

PISTAS ACÚSTICAS E A PERCEPÇÃO DO ACENTO LEXICAL EM PORTUGUÊS BRASILEIRO

Mario A. S. FONTES
(Pontifícia Universidade Católica de São Paulo)
fontes@pucsp.br

RESUMO: O objetivo deste trabalho é avaliar o papel de dois parâmetros acústicos (duração e frequência de formantes) na percepção do acento lexical no Português Brasileiro. Técnicas de manipulação de valores de duração e procedimentos de síntese foram utilizados para gerar os estímulos submetidos à avaliação perceptiva. Os resultados reforçam a importância da duração como um correlato do acento lexical em Português Brasileiro bem como apontam a relevância do padrão formântico para sinalizar o padrão acentual em Português.

PALAVRAS CHAVE: acento lexical, correlatos acústicos, síntese de fala e percepção.

ABSTRACT: The objective of this work is evaluating the role of two acoustic parameters (duration and formant frequency) in the perception of the lexical stress in Brazilian Portuguese. Manipulation of duration values and synthesis procedures have been used. The results reinforce the importance of duration as an acoustic correlate of lexical stress in Brazilian Portuguese as well as point out the importance of the formant pattern in signaling stress pattern in Portuguese.

KEY WORDS: lexical stress, acoustic correlates, synthesis speech, perception

0. Introdução

Este trabalho tem como objetivo avaliar o papel do parâmetro da duração em relação ao da configuração formântica na percepção do acento lexical em Português Brasileiro (doravante PB).

Neste estudo conjugamos procedimentos de análise acústica para a medição dos parâmetros de duração em ms, de frequência fundamental (doravante f_0) em Hz, de intensidade em dB e de frequência de formantes (F_1 , F_2 e F_3) em Hz e teste de avaliação perceptiva das proeminências acentuais.

Foram também utilizados procedimentos de manipulação da duração dos segmentos vocálicos em uma palavra trissílaba e efetuados procedimentos de síntese para a geração dos estímulos a serem submetidos ao teste de percepção.

Os correlatos acústicos do acento lexical variam de acordo com a língua e é necessária a conjugação de procedimentos de análise acústica, síntese de fala e testes perceptivos para se investigar quais parâmetros têm maior relevância dada determinada língua.

1. Fundamentação Teórica

O uso de instrumentos e técnicas de análise acústica permite analisar o sinal acústico e a partir dessa análise é possível inferir aspectos sobre a produção dos sons e verificar correlações entre produção e percepção.

A análise acústica tem como fundamento a teoria acústica de produção da fala (FANT, 1960) que analisa os sons da fala como resultado da ação de transferência do trato vocal sobre as fontes geradoras de voz e ruído.

As formas que a língua assume para a produção dos diversos sons nas línguas naturais determinam tubos de diferentes configurações ao longo do trato vocal, o qual compreende o espaço entre a glote fechada e os lábios na cavidade oral ou as narinas na cavidade nasal. Esses tubos modulam o som laríngeo para a produção dos sons vozeados. Além da fonte de voz, podem haver também fontes de ruído que também são moduladas pelo trato vocal. Sons fricativos não vozeados, por exemplo, têm fonte de ruído enquanto fricativos vozeados possuem fontes de voz e ruído.

A modulação dos sons acontece graças ao fenômeno da ressonância. Nos sons vozeados, a onda direta e a refletida formam dentro do trato vocal ondas estacionárias para determinadas frequências do espectro da onda glotal e essas frequências são assim reforçadas.

As ressonâncias do trato são chamadas de formantes. Para a caracterização das vogais, os três primeiros formantes (F1, F2 e F3) são suficientes. As frequências desses formantes determinam a qualidades das vogais. As vogais podem ainda ser caracterizadas em termos de sua duração, frequência fundamental e intensidade.

A duração das vogais reduzidas é muito menor do que as plenas correspondentes. Em posição silábica pós-tônica no PB as vogais tendem a ter a duração reduzida e a amplitude diminuída. A

duração das vogais é influenciada pelas consoantes no contexto imediato, sobretudo da consoante que as seguem.

As vogais carregam o acento, dimensão prosódica que é descrita geralmente como envolvendo proeminência gerada por suas dimensões acústicas. As dimensões acústicas envolvidas na realização do acento, ou seja, a duração a frequência fundamental (doravante f_0) e a intensidade correspondem, do ponto de vista perceptivo, aos fenômenos de alongamento, *pitch* e *loudness*. A qualidade vocálica também é considerada entre as dimensões que sinalizam o acento (LAVER, 1994, HAYWARD, 2000).

Segundo HAYWARD (2000), uma sílaba acentuada é, de uma maneira geral, mais longa, com *pitch* mais elevado, com *loudness* aumentada, mais hiper-articulada e com pressão subglotal aumentada do que uma não acentuada.

O acento, um dos aspectos prosódicos da fala, do ponto de vista perceptivo, implica em graus de proeminência: os segmentos percebidos como mais proeminentes apresentam acento forte. Do ponto de vista acústico, os elementos com maior grau de força apresentam maior duração, intensidade, elevação de f_0 e alteração de qualidade vocálica que é denotada pelas alterações de padrão formântico. Dependendo da língua há predominância de alguns desses parâmetros na sinalização do acento (LAVER, 1994).

Na literatura sobre dados de produção em estudos de natureza fonético-acústica, a duração é apontada como o principal correlato do acento lexical (MASSINI, 1991; GAMA-ROSSI, 1998; BARBOSA, 2000, 2002) e do acento frasal (BARBOSA, 2000, 2002) em PB.

Na língua inglesa foram realizados estudos na década de 50 e 60 onde mostraram que o principal correlato acústico e perceptivo são f_0 e duração LAVER (1994).

No Português Europeu (PE) DELGADO-MARTINS (1986) mostra que os principais correlatos são a duração e a energia. Em relação à língua espanhola LLISTERRI et al. (2003) mostra que os principais correlatos acústicos são o f_0 e a duração e ALFANO (2005), com base nos trabalhos desenvolvidos por LLISTERRI et al (op. cit), ao pesquisar os correlatos acústicos do acento em língua italiana, concluiu que a duração é o parâmetro mais importante para as proparoxítonas e a intensidade mais relevante no caso das oxítonas.

2. Metodologia

A metodologia empregada neste trabalho é de caráter experimental, ou seja, implica no uso de instrumentos para a

manipulação e análise de dados e de critérios rigorosos para o estabelecimento dos procedimentos de elaboração de *corpora* de análise, de execução da coleta de dados e de escolha de parâmetros de análise.

São contemplados dois tipos de tarefas: tarefa de produção e tarefas de percepção. Os procedimentos utilizados em ambas as tarefas são delineadas a seguir.

3. Tarefa de produção: *corpus*, gravação e análise fonético-acústica

O *corpus*, utilizado neste trabalho, compreende 03 frases veículo (Digo _____ baixinho) nas quais foram inseridos os vocábulos trissílabos "VALido, vaLIdo e valiDO", respectivamente proparoxítono, paroxítono e oxítono. As frases foram digitadas em fichas e apresentadas aleatoriamente na sessão de gravação. As frases foram repetidas 10 vezes por um sujeito do sexo masculino de origem paulistana em taxa de locução normal, totalizando 30 itens para análise.

Os estímulos foram digitalizados na frequência de 22 kHz e analisadas no *software* Praat de código fonte aberto, desenvolvido na Universidade de Amsterdam, com distribuição livre, de autoria de Paul Boersna e David Weenink, rodando em plataforma Windows em *Hardware* DELL.

A gravação foi realizada em ambiente com tratamento acústico, controle da intensidade e da taxa de locução no Laboratório de Rádio da Faculdade de Filosofia, Comunicação, Letras e Artes da Pontifícia Universidade Católica de São Paulo com gravador um gravador digital da SONY, modelo TCDD8.

Foram medidos os valores de f_0 em HZ, intensidade em dB, duração em ms e frequência de formantes (F1, F2 e F3) de todos os segmentos vocálicos dos três vocábulos.

Os valores dos formantes foram medidos na parte estacionária dos núcleos vocálicos, por meio da conjugação das técnicas de Linear Predictive Code (LPC) e Fast Fourier Analysis (FFT). O LPC é uma técnica de codificação para reduzir o número de amostras que representam o sinal e baseia-se na possibilidade de prever uma amostra do sinal a partir de uma combinação linear de um número fixo e pré-definido de amostras anteriores. O FFT é um algoritmo que rapidamente converte tempo em frequência e vice-versa.

Na Figura 1, a seguir, apresentamos o gráfico referente aos espectros de FFT (em preto) e de LPC (em vermelho) obtidos em

porção estacionária da vogal [a] em válido. Podem ser visualizados os três primeiros formantes (F1, F2 e F3).

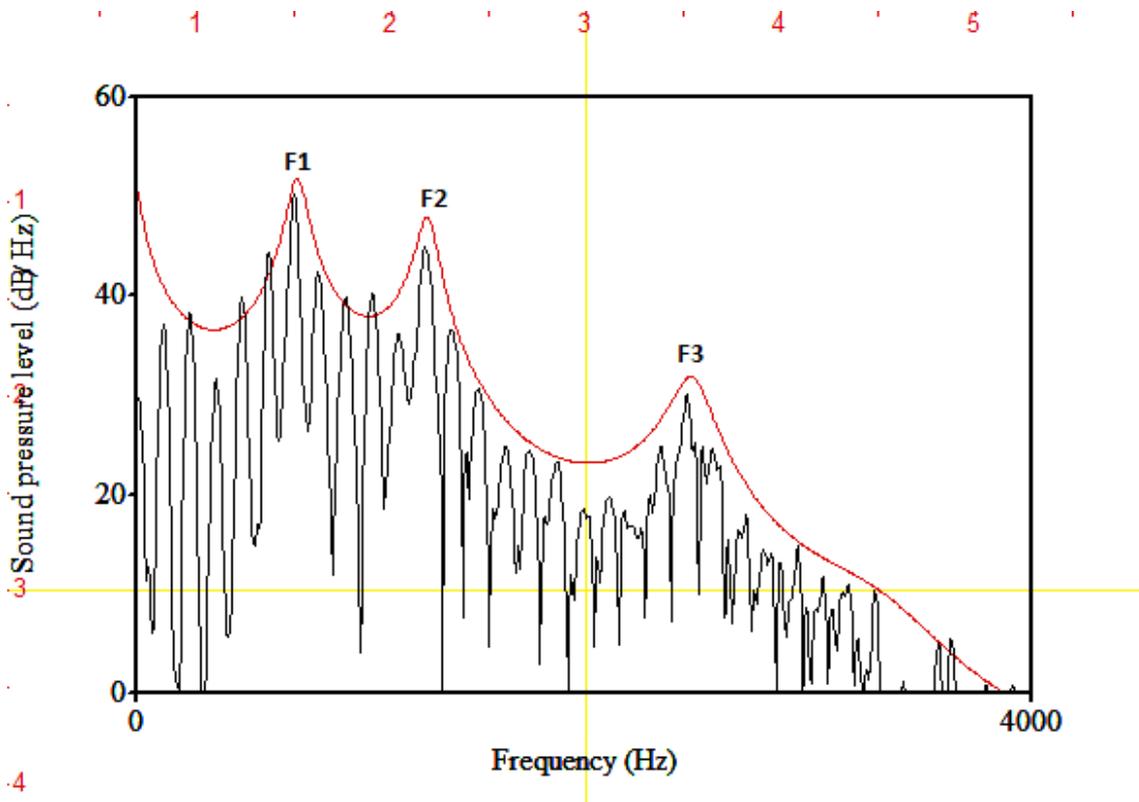


Figura 1. Espectros de FFT (em preto) e de LPC (em vermelho) em porção estacionária da vogal [a] em válido.

Extraídas as medidas referentes às 10 repetições foram calculados os valores médios para cada parâmetro. Esses valores médios são apresentados na Tabela 1, a seguir.

Tabela 1. Valores médios dos parâmetros acústicos

Vogal	Dur (ms)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F0 (Hz)	Int (dB)
A	139	692	1267	2457	112	78
I	54	374	1893	2980	117	72
O	52	438	1394	2738	115	72
A	107	663	1295	2387	111	79
I	100	327	2241	3121	112	69
O	46	402	1453	2764	120	70
A	83	626	1296	2366	115	78
I	62	358	2095	3029	108	69
O	100	451	1096	2793	110	76

Obtidos os valores médios em relação a cada parâmetro acústico, foram sintetizados os estímulos para aplicação do teste de percepção. O primeiro passo, para a preparação dos estímulos, consistiu em substituir para cada uma das vogais das 3 palavras do *corpus*, os valores médios das 10 repetições de cada palavra. Esses valores médios constituiram os estímulos bases que foram etiquetados com como “Valido_MP, vaLIdo_MP e valiDO_MP”. Os valores médios foram incorporados e os estímulos sintetizados com o uso do algoritmo PSOLA (Pitch-Synchronous Overlap and Add).

Os estímulos de base foram manipulados para criar novos padrões. A manipulação consistiu em substituir os valores próprios de um determinado padrão pelo de outro.

Para exemplificar os procedimentos utilizados, registramos, a seguir, os tipos de estímulos produzidos:

- 1) Um estímulo com os valores de f0, duração e intensidade do padrão de *VALido* (estímulo de base Valido_MP).
- 2) Um estímulo com os valores de f0, duração e intensidade do padrão de *vaLIdo* (estímulo de base vaLIdo_MP).
- 3) Um estímulo com os valores de f0, duração e intensidade do padrão de *valiDO* (estímulo de base valiDO_MP).

- 4) Um estímulo com os valores de f_0 e intensidade do padrão de *VALido* e os valores de duração do padrão de *vaLIdo* (estímulo de base *VALido_MP_DurPA*).
- 5) *Um estímulo com os valores de f_0 e intensidade do padrão de VALido e os valores de duração do padrão de valiDO (estímulo VALido_MP_DurOX).*
- 6) *Um estímulo com os valores de f_0 e intensidade do padrão de vaLIdo e os valores de duração do padrão de Valido (estímulo vaLIdo_MP_DurPRO).*
- 7) Um estímulo com os valores de f_0 e intensidade do padrão de *vaLIdo* e os valores de duração do padrão de *valiDO* (estímulo *vaLIdo_MP_DurPA*).
- 8) *Um estímulo com os valores de f_0 e intensidade do padrão de valiDO e os valores de duração do padrão de VALido (estímulo valiDO_MP_DurPRO).*
- 9) *Um estímulo com os valores de f_0 e intensidade do padrão de valiDO e os valores de duração do padrão de VaLIdo (estímulo valiDO_MP_DurPA).*

Com os procedimentos relatados acima foram constituídos os estímulos a serem submetidos ao teste de percepção.

3.1. Tarefa de percepção

Os estímulos foram apresentados a um grupo de 30 juízes em computadores do Laboratório de Informática da Pontifícia Universidade Católica de São Paulo com o uso de fones de ouvido. Os juízes puderam ouvir um mesmo estímulo mais de uma vez.

A apresentação dos estímulos foi aleatorizada para evitar o erro sistemático conhecido como efeito de lista.

Na Figura 2, a seguir, apresentamos dois dos estímulos produzidos. Um deles é o estímulo de base e o outro o gerado com a manipulação da duração.

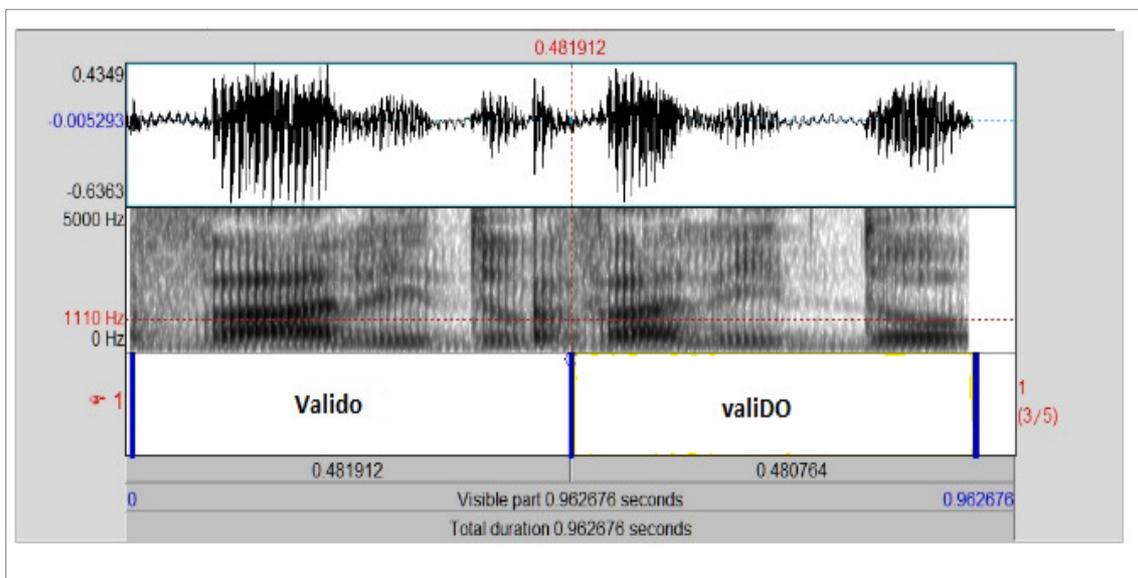


Figura 2. Concatenação do estímulo “VALido” com o estímulo resultante da incorporação dos seus valores de duração no estímulo base “valiDO”.

4. Resultados

Os resultados da tarefa de percepção encontram-se na tabela a seguir e indicam que a alteração da duração afetou a identificação da proeminência acentual.

Tabela 2. Tipo de estímulo e percentagens de identificação do padrão acentual.

ESTÍMULOS	PR	PA	OX
VALido_MPDur_OX	64%	3%	33%
VALido_MPDur_PA	93%	7%	0%
VALido_MP	80%	7%	13%
valiDO_MP	30%	10%	60%
valiDO_MP_DurPA	53%	17%	30%
valiDO_MP_DurPRO	33%	20%	47%
vaLIdo_MP	27%	57%	16%
vaLIdo_MPDur_OX	70%	7%	23%
vaLIdo_MPDur_PRO	83%	10%	7%

Quando a transformação envolveu a mudança de proparoxítona (vaLIdo) para oxítona (valiDO) ou para a paroxítona (vaLIdo), as alterações dos padrões acentuais não foram percebidas, o que indica que o padrão formântico teve importante papel na identificação.

Quando a transformação envolveu mudança de paroxítona (vaLIdo) para a proparoxítona (Valido) a alteração de padrão foi percebida, mas quando envolveu a mudança de paroxítona (vaLIdo) para oxítona (valiDO), o padrão percebido foi de proparoxítona, o que pode ser interpretado como resultado da influência do alongamento da pré-tônica e da configuração formântica da pós-tônica.

Quando a transformação envolveu a mudança de oxítona (valiDO) para proparoxítona (Valido), não foi percebida a alteração de padrão acentual, o que indica que o padrão formântico teve importante papel na identificação, mas quando a mudança foi da oxítona (valiDO) para a paroxítona (vaLIdo), a duração não foi suficiente para a identificação do padrão paroxítono e a estrutura formântica não foi suficiente para a identificação do padrão acentual oxítono.

Na pronúncia da palavra “valido” com três padrões acentuais, a vogal final é produzida como tônica [o] no padrão oxítono. Em formas enfáticas, o [o] ocorrer na posição átona. Nos padrões paroxítonos e proparoxítonos a vogal pós-tônica é geralmente [ɔ].

Um fator interveniente pode ser a frequência de uso do adjetivo em relação ao verbo. Acrescente-se o fato de que diante de fronteira prosódica o alongamento pré-pausal é comum na grande maioria das línguas, inclusive no PB, o que pode ter influenciado no julgamento dos juízes, ou seja, o alongamento pode ser percebido como índice de fronteira e não de acento.

5. Considerações finais

Conclui-se que a frequência dos formantes é uma pista importante na percepção do acento em Português Brasileiro. O aumento da duração foi suficiente para transformar a paroxítona em proparoxítona, mas não foi suficiente para alterar outros padrões, o que mostra a força da estrutura formântica.

Referências Bibliográficas

- ALFANO, I. La percezione dell'accento lessicale: un esperimento sull'italiano a confronto con lo spagnolo. [Tesi di laurea in linguistica generale], 2005.
- BARBOSA, P. A. "Syllable-timing in Brazilian Portuguese": uma crítica a Roy Major. *DELTA*, 16:369-402, 2000.
- BARBOSA, P. A. Explaining Cross-Linguistic Rhythmic Variability via a Coupled-Oscillator Model of Rhythm Production. BEL, B. & MARLIEN, I. (eds.) In: *Proceedings of the Speech Prosody 2002 Conference*, 11-13 April, Aix-en-Provence: Laboratoire Parole et Langage, 163-166, 2002.
- DELGADO-MARTINS, M. R. *Ouvir Falar. Introdução à Fonética do Português*. Lisboa: Editora Caminho, 1986.
- FANT, G. *Acoustic Theory of Speech Production*. Mouton, The Hague, 1960.
- GAMA-ROSSI, A. A. J. Qual é a natureza do acento secundário no português brasileiro? *Cadernos Centro Universitário São Camilo*, 4 (1): 77-92, 1998.
- HAYWARD, K. *Experimental Phonetics*. Harlow, England: Longman, 2000.
- LAVIER, J. *Principles of Phonetics*. Cambridge: Cambridge University Press, 1994.
- LLISTERRI, J., MACHUCA, M. J., de la MOTA, C., RIERA, M., RÍOS, A. The perception of lexical stress in Spanish. *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, 2023-2026, 2003.
- MASSINI, G. A duração no estudo do acento e do ritmo em português. [Dissertação de mestrado]. Universidade Estadual de Campinas, 1991.