

## **Veredas atemática**

### **Volume 18 nº 2 – 2014**

---

#### **A frequência de ocorrência lexical e a realização das vogais anteriores do inglês língua estrangeira**

Clerton Luiz Felix Barboza (UERN)  
Wilson Júnior de Araújo Carvalho (UECE)

**RESUMO:** Analisamos a influência da variável frequência de ocorrência de itens lexicais na realização de exemplares vocálicos anteriores do Inglês Língua Estrangeira (ILE). As teorias fonológicas baseadas no uso apontaram a variável frequência de ocorrência como relevante na aquisição de línguas estrangeiras. Realizamos uma análise acústica de F1, F2, F3 e da distância euclidiana das vogais [i, ɪ, eɪ, ε, æ] por professores brasileiros de ILE. Os dados apontaram pouca influência da frequência de ocorrência, indicando sua irrelevância na construção da fonologia do ILE dos informantes.

Palavras-chave: fonologia de uso; modelo de exemplares; item lexical.

#### **Introdução**

O presente artigo tem por objetivo analisar os efeitos da frequência de ocorrência de diversos itens lexicais na realização do detalhe fonético dos exemplares vocálicos anteriores [i, ɪ, eɪ, ε, æ], coletados junto a professores brasileiros de Inglês Língua Estrangeira (ILE).

Efeitos de frequência de ocorrência e sua influência na realização de diversos fenômenos fonético-fonológicos do inglês vêm sendo descritos pela literatura (BYBEE, 2001, 2003, 2008; PIERREHUMBERT, 2001, 2002a; BUSH, 2001; FAULKES; DOCHERTY, 2006; JOHNSON, 2007; CLOPPER; PIERREHUMBERT, 2008; ELLIS, 2008). No entanto, pesquisas envolvendo a variável na produção de falantes do ILE restringiram-se a discussões teóricas (BYBEE, 2008). Poucos estudos de campo envolvendo a variável frequência foram

desenvolvidos, em especial envolvendo falantes brasileiros de ILE (FERREIRA, 2011; BARBOZA, 2013). O presente trabalho justificou-se pela busca do melhor entendimento da influência da variável em questão na produção de professores brasileiros de ILE. Reanalisamos neste momento dados coletados anteriormente (BARBOZA, 2008) na busca de indícios da relevância da variável frequência de ocorrência.

Temos como pergunta-problema: de que maneira a frequência de ocorrência de palavras do inglês influi na realização dos exemplares vocálicos anteriores [i, ɪ, eɪ, ε, æ]?

Considerando os modelos fonológicos multirrepresentacionais adotados, a fonologia de uso (BYBEE, 2001) e o modelo de exemplares (JOHNSON, 1997), tomamos por hipótese básica que os exemplares vocálicos anteriores em palavras do ILE de maior frequência de ocorrência apresentam detalhe fonético distinto dos exemplares vocálicos realizados em palavras de menor frequência. A distinção decorreria da maior exposição aos itens de alta frequência, o que implicaria na aquisição mais acurada do detalhe fonético associado a tais exemplares lexicais (BYBEE, 2008). Desdobramos a hipótese básica nas três hipóteses secundárias a seguir:

- a) Exemplares vocálicos em palavras de maior frequência de ocorrência do ILE são realizados num espaço acústico de maiores dimensões quando comparados aos exemplares vocálicos realizados em palavras de menor frequência de ocorrência, i.e., os exemplares vocálicos anteriores de palavras mais e menos frequentes apresentam valores distintos de F1, F2 e F3;
- b) Exemplares vocálicos em palavras de maior frequência de ocorrência são realizados com duração distinta dos exemplares vocálicos realizados em palavras de menor frequência, i.e., os exemplares vocálicos [i, æ] das palavras mais frequentes apresentam uma maior duração (em milissegundos), enquanto os exemplares [ɪ, ε] das palavras mais frequentes apresentam uma menor duração, quando comparados aos exemplares de palavras menos frequentes.
- c) Exemplares vocálicos em palavras de maior frequência de ocorrência apresentam maior distância euclidiana quando comparados aos exemplares realizados em palavras de menor frequência de ocorrência, i.e., as realizações acústicas dos exemplares vocálicos das palavras mais frequentes são realizadas num espaço acústico mais amplo.

Além desta introdução, o presente artigo é composto por mais três seções. Na primeira seção realizamos discussão sobre o papel da frequência de ocorrência e tipo na produção de itens lexicais, considerando estudos sobre o inglês como língua materna e estrangeira, segundo os modelos fonológicos multirrepresentacionais. Em sequência, na segunda seção, apresentamos pormenores referentes à metodologia de coleta e análise de dados do estudo. Na terceira seção, expomos resultados e discussões propiciados pela análise do detalhe fonético dos sons vocálicos anteriores [i, ɪ, eɪ, ε, æ] com relação às hipóteses apresentadas anteriormente. Finalmente, concluímos o estudo e apontamos possíveis desdobramentos da pesquisa.

## **1. O papel da frequência nos modelos fonológicos multirrepresentacionais**

A análise do papel da frequência de ocorrência na produção de fenômenos linguísticos vem sendo propiciada pela visão de gramática baseada no uso. O paradigma baseou-se na

ideia de que a gramática, ao invés de um conjunto abstrato de regras e estruturas indiretamente associadas às experiências linguísticas, constituiu-se de forma emergente a partir da categorização promovida pelo uso da língua (BECKNER et. al., 2009). A gramática refletiria detalhes associados à probabilidade de ocorrência e coocorrência dos fenômenos linguísticos. A frequência de uso foi considerada um fator impactante na estrutura linguística, com efeitos comprovados em diversos níveis gramaticais, como no fato de que itens lexicais mais frequentes serem resistentes à regularização (BYBEE, 2001), na observação de características probabilísticas no processamento sintático e lexical (ELLIS, 2002), bem como o papel da frequência de uso em fenômenos de gramaticalização (BYBEE, 2003).

Em se tratando de modelos fonológicos multirrepresentacionais, a fonologia de uso (BYBEE, 2001) e o modelo de exemplares (JOHNSON, 1997) constituem os principais paradigmas a irem ao encontro da visão de gramática baseada no uso. Para ambos os modelos, discutidos a seguir, os papéis da frequência de ocorrência e de tipo são centrais.

Na fonologia de uso, a frequência de ocorrência (ou de *token*) refere-se ao número de vezes em que uma unidade linguística, geralmente uma palavra, ocorre num *corpus*. A importância de analisar a frequência de ocorrência associa-se a seus efeitos na emergência de fenômenos de mudança linguística:

- a) em itens lexicais de maior frequência de ocorrência, seriam favorecidos os mecanismos de mudança linguística de origem fonética, decorrentes da maior automação/sobreposição de sequências motoras justapostas. Desse modo, palavras com maior frequência de ocorrência apresentariam mudança linguística de origem fonética primeiramente do que palavras com menor frequência de ocorrência, como observado nos estudos de Guimarães (2004) e Leite (2006);
- b) em itens lexicais de menor frequência de ocorrência, seriam favorecidos os mecanismos de mudança linguística por generalização ou analogia a itens mais frequentes. Dada a menor familiaridade com itens menos frequentes, tenderíamos a favorecer mecanismos de mudança sonora em direção a um sistema linguístico mais homogêneo.

Ambos os mecanismos podem afetar um único elemento linguístico (BYBEE, 2001). Mudanças sonoras diacrônicas poderiam começar pela maior sobreposição dos gestos articulatórios, fenômeno de origem fonética, e posteriormente serem simplificadas através da perda dos gestos, por meio de analogia. Foi o caso da monotongação da sequência *eo* em /œ/ do inglês antigo resultante de uma sobreposição dos gestos vocálicos, que afetou primeiramente palavras de maior frequência de ocorrência. Posteriormente, a existência da vogal anterior arredondada /œ/, incomum no sistema vocálico do inglês antigo, acarretou uma segunda mudança, por analogia, no inglês médio. A perda da sobreposição do gesto anterior arredondado /œ/ em /e/ emergiu primeiramente nas palavras menos frequentes para só então afetar as palavras mais frequentes, e.g. *deop*, *dep*, ‘deep’.

A frequência de tipo, por sua vez, refere-se ao número de ocorrências de um determinado padrão linguístico. A frequência de tipo é de grande importância na fonologia de uso, uma vez que determina a produtividade. Quanto mais frequente um determinado padrão, maior a sua produtividade, implicando numa maior probabilidade de uso do padrão.

Uma maior frequência de tipo do afixo do passado regular *-ed*, no inglês, contribuiu para a melhor análise de palavras como *played* e *spilled*, em oposição a palavras como *broke* e *went*. A recorrência do sufixo regular, quando comparado a outros padrões do tempo passado, implicou na maior produtividade do padrão em novos verbos do inglês (BYBEE, 2001).

Neologismos tenderiam a usar o padrão mais recorrente na língua, por analogia às formas de maior frequência de tipo.

A relevância das características fundamentais da gramática e da fonologia de uso e suas implicações no ensino de ILE foram discutidas por Bybee (2008). A autora apresentou os resultados esperados dos efeitos de frequência nos aprendizes de línguas estrangeiras. No que tange à frequência de ocorrência, enfatizou os efeitos da repetição dos itens de alta frequência, bem como sua autonomia e os efeitos de redução sonora.

Itens lexicais cujos aprendizes do ILE fossem expostos com maior frequência apresentariam uma produção mais acurada. Tal premissa fundamentou as hipóteses levantadas neste estudo.

A constante repetição dos itens mais frequentes implicaria, todavia, num efeito negativo para o aprendiz de línguas estrangeiras: a redução fonética. Causado pela automação de rotinas neuromotoras, decorrente da maior frequência de ocorrência, o efeito facilitaria a produção linguística dos nativos, ao passo que constituiria um desafio para não-nativos. Uma frase como *I am going to* poderia ser pronunciada como [aimgoɪntu] em sua forma canônica, ou como [aimənə] em sua forma mais reduzida. O domínio de ambas as formas por parte de um aprendiz estrangeiro seria dificultada, uma vez que sua língua materna determinaria mecanismos de redução e coarticulação distintos da língua-alvo.

Efeitos da frequência de tipo seriam também importantes para os aprendizes de línguas. Uma maior frequência de tipo implicaria uma representação mais robusta das estruturas linguísticas. A familiaridade do tipo linguístico garantiria ao aprendiz os efeitos de produtividade, por meio da analogia, discutidos anteriormente.

De forma geral, Bybee (2008) concluiu que a compreensão da gramática enquanto fenômeno emergente do uso implica boas e más notícias para o aprendiz de língua estrangeira, pois, por um lado

[...] é difícil para o aprendiz adulto ou o aprendiz em sala de aula ter uma exposição semelhante ao nativo de sua língua alvo e, além disso, muitos adultos não possuem a plasticidade necessária ao estabelecimento de rotinas neuromotoras para a nova língua. (BYBEE, 2008, p.232-233)<sup>1</sup>

Por outro lado, a autora enfatizou que as mesmas habilidades cognitivas utilizadas na construção da língua materna, como a analogia, a categorização e a automatização, poderiam ser utilizadas na construção de uma língua estrangeira. Tais habilidades, já desenvolvidas, permitiriam aos aprendizes adquirir línguas estrangeiras sem uma reprodução fiel da frequência a que estaria sujeito o falante nativo.

Os trabalhos de Pierrehumbert (1999, 2001, 2002a, 2002b, 2003, 2006) ajudaram a desenvolver o modelo de exemplares. Pierrehumbert (2001) apontou que exemplares seriam armazenados na forma de conjuntos ou nuvens, cujas propriedades aproximar-se-iam em maior ou menor grau de um exemplar prototípico, em um mapa cognitivo. A frequência de uso seria importante para determinar a robustez dessas nuvens. A autora (PIERREHUMBERT, 2002b) advogou a favor do caráter probabilístico da aquisição do detalhe fonético. Assim, o detalhe fonético seria adquirido de forma gradual, associado diretamente à recorrência dos padrões no sistema. Apesar de a observação referir-se à

---

<sup>1</sup> “it is difficult for the adult learner or the classroom learner to get completely native-like exposure to the target language. In addition, many adults lack the plasticity needed to set up native-like neuromotor routines for the new language.”

aquisição do inglês língua materna, testamos neste artigo sua validade no que tange ao percurso de construção da fonologia do inglês língua estrangeira.

Grande parte dos estudos supramencionados abordou a frequência de uso como um fator determinante no entendimento de fenômenos fonológicos em situação de primeira língua. São escassos os estudos envolvendo o controle explícito da variável frequência de ocorrência no percurso de construção de línguas estrangeiras de aprendizes brasileiros.

O estudo de Ferreira (2011) reanalisou dados de palavras CVC de um estudo anterior (FERREIRA, 2007) à luz da fonologia de uso, com especial atenção ao papel da frequência de ocorrência das palavras analisadas. Ao analisar efeitos de frequência de itens lexicais em posição de *coda*, a autora concluiu que desvios da forma-alvo foram mais recorrentes nos itens de baixa frequência de ocorrência. A busca pela frequência de ocorrência das palavras *dad* e *pet* indicou que o primeiro item era mais frequente e apresentava uma realização com menor número de desvios. A confirmação da hipótese de que palavras menos frequentes apresentaram um maior número de desvios da forma-alvo levaram Ferreira (2011) a concluir que a frequência de ocorrência seria uma variável importante em estudos envolvendo o ILE de aprendizes brasileiros.

Por outro lado, o estudo de Barboza (2013) analisou efeitos da frequência de ocorrência de itens lexicais na produção do ILE de aprendizes brasileiros de falares palatalizador e não-palatalizador. O estudo focou na emergência da palatalização das oclusivas alveolares /t, d/ do ILE. Os dados apontaram a irrelevância do fator frequência de ocorrência na maior ou menor emergência do fenômeno nos aprendizes de ambos os falares.

Observamos, portanto, pouca discussão do fator frequência de ocorrência na literatura envolvendo o ILE de aprendizes brasileiros. O presente estudo pretende oferecer subsídios para o melhor entendimento da variável no percurso de construção da fonologia de línguas estrangeiras, focando na influência da variável na realização do detalhe fonético de exemplares vocálicos. Na seção a seguir, descrevemos os pormenores metodológicos de coleta e análise de dados deste estudo.

## 2. Metodologia

A presente pesquisa pode ser delimitada como indutiva e de cunho transversal. Adicionalmente, faz uso de uma metodologia experimental, analisando dados quantitativos.

Os informantes foram professores de ILE de três cidades do oeste do Rio Grande do Norte: Açu, Pau dos Ferros e Mossoró. As primeiras cidades contribuíram com 4 informantes cada, enquanto a última contribuiu com 12 informantes, totalizando 20 informantes. O número de informantes em cada localidade visou refletir diferenças demográficas entre as localidades.

Todos os informantes eram profissionais de ensino que faziam uso cotidiano das habilidades orais do inglês em suas salas de aula. A média de idade dos sujeitos foi de 31,1 anos. O tempo de experiência médio no ensino de ILE foi de 8,4 anos. Entre os informantes encontramos desde recém-formados até professores universitários.

Os sujeitos foram entrevistados entre os meses de setembro a novembro de 2007, no decorrer de pesquisa anterior (BARBOZA, 2008). Dados daquele estudo permitiram uma

reanálise em que tomamos como variável dependente a frequência de ocorrência dos itens lexicais utilizados naquele momento.

As palavras utilizadas na pesquisa foram monotongos do inglês cuja combinação fonotática utilizou o padrão CVC. A primeira posição C foi ocupada por consoante oclusiva, vozeada ou desvozeada. A posição V foi ocupada por um dos cinco exemplares vocálicos anteriores objeto do estudo [i, ɪ, eɪ, ε, æ]. A segunda posição C foi ocupada apenas por consoante oclusiva desvozeada. Tal combinação fonotática foi selecionada por facilitar os procedimentos de coleta de dados espectrais e de duração. Um total de 25 palavras foi analisado. Selecionamos 10 palavras para análise nesta pesquisa, a partir da busca de pares de palavras contendo o mesmo exemplar vocálico anterior, com marcante variação no que tange à frequência de ocorrência de cada item lexical. Apresentamos no Quadro 1 o conjunto, juntamente com a frequência de ocorrência de cada item lexical.

beat (40.772)	teat (91)	0,22%
bit (83.131)	tit (410)	0,49%
date (32.471)	bate (82)	0,25%
debt (22.401)	Beck (3.183)	14%
back (569.803)	tat (426)	0,07%

**Quadro 1** – Palavras utilizadas no estudo (frequência de ocorrência entre parênteses), seguidas do percentual de ocorrência das palavras menos frequentes com relação às mais frequentes.

Fonte: *Corpus of Contemporary American English - COCA*<sup>2</sup>

Observamos no quadro anterior que há grande diferença quantitativa entre as palavras de alta e baixa frequência de ocorrência. Dos 5 pares anteriores, 4 apresentaram uma relação de frequência inferior a 0,5%. A única exceção ocorreu no par *debt-Beck*, em que a frequência do item de menor recorrência correspondeu a 14% do item de maior recorrência. Utilizamos o referido par devido ao fato de que tínhamos por opção substituir o item *debt* pelo item *get*. Todavia, observamos que a oclusiva velar vozeada [g], em posição de *onset*, influenciou a realização do detalhe fonético do exemplar vocálico [ε]. O efeito foi reportado na literatura sobre coarticulação (HILLENBRAND; CLARK; NEAREY, 2001) e foi corroborado nos dados através da realização de uma ANOVA de medidas repetidas. Os resultados apontaram diferenças estatísticas significativas entre as variáveis de F1 ( $F(3, 53) = 17,7; p < 0,001$ ) e F2 ( $F(3, 61) = 31,9; p < 0,001$ ) do exemplar vocálico [ε] na palavra *get* e das quatro outras palavras que fizeram uso do mesmo exemplar.

Adicionalmente, a literatura aponta que os exemplares consonantais em posição de *onset* ou *coda*, nas palavras apresentadas no Quadro 1, não influenciaram de forma significativa a realização de exemplares vocálicos do núcleo silábico, tanto no detalhe fonético relativo ao espectro (KEWLEY-PORT, 1982; HILLENBRAND; CLARK; NEAREY, 2001), quanto no detalhe relativo à duração (HOUSE, 1961; UMEDA, 1975).

Os itens lexicais foram produzidos em posição medial na frase-veículo *Say \_\_\_ again*. As frases foram apresentadas na forma de cartões individuais, com uma frase cada. Diversos distratores foram apresentados concomitantemente. Os informantes foram solicitados a lerem os cartões em velocidade e altura de voz normal. Após a primeira leitura de todos os cartões, solicitamos duas leituras adicionais. Cada item lexical foi realizado em três momentos distintos, o que gerou um total de 600 *tokens* efetivamente analisados. Palavras com o

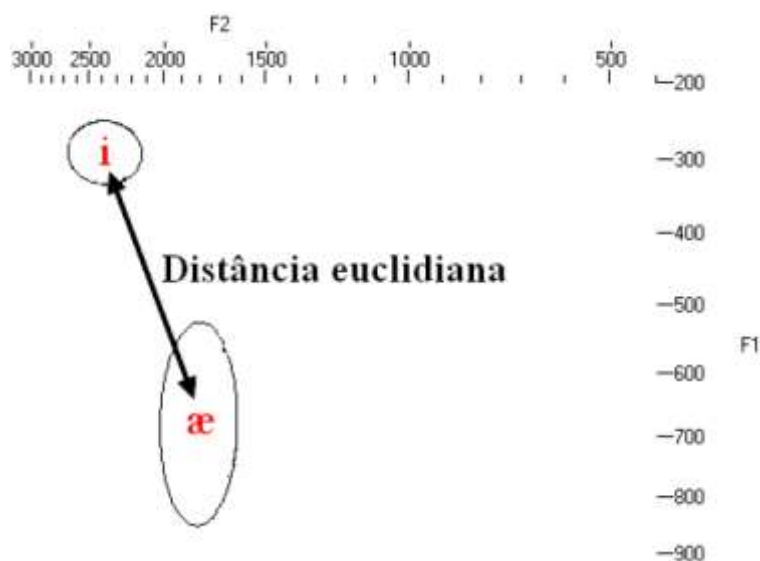
<sup>2</sup> Consulta realizada em janeiro/2013. Consultas posteriores podem apresentar resultados distintos, visto que o referido corpus é constantemente atualizado.



exemplar vocálico [eɪ] foram analisados em 3 variáveis espectrais (F1, F2, F3) e as palavras com os exemplares [i, ɪ, ε, æ] contaram com 4 variáveis (F1, F2, F3, duração). Assim, analisamos estatisticamente 2.280 variáveis quantitativas neste estudo.

As medições dos valores de F1, F2, F3 e da duração de cada exemplar vocálico foram realizadas conforme metodologia apresentada por Ladefoged (2003). Buscamos realizar as aferições dos valores dos formantes (em Hz) dos exemplares [i, ɪ, ε, æ], em um ponto em que os formantes estivessem particularmente estáveis, sem a influência de coarticulação das consoantes em posição de *onset* e *coda*. As medidas de duração vocálica foram facilitadas pelas consoantes oclusivas, sem a adição de eventuais VOTs nos valores de duração. Tendo em vista o ditongo no exemplar [eɪ], optamos por realizar as medições dos formantes num ponto posterior à coarticulação da consoante em posição de *onset* e anterior ao movimento do formante F2 em direção à segunda vogal do ditongo. Medidas de duração não foram realizadas para o exemplar [eɪ] devido ao fato deste som constituir ditongo.

Para a análise quantitativa dos dados apresentamos tabelas com as médias, medianas e desvios-padrão das diversas variáveis analisadas. Tecemos comentários acerca da estatística descritiva principalmente no que tange à proximidade das médias em variáveis semelhantes, bem como quanto à assimetria dos dados, revelada pela mediana. Comentários sobre a variação foram realizados levando em conta o desvio-padrão. Gráficos foram elaborados mostrando o ponto médio de realização de cada exemplar vocálico com relação aos valores de F1 e F2. Os *plots* vocálicos apresentam adicionalmente uma elipse, de um desvio-padrão envolvendo os pontos médios de realização das vogais, como na figura 1.



**Figura 1** – *Plot* (valores de F1 e F2) dos exemplares vocálicos [i] e [æ], realizados nas palavras *beat* e *back*, enfatizando a distância euclidiana.  
Fonte: autoria própria.

A distância euclidiana foi definida como a distância entre dois pontos passíveis de observação. Uma vez que obtivemos os pontos médios de realização dos exemplares vocálicos em Hz, calculamos a distância euclidiana de pontos bidimensionais por meio da fórmula:  $\sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$ . Para aplicar corretamente a equação, enfatizamos que  $p_x$  e  $p_y$  foram associados aos valores de F1 e F2 do primeiro exemplar analisado, enquanto  $q_x$  e  $q_y$  foram associados aos valores de F1 e F2 do segundo exemplar. Os valores foram

subtraídos, elevados ao quadrado e somados. Finalmente, obtivemos a raiz quadrada do valor resultante, que constituiu a distância euclidiana entre os exemplares vocálicos.

Os procedimentos estatísticos utilizados na análise de dados envolveram o cálculo do tamanho de efeito ( $d$ ) e a realização de diversos testes  $t$  pareados (DANCEY; REIDY, 2006). O primeiro procedimento retornou valores que refletiram desvios-padrão entre as variáveis, indicando o grau de sobreposição entre os dois grupos de dados. Quanto menor o valor  $d$  maior a sobreposição, e vice-versa. O teste  $t$  pareado foi usado para calcular a probabilidade de as diferenças entre as variáveis analisadas terem ocorrido devido ao erro amostral. Estabelecemos como ponto em que refutamos a hipótese nula  $p < 0,05$ .

Finalmente, utilizamos como *hardware* de coleta de dados um gravador digital *M-Audio Microtrack 24/96*, gravando diretamente em formato *wave* em frequência mono de 44.100 Hz. Utilizamos ainda um microfone *Shure SM58* unidirecional dinâmico, montado em um pedestal para evitar ruídos de manuseio. O software utilizado na análise dos dados acústicos foi o PRAAT (BOERSMA; WEENINK, 2012), enquanto na análise estatística utilizamos o SPSS (POLAR ENGINEERING AND CONSULTING, 2008).

Apresentamos a seguir seção envolvendo os resultados e a discussão.

### 3. Resultados e discussão

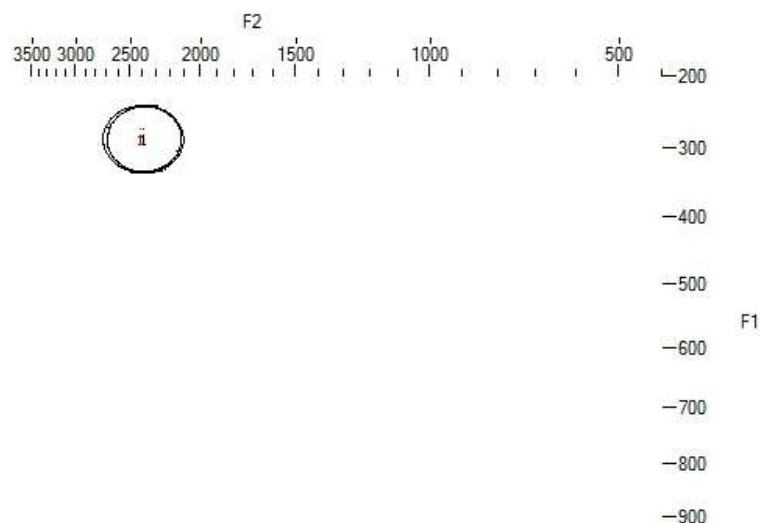
Iniciamos a seção de resultados e discussão apresentando a análise de palavras contendo os exemplares vocálicos anteriores [i, ɪ, eɪ, ε, æ] com relação aos valores de F1, F2 e F3, além da duração. Num segundo momento, fazemos a análise dos dados de distância euclidiana entre os referidos exemplares. Lembramos que todas as comparações envolveram exemplares vocálicos realizados em palavras de maior ou menor frequência de ocorrência, dada a natureza da variável dependente.

		F1	F2	F3	Dur.
<b>beat (af)</b>	<b>Média</b>	288	2389	3153	158
	<b>Mediana</b>	293	2399	3165	151
	<b>Desvio-padrão</b>	24	142	180	34
<b>teat (bf)</b>	<b>Média</b>	289	2398	3128	154
	<b>Mediana</b>	282	2445	3142	145
	<b>Desvio-padrão</b>	24	155	218	43

**Tabela 1** – Análise espectral e de duração dos exemplares vocálicos [i] nas palavras *beat* (alta frequência-af) e *teat* (baixa frequência-bf)

Observamos nos dados da tabela 1 que os valores médios realizados nas duas condições da variável dependente foram bastante aproximados, principalmente com relação às médias das variáveis, à exceção de F3. As medianas apresentaram comportamento diferenciado apenas no caso de F2, indicando pequena assimetria negativa da vogal da palavra *teat*. O desvio-padrão indicou uma maior dispersão dos dados de F3 e da duração dos exemplares vocálicos da palavra menos frequente.





**Figura 2** – Plot (valores de F1 e F2) dos exemplares vocálicos [i] realizados nas palavras *beat* (em preto) e *teat* (em vermelho).  
Fonte: autoria própria.

A figura 2 apresentou grande sobreposição espectral entre as realizações dos exemplares vocálicos [i] em palavras de alta e baixa frequência de ocorrência. O exemplar menos frequente apresentou uma pequena tendência por uma realização mais anterior.

Realizamos uma série de testes estatísticos (tamanho de efeito e teste *t* pareado) visando uma confirmação quantitativa das impressões iniciais. Observamos tamanhos de efeito irrelevantes (F1  $d = 0,01$ ) ou muito pequenos (F2  $d = 0,12$ ; F3  $d = 0,25$ ; DUR  $d = 0,24$ ). Os valores indicaram a grande sobreposição espectral entre os exemplares de maior e menor frequência de ocorrência, confirmando as observações anteriores.

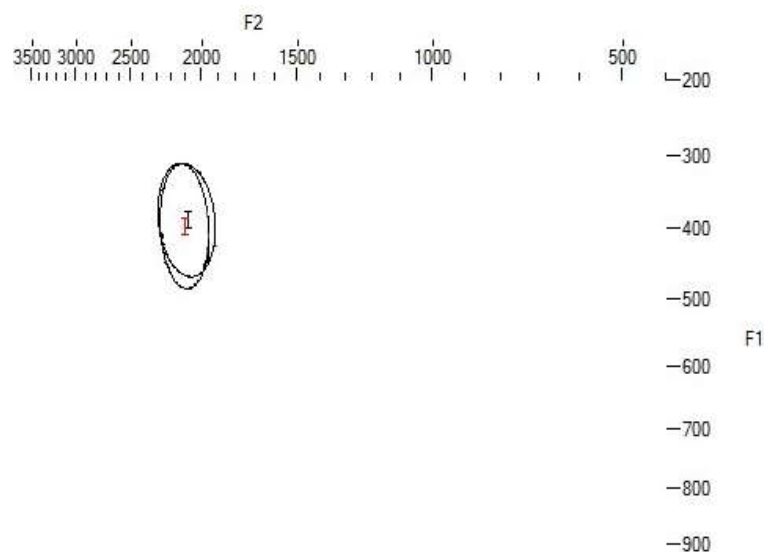
De modo semelhante, os resultados dos testes *t* pareados indicaram grande probabilidade de a variação entre os exemplares vocálicos decorrer de erro amostral. Observamos diferenças estatísticas não significativas quando comparamos os valores de F1 ( $t(19) = -0,02$ ;  $p = 0,981$ ), F2 ( $t(19) = -0,66$ ;  $p = 0,52$ ), F3 ( $t(19) = -0,74$ ;  $p = 0,47$ ) e da duração ( $t(19) = -0,82$ ;  $p = 0,42$ ) dos exemplares vocálicos [i] realizados em palavras de maior ou menor frequência de ocorrência. Concluímos que a frequência de ocorrência de itens lexicais foi irrelevante na realização do detalhe fonético pertinente aos exemplares vocálicos [i] do ILE dos informantes. Passamos à análise dos dados dos exemplares [i] a seguir.

		F1	F2	F3	Dur.
bit (af)	<b>Média</b>	388	2094	2670	140
	<b>Mediana</b>	381	2093	2665	135
	<b>Desvio-padrão</b>	40	97	139	41
tit (bf)	<b>Média</b>	398	2107	2692	139
	<b>Mediana</b>	403	2110	2704	138
	<b>Desvio-padrão</b>	45	81	134	35

**Tabela 2** – Análise espectral e de duração dos exemplares vocálicos [i] nas palavras *bit* (alta frequência) e *tit* (baixa frequência).

Observamos na tabela 2 que os valores médios foram aproximados nas variáveis analisadas, com pequena variação nos valores de F3. A mediana foi próxima à média na

maioria dos casos, indicando pequena assimetria. O desvio-padrão apresentou uma variação maior apenas no caso do F2, indicando uma maior dispersão no caso do exemplar realizado na palavra *bit*.



**Figura 3** – Plot (valores de F1 e F2) dos exemplares vocálicos [ɪ] realizado nas palavras *bit* (em preto) e *tit* (em vermelho).  
Fonte: autoria própria.

Visualizamos na figura 3 um alto grau de sobreposição vocálica entre os exemplares [ɪ] em palavras de alta e baixa frequência. O exemplar menos frequente apresentou leve tendência por uma realização mais anterior e baixa do que o exemplar mais frequente.

Os tamanhos do efeito entre os exemplares foram irrelevantes quanto à duração ( $d = 0,02$ ), mas apresentaram valores entre pequenos e médios nas variáveis espectrais (F2  $d = 0,29$ ; F3  $d = 0,32$ ; F1  $d = 0,43$ ). Os valores revelaram uma maior influência da variável dependente frequência de ocorrência no caso do exemplar [ɪ], quando comparado a [i], resultando em menor sobreposição espectral.

Os resultados dos testes  $t$  pareados indicaram grande probabilidade de as diferenças entre as condições mais e menos frequentes serem decorrentes de erro amostral nas variáveis de F1 ( $t(19) = -1,34$ ;  $p = 0,2$ ), F2 ( $t(19) = -0,57$ ;  $p = 0,58$ ), F3 ( $t(19) = -1,187$ ;  $p = 0,25$ ) e da duração ( $t(19) = -0,11$ ;  $p = 0,92$ ). Os testes retornaram diferenças estatísticas não significativas no caso dos exemplares vocálicos [ɪ] oriundos de palavras mais e menos frequentes. Concluímos que o fator frequência de ocorrência foi irrelevante na realização do detalhe fonético associado ao exemplar [ɪ] no estudo. Passamos a seguir à análise do primeiro elemento dos exemplares vocálicos [eɪ] deste estudo.

		F1	F2	F3
<b>date</b> <b>(bf)</b>	<b>Média</b>	413	1992	2723
	<b>Mediana</b>	408	1976	2710
	<b>Desvio-padrão</b>	23	237	86
<b>bate</b> <b>(bf)</b>	<b>Média</b>	427	1907	2548
	<b>Mediana</b>	426	1880	2537
	<b>Desvio-padrão</b>	30	108	110

**Tabela 3** – Análise espectral dos exemplares vocálicos [eɪ] nas palavras *date* (alta frequência-af) e *bate* (baixa frequência-bf).

A tabela 3 não apresentou dados de duração, devido ao fato de não terem sido realizadas medições da variável nestes exemplares. Diferentemente do observado nos dados relativos aos exemplares vocálicos [i, ɪ] em palavras mais ou menos frequentes, os primeiros elementos dos exemplares [eɪ] apresentaram médias distintas, especialmente dos valores de F2 e F3. Os valores das medianas indicaram leve assimetria positiva em ambas as condições de frequência. Os desvios-padrão apresentaram uma menor variação no caso do exemplar mais frequente, à exceção do valor de F2, que apontou grande variação.

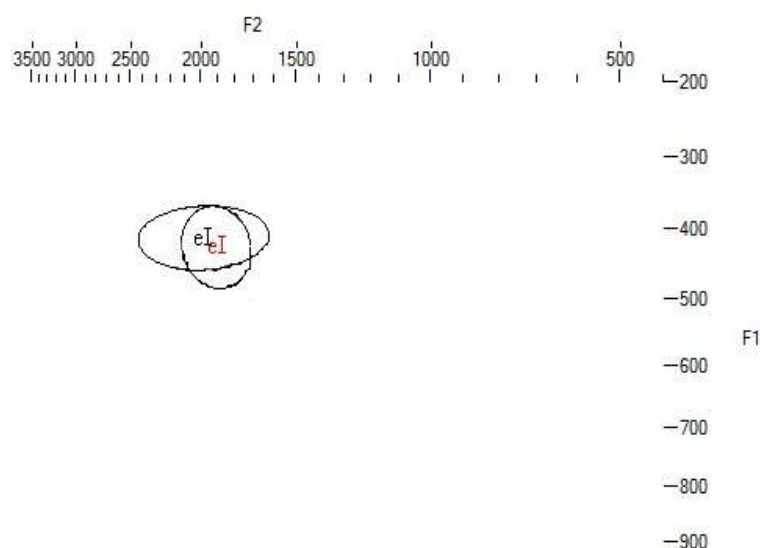


Figura 4 – *Plot* (valores de F1 e F2) dos exemplares vocálicos [eɪ] realizados nas palavras *date* (em preto) e *bate* (em vermelho).  
Fonte: autoria própria.

A observação da figura 4 confirmou a existência de uma grande variação no eixo F2 (anterior-posterior) do exemplar mais frequente. Em se tratando do eixo F1 (alto-baixo), observamos que ambos os exemplares apresentaram alturas semelhantes, apesar de o exemplar mais frequente ter sido realizado ligeiramente mais baixo. A sobreposição espectral entre os exemplares mais e menos frequente de [eɪ] foi menor que a realizada entre os exemplares [i, ɪ].

A análise estatística indicou que os tamanhos do efeito entre as variáveis do primeiro elemento do exemplar [eɪ], como antecipado pela observação do *plot* vocálico, foram grandes (F1  $d = 1$ ; F2  $d = 0,99$ ). Os dados indicaram cerca de 45% de sobreposição entre as variáveis.

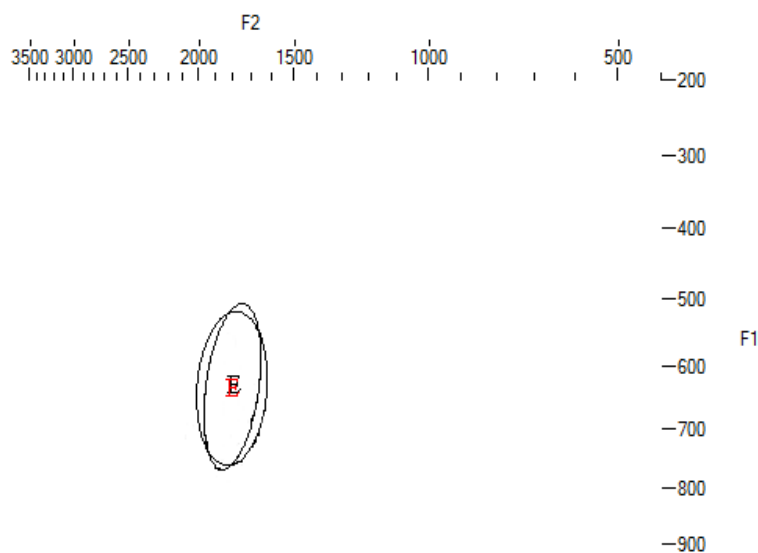
No caso de F3, o efeito foi muito maior ( $d = 3,58$ ), indicando um baixíssimo grau de sobreposição espectral. As medidas de efeito refletiram a pequena sobreposição espectral nos dados do exemplar [e].

Os testes  $t$  pareados indicaram pequena probabilidade de as diferenças entre as condições serem decorrentes de erro amostral no caso das variáveis F1 ( $t(19) = -2,13$ ;  $p = 0,05$ ) e F3 ( $t(19) = 10$ ;  $p < 0,001$ ), mas não na variável F2 ( $t(19) = 1,611$ ;  $p = 0,12$ ). Concluimos que o fator frequência de ocorrência foi relevante na realização do detalhe fonético do primeiro exemplar do som [e] no conjunto de dados. Apresentamos a seguir dados relativos aos exemplares vocálicos [e].

		F1	F2	F3	Dur.
<b>debt (af)</b>	<b>Média</b>	634	1803	2617	181
	<b>Mediana</b>	611	1805	2634	180
	<b>Desvio-padrão</b>	67	80	140	41
<b>Beck (bf)</b>	<b>Média</b>	637	1807	2605	167
	<b>Mediana</b>	627	1809	2610	166
	<b>Desvio-padrão</b>	62	101	112	36

**Tabela 4** – Análise espectral e de duração dos exemplares vocálicos [e] realizados nas palavras *debt* (alta frequência-af) e *Beck* (baixa frequência-bf).

As médias das duas condições na tabela 4 revelaram medidas espectrais bastante aproximadas. Dados de duração, todavia, apontaram maior duração do item de alta frequência. Quanto às medianas, encontramos apenas pequenas assimetrias positivas no F1 nos dois casos, além de leve assimetria negativa no F1 de *debt*. Os desvios-padrão apontaram maior variação no caso de F3 do conjunto de exemplares de maior frequência.



**Figura 5** – Plot (valores de F1 e F2) dos exemplares vocálicos [e] realizados nas palavras *debt* (em preto) e *Beck* (em vermelho).

Fonte: autoria própria.

Observamos na figura 5 o alto grau de sobreposição espectral nos eixos de F1 e F2. O ponto de articulação médio do exemplar menos frequente foi ligeiramente mais anterior e baixo do que do exemplar mais frequente.

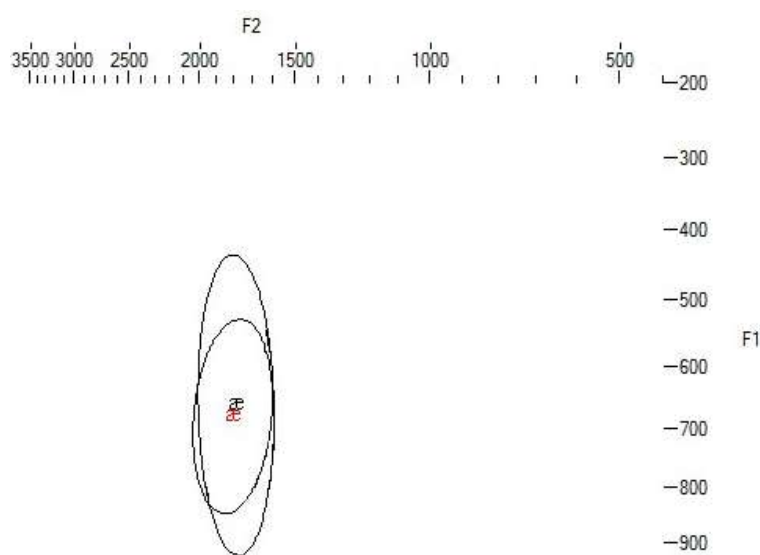
Os tamanhos de efeito entre as variáveis espectrais do exemplar [ɛ] foram pequenos (F1  $d = 0,08$ ; F2  $d = 0,09$ ; F3  $d = 0,19$ ). Os valores indicaram a existência de uma grande sobreposição espectral entre as variáveis. Por outro lado, o tamanho de efeito quando as variáveis de duração foram comparadas foi relativamente grande (DUR  $d = 0,72$ ). O valor indicou uma sobreposição da ordem de 57% entre as condições.

Os resultados dos testes  $t$  pareados indicaram grande possibilidade de as diferenças espectrais serem decorrentes de erro amostral nas variáveis de F1 ( $t(19) = -0,3$ ;  $p = 0,78$ ), F2 ( $t(19) = -0,25$ ;  $p = 0,81$ ) e F3 ( $t(19) = 0,62$ ;  $p = 0,54$ ). Os dados da duração ( $t(19) = 3,03$ ;  $p = 0,01$ ) apontaram diferença significativa, indicando pequena probabilidade de os resultados terem ocorrido devido ao erro amostral. Concluímos que apenas a duração apresentou indícios da influência da variável frequência de ocorrência no caso do exemplar [ɛ]. Passamos à análise dos dados dos exemplares vocálicos [æ] a seguir.

		F1	F2	F3	Dur.
<b>back (af)</b>	<b>Média</b>	666	1791	2607	185
	<b>Mediana</b>	681	1799	2603	181
	<b>Desvio-padrão</b>	117	104	128	39
<b>tat (bf)</b>	<b>Média</b>	683	1809	2602	190
	<b>Mediana</b>	674	1829	2580	189
	<b>Desvi-padrão</b>	81	115	216	37

**Tabela 5** – Análise espectral e de duração dos exemplares vocálicos [æ] realizados nas palavras *back* (alta frequência-af) e *tat* (baixa frequência-bf).

Observamos na tabela 5 que os valores médios, à exceção de F1, foram semelhantes em todos dos casos. As medianas indicaram leves assimetrias, positivas ou negativas, nos casos de F1 e F3. As diferenças entre os desvios-padrão chamaram a atenção no caso dos valores de F1 e F3, indicando uma maior variação.



**Figura 6** – Plot (valores de F1 e F2) dos exemplares vocálicos [æ] realizados nas palavras

*back* (em preto) e *tat* (em vermelho).  
Fonte: autoria própria.

Observamos na figura 6 o maior desvio-padrão do eixo F1 do exemplar mais frequente. Não obstante, observamos que os pontos médios de realização de ambos os exemplares foram aproximados, com o exemplar menos frequente tendendo a uma realização mais baixa e anterior que o exemplar mais frequente.

Os tamanhos do efeito foram relativamente pequenos nas variáveis relativas ao exemplar [æ] (F1  $d = 0,33$ ; F2  $d = 0,33$ ; F3  $d = 0,06$ ; DUR  $d = 0,27$ ). Os maiores efeitos ocorreram nas variáveis F1 e F2, mas indicaram um grau de sobreposição da ordem de 79%.

Os testes  $t$  pareados apontaram alta probabilidade de as diferenças entre as condições decorrerem de erro amostral em todas as variáveis analisadas (F1  $t(19) = -0,82$ ;  $p = 0,42$ ; F2  $t(19) = -0,93$ ;  $p = 0,37$ ; F3  $t(19) = 0,12$ ;  $p = 0,91$ ; DUR  $t(19) = -1,16$ ;  $p = 0,26$ ). Concluimos que existiram diferenças não significativas entre os exemplares [æ] realizados em palavras de maior ou menor frequência de ocorrência nos dados.

Encerramos a análise dos dados relativos ao detalhe fonético espectral e de duração dos exemplares vocálicos anteriores do ILE [i, ɪ, eɪ, ε, æ] realizados em palavras com maior e menor frequência de ocorrência. Retomamos brevemente os resultados, além das hipóteses pertinentes ao conjunto de dados.

Os exemplares vocálicos anteriores do ILE foram pouco influenciados pela variável frequência de ocorrência. O detalhe fonético espectral e de duração dos exemplares [i, ɪ, æ] apresentaram tamanho de efeito entre insignificante e médio, além de retornarem diferenças estatísticas não significativas nos diversos testes  $t$  pareados realizados.

Tamanhos de efeito mais robustos foram observados no caso dos exemplares [eɪ] e [ε]. O primeiro caso apresentou efeitos relevantes em todas as variáveis (F1  $d = 1$ ; F2  $d = 0,99$ ; F3  $d = 3,58$ ). No segundo, encontramos um efeito relevante na análise da variável duração ( $d = 0,72$ ).

Os resultados dos diversos testes  $t$  pareados confirmaram a impressão inicial de independência entre os exemplares de [eɪ] apenas no caso de F1 ( $t(19) = -2,13$ ;  $p = 0,05$ ) e F3 ( $t(19) = 10$ ;  $p < 0,001$ ), com diferenças estatísticas significativas entre as palavras de maior e menor frequência de ocorrência. Concluimos que o exemplar mais frequente foi realizado com um menor deslocamento no eixo alto-baixo.

No caso da duração de [ε], o teste  $t$  pareado retornou diferença estatística significativa entre as condições de maior e menor frequência de ocorrência ( $t(19) = 3,03$ ;  $p = 0,01$ ). O exemplar menos frequente tendeu a uma realização com maior duração do que o exemplar mais frequente.

Lembramos que as hipóteses apontavam:

- a) exemplares vocálicos em palavras de maior frequência de ocorrência do ILE são realizados num espaço acústico de maiores dimensões quando comparados aos exemplares vocálicos realizados em palavras de menor frequência de ocorrência;
- b) exemplares vocálicos em palavras de maior frequência de ocorrência são realizados com duração distinta dos exemplares vocálicos realizados em palavras de menor frequência.

A maioria dos testes estatísticos retornaram diferenças não significativas entre os exemplares associados a palavras de menor e maior frequência de ocorrência, desse modo, refutamos ambas as hipóteses secundárias supracitadas. Não encontramos evidências de que exemplares vocálicos anteriores de palavras de maior ou menor frequência de ocorrência



apresentaram detalhe fonético de realização distinto, no que tange aos valores de F1, F2, F3 e da duração.

Passamos à análise e discussão dos dados envolvendo as distâncias euclidianas do grupo de exemplares vocálicos supracitado. Lembramos que as distâncias foram obtidas a partir de uma equação envolvendo os valores de F1 e F2 dos exemplares analisados.

	<b>Distância Euclidiana <i>beat-bit (af)</i></b>	<b>Distância Euclidiana <i>teat-tit (bf)</i></b>
<b>Média</b>	316	318
<b>Mediana</b>	286	344
<b>Desvio-padrão</b>	131	140

**Tabela 6** – Análise da distância euclidiana entre os exemplares vocálicos [i, ɪ] realizados nas palavras *beat-bit* (alta frequência-af) e *teat-tit* (baixa frequência-bf).

Observamos na tabela 6 que os dados médios da distância euclidiana entre os exemplares vocálicos [i, ɪ] foram bastante aproximados em ambas as condições. Os dados da mediana apresentaram uma divergência mais marcante. A condição de alta frequência tendeu a uma assimetria positiva, enquanto a condição de baixa frequência tendeu a uma assimetria negativa. Os desvios-padrão de ambas as condições apontaram grande dispersão.

No caso da distância euclidiana entre os exemplares [i, ɪ] de palavras mais e menos frequentes, o tamanho de efeito foi insignificante ( $d = 0,02$ ), indicando uma sobreposição de quase 100% entre os grupos. Tendo em vista o tamanho de efeito, o teste *t* pareado retornou diferença não significativa ( $t(19) = -0,09$ ;  $p = 0,93$ ), indicando grande a possibilidade de as diferenças entre os grupos decorrerem de erro amostral. Concluímos que o fator frequência de ocorrência foi irrelevante para a variável distância euclidiana entre os exemplares vocálicos [i, ɪ] de palavras mais e menos frequentes. Passamos à análise da distância euclidiana entre os exemplares [i, eɪ].

	<b>Distância Euclidiana <i>beat-date (af)</i></b>	<b>Distância Euclidiana <i>teat-bate (bf)</i></b>
<b>Média</b>	461	513
<b>Mediana</b>	465	537
<b>Desvio-padrão</b>	116	133

**Tabela 7** - Análise da distância euclidiana entre os exemplares vocálicos [i, eɪ] realizados nas palavras *beat-date* (alta frequência-af) e *teat-bate* (baixa frequência-bf).

Observamos na tabela 7 marcante diferença entre as médias da condição mais e menos frequente. A distância euclidiana foi maior entre os exemplares vocálicos da condição menos frequente. As medianas indicaram assimetria negativa, com a condição menos frequente apresentando maior assimetria. Dados relativos aos desvios-padrão de ambos os grupos foram relativamente aproximados, com a condição menos frequente apresentando maior variação.

Os testes estatísticos indicaram um tamanho de efeito médio ( $d = 0,42$ ), apontando que cerca de 73% dos dados de ambos os grupos encontravam-se sobrepostos. O teste *t* pareado retornou uma diferença significativa entre as médias analisadas ( $t(19) = -2,09$ ;  $p = 0,05$ ), que indicou pequena probabilidade de a diferença ser decorrente de erro amostral. Concluímos que o fator frequência de ocorrência afetou de modo significativo a distância euclidiana entre os exemplares [i] e [eɪ] em palavras de maior e menor frequência de ocorrência no estudo. A

condição menos frequente apresentou maior distância euclidiana do que a condição mais frequente. Passamos à discussão dos dados da distância euclidiana entre as realizações dos exemplares [ɛ, æ].

	<b>Distância Euclidiana <i>debt-back (af)</i></b>	<b>Distância Euclidiana <i>Beck-tat (bf)</i></b>
<b>Média</b>	117	119
<b>Mediana</b>	102	89
<b>Desvio-padrão</b>	86	110

**Tabela 8** – Análise da distância euclidiana entre os exemplares vocálicos [ɛ, æ] realizados nas palavras *debt-back* (alta frequência-af) e *Beck-tat* (baixa frequência-bf).

Observamos na tabela 8 a proximidade entre as distâncias euclidianas médias de realização de ambas as condições. O mesmo não ocorreu com as medianas. Apesar de ambos os grupos tenderem a uma assimetria positiva, constatamos que os exemplares de baixa frequência apresentaram maior assimetria. Chamamos a atenção para os valores de desvio-padrão em ambas as condições, apesar de o grupo *Beck-tat* apresentar maior variação.

Os testes estatísticos revelaram um tamanho de efeito mínimo ( $d = 0,02$ ), indicando quase 100% de sobreposição entre os grupos. O teste *t* pareado retornou um valor não-significativo ( $t(19) = -0,1$ ;  $p = 0,92$ ). A diferença entre a distância euclidiana dos exemplares [ɛ, æ] em palavras de maior e menor frequência de ocorrência decorreu de erro amostral. Passamos à análise da distância euclidiana entre os exemplares [ei, æ].

	<b>Distância Euclidiana <i>date-back (af)</i></b>	<b>Distância Euclidiana <i>bate-tat (bf)</i></b>
<b>Média</b>	371	306
<b>Mediana</b>	330	295
<b>Desvio-padrão</b>	183	109

**Tabela 9** – Análise da distância euclidiana entre os exemplares vocálicos [ei, æ] realizados nas palavras *date-back* (alta frequência-af) e *bate-tat* (baixa frequência-bf).

Observamos na tabela 9 variação entre as médias das condições mais e menos frequente, com o primeiro grupo apresentando um valor mais elevado. Em se tratando das medianas, observamos algumas diferenças. Ambas as condições apresentaram assimetrias positivas, mas o grupo de maior frequência atingiu uma assimetria mais acentuada. O mesmo ocorreu com relação ao desvio-padrão, o grupo mais frequente apresentou uma variação mais marcante.

O tamanho de efeito retornou um valor de ( $d = 0,44$ ), considerado médio. O resultado indicou que cerca de 70% dos grupos encontravam-se sobrepostos. O teste *t* pareado apontou diferença estatística não-significativa entre as médias ( $t(19) = 1,56$ ;  $p = 0,14$ ), sendo grande a probabilidade de as diferenças serem decorrentes do erro amostral. Não podemos afirmar que a diferença entre as distâncias euclidianas dos exemplares [ei, æ] decorreu da frequência de

ocorrência. Finalmente, apresentamos dados referentes às distâncias euclidianas dos exemplares [i, æ].

	<b>Distância Euclidiana</b> <i>beat-back (af)</i>	<b>Distância Euclidiana</b> <i>teat-tat (bf)</i>
<b>Média</b>	719	718
<b>Mediana</b>	684	703
<b>Desvio-padrão</b>	155	159

**Tabela 10** – Análise da distância euclidiana entre os exemplares vocálicos [i, æ] realizados nas palavras *beat-back* (alta frequência-af) e *teat-tat* (baixa frequência-bf).

Os dados da tabela 10 indicaram grande proximidade dos valores médios da distância euclidiana. As medianas indicaram leve assimetria positiva. Os desvios-padrão indicaram variação semelhante em ambos os grupos.

Com relação à estatística, o tamanho de efeito foi insignificante ( $d = 0,01$ ), indicando sobreposição marcante dos grupos analisados. O teste  $t$  pareado ( $t(19) = -0,05$ ;  $p = 0,96$ ) indicou grande possibilidade de a diferença entre os grupos decorrer do erro amostral. A diferença entre as distâncias euclidianas das condições mais e menos frequente nos dados foi não significativa.

Tendo em vista os dados apresentados anteriormente, relativos às distâncias euclidianas entre pares de exemplares vocálicos de palavras mais ou menos frequente, foi pequena a influência da variável dependente frequência de ocorrência. Apenas no caso do par [i, eɪ] encontramos uma diferença estatística significativa entre os grupos analisados ( $t(19) = -2,09$ ;  $p = 0,05$ ). No caso, o par menos frequente apresentou maior distância euclidiana. Nenhum outro par retornou diferenças significativas, apesar de o par [eɪ, æ] ter atingido um tamanho de efeito considerado médio ( $d = 0,44$ ).

Tendo em vista os diversos indícios da irrelevância do fator frequência de ocorrência nos dados anteriores, optamos por refutar a terceira hipótese secundária que afirmava que exemplares vocálicos em palavras de maior frequência de ocorrência apresentam maior distância euclidiana quando comparados aos exemplares realizados em palavras de menor frequência de ocorrência.

### **Conclusão**

Objetivamos neste estudo analisar a realização do detalhe fonético dos exemplares vocálicos anteriores [i, ɪ, eɪ, ε, æ] do ILE. Constatamos diferenças estatísticas não significativas em grande parte das variáveis de F1, F2, F3, duração e distância euclidiana analisadas. A variável dependente frequência de ocorrência dos itens lexicais atingiu níveis de significância estatística apenas no caso de F1 ( $d = 1$ ;  $t(19) = -2,13$ ;  $p = 0,047$ ) e F3 ( $d = 3,58$ ;  $t(19) = 10$ ;  $p < 0,001$ ) dos exemplares [eɪ], da duração ( $d = 0,72$ ;  $t(19) = 3,03$ ;  $p = 0,01$ ) dos exemplares [ε], e da distância euclidiana ( $d = 0,42$ ;  $t(19) = 2,09$ ;  $p = 0,05$ ) entre os exemplares [i, eɪ].

Consideramos, portanto, refutada a hipótese básica, de que os exemplares vocálicos anteriores em palavras do ILE de maior frequência de ocorrência apresentam detalhe fonético distinto dos exemplares vocálicos realizados em palavras de menor frequência. A conclusão vai ao encontro do estudo desenvolvido por Barboza (2013), que apontou a irrelevância do fator frequência de ocorrência de itens lexicais na aquisição do ILE por aprendizes brasileiros. Naquele estudo, indícios apontaram a relevância da variável frequência de tipo na construção

da fonologia do ILE nos nossos aprendizes. Este pode ser um caminho a ser seguido na busca pelo papel da frequência de uso na aquisição de línguas estrangeiras.

Adicionalmente, algumas limitações podem ter contribuído para refutarmos a hipótese básica deste trabalho: a) o alto nível de proficiência dos informantes, profissionais do ensino de ILE, pode ter influenciado no resultado; b) o tipo de experimento utilizado no estudo, que consistiu na leitura de palavras inseridas em frases-veículo, pode ter sido um fator homogeneizador dos exemplares vocálicos analisados; c) o corte utilizado nesta pesquisa, que optou por trabalhar apenas com exemplares vocálicos anteriores, pode ter encoberto possíveis associações entre a frequência de ocorrência e o detalhe fonético/distância euclidiana entre exemplares não analisados neste estudo. As limitações anteriores constituem sugestões para a realização de futuras pesquisas.

Em suma, apesar de termos detectado eventuais efeitos da frequência de ocorrência na realização de exemplares vocálicos do ILE, os dados indicaram pequena influência da variável dependente na realização do detalhe fonético/distância euclidiana. Estudos adicionais se fazem necessários para aprofundar o conhecimento acerca da variável na produção de aprendizes brasileiros de ILE.

### **Lexical token frequency and the realization of front vowels in English as a foreign language**

**ABSTRACT:** The influence of token frequency of lexical items in the realization of front vowels in English as a foreign language (EFL) was analyzed. Usage-based phonological theories consider token frequency a relevant factor for foreign languages acquisition. We carried out an acoustical analysis of F1, F2, F3 and Euclidian distance of [i, ɪ, eɪ, e, æ] vowels as realized by Brazilian EFL teachers. Data revealed limited token frequency effects, indicating the variable was irrelevant for informants' construction of EFL phonology.

Keywords: usage-based phonology; exemplar model; lexical item.

### **REFERÊNCIAS**

BARBOZA, C. L. F. *Descrição acústica dos sons vocálicos anteriores do inglês e do português realizados por professores de inglês língua estrangeira no oeste Potiguar*. 2008. 182 f. Dissertação (Mestrado em Linguística Aplicada) - Universidade Estadual do Ceará - UECE, Fortaleza, 2008.

BARBOZA, C. L. F. *Efeitos da palatalização das oclusivas alveolares do português brasileiro no percurso de construção da fonologia da inglês língua estrangeira*. 2013. 263f. Tese (Doutorado em Linguística) – Universidade Federal do Ceará – UFC, Fortaleza, 2013.

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M. H.; CROFT, W.; ELLIS, N.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. Language is a complex adaptive system: position paper. *Language Learning*, Michigan, v. 51, n. 1, p.1-26, dec. 2009.

BOERSMA, P.; WEENIK, D. *Praat: doing phonetics by computer*. Version 5.1.43. Disponível em: <<http://www.praat.org>. 2012.>

BYBEE, J. *Phonology and language use*. Cambridge: Cambridge University Press, 2001.

BYBEE, J. Mechanisms of change in grammaticalization: the role of frequency. In: JANDA, R.D.; JOSEPH, B.D. *Handbook of historical linguistics*. Oxford: Blackwell, 2003. p. 602-623.

BYBEE, J. Usage-based grammar and second language acquisition. In: ROBINSON, P.; ELLIS, N. C. *Handbook of cognitive linguistics and second language acquisition*. New York: Routledge, 2008. p. 216-236.

BUSH, N. Frequency effects and word-boundary palatalization in English. In: BYBEE, J.; HOPPER, P. (Comp.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001. p. 255-280.

CLOPPER, C.; PIERREHUMBERT, J. Effects of semantic predictability and regional dialect on vowel space reduction. *Journal of the Acoustical Society of America*, v. 124, n. 3, p. 1682-1688, 2008.

DANCEY, C. P.; REIDY, J. *Estatística sem matemática para psicologia: usando SPSS para Windows*. Porto Alegre: Artmed, 2006.

ELLIS, N. C. Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, v. 24, n. 2, p.143-188, 2002.

ELLIS, N. C. The dynamics of second language emergence: cycles of language use, language change, and language acquisition. *The Modern Language Journal*, v. II , n. 92, p.232-249, 2008.

FERREIRA, A. P. P. *Pet or petty? Diferenças entre palavras CVC e CVCV do inglês por aprendizes brasileiros: uma análise acústica*. 2007. 131 f. Dissertação (Mestrado em Letras) – UFPR. Curitiba, 2007.

FERREIRA, A. P. P. Aquisição de codas em inglês como L2: uma análise à luz da Fonologia de Uso. In: CONGRESSO INTERNACIONAL DA ABRALIN, 7. *Anais...* Curitiba: Abralin, 2011. p. 267 - 275. Disponível em: <[http://www.abralin.org/abralin11\\_cdrom/](http://www.abralin.org/abralin11_cdrom/)>. Acesso em: 01 abr. 2011.

FAULKES, P.; DOCHERTY, G.. The social life of phonetics and phonology. *Journal of Phonetics*. v. 34, p. 409-438, 2006.

GUIMARÃES, D. M. L. O. *Sequências de (sibilante+africada alveolopalatal) no português falado em Belo Horizonte*. 2004. 122 f. Dissertação (Mestrado em Linguística) - UFMG, Belo Horizonte, 2004.

HILLENBRAND, J. M.; CLARK, M. J.; NEARY, T. M. Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America*, v. 109, n. 2, p. 748-763, 2001.

HOUSE, A. On vowel duration in English. *Journal of the Acoustical Society of America*, v. 33, n. 9, p. 1174-1178, 1961.

JOHNSON, K. Speech perception without speech normalization. In: JOHNSON, K.; MULLENIX, J. W. (Ed.). *Talker variability in speech perception*. San Diego: Academic Press, 1997. p. 145-165.

JOHNSON, K. Decisions and mechanisms in exemplar-based phonology. In: SOLE, M. J.; BEDDOR, P.; OHALA, M. (Ed.). *Experimental approaches to phonology in honor of John Ohala*. Oxford: Oxford University Press, 2007. p. 25-40.

KEWLEY-PORT, D. Measurement of formant transitions in naturally produced stop consonant-vowel syllables. *Journal of the Acoustical Society of America*, v. 72, n. 2, p. 379-1389, 1982.

LADEFOGED, P. *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Malden: Blackwell, 2003.

LEITE, C. T. *Seqüências de (oclusiva alveolar + sibilante alveolar) como um padrão inovador no português de Belo Horizonte*. 2006. 125 f. Dissertação (Mestrado em Linguística) - UFMG, Belo Horizonte, 2006.

PIERREHUMBERT, J. B. What people know about sounds of language. *Studies in the linguistic sciences*. v. 29, n. 2, p. 111-120, 1999.

PIERREHUMBERT, J. B. Exemplar dynamics: word frequency, lenition and contrast. In: BYBEE, J.; HOPPER, P. (Comp.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001. p. 137-158.

PIERREHUMBERT, J. B. Word-specific phonetics. In: GUSSENHOFEN, C.; WARNER, N. (Ed.). *Laboratory phonology 7*. Berlin: Mouton de Gruyter, 2002a. p. 101-139.

PIERREHUMBERT, J. B. Probabilistic phonology: discrimination and robustness. In: BOD, R.; HAY, J.; HOPPER, P. *Probability theory in linguistics*. Cambridge, MA: The MIT Press, 2002b. p. 175-228.

PIERREHUMBERT, J. B. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech*. v. 46, n. 2-3, p. 115-154, 2003.

PIERREHUMBERT, J. B. The next toolkit. *Journal of phonetics*. v. 34, p. 516-530, 2006.



POLAR ENGINEERING AND CONSULTING. *SPSS statistics*. Version 17.0. [S.l.]: Polar engineering and consulting, 2008.

UMEDA, N. Vowel duration in American English. *Journal of the Acoustical Society of America*, v. 58, n. 2, p. 434-445, 1975.

Data de envio: 07/04/2014

Data de aceite: 03/02/2015

Data de publicação: 23/04/2015