

# Organização temporal de encontros vocálicos no Português Brasileiro e a relação entre Fonética e Fonologia

*Temporal organization of vowel clusters in Brazilian Portuguese and the relationship between Phonetics and Phonology*

Adelaide H. P. Silva

Universidade Federal do Paraná – Curitiba – Paraná – Brasil



**Resumo:** A discussão acerca da diferença entre ditongos e hiatos no português brasileiro (PB) ocupa há tempos os fonólogos, mas não há análises fonéticas disponíveis desses encontros vocálicos no PB. Num momento em que a teoria fonológica tem-se valido cada vez mais frequentemente dos dados fonéticos, como ficam evidentes os trabalhos conduzidos à luz da Fonologia de Laboratório, é de grande importância que haja caracterizações acústicas sobre hiatos e ditongos no PB. Objetiva-se, então, neste estudo, a partir de um breve panorama sobre estudos fonológicos dos encontros vocálicos nesta língua, oferecer uma descrição acústica inicial dos encontros vocálicos no PB. Essa descrição aponta para uma diferença na duração da transição do primeiro para o segundo evento acústico desses encontros. Os achados, argumenta-se então, sugerem que a variável tempo é crucial para a organização da cadeia da fala, o que leva à proposta de uma análise fonológica que incorpore o dado fonético.

**Palavras-chave:** Encontros vocálicos; Análise acústica; Duração; Organização temporal

**Abstract:** Phonologists have been discussing for a long time the difference between hiatus and diphthongs in Brazilian Portuguese, but until now there are no phonetic analyses for these vowel clusters in BP. Nevertheless, such analyses are crucial in a scenario, triggered by Laboratory Phonology, where phonetic data play a special role to phonological analysis: Incorporating phonetic data can yield more adequate and parsimonious phonological explanations for facts of the speech chain. Departing from a brief overview on phonological analyses on vocalic clusters in BP, this paper presents a preliminary acoustic description for diphthongs and hiatus in BP. The description points to a distinct temporal organization among the acoustic events that constitute the clusters. The findings, then, suggest timing is crucial for the organization of the speech chain. This argument leads to the proposition of a phonological analysis that incorporates phonetic data, based on the temporal organization of speech units.

**Keywords:** Vowel clusters; Acoustic analysis; Duration; Temporal organization

## 1 Introdução

O estatuto fonológico dos encontros vocálicos no português brasileiro (PB) é discutido há tempos na literatura fonológica, desde Câmara Jr. (1969).

Ao tratar da estrutura da sílaba no PB, o autor questionava a própria existência dos ditongos na língua<sup>1</sup>,

ou seja, haveria uma sílaba constituída de uma vogal seguida por uma “parte decrescente” ou o que existe na língua são duas vogais que constituem núcleos de sílabas contíguas? Argumentando que pares como “sai/sai” não constituiriam pares mínimos, em razão da diferença na posição do acento, Câmara Jr. (op. cit.) focaliza sequências do tipo rio/riu para sustentar o argumento de que em alguns dialetos do PB, como por exemplo o carioca, haveria a oposição fonológica entre ditongo e hiato. Note-se que, no caso deste último par, o acento ocuparia o mesmo lugar na palavra. O argumento de Câmara Jr., então, se relaciona ao fato de que uma sequência como “rio” seria

<sup>1</sup> Essa pergunta parece ter uma motivação histórica: na evolução do latim para o português, houve o surgimento de encontros vocálicos em razão, por exemplo, do apagamento de uma consoante. Assim, por exemplo, de “fidelem” tem-se “fiel”, em que as duas vogais são heterossilábicas. Teríamos, então, apenas hiatos na língua portuguesa?

produzida como hiato, ainda que com os mesmos sons que uma sequência como “riu”, produzida como ditongo.

Assumindo a existência de ditongos e hiatos na língua, há uma outra questão a se resolver, qual seja, o estatuto da unidade seguinte à vogal nuclear da sílaba, ou da vogal presente na “parte decrescente” da sílaba, em palavras como “riu” ou “pai”, isto é, seria essa unidade uma consoante ou uma vogal?

Baseando suas observações em fatos frequentes na língua, como a monotongação dos ditongos e a variação presente na divisão silábica de sequências de vogal + vogal alta em posição átona como, por exemplo, “vai-da-de/va-i-da-de”, Câmara Jr. afirma que a interpretação mais adequada é a de que a unidade à direita da vogal em encontros vocálicos seja uma vogal assilábica, ou uma semivogal. A estrutura silábica contendo encontros vocálicos no PB, seria, então, por essa proposta, do tipo VV.

Cabe ainda comentar, no que concerne à tipologia dos ditongos, que Câmara Jr. preconiza a existência majoritariamente de ditongos decrescentes na língua – ao todo seriam onze. O ditongo crescente ocorreria num ambiente específico, a saber, depois de oclusiva bilabial e antes da vogal nuclear da sílaba, como em “quais”.

Bisol (1989), por sua vez, argumenta sobre a existência de ditongos falsos e ditongos verdadeiros no PB. Para a autora, os ditongos verdadeiros são os que se constituem de uma sílaba pesada na forma subjacente, travada por uma semivogal, como, por exemplo, “sábia”. Os ditongos falsos, por outro lado, se constituem, na subjacência, de duas sílabas leves, como em “sabia”.

Em estudo posterior, e à luz da Fonologia Lexical, Bisol (1999) chama a atenção para o fato de que os ditongos verdadeiros (decrescentes) se formam num nível lexical; os ditongos falsos (crescentes) se formariam no nível pós-lexical, por ressilabificação. Assim, “história”, por exemplo, que no nível lexical teria duas sílabas leves, depois da tônica, sofreria um processo de ressilabificação, no nível pós-lexical, donde resultaria, depois da sílaba tônica, uma sílaba constituída de *onset* ramificado precedendo o núcleo<sup>2</sup>.

É interessante notar que a proposta de Bisol (1989, 1999) converge com a análise apresentada por Câmara Jr. (op. cit.) no que concerne aos tipos de ditongos existentes no PB: como observa Tornquist (2013), ambos

concordam com a inexistência de ditongos crescentes no PB, apontando para uma distinção entre ditongos e hiatos.

Uma pergunta possível, e decorrente destas análises, é: se há uma diferença fonológica entre ditongos e hiatos, a expectativa é que essa diferença seja observada também no nível fonético. Assim sendo, em que consistiria? Ou, colocando de outro modo: como se diferenciam, foneticamente, hiatos de ditongos<sup>3</sup>?

Chitoran e Hualde (2007) investigam sequências do tipo *iV* que podem apresentar realização variável, i.e., podem ser produzidas tanto como hiatos quanto como ditongos em línguas românicas. Baseados em dados diacrônicos e em dados sincrônicos de cinco dessas línguas – francês, espanhol, romeno, português europeu e português brasileiro –, argumentam que sincronicamente há nelas uma tendência geral de que hiatos, sobretudo aqueles em posições átonas, se realizem como ditongos. Essa tendência se explicaria pela estabilidade da articulação dos ditongos. Argumentam também que haveria: a) um “efeito atrator estrutural” de ditongos do tipo [jV] pré-existentes nessas línguas e b) efeitos prosódicos de alongamento que resultariam em padrões duracionais distintos para hiatos e ditongos nessas línguas, uma vez que sequências resistentes à ditongação favoreceriam vogais longas. Esta poderia ser, segundo os autores, uma das razões pelas quais se verificaria a variabilidade entre ditongos e hiatos no interior das línguas investigadas.

À luz do que foi exposto, neste estudo perseguimos a hipótese de que a diferença entre ditongos e hiatos preconizada por autores como Câmara Jr. (1969) ou Bisol (1989) se pode explicar em termos de padrões duracionais distintos, os quais, por sua vez, se podem relacionar ao nível prosódico da língua, como acento, a exemplo do que argumentam Chitoran e Hualde (op. cit.).

Diferentemente, porém, desses autores, nosso objetivo aqui, conforme se exporá na seção 2, é investigar sequências de vogais que tradicionalmente seriam tipificadas como opondo hiatos a ditongos, a exemplo de “sai/sai” ou “salde/saúde”, ou seja, não se tratará exclusivamente de sequências *iV* como fizeram esses autores, mas de sequências que são objeto de estudo de Câmara Jr. ou Bisol.

Nossa hipótese de trabalho encontra respaldo num programa de pesquisa que se inicia na década de 1980 e que visa a integrar Fonética e Fonologia, de modo a desfazer o abismo que se criara entre ambas as disciplinas, sobretudo no estruturalismo linguístico. Partindo da observação de que a organização temporal tem papel fundamental na organização do nível sonoro da fala (cf., e.g., FOWLER, 1980; BROWMAN e GOLDSTEIN, 1992; BROWMAN e GOLDSTEIN, 1995), esse programa traz o dado fonético para as representações fonológicas, o que pode

<sup>2</sup> Estamos aqui assumindo a estrutura interna da sílaba proposta por Selkirk (1984) e que prevê que uma sílaba é constituída de *onset* e rima. Esse constituinte, por sua vez, se desdobra em núcleo e coda. Não é objetivo deste trabalho confrontar e discutir diferenças na estrutura interna da sílaba concebida pelos diferentes autores mencionados, daí apenas a menção à estrutura silábica prevista por eles.

<sup>3</sup> Frise-se que se assumirá, aqui, a mesma perspectiva adotada por Câmara Jr. e Bisol, comentada anteriormente e segundo a qual não existem ditongos crescentes no PB.

garantir a essas representações maior parcimônia e maior poder explicativo, segundo advoga, por exemplo, Ohala (2007), e como tentaremos argumentar nas últimas seções deste estudo.

## 2 Experimento de produção

Assumindo – como mencionado na Introdução deste estudo – que a incorporação do dado fonético aos estudos fonológicos pode trazer a estes importantes contribuições, no sentido de fornecer subsídios para representações mais parcimoniosas, e perseguindo a hipótese de Chitoran e Hualde (2007), segundo a qual a diferença entre os tipos de ditongos, ou entre ditongos e hiatos, é dada pela organização temporal das unidades constituintes da cadeia da fala, realizou-se um experimento de produção em que se analisou a produção de pares como “cais/caís”; “rói/roí”.

O *corpus* do experimento foi elaborado de modo a contemplar sequências do tipo VA, onde “V” é qualquer vogal oral do português e “A”, aproximante, tanto a palatal [j] como a lábio-velar [w]. Tomando as considerações de Câmara Jr. (1969) sobre a distinção entre hiatos e ditongos, teríamos, no *corpus*, encontros que se diferenciariam quanto à posição do acento. Mas, dado que 1) no dialeto tomado para a investigação não se produzem distinções como aquela que Câmara Jr. apontava para o dialeto carioca da metade do século XX e mencionada na Introdução; 2) há uma certa convenção, em grande parte alimentada pelas descrições tradicionais da gramática, que estabelece uma distinção entre ditongos e hiatos, optou-se por partir desta distinção mais geral para investigar a produção dessas sequências de vogais. Desta forma, o *corpus* do experimento contém as seguintes palavras:

**Quadro 1** – *Corpus* do experimento

“Ditongos”	“Hiatos”
cais	caís
mais	maís
mói	moí
pais	país
reunião	reúne
rói	roí
arruinar	ruína
sábua	sabia
sai	saí
salde	saúde

Ainda sobre a constituição do corpus, há que se notar três pontos:

- a) embora “maís” não seja uma palavra frequente, ela é dicionarizada para o português (vide, e.g.,

Houaiss, Villar e Franco, 2009) e se refere a uma espécie particular de milho;

- b) Os pares arruinar/ruína e reunião/reúne se diferenciam dos demais pares do *corpus* quanto à posição que os encontros vocálicos ocupam relativamente ao acento lexical, pois tanto em “arruinar” como em “reunião<sup>4</sup>”, o acento lexical encontra-se na última sílaba da palavra, depois do encontro vocálico. Além disto, as palavras desses pares têm um número maior de sílabas que as demais palavras do conjunto de dados. Apesar desta diferença, optou-se por manter os pares no *corpus*, inclusive para verificar se haveria alguma diferença entre os padrões duracionais dos encontros vocálicos nesses pares, quando comparados aos outros dados;
- c) O par “salde/saúde” foi incorporado considerando-se que, no dialeto dos sujeitos tomados para este estudo, a lateral sofreu vocalização o que, consequentemente, gera um encontro contendo uma aproximante lábio-velar.

Elaborado o *corpus* do experimento, introduziram-se as palavras-alvo na sentença-veículo “Digo \_\_\_\_\_ cada vez”. O conjunto total de dados, constituído pelas palavras-alvo do Quadro 1 e por distratores<sup>5</sup> na razão de dois distratores para um alvo, foi lido cinco vezes por três informantes do sexo feminino, todas com terceiro grau completo e naturais de Curitiba. Os dados foram apresentados em ordem aleatória a cada uma das informantes e durante cada uma das cinco repetições. Tem-se, então, um total de 300 dados para análise (20 palavras-alvo × 3 informantes × 5 repetições).

A coleta de dados foi feita numa cabine com tratamento acústico, no Laboratório de Estudos Fônicos da Universidade Federal do Paraná, utilizando-se um microfone Shure unidirecional e o *software* Audacity. Os dados foram amostrados a uma taxa de 44100Hz.

Colhidos os dados, eles foram submetidos à análise acústica através do *software* Praat (BOERSMA e WEENINK), conforme será exposto na próxima seção.

## 3 Análise acústica e quantificação dos dados

A análise acústica consistiu na extração das medidas de duração da transição que se verifica no encontro vocálico. A escolha desse parâmetro em detrimento

<sup>4</sup> Frise-se que, nesta palavra, o encontro vocálico focalizado é [ew] e não o encontro vocálico nasal ao final da palavra. Verificar que o comportamento do encontro vocálico nasal foge ao escopo deste estudo.

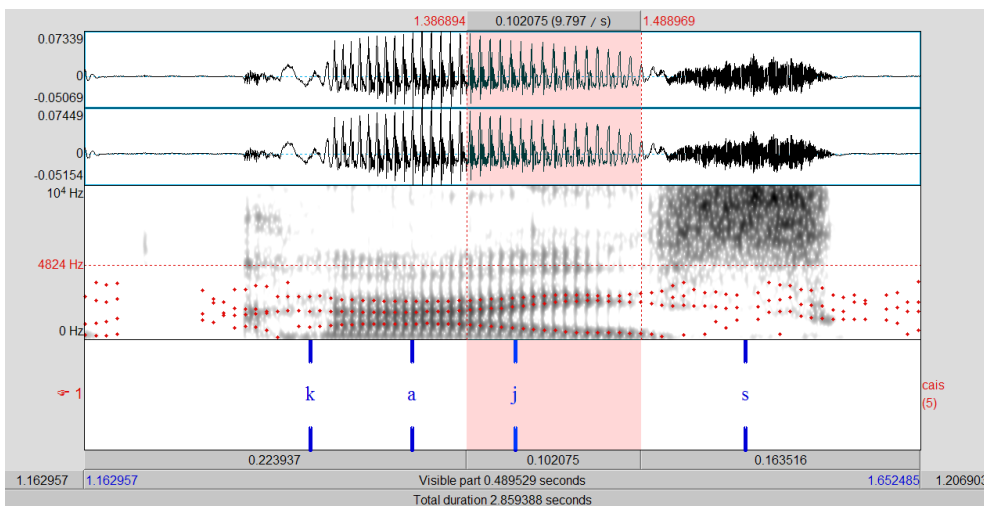
<sup>5</sup> Palavras de mesma estrutura silábico-acental dos alvos, que têm apenas o objetivo de desviar a atenção dos sujeitos sobre o que investigava o experimento como, por exemplo, “vez”.

de outros, como a frequência dos formantes de vogais e aproximantes presentes nos encontros vocálicos investigados, justifica-se pela hipótese deste estudo de que a diferença entre ditongos e hiatos pode ser explicada em termos de padrões duracionais distintos, os quais, por sua vez, se podem relacionar ao nível prosódico da língua. Essa hipótese, como observado na Introdução, baseia-se em estudos como os de Fowler (1980), Browman e Goldstein (1992) e Browman e Goldstein (1995), por exemplo, que argumentam sobre o papel primordial da variável tempo na organização do nível sonoro das línguas.

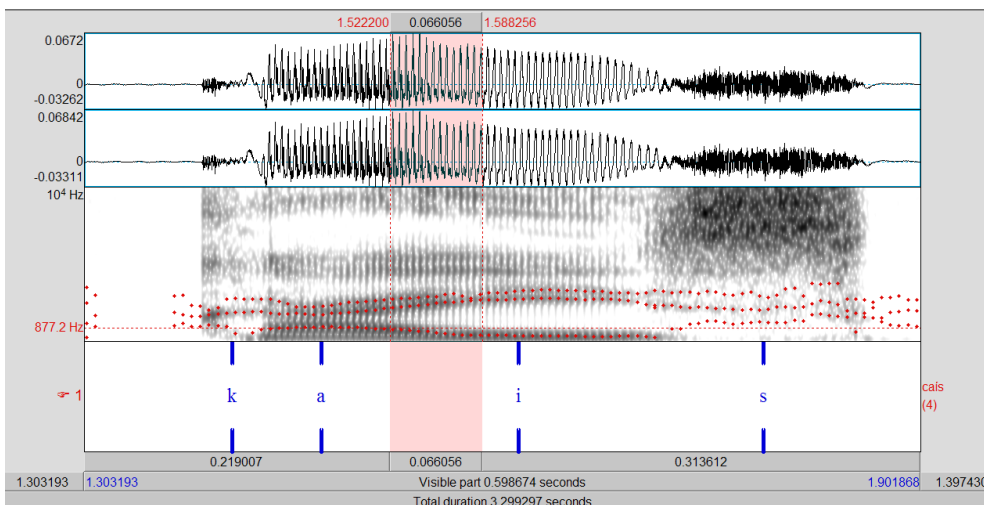
Além disto, se observamos as palavras-alvo do *corpus*, vemos que os eventos acústicos que constituem os dois membros de cada par são os mesmos, conforme se verifica nas Figuras 1 e 2. Elas trazem a forma de onda e o espectrograma do par “cais/caís” produzido por uma das três informantes do experimento. Os espectrogramas evidenciam a presença de uma vogal seguinte à oclusiva e, na sequência, uma transição longa e com trajetórias dos formantes muito marcadas entre essa vogal e o evento acústico seguinte, que se assumirá aqui tratar-se de uma aproximante. Isto porque descrições acústicas de

aproximantes, como as de Kent e Read (1992), apontam justamente para configurações análogas a esta como característica acústica dessa classe de sons, ou seja, a presença da transição é *conditio sine qua non* para a existência desses encontros vocálicos. Deve ficar claro, portanto, que frequência de formantes não distinguiria os sons dos encontros vocálicos dos pares.

Para medirmos a duração da transição da vogal para a aproximante procedemos da seguinte maneira: posicionamos o primeiro cursor no pico correspondente ao momento em que, no espectrograma, começa a haver mudança na trajetória dos formantes – em especial do primeiro (F1) e do segundo (F2) – e medimos até o último pulso da forma de onda em que se verifica movimento na trajetória dos formantes. Nas Figuras 1 e 2, as linhas verticais pontilhadas selecionam as transições focalizadas. Note-se que, no dado da Figura 1, considerou-se que a transição se estende até o final do encontro vocálico porque a trajetória dos formantes não se estabiliza, ao contrário do dado da Figura 2, em que se vê um ponto em que a trajetória fica estável. Esse ponto foi considerado o final da transição para casos análogos.



**Figura 1** – Informante AP\_5ª repetição, dado “cais”



**Figura 2** – Informante AP\_5ª repetição, dado “caís”



Obtidos os valores da duração absoluta das transições de cada um dos membros dos pares examinados, procedeu-se ao cálculo da duração relativa. A duração relativa consiste numa normalização dos dados e é calculada dividindo-se o valor da duração absoluta do evento acústico sob análise pelo valor da duração absoluta da palavra em que o evento se encontra. Em seguida, multiplica-se o resultado obtido por 100, ou seja, tem-se, como resultado final do cálculo, o percentual ocupado por um dado evento acústico no interior da palavra. Esta normalização é indicada para evitar que fatores prosódicos, como velocidade de fala, possam enviesar os resultados.

A Tabela 1 expõe os valores médios da duração relativa para ditongos e hiatos produzidos por cada uma das três informantes do estudo. Traz também a média da duração relativa, tanto de ditongos como de hiatos, considerando todos os dados de todas as três informantes. Finalmente, fornece o desvio padrão para os dados de todas as informantes.

**Tabela 1** – Valores da duração relativa para os dados de cada informante, média geral e desvio padrão

Dur. Relativa (em %)	AP	JI	MJ	N	Média	Desvio Padrão
Ditongos	29,72	27,8	26,04	150	27,85	7.95
Hiatos	13,52	14,15	13,3	150	13,65	3.47

Os dados da Tabela 1 apontam sobretudo para uma duração maior das transições dos ditongos do que das transições dos hiatos. Note-se que, para cada uma das três informantes, a duração relativa da transição dos ditongos é praticamente igual ao dobro da duração relativa da transição dos hiatos – nos dados da informante JI – e chega a ser maior que o dobro da duração relativa dos hiatos – nos dados das informantes AP e MJ. A média geral da duração relativa, que considera os dados de todas as três informantes, confirma essa razão que diferencia a duração dos encontros vocálicos.

O desvio padrão, por sua vez, aponta para uma menor variação dos valores de duração dos hiatos, comparativamente aos ditongos. Esse achado pode ser interpretado como indicador de uma maior estabilidade dos hiatos do que dos ditongos, ou seja, os hiatos necessitam ter a duração da transição da vogal para a aproximante mais precisa, mais pontual, ao contrário do que acontece com os ditongos.

Os dados foram também submetidos à análise estatística. Para tanto, rodou-se um teste-T de Student, mas o resultado obtido ( $T=18,511$ , para  $p<0,001$ ) mostrou-se não significativo. Uma das razões prováveis para esse resultado é a quantidade restrita de dados submetidos à análise.

É mister tecer uma observação específica sobre três pares: “reunião/reúne”; “arruinar/ruína” e “sábida/sabia”. Na seção 2 comentava-se sobre a diferença destes três pares com relação aos demais, tanto no que diz respeito ao fato de haver um número maior de sílabas nas palavras constituintes desses pares, em comparação com as demais palavras do *corpus*, quanto com relação à distância que os encontros vocálicos nessas palavras guardam do acento lexical. Nos dados da informante AP, três pares exibiram um comportamento distinto dos demais em quatro das cinco repetições do conjunto de dados. Tal distinção consistia numa maior duração relativa das transições dos hiatos do que dos ditongos. Entretanto, a diferença era pouca, não chegava ao dobro como se nota para os dados como um todo e como revela a Tabela 1. Por vezes era da ordem de 1 ou 2 pontos percentuais. Nos dados produzidos por JI e MJ, por outro lado, verificou-se o mesmo comportamento desses pares relativamente aos demais, isto é, também para eles a duração das transições de ditongos mostrou-se maior que a de hiatos. Por estas razões elencadas, optou-se por reunir todos os dados no mesmo cômputo. Esses achados parecem sugerir que o maior número de sílabas dessas palavras ou a maior distância dos encontros vocálicos quanto ao acento lexical não desempenham papel relevante para a diferenciação entre hiatos e ditongos quanto ao seu padrão duracional.

De qualquer forma, deve-se frisar que a descrição acústica aqui exposta resulta de um estudo inicial sobre ditongos no PB e que, por isso, o estudo aqui apresentado carece, ainda, de dados mais robustos: é preciso incluir mais informantes nas análises, além de introduzir informantes do sexo masculino. Estas medidas proporcionarão um universo maior de dados, o que provavelmente possibilitará uma análise estatística confiável dos dados.

Ainda assim, e considerando-se os resultados das análises, expostos nesta seção, cabe perguntar em que medida eles reforçam os argumentos deste estudo, segundo os quais a organização temporal desempenha papel importante na diferenciação entre os tipos de encontros vocálicos do português brasileiro e a incorporação do dado fonético pode trazer contribuições importantes às análises fonológicas. É sobre esses pontos que o estudo se debruça na sequência.

#### 4 Discussão dos resultados

Conforme mencionado na seção 3, a propósito da análise dos dados, além de a duração relativa dos hiatos ser bem menor que a dos ditongos, os hiatos parecem mais estáveis, dado que o desvio padrão da duração relativa para eles é menor do que para os ditongos. Essa menor variação duracional dos hiatos pode estar relacionada à

sensibilidade dos encontros vocálicos a fatores prosódicos. A questão é: como se daria essa relação?

Sabe-se, através de estudos como o de Chitoran e Hualde (2007), por exemplo, que os encontros que se convencionou denominar “ditongos” evoluíram, nas línguas românicas, a partir de hiatos<sup>6</sup> e esta evolução pode ter acontecido orientada, por exemplo, pela estrutura rítmica da língua em questão.

O fato histórico parece encontrar eco sincronicamente nas flutuações que observamos em dados do português brasileiro e que, com frequência, causam dúvida aos usuários da língua quanto à alocação de fronteiras silábicas: é frequente, por exemplo, a dúvida de alunos que têm de fazer transcrição fonética, sobretudo de dados de fala espontânea, quanto ao estatuto dos encontros vocálicos, se ditongos ou hiatos. Aparentemente, quanto maior a velocidade de fala do sujeitos, maior a tendência de um encontro vocálico se realizar como ditongo. Logo, o hiato precisa ser mais estável para ser interpretado como tal. Se a transição da vogal para a aproximante é o evento acústico característico desses encontros, sem o qual teríamos apenas uma vogal (conforme se comenta na seção 3 a propósito das Figuras 1 e 2), quanto menos sua duração varia mais o encontro se aproxima de um hiato. Cabe acrescentar que a duração da transição do hiato pode variar, mas essa variação tem limites: é preciso atingir um momento estável na realização da aproximante, quando a trajetória dos formantes já não varia tanto em função do tempo, a exemplo do que ilustra o dado da Figura 2.

Por outro lado, quanto mais a duração da transição varia, mais ela tenderá a um ditongo. A rigor, é possível encontrar dados nos quais a transição da vogal para a aproximante se estende até o início do som seguinte, como acontece com o dado ilustrado na Figura 1.

Em suma, um hiato aparentemente necessita ter a transição da vogal para a aproximante ladeada por dois momentos estáveis, daí a menor variação da duração da transição que os valores de desvio padrão sugerem.

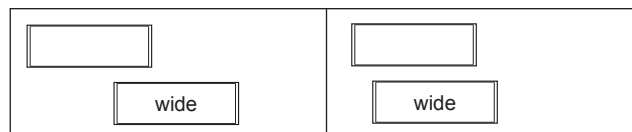
Do que se disse até aqui, deve ficar claro o papel da organização temporal sobre a distinção entre hiatos e ditongos no português brasileiro, afinal temos exatamente os mesmos eventos acústicos nos encontros vocálicos dos pares aqui investigados, alterando-se a duração da transição entre o que se denomina hiato e o que se denomina ditongo. Deve igualmente ter ficado óbvio que não se fez qualquer menção, neste estudo, a organizações silábicas distintas ou a afiliações silábicas deste ou daquele segmento.

A menção à estrutura silábica não se aplica neste estudo. Isso porque, ao assumirmos o papel fundamental da organização temporal sobre a cadeia da fala, assumimos inevitavelmente que a variável tempo deva ser intrínseca aos modelos de análise fonológica, tal como

já proposto por Fowler (1980), e tal como implementado por Browman e Goldstein (1992) em sua *Fonologia Gestual*.

A *Fonologia Gestual* adota um primitivo de análise de tempo intrínseco, o gesto articulatório<sup>7</sup>, e prevê que essas unidades organizam a cadeia da fala coordenando-se no tempo através de mecanismos de sobreposição. A noção de sobreposição gestual é crucial ao modelo, uma vez que permite prever consequências sobre a realização dos gestos, conforme se sobreponham mais ou menos. Assim, os gestos articulatórios são, a um só tempo, unidades simbólicas e discretas, isto é, estão presentes na representação da cadeia da fala e podem ser implementados nessa cadeia, no limite podem-se sobrepor totalmente, o que acarreta o aparente apagamento de um deles ou, no caso extremo oposto, podem-se organizar de modo que o *offset* de um coincida temporalmente com o onset de outro, de onde resulta um pequeno efeito de coarticulação entre ambos.

Considerando os resultados deste estudo à luz da *Fonologia Gestual*, pode-se prever que a maior ou menor sobreposição dos gestos articulatórios engendre a distinção entre hiatos e ditongos. Assim, considerando a Figura 3, temos ali uma representação preliminar para a distinção entre hiatos e ditongos. Admitindo-se que os gestos sejam representados por caixas<sup>8</sup>, e que a extensão horizontal delas traduza o tempo de ativação dos gestos, temos na pauta gestual à esquerda uma organização possível para encontros vocálicos do tipo ditongo e, na pauta gestual à direita, uma organização possível para hiatos.



**Figura 3** – Pauta gestual para hiatos (esquerda) e ditongos (direita)

<sup>6</sup> Em português, historicamente houve aproximantes que se introduziram entre duas vogais para separar um hiato, o que gerou uma sílaba (C)VC travada por aproximante e seguida de uma sílaba V, como em “ceia”. (Agradeço ao prof. dr. Márcio Renato Guimarães (DLLCV/UFPR) este dado, em comunicação pessoal.)

<sup>7</sup> Não é objetivo deste estudo expor as bases da *Fonologia Gestual* ou tampouco explicar como se definem os gestos articulatórios. Estes aspectos do modelo podem ser encontrados, por exemplo, em Silva (2002).

<sup>8</sup> A representação dos gestos articulatórios em caixas é proposta por Browman e Goldstein (1992). A dimensão horizontal da caixa representa, como mencionado, o tempo durante o qual o gesto permanece ativado, enquanto a dimensão vertical representa a magnitude do gesto. A organização coordenada entre dois ou mais gestos recebe o nome de “pauta gestual” no modelo. Frise-se, a exemplo do que fazia a nota 7, que o objetivo deste estudo não é apresentar as bases deste modelo, que o leitor pode encontrar também em Silva, Oliveira e Pacheco (2001).

Nas duas pautas gestuais da Figura 3, o gesto na porção superior da figura corresponde ao gesto de dorso da língua, responsável pela representação e produção de uma vogal. O gesto na porção inferior, com o descritor gestual “wide”, é responsável pela representação e implementação da aproximante. O descritor gestual específica como deve ser o gesto e é através dos descritores gestuais previstos pelo modelo que se podem diferenciar modos de articulação. O descritor *wide*, empregado aqui, prevê um grau de constrição largo para o trato, daí se afirmar que ele representa e implementa a aproximante.

Como mencionado anteriormente, estas são pautas gestuais preliminares, portanto não têm o intuito de representar diferentes vogais ou mesmo as diferentes aproximantes que podem figurar nos encontros vocálicos do português. O objetivo das pautas gestuais da Figura 3 é mesmo ilustrar as diferentes maneiras como os mesmos gestos podem se organizar no tempo. Assim, na pauta à esquerda nota-se pouca sobreposição entre gestual, já que o gesto que representa a aproximante se inicia mais próximo do final do tempo de ativação do gesto que representa a vogal. Por outro lado, na pauta à direita da mesma figura vê-se uma grande sobreposição entre os gestos, de modo que o gesto que representa a aproximante se inicia logo depois do início do tempo de ativação do gesto da vogal.

A consequência dessa diferença entre as pautas é que, na pauta à esquerda, prevê-se um período de sobreposição mais curto, que corresponderia à transição mais breve dos hiatos. O período de sobreposição gestual mais longo, na pauta à direita da figura, representa e implementa uma transição mais longa dos ditongos. Note-se que, no caso de uma representação gestual, dinâmica, como a sugerida aqui, não são necessárias regras lexicais ou pós-lexicais para gerar ditongos ou hiatos; a Fonologia Articulatória é um modelo implementacional, o que significa dizer que a “tradução” do nível fonológico no nível fonético se faz de maneira direta<sup>9</sup>, portanto mais parcimoniosa.

Cabe comentar, finalmente, que o fato de os gestos articulatórios constituintes de ditongos apresentarem maior sobreposição poderia ser explicado admitindo-se, como Browman e Goldstein (1995), que gestos consonantais de *onset* se coordenam sincronicamente com o gesto vocálico seguinte, mas gestos consonantais de coda se coordenam com o gesto vocálico precedente. Através dessa previsão, poder-se-ia, em última instância, estabelecer as fronteiras silábicas como resultado da coordenação temporal entre os gestos. E esta coordenação pode, em última instância, sofrer influência de fatos como o ritmo, o que explicaria a evolução histórica dos hiatos

para os ditongos e a flutuação entre eles em vários dados da língua portuguesa, que acabam resultando em dúvidas sobre as fronteiras silábicas, como exemplificado por Câmara Jr. (1969). Não se busca, aqui, uma discussão detida acerca da previsão da Fonologia Gestual de que a organização entre gestos pode resultar em estruturas silábicas distintas. Pretende-se tão somente chamar a atenção para esse ponto, que pode ser matéria para outros estudos.

## 5 Considerações finais

Deve-se deixar claro que o objetivo deste estudo não é uma discussão aprofundada sobre como é a representação gestual dos hiatos e ditongos. Apresentou-se, na seção 4, uma representação muito preliminar para essa diferença, porque tal representação servia aos propósitos deste estudo, quais sejam, chamar a atenção para contribuições que o dado fonético pode trazer para análises fonológicas e, até, em alguns casos, promover análises mais parcimoniosas.

Pretende-se, por outro lado, que fique claro que o avanço tecnológico acontecido já há algumas décadas proporcionou o surgimento de ferramentas de análise que podem contribuir muito para análises fonológicas. Se estas ferramentas estão disponíveis, como no caso da análise acústica, em função sobretudo do surgimento de programas de análise gratuitos, não há razão para não usá-las. Aliás, a utilização dessas ferramentas pode trazer um grande ganho às análises fonológicas e pode, no limite, desfazer o abismo entre Fonética e Fonologia, que o estruturalismo linguístico se encarregou de impor.

## Referências

- BISOL, Leda. O ditongo na perspectiva da fonologia atual. *D.E.L.T.A.*, São Paulo, v. 5, n. 2, p. 185-224, 1989.
- BISOL, Leda. A sílaba e seus constituintes. In: NEVES, Maria Helena Moura (Org.). *Gramática do português falado*. Campinas: Ed. da UNICAMP, 1999. v. 7.
- BOERSMA, Paul; WEENINK, David. *Praat: doing phonetics by computer*. Disponível em: <<http://www.fon.hum.uva.nl/praat/>>. Acessado em: 30 jul. 2013.
- BROWMAN, Catherine; GOLDSTEIN, Louis. Articulatory Phonology: an overview. In: *Phonetica*, v. 49, p. 155-180, 1992.
- BROWMAN, Catherine; GOLDSTEIN, Louis. Gestural syllable position effects in American English. In: BELL-BERTI, F.; RAPHAEL, L. J. (Ed.). *Producing Speech: Contemporary Issues*. Woodbury, NY: AIP Press, 1995.
- CÂMARA JR., Joaquim Mattoso. *Estrutura da língua portuguesa*. Petrópolis: Vozes, 1969.
- CHITORAN, I.; HUALDE, J. I. From hiatus to diphthong: the evolution of vowel sequences in Romance. In: *Phonology*, v. 24, p. 37-75, 2007.

<sup>9</sup> Para maiores detalhes acerca da natureza implementacional da Fonologia Gestual remete-se o leitor mais uma vez a Silva (2002).

FOWLER, Carol. Coarticulation and theories of extrinsic timing control. *Journal of Phonetics*, v. 8, p. 113-133, 1980.

KENT, Ray; READ, Charles. *The acoustic analysis of speech*. San Diego: The Singular Publishing, 1996.

HOUAISS, Antonio; VILLAR, Mauro Sales; FRANCO, Francisco Manoel de Mello. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2009.

OHALA, John. Methods in phonology; SOLÉ, Maria Josep; BEDDOR, Patricia; OHALA, Manjari. (Ed.). *Experimental approaches to phonology*. Oxford: Oxford University Press, 2007. p. 3-6.

SELKIRK, Elisabeth. *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge: The MIT Press, 1984.

SILVA, Adelaide Hercília Pescatori; PACHECO, Vera; OLIVEIRA, Leonardo. Por uma abordagem Dinâmica

dos processos Fônicos. *Revista de Letras*, Curitiba, v. 55, p. 96-113, 2001.

SILVA, Adelaide Hercília Pescatori. *As fronteiras entre Fonética e Fonologia e a alofonia dos róticos iniciais em PB: dados de dois informantes do sul do país*, 2012. 213 p. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2002.

TORNQUIST, Gabriela. *Ditongos no português e no espanhol – análise sincrônica e diacrônica*, 2013. 100 p. Dissertação (Mestrado em Linguística) – Universidade Católica de Pelotas, Pelotas, 2013.

Recebido: 30 de agosto de 2013

Aprovado: 15 de novembro de 2013

Contato: adelaidehpsilva@gmail.com