

---

## PERCEPÇÃO DO TIMBRE VOCÁLICO X FREQUÊNCIA FUNDAMENTAL: ANÁLISE EXPERIMENTAL<sup>11</sup>

Priscila de Jesus Ribeiro\*  
(UESB)

Vera Pacheco\*\*  
(UESB)

### RESUMO

Este trabalho propõe analisar se a frequência fundamental corrobora para a percepção das vogais médias não arredondadas no dialeto de Vitória da Conquista – BA (V/C). Para isso, foi elaborado um corpus constituído pelas vogais em questão, cada palavra foi inserida em uma sentença, em seguida, um falante conquistense fez uma leitura em situação normal de fala corrente, posteriormente a  $F_0$  foi manipulada tanto para valores ascendentes quanto descendentes em 75%, 50% e 25% do valor original. Os testes foram aplicados com três informantes. Os dados apontam que a manipulação da  $F_0$  afeta a percepção da vogal média não arredondada.

**PALAVRAS-CHAVE:** frequência fundamental-percepção-timbre vocálico

---

<sup>11</sup>Trabalho vinculado ao projeto de pesquisa “investigação acústico-perceptual de aspectos segmentais e suprasegmentais de línguas naturais”. Coordenado pela Prof.<sup>a</sup> Dr.<sup>a</sup> Vera Pacheco / UESB

\*Aluna do curso de mestrado em Linguística e bolsista CAPES

\*\*Orientadora. Doutora em Linguística pela Unicamp. Professora no Departamento em Estudos Linguísticos e Literários (DELL)/UESB.

## INTRODUÇÃO

Uma das questões de principal interesse da fonética experimental consiste em compreender quais as informações utilizadas pelos ouvintes para a identificação das vogais. Alguns estudiosos (JENKINS; STRANGE; EDMAN, 1983; GAYVERT, 1993; HILLENBRAND, 1993; TUERNER, 2003; PATERSON, 2003) propõem que a percepção dos segmentos vocálicos está diretamente relacionada às frequências dos dois formantes mais baixos ( $F_1$  e  $F_2$ ). Esses autores propõem que há, no estado estacionário, um espaço acústico entre as frequências formânticas que determina a percepção vocálica.

Outros pesquisadores (BEINUM, 1988; BERGEM, 1988; POLS, 1988; BOSCH, 1998; CLAES, 1998; COMPERNOLLE, 1998; DOLOGLOU, 1998) adotam a hipótese de que, além das frequências dos dois formantes mais baixos, a discriminação de vogais está também relacionada a outras informações espectrais. Dentre esses estudiosos, pode-se citar Strange (1995) que propõe integrar o estado estacionário e a natureza da transição formântica com a informação espectral. A autora propõe que juntamente com as frequências formânticas, a duração e a frequência fundamental corroboram para a percepção vocálica.

Estudos que, além da frequência formântica, consideram outros aspectos do espectro vocálico, apontam que as demais informações espectrais podem ser igualmente importantes para a compreensão de como se dá a discriminação vocálica. Considerando que a  $F_0$  pode ser uma importante informação utilizada pelo ouvinte para a percepção da qualidade vocálica e considerando que as vogais médias possuem padrões formânticos com valores aproximados, o presente trabalho propõe investigar se a  $F_0$  interfere na percepção das vogais médias não arredondadas em ouvintes naturais de Vitória da Conquista – BA.

## MATERIAL E MÉTODOS

Foi elaborado um *corpus* constituído por palavras que continham em sua composição vogais médias abertas e fechadas arredondadas e não arredondadas em relação de oposição.

<b>Vogal média alta arredondada</b>	<b>Vogal média baixa arredondada</b>	<b>Vogal média alta não arredondada</b>	<b>Vogal média baixa não arredondada</b>
Côrte	Cóрте	Zêlo	Zélo
Chôro	Chóro	Pêso	Péso
Acôrdο	Acórdο	Sêde	Séde
Fôrma	Fórma	Vê-lo	Vélo
Sou	Sol	Pê	Pé

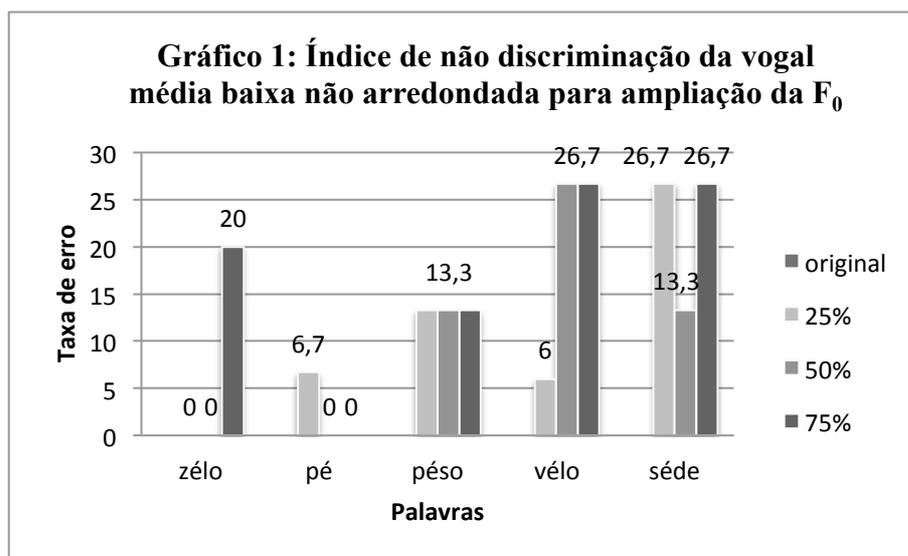
As palavras cujas sentenças foram classificadas pelos julgadores como inadequadas foram inseridas na frase veículo “Digo \_\_\_\_\_ baixinho”.

A frequência fundamental das vogais em questão foi manipulada utilizando o programa *Praat* de forma a obter 25%, 50% e 75% do valor original, tanto para valores ascendentes quanto para valores descendentes. Elaboramos, portanto, diferentes sinais sonoros com a  $F_0$  manipulada tanto para valor inferior ao original, quanto superior nas taxas discriminadas acima.

Para a elaboração dos testes de percepção, utilizamos o programa *Transcriber*, no qual foi possível inserir os sinais de fala, tanto o original quanto os manipulados ( $F_0$  acrescida e diminuída em 25%, 50% e 50%) nas diferentes taxas aqui avaliadas.

## RESULTADOS E DISCUSSÃO

Abaixo seguem os gráficos correspondentes aos resultados parciais dos testes de discriminação e identificação das vogais médias baixas não arredondadas.



A taxa de perda de recuperação da qualidade vocálica correspondente ao sinal sem manipulação de  $F_0$  varia razoavelmente de palavra para palavra. Os dados apontam que na palavra *zélo*, a ampliação da  $F_0$  a 25% e a 50% não interferiu na identificação das vogais, ou seja, os ouvintes obtiveram 0% de erro, o que significa que a sua taxa de recuperação do timbre vocálico foi de 100% em relação ao original.

---

Taxas de manipulação, contudo, altera essa tendência e compromete a inteligibilidade do timbre vocálico. Quando, por exemplo, a manipulação foi de acréscimo de 75%, o ouvinte não recuperou o sinal em 20% dos casos. Sendo assim, a discriminação da vogal /E/ na palavra *zélo* foi afetada em 20% quando a  $F_0$  foi ampliada para o valor de 75%.

Na palavra *pé*, os índices de 75 e 50% de ampliação da  $F_0$  não interferiram na identificação da vogal, ou seja, o ouvinte conseguiu recuperar o sinal em 100% dos casos, no entanto, para valores de 25% a ampliação da  $F_0$  acarretou 6,7% de erro. A manipulação da  $F_0$  parece acarretar, portanto, uma pequena interferência na inteligibilidade da vogal em questão.

Na ampliação da  $F_0$  para a palavra *pés*, o índice de interferência é de 13,3% para os três casos de manipulação, 25%, 50% e 75%, em relação ao original. A ampliação da  $F_0$ , portanto, interfere a inteligibilidade da vogal /E/ no índice de 13,3%.

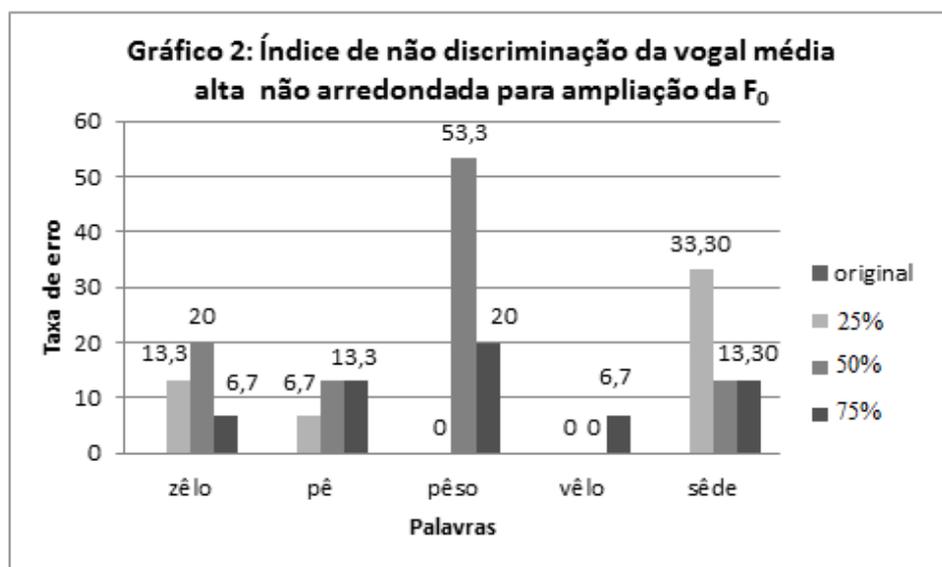
Na palavra *vélo*, o índice percentual 25 gera 6% de erro na inteligibilidade do sinal. Para os valores de 50% e 75%, a ampliação da  $F_0$  suscita uma taxa de erro de 26,7% em relação ao sinal original. A percepção da vogal /E/ parece sofrer, portanto, uma pequena interferência para valores da  $F_0$  ampliados em 25%, a qual se amplia com a ascendência da manipulação para os valores de 50 e 75%.

A identificação da vogal /é/ na palavra *séde* é afetada em 26,7% em relação a original. Para a taxa de 50%, o índice de não discriminação cai para 13,3%. Nesse caso, portanto, as taxas de 25 e 75% parecem afetar a percepção da vogal /E/, a taxa de 50%, índice intermediário, a não discriminação cai para metade do valor, ou seja, a identificação foi menos afetada para o valor de ampliação intermediário.

A ampliação da  $F_0$  parece, portanto, afetar a discriminação das vogais médias baixas não arredondadas, uma vez que em todos os casos

analisados houve taxa de erro para a discriminação da vogal /E/, ainda que em algumas ocorrências a interferência parece pequena, a manipulação suscitou erros.

Resultado muito semelhante é também encontrado para a média alta não arredondada, como podemos observar no gráfico 2.



É possível observar a partir do gráfico 2 que a discriminação da palavra *zêlo* foi afetada para os três índices de manipulações. Para os valores manipulados, a maior taxa de erro foi para o índice de ampliação de 50%, a menor para 75%, e interferência intermediária para a taxa de 25%. A ampliação da F<sub>0</sub> parece interferir a inteligibilidade da vogal média alta não arredondada.

A inteligibilidade da palavra *pê* sofreu interferência nos três valores manipulados, assim, os índices de 50% e 75% alteraram a inteligibilidade do sinal em 13,3%, o índice de 25% gerou interferência na discriminação da vogal em 6,7%. Observa-se, portanto, que a ampliação da F<sub>0</sub> em 25% interferiu a discriminação das vogais, e tal discriminação foi ainda mais afetada e em igual proporção quando essa foi ampliada para os índices de 50% e 75%.

---

Na palavra *pêso*, a ampliação da  $F_0$  em 25% não ocasionou alteração na discriminação da vogal em relação ao sinal original, no entanto, quando essa foi manipulada a taxa de 50%, a discriminação da vogal em questão foi afetada em 53,3% em relação ao original, e quando manipulada em 75%, a taxa de erro de discriminação foi de 20%. Os dados apontam, portanto, que o ouvinte recuperou a vogal média alta não arredondada em 46,7% dos casos, ou seja, em mais de 50% das ocorrências o ouvinte não conseguiu recuperar o sinal quando esse foi manipulado para a taxa percentual de 50% do valor original.

A discriminação da palavra *velo* não sofreu interferência quando manipulada para os valores de 25 e 50%, observa-se uma pequena interferência quando o sinal é manipulado para o valor de 75%.

A palavra *sêde* sofreu interferência para os três valores de ampliação, assim, em 25% a interferência foi de 33,3% e, para as taxas percentuais de 50 e 75, a interferência foi de 13,3%. Observa-se, portanto, que a discriminação da vogal média baixa não arredondada, quando ampliada em 25% do sinal original, sofreu interferência em mais de 30% dos casos, quando a manipulação foi ampliada em 50 e 75%, o índice de erro caiu em 20%. Nesse contexto, portanto, a ampliação da  $F_0$  interferiu a discriminação da vogal média baixa não arredondada.

## **CONCLUSÕES**

Os resultados apontam que a manipulação da frequência fundamental tanto para valores ascendentes quanto para valores descendentes afetam a percepção da vogal média. Em muitos casos o ouvinte identificou a vogal média alta como sendo uma vogal média baixa e uma vogal média alta como sendo uma vogal média alta.

Os dados sinalizam, portanto, que a frequência fundamental parece ser um importante recurso utilizado pelo ouvinte para a discriminação da vogal média, uma vez que, em alguns contextos, a  $F_0$  afetou diretamente a percepção das vogais em questão. Assim, vogais em que a configuração formântica dos dois formantes mais baixos (F1 e F2) possuem valores aproximados, como é caso das vogais por nós investigadas, médias altas e médias baixas, a  $F_0$  parece constituir um importante recurso utilizado pelo ouvinte para a discriminação das vogais médias não arredondadas.

## REFERÊNCIAS

- BERGEM, D. R., POLS, L. C. W; BEINUM, F. J. Perceptual normalization of the vowels of a man and a child in various contexts. *Speech Communication*, v. 7, p. 1-20, 1988.
- BOERSMA, P; WEENINK, D. **Praat**: Doing phonetics by computer. Version 5.1.01, <http://www.praat.org>, 2009.
- CLAES, T.; DOLOGLOU, I; BOSCH, L. T; COMPERNOLLE, V. D. A novel feature transformation for vocal tract length normalization in automatic speech recognition. **IEEE Transactions on Speech and Audio Processing**, v. 6, p. 549–557, 1998.
- HILLENBRAND, J; GAYVERT, R. T. Identification of steady-state vowels synthesized from the Peterson and Barney measurements. **Journal of the Acoustical Society of America**, Kalamazoo, v. 94, n. 2, p. 668-674, 1993.
- JENKINS, J. J.; STRANGE, W.; EDMAN, T. R. Identification of vowels in "vowelless" syllables. **Perception and Psychophysics**, Minneapolis, v. 34, p. 441-450, 1983.

---

STRANGE, W. Cross-language studies of speech perception. In: Winifred Strange. **Speech Perception and Linguistic Experience**: Issues in Cross-Language Research, Baltimore: York Press, 1995.

TUERNER, R. E; PATTERSON, R. D. An analysis of the size information in classical formant data: Peterson and Barney (1952) revisited. **The Acoustical Society of Japan**, Cambridge, v. 33, n. 9, p. 585-589, 2003.