

AS FRICATIVAS [s], [z], [ʃ] E [ʒ] DO PORTUGUÊS BRASILEIRO

Carine HAUPT*

Resumo: *No presente experimento, analisamos as características acústicas das fricativas sibilantes [s, z, ʃ, ʒ] do português brasileiro em início e final de sílaba em um corpus gravado por dois informantes (um que tinha como fricativa em coda em seu dialeto a variante alveolar, e outro a variante palato-alveolar) e em um corpus de fala espontânea, gravado por outros dois informantes com a mesma variação dialetal. Os parâmetros analisados foram os picos espectrais obtidos através do cálculo de 4 formantes, a fim de caracterizar alveolares e palato-alveolares, e a duração da fricção, com o intuito de diferenciar as fricativas surdas das sonoras. Além disso, apresentamos uma descrição fonética desses segmentos segundo diferentes autores. Concluímos que as alveolares têm seu pico mais proeminente em regiões de frequência mais altas que as palato-alveolares e que as surdas têm a fricção mais longa que as sonoras, diferença que, no entanto, diminui quando essas consoantes se encontram em final de sílaba.*

Palavras-chave: *Sibilantes; picos espectrais; formantes; vozeamento; duração da fricção.*

Introdução

Muitos estudos acústicos sobre fricativas já foram feitos, porém poucos tratando especificamente das do português do Brasil (doravante PB), fato que nos motivou a escolher esses segmentos para o presente estudo. Com base na literatura e estudos existentes em outras línguas, como o trabalho de Hughes e Halle (1956) e de Shadle (1995) no inglês, e de Jesus (1999) no português de Portugal, esperamos, com esse experimento, poder fazer algumas afirmações sobre as fricativas do PB, de modo a enriquecer a caracterização desses segmentos.

No presente estudo, temos como objetivo caracterizar acusticamente apenas as fricativas alveolares e palato-alveolares. Limitamos-nos a essas duas classes de fricativas, pois pretendemos verificar como elas se comportam tanto em posição de ataque quanto de coda, e as fricativas labiodentais não ocorrem nessa última posição em língua portuguesa.

Considerando que o ponto de articulação determina as ressonâncias das frequências da fonte de ruído, tomamos como um dos parâmetros para

* Mestre pelo programa de Pós-Graduação em Lingüística, UFSC. E-mail: ka_haupt@hotmail.com

analisar essas fricativas os picos espectrais, que serão definidos através do cálculo dos formantes. A duração é outro parâmetro que nos interessa, pois queremos também diferenciar as fricativas surdas e sonoras, além de verificar se esse parâmetro varia de acordo com a posição na sílaba. O interesse desse experimento, portanto, é averiguar como esses dois parâmetros caracterizam cada um desses segmentos e até que ponto podem distingui-los entre si.

1. Caracterização acústica das fricativas

As fricativas são sons produzidos por uma fonte de ruído resultante da turbulência de ar gerada pela constrição do trato vocal. As fricativas /s, z, ʃ, ʒ/ são conhecidas como sibilantes por terem como sua principal fonte a turbulência produzida quando o jato de ar criado pela constrição encontra um obstáculo: os dentes. Fricativas dentais e alveolares (incluindo as palato-alveolares) são sibilantes e as demais têm a turbulência produzida pela constrição em si (LADERFOGED & MADDIESON, 1993).

Quanto à intensidade, as fricativas podem ser consideradas um dos sons mais fracos do PB. Podemos observar isso na forma de onda das Fig. 1 e 2. Na Fig. 1, temos o oscilograma de uma fricativa surda e na Fig. 2, temos uma fricativa sonora. Observando essas duas figuras, percebemos que essas consoantes são bem menos intensas que suas vogais adjacentes. Comparando-as, percebemos também uma diferença entre elas: a surda constitui-se apenas de um ruído e a sonora apresenta também pulsos glotais.

As fricativas vozeadas são formadas por duas fontes: a fonte glótica, responsável pelo vozeamento, e a fonte de ruído resultante da constrição do trato, enquanto as surdas são formadas apenas por uma fonte de ruído, que é mais forte do que nas sonoras. Segundo Shadle (1995), a fonte de ruído das sonoras é mais fraca devido à necessidade de se manter uma queda de pressão transglotal a fim de manter o vozeamento.

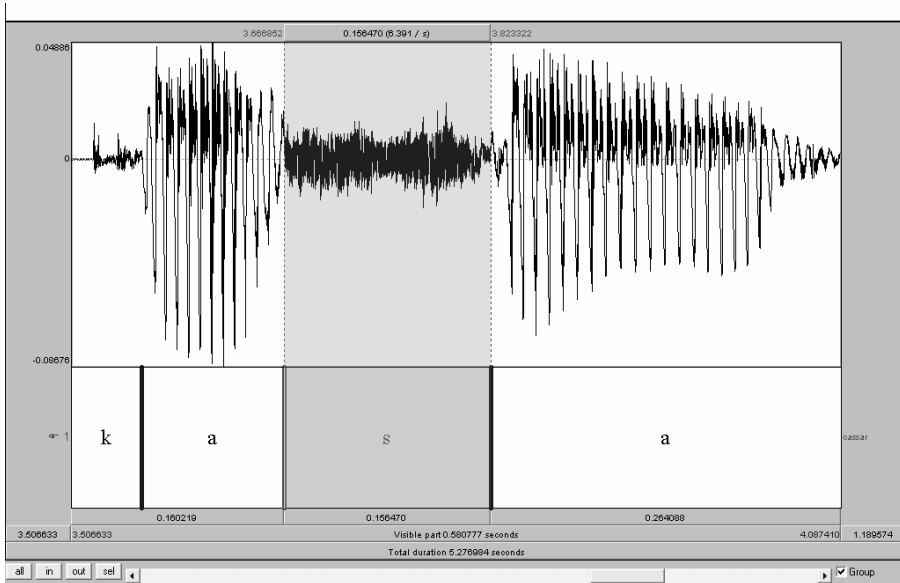


Fig. 1. Forma de onda da palavra [ka 'sar]

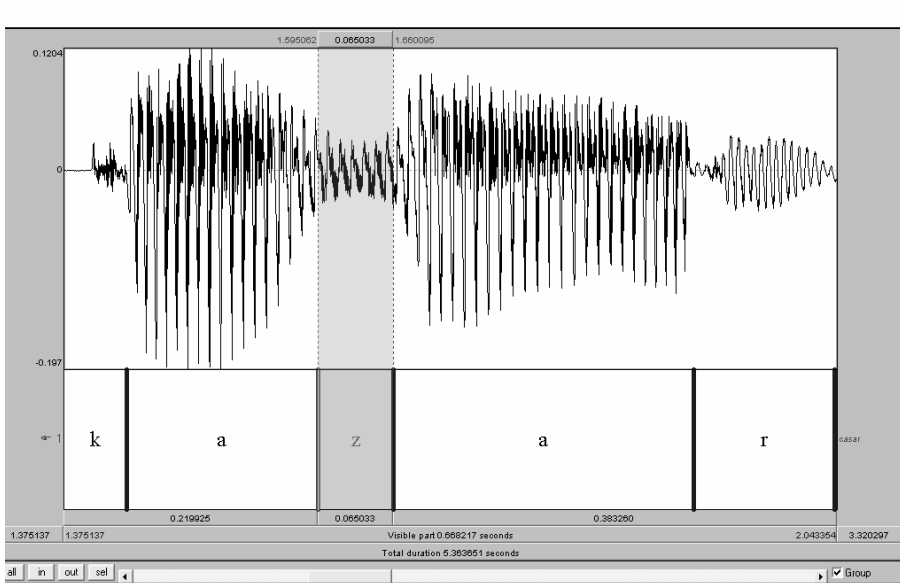


Fig. 2. Forma de onda da palavra [ka 'zar]

Além das diferenças da intensidade da fricção, Laderfoged e Maddieson (1993) nos apresentam diferenças de natureza articulatória entre o [s] e o [ʒ],

que irão influenciar no espectro da fonte de ruído, uma vez que este depende da forma do trato pelo qual a corrente de ar passa, além da velocidade do fluxo através da constrição:

- o [ʃ] tem o traço de arredondamento ou protusão labial;
- o [ʃ] é mais posterior e a área de constrição é mais extensa;
- a parte imediatamente posterior à constrição do [s] é levantada;
- no [ʃ] a frente da língua é erguida, com a parte central acima do nível das laterais da língua;
- na produção do [s] ocorre um abaixamento da parte central da língua em relação às laterais dela;
- para a produção do [s] se forma um canal mais estreito. O canal mais largo para o [ʃ] faz com que o ar tenha menos velocidade. Devido à diminuição da velocidade da corrente de ar, da extensão maior do trato vocal e do arredondamento, presume-se que [ʃ] tenha um *pitch* mais baixo que [s].

Segundo Russo e Behlau (1993, p. 46), as fricativas “/s, z/ são sons fortes, agudos, com uma faixa de frequência acima de 4500 Hz, chegando a 8000 Hz. As posteriores /ʃ, ʒ/ também apresentam uma faixa de frequência ampla, semelhante a /s, z/, porém mais grave, entre 2500 e 6000 Hz”. Observando os espectros abaixo (Fig. 3 e 4), podemos verificar que as fricativas [s, z] (linhas grossas) têm seu pico de amplitude em regiões de frequência mais alta do que [ʃ] e [ʒ] (linhas finas).

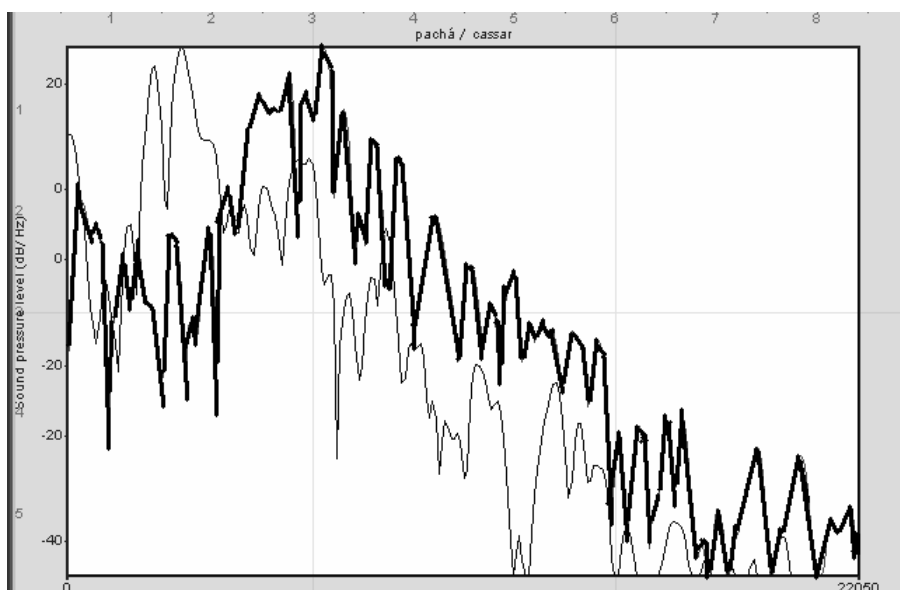


Fig. 3. Espectro das fricativas surdas [ʃ] (linha fina) e [s] (linha grossa)

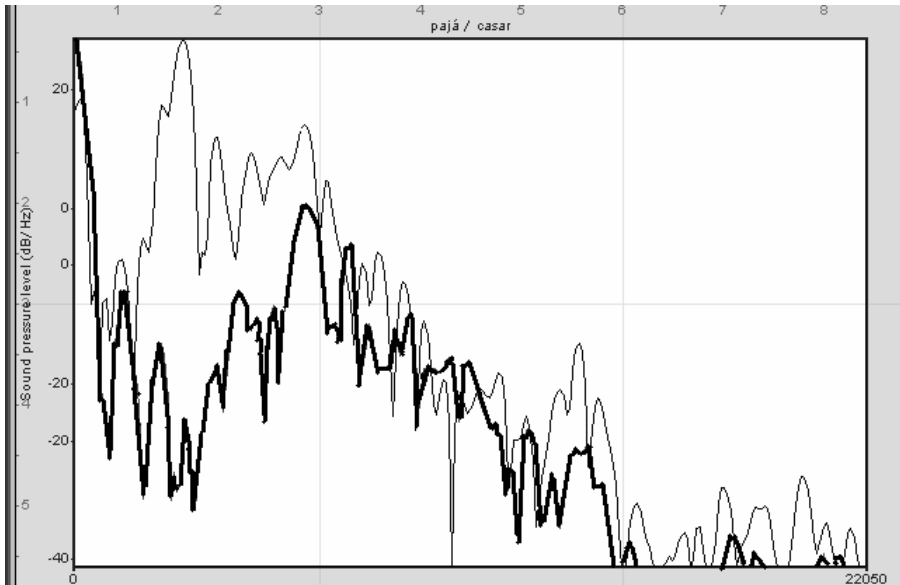


Fig. 4. Espectro das fricativas sonoras [ʒ] (linha fina) e [z] (linha grossa)

Outros autores, em seus trabalhos, confirmam essa diferença. Lacerda (1982), em seus estudos sobre as fricativas surdas do português europeu, fez vários testes perceptuais e concluiu que o /s/ é mais bem percebido quando o estímulo tem altos níveis de intensidade e picos espectrais na região de 5 kHz, e o /ʃ/ é normalmente associado com altos níveis de intensidade juntamente com picos espectrais na região dos 3 kHz. Jesus (1999), ao analisar as fricativas do português europeu em um *corpus* em que elas foram sustentadas por alguns segundos, obteve os seguintes resultados: o /s/ tem seu pico principal na região dos 5 kHz e picos secundários entre 10 e 15 kHz; o /z/ também tem seu pico principal na região dos 5 kHz, enquanto o /ʒ/ tem seu pico mais proeminente em torno de 2,5 kHz, e o /ʃ/ na faixa dos 2,7 kHz, com um pico secundário na região dos 11 kHz. A partir desses dados, parece claro que a localização da constricção determina as ressonâncias das frequências da fonte de ruído, não havendo diferenças em frequência entre surdas e sonoras.

O mesmo autor observou também que o contexto vocálico interfere na região de proeminência espectral de algumas fricativas. O pico do espectro do /s/ em início de palavra em contexto /ɔ/ é 1 kHz mais baixo em relação ao pico no espectro de um /s/ em contextos de /i/, /ɛ/ e /e/. O pico do /z/ em início de palavra é 1,5 kHz mais baixo em contextos vocálicos /ɔ/ e /o/, e 500 Hz mais baixo quando seguido de /u/, em relação ao pico do espectro de /z/ no contexto vocálico de /ɛ/ e /e/. O espectro do /ʒ/ em início de palavra seguido de /ɛ/ tem seu pico na faixa dos 4 kHz, com amplitude mais alta que o mesmo pico nos contextos de /ɔ/, /o/ e /u/ (JESUS, 1999).

Para finalizar, temos ainda mais um parâmetro a considerar: a duração da fricção. As fricativas posteriores /ʃ, ʒ/ são mais longas que as mediais /s, z/, e as surdas são mais longas que as sonoras (SANTOS, 1987), fato que iremos confirmar observando as durações apresentadas por nossos dados. A duração da fricção contribui na discriminação do traço de sonoridade das fricativas.

2. Metodologia

As análises foram feitas a partir de dois *corpora*: um controlado e outro de fala espontânea. Ambos foram gravados a uma taxa de amostragem de 44,1 kHz, taxa necessária para analisar as fricativas, já que elas são sons que apresentam formantes em altas frequências. O *corpus* controlado foi gravado no LINSE (Laboratório de Circuitos e Processamento de Sinais da UFSC), com o uso de um *software* de análise de fala, o Praat, o mesmo usado para fazer as análises. As gravações foram feitas com um microfone unidirecional, para evitar a captação de ruídos. O *corpus* de fala espontânea não foi gravado especialmente para esse estudo, ele já fora utilizado para a caracterização das vogais tônicas de falantes florianopolitanos e nos foi cedido para analisar as fricativas. Consideramos o material apropriado por ter sido gravado a uma alta taxa de amostragem e em estúdio com isolamento acústico.

O *corpus* controlado foi elaborado de modo a abranger as fricativas [s, z, ʃ, ʒ] em início e em final de sílaba, ou seja, em *onset* e em *coda*. Para tanto, foram criadas pseudopalavras que foram repetidas cinco vezes isoladamente, e três vezes dentro de uma frase-veículo por dois informantes. Foram também gravadas palavras reais, nas quais a fricativa se encontrava nos mesmos contextos.

Para a elaboração do *corpus*, todas as fricativas foram colocadas em contexto tônico e seguidas de uma oclusiva. Escolhemos a consoante oclusiva devido ao silêncio característico no início de sua produção, o que evita traços decorrentes da coarticulação. O ambiente tônico nos pareceu ser o mais adequado pelo fato de essa sílaba normalmente ser pronunciada com mais intensidade e mais plenamente. Além disso, não levamos em consideração o contexto vocálico, usando sempre a vogal [a], salvo nos casos em que não foi possível encontrar uma palavra com essa vogal que atendesse às exigências de tonicidade e contexto seguinte. Partindo desses critérios, o *corpus* gravado foi o apresentado na Tabela 1.

Tabela 1. *Corpus* para o experimento controlado

Onset silábico		Coda silábica	
Pseudopalavras	Palavras Reais	Pseudopalavras	Palavras Reais
[pa 'sa]	[ka 'sar]	['paspe]	['kaspe]
[pa 'za]	[ka 'zar]	['pazbe]	['lezbike]
[pa 'ʃa]	[ta 'ʃar]	['paʃpe]	['kaʃpe]
[pa 'ʒa]	[vja 'ʒar]	['paʒbe]	['leʒbike]
		['pastɐ]	['pastɐ]
		['pazde]	['dezdi]
		['paʃtɐ]	['paʃtɐ]
		['paʒde]	['deʒdi]
		['paske]	['kaske]
		['pazge]	['xazge]
		['paʃke]	['kaʃke]
		['paʒge]	['xaʒge]

Os informantes para este *corpus* foram dois homens: um que tinha como fricativa em final de sílaba em seu dialeto a variante alveolar, e outro a variante palato-alveolar, a fim de satisfazer a necessidade de ter as duas fricativas em posição de coda. Os informantes das entrevistas, que caracterizam o *corpus* de fala espontânea, também foram dois homens com as mesmas características dialetais em relação às variantes em coda silábica.

As análises foram feitas da mesma forma para ambos os *corpora*. Os parâmetros avaliados foram os picos espectrais e a duração. Os picos espectrais foram obtidos através do cálculo dos formantes. Esse cálculo foi feito através da média de cada um dos 4 formantes encontrados em uma faixa de 16 kHz e situados na parte mais estável da consoante. Através dos picos espectrais, é possível obter informações quanto ao ponto de articulação, enquanto a duração auxilia na diferenciação das fricativas surdas e sonoras.

3. Análise dos dados

Antes de apresentarmos os resultados da análise dos dados, convém fazermos algumas observações. Ao analisarmos a duração da fricção, encontramos dificuldades em precisar a extensão do [z] na palavra ['dezɔ], devido à consoante seguinte, que, no dialeto do informante 1, foi realizada como uma africada e de forma bastante rápida juntamente com a vogal seguinte, tornando-se impossível distinguir os três segmentos finais, dada a forte coarticulação entre eles em quase todas as amostras. A mesma dificuldade foi encontrada em relação às palavras seguidas de [g]. A fricativa [z] era quase imperceptível, misturando-se com a vogal anterior e com o vozeamento da

consoante seguinte que, embora seja uma oclusiva, na maior parte das amostras não teve o silêncio e a explosão bem definidos, de modo que não foi possível estabelecer a duração da fricção nesses casos. Quanto ao cálculo dos formantes, o [z] não apresentou 4 formantes em todas as amostras. Isso ocorreu na palavra [ˈlɛzɓikə] do corpus controlado e em vários casos na fala espontânea.

A média dos 4 formantes das fricativas resultou nas seguintes tabelas:

Tabela 2. Média dos formantes das fricativas [s, z] em *onset* e coda silábica

Informante 1	F1	F2	F3	F4
[s] em <i>onset</i>	5438	6701	8998	14190
[s] em coda	5377	6897	9201	12314
[z] em <i>onset</i>	5906	7223	9244	12906
[z] em coda	5656	7179	9315	12489
Informante 2	F1	F2	F3	F4
[s] em <i>onset</i>	4411	6967	9207	12492
[z] em <i>onset</i>	4697	7118	9416	12827
Média	5247	7015	9231	12870

Tabela 3. Média dos formantes das fricativas [ʃ, ʒ] em *onset* e coda silábica

Informante 1	F1	F2	F3	F4
[ʃ] em <i>onset</i>	3245	5497	8052	11191
[ʒ] em <i>onset</i>	3230	5602	8496	11879
Informante 2	F1	F2	F3	F4
[ʃ] em <i>onset</i>	2977	6143	8451	12333
[ʃ] em coda	3336	6295	8358	12263
[ʒ] em <i>onset</i>	2769	5994	8049	12454
[ʒ] em coda	3377	6305	8500	12477
Média	3156	5973	8318	12099

Observando estas tabelas, podemos concluir que a fricativa alveolar surda, tanto em posição de *onset* quanto em coda, tem seu primeiro formante na faixa entre 4,5 e 5,5 kHz para o informante 1 e a 5 kHz para o informante 2. Essa diferença entre os dois informantes mostrou-se regular em todas as amostras, sendo, porém, mais significativa apenas no primeiro formante. Hughes e Halle (1956) observaram em seu estudo sobre as propriedades espectrais das fricativas discrepâncias entre os espectros de diferentes falantes, no entanto, as diferenças entre cada classe de fricativa (labial, dental e palatal) eram consistentes, particularmente para sons falados por um único falante. Observando nossos dados, vemos que as diferenças entre as duas classes de

fricativas são constantes para cada informante. A diferença do F1 entre [s, z] e [ʃ, ʒ] gira em torno de 1 e 2 kHz para o informante 2 em todas as amostras. Já para o informante 1, essa diferença é um pouco maior, na faixa dos 2,5 kHz, chegando até a 3 kHz nas sonoras. Observamos também que há uma certa regularidade na distância entre um formante e outro para cada informante. A distância entre o F1 e o F2 das fricativas palato-alveolares, por exemplo, é maior para o informante 2 do que para o informante 1. Para este, a diferença fica em torno de 2 e 2,5 kHz, enquanto que para o informante 2 chega até a passar de 3 kHz. Quanto às alveolares, essa diferença gira em torno de 2 e 3 kHz para o informante 2 e fica em torno de 1 kHz para o informante 1, às vezes um pouco mais.

O primeiro formante corresponde ao pico mais proeminente. Esses resultados estão de acordo com os estudos brevemente apresentados no item 1. O F2 situa-se na região dos 7,5 kHz, o F3 na de 9,5 a 10 kHz, e o F4 na faixa dos 14 kHz. A fricativa alveolar sonora apresenta os formantes nas mesmas regiões, confirmando que a forma do trato vocal é decisiva na ressonância das freqüências.

A fricativa palato-alveolar é menos aguda, tem seu pico mais proeminente, isto é, o F1, na região dos 3 kHz, o F2 em regiões próximas a 6 kHz, o F3 na faixa dos 8 kHz e o F4 na região dos 11 kHz para o informante 1 e 12 kHz para o informante 2.

Analisando o *corpus* de fala espontânea, também encontramos a mesma diferença entre as alveolares e as palatais, ou seja, o [ʃ] e o [ʒ] têm o F1 entre 2,7 e 3,5 kHz, e o [s] e o [z], com uma variação maior, entre 4,5 e 6 kHz. Essa variação talvez se deva ao fato de não termos encontrado nesse *corpus* muitas amostras com a mesma vogal na sílaba tônica ([a]) e termos analisado palavras com outras vogais. Como alguns autores sugerem, o contexto vocálico pode influenciar nos espectros das fricativas, assunto que poderemos aprofundar em um próximo estudo. Mesmo com essas diferenças, podemos concluir que as fricativas alveolares são mais agudas do que as palato-alveolares, ou seja, “quanto mais próxima for a fonte de ruído dos lábios maior será a freqüência natural de ressonância do trato vocal” (RUSSO & BEHLAU, 1993, p.43). Podemos afirmar também que esse parâmetro é válido para distinguir essas fricativas entre si, pois, mesmo que haja diferenças entre falantes, elas são constantes entre as duas classes de fricativas analisadas e não ocorrem diferenças intrafalantes.

Para distinguir as fricativas vozeadas das desvozeadas nos valem da duração da fricção. Nossos dados confirmam que as surdas são mais longas que as sonoras. Em início de sílaba, a consoante surda tem aproximadamente o dobro da duração da sonora, como podemos verificar na Tabela 4. Em coda, embora a surda continue sendo mais longa que a sonora, a diferença entre elas diminui bastante (Tabela 5). Quanto aos dados de fala espontânea, percebemos o mesmo comportamento em relação à duração das fricativas surdas e sonoras em *onset* e em coda.

Tabela 4. Duração das fricativas em onset em ms¹

[pa 'sa]	[ka 'sar]	[pa 'za]	[ka 'zar]	[pa 'ʃa]	[ta 'ʃar]	[pa 'ʒa]
138	155	72	74	130	124	85
133	172	78	89	138	131	82
124	163	80	86	145	125	75
102	116	66	57	111	107	80
106	123	70	58	116	111	76
098	117	70	52	107	106	69
117	141	72	69	125	117	78
[vja 'ʒar]						
57						
71						
59						
67						
73						
71						
66						

Tabela 5. Duração das fricativas em coda em ms

['paspe]	['pazbe]	['kaʃpe]	['xaʒpe]
76	67	72	52
61	45	62	40
88	54	57	55
71	40	47	64
70	53	63	53
64	38	47	49
72	49	58	52

¹ As últimas linhas nas tabelas 4, 5 e 6 correspondem às médias.

Tabela 6. Duração de algumas amostras do *corpus* de fala espontânea dos dois informantes em ms

Informante B				Informante H			
Onset	Onset	Coda	Coda	Onset	Onset	Coda	Coda
[s, ʃ]	[z, ʒ]	[ʃ]	[ʒ]	[s, ʃ]	[z, ʒ]	[s]	[z]
96	42	105	52	106	67	80	68
123	57	70	63	132	66	66	68
104	75	54	56	123	60	93	68
108	36	53	64	92	61	102	68
95	45	55	52	102	57	58	68
105	51	67	57	111	62	80	68

Conclusão

Nesse experimento, analisamos dados de um *corpus* controlado e de um *corpus* de fala espontânea a partir de dois parâmetros: os picos espectrais e a duração da fricção. Escolhemos esses dois parâmetros a fim de diferenciar as fricativas alveolares das palato-alveolares, e as surdas das sonoras. A partir da análise dos dados, constatamos que há diferenças entre os picos espectrais dessas duas classes de fricativas. As alveolares têm seus formantes em regiões de frequências mais altas do que as palato-alveolares. O F1, por exemplo, situa-se na região dos 5,2 kHz para [s, z] e em torno de 3,1 kHz para [ʃ, ʒ]. Observamos também que esses picos são constantes tanto para as surdas quanto para as sonoras, em *onset* e coda silábicos. Podemos, então, concluir que esse parâmetro distingue alveolares das palato-alveolares, confirmando que a configuração do trato é decisiva para as ressonâncias das frequências.

Quanto à duração, observamos em nossos dados que as fricativas surdas são mais longas que as sonoras. Em *onset*, a surda é aproximadamente duas vezes mais longa que a sonora. Essa diferença, no entanto, diminui quando essas consoantes se encontram em coda, chegando a ter médias muito próximas, especialmente as palato-alveolares. A partir disso, podemos concluir que a duração é um parâmetro adequado para distinguir as sibilantes surdas das sonoras quando elas se encontram em *onset*, porém esse parâmetro não se mostrou tão claro quando essas mesmas consoantes se encontram em coda silábica. Para podermos fazer afirmações mais categóricas, parece-nos necessário fazer estudos mais aprofundados, com a utilização de outras técnicas de medição, não apenas a feita manualmente no Praat.

Lembramos que esse trabalho caracteriza-se como um trabalho piloto e deixa muitos pontos em aberto para estudos subseqüentes, como por exemplo, a influência do contexto vocálico e o uso de outros parâmetros para

distinguir as fricativas surdas e sonoras. Muito ainda pode ser estudado sobre as fricativas do português, mas esperamos ter de alguma forma colaborado para a descrição desses segmentos.

CARINE, H. THE SIBILANTS FRICATIVES [s, z, ʃ, ʒ] OF THE BRAZILIAN PORTUGUESE

Abstract: *In the present experiment, we analyze the acoustics characteristics of the sibilants fricatives [s, z, ʃ, ʒ] of the Brazilian Portuguese in onset and coda produced by two speakers (one that had as fricative in coda in its dialect the alveolar variant, and another one that had the palato-alveolar variant) and in a corpus of spontaneous speech, recorded by two others informers with the same dialectal variation. The analyzed parameters had been the spectral peaks and the duration of the friction. Moreover, we present a phonetic description of these segments according to different authors. We conclude that the alveolar fricatives have its more prominent peak in regions of higher frequency than the palato-alveolar ones and that the voiceless have the friction longer than the voiced ones.*

Key-words: *Sibilants; spectral peaks; duration of the friction; vozeamento; formants.*

Referências

HUGHES, George W. & HALLE, Morris. Spectral properties of Fricative Consonants. *The journal of the acoustical society of America*. v. 28, n. 2, p. 303-310, março, 1956

JESUS, Luís Miguel Teixeira. *Analysis of Portuguese Fricative Consonants*. 70f. Mini Thesis- Department of Electronics and Computer Science , University of Southampton, 1999.

LACERDA, Francisco P. Acoustic perceptual study of the Portuguese voiceless fricatives. *Journal of Phonetics*. London, v. 10, p. 11-22, 1982.

LADERFOGED, P & MADDIESON, I. *The sounds of the world's language*. Oxford, Blackwell, 1993, p. 137-181.

RUSSO, Iêda. & BEHLAU, Mara. *Percepção da fala: análise acústica do português brasileiro*. São Paulo: Editora Lovise Científica, 1993.

SANTOS, M. T. *Uma análise espectrográfica dos sons fricativos surdos e sonoros do português brasileiro*. Monografia de especialização. Escola Paulista

de Medicina, 1987.

SHADLE, Christine H. Modelling the noise source in voiced fricatives. *Proceedings of the 15th International congress on Acoustics (ICA 95)*. Trondheim, Norway, p. 145-148, 1995.