

## Estratégias acústico-articulatórias empregadas por anglofalantes na pronúncia do *tap* alveolar no português brasileiro

Cirineu Cecote Stein\*

**Resumo:** o aprendizado de uma língua estrangeira implica não apenas a aquisição de uma nova estrutura morfossintática e lexical, como também a de estratégias articulatórias necessárias à pronúncia das palavras dessa língua, especialmente nos casos em que não há coincidência entre o quadro fonêmico da L2 e o da L1. Este artigo focaliza algumas estratégias utilizadas por aprendizes anglofalantes do português brasileiro para a pronúncia da consoante *tap* alveolar [ɾ], tanto em posição intervocálica quanto em posição de encontro consonantal, apresentando a fundamentação acústico-articulatória envolvida em cada uma delas. Como estratégias mais produtivas, identificaram-se os usos da consoante aproximante alveolar [ɹ]; da qualidade lateral; de uma vogal epentética entre a consoante obstruinte e o *tap* alveolar, nos encontros consonantais; e de uma consoante vibrante alveolar [r].

**Palavras-chave:** *Tap* alveolar; vogal epentética; fonética acústico-articulatória; português como língua estrangeira

**Abstract:** learning a foreign language implies not only the acquisition of a new morphosyntactic and lexical structure, but also the acquisition of the necessary articulatory strategies involved in the pronunciation of words in that language, especially when there is no coincidence between the phonemic tables of L2 and L1. This paper focuses on some strategies used by English-speaking learners of Brazilian Portuguese in pronouncing the alveolar tap [ɾ], both in intervocalic and in cluster positions. It discusses the acoustic and articulatory principles involved in each of them. The most useful strategies identified were the use of an alveolar approximant consonant [ɹ]; the lateral quality of the segment; an epenthetic vowel between the cluster consonants; and an alveolar trill [r].

**Key-words:** Alveolar tap; epenthetic vowel; acoustic and articulatory phonetics; Portuguese as foreign language

### Introdução

Uma das grandes dificuldades no aprendizado de uma língua estrangeira é a realização apropriada de seus fonemas. Frequentemente, a língua a ser apreendida conta com segmentos que não correspondem aos existentes na língua materna, e a sua articulação torna-se, por vezes, um grande desafio, tanto para o aprendiz (que precisa dominar os mecanismos acústico-articulatórios típicos desses segmentos) quanto para o professor (que precisa encontrar uma metodologia eficiente para o domínio dessa gesticulação).

---

\* Doutor em Letras Vernáculas pela UFRJ, Professor Adjunto do Departamento de Letras Clássicas e Vernáculas da UFPB.

A realização da consoante *tap* alveolar [r] é considerada de grande complexidade, devido à especificidade de seus gestos articulatórios. Isso pode ser evidenciado mesmo entre muitos falantes nativos do português brasileiro (PB), que, por não conseguirem a exatidão desses gestos, acabam, muitas vezes, substituindo essa consoante por outra da mesma classe das líquidas, a consoante lateral alveolar [l], caracterizando-se o que se conhece como “lambdacismo”. A aquisição do *tap* alveolar por um estrangeiro cujo sistema lingüístico materno não o incorpora como fonema, por conseguinte, tende a ser um grande desafio, como ocorre com os anglofalantes. Por vezes, mesmo que o *tap* alveolar seja um alofone utilizado em algumas variedades dialetais da língua materna do aprendiz (como é o caso de alguns dialetos do inglês americano), sua ocorrência não coincide com o posicionamento verificado na cadeia fônica da língua a ser adquirida, sendo necessário um redimensionamento também no nível fonológico.

Como o primeiro passo para o desenvolvimento de qualquer atividade de ensino nesse sentido é conhecer os mecanismos acústico-articulatórios envolvidos na realização de cada fonema, tanto da língua materna do aprendiz quanto da língua que está sendo adquirida, o presente trabalho pretende caracterizar acústica e articulatoriamente o *tap* alveolar, assim como as estratégias articulatórias que anglofalantes, ao adquirirem o PB, podem utilizar como forma de atingir a realização apropriada desse fonema.

### **Considerações sobre a articulação de fonemas estranhos à língua materna**

A eficiência no aprendizado de uma língua estrangeira (L2<sup>1</sup>) implica, entre outras habilidades, a aquisição do quadro fonêmico dessa L2 e a capacidade de realizar esses fonemas em cadeias fônicas de forma eficiente. O nível dessa eficiência será considerado tanto mais elevado quanto mais próxima a realização acústica dos fonemas estiver da realização observada nos falantes nativos, considerando não apenas a qualidade sonora de cada fone individualmente, mas também a conformação melódica impressa aos constituintes sintáticos e às modalidades frasais como um todo.

---

<sup>1</sup> Assumiremos a sigla ‘L2’ da mesma forma que Lima Jr. (2010, p. 5) o faz: para designar “uma língua que não a nativa, não importando a modalidade de aquisição (LE ou imersão)”. Neste trabalho, foram analisados informantes aprendizes do português no Brasil: trata-se, portanto, de situações de imersão, e não da “aprendizagem de uma língua que não a nativa em um país onde essa língua não é falada como L1” (o que caracterizaria a sigla “LE”).

Ao se pensar um nível de excelência na aquisição de uma L2, pensa-se a capacidade de o aprendiz se manifestar oralmente de uma forma tal que não seja reconhecido por um falante nativo dessa L2 como sendo estrangeiro. O domínio dessa habilidade está diretamente relacionado a fatores tanto de ordem cognitiva quanto de ordem física. Cognitivamente, é de esperar que sistemas lingüísticos aparentados ofereçam menor dificuldade de aquisição (no entanto, não se deve desprezar a considerável problemática envolvida no que se conhece como estágio de “interlíngua”, por exemplo, na interface português-espanhol; cf. CARVALHO, 2002) do que os que não apresentam uma origem comum, como é o caso do alemão e do português. Essa menor dificuldade de aquisição se manifestaria, em princípio, nos níveis semântico, sintático e morfológico, com ordenação sintagmática e identidade lexical semelhantes em ambas as línguas.

Quanto aos fatores de ordem física (observando-se, evidentemente, um viés cognitivo), devem-se considerar as habilidades motoras envolvidas na realização fonética dos fonemas da L2. Embora o ser humano tenha a capacidade de articular qualquer fonema de qualquer língua, o domínio dessa habilidade tende a decrescer com o seu amadurecimento. Dificuldades articulatórias se manifestam desde casos bem dissonantes, como a pronúncia de um clique por uma pessoa cuja língua não possua fonemas desse tipo (considerados “exóticos” pelo senso comum vigente entre os usuários das línguas de maior prestígio socioeconômico mundial), até casos mais familiares (principalmente pela exposição midiática), como a pronúncia de uma consoante fricativa dental por um falante do português. É necessário que o aprendiz da L2 adquira consciência não apenas da existência de fonemas diferentes dos fonemas de sua língua materna (L1), como também dos gestos articulatórios envolvidos em sua realização e, por meio da prática (por vezes exaustiva), automatize esses gestos.

As dificuldades encontradas pelos aprendizes não ocorrem apenas na realização de um fonema individualmente. A complexidade aumenta consideravelmente quando, além de articular um fonema inexistente em sua L1, é necessário coarticulá-lo com outro ou outros fonemas que também são específicos da L2. Arrisca-se afirmar que os gestos articulatórios, nesses casos, são muito mais complexos que os envolvidos na pronúncia dos conhecidos trava-línguas até mesmo por falantes nativos. O componente fonético-fonológico de uma L2 pode ser responsável, portanto, por frustrar

consideravelmente o seu aprendiz, que, se não for bem orientado, poderá se sentir impotente ao conseguir articular um fonema específico da L2 isoladamente (ou em contextos fonológicos mais simples), mas não fazê-lo quando inserido em uma cadeia fônica mais complexa.

### **Não-interseções entre os quadros fonêmicos da L1 e da L2**

A sobreposição dos quadros fonêmicos de duas línguas distintas evidenciará várias células em interseção (modo de articulação / ponto de articulação), sendo que algumas serão exclusivas de uma língua ou de outra. Em princípio, os fonemas comuns às duas línguas não apresentariam problemas de aquisição articulatória da L2, muito embora isso nem sempre seja o que ocorre. Tanto no inglês quanto no português, por exemplo, manifestam-se as mesmas consoantes oclusivas; no entanto, as consoantes surdas [p, t, k], no inglês, são aspiradas em posição inicial de palavra e em outras posições quando em sílaba tônica, mas são sempre produzidas sem a aspiração quando seguindo [s] na mesma sílaba (LADEFOGED, 1999, p. 43). Ao aprenderem o português, língua em que essa aspiração não ocorre, os aprendizes anglofalantes fazem a transferência dessa aspiração para as consoantes oclusivas da L2, o que necessariamente os identifica como estrangeiros.

Caso a L2 apresente fonemas que lhe são específicos, comparativamente à L1 do aprendiz, o processo de aquisição articulatória desses fonemas percorrerá, em muitos casos, uma trajetória de natureza adaptativa. Na tentativa de articular esses fonemas, o aprendiz se apoiará naqueles de sua L1 que apresentem características articulatórias mais próximas dos fonemas a que está sendo apresentado na L2. Por meio de estratégias de erro-e-acerto, o aprendiz testará diferentes possibilidades articulatórias, até encontrar a que seja idêntica ou mais próxima da utilizada pelos falantes nativos na pronúncia desses segmentos. Dutra (2008, p. 51) identificou, em aprendizes estadunidenses e espanhóis do português, que a dificuldade de articulação, por exemplo, da consoante lateral palatal [ʎ] foi contornada com o uso de [l], [l<sup>j</sup>] ou [n]; e que a da consoante nasal palatal [ɲ] foi contornada com [ỹ] ou [y]. Esses dados refletem um

comportamento pragmático: o aprendiz tem a necessidade de pronunciar as palavras que está adquirindo e, para isso, vale-se das estratégias que lhe sejam mais acessíveis<sup>2</sup>.

A comparação dos quadros fonêmicos da L2 e da L1 de um aprendiz se coloca como um dos primeiros passos necessários à consciência do que será o processo de aquisição articulatória da L2. A partir dessa comparação, será possível estabelecer uma hierarquia para os níveis de complexidade articulatória envolvidos nos fonemas exclusivos da L2 e considerar, por exemplo, que embora os róticos estejam presentes em diversas línguas naturais<sup>3</sup>, podem pertencer ao último grupo de consoantes adquiridas pelos falantes dessas línguas, como ocorre no caso do inglês (BALL *et al.*, 2001) e do PB (HERNANDORENA, 1990). Mas uma classe de fonemas ou um fonema específico integrar o quadro fonemático de duas línguas distintas não implicará necessariamente uma facilitação articulatória na L2.

Além da qualidade segmental propriamente dita, é necessário considerar os posicionamentos fonológicos possíveis para esse segmento na L1 e na L2. Como indica Face (2006, p. 48), a pronúncia do *tap* alveolar [ɾ] por aprendizes estadunidenses do espanhol apresenta dois fatores complicadores. No inglês, esse segmento ocorre como alofone de /t/ e /d/ em posição pós-tônica, como em ‘later’ e ‘ladder’. Esse alofone realiza-se de forma praticamente idêntica à do *tap* alveolar espanhol. No entanto, os falantes nativos do inglês americano não estabelecem a associação desse alofone com um rótico, e isso implica uma dificuldade na pronúncia desse segmento no espanhol, pois é necessário recategorizá-lo, uma vez que no espanhol trata-se efetivamente de um rótico, não de um alofone. O segundo fator complicador diz respeito ao seu posicionamento na estrutura silábica em cada uma dessas línguas: no inglês americano, o *tap* alveolar ocorre exclusivamente em posição pós-tônica intervocálica; no espanhol, ele ocorre entre vogais em posição tanto pré-tônica quanto pós-tônica, e também em

---

<sup>2</sup> Um interessante exercício de identificação das possíveis estratégias articulatórias envolvidas na aquisição de fonemas é observar uma pessoa cantando uma música em um idioma que não domina. Ela pronunciará as palavras segundo a sua percepção, e os fonemas que não forem comuns à sua L1 serão articulados de acordo com o referencial do seu quadro fonemático. Uma das premissas da fonêmica diz que os sons tendem a flutuar. Assim, como no espanhol, por exemplo, não se verificam as vogais orais médias-baixas, um hispano-falante ouvirá ambas as palavras ‘vovô’ e ‘vovó’ como sendo a mesma, e as pronunciará como [vo"vo], projetando a vogal posterior média-alta – existente em sua L1 – sobre a vogal posterior média-baixa do português. Esse mesmo procedimento será utilizado ao ser cantada uma música em um idioma desconhecido.

<sup>3</sup> Maddieson (1980) indica que 189 línguas (59%) das 321 que compõem o corpus do UCLA Phonological Segment Inventory Database (UPSID) apresentam algum tipo de rótico.

final de palavra ou de sílaba, além de poder ocupar a posição de segundo membro de um ataque silábico complexo. O aprendiz estadunidense do espanhol, portanto, embora articule o *tap* alveolar corriqueiramente, encontra dificuldade na sua articulação em vocábulos espanhóis como ‘caridad’, ‘cítara’, ‘morder’ e ‘claustro’, tanto pela necessidade de sua recategorização segmental quanto pelos seus diferentes posicionamentos fonológicos. O mesmo ocorrerá considerando-se o português como L2, como se pode deduzir da semelhança na ordenação segmental desses significantes espanhóis comparativamente aos seus equivalentes portugueses.

Ainda considerando posicionamentos fonológicos, é pertinente o exemplo apresentado em Dutra (2008, p. 80-81), evidenciando que o contexto de uso de um fonema na L1 (inglês ou espanhol) motivará o seu uso na L2 (PB). A articulação de uma consoante aproximante alveolar [ɹ] em início de sílaba, em início de vocábulo, pelos estadunidenses, projeta-se sobre essas posições preenchidas, nas palavras portuguesas, por uma consoante fricativa glotal [h]. Assim, para a palavra ‘redondo’, ao invés da pronúncia [he'dõdu], é comum observar-se [ɹe'dõdu], o mesmo ocorrendo em ‘prova’, com [ˈpɹɔvə] no lugar de [ˈpɾɔvə].

### **A percepção de fonemas que não integram a L1 pelo aprendiz de L2**

Considerando-se que o aprendiz de L2, valendo-se da proximidade categorial, buscará uma articulação para um fonema que é estranho ao quadro fonêmico de sua L1, infere-se sua capacidade de reconhecer sons discretos na L2, atribuindo-lhes uma identidade. Essa capacidade está diretamente relacionada à percepção (inconsciente) que o aprendiz tem dos movimentos articulatorios envolvidos na realização de cada fonema. Best (1995) reconhece o primitivo perceptual da fala como sendo o gesto articulatorio, que, assumindo-se a definição de Browman e Goldstein (1992), mostra-se simultaneamente categórico e gradiente, sendo capaz, portanto, de construir a identidade única de um segmento fônico como também revelar uma transição gradativa entre as categorias fonemáticas adjacentes. O gesto articulatorio é caracterizado, assim, pelo movimento dos articuladores, mas coloca-se também como a própria representação desses gestos. Reconhecendo esse gestual, o aprendiz de L2 buscará reproduzi-lo.

Perceptualmente, é possível ocorrerem dois comportamentos distintos. O aprendiz de L2 poderá, por um processo assimilatório, categorizar fonemas da L1 e da

L2 como sendo idênticos, mesmo que um deles apresente alguma característica específica (ECKMAN, 2004). No exemplo já citado das consoantes oclusivas no inglês, percebe-se uma aspiração imediatamente após a abertura da oclusão típica desses segmentos. No português, isso não se verifica como regra. O anglofalante aprendiz do português tenderá a assimilar essas duas realizações distintas como idênticas, não discriminando, portanto, [p, t, k] de [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>] (e pronunciará essas oclusivas, no português, também com aspiração). Outro comportamento identifica-se com a afirmação de Zimmer (2007, p. 110): se não houver assimilação perceptual com uma categoria existente na L1, os fonemas da L2 serão mais bem discriminados e, segundo ela, produzidos corretamente. Efetivamente, a não-assimilação categorial por via perceptiva parece permitir, via identificação do gesto articulatório, uma maior consciência sobre a especificidade de um fonema estranho à L1, mas isso não implica necessariamente uma produção articulatória correta. Essa produção dependerá de vários fatores, cognitivos e motores. Da mesma forma que, como indica Eckman (2004), a capacidade de recategorizar fonemas da L2 que sejam diferentes da L1 pode diminuir com o avanço da idade, a capacidade motora para novos gestos articulatórios também poderá diminuir.

Percepção e realização acústica, embora estejam diretamente relacionadas, de certa forma podem dissociar-se, dependendo da complexidade articulatória envolvida. Alguns fonemas, para serem realizados foneticamente, dependem de uma gesticulação complexa a ser desenvolvida em uma duração temporal muito breve. Nishida (2005, p. 7) exemplifica essa complexidade por meio do *tap* alveolar. A realização desse rótico, em si mesma, demanda uma habilidade motora considerável (cf. mais à frente). Se ela ocorrer com um ritmo acelerado de fala, ou coarticular-se com um segmento adjacente, poderão ser produzidos segmentos que, perceptualmente, se identifiquem como *taps*, embora não se trate da realização desse segmento em sua forma canônica. E essa produção, frequentemente, passará despercebida aos usuários da língua. No entanto, se esses segmentos sonoros forem submetidos a uma análise espectrográfica, poderão ser reveladas diferenças acústicas produzidas pelos próprios falantes nativos. Se um falante nativo produz variações acústicas em diferentes realizações de um mesmo fonema, é esperado que o aprendiz dessa L2 também as produza. Se essas variações acústicas realizadas pelo falante nativo não comprometem a percepção e a identificação categorial

desse segmento, é possível que também não comprometam quando realizadas pelo aprendiz estrangeiro, desde que, naturalmente, estejam dentro de certos limites. Algumas imperfeições nos gestos articulatórios podem passar despercebidas, inclusive por conta de o segmento sobre o qual elas incidem estar inserido em uma cadeia fônica, e a percepção não se dar no nível segmental, mas no nível de agrupamentos segmentais.

A gradiência do gesto articulatório, como caracterizada por Browman e Goldstein (1992, *apud* BEST, 1995), implicando também nuances perceptivas, indica a necessidade de um treinamento específico de pronúncia para o aprendiz de L2. Como prática institucionalizada, normalmente assume-se que esse aprendiz, como ser humano dotado de um sistema lingüístico próprio e capaz de interagir com outros sistemas, internalizará espontânea e automaticamente os novos gestos articulatórios que se lhe apresentarem. Embora isso, em parte, possa ser pertinente, uma metodologia de ensino deve, por princípio, prever procedimentos facilitadores e aceleradores da aquisição da L2, inclusive em seus aspectos fonético-fonológicos. Como indica o estudo de Dutra (2008, p. 73), alguns resultados obtidos junto a estadunidenses apontam uma tendência à redução dos índices da consoante retroflexa (marca transposta do sistema inglês para o PB) e à elevação nos índices do *tap* alveolar – à medida que ocorre uma maior reflexão por parte dos informantes quando falam.

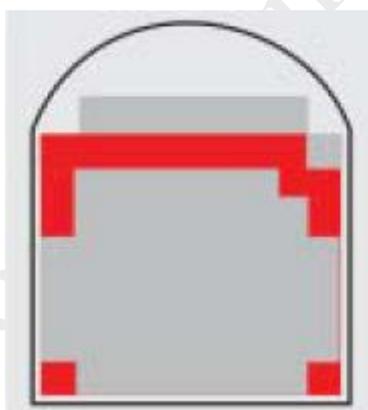
### **Caracterização acústico-articulatória do *tap* alveolar [ɾ]**

O *tap*<sup>4</sup> insere-se em uma classe de sons tradicionalmente conhecida como róticos. Embora as propriedades articulatórias ou auditivas sejam o critério normalmente utilizado para a categorização das classes sonoras, o termo ‘róticos’ é baseado no fato de, nos sistemas ortográficos, os sons que se incluem nessa categoria serem representados pela letra ‘r’ (LADEFOGED & MADDIESON, 1996, p. 215). Entre os róticos, o estabelecimento de um grupo nomeado “Tap or Flap” no Alfabeto Fonético Internacional sugere que os dois termos ‘tap’ e ‘flap’ referem-se a uma mesma categoria sonora. E normalmente não se estabelece uma distinção entre eles. Ladefoged & Maddieson (1996, p. 231), no entanto, propõem que sejam diferenciados pelo tipo de contato feito pelo articulador ativo com o passivo durante a sua produção. Durante a

---

<sup>4</sup> A tradição descritiva que caracteriza o *tap* como “vibrante simples”, seguindo Hoyos-Andrade (2002, p. 56), é inapropriada, uma vez que, como se trata de um ligeiro toque do articulador ativo no passivo, não se caracteriza uma vibrante, que suporia, pelo menos, dois toques sucessivos.

produção do *flap*, o breve contato entre os articuladores se daria de forma tangencial, ou seja, a superfície superior do trato vocal seria tocada de passagem. Na produção do *tap*, o articulador ativo move-se diretamente para essa superfície superior, caracterizando um movimento balístico (*idem*, p. 243), em um contato também breve. Os dois tipos são normalmente coronais. Esse gesto coronal pode ser evidenciado palatograficamente. Reis & Espesser (2006, p. 195) apresentam a representação palatográfica do *tap* na palavra ‘garapa’, reproduzida na figura 1, abaixo. Após terem indicado que a caracterização do *tap* se faz a partir de um contato estreito tanto no eixo longitudinal lateral como no eixo transversal, observam que, nos dados obtidos, a interrupção do contato longitudinal lateral, em ambos os lados, não seria esperada, e que isso se deveria a problemas de adaptação do palato artificial em contato com o palato do informante. O interesse da observação se volta, especificamente, para o contato transversal, indicando o contato do articulador ativo com a região alveolar.



**Figura 1** Representação eletropalatográfica do *tap* alveolar [r], na palavra ‘garapa’ (REIS & ESPESSER, 2006, p. 195).

Acusticamente, a inspeção espectrográfica indica, para o *tap*, uma descontinuidade espectral. Essa descontinuidade representa o momento em que ocorre uma rápida oclusão no trato vocal, decorrente do contato do articulador ativo (ápice da língua) com a região dental ou alveolar. Como ocorre o contato direto do articulador ativo com o passivo, verifica-se a diminuição do trato vocal. A intensidade também tende a ser menor, com a diminuição de energia representada no espectrograma por tons mais claros de cinza ou, mesmo, uma lacuna em branco. Se se considerar um *tap* intervocálico, percebe-se que a concentração de energia ao longo das vogais é elevada

(tons bastante escuros); a representação espectrográfica da energia diminui durante a realização do *tap* (tons bem mais claros, devido à oclusão) e, imediatamente após o contato do ápice da língua com a região dental ou alveolar, verifica-se uma barra de explosão, representando a liberação da energia retida com a oclusão; a partir de então é retomado o tom escuro gerado pela intensidade maior da segunda vogal. Silva (1996, p. 40), baseando-se em Recanses (1991b), sugere, para caracterizar esses dois momentos que constituem a produção do *tap*, as nomenclaturas de “início do fechamento” (correspondente ao período de “quase silêncio”) e “final do fechamento” (momento em que o articulador ativo toca o passivo). Sua descrição dos gestos articulatórios envolvidos na produção de um *tap* se dá nos seguintes termos: “Pode-se dizer que esta estrutura do *tap* se deve ao fato de que, quando a língua bate nos alvéolos, há uma interrupção momentânea de saída de ar pela boca, de tal forma que a voz quase desaparece, devido ao aumento da pressão supraglótica. Antes que a voz desapareça, no entanto, a constrição se desfaz e a laringe retoma a produção da voz, ainda que com pouca energia (comparando sempre com vogais ou outras aproximantes)”.

Ladefoged & Maddieson (1996, p. 231) verificam, para a palavra espanhola ‘caro’ [‘karo], dois comportamentos formânticos distintos. Na pronúncia realizada por uma informante espanhola, o segundo formante (F2) apresenta uma elevação acentuada enquanto a consoante é formada (facilmente verificável, via coarticulação, pela elevação de F2 na segunda metade da primeira vogal). Já na realização por um informante peruano, essa elevação não se verificou. Esses dados indicam a instabilidade de produção do *tap* (e dos róticos, de uma forma geral), com possível diferenciação nos pontos de articulação. Os autores também indicam que, embora o *tap* seja considerado uma consoante vozeada, nessa segunda pronúncia não se observou vibração das pregas vocais. Isso é decorrência do potencial conflito (também verificado na produção dos “trills”, as vibrantes) entre uma definição acústica (mais de um período de vibração efetiva) e uma definição articulatória (posicionamento dos articuladores em uma configuração tal que, dadas as condições aerodinâmicas apropriadas, a vibração deveria ocorrer) (*idem*, p. 217). Essa não-vibração das pregas vocais durante a realização dos *taps* não seria, segundo os autores, incomum.

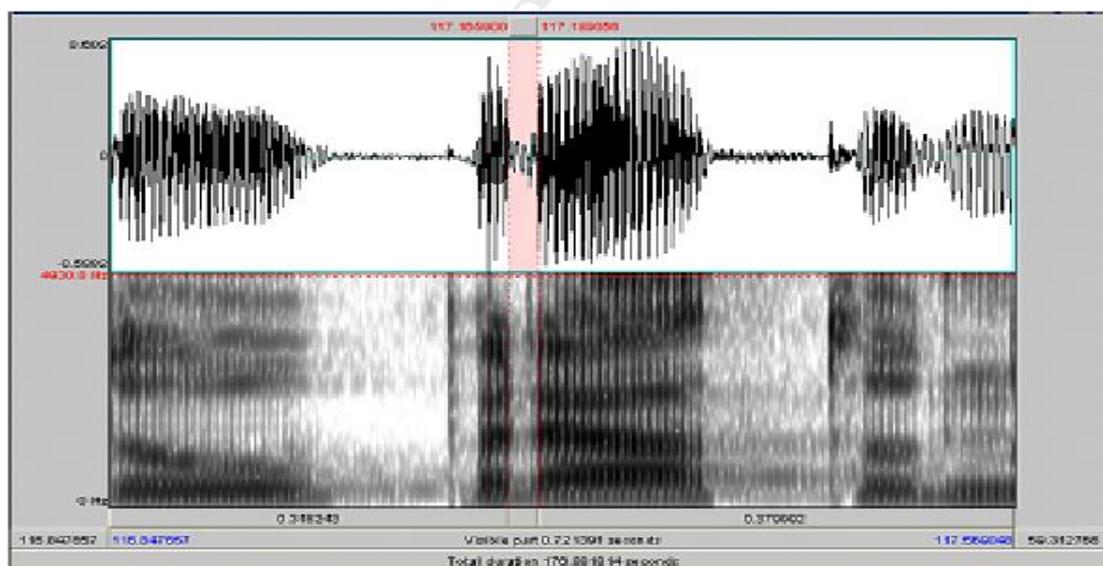
No português (e em outras línguas, como o espanhol), o *tap*, além de ocorrer em ambiente intervocálico, pode ocorrer também em grupos consonantais<sup>5</sup> (ou “clusters”), cuja sequência fonológica deve ser constituída de uma consoante obstruinte + *tap* + vogal, como em ‘crase’, ‘frase’, ‘livro’. O ambiente intervocálico pode ser exemplificado com ‘caro’ ou ‘para’. Há, ainda, a possibilidade de ambientação em coda silábica (como na pronúncia de ‘par’, por um informante paulistano, em Silva, 1996, p. 68-9). A caracterização articulatória para o *tap* em posição intervocálica segue a referenciada anteriormente. Acusticamente, Silva (1996, p. 73-4) indica que, em seus dados, o *tap* apresenta os dois primeiros formantes (F1 e F2) distantes entre si, com o terceiro formante (F3) alto e próximo de F2, sendo que a diferença entre esses formantes, no início e no final do fechamento (breve oclusão) não é significativa. Em associação com o aspecto duracional, em que foi detectada sua estabilidade em qualquer posição (intervocálica, em encontro consonantal ou em coda silábica), esses dados confirmariam o comportamento do *tap* como sendo uma unidade (distinta de outras realizações róticas, como a aproximante alveolar). Em grupos consonantais, no entanto, há um item adicional a ser considerado: um elemento vocálico, recorrentemente reconhecido como vogal epentética.

No português, as possibilidades de ataque complexo silábico se configuram na combinação de uma consoante obstruinte e uma consoante líquida, nesta sequência (/pr, pl, br, bl, tr, tl, dr, kr, kl, gr, gl, fr, fl, vr, vl/). Como indica Silva (1996, p. 17), a tendência nas línguas com encontros consonantais é a de que as líquidas ocorram em posição próxima ao núcleo silábico. Observe-se que as consoantes líquidas (entre as quais se inclui o *tap* alveolar), apresentam características acústicas que as localizam intermediariamente entre os sons vocálicos e os consonânticos. Na observação de Quilis (1981, p. 274), as consoantes líquidas e as vogais apresentam estrutura formântica muito semelhante entre si, diferindo quanto à intensidade global (menor nas líquidas) e à frequência fundamental (também mais baixa nas líquidas).

---

<sup>5</sup> Existe a possibilidade de o *tap*, em encontros consonantais, ser pronunciado como uma vibrante múltipla. Isso equivale a dizer que a sequência “oclusão entre ápice da língua e região alveolar com posterior liberação da oclusão” ocorre, nesses casos, pelo menos duas vezes. Espectrograficamente, esse tipo de ocorrência seria caracterizado, então, por pelo menos duas sequências de diminuição de energia (tom mais claro de cinza) seguida por barra de explosão. A repetição desses eventos é que permitiria caracterizar o segmento como produzido com vibração múltipla. Como é possível observar na figura 2 (realização de ‘prato’, por um brasileiro) e nas figuras 6 e 7 (realizações, respectivamente, de ‘pedra’ e de ‘empregado’, por um estadunidense), houve apenas uma vibração, daí assumir-se, aqui, o *tap* como o segmento realizado em encontros consonantais, ao invés de uma vibrante múltipla.

O sistema fonológico do português (brasileiro), diferentemente do de outras línguas, não sustenta o encontro entre consoantes oclusivas e consoantes fricativas ou nasais, por exemplo. Isso é confirmado pelo uso comum que os brasileiros normalmente fazem de uma vogal epentética entre essas consoantes, quer heterossilabicamente, como em ‘advogado’ [adivɔ'gadu], quer tautossilabicamente, como em ‘pneu’ [pi'neʊ]. Percebe-se que, por conta de sua natureza acústico-articulatória, a consoante oclusiva, para realizar-se, precisa estar apoiada em uma vogal. O sistema fonológico admitir a combinação tautossilábica de uma consoante obstruinte com uma consoante líquida, então, parece dever-se ao fato de as consoantes líquidas apresentarem-se acusticamente de forma muito semelhante às vogais. Silva (1996, p. 69-70), inclusive, sugere que o *tap* é uma consoante “vocalizável”. Seus dados indicam, especialmente nos obtidos a partir de encontros consonantais, que o início do fechamento (para a ligeira oclusão articulatória) foi possível de ser identificado apenas através da amplitude de seus períodos, menor que a dos períodos da vogal tônica, por conta do aspecto semelhante entre eles, o que seria mais um indício do caráter vocálico do *tap*.



**Figura 2** Oscilograma e espectrograma evidenciando a pronúncia da palavra ‘prato’ (NISHIDA, 2005, p. 16). Observe-se que o *tap*, isolado pelas linhas verticais, está precedido por um elemento vocálico que apresenta a mesma conformação formântica da vogal que o sucede.

Essa natureza vocalizável do *tap*, no entanto, não seria suficiente para sustentar seu posicionamento posterior imediato a uma consoante oclusiva, ele próprio

necessitando, também, apoiar-se em duas vogais, como sugere Silva (1996, p. 68). A autora detectou que, nos encontros consonantais, surge um elemento vocálico (que denominou “abertura oral”) anterior ao *tap*, semelhante à vogal tônica quanto a sua natureza formântica. Esse mesmo elemento vocálico ocorreu também em posição final<sup>6</sup>. Em estudo posterior, Nishida (2005) propõe que, no PB, esse elemento vocálico seja efetivamente a mesma vogal nuclear do grupo, entrecortada pelo *tap*, com duração menor que a dela<sup>7</sup>. Esse elemento vocálico seria, então, a primeira (parte da) vogal em que o *tap* se apoiaria (cf. figura 2), rompendo-se a seqüência fonológica consoante obstruente + consoante líquida tipificada para os ataques silábicos complexos.

### **Metodologia**

Os dados que serão discutidos a seguir foram colhidos junto a quatro informantes de L1 inglesa, aprendizes do PB. Esses estudantes estavam, à época, filiados ao Programa Lingüístico-Cultural para Estudantes Internacionais (PLEI), em realização na Universidade Federal da Paraíba, e cursavam, respectivamente, os níveis básico, pré-intermediário, intermediário e avançado. Os dois primeiros eram masculinos, nascidos nos Estados Unidos da América; as duas outras, femininas, eram nascidas na Inglaterra.

Com o objetivo de detectar os fonemas do português que ofereceriam maior complexidade articulatória a esses informantes, foi solicitada a leitura, com registro sonoro concomitante, de textos e frases, sem recorrer a frases-veículo. Isolados os vocábulos em que fonemas apresentavam impropriedades articulatórias, procedeu-se à identificação das estratégias utilizadas para a articulação desses fonemas. A inspeção acústica dessas estratégias se deu com o auxílio do programa computacional para análise acústica Praat (BOERSMA & WEENINK, 2006).

---

<sup>6</sup> Dados semelhantes foram encontrados também para o norueguês, segundo Kvale e Foldvik (1995); para o espanhol, segundo Quilis (1993) e Almeida & Dorta (1993); para o búlgaro e o tcheco, segundo Jetchev (1995), *apud* Nishida (2005).

<sup>7</sup> Ramírez (2006, p. 51) detectou que, no espanhol, a duração desse elemento vocálico esteve na média de 32% da duração da vogal total, no caso dos encontros consonantais.

## **Estratégias empregadas na articulação do *tap* alveolar**

### **a) A consoante aproximante alveolar [ɹ]**

O sistema lingüístico do inglês não utiliza o *tap* alveolar com valor fonêmico. Como discutido, algumas variedades dialetais do inglês americano incorporam o *tap* como alofone de /t/ e /d/ em posição pós-tônica, como em ‘later’ e ‘ladder’, o que significa que esses anglofalantes são familiarizados com os gestos articulatórios necessários a essa realização. Contudo, a posição fonológica em que essa variante pode ocorrer não é semelhante àquela em que o *tap* ocorre na língua portuguesa: entre vogais, com valor fonêmico, e em encontros consonantais, como segundo elemento de um ataque silábico complexo.

Não possuindo o sistema lingüístico de sua L1 esse fonema da L2, o aprendiz buscaria, como primeira estratégia articulatória para realizá-lo, aquele de sua L1 com características mais próximas do fonema-alvo da L2. Como se trata de um *tap* alveolar, o anglofalante poderia valer-se de uma consoante aproximante alveolar [ɹ] ou de uma consoante lateral alveolar [l] (o *tap* classifica-se tanto como rótico como consoante líquida). O uso da consoante aproximante alveolar mostrou-se uma estratégia bastante produtiva, especialmente para os informantes dos níveis pré-intermediário (masculino, estadunidense) e intermediário (feminina, inglesa). Deve-se considerar, no entanto, que mesmo entre falantes nativos do inglês americano, a consoante aproximante alveolar varia quanto aos detalhes da articulação específica, podendo ser retroflexa e verdadeiramente alveolar para alguns e não-retroflexa e pós-alveolar para outros, apresentando também diferentes graus de arredondamento dos lábios (FACE, 2006, p. 47-48).

Ao caracterizarem a família dos róticos, Ladefoged & Maddieson (1996, p. 232) mencionam incluírem-se nela membros em cuja articulação não existe contato, mas apenas uma aproximação entre os articuladores. Nos casos em que essa produção típica é acompanhada por fricção, tem-se, no sistema fonológico do inglês, a consoante fricativa glotal surda [h]; em outros, uma aproximante é produzida.

O termo “aproximante” é utilizado por haver apenas uma aproximação entre os articuladores, sem contato direto entre eles (como ocorre, por exemplo, na realização do *tap*). As constrictões ocorrem na parte mais baixa da faringe e no centro do palato, mas

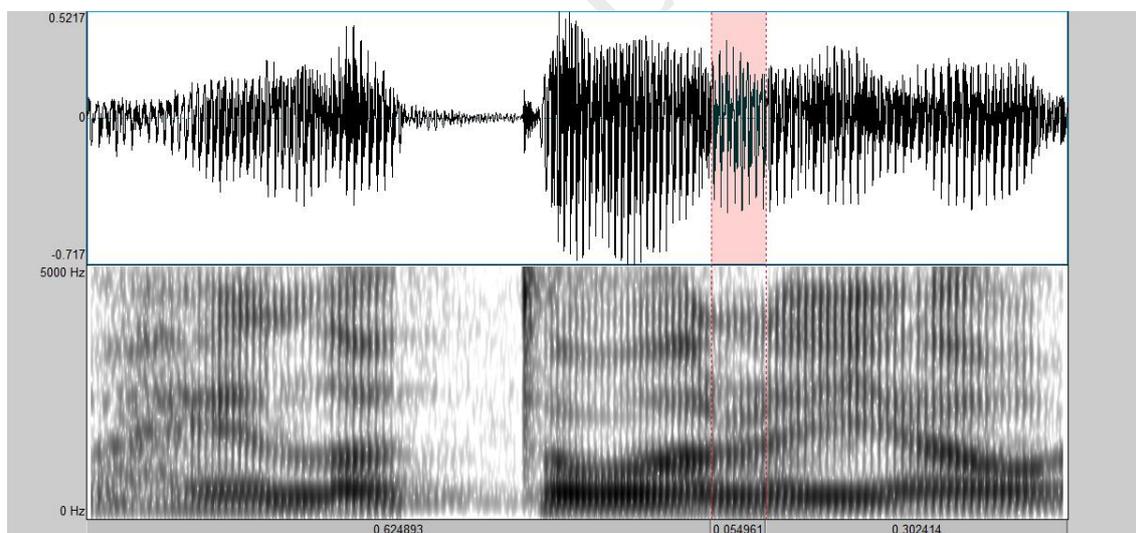
sem elevação da ponta ou da lâmina da língua (ULDALL, 1958, *apud* LADEFOGED & MADDIESON, 1996, p. 234). Assim como as fricativas, as aproximantes são sons contínuos, o que implica dizer que o ar não sofre interrupção completa ao passar pelo trato vocal, e nem mesmo uma constrição intensa, como ocorre nas fricativas (essa constrição severa, nas fricativas, é responsável pela produção do ruído que lhes é peculiar). Esse caráter contínuo do som reflete-se, visualmente, numa continuidade espectral. A relativa regularidade que se observa nas formas de pulsos das ondas permite a observação de uma certa estrutura formântica. Embora esses sons se assemelhem, de certa forma, às vogais, apresentam uma amplitude mais baixa que elas. Comparativamente ao *tap* alveolar, a amplitude é mais alta, como decorrência do contraste entre o caráter contínuo das aproximantes e interrupto do *tap*. Ladefoged & Maddieson (1996, p. 234) reconhecem como característica acústica mais marcante desse som a frequência muito baixa do terceiro formante (F3). Há também a possibilidade de esse som ser articulado com considerável estreitamento da abertura labial. Referindo-se a Fant (1968), indicam que, havendo constrições na região mais baixa da faringe ou na região palatal, prevê-se F3 relativamente baixo, próximo a F2<sup>8</sup>. No caso de arredondamento dos lábios, ambos F2 e F3 são rebaixados. É possível que falantes do inglês americano utilizem, para [ɹ], uma articulação mais ou menos retroflexa, também combinada com uma constrição na região mais baixa da faringe e com o arredondamento dos lábios, produzindo, essa combinação articulatória, F3 baixo. Como apontam os autores, a produção de um F3 baixo pode ser obtida por meio da combinação de múltiplos mecanismos articulatórios, empregados para as possíveis variantes desse segmento. Hashi *et al.* (2003, p. 4) indicam que muitos estudos parecem concordar que o valor baixo de F3 está associado à cavidade frontal, embora não haja evidência conclusiva para a presença de correlatos acústicos das várias posturas articulatórias envolvidas na produção do [ɹ] no inglês americano, pelo menos quanto aos três primeiros formantes.

---

<sup>8</sup> Espy-Wilson *et al.* (2000), ao indicarem que o /r/ do inglês americano ocorre tanto como um núcleo silábico como em posição consonantal, evidenciam que, para ambos os casos, o padrão formântico característico envolve um padrão de F1 – F2 semelhante ao de uma vogal central arredonda canônica, acompanhado de F3 baixo. Referindo-se a Hagiwara (1995), assumem que F3 sofre um rebaixamento de 60% a 80% em relação ao valor de F3 na vogal neutra, o que localiza F3 abaixo dos 2000 Hz. Segundo os dados disponíveis na literatura que citam, F1 situa-se entre 250-550 Hz; F2, 900-1500 Hz; F3, 1300-1950 Hz.

Como estratégia articulatória para atingir a produção do *tap* alveolar no PB, é interessante notar que a gesticulação envolvida na realização da aproximante alveolar é menos complexa que a daquele segmento (seria uma hipoarticulação do *tap*, nas palavras de Nishida, 2005, p. 22). O *tap* combina, além de um posicionamento específico do dorso da língua, um movimento de seu ápice em direção à região alveolar, que será tocada (gerando uma breve oclusão). Assim, o uso da aproximante alveolar poderia ser percebido como um estágio preparatório para a articulação precisa do *tap*.

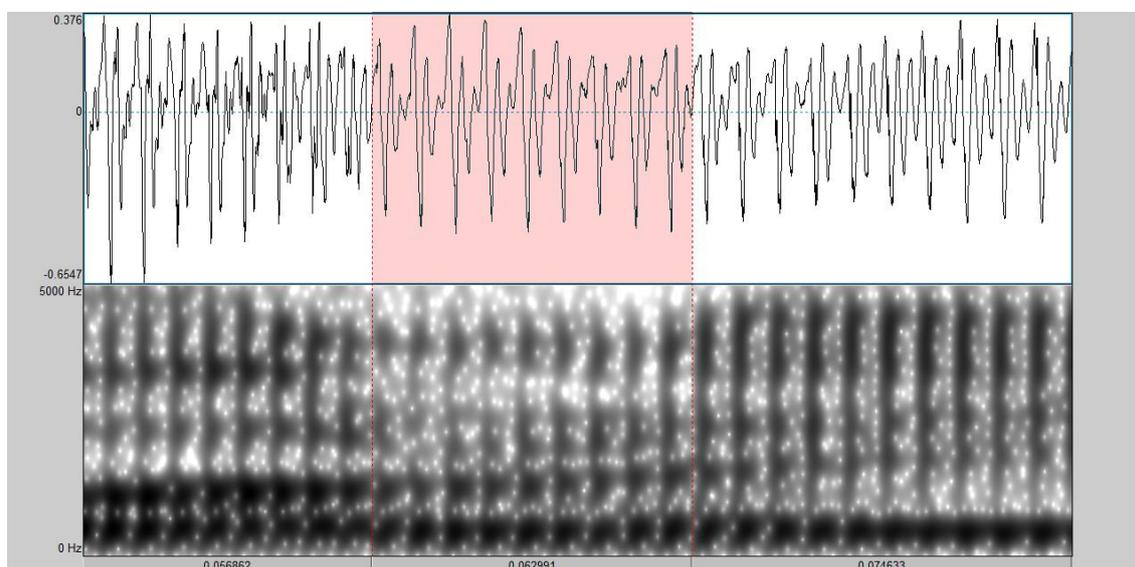
A figura 3 apresenta o oscilograma e o espectrograma referentes à pronúncia da palavra ‘relatório’ pelo informante estadunidense de nível pré-intermediário. O primeiro indício para não se tratar da realização de um *tap* é a continuidade espectral verificada na área delimitada: o *tap* produziria um espaço não preenchido ou esparsamente preenchido, como decorrência da ligeira obstrução verificada com o contato estabelecido entre o articulador ativo e o passivo (cf. figura 2). Como uma consoante aproximante realiza-se com um som contínuo, a descontinuidade espectral não ocorre.



**Figura 3** Oscilograma e espectrograma representando a pronúncia da palavra ‘relatório’ por informante estadunidense de nível pré-intermediário no aprendizado do PB. A área destacada corresponde à consoante aproximante alveolar [ɹ].

A ampliação visual da área delimitada, exibida na figura 4, evidencia, nesse caso específico, uma grande regularidade dos pulsos das ondas. Essa regularidade contribui para a manifestação de uma estruturação formântica no segmento. Observa-se que F3 mostra-se mais baixo na região delimitada do que nas vogais que a envolvem. Os

valores de F3, considerando sua média nos segmentos, o ponto de estabilidade articulatória (medido na duração média), e o ponto de fronteira com a área delimitada, nos segmentos vocálicos, são detalhados na tabela 1.



**Figura 4** Expansão espectrográfica e de forma de onda do trecho correspondente à consoante aproximante alveolar [ɹ], na pronúncia da palavra ‘relatório’ por informante estadunidense de nível pré-intermediário no aprendizado do PB.

	[ɔ]	[ɹ]	[i]
F3 <sub>média</sub>	2708	2346	2583
F3 <sub>estável</sub>	2794	2279	2771
F3 <sub>fronteira</sub>	2400		2489

**Tabela 1** Valores do terceiro formante (F3) para a pronúncia dos segmentos [ɔ], [ɹ] e [i], na palavra ‘relatório’, por informante estadunidense de nível pré-intermediário no aprendizado do PB. Valores expressos em Hertz.

Como referenciado, esse rebaixamento de F3 indica ter havido alguma constrição na região mais baixa da faringe ou na região palatal. Deve-se levar em consideração que F3, na área delimitada, não está realmente próximo de F2. Visualmente, percebe-se que F2 encontra-se a meia distância entre F1 e F3 (no ponto de estabilidade, F1 encontra-se a 409 Hz; F2, a 1332 Hz; F3, a 2279 Hz. A média entre F1 e F3 estaria em 1344Hz.). Essa não-proximidade entre F3 e F2 deve decorrer do fenômeno coarticulatório que prevê uma transição formântica que eleve F2 de sua

posição mais próxima de F1, numa vogal posterior, para uma posição mais distante de F1, numa vogal anterior. Assim, F2, na área delimitada, mostra-se com um movimento ascendente. Esse não-rebaixamento de F2, visualmente perceptível, parece também indicar não ter sido realizado o arredondamento dos lábios.

Quanto à tonalidade percebida no espectrograma, a área delimitada apresenta-se mais clara, comparativamente à tonalidade observada nas duas vogais que a envolvem, o que indica que esse segmento possui uma intensidade menor que a das vogais.

Esses dados indicam, em seu conjunto, que, ao buscar articular o *tap* em posição intervocálica, esse anglofalante produziu uma consoante aproximante alveolar. É importante notar que houve variantes gradientes dentro dessa estratégia, desde a realização de uma consoante aproximante alveolar plena até um *tap* com resquícios articulatorios da aproximante.

#### **b) A lateralização**

Vários sistemas lingüísticos não apresentam a consoante *tap* alveolar [ɾ] integrando o seu quadro fonêmico. Os falantes nativos dessas línguas, como o inglês, o árabe, o búlgaro, o japonês e o chinês, para citar apenas algumas, encontram dificuldades na articulação desse fonema em palavras de línguas que o possuem em seu sistema, como é o caso do português. Uma estratégia muito produtiva no preenchimento articulatorio do *locus* fonológico destinado a ele é a realização da consoante lateral aproximante [l]. A substituição de um fonema pelo outro fundamenta-se em algumas similaridades fonéticas e fonológicas que compartilham entre si, constituindo um grupo que, tradicionalmente, é conhecido como “líquidas”. Toda a caracterização feita a seguir, tanto para a consoante lateral aproximante, individualmente, quanto para o grupo das líquidas, será baseada em Ladefoged & Maddieson (1996), especialmente o capítulo 6.

Ladefoged & Maddieson (*op. cit.*, p. 182) propõem que os sons laterais são produzidos com uma contração da língua de forma a estreitar o seu contorno de lado a lado. Esse estreitamento permite que um volume maior de ar flua por um ou por ambos os lados, volume esse maior que o que flui pelo centro da língua. Foi demonstrado, por meio de estudos palatográficos e de raios-x de várias línguas, que a maioria dos segmentos laterais é produzida com uma oclusão na região dental ou alveolar, e que

essa oclusão é limitada a poucos milímetros na região alveolar, na área atrás dos dentes incisivos, podendo estender-se até os pré-molares. A oclusão não se estende posteriormente, até as regiões molares<sup>9</sup>; ao invés disso, verifica-se que o corpo da língua posiciona-se relativamente baixo na boca atrás da oclusão, o que permite que o ar escape lateralmente de forma anteriorizada até a região palatal. Mesmo nos casos em que não há oclusão central completa (pelo contato do corpo da língua com o palato duro), o que permite que o ar passe por essa região central, o fluxo de ar principal tem um canal de escape maior mais para trás na região bucal. O número máximo de variantes da lateral aproximante vozeada simples já registrado é quatro: a língua Kaitij, por exemplo, apresenta laterais lâmino-dentais, ápico-alveolares, ápico-pós-alveolares e lâmino-pós-alveolares (*op. cit.*, p. 185). No inglês, Dart (1991, *apud op. cit.*, p. 191-192) identificou uma grande variedade de articulações para a aproximante lateral dentro da região dental / alveolar, sendo que a tendência maior é para a realização apical.

No caso do inglês, a consoante aproximante lateral [l] pode articular-se de formas diferentes por influência de consoantes desvozeadas adjacentes (especialmente oclusivas que as precedam), de efeitos coarticulatórios de vogais adjacentes e, ainda, por conta de seu posicionamento na sílaba ou no morfema (*op. cit.*, p. 192). Quanto à articulação de [l] em posição de coda silábica, no PB, os autores indicam, com base em Feldman (1972), que é possível ocorrerem ‘laterais’ completamente sem oclusão, em posições pós-vocálicas (*op. cit.*, p. 193). Possivelmente, embora não houvesse, nesses casos, qualquer tipo de oclusão, haveria uma elevação do ápice da língua em direção à borda alveolar, e esse vestígio de elevação, associado à elevação da parte posterior da língua (produzindo um segmento similar a [ʊ]), assumindo-se o critério de estreitamento no gesto articulatório da língua, seria suficiente para, também nessa posição fonológica e com esse tipo de realização acústica, essas variações serem classificadas como laterais.

Acusticamente, as consoantes laterais aproximantes canônicas são caracterizadas por ressonâncias formânticas bem definidas. Normalmente, F1 é baixo em frequência. A frequência central de F2 pode estar localizada em uma gama larga de variação, dependendo do local da oclusão e do contorno da língua. F3, tipicamente, apresenta uma

---

<sup>9</sup> Em Mid-Waghi, no entanto, uma consoante lateral velar contrasta fonemicamente com uma lateral lâmino-dental e com uma lateral ápico-alveolar. Palatograficamente, a oclusão velar não foi detectável, devido ao seu nível de posteriorização (LADEFOGED & MADDIESON, 1996, p. 185-186).

amplitude forte e alta frequência. Há ainda a possibilidade de vários formantes adicionais muito próximos acima da frequência de F3. Quando se trata de uma realização apical, se a lateral for adjacente a vogais, é possível que se observe uma mudança abrupta na localização do formante, tanto quando a oclusão medial para a lateral é formada quanto para quando é liberada (*op. cit.*, p. 193).

Essas características da consoante aproximante lateral [l] a aproximam, em alguns aspectos, do *tap* alveolar [r]. Essa aproximação é que permite o estabelecimento de uma classe de consoantes conhecidas como “líquidas”. Foneticamente, essas consoantes estão entre as mais sonoras das consoantes orais<sup>10</sup> e, fonotaticamente, são as que têm maior liberdade de ocorrer em encontros consonantais (como é o caso do português, em que uma consoante obstruinte só pode ser seguida por uma consoante líquida, em encontros tautossilábicos). Observe-se, ainda, que algumas poucas línguas apresentam um único fonema líquido cuja pronúncia varia entre a de uma lateral e a de um rótico (*op. cit.*, p. 182).

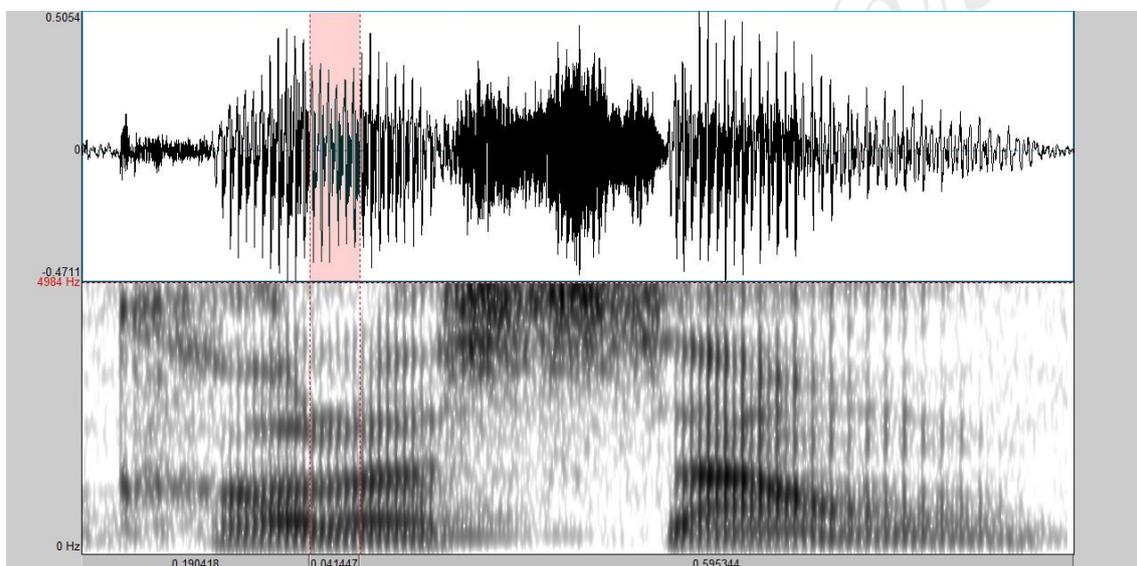
Como Ladefoged & Maddieson (*op. cit.*, p. 243) indicam, existe um membro comum entre os róticos e as laterais: o flap lateral. Como já referenciado, a diferença proposta por esses autores entre o flap e o *tap* se baseia no tipo de contato que se estabelece entre o articulador ativo e o passivo: enquanto, para a produção do flap, esse contato se dá de forma tangencial, para a do *tap*, ocorre um breve contato balístico, geralmente com a região pós-alveolar. No caso do flap lateral, durante sua articulação, um lado da língua permanece baixo, permitindo que o ar flua de forma contínua através desse canal lateral que se forma. Como os autores indicam, é possível que a parte mais alta da língua produza ou não um contato firme o suficiente para impedir o fluxo de ar pelo lado oposto. O som resultante dessa gesticulação lembra, auditivamente, ambos [r] e [l]. A percepção desse som como [r] ou [l] pode estar diretamente relacionada ao que se consideraria uma articulação consistente: vogais posteriores predisporiam a produção (ou percepção) de variantes laterais, enquanto que vogais anteriores sugeririam variantes róticas.

Considerando-se os aspectos da aquisição da linguagem, é interessante notar, com Hernandorena (1990), que a aprendizagem das líquidas ocorre tardiamente. No PB,

---

<sup>10</sup> Esse nível tão elevado de sonoridade é responsável pela semelhança visual observada, espectrograficamente, entre algumas das consoantes líquidas e as vogais, que ocupam o ponto mais elevado na Escala de Sonoridade.

estão entre os últimos sons a serem adquiridos (no inglês, como indicam BALL *et al.*, 2001, os róticos assumem essa posição), e é muito comum a troca de /r/ por /l/ por crianças nessa fase. Essa troca pode ocorrer também em adultos. Embora essa informação diga respeito à aquisição em língua materna, o indício da complexidade dos gestos articulatórios envolvidos na realização das líquidas, mesmo para falantes nativos de uma língua cujo quadro fonêmico inclui o *tap* alveolar, como é o caso do português, parece permitir sugerir que também para os anglofalantes, em cujo sistema lingüístico materno esse fonema não ocorre, a dificuldade articulatória ao tentarem realizá-lo na L2 seja considerável.



**Figura 5** Oscilograma e espectrograma representando a pronúncia da palavra ‘coração’ por informante estadunidense de nível pré-intermediário no aprendizado do PB. A área destacada corresponde à lateralização de uma consoante aproximante alveolar [ɹ].

O mesmo informante estadunidense de nível pré-intermediário que utilizou como estratégia articulatória para a produção do *tap* alveolar a articulação de uma aproximante alveolar valeu-se, em outros momentos, da articulação de um segmento que se assemelha auditivamente a uma consoante lateral, mas conserva semelhança com a consoante aproximante alveolar. Nesse caso, parece mais apropriado sugerir uma qualidade lateral para o segmento, provavelmente na linha do mesmo fenômeno observado por Ladefoged & Maddieson (*op. cit.*, p. 243) e discutido dois parágrafos atrás. A figura 5 apresenta o oscilograma e o espectrograma representando a pronúncia de ‘coração’ por esse informante.

O primeiro ponto a observar na área delimitada são as ressonâncias formânticas: estão bem definidas. É possível perceber, no espectrograma, que o tom de cinza é mais claro que o das vogais que fazem fronteira com esse segmento. A contraparte a ser observada, no oscilograma, é a menor amplitude dos períodos.

Enquanto F1 mostra-se baixo em sua frequência, comparativamente às vogais (valores médios: [o], 578 Hz; [ɪ], 467 Hz; [a], 532 Hz), dentro da expectativa para uma consoante lateral aproximante canônica, F3 não evidencia uma amplitude forte. Visualmente, essa configuração lembra a de uma consoante aproximante alveolar (cf. Figura 3), mas a sensação auditiva oscila entre [l], [ɹ] e [r]. Essa mesma estratégia articulatória verificou-se também nas realizações das informantes de nível intermediário e avançado, ambas de origem inglesa.

### c) O uso de uma vogal epentética

Como estratégia articulatória para a produção da consoante *tap* alveolar, o uso de uma vogal epentética entre ela e a consoante obstruinte que a precede parece ser recorrente mesmo entre falantes nativos das línguas cujo sistema fonológico apresenta esse tipo de encontro consonantal, como se verifica no PB e no espanhol (cf. SILVA, 1996; RAMÍREZ, 2006). O apoio do *tap* entre duas vogais (ou entre um elemento vocálico e uma vogal), interrompendo a seqüência consoante obstruinte + *tap*, possível para os ataques silábicos, coloca-se como um indício da complexidade envolvida na gesticulação articulatória desse segmento, confirmada também em posição final (cf. SILVA, 1996, p. 98, para o PB; cf., também, a nota 6). Um aprendiz do português como L2 cujo sistema lingüístico materno não incluía o *tap*, portanto, previsivelmente recorrerá a essa mesma estratégia<sup>11</sup> para a produção desse segmento em encontros consonantais.

---

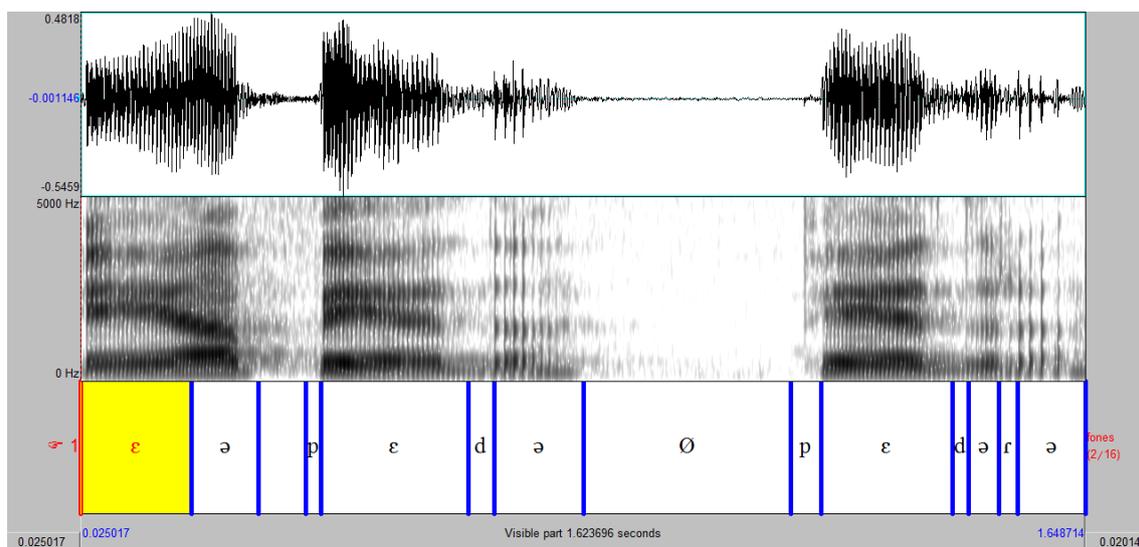
<sup>11</sup> Considera-se, assim, que se trata de uma estratégia articulatória também para o falante nativo do português brasileiro. Observe-se que a pronúncia de formas sincopadas como 'pra' podem ser percebidas como a pronúncia em ritmo acelerado da forma plena equivalente ('para'). Nos casos de problemas de aquisição do *tap* em encontros consonantais na língua materna, seria possível desenvolver um exercício articulatório que inserisse, entre a consoante obstruinte e o *tap*, uma vogal idêntica à do núcleo silábico (o que, naturalmente, produziria uma ressilabificação). Por exemplo, 'prato' seria inicialmente pronunciado [pa.'ra.tu]. Sequencialmente, essa pronúncia seria feita em ritmo mais acelerado, o que permitiria que a vogal (átônica, pretônica) da sílaba inicial tivesse sua duração consideravelmente reduzida, assumindo a configuração de um elemento vocálico e, por conseguinte, a retomada da silabificação original ['pra.tu].

O que permitiria diferenciar um falante nativo do PB de um anglofalante ao utilizarem essa estratégia articulatória na pronúncia de uma palavra do PB possivelmente seria o valor duracional dessa vogal epentética, ou elemento vocálico. Quanto a sua duração, Ramírez (2006, p. 51) detectou, no espanhol, que ela esteve na média de 32% da duração da vogal total, considerando-se os encontros consonantais. Nos dados de Almeida e Dorta (1993, *apud* CLEMENTE & NISHIDA, 2007, p. 77), para falantes da Ilha de Tenerife, ela poderia superar a do *tap*. Silva (1996, p. 69) identificou que a duração da vogal (tônica) seria bem maior que a do elemento vocálico, no PB. Nishida (2005, p. 21), ao caracterizar a estrutura desse segmento, também no PB, indica haver uma estrutura formântica bem definida, com amplitude de onda semelhante à da vogal nuclear do grupo, variando entre 3 e 5 pulsos. Também indicou que, aparentemente, no caso de sílabas átonas, vogal e elemento vocálico teriam duração semelhante, enquanto que, em sílabas tônicas, a duração do elemento seria mais breve que a da vogal.

Essa duração reduzida da vogal epentética, quando produzida por um falante nativo do PB, faz com que ela normalmente passe despercebida auditivamente, causando a impressão, em um encontro consonantal, de que a consoante obstruinte é imediatamente seguida pelo *tap* (CLEMENTE & NISHIDA, 2007, p. 79-80). No entanto, impressionisticamente, esse segmento já teria sido notado no espanhol (peninsular e americano) por Lenz (1892), como indica Nishida (2005, p. 8), citando Quilis (1993), caracterizado como a produção de um som glótico entre a consoante obstruinte e o *tap*. Caso a duração desse segmento seja maior, sua realização tenderá a ser percebida auditivamente, e parece ser esse o caso de alguns anglofalantes pronunciando a seqüência consoante obstruinte + *tap* alveolar + vogal.

É possível que um anglofalante também omita a realização do *tap* alveolar, produzindo o seu apagamento fonético. A figura 6 representa a pronúncia de ‘e a pedra’, pelo informante estadunidense de nível básico. A palavra ‘pedra’ foi pronunciada duas vezes. Na primeira, o informante não realizou o *tap*. Percebendo isso, corrigiu sua pronúncia, inserindo, na segunda tentativa, esse segmento. Entre a consoante oclusiva [d] e ele, verifica-se a presença da consoante epentética. A observação espectrográfica da vogal nuclear indica, na primeira metade, um espaçamento bem marcado entre os ciclos das pregas vocais, caracterizando uma realização vibrante. Não se trata, nesses

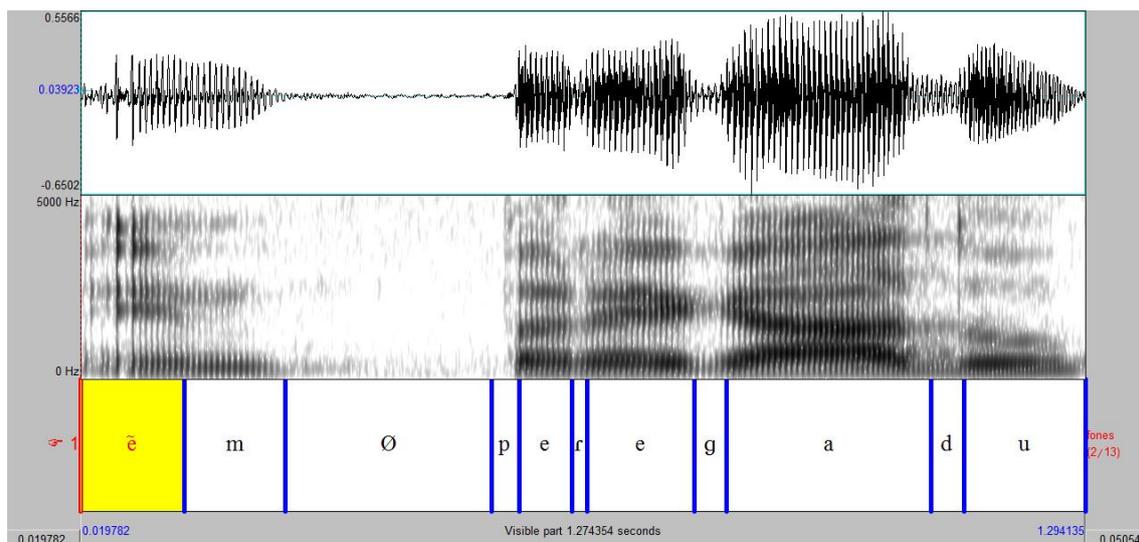
dois casos, da realização de uma consoante vibrante (múltipla) (cf. essa caracterização mais à frente), que poderia ser percebida como outra estratégia utilizada pelo informante para a produção do *tap* alveolar, mas de uma vibração no próprio núcleo vocálico. Observe-se que essa ocorrência se deu em uma sílaba pós-tônica.



**Figura 6** Oscilograma e espectrograma representando a pronúncia de ‘e a pedra’, com alinhamento sonoro<sup>12</sup>, por informante estadunidense de nível básico no aprendizado do PB. Observe-se a não-realização do *tap* alveolar na primeira ocorrência de ‘pedra’, e o uso de uma vogal epentética entre [d] e [r], na segunda. A caracterização espectrográfica da vogal schwa [ə], em posição átona final, indica ter havido uma vibração durante a realização desse segmento, tanto na primeira quanto na segunda ocorrências.

Na figura 7, representa-se a pronúncia de ‘empregado’ pelo mesmo informante estadunidense de nível básico. O silêncio observado entre a primeira e a segunda sílabas da palavra (261ms) indica a hesitação do informante em pronunciar a segunda sílaba (que apresenta um ataque silábico em que o *tap* se faz presente), como em uma preparação para os gestos articulatórios necessários. Também nesse caso, verifica-se a manifestação da vogal epentética.

<sup>12</sup> Por conta dos efeitos de coarticulação, optou-se, para o alinhamento sonoro, pelo estabelecimento da fronteira entre a vogal tônica [ε] e a consoante oclusiva alveolar sonora [d] nesses pontos.



**Figura 7** Oscilograma e espectrograma representando a pronúncia de ‘empregado’, com alinhamento sonoro, por informante estadunidense de nível básico no aprendizado do PB. Observe-se o uso de uma vogal epentética entre [p] e [r].

	[ə]	[ɹ]	[ə]	[e]	[r]	[e]
<b>duracão</b>	49	29	126	67	19	136
<b>F1<sub>estável</sub></b>	412	339	461	466	393	471
<b>F2<sub>estável</sub></b>	1437	1469	1388	1394	1467	1776
<b>F3<sub>estável</sub></b>	2520	2657	2500	2376	2439	2309

**Tabela 2** Valores duracionais e dos três primeiros formantes (F1, F2, F3, medidos no ponto de estabilidade acústica) para a realização dos segmentos [ə], [ɹ] e [ə], na palavra ‘pedra’ (segunda ocorrência, na figura 6), e [e], [r] e [e], na palavra ‘empregado’, por informante estadunidense de nível básico no aprendizado do PB. Valores expressos em milissegundos e em Hertz.

Como se pode observar na tabela 2, os três primeiros formantes das vogais epentéticas apresentam valores muito próximos dos verificados nas vogais nucleares, tanto na pronúncia de ‘pedra’ (segunda ocorrência) quanto na de ‘empregado’. A diferença mais acentuada para os valores de F2 da vogal nuclear em relação à vogal epentética (1776 Hz contra 1394 Hz, respectivamente, em ‘empregado’), provavelmente se deve ao efeito coarticulatório induzido pela consoante oclusiva velar sonora [g], no ataque da sílaba seguinte. É interessante observar, portanto, que tanto em posição pré-tônica quanto em posição pós-tônica, a vogal epentética assume configuração semelhante à da vogal nuclear. Esses dados estão em sintonia com a indicação de Nishida (2005, p. 34), segundo a qual a estrutura formântica desses elementos vocálicos

seria condicionada pela qualidade da vogal nuclear do grupo. Ressalte-se que, aqui, trata-se da pronúncia de um anglofalante.

Quanto aos aspectos duracionais, detalhados na tabela 2, a vogal epentética em ‘pedra’ (segunda ocorrência) realizou-se em 49ms, o que corresponde a 28% da duração total (vogal epentética + vogal nuclear); em ‘empregado’, realizou-se em 67ms, correspondendo a 33% da duração total. Como já referido, Ramírez (2006, p. 51), analisando dados do espanhol, identificou que a produção do elemento vocálico corresponderia a cerca de um terço da duração da vogal total, proporção essa que se assemelha a esses dados do informante estadunidense. No entanto, o autor não menciona se, auditivamente, foi possível perceber a realização desse elemento vocálico. No caso das realizações aqui apresentadas, essa percepção é uma realidade em ambas as ocorrências, um pouco menos acentuada para ‘pedra’ do que para ‘emprego’, mas plenamente identificáveis auditivamente. Essa marca auditiva, nesses casos, é responsável pelo reconhecimento do informante como estrangeiro ao PB.

#### **d) O uso de uma consoante vibrante**

Dentre as consoantes vibrantes, a que ocorre como alofone do *tap* alveolar, no PB, é a vibrante alveolar, identificada como “trill” alveolar no Alfabeto Fonético Internacional e representada pelo símbolo [r]. No PB, ela não possui caráter fonêmico, diferentemente do que ocorre no espanhol (‘perro’ [ˈpero] vs. ‘pero’ [ˈpero]).

Como indicam Ladefoged & Maddieson (1996, p. 217), uma vibrante tem como característica primária a vibração de um órgão da fala contra outro, em condições aerodinâmicas apropriadas. Essa vibração será resultado da passagem da corrente de ar que força a abertura de um canal entre uma das partes macias e móveis do trato vocal em contato com outra superfície, causando um padrão repetitivo de abertura e fechamento do fluxo desse canal. Como os autores sugerem, esse movimento é essencialmente semelhante à vibração das pregas vocais durante o vozeamento. Tanto na produção de uma vibrante quanto na vibração das pregas vocais, não há ação muscular controlando cada vibração individualmente e, para que a vibração ocorra, é necessário haver um tamanho específico para a abertura, e o fluxo de ar deve estar dentro de limites adequados. Como essas condições podem variar consideravelmente, é possível que a consoante vibrante, canonicamente vozeada, seja realizada total ou parcialmente sem vozeamento (especialmente se acompanhada de um segmento

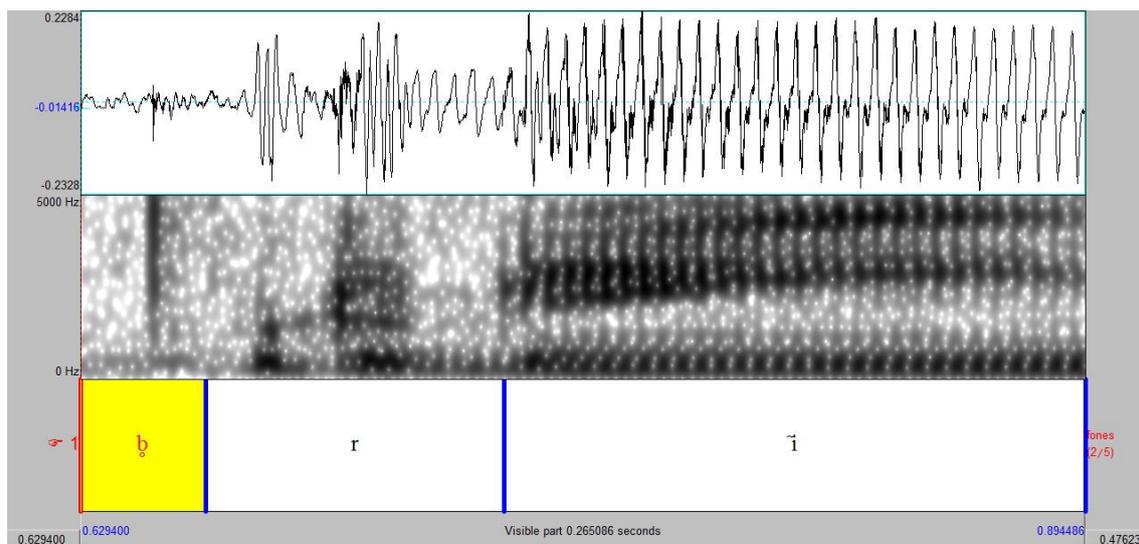
desvozeado) e, numa língua em que possua caráter fonêmico, um seu alofone seja produzido sem a vibração. Isso decorre de um potencial conflito entre uma definição acústica, que prevê mais do que um período de vibração efetiva, e uma definição articulatória, que prevê um posicionamento tal dos articuladores que a vibração ocorra, consideradas as condições aerodinâmicas apropriadas.

Ainda segundo Ladefoged & Maddieson (*op. cit.*), uma vez que a gesticulação envolvida na produção de uma vibrante envolve o contato entre dois articuladores, formando-se uma oclusão, posteriormente liberada, o aspecto visual do espectrograma indicará a intercalação dessas fases de oclusão e liberação, representadas, respectivamente, por áreas claras (em que a energia formântica é ausente ou muito fraca) e escuras. As áreas escuras apresentam características acústicas semelhantes às das vogais, e sua duração é da ordem de 25ms cada, a mesma normalmente observada também para as áreas claras. Linguisticamente, as vibrantes normalmente são formadas por dois a cinco períodos (ciclos completos de oclusão e liberação da oclusão).

Ao buscar uma articulação para o *tap* alveolar, a informante inglesa de nível intermediário produziu uma vibrante na sílaba ‘bri’ da palavra ‘abobrinha’, como representado na figura 8. Nessa realização específica, a informante desvozeou a consoante oclusiva bilabial sonora, o que causou o desvozeamento tanto do primeiro quanto do segundo intervalos de oclusão da consoante vibrante<sup>13</sup>. É interessante observar que, imediatamente após a liberação da oclusão do segundo ciclo da vibrante, manifesta-se uma aproximante, com duração de 14ms (cf. ocorrência semelhante nos dados discutidos em Ladefoged & Maddieson, 1996, p. 219, para realizações do italiano padrão). Essa oscilação articulatória envolvendo tanto uma vibrante quanto uma aproximante evidencia a complexidade na produção do *tap* alveolar, especialmente para uma informante cujo sistema lingüístico materno não possui esse segmento em seu quadro fonêmico.

---

<sup>13</sup> O alinhamento entre as consoantes oclusiva e vibrante, nesse caso, considerou apenas a mudança no movimento formântico entre elas, uma vez que, tendo ocorrido desvozeamento no primeiro ciclo da vibrante, não foi possível determinar o VOT (*Voice Onset Time*, correspondente à duração temporal decorrida entre a explosão de uma consoante oclusiva e o início da vibração das pregas vocais) que se segue à explosão da consoante oclusiva. O espaçamento claro no espectrograma, entre a explosão da oclusiva e a liberação do primeiro contato na produção da vibrante deve ser dividido, portanto, entre o VOT da oclusiva e a fase de oclusão da vibrante.



**Figura 8** Oscilograma e espectrograma representando a pronúncia da sílaba ‘bri’, na palavra ‘abobrinha’, com alinhamento sonoro, por informante inglesa de nível intermediário no aprendizado do PB.

### Considerações Finais

Os dados aqui apresentados não devem ser percebidos como estratégias únicas utilizadas por anglofalantes na realização do *tap* alveolar no PB. É possível que existam outras, influenciadas inclusive pelas diversas variedades dialetais de que eles provenham. O que se deve observar é que, em face de um desafio articulatório, o aprendiz da L2 experimenta (no sentido restrito do termo) possibilidades intermediárias que, mais próximas do seu sistema fonológico nativo, conduzam o seu sistema articulatório na aproximação dos (novos) gestos necessários a uma realização semelhante ou muito próxima à de um falante nativo.

Cada uma dessas estratégias utilizadas pode apresentar variações, dependendo inclusive da forma como segmentos adjacentes são produzidos na língua materna do aprendiz. Uma vez que, no inglês, a consoante oclusiva velar surda [k] é produzida com aspiração em posição inicial de palavra e em outras posições quando em sílaba tônica, a pronúncia da palavra ‘cravo’, por exemplo, envolverá estratégias para a produção do *tap* alveolar e, ao mesmo tempo, para que a consoante oclusiva não seja aspirada.

Qualquer que seja a situação, deve-se considerar que a consciência quanto à gesticulação envolvida na realização de cada fonema, mais do que uma teorização, pode se colocar como mecanismo de aceleração do processo de aquisição da L2.

## Referências Bibliográficas

- BALL, M. J.; MULLER, N.; MUNRO, S. The acquisition of the rhotic consonants by Welsh-English bilingual children. **International Journal of Bilingualism**, v.5, n.1, jan 2001.
- BEST, C. T. A direct realist view of cross-language speech perception research with adults. In: STRANGE, W. **Speech perception and linguistic experience: issues in cross-language research**. Baltimore: York Press, 1995.
- BOERSMA, P.; WEENINK, D. (2006). **Praat: doing phonetics by computer** (Version 4.5.01) [Computer program]. Disponível em: <<http://www.praat.org/>>.
- CARVALHO, A. M. Português para falantes de espanhol: perspectivas de um campo de pesquisa. **Hispania**, v. 85, n. 3, p. 597-608, 2002.
- CLEMENTE, F. C.; NISHIDA, G. Características acústicas do *tap* em coda: dados do português de Curitiba e do espanhol de Buenos Aires. **Revista Letras**. Curitiba: Editora UFPR, n. 73, set./dez. 2007.
- DUTRA, A. **Aquisição do português como língua estrangeira: fenômenos de variações no âmbito fonológico**. Araraquara, 2008. Tese (Doutorado) – Faculdade de Ciências e Letras, Universidade Estadual Paulista.
- ECKMAN, F. R. From phonemic differences to constraint rankings: research on second language phonology. **Studies in Second Language Acquisition**, Cambridge University Press, v. 26, n.4, p. 513-549, 2004.
- FACE, T. L. Intervocalic rhotic pronunciation by adult learners of Spanish as a second language. In: KLEE, C. A.; FACE, T. L. **Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages**, Somerville, MA: Cascadilla Proceedings Project, 2006.
- HASHI, M.; HONDA, K.; WESTBURY, J. R. Time-varying acoustic and articulatory characteristics of American English [ɹ]: a cross-speaker study. **Journal of Phonetics**, v. 31-1, 2003.
- HERNANDORENA, C. L. **Aquisição da fonologia do português: estabelecimento de padrões com base em traços distintivos**. Porto Alegre, 1990. Tese (Doutorado) – Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul.
- LADEFOGED, P. American English. **Handbook of the International Phonetic Association**. A guide to the use of the International Phonetic Alphabet, Cambridge: Cambridge University Press, 1999.
- LADEFOGED, P.; MADDIESON, I. **The sounds of the world's languages**. Oxford: Blackwell Publishers, 1996.

MADDIESON, I. A survey of liquids. **UCLA working papers in Phonetics**. N. 50, 1980.

NISHIDA, G. **Análise acústica do tap em grupos no PB**. Curitiba, 2005. Monografia. Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná.

QUILIS, A. **Fonética acústica de la lengua española**. Madrid: Gredos, 1981.

RAMÍREZ, C. J. Acoustic and perceptual characterization of the epenthetic vowel between the clusters formed by consonant + liquid in Spanish. In: DÍAS-CAMPOS, M. (ed.) **Selected proceedings of the 2<sup>nd</sup> Conference on Laboratory Approaches to Spanish Phonetics and Phonology**, Somerville, MA: Cascadilla Proceedings Project, 2006.

REIS, C.; ESPESSER, R. Estudo eletropalatográfico de fones consonantais e vocálicos do português brasileiro. **Estudos da Língua(gem)**. N. 3, junho de 2006.

SILVA, A. H. P. **Para a descrição fonético-acústica das líquidas no português brasileiro: dados de um informante paulistano**. Campinas, 1996. Dissertação (Mestrado) – Instituto de Estudos da Linguagem, Universidade de Campinas.

ZIMMER, M. C. Um estudo conexional da transferência do conhecimento fonético-fonológico do PB (L1) para o inglês (L2) na leitura oral. In: POERSCH, M.; ROSSA, A. **Processamento da linguagem e conexionalismo**. Santa Cruz: EDUNISC, 2007.

Domínios de Lingu@gem