

# Análise acústica de características temporais de consoantes no Português Brasileiro

## *Acoustic analysis of temporal characteristics of consonants from Brazilian Portuguese*

Cristina Ferraz Borges Murphy<sup>1</sup>, Luciana de Oliveira Pagan-Neves<sup>2</sup>, Eliane Schochat<sup>3</sup>, Haydée Fiszbein Wertzner<sup>4</sup>

### RESUMO

**Objetivo:** Analisar acusticamente as consoantes plosivas /p, b, t, d, k, g/ e fricativas /f, v, s, z, ʃ, ʒ/ do Português Brasileiro em relação às variáveis duração, frequência e variação de oitavas por segundo. **Métodos:** A produção de palavras contendo os sons-alvo foi analisada por meio de uma prova de imitação de palavras. Cada participante, com idades entre oito e dez anos repetiram uma série de palavras isoladas retiradas de uma lista foneticamente balanceada para o Português Brasileiro. Os parâmetros analisados foram: duração do som-alvo (fricativo ou plosivo) e valores das frequências inicial e final do F2 medidos em Hz. A partir destes valores foi calculada a variação de oitavas por segundo presente no F2 de cada consoante e, posteriormente, foi realizada a média para todos os sons. **Resultados:** Quanto à duração das consoantes, os valores variaram entre 31 ms e 170,5 ms. Em relação ao F2, os valores variaram entre 2617 Hz na frequência inicial da plosiva sonora /b/, até 4557 Hz, valor médio da frequência inicial observado na fricativa sonora /z/. Em relação à variação de oitavas por segundo presentes no F2, a maior variação foi encontrada para a plosiva surda /p/ (16,23 oitavas por segundo) e a menor para a fricativa surda /ʃ/ (1,624 oitavas por segundo). Considerando a média para todos os sons-alvo analisados, o valor encontrado foi de aproximadamente 6,8 oitavas por segundo. **Conclusão:** Os valores observados serão úteis para a elaboração de testes e de treinamentos auditivos contendo estímulos auditivos não-verbais semelhantes às características temporais presentes no Português Brasileiro.

**Descritores:** Percepção auditiva; Linguagem; Acústica; Fala

### INTRODUÇÃO

Pesquisas ressaltam a importância da utilização de estímulos não-verbais análogos às características acústicas da fala em teste de processamento auditivo. Segundo os pesquisadores, desta forma é possível investigar a existência de um déficit envolvendo o processamento temporal de um estímulo que

representa a fala, sem a influência dos aspectos linguísticos<sup>(1-3)</sup>.

Um exemplo da utilização deste tipo de estímulo foi publicado por Rosen e Manganari<sup>(3)</sup>. Com o intuito de analisar a relação entre alterações perceptivas e o aprendizado da leitura, pesquisadores aplicaram testes auditivos verbais e não-verbais, em crianças com e sem dislexia. Os testes não verbais apresentavam características análogas às sílabas do teste verbal (/ba/ e /da/). Para isso, o segundo formante contido nas sílabas e considerado crucial para a discriminação dos pares, foi isolado e manipulado para que pudesse apresentar a mesma duração e magnitude do estímulo verbal<sup>(3)</sup>.

Além dos testes, treinamentos auditivos computadorizados também são desenvolvidos com a utilização de estímulos não-verbais. Este tipo de estímulo pode ser facilmente manipulado para o treinamento de características perceptivas específicas bem como habilidades auditivas relacionadas aos processos que envolvem a linguagem<sup>(4)</sup>.

Geralmente, os estímulos não-verbais utilizados em treinamentos são formados por uma varredura de frequências análogas à transição do segundo formante da consoante<sup>(4)</sup>.

O programa americano *Fast ForWord* é um exemplo deste tipo de treinamento<sup>(5)</sup>. Em um dos jogos que compõem o programa, os estímulos são compostos por uma varredura de fre-

Trabalho realizado no Departamento de Fisioterapia, Fonoaudiologia e Terapia Ocupacional da Faculdade de Medicina da Universidade de São Paulo – USP – São Paulo (SP), Brasil.

(1) Pós-doutoranda em Ciências da Reabilitação / Área Comunicação Humana da Faculdade de Medicina da Universidade de São Paulo – USP – São Paulo (SP), Brasil.

(2) Doutora, Pesquisadora colaboradora do Laboratório de Investigação Fonoaudiológica em Fonologia do Departamento de Fisioterapia, Fonoaudiologia e Terapia Ocupacional da Universidade de São Paulo – USP – São Paulo (SP), Brasil.

(3) Livre-docente do Curso de Fonoaudiologia da Faculdade de Medicina da Universidade de São Paulo – USP – São Paulo (SP), Brasil.

(4) Livre-docente do Curso de Fonoaudiologia da Faculdade de Medicina da Universidade de São Paulo – USP – São Paulo (SP), Brasil.

**Endereço para correspondência:** Cristina Ferraz Borges Murphy, Av. Padre Antonio José dos Santos, 313/134, São Paulo (SP), Brasil, CEP: 04563-010. E-mail: crist78@yahoo.com

**Recebido em:** 4/12/2007; **Aceito em:** 18/3/2009

quências que varia 16 oitavas por segundo. Este valor, segundo os criadores do programa, representa a velocidade da transição dos formantes para as consoantes do inglês e, geralmente, as crianças com déficit de linguagem têm dificuldade em sua discriminação. Outro exemplo nesta linha de pesquisa, seria um programa alemão<sup>(6)</sup>. Neste programa, os estímulos apresentam uma varredura de frequências que varia quatro oitavas por segundo, valor correspondente às características da língua alemã.

Conforme as pesquisas citadas, para que os testes temporais bem como os treinamentos desenvolvidos a partir deles utilizem os parâmetros temporais apropriados e análogos às características da fala, a análise das características temporais da língua em questão é fundamental<sup>(7)</sup>.

No Brasil, não há estudos científicos que investigaram parâmetros temporais como a variação de oitavas por segundo contida em cada uma das consoantes analisadas.

Assim, o objetivo desta pesquisa foi analisar acusticamente as consoantes plosivas /p, b, t, d, k, g/ e fricativas /f, v, s, z, ʃ, ʒ/ do Português Brasileiro a partir das variáveis duração, frequência e variação de oitavas por segundo contidas nesses sons.

## MÉTODOS

O presente estudo foi aprovado pela Comissão de Ética para Análise de Projetos de Pesquisa – CAPPesq da Faculdade de Medicina da Universidade de São Paulo, em 9/08/2006, sob o Protocolo de Pesquisa nº 551/06.

### Sujeitos

A coleta de dados foi realizada no período de agosto a dezembro de 2006. Inicialmente, os critérios de inclusão adotados foram: idade entre sete e 12 anos e 11 meses; ausência de queixas envolvendo linguagem oral e audição. Esta faixa etária foi escolhida baseada em estudos similares realizados anteriormente<sup>(6)</sup>.

De acordo com estes critérios, dez participantes foram recrutados em escolas próximas à região na qual os dados foram coletados. Todos os responsáveis foram solicitados a assinar o Termo de Consentimento Livre e Esclarecido comprovando assim, o interesse e concordância pela participação na pesquisa e divulgação em meios científicos.

Para a investigação de possíveis alterações envolvendo aspectos relacionados à linguagem e à audição, os dez participantes foram submetidos, inicialmente, a uma série de procedimentos: história clínica, avaliação de linguagem infantil (ABFW), avaliação do sistema sensorio motor oral e avaliação da consciência fonológica. Além disso, também foram submetidos à avaliação audiológica básica (audiometria, logoaudiometria e imitanciométrica). Por meio desta bateria de testes foi possível investigar e excluir participantes que apresentassem alterações nos testes aplicados. Dos dez participantes avaliados, dois apresentaram alterações na avaliação audiológica básica e no teste de linguagem infantil e, portanto, foram excluídos e encaminhados para tratamento.

Assim, o grupo estudo foi formado por oito crianças, sendo quatro do sexo masculino e quatro do feminino, com idade entre oito e dez anos e 11 meses, apresentando audição dentro

dos padrões da normalidade e resultados dentro do esperado para a idade, em relação às outras provas aplicadas.

### Procedimentos

Selecionado o grupo de estudo definitivo, foi realizada a coleta dos dados. Foi explicado à criança que ela iria ouvir algumas palavras e deveria repetir cada uma delas da forma como ouviu. Assim, a primeira autora deste trabalho leu uma lista de palavras isoladas adaptadas de uma lista foneticamente balanceada para o Português Brasileiro<sup>(8)</sup>, que foram repetidas uma a uma pelo participante. Ao todo, foram contabilizadas dez palavras apresentando um dos sons alvo (/p, b, t, d, k, g, f, v, s, z, ʃ, ʒ/) em *onset* inicial. Este procedimento foi realizado individualmente, em uma sala vedada acusticamente, com um microfone posicionado a 5 cm da boca e conectado a um computador.

O *software* utilizado para a gravação e posterior análise acústica foi o CSL (*Computerized Speech Laboratory*, 4300-B) fabricado pela Kay Elemetrics®. Os parâmetros acústicos analisados foram: a duração do som-alvo (fricativa ou plosiva) medida a partir do primeiro ponto espectral da consoante até o último ponto anterior à coarticulação com a vogal seguinte e o valor inicial e final do F2 medidos individualmente no primeiro e último pontos espectrais da segunda frequência formante respectivamente. As vogais subsequentes à consoante não foram analisadas, assim como nos estudos realizados anteriormente e que serviram de orientação teórica para este<sup>(2-3,5)</sup>. A escolha específica da análise do F2 foi baseada em autores que referem ser este formante o principal responsável pela inteligibilidade de fala<sup>(3,9)</sup>.

#### *Cálculo do número de oitavas por segundo presente no F2 de cada consoante*

Para este cálculo, primeiramente foram considerados dois conceitos: a definição de uma oitava (quando a razão entre duas frequências é igual a dois) e o fato de as frequências do som variarem em uma escala logarítmica.

A seguir, são descritos todos os passos necessários para o cálculo:

- 1- Cálculo do número de oitavas, presente no F2 de cada consoante, para cada participante: para o cálculo do número de oitavas contidas entre a frequência final (Freq.Final) e inicial (Freq.Inicial) de cada consoante, para cada um dos oito participantes, utilizou-se a função logarítmica a seguir:  $\log_2 \frac{\text{FreqFinal}}{\text{FreqInicial}} = X$ , sendo X=variação de oitavas. Assim, obteve-se o número de oitavas, presente em cada consoante, considerando cada participante da pesquisa.
- 2- Cálculo do número de oitavas por segundo, de cada consoante, para cada participante: para transformar a variação de oitavas obtida em cada consoante (passo 2) na unidade oitavas por segundo, foi considerada a duração de cada consoante para cada participante da pesquisa. Assim, a partir de uma regra de três simples, calculou-se qual seria a variação de oitavas, presente em cada consoante, se cada uma apresentasse a duração de 1000 ms.

**Tabela 1.** Média do grupo para os valores de frequência inicial, frequência final e oitavas por segundo

Fonemas	Média de duração (ms)	Média de frequência inicial (Hz)	Média de frequência final (Hz)	Média das oitavas por segundo
p	31	3533	3634	16,23
b	59	2617	3114	6,82
t	36,12	4313	3675	13,50
d	67,87	4099	3333	4,55
k	55,12	3694	2969	7,31
g	63,75	3117	2900	3,96
f	115,75	3377	2793	12,86
v	113,5	3300	3221	2,88
s	149,87	3922	3409	4,41
z	123,37	4557	2648	5,33
ʃ	170,5	3866	3285	1,62
ʒ	96,25	4460	3254	2,92

- 3- Cálculo da média de oitavas por segundo do grupo para cada consoante: calculou-se a média aritmética do grupo, para cada consoante, a partir da variação de oitavas por segundo obtida para cada participante (passo 2).
- 4- Cálculo da média de oitavas por segundo para todas as consoantes: calculou-se a média aritmética do grupo em relação ao número de oitavas por segundo, considerando todas as consoantes analisadas.

## RESULTADOS

Na Tabela 1, é possível visualizar as médias do grupo considerando as variáveis duração, frequências inicial e final do F2 e número de oitavas por segundo, para cada consoante analisada.

Em relação à variável duração, a média dos valores variou entre 31 ms e 170,5 ms sendo que os menores valores encontrados foram para as plosivas surdas /p/ (31 ms) e /t/ (36,12 ms) e os maiores para as fricativas surdas /ʃ/ (170,5 ms) e /s/ (149,87 ms).

Em relação à variável frequência inicial e final, os valores médios variaram entre 2617 Hz na frequência inicial da plosiva sonora /b/ e 4557 Hz, valor médio da frequência inicial observado na fricativa sonora /z/.

Quanto à variação de oitavas por segundo, a maior variação foi encontrada para a plosiva surda /p/ (16,23 oitavas por segundo) e a menor para a fricativa surda /ʃ/ (1,624 oitavas por segundo). A média, considerando todas as consoantes analisadas, foi de 6,87 oitavas por segundo.

## DISCUSSÃO

O presente estudo teve como principal objetivo analisar acusticamente as consoantes plosivas e fricativas do Português Brasileiro considerando as variáveis: duração, frequência e variação de oitavas por segundo.

Por meio da Tabela 1 foi possível analisar os valores médios de duração considerando o grupo estudado. Comparando os valores dos pares mínimos plosivos, por exemplo, nota-se que

todos os sons surdos apresentaram valores de duração menores do que os sonoros (/p/=31 ms; /b/=59 ms; /t/=36,12 ms; /d/=67,87 ms; /k/=55,12 ms; /g/=63,75 ms). Estes resultados coincidem com outro estudo recente que também analisou os valores médios da duração das consoantes plosivas e fricativas do português brasileiro<sup>(10)</sup>. Este outro estudo mostrou que, ao contrário do que ocorre com o inglês norte americano<sup>(11)</sup>, os valores do VOT das plosivas sonoras são maiores que os apresentados pelas plosivas surdas. Os valores do VOT para oclusivas surdas estão entre +25 ms e +100 ms, enquanto para as oclusivas sonoras estes valores estão entre -20 ms e +20 ms.

Por outro lado, considerando-se os pares mínimos fricativos, o resultado observado foi contrário, ou seja, todos os sons surdos apresentaram valores médios de duração maiores do que os sons sonoros (/f/=115,75 ms; /v/=113,5 ms; /ʃ/=170,5 ms; /ʒ/=96,25 ms; /s/=149,87 ms; /z/=123,37 ms). Estes resultados corroboram os achados de outro autor<sup>(12)</sup> cujo estudo mostra que as fricativas surdas, em posição inicial (*onset*) de sílaba, são quase duas vezes mais longas do que as sonoras na mesma posição. Este estudo conclui ainda que a duração da fricção contribui para a discriminação do traço de sonoridade das fricativas. Os mesmos achados também foram observados por outros autores<sup>(13)</sup>.

As diferenças encontradas nos valores de duração das consoantes surdas e sonoras são pistas auditivas importantes para a discriminação destes sons<sup>(14)</sup>. Assim, a partir destes valores, testes e treinos auditivos poderão avaliar e treinar a discriminação temporal por meio de estímulos não verbais que apresentem durações com valores semelhantes. Ainda não existem testes e treinamentos auditivos para o Português Brasileiro considerando o parâmetro de duração das consoantes. Em um dos testes americanos que avaliam o padrão de duração<sup>(15)</sup>, os pares de estímulos a serem discriminados apresentam duração entre 250 e 500 ms. Já em relação aos treinamentos, um dos jogos presente no programa americano *Fast ForWord* apresenta estímulos não verbais que variam de 60 a 20 ms, em relação à variável duração<sup>(5)</sup>.

O segundo parâmetro estudado foi a variação de frequências contidas no F2 (Tabela 1). Conforme explicado

anteriormente, a escolha da análise deste formante foi baseada em autores que referem que o F2 é o principal responsável pela inteligibilidade de fala<sup>(3,9)</sup>. Os valores encontrados na pesquisa atual, relacionados às frequências de cada consoante, foram utilizados para a investigação do próximo parâmetro, a variação de oitavas por segundo, que será discutida a seguir.

A variação de oitavas por segundo presente em cada uma das consoantes foi calculada de acordo com a descrição apresentada no método. Conforme demonstrado na Tabela 1, a plosiva surda /p/ apresentou a maior variação de oitavas por segundo (16,23 oitavas por segundo) enquanto a fricativa surda /ʃ/ apresentou a menor variação (1,62 oitavas por segundo). Considerando o Português Brasileiro, não foram encontrados outros estudos que investigassem este parâmetro.

Posteriormente, foi analisada a média para todas as consoantes analisadas. O valor encontrado foi de 6,87 oitavas por segundo. A importância desta variável diz respeito à possível criação, a partir da medida encontrada, de estímulos não verbais compostos por uma varredura de frequência, que variem no tempo de forma semelhante aos sons das consoantes do Português Brasileiro.

Esta análise já foi realizada para falantes do inglês norte americano, no qual foi encontrada uma variação de 16 oitavas por segundo<sup>(5)</sup> e, em falantes do alemão, com valores de quatro oitavas por segundo<sup>(6)</sup>.

A diferença entre os valores encontrados pode estar associada à maneira como falantes de diferentes línguas processam auditivamente estímulos não verbais. Talvez as características acústicas próprias de cada língua influenciem no processamento temporal deste tipo de estímulo. Um estudo<sup>(16)</sup> feito com crianças inglesas mostra que estas são capazes de utilizar a pista da transição do F1 aos quatro anos de idade. Já as

francesas só a utilizam por volta dos nove anos. Estes achados permitiram aos autores concluir que o ambiente linguístico interfere na habilidade de discriminação e de nomeação de padrões acústicos.

Portanto, embora sejam utilizados estímulos não verbais em pesquisas como as citadas, o modo como esse estímulo é processado por falantes de diferentes línguas pode variar<sup>(17)</sup>. Além disso, esta diferença entre as línguas comprova a importância de considerarmos os parâmetros acústicos específicos de cada língua na elaboração de testes e treinamentos temporais auditivos. Pesquisas futuras são necessárias para que os valores encontrados possam ser revalidados a partir de uma amostra maior de participantes.

## CONCLUSÃO

O estudo mostra que, a duração das consoantes plosivas /p, b, t, d, k, g/ e fricativas /f, v, s, z, ʃ, ʒ/ do Português Brasileiro variou entre 31 ms e 170,5 ms. Os valores do F2 apresentaram uma variação de 2697 Hz, considerando-se a média da frequência inicial da fricativa surda /f/ e 4916 Hz para a fricativa surda /s/. A média de variação das oitavas por segundo foi de aproximadamente 6,8. Este trabalho salienta a necessidade de estudos mais amplos na área para que os resultados encontrados possam auxiliar na elaboração de testes e treinamentos auditivos apresentando estímulos não verbais análogos às características acústicas da língua em questão.

## AGRADECIMENTOS

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio concedido.

## ABSTRACT

**Purpose:** To acoustically analyze the stop consonants /p, b, t, d, k, g/ and the fricatives /f, v, s, z, ʃ, ʒ/ from Brazilian Portuguese, regarding the variables duration, frequency and octave variation per second. **Methods:** The production of words with the target phonemes was analyzed by using an imitation task. Each participant, with age between eight and ten years, repeated a series of isolated words selected from a phonetically balanced list for Brazilian Portuguese. Words were acoustically analyzed considering duration of the target sound (stop or fricative consonant), and initial and final frequency values of F2, measured in Hz. Based on these values, the variation of octaves per second present in each consonant's F2 was calculated, as well as the average value considering all consonants analyzed. **Results:** Consonants' duration varied from 31 ms to 170.5 ms. Values of F2 varied from 2617 Hz, for the initial frequency of the voiced stop /b/, to 4557 Hz, average value for the initial frequency of the voiced fricative /z/. Regarding the variation of octaves per second present in F2, the higher variation was found for the unvoiced stop /p/ (16.23 octaves per second), and the lower variation, for the unvoiced fricative /ʃ/ (1.624 octaves per second). The average value found for all target-sounds analyzed was of approximately 6.8 octaves per second. **Conclusion:** The values observed in the study will be useful to the elaboration of auditory tests and trainings including nonverbal stimuli analogous to the temporal characteristics of speech in Brazilian Portuguese.

**Keywords:** Auditory perception; Language; Acoustics; Speech

**REFERÊNCIAS**

1. Mody M, Studdert-Kennedy M, Brady S. Speech perception deficits in poor readers: auditory processing or phonological coding? *J Exp Child Psychol.* 1997;64(2):199-231.
2. Temple E, Poldrack RA, Protopapas A, Nagarajan S, Salz T, Tallal P, et al. Disruption of the neural response to rapid acoustic stimuli in dyslexia: evidence from functional MRI. *Proc Natl Acad Sci U S A.* 2000;97(25):13907-12.
3. Rosen S, Manganari E. Is there a relationship between speech and nonspeech auditory processing in children with dyslexia? *J Speech Lang Hear Res.* 2001;44(4):720-36.
4. Thibodeau LM. Computer-based auditory training for (central) auditory processing disorders. In: Chermak GD, Musiek FE. *Handbook of (central) auditory processing disorder. Volume 2: Comprehensive intervention.* San Diego: Plural Publishing; 2006. v. 2. p.167-206.
5. Merzenich MM, Jenkins WM, Johnston P, Schreiner C, Miller SL, Tallal P. Temporal processing deficits of language-learning impaired children ameliorated by training. *Science.* 1996;271(5245):77-81. Comment in: *Science.* 1996;271(5245):27-8.
6. Strehlow U, Haffner J, Bischof J, Gratzka V, Parzer P, Resch F. Does successful training of temporal processing of sound and phoneme stimuli improve reading and spelling? *Eur Child Adolesc Psychiatry.* 2006;15(1):19-29.
7. Schochat E, Musiek FE. Maturation of outcomes of behavioral and electrophysiologic tests of central auditory function. *J Commun Disord.* 2006;39(1):78-92.
8. Zaidan E, Sameshima K. Desenvolvimento de listas de palavras foneticamente balanceadas para avaliação do processamento auditivo. *Pró-Fono.* 2003;15(3):307-16.
9. Borden GJ, Havies KS, Raphael LJ. *Speech science primer: physiology, acoustics, and perception of speech.* 3a ed. Baltimore: Williams & Wilkins;1994.
10. Gurgueira AL. Estudo acústico do “voice onset time” e da duração da vogal na distinção da sonoridade dos sons plosivos em crianças com transtorno fonológico [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo; 2006.
11. Kent RD, Read C. *The acoustic analysis of speech.* San Diego, Calif.: Singular Pub. Group; c1992.
12. Haupt C. As fricativas [s], [z], [Z] [X] do Português Brasileiro. *Estudos Linguísticos.* 2007;36(1):37-46.
13. Baum SR, Blumstein SE. Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English. *J Acoust Soc Am.* 1987;82(3):1073-7.
14. Wright BA, Buonomano DV, Mahncke HW, Merzenich MM. Learning and generalization of auditory temporal-interval discrimination in humans. *J Neurosci.* 1997;17(10):3956-63.
15. Musiek FE, Baran JA, Pinheiro ML. Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. *Audiology.* 1990;29(6):304-13.
16. Simon C, Fourcin AJ. Cross-language study of speech-pattern learning. *J Acoustic Soc Am.* 1978;63(3):925-35.
17. Chermak GD, Musiek FE. *Central auditory processing disorders: new perspectives.* San Diego: Singular Publisher Group; c1997.