquisa está relacionado à inclusão de pausa em contextos caracterizados por palavras que finalizam em consoante plosiva alveolar, /t, d/, e que são sucedidas por palavras que iniciam em consoantes que possuem as mesmas características, como no contexto “almost daily”.

Esses hábitos diferem das estratégias utilizadas por falantes nativos do inglês uma vez que eles tendem a não inserir vogais em palavras finalizadas por consoantes e, além disso, frequentemente coarticulam consoantes homorgânicas imediatamente adjacentes.

Entender os processos de alteração dos segmentos é importante para o desenvolvimento da percepção (FIELD, 2008), além de proporcionar uma produção de fala mais inteligível (REED; MICHAUD, 2005). Além disso, Alameen e Levis (2015) afirmam ser necessário que falantes não nativos de inglês compreendam a produção de fala de nativos, especialmente, no que tange à produção não canônica.

Devido à importância da compreensão das características que abrangem a fala concatenada, desenvolvemos uma pesquisa que analisa os fenômenos coarticulatorios com base na Fonologia Articulatória (FAR) de Browman e Goldstein (1986; 1989; 1990a, b; 1992)

Esperamos construir conhecimento sobre os fenômenos coarticulatorios que ocorrem na fala concatenada que servirão como subsídios para o desenvolvimento não apenas da produção, mas da percepção de fala em língua inglesa por ABLI.

Na próxima seção, trataremos das teorias nas quais as análises deste trabalho foram baseadas.

I. FUNDAMENTAÇÃO TEÓRICA

Devido aos diferentes processos coarticulatórios existentes na produção de fala, surgiram teorias e modelos que tratam da coarticulação. As teorias sobre a coarticulação segundo Farnetani e Recasens (1999, p. 31) têm o propósito de explicar origem, natureza e função da coarticulação, além de explicar como ouvintes conseguem perceber variantes articulatórias e recuperam a mensagem subjacente. Os modelos coarticulatórios tendem a prever os detalhes do processo de ligação entre as unidades distintas e invariantes de representação para a articulação e a acústica.

Apesar da existência de diversas teorias e modelos coarticulatórios, havia uma lacuna entre os aspectos físicos e cognitivos da linguagem até o surgimento da Fonologia Articulatória (FAR), elaborada por Browman e Goldstein (1986; 1989; 1990a, b; 1992). Tal teoria surgiu como uma nova abordagem fonológica com o intuito de tratar de questões relacionadas à linguagem. Essa abordagem serviu como uma ponte que uniu os aspectos físicos e cognitivos da linguagem, os quais não haviam sido contemplados por teorias como a gerativa (CHOMSKY, 1985).

A fonologia articulatória parte da hipótese de que os dois domínios físico e cognitivo fazem parte da descrição dimensional de um sistema único e complexo.

Desta forma, na FAR, as unidades fonológicas abarcam ações articulatórias específicas, denominadas de “gestos”. Em outras teorias, as estruturas linguísticas da fala eram analisadas por uma sequência de unidades segmentais que correspondiam a uma característica de um todo.

Teorias que antecederam a FAR, como a gerativa (CHOMSKY, 1985), foram criticadas por fonologistas e foneticistas: aqueles defendiam a visão de que as restrições impostas por sequências lineares de segmentos não sobrepostos eram demasiadamente extremas para explicar a variabilidade de fatos fonológicos: estes defendiam a relevância linguística da estrutura temporal detalhada da fala (BROWMAN; GOLDSTEIN, 1986). Por exemplo, a organização interarticulatória temporal pode variar entre as línguas de maneira que não podem ser previstas por princípios universais. Sendo assim, devem ser especificadas em descrição de línguas.

Devido a esses pontos de vista, surgiu a necessidade de uma revisão de estruturas fonético/fonológicas para que unidades fonológicas que se sobrepõem fossem implantadas permitindo que relações temporais entre estruturas articulatórias surgissem desta descrição (BROWMAN e GOLDSTEIN, 1986).

Na fonologia articulatória as unidades de produção de fala são ações dinâmicas cujas funções fonológicas, como as de contraste, estão em um nível dimensional inferior. Ao que concerne as unidades básicas entre a articulação e a acústica estas não são neutras, mas articulatórias por natureza. Portanto, na fonologia articulatória, a unidade fonológica básica é o gesto articulatório, o qual é definido como um sistema dinâmico específico com uma característica de ajuste de parâmetros de valores (BROWMAN e GOLDSTEIN, 1986).

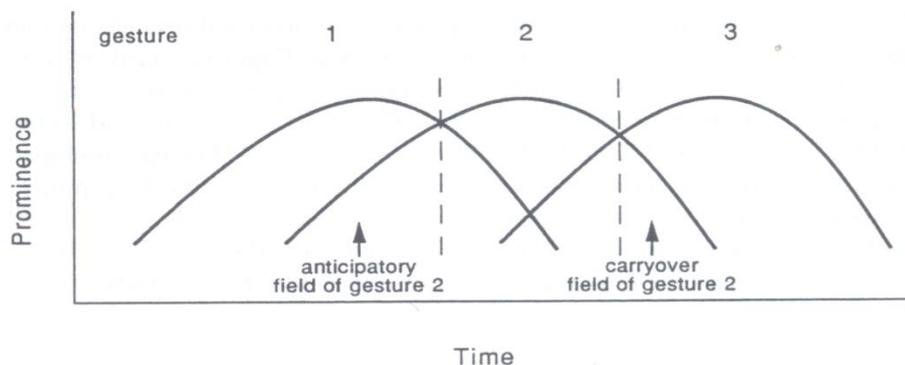
Nessa teoria, pelo fato de as tarefas serem distribuídas entre vários ajustes dos articuladores do trato vocal (lábios, língua, glotes, véu palatino etc.), um enunciado é modelado como um conjunto, ou constelação, de um número pequeno de unidades gestuais potencialmente sobrepostos.

Para Browman e Goldstein (1989, p. 202), os gestos articulatórios como unidades fonológicas podem ser identificadas pela observação dos movimentos coordenados pelos articuladores do trato vocal. Nessa abordagem, gestos são as unidades básicas das quais estruturas fonológicas são formadas. Além disso, segundo Browman e Goldstein (1989) e Albano (2001) o gesto articulatório é uma oscilação abstrata através da qual as constrictões no trato vocal são especificadas de modo a induzir o movimento dos articuladores.

O gesto articulatório não é correspondente a um segmento ou a um traço, de modo a estar localizado em uma região intermediária entre este e aquele. Um [p] envolve somente um gesto de oclusão labial, enquanto um [b] envolve a) um gesto de oclusão labial e b) um gesto de adução das pregas vocais. Um [m] assim como um [b] engloba os mesmos gestos mais um gesto de abertura do véu palatino para a produção de tal som nasal.

Por meio da análise da organização gestual (vide figura 1), é possível explicar como ocorre a sobreposição entre os gestos no tempo.

Figura 1. A organização espaço temporal dos gestos



Fonte: Browman e Goldstein (1989)

A figura 1 nos mostra que a proeminência dos gestos aumenta ou diminui no tempo. Dessa forma, quanto maior a proeminência de um certo gesto, maior será sua influência no ajuste do trato vocal.

As linhas pontilhadas delimitam o intervalo de tempo no qual o gesto 2 é proeminente de maneira a ser menos influenciado pelos gestos 1 e 3. Antes desse intervalo, a influência do gesto 1 é soberana, o que gera a denominada *anticipatory coarticulation* (nível antecipatório de coarticulação). Posterior ao intervalo de

proeminência do gesto 2, temos o gesto 3, o qual é influenciado pelo gesto anterior de maneira a gerar a denominada *carryover coarticulation* (nível posterior de coarticulação).

Tanto o nível de coarticulação antecipatório quanto o nível de coarticulação posterior abrangem o princípio de sobreposição gestual.

A sobreposição entre os gestos reflete suas coordenações temporais. Nessa visão, é levado em conta o fato de que o *onset* de um gesto ocorre em um estágio específico de ciclo do gesto ativo precedente. Ou seja, em uma palavra como *got* (consideramos /g/ como gesto 1, /o/ como gesto 2 e /t/ como gesto 3) em algum momento durante a produção do gesto 1 o gesto 2 passou a ser acionado. Assim, na evolução espaço-temporal, conforme o nível de proeminência do gesto 1 tornar-se menos evidente, mais evidente tornar-se-á o nível de proeminência do segundo gesto, e assim sucessivamente.

O estágio entre os gestos é controlado no nível do plano: um aumento na sobreposição dos gestos aumenta a quantidade de efeitos coarticulatórios. As variações no nível de sobreposição gestual se realizam de acordo com o estilo da fala (fala rápida, fala alta, fala cuidada etc.).

O fato de que os gestos possuem uma extensão temporal de maneira a se sobreporem a outros gestos dá possibilidade a uma série de consequências fonético-fonológicas (BROWMAN; GOLDSTEIN, 1989). Desse modo, para tais autores a sobreposição gestual pode gerar variação contextual de diferentes tipos: a) a não existência de invariância acústica, como diferentes transições de formantes resultantes da sobreposição de um gesto consonantal invariante sobre diferentes gestos vocálicos; b) variação alofônica, tais como vogais nasalizadas produzidas pelas sobreposições entre a abertura do véu palatino em um gesto final e o gesto vocálico; e, c) vários tipos de fenômenos articulatórios dependentes do contexto (214).

Ao que concerne à grande variância entre a pronúncia canônica e a pronúncia na fala fluente Brown (1977); e Shockey (1974) argumentam que essas alterações de pronúncia resultam a) na redução da magnitude de gestos específicos (no tempo e espaço) e b) no aumento da coprodução intergestual.

Para exemplificar o que ocorre em contextos como “the cat died”, no qual a consoante /t/ da palavra *cat* pode ser coproduzida com a consoante /d/ da palavra seguinte, Browman e Goldstein (op. cit.) propõem a observação do que ocorre quando o nível de coarticulação é alterado pelo aumento de sobreposição entre dois gestos de constrição oral. O fruto de tal aumento produzirá diferentes consequências no resultado acústico e articulatório do modelo dependendo se os gestos “contemplam as mesmas ou diferentes variáveis do trato” (BROWMAN; GOLDSTEIN, 1989, p. 214). Se os gestos contemplarem articuladores diferentes, o resultado na variável do movimento do trato não será alterado.

Quando há um alto nível de sobreposição um gesto pode tornar-se inaudível. Browman e Goldstein (op. cit.) referem-se a tal fenômeno como *gestual hiding*.

Em um contexto como “ten things”, no qual as consoantes /n/ e /θ/ das palavras “ten” e “things”, respectivamente, estão na mesma camada da pauta gestual e que são associados a gestos de constrição de ponta de língua, eles não podem sofrer sobreposição sem que haja perturbação na variável do movimento do trato. Em tal exemplo, em que os gestos possuem tarefas distintas utilizando o mesmo articulador, ocorre conforme o fenômeno de *blending* (SALTZMAN *et al.*, 1988).

Fenômenos da fala, como o apagamento de segmentos plosivos alveolares finais em encontros consonantais (*clusters*), assimilação de segmentos plosivos alveolares nasais que ocorrem em posição final seguidos por segmentos labiais e velares e assimilação de sons plosivos alveolares em posição final para consoantes que têm como ponto de constrição a ponta da língua podem ser considerados como o fruto do aumento da sobreposição entre gestos na fala fluente.

II. METODOLOGIA

A seguir, apresentaremos o *corpus* utilizado nas investigações, o perfil do sujeito cuja produção de fala foi analisada e apresentaremos os critérios de inspeção acústica dos dados.

II.1 *Corpus* de pesquisa

Com o intuito de realizar as análises, selecionamos um fragmento de uma entrevista pertencente ao programa *Inside the Actors' Studio*, a qual foi lançada em 9 de julho de 2006.

Apresento, a seguir, o fragmento da entrevista que foi selecionado:

Quadro 1. *Corpus* de pesquisa

It's funny, you know, someone comes into your life at a certain time and one of the great things that happens on earth is you're mysteriously guided towards these people that you get to dance with, you know. And I thought "How great is that", he's kind of, like, I don't want to say an angel to her, but he's someone who needs as much as he's prepared to offer, and he has seen a lot of life, and he's not a typical lawyer-type.

Fonte: Elaborado pelo autor

O *corpus* apresentado acima deu origem a 6 contextos de análise, os quais apresentamos a seguir:

- 1) *certain time*: nasal alveolar seguida por plosiva alveolar
- 2) *guided towards*: plosiva alveolar vozeada seguida por plosiva alveolar não vozeada
- 3) *prepared to*: plosiva alveolar vozeada seguida por plosiva alveolar não vozeada

- 4) *these people*: fricativa alveolar seguida por plosiva bilabial
- 5) *as much*: fricativa alveolar seguida por nasal bilabial
- 6) *I don't want to*: plosiva alveolar seguida por aproximante bilabial

II. 2 Sujeito

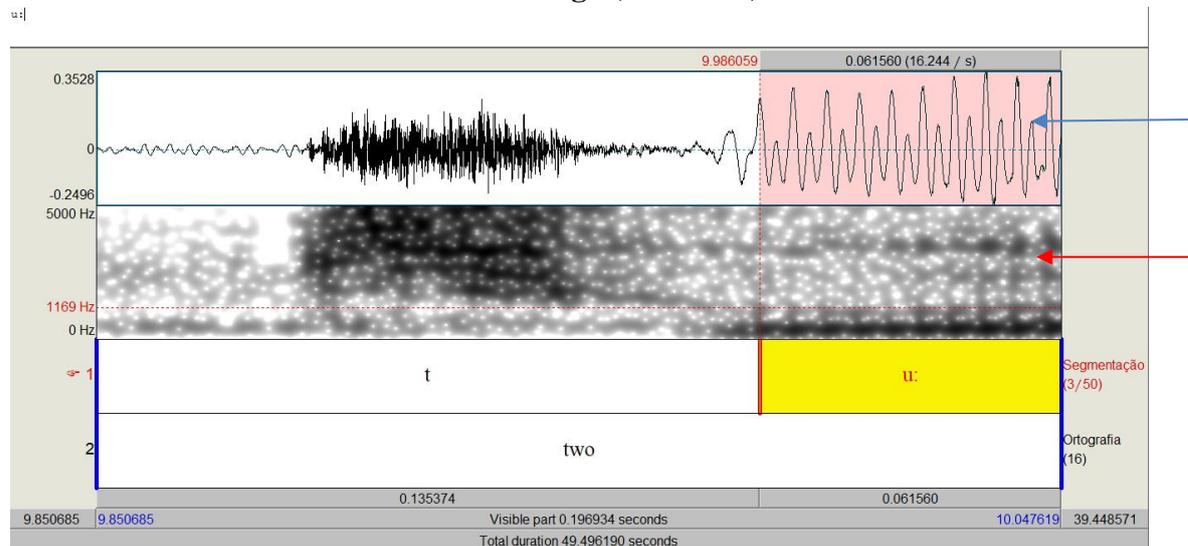
Participou da gravação do *corpus*, um falante estadunidense do sexo masculino, nascido em Nova Iorque, e com 41 anos na ocasião da entrevista.

II. 3 Segmentação

Para realizar as inspeções acústicas, utilizamos o software livre *Praat*, versão 6.0.39, que foi desenvolvido por Boersma e Weenink (2018).

As segmentações dos segmentos fônicos foram realizadas por meio da observação da configuração da forma da onda e do espectrograma, conforme figura a seguir.

Figura 2. Software Praat: setas indicam a forma da onda (azul) e o espectrograma de banda larga (vermelha)



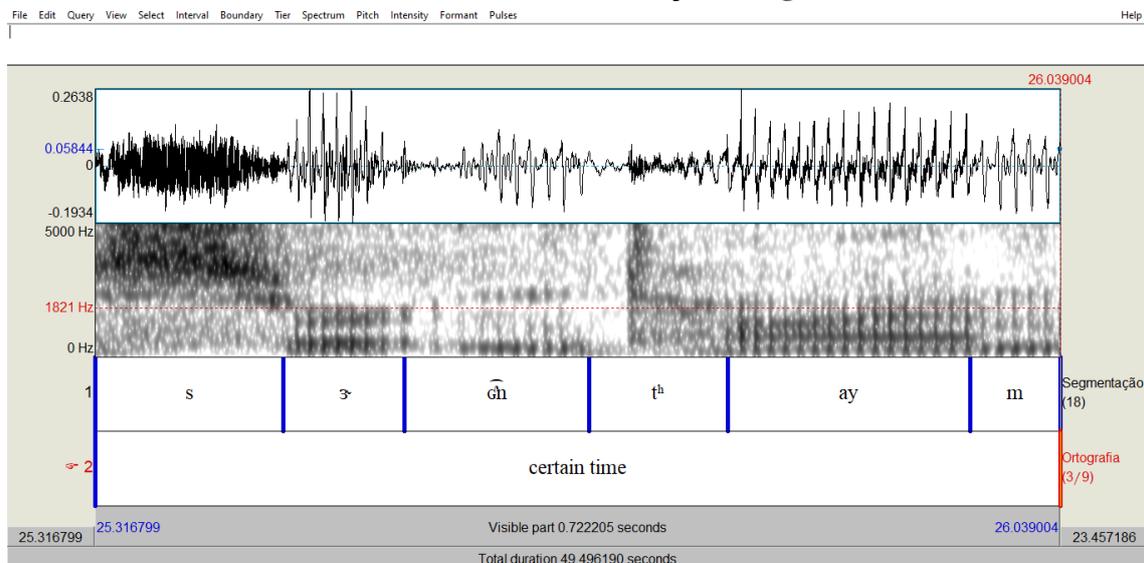
Fonte: elaborado pelo autor

III. ANÁLISE E INTERPRETAÇÃO DOS DADOS

A seguir, apresentamos as inspeções acústicas referentes aos 6 contextos de análise.

Contexto de análise 1, *certain time*: nasal alveolar seguida por plosiva alveolar

Figura 3. Segmentação de “certain time”: forma da onda, espectrograma de banda larga, camada de segmentação com transcrição fonética dos segmentos fônicos e camada de transcrição ortográfica

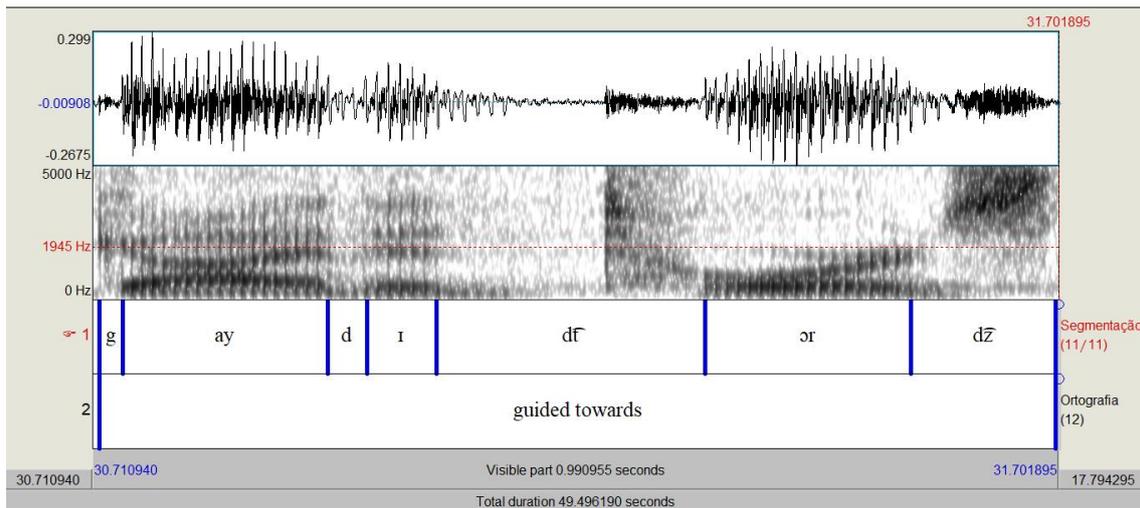


Fonte: Criado pelo autor com a utilização do *software* "Praat"

No contexto “at a certain time”, podemos observar, por meio da inspeção da imagem acima, que não houve pausa silenciosa entre a produção das consoantes /n/ e /t/ das palavras “certain” e “time”, respectivamente, de modo que elas foram coarticuladas. Além disso, como tais consoantes estão na mesma camada da pauta gestual, o fenômeno de “blending” foi observado.

Contexto de análise 2, *guided towards*: plosiva alveolar vozeada seguida por plosiva alveolar não vozeada

Figura 4. Segmentação de “guided towards”: forma da onda, espectrograma de banda larga, camada de segmentação com transcrição fonética dos segmentos fônicos e camada de transcrição ortográfica

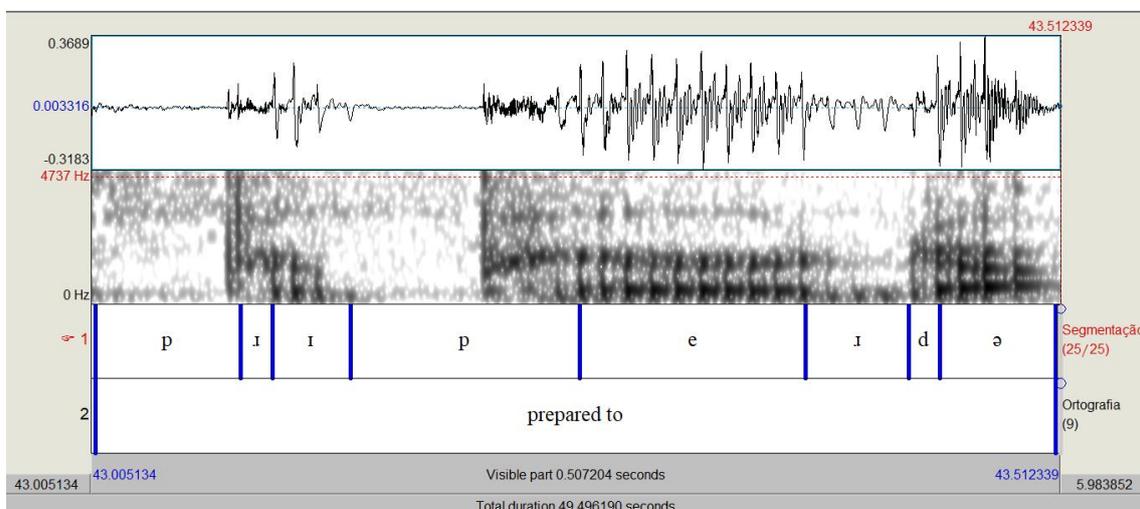


Fonte: Criado pelo autor com a utilização do *software* "Praat"

No contexto acima, “guided towards”, observamos a ausência de pausa silenciosa entre as consoantes adjacentes e homorgânicas /d/ e /t/, pois elas foram coproduzidas, gerando o fenômeno de “blending”

Contexto de análise 3, *prepared to*: plosiva alveolar vozeada seguida por plosiva alveolar não vozeada

Figura 5. Segmentação de “prepared to”: forma da onda, espectrograma de banda larga, camada de segmentação com transcrição fonética dos segmentos fônicos e camada de transcrição ortográfica



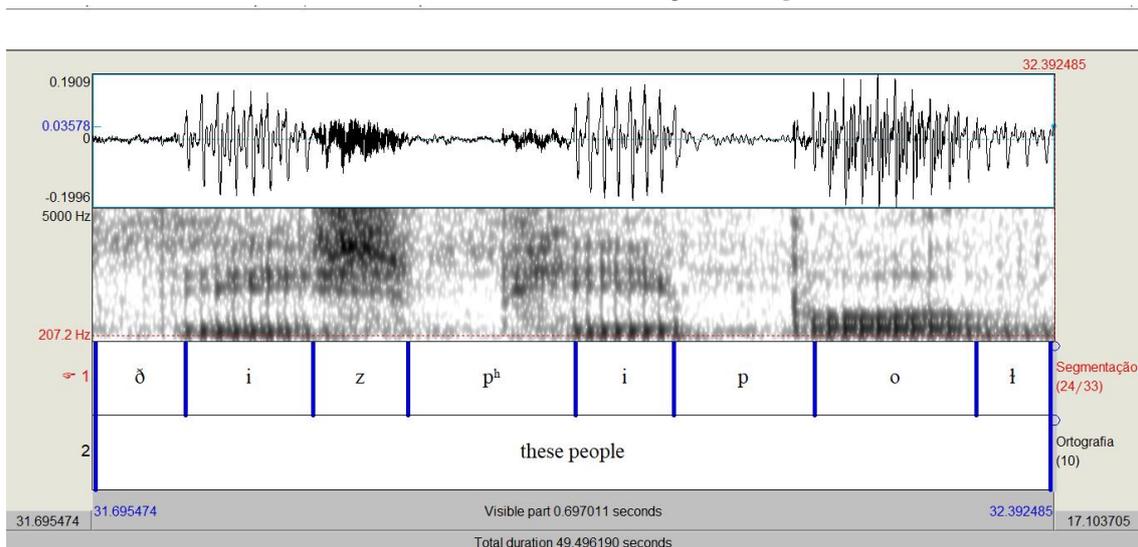
Fonte: Criado pelo autor com a utilização do *software* "Praat"

Em “prepared to”, por meio da inspeção do espectrograma, notamos que não houve a soltura do articulador da consoante /d/ da palavra “prepared” de modo que ela

foi coproduzida com a consoante /t/ da palavra “to”, dando origem ao fenômeno de “blending”. Além disso, não houve a interrupção da vibração das pregas vocais para a produção da consoante /t/ da palavra “to”, ou seja, foi produzida a consoante [d].

Contexto de análise 4, *these people*: fricativa alveolar seguida por plosiva bilabial

Figura 6. Segmentação de “these people”: forma da onda, espectrograma de banda larga, camada de segmentação com transcrição fonética dos segmentos fônicos e camada de transcrição ortográfica

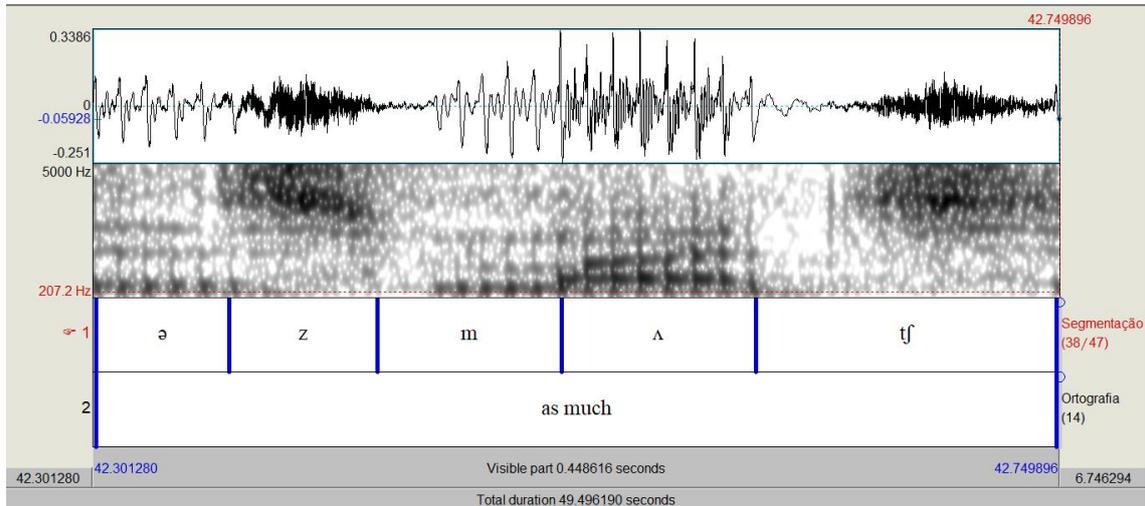


Fonte: Criado pelo autor com a utilização do *software "Praat"*

Na produção de “these people” a não presença de pausa entre as consoantes /z/ e /p/ indica que tais consoantes foram coarticuladas. A coarticulação foi dada pelo gesto de oclusão labial para a produção de /p/ enquanto a consoante /z/ ainda estava sendo produzida, ocasionando a presença do fenômeno de “hiding”.

Contexto de análise 5, *as much*: fricativa alveolar seguida por nasal bilabial

Figura 7. Segmentação de “as much”: forma da onda, espectrograma de banda larga, camada de segmentação com transcrição fonética dos segmentos fônicos e camada de transcrição ortográfica

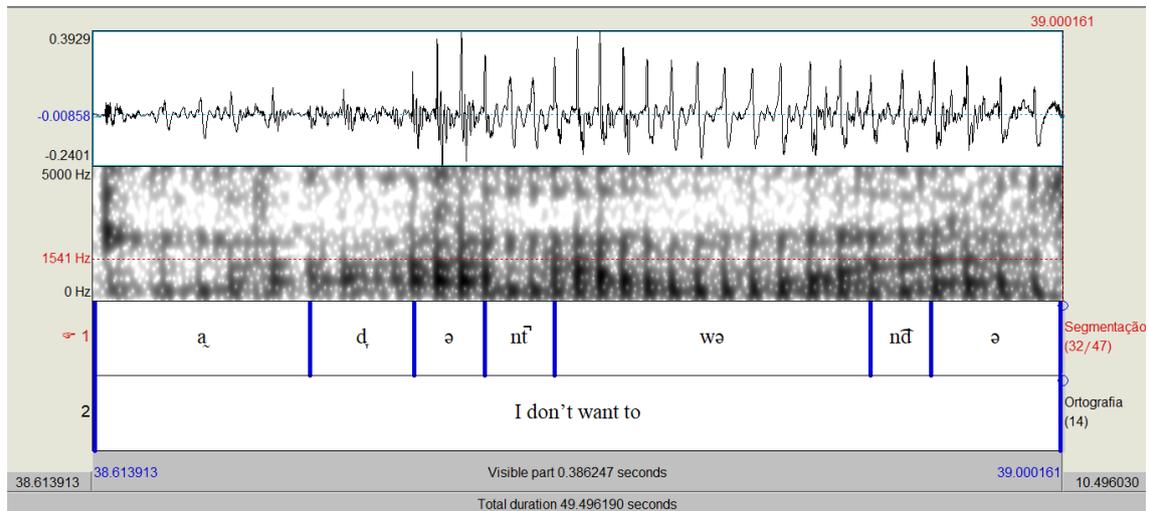


Fonte: Criado pelo autor com a utilização do software "Praat"

Em “as much”, as consoantes /z/ da palavra “as” e /m/ da palavra “much” foram coarticuladas, fato que pode ser observado pela ausência de pausa entre tais consoantes. Além disso, durante a produção de /z/ iniciou-se o gesto de oclusão labial para a produção da consoante bilabial e plosiva /m/. O gesto de oclusão ocorreu de modo a se sobrepor, gradativamente, aos gestos necessários para a produção de /z/ gerando o processo de coarticulação denominado de “hiding”.

Contexto de análise 6, *I don't want to*: plosiva alveolar seguida por aproximante bilabial

Figura 8. Segmentação de “I don't want to”: forma da onda, espectrograma de banda larga, camada de segmentação com transcrição fonética dos segmentos fônicos e camada de transcrição ortográfica



Fonte: Criado pelo autor com a utilização do software "Praat"

No contexto “I don’t want to”, o gesto de soltura da constrição da articulação da consoante /t/ foi sobreposto pelo gesto de protusão labial para a protusão da aproximante /w/ da palavra “want”, ocasionando o processo coarticulatório denominado de “hiding”.

IV. DISCUSSÃO

Por meio das investigações, foi possível observar que o sujeito analisado utilizou estratégias de fala caracterizadas pela coarticulação de segmentos fônicos imediatamente adjacentes interpalavras, fenômenos contemplados pela FAR (BROWMAN; GOLDSTEIN, 1986).

Desse modo, acreditamos ser importante que o conhecimento de tais estratégias seja considerado por professores de inglês ao escolherem amostras de fala que serão utilizadas em tarefas de percepção. Assim, seria possível chamar a atenção para a ocorrência dessas estratégias possibilitando que os ABLI comecem a se conscientizar da existência desses fenômenos, o que resultaria em um refinamento na percepção. Além disso, por estarem cientes da existência desses fenômenos coarticulatórios, os ABLI poderiam incorporar em suas produções estratégias caracterizadas pela coprodução de segmentos fônicos, como a apresentada no contexto “guided towards”, ou a não utilização de pausa, conforme apresentado no contexto “certain time”, a fim de evitar a hiperarticulação de segmentos com o intuito de reduzir esforço articulatório, resultando em maior fluidez no fluxo de fala.

REFERÊNCIAS BIBLIOGRÁFICAS

ALAMEEN, G.; LEVIS, J. Connected Speech. In: LEVIS, J.; REED, M. (Ed.) *The Handbook of English Pronunciation*. New Jersey, USA: Wiley Blackwell, 2015, p. 159-174.

ALBANO, E. C. *O gesto e suas bordas: esboço de uma fonologia acústico-articulatória do português brasileiro*. São Paulo, Campinas: Mercado de Letras: Associação de Leitura do Brasil – ALB, Fapesp, 2001

BOERSMA, P., WEENINK, D. **Praat: doing phonetics by computer**. Versão 6.0.39, Amsterdã, Universidade de Amsterdã, 2018.

BROWN, G. *Listening to spoken English*. London: Longman, 1977.

BROWMAN, C. P., GOLDSTEIN, L Towards an articulatory phonology. *Phonology*, v. 3, 1986, p. 219-252.

_____. Articulatory gestures as phonological units. *Phonology*, v. 6, 1989, p. 201-251.

_____. Gestural specification using dynamically defined articulatory gestures. *Journal of Phonetics*, v. 18, 1990a, p. 299-320.

_____. Tiers in articulatory phonology with some implicants for casual speech. In: KINGSTON, J.; BECKMAN, M. E. (Ed.). *Papers in Laboratory Phonology I*, Cambridge, RU: Cambridge University Press, p. 341-376, 1990b

_____. Articulatory phonology: an overview. *Phonetica*, v. 49, 1992, p. 155-180.
CHOMSKY, N. *Syntactic Structures*. Berlin and New York: The Hague: Mouton, 1985. [1957], Data de publicação original.

FARNETANI, E.; RECASENS, D. Coarticulation models in recent speech production theories. In: HARDCASTLE, W.J.; NIGEL, H. (Ed.). *Coarticulation: Theory, Data and Techniques*. Cambridge, RU: Cambridge University Press, 1999, p. 1-31

FIELD, J. Bricks or mortar: Which parts of the input does a second language listener rely on? *TESOL Quarterly*, 2008, p. 411-432.

REED, M., MICHAUD, C. *Sound concepts: An integrated pronunciation course*. New York: McGraw-Hill, 2005.

SALTZMAN, E.; L. GOLDSTEIN; C. P. BROWMAN; P. E. RUBIN. Dynamics of gestural blending during speech production. Paper presented at 1ST ANNUAL INTERNATIONAL NEURAL NETWORK SOCIETY (INNS), Boston, 1988.

SHOCKEY, L. Phonetic and Phonological properties of connected speech. Ohio State Working Papers in Linguistics, v. 17, 1974, p. iv-143.

SILVA, A., F. Contrastes entre estratégias de falantes bilíngues na produção de um diálogo e um monólogo em inglês. Dissertação de Mestrado, PUC-SP, 2009

_____ Percepção de reduções em inglês como L2. Tese de Doutorado, PUC-SP, 2016