

ANDRÉ SCHNEIDER

**A EPÊNTESE MEDIAL EM PB E NA AQUISIÇÃO DE
INGLÊS COMO LE: UMA ANÁLISE
MORFOFONOLÓGICA**

**PORTO ALEGRE
2009**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ÁREA: ESTUDOS DA LINGUAGEM
ESPECIALIDADE: TEORIA E ANÁLISE LINGUÍSTICA
LINHA DE PESQUISA: FONOLOGIA E MORFOLOGIA**

**A EPÊNTESE MEDIAL EM PB E NA AQUISIÇÃO DE
INGLÊS COMO LE: UMA ANÁLISE
MORFOFONOLÓGICA**

ANDRÉ SCHNEIDER

ORIENTADOR: PROF. DR. LUIZ CARLOS DA SILVA SCHWINDT

Dissertação de Mestrado em Teoria e Análise Linguística, apresentada como requisito parcial para a obtenção do título de Mestre pelo Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul.

**PORTO ALEGRE
2009**

AGRADECIMENTOS

Dedico este trabalho a todas as pessoas que depositaram em mim a confiança necessária para que, após dias e noites incansáveis de trabalho e de estudo, eu pudesse concluir com êxito essa tarefa. Destaco, dentre elas:

- meus pais, Luiz e Ordete, meu irmão, Cristiano, Janaína e demais familiares, que acompanharam a distância minhas angústias na elaboração da dissertação;

- meu orientador e amigo, Luiz Carlos, pessoa de inigualável conhecimento linguístico com quem tive a honra de trabalhar desde o primeiro ano de meu curso de graduação e com quem aprendi, entre outros tantos ensinamentos proferidos durante o horário do cafezinho, que uma coisa é uma coisa e que outra coisa é outra coisa;

- Alines (Grodt, Lorandi e Padilha), Emanuel, Evelyne, Gabriela, Laura, Simone, Taís e Taíse; Gisela, Laura, Valéria e demais membros do Círculo Linguístico, que me auxiliaram no esclarecimento das muitas dúvidas que surgiram durante o desenvolvimento de minha pesquisa;

- Berna, Cíntia, Clarice, Dienstmann, Fabi, Ísis, Karin, Lis, Lu, Nodari, Paulo, Rafa, Zé e demais integrantes do grupo dos ‘mesmos’, amigos especiais do Coral da UFRGS que, com muita diversão, aliviaram o estresse de minha rotina de trabalho;

- minhas colegas de trabalho, Mara, Margarida e Nice, que compreenderam os momentos em que ‘trabalhar’ e ‘dissertar’ deixaram de ter uma relação de distribuição complementar e entraram em variação livre.

Registro também meus agradecimentos:

- à coordenação do setor de inglês da escola de idiomas onde recebi autorização para proceder à coleta de grande parte dos dados que compõem esta pesquisa;

- aos professores de inglês que concederam seu precioso tempo de aula e consentiram a participação de alguns de seus alunos na coleta de dados;

- a todos os 56 participantes submetidos à coleta de dados, que aceitaram contribuir voluntariamente para o desenvolvimento desta pesquisa.

Agradeço, ainda, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pela bolsa de estudos concedida entre março de 2007 e julho de 2008 e, finalmente, à Universidade Federal do Rio Grande do Sul, meu segundo lar, que me acolheu, em 2002, como aluno de graduação; em 2006, como aluno de segunda graduação; em 2007, como aluno de mestrado acadêmico; e em 2008, como servidor técnico-administrativo. Que venham os próximos acolhimentos!

RESUMO

No presente estudo, fazemos uma análise quantitativa do fenômeno fonológico de epêntese vocálica medial, que é atestado na fala do português brasileiro (PB) (por exemplo, [i] em *mag[i]ma*) e do inglês aprendido por brasileiros (por exemplo, [i] em *mag[i]net*). As amostras de fala utilizadas para essa análise foram obtidas a partir da gravação de frases lidas por 16 porto-alegrenses falantes de inglês como língua estrangeira. Partimos do pressuposto de que a produção de epêntese na fala do inglês seja resultado da transferência linguística que ocorre do PB para a interlíngua dos aprendizes. COLLISCHONN (2002) levantou a suspeita de que o número de ocorrências de epêntese após prefixos como *sub-* seria maior do que o número de ocorrências de epêntese em encontros consonantais localizados no interior de vocábulos que não são prefixados. Na tentativa de averiguar essa suspeita, incluímos em nossa análise a verificação do papel do contexto morfológico na aplicação da epêntese. Partimos da hipótese de que os prefixos que podem favorecer a epêntese em PB e em inglês têm diferentes graus de transparência, sendo que os prefixos mais transparentes parecem estar mais relacionados à aplicação de epêntese, em razão de uma possível independência fonológica que eles possam ter. Constatamos que, em PB, os prefixos *sob-* e *sub-* são os mais transparentes, ao passo que, em inglês, *arch-*, *out-*, *post-* e *sub-* são os mais transparentes. Nos resultados das análises estatísticas que realizamos com os dados de cada língua, vimos, no entanto, que o papel morfológico só tem maior relevância estatística nos dados do PB, nos quais foi possível confirmar nossa hipótese de que a epêntese seria mais frequente após *sob-* e *sub-*. Os resultados das análises estatísticas nos permitiram ainda constatar que todos os grupos de fatores selecionados nas rodadas com dados do inglês ('tipo de consoante perdida', 'vozeamento da consoante perdida', 'acento' e 'informante') haviam sido igualmente selecionados nas rodadas com dados do PB. A partir dessa constatação, realizamos algumas comparações entre os resultados obtidos para os grupos de fatores selecionados nas duas línguas, o que nos permitiu verificar que, para os dados de nosso estudo, existe uma relação estreita entre os fatores muito/pouco relacionados à epêntese em português e os fatores muito/pouco relacionados à epêntese em inglês. Isso parece confirmar a atuação da transferência linguística que assumimos existir entre a L1 e a interlíngua.

ABSTRACT

In this research, we aim at analyzing the recurrence of the epenthetic vowel found in word-medial clusters both in Brazilian Portuguese (BP) (e.g. [i] in *mag[i]ma*) and in English learnt by Brazilians (e.g. [i] in *mag[i]net*). The samples of oral production that we analyzed consist of a set of recorded sentences that 16 Brazilian learners of English (from Porto Alegre/RS) were asked to read. We assume that epenthesis found in English is a consequence of language transfer from BP to the learners' interlanguage. COLLISCHONN (2002) argued that the frequency of epenthesis found after prefixes such as *sub-* seemed to be higher than the frequency of epenthesis found in medial-word *clusters* of nonprefixed words. In order to investigate such statement, we included in our study the analysis of the role of morphological aspects in the production of epenthetic vowels. Firstly, we assumed that the prefixes that may favor vowel epenthesis in BP and in English are given different gradients of transparency, and we also assumed that the more a prefix was transparent, the more likely it would motivate epenthesis, since transparency may mean some phonological independence for these prefixes. We confirmed through data collection that *sob-* and *sub-* are the most transparent prefixes in BP, and that *arch-*, *out-*, *post-* and *sub-* are the most transparent ones in English. The statistical results for the analysis of oral production however showed us that only BP prefixes *sob-* and *sub-* were significant to vowel epenthesis. These statistical results also showed us that all the groups of factors considered to be significant to vowel epenthesis in English ('type of stray consonant', 'voicing of the stray consonant', 'stress', and 'participant') were significant in vowel epenthesis in BP, too. Such results motivated us to compare the frequency of epenthesis in the two languages for each of the factors comprised by these four groups considered significant to vowel epenthesis. The results of this comparison showed us that there is a straightforward relation between the factors that motivate vowel epenthesis in BP and the factors that motivate vowel epenthesis in English. Such results seem to confirm that language transfer is operative between L1 and the interlanguage.

LISTA DE QUADROS

Quadro 1 - Vocábulos integrantes da amostra de fala em PB	64
Quadro 2 - Células sociais da amostra de fala em PB e da amostra de fala em inglês	65
Quadro 3 - Vocábulos integrantes da amostra de fala em inglês	67
Quadro 4 - Distribuição das variáveis por rodada e por análise realizada	89
Quadro 5 - Resultado do teste de transparência de prefixos em PB	91
Quadro 6 - Resultado do teste de transparência de prefixos em inglês	98
Quadro 7 - Grupos analisados e grupos selecionados por rodada (PB)	104
Quadro 8 - Grupos analisados e grupos selecionados por rodada (inglês)	125
Quadro 9 - Grupos selecionados por rodada em PB e em inglês	133

LISTA DE GRÁFICOS

Gráfico 1 - Isolamento de prefixos em vocábulos não prefixados do PB	93
Gráfico 2 - Transparência de <i>sob-</i> e <i>sub-</i> em PB	94
Gráfico 3 - Transparência de <i>ab-</i> , <i>ad-</i> , <i>ec-</i> e <i>ob-</i> em PB	95
Gráfico 4 - Isolamento de prefixos em vocábulos não prefixados do inglês	99
Gráfico 5 - Transparência de <i>arch-</i> , <i>out-</i> , <i>post-</i> e <i>sub-</i> em inglês	100
Gráfico 6 - Percentual geral de epêntese medial em PB	102
Gráfico 7 - Frequência de epêntese em PB por contexto morfológico	117
Gráfico 8 - Percentual geral de epêntese medial em inglês	122
Gráfico 9 - Frequência de epêntese em inglês por contexto morfológico	123
Gráfico 10 - Frequência de epêntese em PB e em inglês por Tipo de consoante perdida	133
Gráfico 11 - Frequência de epêntese em PB e em inglês por Informante	134

LISTA DE TABELAS

Tabela 1 - Prefixos relacionados à epêntese vocálica em PB	56
Tabela 2 - Vocábulo prefixado submetido à análise de transparência em PB	58
Tabela 3 - Vocábulo com grafia inicial idêntica a prefixos do PB	58
Tabela 4 - Prefixos relacionados à epêntese vocálica em inglês	60
Tabela 5 - Vocábulo prefixado submetido à análise de transparência em inglês	61
Tabela 6 - Vocábulo com grafia inicial idêntica a prefixos do inglês	61
Tabela 7 - Tipos de consoantes perdidas analisados	73
Tabela 8 - Tipos de contextos seguintes analisados	75
Tabela 9 - <i>Clusters</i> relacionados à epêntese vocálica em PB inexistentes em inglês	78
Tabela 10 - <i>Clusters</i> relacionados à epêntese vocálica em inglês inexistentes em PB	79
Tabela 11 - <i>Clusters</i> relacionados à epêntese vocálica existentes em PB e em inglês	80
Tabela 12 - Resultado para Vozeamento da consoante perdida (PB)	105
Tabela 13 - Resultado para Tipo de consoante perdida (PB)	107
Tabela 14 - Resultado de amálgama para Tipo de consoante perdida (PB)	109
Tabela 15 - Resultado para Vozeamento do contexto seguinte (PB)	110
Tabela 16 - Resultado para Tipo de contexto seguinte (PB)	111
Tabela 17 - Resultado para Informante (PB)	115
Tabela 18 - Resultado para Contexto morfológico (PB)	117
Tabela 19 - Resultado para Tipo de <i>cluster</i> (PB)	119
Tabela 20 - Resultado para Acento (PB)	120
Tabela 21 - Resultado para Vozeamento da consoante perdida (inglês)	126
Tabela 22 - Resultado para Tipo de consoante perdida (inglês)	127
Tabela 23 - Resultado para Acento (inglês)	129
Tabela 24 - Resultado para Informante (inglês)	131

SUMÁRIO

INTRODUÇÃO	12
1 REVISÃO DA LITERATURA	14
1.1 Pressupostos sobre aquisição/aprendizagem de L2/LE e sobre variação	14
1.2 Pressupostos sobre morfologia	17
1.3 Pressupostos sobre fonologia: a sílaba	20
1.3.1 Teorias gerais da sílaba	21
1.3.1.1 Kahn (1976)	22
1.3.1.2 Selkirk (1982)	23
1.3.1.3 Clements e Keyser (1983)	25
1.3.1.4 Itô (1986)	28
1.3.2 A sílaba em português brasileiro (PB)	29
1.3.3 A sílaba em inglês	33
1.3.4 A epêntese como estratégia de reparo à malformação da sílaba	36
1.3.4.1 Análises sobre o fenômeno variável de epêntese	40
1.3.4.1.1 A epêntese em PB (Collischonn, 2002)	41
1.3.4.1.2 A epêntese na aquisição do inglês como LE (Pereyron, 2008)	45
2 METODOLOGIA	54
2.1 Instrumentos para a coleta de dados	55
2.1.1 Teste de verificação de transparência dos prefixos em PB	55

2.1.2	Teste de verificação de transparência dos prefixos em inglês	59
2.1.3	Amostra de Fala 1 (PB)	62
2.1.4	Amostra de Fala 2 (inglês)	65
2.2	Variáveis submetidas à investigação	68
2.2.1	Variável dependente	68
2.2.2	Variáveis independentes	70
2.2.2.1	Variáveis independentes linguísticas	70
2.2.2.1.1	Contexto morfológico	70
2.2.2.1.2	Tipo de consoante perdida	72
2.2.2.1.3	Vozeamento da consoante perdida	73
2.2.2.1.4	Tipo de contexto seguinte	74
2.2.2.1.5	Vozeamento do contexto seguinte	76
2.2.2.1.6	Acento	76
2.2.2.1.7	Tipo de <i>cluster</i>	77
2.2.2.2	Variáveis independentes extralinguísticas	82
2.2.2.2.1	Sexo	82
2.2.2.2.2	Idade	83
2.2.2.2.3	Informante	83
2.2.2.2.4	Proficiência em inglês	84
2.3	Instrumentos para a interpretação dos dados coletados	84
3	ANÁLISE E DISCUSSÃO DOS RESULTADOS	90
3.1	Teste de verificação de transparência dos prefixos em PB	90
3.2	Teste de verificação de transparência dos prefixos em inglês	98
3.3	Amostra de Fala 1 (PB)	101
3.3.1	Percentual geral de epêntese medial em PB	102
3.3.2	Resultados das análises estatísticas para dados do PB	103
3.3.2.1	Vozeamento da consoante perdida e Tipo de consoante perdida	105
3.3.2.2	Vozeamento do contexto seguinte e Tipo de contexto seguinte	110
3.3.2.3	Informante	114
3.3.2.4	Contexto morfológico e tipo de <i>cluster</i>	116
3.3.2.5	Acento	120

3.4 Amostra de Fala 2 (inglês)	121
3.4.1 Percentual geral de epêntese medial em inglês	122
3.4.2 Resultados das análises estatísticas para dados do inglês	124
3.4.2.1 Vozeamento da consoante perdida e Tipo de consoante perdida	126
3.4.2.2 Acento	129
3.4.2.3 Informante	130
3.5 Comparações entre a Amostra de Fala 1 e a Amostra de Fala 2	132
4 CONCLUSÕES	136
REFERÊNCIAS BIBLIOGRÁFICAS	141
ANEXOS	147

INTRODUÇÃO

O interesse pelo desenvolvimento da presente pesquisa surgiu durante o segundo semestre do ano de 2006, quando este pesquisador iniciou suas atividades na docência da língua inglesa para aprendizes brasileiros. A experiência adquirida durante as aulas ministradas permitiu ao pesquisador constatar que muitos dos ‘erros’ de fala produzidos pelos aprendizes eram decorrentes da aplicação de fenômenos linguísticos existentes na língua portuguesa, (i.é, sua língua materna). Um desses fenômenos é a epêntese vocálica medial, que corresponde ao foco de análise do presente estudo.

Resumidamente, a epêntese vocálica medial consiste na inserção de uma vogal entre duas consoantes presentes no interior de um vocábulo. Exemplos de aplicação de epêntese vocálica medial são atestados nos vocábulos *mag[i]ma* (PB) e *mag[i]net* (‘ímã’ em inglês). COLLISCHONN (2002) realizou uma análise quantitativa sobre o fenômeno variável da epêntese vocálica medial em PB. Em seu estudo, a autora (op. cit.) desconsiderou os casos de aplicação de epêntese após o prefixo *sub-* por suspeitar que esse prefixo poderia estar relacionado a uma frequência elevada de aplicação de epêntese (em virtude de uma aparente independência fonológica que esse prefixo tem), o que poderia obscurecer os resultados finais do estudo. Em razão da suspeita levantada por COLLISCHONN (2002), decidimos averiguar se, de fato, vocábulos com prefixos (como *sub-*) podem estar relacionados com frequências mais elevadas de aplicação de epêntese do que vocábulos sem prefixos.

Nosso estudo tem como principais objetivos:

(i) verificar o grau de transparência dos prefixos que podem motivar epêntese vocálica em PB e em inglês;

(ii) investigar o papel do contexto morfológico e de outros fatores linguísticos e extralinguísticos na aplicação da epêntese medial em PB e em inglês; e

(iii) comparar a aplicação de epêntese medial em PB com a epêntese na aquisição de inglês como língua estrangeira, pressupondo-se a atuação de transferência linguística entre L1 e interlíngua.

O presente trabalho está subdividido em quatro capítulos. No primeiro capítulo, Revisão da Literatura, apresentamos os principais aspectos teóricos que norteiam nossa análise. Tais aspectos teóricos tratam: (i) da ideia de que, no processo de aquisição de uma segunda língua, ocorra transferência linguística da língua materna para a interlíngua de um aprendiz, o que permite associar a variação da epêntese que ocorre em PB com a variação da epêntese que ocorre em inglês como LE (cf. 1.1); (ii) da hipótese de que, sob o ponto de vista morfológico, alguns prefixos podem ser mais transparentes do que outros prefixos, o que pode também significar que prefixos mais transparentes podem apresentar configuração prosódica diferente da configuração prosódica de prefixos menos transparentes (cf. 1.2); e (iii) da ideia de que a epêntese vocálica esteja diretamente relacionada com o constituinte prosódico 'sílabas', uma vez que tal fenômeno é entendido por nós como uma estratégia de reparo a sílabas malformadas (cf. 1.3).

No segundo capítulo, apresentamos a metodologia que adotamos para investigar (i) a transparência de prefixos que podem motivar a epêntese em PB e em inglês e (ii) os fatores linguísticos e extralinguísticos que podem estar relacionados à aplicação de epêntese em dados de fala dessas duas línguas.

O terceiro capítulo contempla a apresentação e a discussão dos resultados obtidos para cada uma das investigações que propusemos no capítulo sobre metodologia. No capítulo 3, apresentamos, portanto, os graus de transparência atribuídos aos prefixos que podem favorecer epêntese em PB e em inglês, bem como apresentamos e discutimos os resultados obtidos para os fatores linguísticos e extralinguísticos que, segundo as análises estatísticas que realizamos, são significantes para o fenômeno de epêntese.

No capítulo 4, finalmente, apresentamos nossas considerações finais sobre a presente investigação. Nesse capítulo, além de resumirmos os principais resultados apresentados e discutidos no capítulo 3, defendemos que tais resultados comprovam que a transferência linguística é bastante operante entre a L1 e a interlíngua dos participantes por nós inquiridos.

1 REVISÃO DA LITERATURA

O presente capítulo tem por objetivo apresentar os principais aspectos teóricos que permeiam a análise quantitativa que fazemos neste trabalho sobre a variação que ocorre na aplicação do fenômeno de epêntese em vocábulos da língua portuguesa e da língua inglesa aprendida por brasileiros. As seções deste capítulo estão distribuídas da seguinte forma: em 1.1, apresentamos as razões que nos motivaram a considerar a epêntese em inglês falado por brasileiros como uma *variação de desenvolvimento* presente na interlíngua (português-inglês) que merece ser comparada com a epêntese que ocorre em português como língua materna; em 1.2, trazemos os principais pressupostos morfológicos considerados no trabalho, que envolvem as noções de *competência lexical*, de *transparência morfológica* e a distinção entre *prefixos incorporados* e *adjuntos*; em 1.3, por fim, trazemos algumas abordagens teóricas sobre o constituinte prosódico *sílaba* e sobre o fenômeno fonológico de epêntese vocálica, aqui entendido como uma estratégia de reparo à malformação de uma sílaba, tanto em português quanto na interlíngua.

1.1 Pressupostos sobre aquisição/aprendizagem de L2/LE e sobre variação

Nesta seção inicial, temos como objetivo principal apresentar as razões por que optamos, neste estudo, por considerar a epêntese medial como um fenômeno sujeito à *variação* tanto na fala de nativos do português brasileiro (PB) quanto na fala de brasileiros em fase de aquisição / aprendizagem de inglês como segunda língua (L2) / língua estrangeira (LE).¹ Para isso, devemos detalhar dois dos pressupostos que assumimos ao longo deste estudo. São eles:

(i) a existência de uma *interlíngua* (e de uma consequente interfonologia) entre o PB e o inglês e a atuação do princípio de *transferência linguística* nessa interlíngua;

(ii) a existência de *variação* na aplicação de epêntese tanto em PB quanto na interlíngua entre o PB e o inglês.

Passemos à discussão do pressuposto (i). LADO (1957) observou que, na aprendizagem de uma língua estrangeira, os indivíduos tendem a transferir as formas e

¹ Neste trabalho, não fazemos distinção no uso dos termos *aquisição* e *aprendizagem*, nem no uso dos termos *segunda língua* (L2) e *língua estrangeira* (LE).

significados de sua língua nativa para a língua estrangeira através de um processo que denominou *transferência linguística*. A transferência linguística corresponde a um dos princípios da Hipótese da Análise Contrastiva - HAC proposta por Lado. De acordo com essa hipótese, na aquisição da fonologia de uma língua estrangeira, por exemplo, é possível prever as dificuldades que os aprendizes têm para produzir as formas-alvo com base nas diferenças fonológicas existentes entre a L1 (língua materna) e a L2. Nesse sentido, é possível prever, por exemplo, que um aprendiz terá maior dificuldade em produzir formas-alvo em uma L2 quando essa língua tem sons inexistentes na L1, ou quando tem dois ou mais sons contrastivos que, na L1, se resumem a um só som. Estudos posteriores ao de LADO (1957) mostraram que os princípios da HAC nem sempre se sustentam. Além disso, alguns estudiosos rejeitam os fundamentos da HAC (cf. SELINKER, 1972), porque entendem que, segundo essa abordagem, a aquisição de uma L2 consiste na mera aquisição de itens da L2 que não existem na L1, sendo ignorados, nesse sentido, os fatores intralinguísticos de aprendizagem, como fatores sociais, linguísticos e psicológicos referentes a cada aprendiz.

SELINKER (1972) propôs a existência, no processo de aquisição de L2, daquilo que denominou *interlíngua*. A interlíngua é definida como um sistema criativo formado pelo aprendiz que engloba o conhecimento desse aprendiz sobre a L2, sobre a L1, sobre a natureza comunicativa da linguagem, sobre as pessoas e sobre o mundo em geral. Esse sistema corresponde, pois, a uma etapa prévia à aquisição plena do sistema da língua-alvo e ocorre de forma gradual.

É importante observar que aceitar a existência de uma interlíngua não significa rejeitar os princípios da Hipótese da Análise Contrastiva. Nesse sentido, o princípio da *transferência linguística* apresentado por LADO (1957) pode ser admitido na interlíngua de um falante sem problemas², já que a transferência a que esse princípio se refere é justamente um reflexo do conhecimento que um indivíduo tem sobre sua L1, conhecimento esse que é entendido como um dos fundamentos do sistema de interlíngua. É essa, pois, a posição que assumimos para este trabalho. Acreditamos que na interlíngua de um indivíduo ocorram processos de ordem morfofonológica resultantes da transferência linguística que ocorre da L1 para essa interlíngua. Mais especificamente, acreditamos que a epêntese vocálica que ocorre no inglês

² Cabe ressaltar que, ao fazermos uso da expressão *transferência linguística* ao longo deste trabalho, não temos por objetivo assumir os pressupostos behavioristas enraizados na Hipótese da Análise Contrastiva defendida por LADO (1957). Nosso objetivo ao usar tal expressão é, pois, admitir que existem estruturas linguísticas que podem ser transferidas da L1 para a interlíngua, mas o modo como tal transferência linguística ocorre é de maior complexidade do que aquele proposto por LADO (1957), uma vez que, ao falarmos de transferência em uma interlíngua, estamos tratando de apenas um de diversos fatores que estão envolvidos no processo de aquisição de língua estrangeira por um indivíduo.

falado por brasileiros aprendizes desse idioma seja reflexo da transferência linguística que ocorre do PB (no qual a epêntese é comprovadamente recorrente, cf. COLLISCHONN, 2002, por exemplo) para a interlíngua desses aprendizes.

Resta-nos discutir o pressuposto (ii) apresentado no início desta seção. Assumimos, neste trabalho, que existe *variação* na aplicação de epêntese tanto em PB quanto no inglês em fase de aquisição por brasileiros. Deve-se destacar, no entanto, que, conceitualmente, a variação que ocorre na interlíngua não é a mesma que ocorre em PB.

Quando tratamos, neste trabalho, da variação que ocorre na aplicação de epêntese em PB, estamos nos referindo a um fenômeno fonológico condicionado socialmente e que está sujeito, portanto, aos pressupostos teóricos da Sociolinguística. Ao tratarmos da variação que ocorre na epêntese da interlíngua, estamos nos referindo ao que BEEBE (1988) denominou *variação de desenvolvimento*. No entendimento de Beebe, variação de desenvolvimento é a variação atestada em fenômenos linguísticos recorrentes na interlíngua. Em nosso trabalho, especificamente, a variação de desenvolvimento é entendida como a alternância que ocorre na frequência de aplicação de epêntese na interlíngua português-inglês.

Essa variação que ocorre na interlíngua não é aleatória. De acordo com ELLIS (1994:134), a variação de desenvolvimento (referida pelo autor simplesmente como *variação linguística*) pode ser *intra-aprendiz* ou *interaprendiz*. A variação *intra-aprendiz* resulta de fatores individuais como motivação, aptidão e personalidade, enquanto a variação *interaprendiz* provém de fatores externos ao indivíduo, sejam eles linguísticos (fonológicos, sintáticos etc), pragmáticos (tempo, tópico e propósito do discurso) ou psicolinguísticos.

Entendemos, a partir do exposto, que tanto a variação de epêntese em PB quanto a variação desse fenômeno no inglês aprendido por brasileiros mereçam ser analisadas. Ainda que se esteja tratando de dois casos de variação distintos para um mesmo fenômeno fonológico, pressupõe-se que a epêntese vocálica, em virtude da transferência linguística que admitimos ocorrer da L1 para a interlíngua, deva se manifestar de modo semelhante em ambos os sistemas, dependendo dos fatores linguísticos e extralinguísticos analisados. É com base nessa pressuposição, portanto, que realizamos neste trabalho a análise conjunta do fenômeno de epêntese medial em PB e na aquisição do inglês com LE.

1.2 Pressupostos sobre morfologia

No presente trabalho, foram igualmente levados em consideração aspectos que dizem respeito à morfologia das línguas. Esses aspectos estão fundamentalmente relacionados à análise que incluímos neste estudo sobre a aplicação do fenômeno fonológico de epêntese vocálica no contexto morfológico de fronteira entre prefixo e base. Em português brasileiro, tal ocorrência pode ser atestada em exemplos como *ad[ɪ]junto* e *sub[ɪ]raças*. No inglês aprendido por brasileiros, tal fenômeno também é observado, tal como nos exemplos *post[ɪ]date* e *sub[ɪ]chief*.

São basicamente três os itens relacionados à morfologia que devem ser detalhados:

- (i) a definição de prefixo e a forma de identificação de um prefixo;
- (ii) a noção de competência lexical e a noção de transparência a ela relacionada;
- (iii) a distinção entre prefixos incorporados e prefixos adjuntos.

A partir do detalhamento dos itens acima, temos melhores condições de explicar o comportamento da inserção vocálica no contexto morfológico referido.

No que diz respeito a (i), partimos da definição de prefixo apresentada por CAMARA JR. (1996) em seu *Dicionário de linguística e gramática: referente à língua portuguesa*. Para o autor, *prefixo* é

“(...) o afixo (v.) que vem na parte inicial do vocábulo. Na língua portuguesa, (...), o prefixo, (...), cria uma nova significação externa para a palavra a que se adjunge, (...). A apreensão dos prefixos é uma técnica de análise descritiva (v.) e se situa na sincronia linguística (v.). Para isso há a considerar três casos: 1) quando o radical a que se acrescenta constitui uma forma livre na língua (ex.: *predizer* = *pré* + *dizer*; *desconsolo* = *des* + *consolo*); 2) quando esse radical é forma livre numa estrutura variante (ex.: *permitir*, cf. *meter*); 3) quando esse radical só é forma presa, mas constitui a base de duas palavras, pelo menos, com prefixos distintos (ex., *colisão*, *elisão*). Fora disso, a palavra é, do ponto de vista descritivo, de radical simples (...)”

CAMARA JR. (1996:198)

Mais do que saber que um prefixo é um afixo que antecede um vocábulo e que atribui a este uma significação externa, é importante observar, na definição de Camara Jr., a forma como o prefixo pode ser apreendido, especialmente porque, neste estudo, tal ‘capacidade de

depreensão' foi colocada em prática por falantes nativos do PB e por aprendizes brasileiros de inglês, de modo que se pudesse verificar o *status* sincrônico (i.é, o grau de transparência) de prefixos relacionados à epêntese em PB e em inglês como LE (cf. seções 2.1.1 e 2.1.2). Essa capacidade de uma pessoa depreender um prefixo está relacionada com as noções de competência lexical e de transparência, que constituem o item (ii) de nossa discussão. Tais noções são apresentadas a seguir.

A competência lexical de um falante nativo compreende, segundo BASILIO (1980):

“a) o conhecimento de uma lista de entradas lexicais; b) o conhecimento da estrutura interna dos itens lexicais, assim como relações entre os vários itens; c) o conhecimento subjacente à capacidade de formar entradas lexicais gramaticais novas (e, naturalmente, rejeitar as agramaticais).”

BASILIO (1980:9)

Nesta pesquisa, o item b) referido na definição de competência lexical de Basilio é o que tem maior importância. Ele diz respeito à habilidade de que um falante dispõe para analisar as estruturas internas de uma palavra na sincronia da língua. A operacionalização do ato de 'tomar conhecimento' da estrutura interna de um vocábulo é estabelecida em um indivíduo, segundo a mesma autora, por meio de regras de análise estrutural (RAEs), que integram o léxico dessa pessoa.

As formas de depreensão de um prefixo apresentadas por CAMARA JR. (1996) estão intimamente relacionadas às RAEs. Casos, por exemplo, em que um falante reconhece o prefixo *ad-* no vocábulo *adjunto* ilustram a atuação da RAE [ad-_{pref.} [junto]_A]_A. Nas situações em que a RAE se aplica, diz-se que o prefixo depreendido é *transparente* (i.é, visível na estrutura interna do vocábulo).

O conhecimento sincrônico do falante sobre a estrutura da língua pode revelar, no entanto, que vocábulos historicamente prefixados não são mais identificados pela estrutura [prefixo [base]]. Num vocábulo como *eclipse*, por exemplo, é provável que um falante de PB não tenha a capacidade de identificar o prefixo *ec-*. Nesse caso, assume-se que o prefixo seja *opaco*, i.é, não transparente.

Os falantes de uma língua podem às vezes oscilar entre si (ainda que minimamente) na identificação de um prefixo da língua que falam, ou seja, nem todos os falantes de PB, por exemplo, necessariamente reconhecem *ad-* em *adjunto*, assim como nem todos os falantes

necessariamente deixam de identificar *ec-* em *eclipse*. Esse fato corrobora a postulação de que, numa determinada língua, deva existir o que chamamos de *continuum de transparência*.³ Admitem-se para esse *continuum* dois extremos: num extremo, estão as formas totalmente visíveis aos falantes (formas que são também produtivas na língua); no outro extremo, encontram-se as formas opacas, que não são mais reconhecidas por nenhum falante (e que tampouco são usadas na formação de novas palavras). É entre esses dois extremos do *continuum* que se encontra a maioria das formas de uma língua.

A cada forma é atribuído um grau de transparência específico e, dependendo do valor do grau dessa forma dentro do *continuum* (se mais próximo ou mais afastado de um dos extremos), as configurações prosódicas que se aplicam à essa forma podem mudar. A partir desse pressuposto, admite-se, neste estudo, que o comportamento fonológico atestado em prefixos que apresentem graus elevados de transparência seja diferente daquele atestado em prefixos com graus médios ou baixos de transparência. Mais especificamente, acredita-se que, quanto maior o grau de transparência de um prefixo, maiores as chances de esse prefixo estar sujeito ao fenômeno de epêntese vocálica. O estabelecimento da distinção entre prefixos incorporados e prefixos adjuntos, cf. item (iii), pode colaborar na defesa dessa hipótese.

SCHWINDT (2008), com base em trabalhos como os de BOOIJ (1996) e VIGÁRIO (2001), observa que os prefixos do PB, sob o ponto de vista prosódico, podem ser classificados em *incorporados*, *adjuntos* ou *composicionais*. Aqui nos interessa apresentar apenas os dois primeiros tipos de prefixos.

O prefixo incorporado, segundo SCHWINDT (2008:392), corresponde a uma sílaba átona à esquerda de outras sílabas, formando com estas uma só palavra fonológica (PW), tal como em [in- + -scrito]_{PW}. O prefixo adjunto, por sua vez, corresponde a uma sílaba átona à esquerda de uma PW, formando com ela uma PW recursiva, como em [des + [atado]_{PW}]_{PW}.

SCHWINDT (2008) demonstra no decorrer de seu trabalho que fenômenos fonológicos que envolvem prefixos incorporados podem não envolver prefixos adjuntos, ou

³ Esse *continuum* está intimamente relacionado à proposta de KATAMBA & STONHAM (2006) (feita a partir do trabalho de ARONOFF, 1976), que prevê a existência de uma escala com gradientes de produtividade para processos de formação de palavras. Num dos extremos dessa escala estão processos ‘congelados’ ou ‘atrofiados’ (i.é. aqueles em que as regras de formação de palavras não são empregadas pelos usuários da língua), como a formação de adjetivos atributivos em inglês a partir do sufixo latino *-id* (*timid*, *morbid*, *splendid*). No outro extremo estão casos em que as regras de formação de palavras estão em plena atuação, como no caso da formação dos agentivos em inglês a partir do sufixo *-er* (*worker*, *writer*, *jogger*). Entre esses dois extremos existe uma ‘área nebulosa’ na qual operam os processos que não chegam a ser plenamente produtivos, mas que tampouco são considerados ‘congelados’. Alguns desses processos são relativamente mais produtivos do que outros.

vice-versa. Um exemplo extraído de seu trabalho (cf. SCHWINDT, 2008:399) é a epêntese que ocorre após o prefixo adjunto *in-NEG* em *inesperado*,⁴ mas que não se aplica ao prefixo incorporado *in-DENTRO* em *inspirar*.

Em nosso estudo, analisamos o comportamento dos prefixos *ab-*, *ad-*, *ec-*, *sob-* e *sub-* para o PB, e dos prefixos *arch-*, *out-*, *post-* e *sub-* para o inglês. Em PB, *ec-* nos parece, numa classificação prévia, ser o único a se configurar como prefixo incorporado (já que parece corresponder sempre a uma sílaba átona à esquerda de outras sílabas, com as quais forma uma PW, p. ex., [ec- + -lipse]_{PW}). Não faremos classificações prévias com relação aos demais prefixos em PB.

Em inglês, vamos admitir que *arch-*, *out-*, *post-* e *sub-* parecem se configurar como prefixos adjuntos, já que parecem sempre se adjungir a uma PW pré-existente, formando com ela uma PW recursiva, como em [arch + [enemy]_{PW}]_{PW}, [out + [run]_{PW}]_{PW}, [post + [date]_{PW}]_{PW} e [sub + [license]_{PW}]_{PW}.

Após estabelecermos essas classificações prévias, resta-nos relacionar as noções de prefixo incorporado e de prefixo adjunto com a ideia de transparência e, finalmente, com a ocorrência da epêntese vocálica. Acreditamos que prefixos incorporados tenham maior tendência a se aproximar do extremo ‘opaco’ do *continuum* de transparência, ao passo que prefixos adjuntos devem estar mais próximos do extremo ‘transparente’. Isso, por consequência, implica prever que os prefixos adjuntos estão mais sujeitos à aplicação de epêntese do que prefixos incorporados.

Passemos agora à apresentação dos aspectos prosódicos que nortearam este trabalho.

1.3 Pressupostos sobre fonologia: a sílaba

A terceira área de estudo sobre o qual nos debruçamos é a Fonologia. O foco de discussão nesta pesquisa, que será detalhado nesta seção e em seus subitens, é a ocorrência de um fenômeno fonológico, qual seja, a epêntese vocálica, que ocorre em PB e em inglês como LE como estratégia de reparo a um constituinte prosódico: a sílaba.

A sílaba, apesar de já ser reconhecida no trabalho de CHOMSKY & HALLE (1968), *The Sound Pattern of English* (considerado um marco nos estudos gerativistas sobre Fonologia), não recebeu, nessa obra, uma discussão pormenorizada das características

⁴ Orientado por COLLISCHONN & SCHWINDT (2005), o autor assume que, em PB, as palavras iniciadas por [esC] são, em sua maioria, produto de epêntese sincrônica.

atinentes a ela. Somente a partir dos anos 70, com a proposição da Teoria Autossegmental, a sílaba passou a ser vista como um elemento fonológico com características particulares.

Os trabalhos de nosso interesse que tratam da estrutura da ‘sílabas’ e das regras ou condições que atuam sobre ela são apresentados nas subseções do item 1.3.1, Teorias gerais da sílaba. Esses trabalhos servem para demonstrar a evolução das discussões sobre o constituinte silábico, bem como para mostrar quais pressupostos foram adotados nas análises teóricas para a epêntese em PB já feitas – cf. COLLISCHONN (1996) e BISOL (1999) na seção 1.3.4.

Em 1.3.2, são apresentados os moldes silábicos e as condições de boa-formação que constituem o mapeamento da sílaba em PB, com base em CAMARA JR. (1969), em COLLISCHONN (1997) e BISOL (1999). A seção 1.3.3, por sua vez, apresenta essa mesma proposição de padrões silábicos, agora para o inglês, com base nos trabalhos de SELKIRK (1982) e HAMMOND (1999). Na última seção, 1.3.4, é aprofundada a discussão sobre o fenômeno de epêntese, foco do presente estudo, aqui visto como um recurso utilizado por falantes nativos de PB e por brasileiros aprendizes de inglês como LE para respeitar condições de boa-formação de sílaba.

1.3.1 Teorias gerais da sílaba

As investigações gerativistas na área de Fonologia passaram a dar maior atenção ao componente ‘sílabas’ a partir da década de 1970, especialmente após a introdução da Fonologia Autossegmental por GOLDSMITH (1976). Existem basicamente duas abordagens principais nos estudos sobre a sílaba: uma que considera que a estrutura silábica seja construída por regras de silabação e outra que não admite a existência de regras, mas, sim, de princípios gerais de estrutura silábica.

Nas próximas quatro subseções, são apresentadas as propostas sobre sílaba consideradas de maior relevância para este trabalho. Os trabalhos de KAHN (1976) e de CLEMENTS & KEYSER (1983) representam abordagens que admitem a existência de regras de silabação, enquanto os de SELKIRK (1982) e ITÔ (1986) representam abordagens que pressupõem a atuação de princípios gerais sobre a estrutura da sílaba. No que diz respeito à explicação da epêntese aqui investigada, as propostas apresentadas por essas duas últimas autoras parecem ser as mais adequadas.

1.3.1.1 Kahn (1976)

A proposta de sílaba apresentada por KAHN (1976) se baseia nos princípios da Fonologia Autossegmental, modelo não linear introduzido por GOLDSMITH (1976).

Para Kahn, os segmentos são diretamente vinculados a sílabas, através de linhas de associação (cf. [1]). Nesse caso, as sílabas são consideradas elementos de um nível autossegmental independente e não possuem uma estrutura interna definida.

[1] m æ g n ə t
 \ | / \ | /
 S₁ S₂

Tal proposta não faz referência à existência de constituintes internos à sílaba. Os segmentos estão diretamente subordinados a ela e possuem, portanto, relação idêntica entre si na hierarquia.

Para que se estabeleça a associação entre os segmentos e as sílabas, devem ser respeitadas convenções universais (cf. [2]), assim como regras específicas.

[2] Convenções universais

- a) cada segmento [+ silábico] é associado a exatamente uma sílaba.
- b) cada segmento [- silábico] é associado a pelo menos uma sílaba.
- c) as linhas que associam sílabas e segmentos não podem se cruzar.

(KAHN, 1976:38)

Como a aplicação dessas convenções acaba permitindo a formação de mais de um tipo de silabação para uma mesma sequência de segmentos, faz-se também necessária a aplicação de regras específicas (particulares à língua ou universais), que proibirão determinadas sílabas.

De acordo com o que diz a convenção b) acima, é lícita a ocorrência de segmentos ambissilábicos (vinculados a mais de uma sílaba). Tal recurso é admitido na proposta de

Kahn, contanto que se observe o critério explicitado pela convenção c), que permite a dupla associação de um segmento somente a sílabas vizinhas (cf. ‘n’ em [3]).

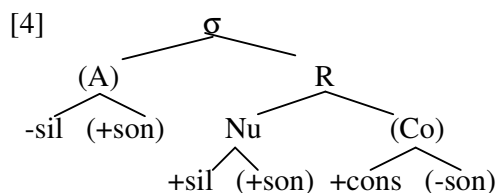
[3] p o n i
 \ | / \ |
 S₁ S₂

O processo de silabação ocorre de dentro para fora. Num primeiro momento, as vogais são individualmente associadas a um elemento S. Em seguida, são associadas as consoantes à esquerda da vogal e, por fim, as consoantes à direita da vogal. O papel das convenções universais é controlar a aplicação das regras de silabação, de modo a impedir a produção de representações de associação malformadas.

Uma das principais contribuições de KAHN (1976) foi seu pioneirismo em relação ao uso de um novo nível de representação fonológica, a saber, o das sílabas. No entanto, seu estudo deu margem para críticas, já que suas representações, por exemplo, não estabeleciam uma distinção precisa entre segmentos nucleares e periféricos dentro da sílaba, tampouco tratavam das possibilidades de ressilabação nas línguas. Tais questões foram contempladas na proposta de SELKIRK (1982).

1.3.1.2 Selkirk (1982)

Na análise proposta por SELKIRK (1982), a sílaba (simbolicamente σ) é vista como um constituinte prosódico de uma hierarquia maior que é composto por subdivisões internas. A representação em [4] mostra a proposta de estrutura da sílaba para o inglês.



Legenda: A: ataque sil: silábico
 R: rima son: sonorante
 Nu: núcleo cons: consonantal
 Co: coda (): opcional
 Adaptado de SELKIRK (1982:344)

A autora defende esse tipo de estruturação silábica para explicar os casos de aplicação de restrições fonotáticas específicas a partes da sílaba. Através dessa representação, o

constituente rima, por exemplo, permite que se estabeleça a distinção entre sílabas pesadas e leves. Sílabas com rima complexa apresentam um elemento vocálico ou consonantal após o núcleo (CVC ou CVV) e são denominadas pesadas, enquanto sílabas com rima simples, em que apenas seu núcleo é preenchido (CVY), são denominadas leves. Essa distinção é de grande relevância nos estudos sobre acento, por exemplo, uma vez que sílabas pesadas tendem a atrair o acento da palavra.

A atribuição da sílaba em SELKIRK (1982) se dá em três partes. Na primeira delas, é feita a caracterização das estruturas possíveis para a sílaba. Isso inclui: a) a caracterização da estrutura interna da sílaba; b) a especificação do número mínimo e máximo de posições terminais; e c) um conjunto de condições sobre os nós terminais. O resultado disso é o que Selkirk denominou *template* (ou molde silábico), tal como o que é mostrado na representação [4] supra.

A segunda parte envolve a caracterização das sequências que são possíveis ou não no nível melódico. Admitindo-se, por exemplo, a existência de um molde silábico CCV para o português brasileiro, é possível formar sequências como *cri* ou *clã*, mas também **vde* ou **bno*, estas últimas inexistentes no idioma. Para que as sequências complexas *vd* ou *bn* sejam evitadas no ataque, a teoria propõe a existência de filtros, que têm poder de restringir os tipos de segmentos ou sequências de segmentos em qualquer subconstituente da sílaba.

Finalmente, a terceira parte consiste na caracterização dos tipos de associação que são possíveis com base nas duas primeiras etapas. Para essa caracterização, é levada em consideração a condição de não cruzamento de linhas de associação, assim como outras condições, universais ou específicas, que determinam o modo como os segmentos são associados aos subconstituintes da sílaba.

SELKIRK (1982) defende que os *templates* ou moldes são estruturas subjacentes e que, portanto, o processo de estruturação silábica não se preocupa em mostrar regras que associam sequências de segmentos a sílabas (como é o caso em KAHN, 1976).⁵ O processo de formação da sílaba em uma língua envolve, pois, princípios de composição (*basic syllable composition*) que determinam a existência de um *template* sobre o qual operam condições de boa-formação (particulares à língua ou universais).

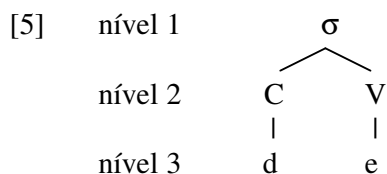
⁵ A ressilabação, no entanto, é tratada pela autora como um processo em que a reassociação de segmentos a sílabas se faz necessária.

Em suma, a proposta de SELKIRK (1982) se destacou pelo fato de considerar a silabação não como um processo de aplicação de regras de formação, mas como um conjunto de condições de boa-formação que devem ser aplicadas a moldes silábicos subjacentes. Além disso, a divisão da sílaba em ataque e rima (e de rima em núcleo e coda) foi uma grande contribuição de SELKIRK (1982) para os estudos subsequentes sobre a sílaba, em especial por possibilitar a explicação de fenômenos fonológicos como a epêntese existente no PB e na aquisição do inglês por brasileiros, que é motivada especificamente por determinados tipos de consoantes que ocupam a posição de coda de uma sílaba.

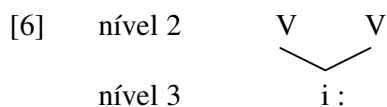
1.3.1.3 Clements e Keyser (1983)

À semelhança de KAHN (1976), CLEMENTS & KEYSER (1983) argumentam em favor de uma estrutura plana para a sílaba. Para eles, não há razão de se considerar uma estrutura complexa para explicar processos que afetem apenas partes do constituinte, como sugere SELKIRK (1982).

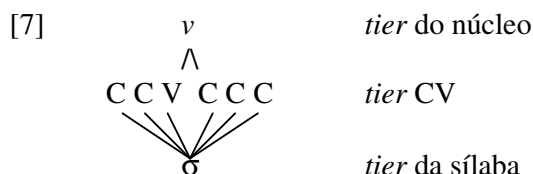
A diferença que os autores trazem com relação a KAHN (1976) está na proposição de um nível intermediário entre o nível dos segmentos e o nível da sílaba. Esse nível corresponde ao denominado esqueleto CV (cf. [5]).



Através da representação do esqueleto CV, é possível diferenciar os segmentos que podem constituir pico de sílaba (V) dos que não podem (C). Além disso, o esqueleto CV representa as unidades de tempo às quais os segmentos estão associados. Isso possibilita que se estabeleça uma projeção clara de como se constituem, por exemplo, as vogais longas, em que duas unidades de tempo estão associadas a um único elemento fonético de duração maior (cf. [6]).



CLEMENTS & KEYSER (1983) propõem, ainda, a existência de um nível silábico denominado núcleo (v) (cf. [7]), especialmente para dar conta da distinção entre sílabas pesadas e sílabas leves (que, em SELKIRK, 1982, é feita através da presença ou da ausência da coda na rima). O núcleo (v) é composto por uma vogal longa ou ditongo, ou por uma vogal curta e uma consoante.



Esse nível é proposto num plano à parte dos níveis da sílaba e do esqueleto CV. Como observa COLLISCHONN (1997:41), essa forma de representação foi criticada porque atribui um poder indesejável à teoria, já que, a partir do momento em que se admite um novo plano de representação (neste caso específico, para o núcleo), é também possível presumir a existência de outros *tiers* para dar conta de outras subdivisões silábicas.

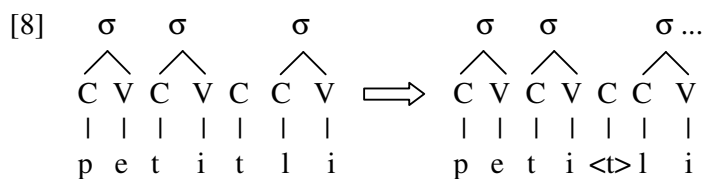
O modo como a estrutura da sílaba é atribuída é semelhante à proposta de KAHN (1976), com a diferença de que a associação agora não se dá diretamente entre os segmentos e a sílaba, mas por intermédio do esqueleto CV. Além disso, os autores consideram que a estrutura silábica já esteja presente na subjacência.

Para CLEMENTS & KEYSER (1983), a silabação deve obedecer a convenções universais de associação e a condições específicas das línguas, essas classificadas em positivas (condições que dizem quais sequências são bem formadas) e negativas (condições que excluem as sequências admitidas pelas condições positivas que não são atestadas na língua). As condições negativas se assemelham aos filtros já propostos em SELKIRK (1982).

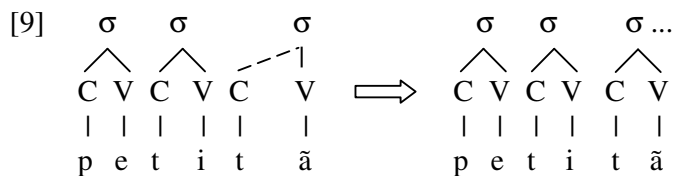
Clements e Keyser discutem, ainda, a noção de extrassilabidade. Os autores defendem que, durante o processo de derivação, todas as linhas de associação entre elementos CV e σ são apagadas para serem reassociadas. Durante essa nova associação, segmentos podem permanecer desassociados a σ , vindo a se associar posteriormente à sílaba adjacente por meio de regras de afiliação.

Um exemplo de segmento extrassilábico dado por Clements e Keyser é o segundo /t/ da palavra francesa *petit* ('pequeno'). Em francês, não são admitidas obstruintes na posição de coda silábica e, por isso, /t/ não é realizado em casos como *petit livre* e *petit garçon*. Observe-se, no entanto, que em casos em que /t/ é seguido de uma vogal, como em *petit enfant*, a consoante se realiza ([petitãfã]).

Admite-se que o segundo /t/ em *petit livre* não possa ser associado a nenhuma sílaba durante a derivação. Em razão disso, esse segmento não será realizado na superfície, mas também não será apagado, uma vez que seu caráter extrassilábico o torna invisível às operações de apagamento (cf. [8]).



Quando estiver diante de vogal, no entanto, o /t/ de *petit* se associa à sílaba à qual essa vogal vizinha está associada, tornando-se o *onset* dela (cf. [9]). Nesse caso, /t/ aparece na superfície.



Em suma, destacam-se, dentre as contribuições de CLEMENTS & KEYSER (1983), a introdução do esqueleto CV nos estudos sobre a sílaba e a defesa da existência de elementos extrassilábicos. Um aspecto, contudo, considerado negativo na proposta é a proposição de um núcleo representado em um nível distinto do plano da sílaba e do esqueleto CV, como foi mencionado anteriormente.

1.3.1.4 Itô (1986)

No estudo desenvolvido por ITÔ (1986), o constituinte sílaba é tratado como parte integrante de um conjunto de constituintes prosódicos (como a frase entonacional, a palavra fonológica, o pé métrico, etc.) introduzido pelos estudos da Fonologia Prosódica, proposta teórica que estipula a hierarquização entre os elementos fonológicos desse conjunto.

Os pressupostos da Fonologia Prosódica preveem que essa hierarquia deva obedecer a três princípios básicos, quais sejam: (i) Licenciamento Prosódico, (ii) Localidade e (iii) Direcionalidade. O Licenciamento Prosódico diz que todas as unidades fonológicas (sílabas, pés, etc.) devem pertencer a estruturas fonológicas superiores (à exceção de elementos extraprosódicos). O princípio de Localidade estipula que a boa-formação de uma estrutura prosódica é determinada localmente (i.é, dentro da própria unidade prosódica, independente das informações externas). Já a Direcionalidade determina a existência de uma direção na ordem do mapeamento prosódico, que pode ser da esquerda para a direita, ou da direita para a esquerda, dependendo da língua.

ITÔ (1986), à semelhança de SELKIRK (1982), argumenta em favor da existência de um molde silábico (*template*), condição que determina o número máximo de elementos C e V ao nível silábico σ . Não existe, no entanto, uma estrutura interna para a sílaba em seu modelo: os elementos silábicos dominam diretamente o nível do esqueleto CV, o que se assemelha, nesse caso, à representação plana proposta por CLEMENTS & KEYSER (1983).

Itô ainda defende em seu estudo que a silabação obedece a condições de boa-formação (*wellformedness conditions*).⁶ Essas condições são específicas em cada língua e determinam a distribuição dos segmentos para cada posição do *template*, de modo a dar conta de fatos distribucionais e de restrições de coocorrência particulares a cada língua. O mapeamento da sílaba é, portanto, resultante da conjunção entre o molde silábico e as condições de boa-formação, sendo sempre respeitados os três princípios básicos da teoria.

A análise feita por ITÔ (1986) vincula a Teoria da Sílaba à Fonologia Lexical. Assim sendo, a silabação é vista como um processo que ocorre dentro do Léxico. A teoria prevê, nesse sentido, que cada elemento C ou V deverá, ao final do processo de aplicação de regras de cada ciclo do Léxico, estar vinculado a seu constituinte imediatamente superior (no caso, σ). Caso algum elemento não seja associado a nenhuma σ , o processo de Apagamento de

⁶ Assim como SELKIRK (1982), a autora argumenta que a silabação não se preocupa com regras de formação da sílaba.

Elemento Perdido (*Stray Erasure Convention*) se aplica, em respeito ao princípio de Licenciamento Prosódico.

Pode-se resumir que o estudo de ITÔ (1986) sobre a sílaba reuniu diferentes propostas teóricas (em especial as de SELKIRK, 1982, e de CLEMENTS & KEYSER, 1983) para chegar a um modelo mais econômico e abrangente sobre o constituinte, válido tanto para os estudos da Teoria da Sílaba como os da Fonologia Prosódica e da Fonologia Lexical.

No que se refere à epêntese em PB, o modelo de sílaba de ITÔ (1986) serviu de base para sua investigação em COLLISCHONN (1996). Esse modelo, no entanto, não foi seguido por BISOL (1999), que, para explicar a epêntese, admitiu a existência de uma estrutura interna para a sílaba, tal como em SELKIRK (1982). As abordagens de COLLISCHONN (1996) e de BISOL (1999) sobre a epêntese em PB são detalhadas em 1.3.4.

Nas duas próximas seções, são apresentados os moldes silábicos da língua portuguesa e da língua inglesa, bem como as condições de boa-formação que atuam sobre esses moldes.⁷ Como será visto, as condições que operam nos *templates* do português são diferentes das que atuam em inglês, fato que justifica a inserção de uma vogal em determinados *clusters* consonantais do PB e do inglês aprendido por falantes nativos de PB.

1.3.2 A sílaba em português brasileiro (PB)

CAMARA JR. (1969) analisa a sílaba como um elemento funcional que, na enunciação, apresenta um momento ascensional, um momento de plenitude e um momento de descensão, ou seja, um aclave, um ápice e um declive. Ao ápice corresponde a emissão de uma vogal, que constitui um segmento ‘silábico’. Os fonemas do aclave e do declive são considerados ‘assilábicos’ e podem estar ausentes na sílaba.

Para CAMARA JR. (1969:26), a sílaba na língua portuguesa se apresenta em três estruturas fundamentais: sílaba simples, V, como em *é*; sílaba complexa aberta, CV, como em *pé*; e sílaba complexa fechada, VC ou CVC, como em *ás* e *par*.

O autor observa que, no aclave simples, todas as consoantes da língua podem aparecer. Pode haver, no entanto, a ocorrência de duas consoantes no aclave; nesse caso, a segunda delas é sempre /r/, /l/, ou um glide. O /N/, em casos como *pneu*, é tratado pelo autor como

⁷ Optou-se por usar no restante do trabalho o termo *condição de boa-formação* (cf. ITÔ, 1986) para se referir a condições/filtros/regras que militam sobre cada posição do molde silábico.

acrive de uma sílaba diferente da consoante antecedente. Nesse exemplo, o /p/ é considerado acrive de uma nova sílaba, cujo ápice é preenchido por uma vogal epentética.

Com relação ao declive da sílaba, CAMARA JR. (1969) diz que sílabas fechadas (ou travadas) são menos frequentes na língua e têm uma limitação muito grande com relação às consoantes que figuram no declive. As consoantes admitidas após o ápice, segundo o autor, são /l/ *mal*, /r/ *mar*, /s/ *paz* e /N/ *sim*, bem como os glides /y/ *sei* e /w/ *seu*. As consoantes oclusivas e as fricativas /f/ e /v/, portanto, não ocorrem no declive. Nesse sentido, o /k/ que ocorre em *pacto* e o /f/ que ocorre em *afta* constituem acrive de uma nova sílaba, cujo ápice corresponde a uma vogal epentética.

COLLISCHONN (1997), com base nas teorias que admitem a existência de *templates* para as sílabas nas línguas, relacionou todos os padrões silábicos encontrados na língua portuguesa. A relação completa desses padrões pode ser verificada em [10].

[10]	V	<u>é</u> , <u>a</u> berto
	CV	<u>cá</u> , <u>pa</u> lito
	CCV	<u>pré</u> , <u>af</u> lito
	CVC	<u>ma</u> r, <u>pe</u> ste
	VC	<u>a</u> r, <u>on</u> tem
	CCVC	<u>três</u> , <u>gra</u> nde
	CVCC	<u>pers</u> picaz
	CCVCC	<u>trans</u> por, <u>trens</u>
	VCC	<u>in</u> stante
	CVV	<u>pa</u> i, <u>rei</u> no
	CCVV	<u>gra</u> u, <u>fre</u> i
	VV	<u>oi</u> to, <u>aur</u> ora
	CCVVC	<u>gra</u> us, <u>clau</u> stro
	CVVC	<u>ma</u> is, <u>cáu</u> stico
	VVC	<u>eis</u> , <u>au</u> stríaco

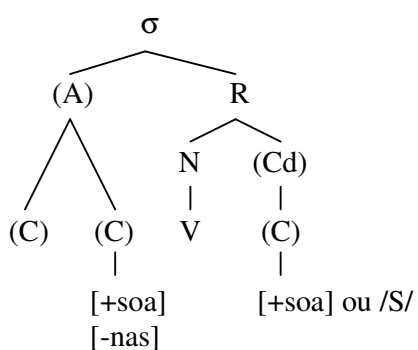
Adaptado de COLLISCHONN (1997:77)

COLLISCHONN (1997) constata que, no ataque da sílaba, pode haver até duas consoantes. Para dar conta das restrições que se aplicam nos casos em que as duas consoantes aparecem simultaneamente na língua em *onset*, a autora postula duas condições de ataque complexo, que atuam em conjunto: 1ª. Se um ataque for constituído de dois elementos, o primeiro deve ser [-soante] (oclusivo) e o segundo [+aproximante] (líquida); 2ª. Se um ataque for constituído de duas casas, então o segmento da primeira casa não pode estar associado a [+coronal] e [+contínuo] simultaneamente (ou seja, não pode ser /s,z,ʃ,ʒ/).

Para a coda da sílaba, COLLISCHONN (1997) admite a existência de duas posições temporais: na primeira podem estar semivogais, líquidas ou nasais, ao passo que, na segunda, aparecem apenas sibilantes (como em *trens*, CCVCC). A autora defende, portanto, a existência de duas posições na coda, que são controladas por três condições que atuam em cooperação: 1ª um elemento na coda não deve ser ao mesmo tempo [-soante,-contínuo] (ou seja, não deve ser oclusivo); 2ª se há um segmento [-vocoide] na coda, este deve ser [+coronal,+anterior] (o que exclui /m,f,v,x,ɣ,ŋ,ʃ,ʒ,ɲ/ e permite apenas glides, nasais, líquidas e sibilantes na coda); 3ª havendo dois elementos na coda, o segundo é obrigatoriamente [-soante] (i.é, /s/). COLLISCHONN (1997:103) destaca que essas condições “(...) determinam que, havendo uma sequência de duas oclusivas (como em *apto*), ou de oclusiva e nasal (como em *ritmo*), a primeira das duas consoantes não pode ser associada, o que cria o contexto para a epêntese.” Essa abordagem sobre epêntese é detalhada em 1.3.4.

BISOL (1999) igualmente pressupõe a existência de *templates* e apresenta um padrão silábico sintético CCVC (C) para o PB (em que o C parentético - /S/ - é resultado de uma regra particular). BISOL (1999) propõe, ademais, que a sílaba (σ) em PB se apresente sob a forma de estrutura arbórea, tal como representado em [11].

[11]



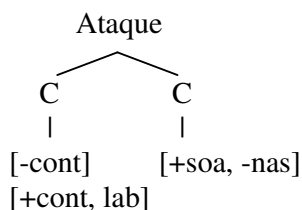
Legenda: A: ataque soa: soante
 R: rima nas: nasal
 N: núcleo /S/: [s,z,ʃ,ʒ]
 Cd: coda (): opcional

Adaptado de BISOL (1999:703)

BISOL (1999) utiliza o mesmo tipo de representação de sílaba proposto por SELKIRK (1982) para o inglês (cf. 1.3.1.2 e 1.3.3). Seu principal argumento em admitir a existência de estruturas internas é que a distinção entre rimas simples e complexas que a ramificação de R representa é de grande importância para a descrição do acento em PB (e em outras línguas sensíveis ao peso da sílaba). Além disso, “(...) certas regras precisam referir-se aos constituintes silábicos ou a limites com eles relacionados, como a vocalização da lateral que ocorre somente em posição de coda (maL~maw), (...).” (BISOL, 1999:704).

A partir da pressuposição de princípios universais que atuam sobre a sílaba, BISOL (1999) propõe a existência de apenas dois tipos de condição de boa-formação: a Condição do Ataque e a Condição da Coda. A Condição do Ataque é representada pela autora da seguinte forma:

[12] Condição Positiva do Ataque Complexo



BISOL (1999:718)

Essa condição determina que, na primeira posição do ataque complexo, a consoante pode ser tanto [-contínua] (/p/ *pra.to*, /t/ *tra.to*, /k/ *cra.vo*, /b/ *bra.ço*, /d/ *dra.ma*, /g/ *gra.ma*) como uma [+contínua, labial] (/f/ *fra.co*, /v/ *li.vro*). Na segunda posição, a consoante deve ser simultaneamente [+soante] e [-nasal] (/l/ *bloco*, /r/ *fre.vo*). São excluídos por essa condição, portanto, os *clusters* [sr], [sl], [zr], [zl], [fr], [fl], [ʒr], [ʒl], [xl] e [xr].⁸ As sequências de segmentos de sonoridade idêntica (obstruinte + obstruinte, *ptose*) ou vizinha (obstruinte + nasal, *pneu*), acabam sendo rejeitadas dentro de uma mesma sílaba da língua, isso porque nela atuam princípios universais (no caso de *ptose*, atua o Princípio de Sonoridade Sequencial, o qual determina que, entre qualquer membro de uma sílaba e o pico silábico, a sonoridade é crescente⁹). A epêntese atestada em *p[i]tose* e *p[i]neu* é, sob esse aspecto, uma forma de reparo à uma sílaba malformada em PB (cf. 1.3.4).

A segunda condição de sílaba segundo BISOL (1999) é a Condição da Coda. A representação dessa condição é apresentada em [13].

⁸ Os encontros /dl/ e /vl/, embora sejam permitidos por essa condição, não são atestados em PB (a não ser em raros nomes próprios, como *Adler* e *Vladimir*, e na fronteira entre prefixo e base, como em *adligação*). BISOL (1999) interpretou esses encontros como grupos vazios no sistema, abertos a empréstimos ou a palavras novas.

⁹ A apresentação desse Princípio (*Sonority Sequencing Principle*) é feita originalmente por CLEMENTS (1990), que propõe a existência da Escala de Sonoridade, na qual os segmentos apresentam diferentes graus de sonoridade, de acordo com a seguinte ordem: vogais > líquidas/glides > nasais > fricativas > oclusivas.

[13] Condição da Coda

$$\begin{array}{c} *C] \sigma \\ | \\ [-soante], \text{ exceto /S/} \end{array}$$

BISOL (1999:720)

Essa condição negativa proíbe a ocorrência de obstruintes (exceto /S/) e permite a ocorrência de glides (*pai*), líquidas (*par*) e nasais (*sem*) na coda. Essa condição, no entanto, não dá conta de casos como *graus*, que apresentam duas posições na coda.

BISOL (1999) postula, pois, a existência de apenas uma posição C na coda do *template* canônico (CCVC). Segundo a autora, a presença do segundo C em sílabas como VCC (*austríaco*)¹⁰, CCVCC (*graus*) etc. corresponde sempre a um /S/ que é adjungido por meio de uma regra adicional, denominada Regra de Adjunção de /S/ (RAS), que determina: “Acrescente /S/ à rima bem formada.” (BISOL, 1999:704). A RAS e a Condição da Coda juntas, portanto, dão conta das possibilidades restritas de segmentos admitidos na coda de sílabas do PB. Segmentos que não respeitam essas condições podem ser ‘salvos’ através de reparos silábicos, como é o caso atestado em vocábulos com epêntese vocálica (cf. 1.3.4).

1.3.3 A sílaba em inglês

A sílaba da língua inglesa apresenta estrutura razoavelmente diferente daquela que atestamos para a língua portuguesa. Uma das principais diferenças que entre as sílabas dessas duas línguas, como veremos nos próximos parágrafos, está no fato de que determinados *clusters* consonantais são admitidos no *onset* ou na coda da sílaba em inglês, mas não são admitidos nessas posições na sílaba em português. Devido a esse tipo de diferença, os falantes nativos de PB podem fazer uso da epêntese na interlíngua para desfazer esses *clusters* consonantais que ocorrem em determinadas posições na sílaba da LE, mas que não ocorrem nessas mesmas posições na sílaba da L1.

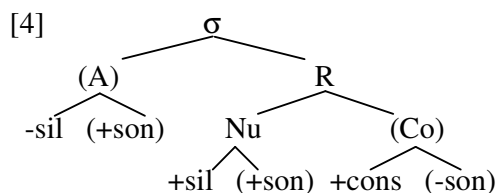
Uma sílaba em inglês pode admitir até seis segmentos. Apresentamos em [14], com base na adaptação do levantamento feito por PEREYRON (2008:31), os padrões silábicos que existem na língua inglesa.

¹⁰ Note-se que glides são considerados elementos C pela autora.

[14]	V	'a'	[ə]
	CV	'to'	[tu]
	CCV	'tree'	[tri:]
	VC	'at'	[æt]
	CVC	'cat'	[kæt]
	CCVC	'black'	[blæk]
	CCCVC	'stress'	[stres]
	VCC	'apt'	[æpt]
	CVCC	'cast'	[kæst]
	CCVCC	'plant'	[plænt]
	CCCVCC	'springs'	[sprɪŋz]
	VCCC	'angst'	[æŋst]
	CVCCC	'text'	[tekst]
	CCVCCC	'sphinx'	[sfɪŋks]
	CCCVCCC	'strenghts'	[strenθs]
	VCCCC	'angsts'	[aŋsts]
	CVCCCC	'texts'	[teksts]

Adaptado de PEREYRON (2008:31)

O trabalho de SELKIRK (1982) apresenta a sílaba em inglês sob a forma de uma estrutura arbórea que contempla os padrões silábicos apresentados em [14], bem como apresenta as restrições (filtros) que existem sobre os tipos de segmento que são admitidos em cada posição da sílaba. Essa estrutura arbórea, que já foi apresentada em [4], é aqui retomada.



Legenda: A: ataque sil: silábico
 R: rima son: sonorante
 Nu: núcleo cons: consonantal
 Co: coda (): opcional

Adaptado de SELKIRK (1982:344)

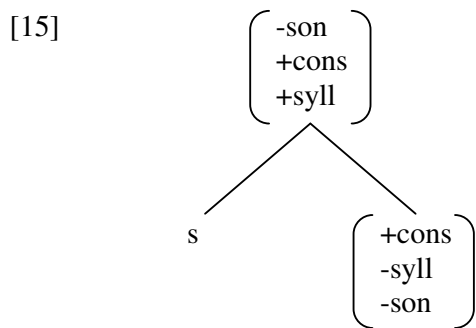
De acordo com o estudo de SELKIRK (1982), o ataque da sílaba em inglês, quando for simples (i.é, não ramificado), pode ser preenchido por qualquer consoante atestada nesse idioma, exceto [3] e [ŋ]. Quando o *onset* apresentar dois segmentos, o primeiro é sempre [-silábico], podendo ser tanto uma consoante oclusiva ([klɪp] *clip*) quanto uma fricativa desvozeada ([θri:] *three*).

O segundo segmento de um *onset* com dois elementos deve ser um segmento sonorante (não obstruinte). Há, no entanto, algumas restrições sobre a ocorrência desses

segmentos sonorantes na segunda posição de *onset*. Uma delas é a de que o segmento sonorante [j] não é permitido nessa posição. Outra restrição é a de que [w] não ocorre após consoantes labiais, tampouco após o *cluster* [st]. No que concerne às líquidas, o segmento [r] nunca ocorre após [s] ou [h], enquanto [l] nunca ocorre depois de [t], [d], [ʃ], [h] de [sk]. Há, ainda, a restrição de que os segmentos nasais [m] e [n] devam ser sempre antecidos por [s], como atestam os exemplos *smart* e *snack*.

SELKIRK (1982) dá atenção especial para os ataques silábicos formados por três consoantes (como [str] em *street*), que apresentam sempre [s] na primeira posição. Note-se inicialmente que, na representação arbórea proposta pela autora (cf. [4]), a presença de três segmentos no ataque nem chega a ser ilustrada.

A autora (op. cit.) propõe a existência de um *template* auxiliar que dê conta dos casos em que [s] antecede duas consoantes em ataque de sílaba. A representação desse *template* auxiliar está expressa em [15].



SELKIRK (1982:347)

SELKIRK (1982) propõe a existência desse *template* auxiliar sob o argumento de que, nos *clusters* /sC/ (que podem ocorrer tanto no ataque quanto na coda de sílaba, como mostram os exemplos *stress* e *angsts* em [14]), a sibilante [s] forma com a consoante obstruinte seguinte uma obstruinte única. A partir disso, a autora conclui que essa sibilante não deva corresponder a um terceiro elemento em posição de ataque silábico (o que justifica a representação de ataque com apenas dois elementos em [4]).

O núcleo da sílaba em inglês, segundo SELKIRK (1982), é formado por até dois elementos. Quando é preenchido por só um elemento, este pode ser uma vogal (*pass*) ou uma consoante sonorante (como [l] em [nu:.dl] *noodle*). Quando o núcleo é preenchido por dois

elementos, podem ocorrer nele ditongos (como [eɪ] em *say*) ou vogais longas (como [i:] em *beat*). Existem ainda casos restritos em que [r], [l] e /N/ podem figurar na segunda posição do núcleo silábico.

Na posição de coda da sílaba em inglês, existem, segundo SELKIRK (1982), até duas posições que podem ser preenchidas. A consoante [s] que aparece em encontros /sC/ na coda deve ser contada juntamente com a C seguinte para efeitos de contagem de posição das posições (i.é, /sC/ ocupa apenas uma posição das duas posições disponíveis na coda em inglês).¹¹

Quando a coda em inglês é preenchida por uma consoante apenas, tal consoante pode ser uma líquida (*car*), uma nasal (*sin*) ou uma obstruinte (*cat*). Quando ela é preenchida por duas consoantes, as restrições sobre os tipos de consoante admitidos na primeira e na segunda posição são maiores.

Na primeira posição de uma coda com dois elementos, não são atestadas as consoantes [b], [g], [v], [ʃ] e [ʒ], tampouco as africadas [tʃ] e [dʒ]. Na segunda posição, a consoante atestada é sempre uma obstruinte (como [t] *best* ou [θ] *oath*).

Como podemos perceber através da descrição das observações feitas por SELKIRK (1982) a respeito da sílaba em inglês, nesse idioma, admite-se a ocorrência de consoantes obstruintes na posição de coda. São exemplos disso as consoantes [g] e [t] que ocorrem em *mag.net*. Essa parece ser uma realidade diferente da constatada para o português (cf. 1.3.2), já que, nesse idioma, as obstruintes (à exceção das sibilantes) não são admitidas na coda silábica. Como veremos na próxima seção, palavras que apresentam obstruintes em posição de coda tendem a ser ressilabadas pelos falantes nativos de PB através de epêntese. A epêntese, nesse sentido, atua como uma estratégia de reparo a uma sílaba malformada.

1.3.4 A epêntese como estratégia de reparo à malformação da sílaba

A epêntese é um fenômeno fonológico registrado em diversas línguas. Na obra de (CRYSTAL, 2003), *A Dictionary of Linguistics and Phonetics*, o vocábulo *epenthesis* (epêntese) é definido como

¹¹ Casos com quatro elementos consonantais em coda (como [teksts] *texts*), são apenas atestados em sílabas que contêm na coda o encontro /sC/ e o segmento marcador de flexão (em *texts*, [s]). SELKIRK (1982) argumenta, adicionalmente à proposta de que /sC/ ocupem uma só posição, que o segmento marcador de plural corresponde a um sufixo inflexível e que, em razão disso, ele é externo à silabação.

“A term used in PHONETICS and PHONOLOGY to refer to a type of INTRUSION, where an extra sound has been inserted in a WORD; often subclassified into PROTHESIS and ANAPTYXIS.”¹²

CRYSTAL (2003:163)

A inserção de um segmento (consonantal ou vocálico) nas línguas pode ocorrer por questões de acomodação articulatória, de eufonia, de analogia, etc. Para o português brasileiro (PB), entendemos que a epêntese seja uma estratégia de reparo à malformação de uma sílaba na língua, i.é, uma forma de se evitar que as condições de boa-formação aplicadas ao molde silábico da língua sejam desrespeitadas. Essa estratégia de reparo pode envolver a inserção de uma vogal no início de um vocábulo (i.é, a *prótese*, como [e] em [e]státua, originalmente *statua* em latim), em seu interior, entre duas consoantes (i.é, o *suarabácti* ou *anaptixe*, como [i] em *p[i]neumonia* ou em *sub[i]marino*) ou em seu final (i.é, a *epêntese* ou *paragoge*, como [i] em *clip[i]*, forma aportuguesada de *clip*).

Os exemplos acima mostram que o PB apresenta tanto vocábulos nos quais a aplicação de epêntese é considerada categórica – como *estátua* e *clipe*, que inclusive já tiveram a vogal incorporada à grafia – quanto vocábulos nos quais a inserção vocálica é um fenômeno sincronicamente variável restrito à fala dos usuários do idioma, como é o caso em *p[i]neumonia* e em *sub[i]marino* (que são frequentemente pronunciados sem a vogal, tais como são ortograficamente registrados). Nesses dois últimos exemplos, é também possível perceber que a inserção da vogal no interior do vocábulo pode ocorrer em dois contextos morfológicos distintos: dentro da base (como em *pneumonia*), quanto na fronteira entre um prefixo e sua base (como em *sub+marino*).

Na aquisição do inglês como língua estrangeira (LE) por brasileiros, a epêntese (inicial, medial ou final) é frequentemente identificada em dados de fala produzidos pelos aprendizes. Na língua inglesa, no entanto, a inserção vocálica não é um fenômeno comumente adotado como estratégia de reparo à sílaba, já que o molde silábico e as condições de boa-formação dessa língua são diferentes e admitem sequências de segmentos que em PB são

¹² Tradução: ‘Um termo usado em Fonética e em Fonologia para se referir a um tipo de intrusão, na qual houve a inserção de um som adicional em uma palavra; geralmente é subclassificada em prótese e anaptixe.’

evitadas.¹³ A realização da epêntese na interlíngua parece indicar, portanto, que o processo de aquisição do inglês como L2 faz uso das mesmas condições de boa-formação para sílaba presentes na L1 (PB).

O foco de investigação deste estudo é, pois, a inserção vocálica que sincronicamente ocorre no interior de vocábulos (*suarabácti*) do PB e do inglês como LE, seja dentro de uma base ou na fronteira entre um prefixo e uma base. Mais especificamente, o estudo busca verificar os fatores linguísticos e extralinguísticos que estão relacionados a esse tipo de epêntese, bem como procura identificar semelhanças e diferenças entre a epêntese que ocorre em PB e a que ocorre na aquisição de inglês como LE por brasileiros.

CAMARA JR. (1970) chama atenção para o fato de que, na língua portuguesa, a epêntese se restringe a vocábulos de origem erudita, introduzidos através da língua escrita a partir do séc. XV com empréstimos do latim clássico.¹⁴ Palavras como *fictício*, *óbvio*, *ritmo* e *afta* apresentam na grafia uma plosiva ou uma fricativa labial seguida imediatamente por uma plosiva, uma fricativa labial ou uma nasal, sendo ambas consoantes pronunciadas. No caso dos encontros consonantais que ocorrem após a sílaba tônica, o autor iguala a duração da vogal epentética à de uma vogal postônica não final de um vocábulo proparoxítono, como em *rpto* e *rápido*, que, segundo o autor, se diferem em pronúncia apenas com relação ao vozeamento da última consoante de cada palavra (surda no primeiro caso e sonora no segundo).

Camara Jr. defende que a duração das vogais epentéticas não deve ser desconsiderada para efeitos da análise das estruturas da sílaba em português, exemplificando com argumento de que, em vocábulos em que há epêntese, como em *D[i]jalma*, o que ocorre é que a primeira consoante do *cluster* passa a constituir em português uma sílaba distinta (cf. CAMARA JR., 1969).

Essa observação feita por Camara Jr. levou linguistas a analisar mais detidamente o modo como se configura a epêntese medial na subjacência da língua portuguesa. Nesse sentido, COLLISCHONN (1997) propõe, com base nos trabalhos de ITÔ (1986) e de PIGOTT (1995) (que acrescenta moras à representação silábica), que a epêntese tem início no

¹³ KATAMBA & STONHAM (2006:84) observam, no entanto, que pode haver epêntese em certos encontros consonantais de palavras estrangeiras emprestadas pelo inglês. Exemplos disso estão nos topônimos *Tblisi* e *Zgusta*, que apresentam os *clusters* [tbl] e [zg], inexistentes nesse idioma. Como esses encontros não respeitam os princípios fonotáticos da língua inglesa, eles tendem a ser desfeitos pelos falantes, resultando nas pronúncias [təblisi] e [zəgustə], respectivamente.

¹⁴ Sabe-se, no entanto, que a epêntese em PB é também resultante de empréstimos recentes de outras línguas modernas, como o inglês (cf. *clipe*).

componente lexical do PB, como resultado do processo de silabação. Durante a silabação, uma consoante não apta a ocupar uma posição de ataque ou de coda (ou seja, que desrespeita as condições de boa formação da língua) não é ligada a nenhum nó silábico. A presença dessa consoante perdida na representação fonológica provoca a criação de uma sílaba estrutural, desprovida de núcleo vocálico, que permite a associação da consoante na posição de ataque. O núcleo dessa sílaba é preenchido por com uma vogal e a mora correspondente apenas no pós-léxico. A ilustração em [16] demonstra isso.

[16]

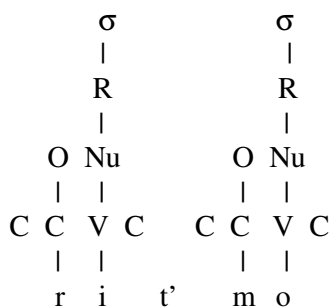
Léxico				
Forma subjacente	Silabação	Criação de sílaba sem núcleo	Acento	
a f t a →	$\begin{array}{cc} \sigma & \sigma \\ & / \\ \mu & \mu \\ & \\ a & t a \end{array}$	$\begin{array}{ccc} \sigma & \sigma & \sigma \\ & / & / \\ \mu & \mu & \mu \\ & & \\ a & f \bullet & t a \end{array}$	(* .)	
	a f t a →	a f • t a →	a f • t a →	

Pós-léxico
Inserção de vogal (* . .)
$\begin{array}{ccc} \sigma & \sigma & \sigma \\ & / & / \\ \mu & \mu & \mu \\ & & \\ a & f \text{ I} & t a \end{array}$

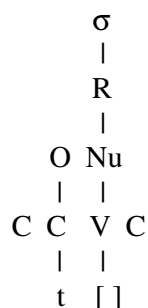
Extraído de COLLISCHONN (2004:63)

BISOL (1999) admite que a epêntese constitui parte da silabação em PB e está presente em todos os níveis lexicais, assim como no pós-léxico. Nessa perspectiva, a epêntese medial é vista como o preenchimento de nós vocálicos vazios que são criados durante a silabação iterativa. Esses nós vocálicos vazios correspondem a núcleos de sílabas que são criadas para salvar consoantes extraviadas, i.é, consoantes inicialmente dessilabadas porque violam as condições de boa-formação de *onset* ou de coda silábica. A partir do padrão canônico CCVC, Bisol representa a epêntese que ocorre em *ritmo* da seguinte forma:

[17] Primeira iteração



Segunda iteração



CV CV CV
Resultado: r i t i m o

Adaptado de BISOL (1999:729-730)

De acordo com essa representação, portanto, a epêntese (entendida como processo) coloca uma consoante extraviada t'^{15} junto a um *slot* V do padrão canônico, formando uma sílaba CV, cujo núcleo, sem traços (cf. [] na segunda iteração), é preenchido posteriormente por regra de redundância (*default*).

Neste trabalho, não é nosso objetivo escolher, dentre as propostas teóricas sobre epêntese em PB, qual delas é a mais apropriada para explicar o fenômeno que ocorre na produção oral dos nativos desse idioma. Nosso foco é, pois, identificar e discutir os possíveis fatores de ordem linguística e extralinguística que se relacionam à epêntese que ocorre no interior de vocábulos do PB e do inglês como segunda língua. Nesse sentido, apresentamos na próxima seção (e subitens) alguns estudos prévios que consideraram a inserção vocálica sob o ponto de vista quantitativo nessas duas línguas.

1.3.4.1 Análises sobre o fenômeno variável de epêntese

Em português brasileiro (PB), existe um número considerável de estudos que focalizam o fenômeno de epêntese, especialmente em dados de fala da Região Sul do Brasil (cf. COLLISCHONN, 1996, 2000, 2002, 2003, 2004, e KINZEL, 2000). Nesta pesquisa, tomou-se como um dos pontos de partida o trabalho de análise quantitativa de COLLISCHONN (2002) (cf. seção 1.3.4.1.1), que investigou a epêntese medial atestada no

¹⁵ Note-se que a consoante t é extraviada porque, na primeira iteração, não satisfaz a condição do ataque (*ri.tmo) nem a condição da coda (*rit.mo) do PB.

português da Região Sul do Brasil sob a perspectiva variacionista (com base na teoria derivacional da sílaba de ITÔ, 1986, e do trabalho de PIGOTT, 1995).

Na aquisição do inglês como L2 por brasileiros, a análise do fenômeno de epêntese igualmente já foi feita em diversos estudos (cf. REBELLO, 1997; CAGLIARI, 2001; KOERICH, 2002; CORNELIAN JR., 2003; ALVES, 2004; CARDOSO, 2007; ALVES, 2008; PEREYRON, 2008). Desses estudos, no entanto, somente o de PEREYRON (2008) investigou a epêntese que ocorre no meio de um vocábulo (os demais tratam da epêntese em início ou em final de palavra). O trabalho de PEREYRON (2008) (cf. seção 1.3.4.1.2), portanto, constituiu o segundo ponto de partida para o desenvolvimento desta pesquisa, principalmente porque ele inicia uma discussão sobre um aspecto linguístico (qual seja, a epêntese medial na aquisição de inglês como L2 por brasileiros) que merece ser aprofundada.

Nas próximas subseções, resumimos as investigações de COLLISCHONN (2002), sobre a epêntese em PB, e de PEREYRON (2008), sobre a epêntese em inglês. Como foi dito, tais trabalhos serviram de base para o desenvolvimento desta pesquisa e, portanto, merecem ser detalhados.

1.3.4.1.1 A epêntese em PB: Collischonn (2002)

COLLISCHONN (2002) apresenta um estudo sobre a realização variável da epêntese vocálica no Sul do Brasil. A partir dos encontros consonantais apontados por CAGLIARI (1981) como favorecedores da epêntese no português brasileiro, a autora investiga a ocorrência do fenômeno em dados de fala dos municípios de Porto Alegre, Florianópolis e Curitiba, que integram o projeto Variação Linguística Urbana no Sul do País - VARSUL.¹⁶

O tratamento teórico dado ao fenômeno nesse estudo se baseia na teoria derivacional da sílaba apresentada por ITÔ (1986) e no trabalho de PIGOTT (1995). COLLISCHONN (2002) assume, a partir desses estudos, que a epêntese ocorre no componente lexical da fonologia do português brasileiro, sendo resultante do processo de silabação. Uma consoante não apta a ocupar uma posição silábica de ataque ou de coda não seria associada a nenhum nó silábico, mas motivaria a criação de uma sílaba estrutural, desprovida de núcleo vocálico, o que permitiria que tal consoante passasse a ocupar o ataque dessa nova sílaba. Apenas no pós-léxico a posição de núcleo vocálico seria preenchida.

¹⁶ VARSUL é um projeto que tem por objetivo geral a descrição do português falado na Região Sul do Brasil. Mais informações sobre o projeto podem ser obtidas em <http://www.pucrs.br/fale/pos/varsul/>.

O levantamento dos dados de COLLISCHONN (2002), que seguiu os moldes metodológicos da pesquisa sociolinguística, levou em consideração as seguintes variáveis independentes como potenciais fatores relacionados à epêntese:

I - Variáveis linguísticas:

- a) posição da consoante perdida – inicial (*psicólogo*) ou medial (*opção*);
- b) tipo de consoante perdida – oclusiva labial (*pneu*), oclusiva alveolar (*atmosfera*), oclusiva velar (*acne*), fricativa labial ou palatal (*afta*) ou nasal labial (*amnésia*);
- c) contexto seguinte à consoante perdida – oclusiva nasal (*hipnose*), oclusiva não nasal (*espectro*), fricativa sibilante (*Pepsi*) ou fricativa não sibilante (*advogado*).
- d) posição da consoante perdida em relação à sílaba tônica – pretônica (*objetivo*) ou postônica (*técnica*);
- e) posição do vocábulo no grupo de força (o ponto em que a palavra-alvo se encontra no espaço entre duas pausas – inicial, medial ou final);
- f) velocidade da fala – rápida ou normal;
- g) origem do vocábulo – estrangeiro (*partner*) ou nativo (*absurdo*).

II - Variáveis extralinguísticas:

- a) sexo – masculino ou feminino;
- b) idade – menos de 50 anos ou mais de 50 anos;
- c) escolaridade – primário, ginásio ou colegial;
- d) grupo geográfico – Porto Alegre, Florianópolis ou Curitiba.

As principais hipóteses apresentadas no trabalho de COLLISCHONN (2002) foram as seguintes:

(i) ‘escolaridade’: informantes com maior nível de instrução devem realizar menos epêntese, uma vez que possuem maior contato com a escrita, que não registra a vogal epentética;

(ii) ‘posição da consoante perdida’: por analogia ao fato de que, em posição final de palavra do PB, a vogal sempre se realiza (cf. CÂMARA JR., p. 28), a epêntese deve ser igualmente frequente na extremidade inicial de palavra, com índices de aplicação mais elevados do que no meio de palavra;

(iii) ‘contexto seguinte à consoante perdida’: a epêntese deve ser desfavorecida nos casos em que a consoante seguinte é uma fricativa sibilante ([s] ou [z]), dada a possibilidade de essa sibilante formar com a consoante perdida uma africada na realização fonética (como em *o[p_s]ional*), e mais favorecida diante de uma oclusiva (como em *espec[t]ro*) ou de uma nasal (como em *hip[n]ose*);

(iv) ‘posição da consoante perdida em relação à sílaba tônica’: com base na afirmação de CÂMARA JR. (1969, p. 28) de que, na posição postônica, a vogal se reduz (como comprovam os vocábulos *rapto* e *rápido*, que possuem pauta prosódica idêntica e se diferenciam apenas pelo vozeamento da consoante seguinte – [t] ou [d]), a epêntese deve ocorrer mais frequentemente no contexto pretônico;

(v) ‘origem do vocábulo’: as palavras já integradas ao português devem apresentar a vogal epentética em maior número do que as estrangeiras, visto que, na literatura sobre o comportamento dos empréstimos, considera-se que o falante tem um cuidado maior na pronúncia de palavras que ele sabe não serem nativas de sua língua, o que sugere que a epêntese seja evitada nesse tipo de palavra.

Os dados obtidos em COLLISCHONN (2002) foram submetidos a análise estatística através dos programas computacionais do pacote VARBRUL. A frequência global de aplicação da epêntese foi de 52% e os grupos de fatores selecionados foram, respectivamente, ‘posição da consoante seguinte em relação à sílaba tônica’, ‘tipo de consoante seguinte’, ‘grupo geográfico’ e ‘tipo de consoante perdida’.

Com esse resultado, algumas das hipóteses apresentadas por COLLISCHONN (2002) – (i), (ii) e (v) – nem chegaram a ser discutidas em seu estudo, já que envolviam fatores considerados não significativos nas rodadas estatísticas. No que diz respeito aos grupos considerados relevantes, foram apresentadas as explicações descritas abaixo.

O fator ‘posição da consoante perdida em relação à sílaba tônica’, considerado o mais significativo na realização da epêntese, teve o contexto pretônico (como em *objeto*) como o que mais favoreceu sua aplicação: o peso relativo¹⁷ para essa categoria foi de 0,68, ao passo que, para o contexto postônico (como em *ético*), o valor foi de 0,24.

Para explicar esse resultado, que mostra a visível relação da epêntese com o acento, COLLISCHONN (2002) primeiramente argumentou não ser possível a interpretação desses dados através da teoria de ITÔ (1986) porque, nessa proposta, a silabação (na qual ocorre a criação da sílaba estrutural sem núcleo, a ser preenchido por epêntese no pós-léxico) ocorre antes da acentuação, ou seja, num momento em que ainda não estão disponíveis as informações sobre o acento. A partir disso, a autora defende que, com base na Teoria da Otimidade (MCCARTHY & PRINCE, 1993; PRINCE & SMOLENSKY, 1993), é possível demonstrar que, para a posição postônica de vocábulos do PB, existe um conjunto de restrições de acento que domina um conjunto de restrições de estrutura silábica, impedindo-se, em razão disso, a aplicação da epêntese nos casos em que o acento ficaria depois da terceira sílaba, a contar da borda final da palavra (como em ?[‘e.t[ɪ.ni.ku], *ético*).¹⁸

O segundo fator selecionado como relevante no estudo de COLLISCHONN (2002) foi ‘contexto seguinte à consoante perdida’. Para esse fator, as sibilantes [s] e [z] foram as que menos favoreceram a epêntese (peso relativo de 0,32), o que corroborou a hipótese de que tais fricativas possam formar com a oclusiva precedente uma espécie de africada, posicionada no ataque da sílaba (como em [‘pe.p_sɪ], *Pepsi*).

A categoria de consoante seguinte que mais demonstrou favorecer a epêntese foi ‘fricativa não-sibilante’ (como em *advogado*), com peso relativo 0,81. A justificativa apresentada no estudo foi a de que 100% dos dados que possuíam consoante seguinte correspondente a uma fricativa não sibilante continham o *cluster* em uma posição pretônica (posição francamente favorável à epêntese).

O terceiro fator selecionado - o único de caráter extralinguístico¹⁹ - foi ‘grupo geográfico’. Os resultados encontrados para esse fator revelam que os falantes de Porto

¹⁷ O peso relativo, na definição de GUY (1998:31), corresponde a um número entre zero e um que caracteriza o efeito de um fator sobre a variável analisada (neste caso, a epêntese). Quanto mais próximo de 1,00, maior a chance de a regra se aplicar quando esse fator estiver presente no contexto. Mais informações sobre o peso relativo são apresentadas em 2.3.

¹⁸ COLLISCHONN (2000) e COLLISCHONN (2004) apresentam uma discussão mais aprofundada sobre a realização da epêntese em PB sob a perspectiva otimalista.

¹⁹ COLLISCHONN (2003) aprofunda a discussão sobre o papel de variáveis extralinguísticas na aplicação da epêntese em dados do Sul do Brasil.

Alegre aplicam mais a epêntese do que os das outras duas capitais da Região Sul: o peso relativo para os dados da capital gaúcha foi de 0,63, enquanto os pesos para os dados de Florianópolis e de Curitiba foram de 0,39 e 0,46, respectivamente.

No que diz respeito ao fator ‘tipo de consoante perdida’, o último selecionado nas análises estatísticas do estudo de COLLISCHONN (2002), concluiu-se que as oclusivas alveolares (/t/ e /d/) formam as piores codas no PB, o que foi atestado pelo peso relativo de 0,71 para a categoria, que indica seu favorecimento à epêntese. Numa tabulação cruzada entre o ‘tipo de consoante perdida’ (labial, alveolar ou velar) e ‘contexto seguinte à consoante perdida’ (fricativa não-sibilante, nasal, oclusiva ou fricativa sibilante), mostrou-se que a epêntese após /t/ e /d/ só é menos frequente (peso relativo 0,52)²⁰ diante de [s] ou [z], o que corrobora, segundo a autora, a tese de que, nesses casos, a consoante perdida e a consoante seguinte tendem a se juntar em uma mesma sílaba, formando o ataque dessa sílaba sob a forma de uma africada (como em [fo.t_saw], *futsal*) e, conseqüentemente, tornando desnecessária a aplicação da epêntese.

O estudo de COLLISCHONN (2002) constitui uma das principais motivações para o desenvolvimento do presente estudo. A autora desconsiderou em sua análise “(...) casos como *subtenente, submarino, subdiretora, subdivisão, suboficial* (...)” porque, segundo ela, “(...) existe a suspeita de que o prefixo atue com uma certa independência fonológica e, nesse caso, a consoante /b/ estará em final de palavra, (...)”. A partir dessa suspeita apresentada pela autora, o presente trabalho procura confirmar, entre outras hipóteses levantadas, se o contexto morfológico de fronteira entre alguns prefixos e suas bases pode ser igualado ao contexto de final de palavra, caso em que a ocorrência de fenômenos como a epêntese seriam mais frequentes.

1.3.4.1.2 A epêntese na aquisição do inglês como LE (Pereyron, 2008)

O estudo de PEREYRON (2008) apresenta uma análise sobre a inserção vocálica em encontros consonantais mediais em dados de fala da língua inglesa produzidos por falantes brasileiros aprendizes dessa língua. A partir da investigação de COLLISCHONN (2002) (cf. seção 1.3.4.1.1), referente à epêntese em português brasileiro (PB), a autora busca verificar se

²⁰ Diante de nasal, de oclusiva e de fricativa não sibilante, os pesos relativos para a aplicação de epêntese após /t/ e /d/ foram de 0,73, 0,75 e 0,82, respectivamente.

os condicionadores que atuam na regra variável de epêntese no PB são igualmente relevantes em dados do inglês como língua estrangeira (LE).

PEREYRON (2008) se baseia na proposta de constituição da sílaba apresentada por SELKIRK (1982), que defende a existência de uma estrutura hierárquica interna (ataque e rima – que, por sua vez, se subdivide em núcleo e coda) para o constituinte prosódico σ , que integra uma estrutura hierárquica maior.

Também baseada em SELKIRK (1982), a autora admite a existência de moldes silábicos (*templates*) para as línguas, através dos quais são especificados, por meio de restrições, os diferentes tipos de segmentos admitidos em cada posição dentro da sílaba (ataque, núcleo e coda) de cada língua. Com base nesses moldes, a autora mostra que determinados segmentos admitidos na coda de sílabas em inglês não são permitidos nas sílabas do português e argumenta que, por tal razão, os falantes brasileiros aprendizes de inglês buscam simplificar tipos silábicos que envolvam tais segmentos.

PEREYRON (2008) equiparou a ocorrência da epêntese vocálica na fala de brasileiros aprendizes de inglês com a aplicação desse fenômeno em dados do português, a partir dos resultados obtidos por COLLISCHONN (2002). A autora, que assume que a epêntese no inglês falado por brasileiros é um fenômeno variável (p. 65), segue em seu trabalho fundamentalmente a linha teórico-metodológica sociolinguística introduzida por William Labov.

O trabalho de PEREYRON (2008) também é amparado pela abordagem psicolinguística de PRESTON (2004). Esse autor propõe um modelo linguístico que admite a existência de um *vernáculo* (aprendido nos primeiros anos de vida através da interação com os pais, outras crianças ou em qualquer contexto que não empregue a língua formal), e de um *pós-vernáculo* (que abrange o que é aprendido depois do vernáculo, seja na L1 ou em uma L2). Com base nessa divisão, PEREYRON (2008) argumentou que “características fonológicas, morfológicas, sintáticas, etc., são adquiridas substancialmente da mesma forma por adultos ao aprenderem o pós-vernáculo de sua língua materna e ao aprenderem uma L2” (p. 71). A partir disso, a autora buscou investigar se, para uma mesma regra variável (no caso, a epêntese), a escolha que o falante faz entre as formas em variação em sua L1 é condicionada pelos mesmos fatores linguísticos e extralinguísticos que influenciam a escolha entre as formas em variação em sua L2.

A metodologia adotada por PEREYRON (2008) envolveu a seleção de 16 aprendizes de língua inglesa como LE do município de Porto Alegre, que procederam à leitura (gravada) de uma lista de palavras e de uma lista de frases em inglês. A identificação da epêntese foi feita através de duas formas de análise: uma perceptual (feita pela simples audição dos arquivos de áudio) e outra acústica (executada com uso do *software* Praat – v.4.5.24 –, que possibilita a identificação visual da epêntese, cf. seção 2.3).

As variáveis independentes consideradas pela autora foram as seguintes:

I - Variáveis linguísticas:

a) qualidade da vogal epentética – [i], [ə], [o] ou não introdução da vogal;

b) tipo de consoante perdida ou contexto precedente – oclusiva bilabial [b] (*absent*), oclusiva bilabial [p] (*aptitude*), oclusiva alveolar [d] (*admire*), oclusiva alveolar [t] (*atmospheric*), oclusiva velar [g], (*magnet*), oclusiva velar [k] (*acne*), fricativa labiodental [f] (*fifty*), fricativa alveopalatal [ʃ] (*pashmina*), nasal bilabial [m] (*gymnastics*), nasal alveolar [n] (*enmity*), fricativa alveolar [z] ([fʌzbi], ocorrência por motivo de alteração da palavra de origem, *fubsy*), africada alveopalatal [dʒ] (*admire*, ocorrência decorrente da palatalização da consoante alveolar [d]), africada alveopalatal [tʃ] (*atmospheric*, ocorrência decorrente da palatalização da consoante alveolar [t]) ou apagamento da consoante perdida (como em [ˈemiti] *enmity*);

c) vozeamento da consoante perdida – vozeado (*obtain*), desvozeado (*acne*) ou apagamento da consoante perdida;

d) contexto seguinte à consoante perdida ou à vogal epentética – oclusiva bilabial [b] (*goodbye*), oclusiva bilabial [p] (*webpage*), oclusiva alveolar [d] (*abdominal*), oclusiva alveolar [t] (*aptitude*), oclusiva velar [k] (*babka*), africada alveopalatal [dʒ] (*object*), africada alveopalatal [tʃ] (*picture*), fricativa labiodental [v] (*advantage*), fricativa alveolar [z] (*fubsy*), fricativa alveolar [s] (*adsorb*), fricativa alveopalatal [ʃ] (*functionality*), nasal bilabial [m] (*chipmunk*), nasal alveolar [n] (*obnoxious*), lateral alveolar [l] (*sublicense*), fricativa alveopalatal [ʒ] (ocorrência decorrente da substituição da africada [dʒ] em *object*) ou apagamento do contexto seguinte (como em [viətəˈmiːz] *Vietnamese*);

e) vozeamento do contexto seguinte – vozeado (*abdicate*), desvozeado (*napkin*) ou apagamento do contexto seguinte;

f) acento – pré-tônico (a consoante perdida antecede a sílaba tônica, como em [æbsə'lu:tli] *absolutely*) ou pós-tônico (a consoante perdida ocorre após a sílaba tônica, como em [pa:rt'nər] *partner*);

g) tipo de *cluster* – encontros consonantais comuns ao português e ao inglês (como [kn], presente tanto em *acne* como em *acknowledge*) ou encontros consonantais que pertencem somente ao inglês (como [pk] em *napkin*);

h) alteração na produção do falante – *cluster* modificado (mudança na consoante perdida, no contexto seguinte ou em ambos, como na inversão em [fʌzbi] *fubsy*), acento modificado (quando o falante produz uma palavra pré-tônica como pós-tônica ou vice-versa, como em [pa:rt'nər] ao invés de [pa:rt'nər] *partner*), duas modificações (ocorrência simultânea de *cluster* modificado e de acento modificado, como em [kæptʊr] ao invés de [kæptʊr] *capture*) ou ‘esperado’, quando nenhuma alteração ocorre.

II - Variáveis extralinguísticas:

- a) sexo – masculino ou feminino;
- b) idade – de 15 a 34 anos ou de 35 a 57 anos;
- c) nível de proficiência – básico (até quatro anos de estudo de inglês em escola de idiomas) ou avançado (mais de quatro anos de estudo de inglês);
- d) tipo de instrumento – lista de palavras ou lista de frases;
- e) informante – informante 1 a informante 16.

As hipóteses formuladas por PEREYRON (2008) para suas variáveis se fundamentaram principalmente nas interpretações que COLLISCHONN (2002) fez sobre os resultados que obteve em seu estudo (cf. seção 1.3.4.1.1). Para as variáveis não presentes em COLLISCHONN (2002), as hipóteses apresentadas pela autora foram as seguintes:

(i) ‘nível de proficiência’: alunos com mais anos de estudo da língua inglesa aplicam menos a regra de epêntese devido ao maior contato com a língua escrita;

(ii) ‘tipo de instrumento’: com base na afirmação de LABOV (1972) de que a leitura de listas de palavras implica uma fala cuidada por parte do entrevistado, pressupõe-se que a

taxa de aplicação da regra de epêntese seja mais baixa nos dados provenientes desse tipo de instrumento em comparação à lista de frases;

(iii) ‘vozeamento da consoante perdida’: partindo-se da constatação feita em CAGLIARI (2001), consoantes perdidas vozeadas devem condicionar a epêntese com taxas mais elevadas do que as consoantes não vozeadas, uma vez que a maior vibração das cordas vocais facilitaria o surgimento de uma vogal, que compartilha com a consoante tal traço;

(iv) ‘vozeamento do contexto seguinte’: à semelhança da hipótese para ‘vozeamento da consoante perdida’, os segmentos vozeados devem condicionar a aplicação da vogal mais frequentemente do que os não vozeados;

(v) ‘tipo de *cluster*’: os *clusters* comuns ao português e ao inglês devem condicionar mais a aplicação da regra, visto que os encontros do inglês devem ser mais observados pelos falantes, implicando menor taxa de epêntese nesses contextos.

Os dados coletados foram codificados e submetidos a análise estatística por intermédio dos programas computacionais do pacote VARBRUL 2S. Foram realizadas três rodadas para cada uma das duas formas de análise realizadas (acústica e perceptual) como forma de evitar a apreciação concomitante de variáveis com baixa ortogonalidade entre si.²¹ Os principais resultados obtidos por PEREYRON (2008) estão descritos abaixo.

No que diz respeito à frequência global de aplicação da epêntese, a taxa verificada na análise acústica (com uso do *software* Praat) foi de 8%, ao passo que, na análise perceptual, foi de 33%. Pereyron afirma que o fato de ter registrado maior frequência de vogal epentética na análise perceptual “(...) parece indicar que o sistema auditivo é sensível a essa.” (p. 100).

Acreditamos que a afirmação da autora vai ao encontro da crença compartilhada entre alguns linguistas de que um fenômeno como a epêntese, ao ser realizado em uma língua – não nativa ao falante – na qual não deveria ocorrer, é possivelmente resultante do que denominam ‘percepção’, isto é, um processo inferencial que o falante realiza através da constante equiparação dos *inputs* da LE com os dados de sua língua materna. Nessa perspectiva, a vogal epentética, ainda que não esteja no *input*, é ilusoriamente percebida pelo falante nos contextos de *clusters* consonantais que não são admitidos na língua materna.

²¹ A definição de ortogonalidade é apresentada na seção 2.3.

As rodadas estatísticas do trabalho de PEREYRON (2008) não julgaram relevante o papel das variáveis ‘alteração na produção do falante’, ‘acento’ e ‘tipo de instrumento’. Por tal razão, a hipótese (ii) anteriormente citada não pôde ser confirmada pela autora.²² Com relação aos resultados referentes às variáveis consideradas estatisticamente significativas no estudo de PEREYRON (2008), resumem-se, nos próximos oito parágrafos, as principais considerações feitas pela autora.²³

‘Vozeamento da consoante perdida’. O fator ‘vozeado’ (*obtain*) apresentou pesos relativos de 0,62 (análise acústica) e 0,52 (análise perceptual), enquanto o fator não vozeado (*aptitude*) teve pesos iguais a 0,33 e 0,48 nessas duas análises. Tal resultado, que indica maior aplicação de epêntese após um contexto vozeado, foi ao encontro da hipótese (iii) de que consoantes vozeadas, por compartilharem a configuração [+ vozeado] com as vogais, estão mais intimamente relacionadas ao fenômeno de epêntese vocálica.

‘Vozeamento do contexto seguinte’. A epêntese é mais uma vez favorecida pelo fator ‘vozeado’ (*grapnel*), que apresentou pesos relativos iguais a 0,71 e 0,53 nas análises acústica e perceptual, respectivamente. Contrastando esses valores com os pesos relativos de 0,22 e de 0,46 para o contexto seguinte desvozeado (*capture*), confirmou-se novamente – cf. a hipótese (iv) – que a informação sobre o vozeamento dos segmentos que ladeiam o ponto em que pode ocorrer a inserção da vogal tem relevância para a efetiva aplicação do fenômeno.

‘Tipo de *cluster*’. Os encontros consonantais comuns ao português e ao inglês (como [pt] em *adaptation* e em *apto*) mostraram favorecer a aplicação da epêntese quando comparados com *clusters* existentes somente em inglês (como [nm] em *enmity*). Os pesos relativos para a primeira categoria foram de 0,54 e de 0,58 (análise acústica e perceptual, respectivamente). Esses valores, embora estejam próximos do ponto neutro (0,50), foram considerados significativos quando foram contrastados com os pesos desfavorecedores de 0,41 e de 0,34 nas mesmas análises para a segunda categoria. Com isso, confirmou-se a hipótese (v) supra.

²² Tampouco pôde ser confirmada sua hipótese de que, à semelhança dos resultados de COLLISCHONN (2002), a posição postônica da consoante perdida inibiria a epêntese.

²³ As variáveis ‘nível de proficiência’ e ‘sexo’ foram consideradas não significativas em metade das rodadas realizadas e significativas na outra metade (‘nível de proficiência’ foi excluída nas três rodadas da análise acústica e incluída nas três da análise oitiva, enquanto ‘sexo’ foi excluída nas três rodadas da análise oitiva e incluída nas três da análise acústica). Em razão disso, a autora julgou importante incluir tais variáveis na discussão dos resultados.

‘Consoante perdida’. A consoante que mais favoreceu a aplicação de epêntese foi a velar [g] (peso relativo igual a 0,84 para a análise acústica e igual a 0,83 para a análise perceptual). Tal resultado foi de encontro à hipótese de que, assim como foi para COLLISCHONN (2002), as velares deveriam ser as menos favorecedoras da inserção vocálica. A justificativa apresentada para o resultado foi a de que 100% dos dados com consoante perdida [g] tinham como contexto seguinte uma nasal, formando, portanto, o *cluster* ‘velar + nasal’, apontado igualmente no estudo de COLLISCHONN (2002) como motivador de epêntese (peso relativo 0,79). A velar [k] apresentou comportamento diverso, mostrando-se a maior inibidora da epêntese na análise acústica (peso de 0,19) e uma das motivadoras do fenômeno na análise oitiva (peso de 0,68). O amálgama das consoantes perdidas nos grupos ‘labial’, ‘alveolar’ e ‘velar’ revelou que as alveolares são as que menos favorecem a epêntese (pesos de 0,38 e 0,17 para as análises acústica e perceptual). A conclusão apresentada foi a de que o valor não marcado de Ponto de Consoante na coda, atribuído por COLLISCHONN (2002) às velares, é, na interlíngua, atribuído às alveolares.

‘Contexto seguinte’. Essa foi a primeira variável a ser selecionada (tanto na análise acústica quanto na perceptual) nas rodadas em que foi incluída. A fricativa sibilante [s] (*absent*) teve comportamento divergente nas duas análises: peso relativo de 0,13 na análise acústica (o menor de todos) e de 0,72 na análise oitiva (o maior de todos). Tal resultado foi justificado com a explicação de que, por questões de espraiamento de traços, tanto o contóide [s] como o vocóide [i] apresentam o traço coronal, e, com isso, “(...) é possível entender que pela análise perceptual tenha se registrado a presença da vogal [i] antes de [s], uma vez que a teoria²⁴ atribui ao conjunto de traços articulatorios a explicação de interação entre consoantes e vogais.” (p. 129). A autora concluiu, após a interpretação individual dos resultados da análise acústica para os vocábulos com os *clusters* [bs], [ps], [ds] e [ks], que a hipótese de que a sibilante tende a formar uma africada com a consoante oclusiva que a antecede é confirmada nos dados do inglês. Quanto à fricativa não sibilante [v] (*obvious*), apontada por COLLISCHONN (2002) como maior favorecedora da epêntese, seus pesos relativos de 0,53 e de 0,40 impediram afirmar que em inglês ela também é motivadora do fenômeno. A nasal [n] (*kidney*), com pesos de 0,81 e 0,71, foi a única a mostrar valores aproximados aos de COLLISCHONN (2002), em que a consoante foi o segundo tipo de contexto seguinte que mais favoreceu a inserção vocálica (peso relativo 0,74).

²⁴ A teoria referida é a Geometria de Traços, proposta por Clements e Hume (1995).

‘Idade’. De acordo com o resultado para a variável, ficou comprovado que falantes de inglês como LE com idade superior a 35 anos realizam mais epêntese do que aqueles com idade entre 15 e 34 anos. Na análise acústica, os pesos relativos foram de 0,67 para os mais velhos e de 0,33 para os mais novos. Resultados semelhantes foram obtidos na análise perceptual, muito embora a diferença entre os dois fatores tenha diminuído: peso relativo de 0,57 para o fator ‘35 a 57 anos’ e de 0,43 para o fator ‘15 a 34 anos’.

‘Nível de proficiência’. Essa variável foi selecionada somente nas rodadas da análise perceptual. O fator ‘básico’ (menos de 4 anos de estudo) obteve peso relativo igual a 0,53, enquanto o fator ‘avançado’ (mais de 4 anos de estudo) apresentou peso relativo igual a 0,47. A autora confirmou, assim, a hipótese de que os informantes com menos anos de estudo de inglês como língua estrangeira aplicariam mais frequentemente a regra de epêntese, muito embora tenha reconhecido que tais valores estão muito próximos ao ponto de referência (0,50).

‘Sexo’. Essa foi a segunda variável social mais relevante nas rodadas da análise acústica (a primeira foi ‘idade’), muito embora tenha sido excluída pelo programa estatístico na análise perceptual. Para as mulheres, a aplicação da epêntese foi mais frequente, com peso relativo de 0,58, ao passo que, para os homens, o peso relativo foi igual a 0,42. O resultado, que foi contrário à hipótese de que as mulheres realizariam menos epêntese, foi justificado com o cruzamento dos fatores ‘sexo’ e ‘idade’, através do qual se constatou que as mulheres mais velhas eram as que provocavam maiores índices de aplicação da epêntese.

A análise geral dos resultados encontrados em PEREYRON (2008) nos permite deduzir que não é possível estabelecer uma relação de biunivocidade entre os fatores envolvidos na epêntese em inglês (realizada por aprendizes brasileiros da língua) e os envolvidos na epêntese em PB (realizada por falantes nativos). Isso pode ser um indício de que a alternância entre a inserção e a não inserção da vogal na interlíngua português-inglês deva ser tratada de modo diverso à variação que ocorre entre os falantes nativos do PB.

Com base nos trabalhos de COLLISCHONN (2002) e de PEREYRON (2008), a presente investigação busca confirmar se de fato a epêntese em PB e a epêntese na aquisição de inglês como LE se comportam de formas distintas. Ademais, este trabalho acrescenta aos estudos sobre a epêntese vocálica a discussão sobre o papel do contexto morfológico em que ela pode ser inserida. No próximo capítulo, Metodologia, são apresentados os instrumentos

utilizados para investigar o fenômeno de inserção vocálica em dados do PB e do inglês, assim como são relacionadas as hipóteses para este trabalho.

2 METODOLOGIA

Este capítulo se preocupa com a descrição dos elementos de ordem metodológica que foram empregados para a realização do estudo de investigação da epêntese medial em dados de fala do português brasileiro (PB) e do inglês.

Na seção 2.1, são apresentados os procedimentos adotados para a elaboração e aplicação de cada um dos quatro instrumentos que foram utilizados para coletar os dados aqui investigados. Além disso, são especificados os critérios adotados na delimitação dos informantes que participaram de cada uma das etapas da pesquisa, bem como são exibidas as células sociais formadas para a obtenção dos dados de fala através da Amostra de Fala 1 (referente ao PB) e da Amostra de Fala 2 (referente ao inglês), principais instrumentos de análise neste trabalho.

Na seção 2.2, são identificados os fatores linguísticos e extralinguísticos que foram considerados relevantes para a análise do fenômeno de epêntese tanto em PB quanto em inglês, de acordo com os objetivos propostos para a coleta da Amostra de Fala 1 e da Amostra de Fala 2. Deve-se deixar claro que a inserção vocálica que ocorre na interlíngua não é tratada aqui como um fenômeno variável tal como de fato o é na língua materna, já que, nos dados do inglês, ela não corresponde a uma mudança linguística em curso que possa implicar uma evolução da língua (como no PB), mas, sim, ao reflexo de que as condições de sílaba particulares da língua materna estão sendo transferidas para a interlíngua do indivíduo (cf. seção 1.1). Optou-se, ainda assim, por investigar a epêntese em inglês por meio dos mesmos padrões metodológicos adotados nas análises variacionistas que tratam de fenômenos legitimamente sujeitos aos preceitos da Teoria da Variação.

Em 2.3, finalmente, são discriminadas as modalidades de análise que foram empregadas no processo de interpretação dos dados coletados nos instrumentos descritos em 2.2.3 e 2.2.4. Para identificar a realização da vogal epentética, fez-se uso de um *software* de análise acústica, ao passo que, para a análise quantitativa das variáveis relacionadas ao fenômeno, foi utilizado um programa computacional de análise estatística.

2.1 Instrumentos para a coleta de dados

Para que se procedesse à coleta dos dados que foram analisados neste estudo, foram elaborados quatro instrumentos. Desses, dois correspondem a testes de transparência (cf. 2.1.1 e 2.1.2), aplicados com o intuito de se verificar até que ponto determinadas palavras, consideradas como prefixadas por alguns gramáticos e dicionaristas da língua portuguesa e da língua inglesa, são igualmente identificadas como tal por falantes do PB e por aprendizes brasileiros de inglês, respectivamente. Tais instrumentos constituem, pois, uma etapa preliminar da pesquisa e servem para consolidar algumas hipóteses a respeito dos contextos morfológicos em nos quais pode ocorrer epêntese.

Os outros dois procedimentos de coleta contemplam o que doravante será denominado Amostra de Fala 1 e Amostra de Fala 2. Através deles, foram realizadas as coletas dos dados de fala do PB (Amostra de Fala 1) e do inglês (Amostra de Fala 2) que serviram para a investigação da epêntese vocálica. Esses instrumentos estão descritos nas seções 2.1.3 e 2.1.4.

2.1.1 Teste de verificação de transparência dos prefixos em PB

O primeiro instrumento (Anexo I) constitui-se de uma lista de 46 palavras pertencentes ao português brasileiro que foi submetida à análise de falantes nativos dessa língua.

O objetivo principal dessa atividade foi investigar o *status* gramatical sincrônico (isto é, o grau de transparência, cf. seção 1.2) de cada um dos prefixos presentes nos vocábulos submetidos à análise dos entrevistados, que tinham como tarefa circular a parte das palavras que possuía o núcleo do significado das mesmas, ou seja, sua raiz ou base (destacando, inconscientemente, seus afixos).

A importância de se verificar a ciência dos falantes sobre a existência dos prefixos investigados é justificada pelo fato de que um dos objetivos deste trabalho é demonstrar que os prefixos transparentes de uma língua apresentam comportamento fonológico diferente de prefixos com menor grau de transparência, o que, neste trabalho, pode significar diferentes taxas de aplicação do fenômeno de epêntese vocálica entre essas duas categorias. Hipotetiza-se que os prefixos considerados transparentes pelos informantes apresentem percentuais mais elevados de aplicação desse fenômeno fonológico se comparados com os prefixos considerados menos transparentes ou opacos.

A hipótese de que prefixos transparentes favorecem a aplicação de um fenômeno como a epêntese se dá pelo fato de se acreditar que esses prefixos apresentem certa independência fonológica, caracterizando-se como prefixos *adjuntos*, em oposição a prefixos *incorporados*, para os quais a epêntese é menos esperada ou inexistente (cf. seção 1.2).

A elaboração do presente teste de transparência foi precedido pela busca de todos os casos de prefixos do PB em que se considerou possível a aplicação da epêntese vocálica. Isso significa dizer que foram selecionados todos os prefixos da língua terminados em consoante obstruinte. A lista dos prefixos investigados está apresentada na Tabela 1, abaixo, e foi obtida a partir do rol de prefixos atestados na língua portuguesa de acordo com a *Nova Gramática do Português Contemporâneo*, de CUNHA & CINTRA (2001) e com a *Moderna Gramática Portuguesa*, de BECHARA (2006).

Tabela 1 - Prefixos relacionados à epêntese vocálica em PB

Origem / Prefixo	Significado	Exemplos	
Latim	<i>ab-</i>	‘distanciamento’; ‘a partir de’; ‘depois de’	<i>abjurar, abdicar</i>
	<i>ad-</i>	‘em direção a’; ‘aproximação’	<i>adjunto, adventício</i>
	<i>ob-</i>	‘diante de’; ‘para’; ‘por causa de’; ‘em troca de’; ‘por’	<i>objeto, obstáculo</i>
	<i>sob-</i> ²⁵	‘em baixo de’; ‘por baixo de’; ‘em consequência de’; ‘em presença de’; ‘perto de’; ‘imediatamente antes de’; ‘durante’; ‘em direção a’; ‘depois de’	<i>sobestar, sobpor</i>
	<i>sub-</i>	‘sotoposição’; ‘ascensão’; ‘subordinação’; ‘substituição’; ‘proximidade’; ‘transmissão’; ‘abrandamento’; ‘ação furtiva’	<i>subalterno, submarino</i>
Grego	<i>ec-</i>	‘fora de’; ‘movimento para fora’	<i>eclipse</i>

Os prefixos acima listados são morfologicamente identificados como tal não somente em CUNHA & CINTRA (2001) e em BECHARA (2006), mas também no Dicionário Aurélio Eletrônico - Século XXI (1999, v. 3.0) e no Dicionário Eletrônico Houaiss da língua portuguesa (2002, v.1.0.5). Através da observação dessa lista, no entanto, levantou-se a suspeita de que a estrutura morfológica de alguns dos exemplos da última coluna da Tabela 1 (igualmente extraídos das gramáticas referidas) se apresentava sincronicamente opaca, ou

²⁵ A forma *sob-* é originária do prefixo *sub-*. No entanto, em virtude de que ambas as formas possuem entradas lexicais distintas nos dicionários consultados neste trabalho, *sob-* e *sub-* são analisadas separadamente no estudo.

seja, poderia não ser identificada pelos falantes nativos do PB. Tal suspeita reforçou a necessidade de se confirmar quais desses prefixos seriam julgados transparentes e quais seriam considerados opacos.

Antes da elaboração do instrumento, procedeu-se, ainda, à conferência da configuração morfológica dos vocábulos exemplificados por CUNHA & CINTRA (2001) e BECHARA (2006) através das informações etimológicas apresentadas pelo Dicionário Eletrônico Houaiss da Língua Portuguesa e pelo Dicionário Aurélio Eletrônico Século XXI (doravante Dicionário Houaiss e Dicionário Aurélio, respectivamente). Verificou-se, a partir disso, que a maioria das 11 palavras listadas (*abjurar, adjunto, adventício, objeto, obstáculo, subalterno, sobestar e eclipse*) já apresentavam, em suas línguas de origem (latim e grego), seus prefixos adjungidos às suas bases. Somente *abdicar, submarino* e *sobpor* são discriminadas morfológicamente pela estrutura prefixo + base.²⁶ Partindo-se do pressuposto de que essa diferença crônica esteja intimamente relacionada à consciência morfológica dos falantes de uma língua, tem-se como hipótese que, dos prefixos da Tabela 1, *sob-* e *sub-* devam ser mais facilmente identificados pelos informantes inquiridos. No que diz respeito aos demais prefixos, preferimos não fazer previsões específicas, ainda que exista a suspeita de que *ec-* seja pouco identificado pelos falantes.

No teste de verificação de transparência elaborado para o PB, trinta das 46 palavras submetidas à análise dos falantes estão registradas no Dicionário Houaiss ou no Dicionário Aurélio como prefixadas. Como demonstra a Tabela 2, esses 30 vocábulos representam o somatório de cinco exemplos para cada um dos seis prefixos do PB que podem motivar a epêntese em PB, à exceção do prefixo *ec-*, para o qual foram listados seis exemplos,²⁷ e do prefixo *ob-*, para o qual selecionaram-se quatro vocábulos.²⁸

²⁶ Com relação ao verbo *abdicar*, apenas um dos dicionários (o Dicionário Houaiss) apresenta a estrutura *ab-* + *dicāre*, muito embora, nesse caso, o sentido de *abdicar* seja ‘tornar público, anunciar’, e não ‘renunciar a direito’. Em razão disso, optou-se por considerar o vocábulo como prefixado, muito embora fosse esperado um baixo índice de reconhecimento desse prefixo pelos informantes.

²⁷ O prefixo *ec-* é de uso bastante restrito na língua portuguesa e, como esclarece o Dicionário Houaiss (no verbete ‘*ec-*’), está presente ‘mormente em cultismos da língua’. Nesse mesmo dicionário também é informado que a consciência do prefixo é inexistente nos usuários em geral. Tais argumentos motivaram o uso de um número maior de palavras com esse prefixo, de modo que se tentasse verificar se a transparência era de fato nula.

²⁸ Originalmente, a lista era composta por cinco vocábulos com o prefixo *ob-*. Constatou-se, após a aplicação do instrumento, que o vocábulo *obturar*, inicialmente considerado como detentor do prefixo *ob-*, não é morfológicamente prefixado, o que motivou sua exclusão desta lista e, automaticamente, sua inclusão na lista de palavras não prefixadas que também integra esse primeiro instrumento (cf. Tabela 3).

Tabela 2 - Vocábulo prefixados submetidos à análise de transparência em PB

Prefixo	Vocábulo prefixados
<i>ab-</i>	<i>abaxial, abdicar, abnegar, ablação, abstenção</i>
<i>ad-</i>	<i>adaptar, advogada, administração, adligação e adjunto</i>
<i>ec-</i>	<i>ecdêmico, eclesiástico, eclético, eclipse, ecnênia, eczema</i>
<i>ob-</i>	<i>obumbrar, obnubilar, obviar, obstruir</i>
<i>sob-</i>	<i>sobalçou, sobescavar, sobestar, sobdominante, sobgrave</i>
<i>sub-</i>	<i>subarquado, subarrendar, subnutrido, sublocar, subsolo</i>

Tomou-se o cuidado de utilizar, para cada prefixo, pelo menos um vocábulo com base iniciada em vogal,²⁹ para verificar se o informante, ao se deparar com esses vocábulos, não circularia o núcleo do significado dos mesmos baseando-se exclusivamente nos critérios de separação silábica. Assim, diante da palavra *subarquado*, por exemplo, esperava-se que o *b* ficasse fora da parte circulada (ou seja, fora da parte da palavra que possui o ‘núcleo’ de seu significado).

Além das palavras prefixadas, foram usadas outras 16 palavras, listadas na Tabela 3, abaixo, que correspondem a formas que não apresentam nenhum dos prefixos sob investigação. Vale observar que o início desses itens lexicais é graficamente idêntico às formas prefixais apresentadas na Tabela 2. Esses vocábulos foram distribuídos aleatoriamente dentro da lista do instrumento e serviram para ver se o informante identificava as formas prefixadas sem se basear exclusivamente nas semelhanças gráficas existentes entre as palavras listadas.

Tabela 3 - Vocábulos com grafia inicial idêntica a prefixos do PB³⁰

Grafia inicial	Vocábulos sem os prefixos analisados
<i>ab-</i>	<i>abacate, abestado, ablitense</i>
<i>ad-</i>	<i>adolescente, adotado, adrenalina</i>
<i>ec-</i>	<i>economia, ecumênico</i>
<i>ob-</i>	<i>obedecer, obediente, obrigou, obturar</i>
<i>sob-</i>	<i>sobejo, sobrado</i>
<i>sub-</i>	<i>subtração, suburitano</i>

²⁹ O prefixo *ec-* só aparece diante de bases iniciadas por consoante. De acordo com o Dicionário Aurélio, para expressar o significado desse prefixo diante de bases iniciadas por vogal, faz-se uso do alomorfe *ex-*.

³⁰ Cabe observar que alguns desses vocábulos, como *abestado*, apresentam uma forma prefixal não submetida à investigação (no caso, *a-*). Esse item morfológico, no entanto, é graficamente diferente da forma prefixal que se está analisando neste trabalho (a saber, *ab-*), e, mesmo que fosse identificado pelo informante, não atrapalharia o objetivo do instrumento, já que sua diferenciação em relação ao prefixo *ab-* se mostraria visível na resposta do entrevistado, uma vez que este incluiria o *b* na parte circulada da palavra.

Chama-se atenção para o fato de que se julgou necessário apresentar a definição, entre parênteses, ao lado de 24 das palavras submetidas à apreciação dos entrevistados. Tal precaução foi tomada por se julgar que o informante, sem saber a definição exata do vocábulo analisado, poderia ficar confuso no momento de dar a resposta, informando, assim, um dado sem fundamentação, o que poderia prejudicar sua análise posterior.

Uma vez elaborado o instrumento, procedeu-se à aplicação do mesmo. O exercício foi aplicado a 20 falantes nativos da língua portuguesa na cidade de Porto Alegre - RS. Todos os informantes possuíam grau de instrução correspondente a, no mínimo, nível superior incompleto. Tais requisitos foram verificados através de perguntas individuais direcionadas aos potenciais participantes.

Além desses critérios, verificou-se se o informante não era ou havia sido aluno do curso de graduação em Letras. Tal cuidado, adotado igualmente na aplicação dos demais instrumentos, foi tomado por assumir-se que os alunos desse curso tendem a ser cautelosos em suas respostas ao participarem de investigações referentes ao estudo de línguas e geralmente têm facilidade na identificação do objetivo desses tipos de instrumentos, o que, no caso, prejudicaria a posterior análise dos resultados.

Após a devolução do exercício pelos 20 informantes, todas as respostas foram tabuladas para que pudessem ser analisadas. Os resultados dessa análise, apresentados na seção 3.1, mostram, em síntese, quais dos prefixos analisados em PB foram considerados transparentes e quais foram considerados pouco transparentes ou opacos, partindo-se da hipótese de que os prefixos *sob-* e *sub-* apresentam índices elevados de transparência, diferentemente de *ab-*, *ad-*, *ec-* e *ob-*. Esses resultados também servem de base para a formulação de algumas das hipóteses referentes aos dados coletados na Amostra de Fala 1 (cf. seção 2.1.3), que inclui a análise desses prefixos no que diz respeito a sua relação com a epêntese na fala do português brasileiro.

2.1.2 Teste de verificação de transparência dos prefixos em inglês

O segundo teste de transparência, (Anexo II), de um modo geral, se assemelha ao Instrumento 1 (cf. seção 2.1.1) no que diz respeito a seus objetivos e procedimentos metodológicos. A principal diferença está no fato de que este teste trata da análise da

transparência de prefixos da língua inglesa, a partir da percepção de aprendizes brasileiros desse idioma.

Para a elaboração do teste, foram levantados, através de consulta à obra *A comprehensive grammar of the English language*, de QUIRK, GREENBAUM, LEECH & SVARTVIK (1985), todos os prefixos da língua inglesa terminados em consoante obstruente, que, em razão dessa característica, são pressupostamente favorecedores da realização da epêntese vocálica na fala de aprendizes brasileiros de inglês. A lista desses prefixos pode ser visualizada na Tabela 4.

Tabela 4 - Prefixos relacionados à epêntese vocálica em inglês

Prefixo	Significado	Exemplos
<i>arch-</i> [ɑ:rtʃ] ³¹ ou [ɑ:rk] ³²	'highest', 'worst' = 'o mais elevado', 'o pior' ³³	<i>archduke</i> = 'arquiduque' <i>archenemy</i> = 'arqui-inimigo'
<i>out-</i> [aʊt]	'to do something faster, longer, etc than' = 'fazer algo mais rápido, por mais tempo etc. que'	<i>outrun</i> = 'correr mais que outrem' <i>outlive</i> = 'viver mais que'
<i>post-</i> [pəʊst]	'after' = 'após'	<i>postclassical</i> = 'pós-classico' <i>postwar</i> = 'pós-guerra'
<i>sub-</i> [sʌb]	'lower than', 'less than', 'beneath', 'lesser in rank' = 'mais baixo que', 'menos que', 'abaixo', 'inferior a'	<i>subhuman</i> = 'subumano' <i>subconscious</i> = 'subconsciente'

A partir dos prefixos apresentados na Tabela 4, formulou-se o segundo teste de transparência, constituído de 28 palavras. Vinte dessas palavras, que estão listadas na Tabela 5, correspondem a formas prefixadas em inglês, de acordo com o *Merriam-Webster Unabridged Dictionary* (2000).

³¹ As transcrições fonéticas foram feitas de acordo com as convenções do International Phonetic Alphabet e contaram com o auxílio do *Cambridge Advanced Learner's Dictionary* (disponível em <http://dictionary.cambridge.org/>) e do *Merriam-Webster Unabridged Dictionary* (versão eletrônica 2.5, 2000).

³² A pronúncia de *arch-* com [k], de acordo com o *Merriam-Webster Unabridged Dictionary* (2000), só é admitida para o vocábulo *archangel* ('arcanjo') e seus derivados. Pressupôs-se, no entanto, que os informantes submetidos ao instrumento 4 (cf. seção 2.1.4) poderiam oscilar entre a pronúncia da obstruente [k] e a da africada [tʃ] diante de *arch-*, o que constituiu uma razão a mais para incluir tal prefixo na lista de potenciais motivadores da epêntese vocálica.

³³ As glosas e traduções foram feitas pelo autor deste trabalho, utilizando-se, quando necessário, o Dicionário Eletrônico Michaelis (v. 5.0, 1998) e o Longman Dicionário Escolar (versão eletrônica 1.0.0.1, 2004).

Tabela 5 - Vocábulo prefixados submetidos à análise de transparência em inglês

Prefixo	Vocábulo prefixados
<i>arch-</i>	<i>archbishop, archdeacon, archduke, archfiend, archpriest</i> = 'arcebispo, arquidiácono, arquiduque, arquí-inimigo, arcipreste'
<i>out-</i>	<i>outbreathe, outdated, outfighting, outmaneuver, outrun</i> = 'exalar, obsoleto, lutando melhor que, manobrar melhor que, correr mais que'
<i>post-</i>	<i>postconsonantal, postdate, postfix, postocular, postwar</i> = 'pós-consonantal, pós-datar, pós-fixar, pós-ocular, pós-guerra'
<i>sub-</i>	<i>subcutaneous, sublet, subnitrate, subarid, subdominant</i> = 'subcutâneo, sublocar, subnitrato, subárido, subdominante'

Os oito vocábulo restantes do teste de transparência correspondem a palavras que apresentam grafia inicial similar a dos prefixos submetidos à investigação, mas que não são prefixadas. À semelhança do cuidado tomado no primeiro instrumento, objetivou-se verificar, com o acréscimo dessas palavras no teste, se os informantes, ao circularem a parte que julgassem ser o núcleo das palavras listadas, adotariam como critério de separação simplesmente a repetição gráfica de seqüências de letras.

Tabela 6 - Vocábulo com grafia inicial idêntica a prefixos do inglês

Grafia inicial	Vocábulo sem os prefixos analisados
<i>arch-</i>	<i>archil, archin</i> = 'urzela (ou tintura obtida desse líquen)', 'unidade russa de comprimento'
<i>out-</i>	<i>outardes, outrage</i> = 'ganso selvagem encontrado no Canadá (pl.)', 'ultraje'
<i>post-</i>	<i>posthitis, postiche</i> = 'postite', 'cabelo falso (peruca)'
<i>sub-</i>	<i>subah, subanum</i> = 'nome de província do Império Mogul', 'povo pagão das Filipinas'

Uma vez elaborado o exercício, o mesmo foi aplicado a 20 informantes³⁴, todos estudantes de língua inglesa de uma escola de idiomas localizada em Porto Alegre - RS e com idade mínima de 16 anos. Para participar do teste, exigiu-se que o estudante fosse falante nativo de português e que estivesse estudando inglês como língua estrangeira há pelo menos

³⁴ Esses informantes não são os mesmos que participaram do Instrumento 1, tampouco são os mesmos que participaram da coleta de dados de fala (Amostras de Fala 1 e 2). A escolha de diferentes informantes para os instrumentos deste estudo foi feita a partir do entendimento de que seria muito exaustiva e prolongada a participação de um informante caso o mesmo tivesse de participar de todas as atividades por nós propostas.

quatro anos. A exigência desse período se deu por se acreditar que, depois de oito semestres de aulas de inglês, o informante já teria condições de identificar morfemas nessa língua por já ter certa ‘intimidade’ com a estrutura da língua.

Hipotetiza-se que, para os resultados do Instrumento 2, todos os quatro prefixos da língua inglesa sejam percebidos pelos informantes. Há três razões que justificam a sustentação de tal hipótese.

Uma delas, que diz respeito a *arch-* e *out-*, está no fato de que a apresentação gráfica de tais prefixos é igual à atestada em inglês para as formas livres *arch* (‘arco’, ‘arquear’) e *out* (que pode ser tanto um adjetivo quanto um advérbio ou preposição). A semelhança dos prefixos com essas outras formas encontradas na língua inglesa pode contribuir para que os índices de isolamento de *arch-* e *out-* sejam elevados.

Outra razão, agora referente aos prefixos *post-* e *sub-*, também está relacionada à possibilidade de os informantes serem influenciados por analogias. Os prefixos *post-* e *sub-* não apenas se assemelham graficamente como também possuem significados muito próximos (para não dizer idênticos) aos prefixos *pós-* e *sub-*, da língua portuguesa. A comparação entre os dados de língua inglesa com vocábulos de sua língua nativa deve motivar um percentual elevado de isolamento desses prefixos.

A terceira e última razão que sustenta tal hipótese está associada a nossa crença de que todos esses quatro prefixos se comportam, sem exceção, como sílabas à esquerda de uma palavra fonológica recursiva, ou seja, são prefixos *adjuntos* (cf. 1.2). Admitindo-se que prefixos adjuntos são mais facilmente identificados pelos falantes, *arch-*, *out-*, *post-* e *sub-* devem, pois, apresentar graus elevados de transparência. A verificação disso é feita em 3.2.

2.1.3 Amostra de Fala 1 (PB)

A Amostra de Fala 1 corresponde ao terceiro instrumento deste trabalho e está diretamente relacionada ao teste de verificação de transparência de prefixos do PB (cf. seção 2.1.1) e à Amostra de Fala 2 (cf. seção 2.1.4).

Através deste terceiro instrumento, objetivou-se fazer o levantamento de dados de fala do português brasileiro para verificar a aplicação da epêntese vocálica medial. A Amostra de Fala 1 foi formada a partir de um conjunto de 13 frases (cf. Anexo III) nas quais estão inseridas 39 palavras em que há contexto para a inserção da vogal, ou seja, uma sílaba com

consoante obstruinte, nasal ou fricativa (exceto [s] e [z]) em posição de coda que anteceda uma sílaba iniciada por consoante.

Uma lista com os contextos apontados como favorecedores da epêntese em PB foi apresentada por CAGLIARI (1981) *apud* COLLISCHONN (2002:206). Essa lista, no entanto, além de não apresentar todos os encontros consonantais internos a uma base em que possa ocorrer epêntese,³⁵ não apresenta os *clusters* em fronteira entre prefixos e bases, nos quais a inserção vocálica também acontece. Em razão disso, elaboramos uma lista atualizada com os encontros em que a epêntese pode ocorrer no interior de um vocábulo (cf. Anexo IV).

A lista do Quadro 1 (cf. próxima página) apresenta os 39 vocábulos que foram cuidadosamente selecionados a partir das características linguísticas presentes em cada um deles (cf. seção 2.2.2). Isso significa dizer que, no momento de seleção dessas palavras, tomou-se o cuidado de separar ao menos um vocábulo para cada um dos tipos de fatores prosódicos e morfológicos considerados relevantes na realização da epêntese vocálica. No caso da divisão feita entre vocábulos com e sem prefixos, foram selecionados pelo menos dois exemplares para cada um dos seis afixos sob investigação (*ab-*, *ad-*, *ec-*, *ob-*, *sob-* e *sub-*). Além disso, para a categoria ‘Tipo de consoante perdida’ (cf. seção 2.2.2.1.2), escolheu-se um mínimo de três palavras sem prefixo para representar cada uma das oito possibilidades de consoantes perdidas (exceto para [m] e [f], para as quais foi selecionado apenas um vocábulo em virtude do número restrito de ocorrências desse contexto na língua).

Na formulação dessa lista, tomou-se igualmente o cuidado de selecionar palavras que apresentam *clusters* consonantais do PB não existentes em língua inglesa (o que explica o fato de haver mais vocábulos sem prefixo - quatro, em vez de três - nas linhas das obstruintes [k] e [g], bem como o fato de haver mais vocábulos - quatro, em vez de dois - com os prefixos *ec-* e *ad-*; e três, em vez de dois, com o prefixo *ob-*). A inclusão dessas palavras no instrumento se justifica pelo fato de se acreditar que os percentuais de aplicação de epêntese em encontros consonantais existentes apenas em uma das duas línguas sob análise sejam diferentes se comparados a contextos comuns aos dois idiomas (cf. seção 2.2.2.1.7).

³⁵ A busca rigorosa de contextos favorecedores de epêntese no Dicionário Eletrônico Houaiss (2002) permitiu identificar encontros consonantais não apontados por CAGLIARI (1981), como [ts] em *futsal*, [tn] em *etnia*, etc.

Quadro 1 - Vocábulo integrantes da amostra de fala em PB³⁶

Consoante perdida	sem prefixo	com prefixo
[p]	optou, apneias, opção	-
/t/	ritmo, bodisatvas, quartzo	-
[k]	ecpiesmas, maregraviáceas, actinites, écdico	éctipo, eczemas, écbase, ecdêmicas ³⁷
[b]	obturação, lambda, absinto	absorveu, abdicou, obfirmam, objeto, obversão, sobgraves, sobpõem, subchefe, sub-raças
/d/	vodca, hidnáceas, Djalma	adjunto, adligação, ad-rogar, ad-digital
[g]	magbás, pigmeus, estagflação, UFRGS	-
[m]	amnésias	-
[f]	aftas	-

A razão de se ter optado pela inclusão desses 39 vocábulos em frases, em vez de simplesmente submetê-los à leitura pelos informantes sob a forma de lista, está no fato de que, como observa LABOV (1972:211), textos parecem propiciar uma leitura menos formal do que listas de palavras. Procurou-se evitar, portanto, que os falantes procedessem a uma leitura muito cuidada, o que poderia gerar dados de fala ‘artificiais’ (i.é., não condizentes com aquilo que é realmente produzido pelo falante em uma situação não formal de fala).

Após a elaboração das frases, procedeu-se à seleção de 16 aprendizes brasileiros de inglês como língua estrangeira em uma escola particular de idiomas do município de Porto Alegre - RS. Esses informantes, que participaram igualmente da coleta dos dados que compõem a Amostra de Fala 2 (cf. seção 2.1.4), foram escolhidos de acordo com os pré-requisitos apresentados nas células sociais especificadas no Quadro 2, respeitando-se o método de seleção aleatório estratificado. Cada célula foi preenchida por dois participantes.

³⁶ Todos os vocábulos listados, à exceção dos nomes próprios, foram selecionados e tiveram sua composição morfológica verificada a partir de consulta ao Dicionário Aurélio Eletrônico (1999) e ao Dicionário Eletrônico Houaiss (2002).

³⁷ Note-se que a classificação de *éctipo* e de *écbase* como vocábulos prefixados pelos dicionários entra em conflito com o fato de que, em PB, um prefixo adjungido a uma base nunca atrai o acento da palavra para si (diferentemente dos casos de sufixação). Ainda assim, optou-se neste trabalho por respeitar a classificação apresentada nos dicionários consultados.

Quadro 2 - Células sociais da amostra de fala em PB e da amostra de fala em inglês

Célula 1	sexo masculino idade \leq 30 anos até 2 anos de estudo de inglês	Célula 5	sexo feminino idade \leq 30 anos até 2 anos de estudo de inglês
Célula 2	sexo masculino idade \geq 38 anos até 2 anos de estudo de inglês	Célula 6	sexo feminino idade \geq 38 anos até 2 anos de estudo de inglês
Célula 3	sexo masculino idade \leq 30 anos mais de 4 anos de estudo de inglês	Célula 7	sexo feminino idade \leq 30 anos mais de 4 anos de estudo de inglês
Célula 4	sexo masculino idade \geq 38 anos mais de 4 anos de estudo de inglês	Célula 8	sexo feminino idade \geq 38 anos mais de 4 anos de estudo de inglês

Uma vez selecionados os informantes, foi-lhes entregue um Termo de Consentimento (Anexo VI) e um Formulário sobre Informações Pessoais (Anexo VII) que foram preenchidos e assinados como condição para sua participação na pesquisa. Em seguida, os estudantes foram convidados a, individualmente, deixarem suas atividades de aula por alguns minutos para, em uma sala isolada, procederem à leitura das 13 frases em português e, posteriormente, das frases em inglês (que compõem a Amostra de Fala 2). Todos os dados foram gravados com equipamento de áudio Foston Digital Music Player F-56.

As hipóteses levantadas para os resultados deste instrumento estão descritas nos subitens da seção 2.2.2. A análise e a discussão desses resultados são apresentadas em 3.3 e em 3.5 (seção em que os resultados da Amostra de Fala 1 são comparados aos resultados da Amostra de Fala 2).

2.1.4. Amostra de Fala 2 (inglês)

O quarto e último instrumento de coleta de dados nesta pesquisa corresponde a um conjunto de quatro parágrafos com frases em língua inglesa, que foi submetido à leitura por parte de aprendizes brasileiros de inglês. Esse material (cf. Anexo V), trata, pois, da segunda parte do trabalho em que se efetuou a gravação de dados de fala para se verificar a realização da epêntese medial (a primeira parte - a Amostra de Fala 1-, referente a dados de fala do PB, está descrita em 2.1.3).

A elaboração dos quatro parágrafos que compõem este instrumento seguiu os mesmos passos adotados na formulação do instrumento referente à Amostra de Fala 1. Primeiramente,

foram localizados, através de consulta ao *Merriam-Webster Unabridged Dictionary* (2000), todos os verbetes que apresentavam *clusters* consonantais nos quais pudesse haver a ocorrência de epêntese por parte de aprendizes brasileiros de inglês. A partir de uma lista maior,³⁸ foram selecionados 61 vocábulos que, juntos, contemplavam todas as categorias de fatores linguísticos (cf. 2.2.2) considerados relevantes para a análise da aplicação do fenômeno fonológico ora investigado.

Esses vocábulos, apresentados no Quadro 3, (cf. próxima página), foram distribuídos dentro das frases que compõem os quatro parágrafos do instrumento de leitura. A razão de se optar pela leitura de frases/parágrafos em vez de uma lista de palavras está no fato de que essa última forma poderia fazer com que os participantes percebessem do que tratava a atividade, o que poderia influir no resultado da coleta.

A distribuição irregular com relação ao número de vocábulos em cada uma das categorias apresentadas no Quadro 3 é explicada a partir dos seguintes outros critérios adotados para a seleção desses vocábulos: (i) selecionou-se um mínimo de cinco vocábulos para representar cada um dos quatro prefixos de língua inglesa que poderiam motivar a epêntese; (ii) selecionou-se um mínimo de duas palavras para cada tipo de consoante perdida (exceto para [m], [n], [ʒ], [tʃ] e [ð], que representam categorias nas quais o número de palavras contempladas pela língua é reduzido); e (iii) selecionaram-se todos os vocábulos que apresentavam *clusters* consonantais existentes em inglês, mas não existentes em português brasileiro (em razão da hipótese apresentada em 2.2.2.1.7). A adoção desse último critério justifica a existência de mais do que dois vocábulos em algumas das categorias sem prefixo, e mais do que cinco vocábulos para a categoria com os prefixos *out-* e *post-*.

Uma vez elaborados os parágrafos, procedeu-se à aplicação do exercício de leitura a 16 aprendizes brasileiros de inglês como língua estrangeira de uma escola particular de idiomas localizada no município de Porto Alegre - RS. Esses informantes, que correspondem aos mesmos participantes da Amostra de Fala 1 (cf. seção 2.1.3), foram selecionados a partir das especificações sociais já apresentadas no Quadro 2.

³⁸ A relação completa desses encontros está contida na lista do Anexo IV.

Quadro 3 - Vocábulo integrantes da amostra de fala em inglês³⁹

Consoante perdida	sem prefixo	com prefixo
[p]	o ptions, ch ipmunk, cap ture, na pkin, na phtha	-
[t]	Latvian, flotsam, portfolio	outnumbered, postgraduate, outrun, outpassed, outbreathing, outshone, postsurgical, postdate, outthink, postmortal, posttest, postvocalic, outjockey, postrecord, postfixes
[k]	pictures, auction	-
[b]	sabzi, o bjects, rhabdom	subzone, subchief, subdivided, subconsciously, sublicense
[d]	dodgasted, adze	-
[g]	Nagpur, suggested [gdʒ], rugby, luxurious [gʒ], eczema	-
[m]	amnesia	-
[n]	enmity	-
[f]	fifty, hafnium, diphthongs	-
[ʃ]	marshmallows, ashlar	-
[ʒ]	cashmere	-
[tʃ]	hotchpotch	archconfraternity, archpriests, archbishops, archsee, archfiend
[dʒ]	hodgepodge, fledglings	-
[θ]	athletes, athbash, arithmetic, ethnic	-
[ð]	rhythmic	-

Para a realização do exercício de leitura, os informantes primeiramente tiveram de assinar um Termo de Consentimento (cf. Anexo VI), no qual declaravam sua concordância com a participação na pesquisa. Além disso, foi-lhes solicitado o preenchimento de um Formulário sobre Informações Pessoais (cf. Anexo VII), através do qual foram feitos questionamentos sobre características individuais do informante, que poderiam servir de suporte no entendimento dos resultados obtidos a partir dos dados coletados.

³⁹ Todos os vocábulo em inglês apresentados neste quadro, à exceção dos nomes próprios e dos adjetivos pátrios, foram selecionados e tiveram sua composição morfológica verificada a partir de consulta ao *Merriam-Webster Unabridged Dictionary* (2000). A transcrição fonética e o significado dos mesmos podem ser verificados no Anexo VIII.

Após a assinatura do Termo de Consentimento e o preenchimento do Formulário sobre Informações Pessoais, os participantes foram individualmente convidados a deixar as atividades de aula por alguns instantes para procederem à leitura das frases em português - cf. seção 2.1.3 - e em inglês, nessa ordem. Os dados foram gravados com equipamento de áudio Foston Digital Music Player F-56.

Todas as hipóteses levantadas para os resultados deste instrumento - assim como para os resultados da Amostra de Fala 1 - estão descritas nos subitens da seção 2.2.2. A análise e a discussão desses resultados estão apresentadas em 3.4 e em 3.5.

2.2 Variáveis submetidas à investigação

De acordo com a metodologia adotada em estudos de cunho variacionista (que buscam seguir os moldes do trabalho de William Labov sobre a ditongação no dialeto do inglês em Martha's Vineyard, cf. LABOV, 1972), deve-se, inicialmente, proceder à delimitação de um fenômeno linguístico variável, i.é., deve-se definir uma variável dependente.

A partir dessa variável, são identificados outros fatores, de ordem linguística ou extralinguística: as variáveis independentes. Essas variáveis podem favorecer ou inibir a realização da variável dependente, ou, ainda, não exercer influência alguma sobre ela.

Na subseção 2.2.1, são definidas as características da variável dependente analisada nesta pesquisa. Já em 2.2.2 e em seus subitens, são descritos os fatores de ordem linguística e extralinguística considerados relevantes para a aplicação ou não do fenômeno ora investigado, bem como são formuladas hipóteses sobre a forma como cada uma dessas variáveis independentes influi na realização da variável dependente.

2.2.1 Variável dependente

A variável dependente considerada no presente estudo corresponde a qualquer elemento vocálico que possa vir a ser inserido por falantes nativos de português brasileiro na pronúncia de determinados *clusters* consonantais atestados no interior de vocábulos do PB ou do inglês (as listas com esses *clusters* são apresentadas na subseção 2.2.2.1.7).

Tais encontros consonantais podem aparecer ou dentro de bases - como, por exemplo, em *acne* e em *option* - ou na fronteira entre alguns prefixos e suas bases - como em *ad+junto* e em *out+run*. Investiga-se aqui, portanto, a modalidade de epêntese denominada suarabácti

(ou anapitxe) e excluem-se os casos de epêntese inicial (como em [ɪ]structure) ou final (como em west[ɪ]).

Há duas possibilidades de classificação para essa variável: a não inserção da vogal epentética nos dados de fala, ou a aplicação do fenômeno. Exemplos de aplicação da variável dependente são verificados entre colchetes em *ac[ə]ne*, *op[ɪ]tion*, *ad[ɪ]junto* e *out[ɪ]run*. Neste estudo, assim como em COLLISCHONN (2002), não se levou em consideração a qualidade da vogal epentética produzida pelo falante, mas apenas sua realização ou não. O trabalho de PEREYRON (2008) demonstra que esse tipo de vogal não tem relação direta com as variáveis linguísticas e extralinguísticas envolvidas na aplicação da epêntese. A título de curiosidade, esse mesmo trabalho demonstra que 84% dos casos de epêntese atestados correspondem à vogal breve [ɪ], o que indica a preferência (quase categórica) por essa qualidade de vogal epentética pelos falantes de Porto Alegre.⁴⁰

É importante esclarecer o porquê de se falar em ‘variável dependente’. Ainda que essa terminologia esteja intimamente associada aos estudos sociolinguísticos desenvolvidos a partir da década de 1970, seu uso pode ser transposto a trabalhos não necessariamente vinculados a essa linha de pesquisa (como no caso do presente estudo, especificamente em relação aos dados de língua inglesa como LE, para os quais a condição de ‘fenômeno em variação’ e os preceitos teóricos da Sociolinguística são diferentes do que se postula para os dados de PB como língua materna). Através do termo ‘variável’, infere-se logicamente que a aplicação do fenômeno investigado (no caso, a inserção de uma vogal) é inconstante, podendo não se aplicar categoricamente. Retomando os exemplos do parágrafo anterior, a não aplicação da epêntese nos contextos apontados resultaria em dados de fala equivalentes a [ˈaknɪ], [ˈɑpʃən], [adˈʒuːtʃ] e [aʊtˈrʌn].

Além disso, o termo adotado pelos estudos sociolinguísticos é, antes, comum na terminologia de análises quantitativas, e faz referência ao caráter ‘dependente’ da variável. Isso significa que essa variável literalmente depende de outros fatores para que seja aplicada. Esses fatores, denominados ‘variáveis independentes’ (ver próxima seção), podem ser de ordem puramente linguística, assim como podem estar relacionados a características extralinguísticas privativas a cada indivíduo que participou como informante nessa etapa da pesquisa.

⁴⁰ Embora provável que [ɪ] seja a vogal mais frequente por questões articulatórias, não está descartada a hipótese de que ela seja também mais recorrente por ser socialmente mais prestigiada, especialmente quando comparada a [e] (vejam-se os exemplos *ad[ɪ]vogado* e *ad[e]vogado*, *p[ɪ]neu* e *p[e]neu*).

2.2.2 Variáveis independentes

As variáveis independentes correspondem aos fatores de natureza gramatical ou extralinguística que podem estar associados ao fenômeno sob investigação - no caso, à epêntese vocálica medial em português brasileiro e em inglês. Nas subseções de 2.2.2.1 são apresentadas as variáveis gramaticais (ou linguísticas), enquanto nas subseções de 2.2.2.2 são definidas as variáveis sociais (ou extralinguísticas) deste estudo. Nelas são igualmente apresentadas as hipóteses formuladas para os dados de fala coletados em português (cf. seção 2.1.3) e em inglês (cf. seção 2.1.4).

2.2.2.1 Variáveis independentes linguísticas

Nas próximas subseções, são discriminados os fatores de ordem linguística que, com base em outros estudos sobre epêntese - em especial os de COLLISCHONN (2002) e de PEREYRON (2008) -, foram considerados relevantes para a investigação da epêntese (inicial, medial ou final) em dados de fala do português e do inglês como LE. O objetivo principal em ‘aproveitar’ as hipóteses e os resultados apontados nesses estudos é o de consolidar certa homogeneidade teórica e buscar definir alguma linearidade nos resultados dos trabalhos que envolvem a epêntese vocálica realizada por brasileiros, seja nos vocábulos de sua língua nativa ou nos da língua estrangeira que estão aprendendo.

Algumas das hipóteses, embora tenham sido obtidas de outros estudos sobre a inserção vocálica, tiveram de ser adaptadas aos propósitos desta investigação. Há, além disso, o caso de uma variável (‘contexto morfológico’) que recebe tratamento exclusivo neste estudo.

2.2.2.1.1 Contexto morfológico

Dos fatores linguísticos relacionados à epêntese medial em PB e em inglês, ‘contexto morfológico’ é o que recebe maior destaque neste trabalho, especialmente porque, nas investigações sobre o fenômeno fonológico de inserção vocálica na fala de brasileiros já realizadas, nenhuma ou pouca atenção lhe foi dada.

COLLISCHONN (2002) excluiu de seu estudo vocábulos com o prefixo *sub-* (*subtenente*, *subterrâneo*, etc.) justificando que “(...) existe a suspeita de que o prefixo atue com uma certa independência fonológica, e nesse caso, a consoante /b/ estará em final de

palavra, (...)” (p. 211). PEREYRON (2008) incluiu em seus dados os vocábulos *subcategory*, em que o *cluster* [bk] está num contexto morfológico de fronteira entre prefixo e base, e *webpage*, em que [bp] está na fronteira entre duas formas livres. A autora, no entanto, não discutiu se o fato de esses encontros consonantais estarem numa posição de fronteira morfológica poderia implicar diferentes taxas de aplicação de epêntese. É em razão da ausência dessa discussão e, principalmente, da suspeita levantada por COLLISCHONN (2002) acerca de *sub-* que a variável ‘contexto morfológico’ é investigada neste trabalho.

Assume-se, pois, que, dentro dos vocábulos, a vogal pode ser inserida em dois contextos morfológicos distintos: no interior da base, como ocorre nos exemplos *pig*[ɪ]*meu*, *am*[ə]*néσίας*, *cash*[ɪ]*mere* e *arith*[ɪ]*methic*, ou na fronteira entre o prefixo e sua base, ou seja, entre a extremidade esquerda da base e a extremidade direita do prefixo (terminado em consoante obstruente ou africada), como nos exemplos *ec*[ə]*dêmicas*, *ad*[ɪ]*ligação*, *arch*[ɪ]*bishops* e *post*[ɪ]*date*.

Para as rodadas estatísticas realizadas com os dados de fala do português brasileiro, foram considerados os seguintes contextos morfológicos:

- interior da base, como em *quartzo*;
- prefixo *ab-* + base, como em *absorveu*;
- prefixo *ad-* + base, como em *adjunto*;
- prefixo *ec-* + base, como em *eczemas*;
- prefixo *ob-* + base, como em *obversão*;
- prefixo *sob-* + base, como em *sobgraves*;
- prefixo *sub-* + base, como em *subchefe*.

Na realização das rodadas estatísticas com os dados de fala em língua inglesa, foram admitidas as seguintes categorias:

- interior da base, como em *diphthongs*;
- prefixo *arch-* + base, como em *archconfraternity*;

- prefixo *out-* + base, como em *outnumbered*;
- prefixo *post-* + base, como em *postmortal*;
- prefixo *sub-* + base, como em *sublicense*.

Hipotetiza-se que, tanto em português quanto em inglês, a epêntese seja mais frequente em *clusters* localizados na fronteira entre prefixos e bases do que em *clusters* localizados no interior das bases. Mais especificamente, busca-se confirmar a suspeita apresentada por COLLISCHONN (2002) de que vocábulos com o prefixo *sub-* (e com o prefixo *sob-*, por extensão) apresentem maiores índices de inserção vocálica, comprovando-se, assim, a existência de certa independência fonológica para esse afixo.

2.2.2.1.2 Tipo de consoante perdida

Este fator linguístico corresponde ao tipo de consoante que antecede o ponto em que ocorre a inserção da vogal epentética. Em outras palavras, corresponde ao contexto que precede o local em que a epêntese pode ser aplicada.

Originalmente, tais consoantes ocupam a posição de coda de sílaba. Com a inserção da epêntese, admite-se que tais consoantes percam tal posição, assumindo o ataque de uma nova sílaba (cf. seção 1.3.4).

A lista com os diferentes tipos de consoantes perdidas que foram analisados neste estudo pode ser visualizada na Tabela 7 (cf. próxima página). Cabe observar que nem todos os tipos de consoantes perdidas que ocorrem na língua inglesa estão presentes em PB (cf. seção 2.2.2.1.7, que discrimina os tipos de *clusters* consonantais analisados em cada língua).

Foram excluídos da análise estatística os dados em que a consoante perdida deixou de ser pronunciada. A justificativa para essa exclusão está no fato de que, com a eliminação dessa consoante e consequente simplificação do *cluster*, não há mais contexto para a aplicação da epêntese.

Tabela 7 - Tipos de consoantes perdidas analisados

Descrição da consoante perdida	Exemplo em português	Exemplo em inglês
[p] - oclusiva bilabial desvozeada	<i>optou</i>	<i>napkin</i>
[b] - oclusiva bilabial vozeada	<i>subchefe</i>	<i>objects</i>
[t] - oclusiva alveolar desvozeada	<i>ritmo</i>	<i>portfolio</i>
[d] - oclusiva alveolar vozeada	<i>vodca</i>	<i>adze</i>
[k] - oclusiva velar desvozeada	<i>actinites</i>	<i>auction</i>
[g] - oclusiva velar vozeada	<i>pigmeus</i>	<i>rugby</i>
[m] - nasal bilabial	<i>amnésias</i>	<i>amnesia</i>
[n] - nasal alveolar	-	<i>enmity</i>
[f] - fricativa labiodental desvozeada	<i>aftas</i>	<i>hafnium</i>
[ʃ] - fricativa alveopalatal desvozeada	-	<i>marshmallows</i>
[ʒ] - fricativa alveopalatal vozeada	-	<i>cashmere</i>
[tʃ] - africada alveopalatal desvozeada	-	<i>archfiend</i>
[dʒ] - africada alveopalatal vozeada	-	<i>fledglings</i>
[θ] - fricativa interdental desvozeada	-	<i>ethnic</i>
[ð] - fricativa interdental vozeada	-	<i>rhythmic</i>

O estudo de COLLISCHONN (2002) apontou que as oclusivas alveolares /t/ e /d/ formam as piores codas no PB, já que favorecem consideravelmente a epêntese, exceto quando seguidas de uma sibilante (caso em que tendem a formar africadas como [tʃ]). Espera-se, pois, que a inserção vocálica seja igualmente favorecida por /t/ e /d/.

PEREYRON (2008), por sua vez, demonstrou que, nos dados do inglês, a velar [g] é a que mais favorece a epêntese medial. A velar [k] também se mostrou motivadora do fenômeno em uma das análises da autora (peso relativo 0,68 na análise perceptual). As alveolares, ao contrário do que foi constatado em COLLISCHONN (2002), foram as que menos favoreceram a epêntese. A partir disso, presume-se que o comportamento da epêntese nos dados coletados deva ser diferente entre os vocábulos do PB e os do inglês, o que pode indicar que a variação da epêntese em PB e a variação da epêntese na interlíngua devam ser tratadas de maneiras distintas.

2.2.2.1.3 Vozeamento da consoante perdida

Outro aspecto linguístico que pode favorecer a inserção de uma vogal entre os *clusters* consonantais investigados é o vozeamento da consoante perdida. Como se observa na seção

anterior, no que diz respeito ao traço [sonoro], as consoantes precedentes podem ser desvozeadas, como em *ri[t]mo* e em *fi[f]ty*, ou vozeadas, como em *pi[g]meus* e em *o[b]jects*.

KINZEL (2000) registrou maior ocorrência de epêntese nos vocábulos do PB que apresentam *clusters* vozeados, como em *magma*. PEREYRON (2008), por sua vez, analisou separadamente o vozeamento da consoante perdida e o da consoante seguinte em dados do inglês. Em seu estudo, as consoantes perdidas com o traço [+ vozeado] se mostraram mais propensas a favorecer a epêntese.

Presume-se, com isso, que a inserção vocálica ocorra com maior frequência após consoantes vozeadas tanto em PB quanto em inglês. Admite-se, pois, que o fato de a consoante perdida compartilhar com as vogais a marcação positiva para o traço [vozeado] implique maior probabilidade de que uma vogal epentética a acompanhe.

2.2.2.1.4 Tipo de contexto seguinte

O contexto seguinte corresponde à consoante que imediatamente sucede a consoante perdida ou, no caso de haver aplicação de epêntese, a consoante que vem imediatamente após a vogal inserida. A Tabela 8 (cf. próxima página) apresenta os tipos de contexto seguinte considerados neste estudo.

Assim como para os casos de apagamento da consoante perdida, foram ignoradas na análise estatística as ocorrências em que não houve produção do contexto seguinte. Com a ocorrência de apenas uma das consoantes integrantes dos *clusters* investigados, não há como constatar a presença da vogal epentética.

A variável ‘contexto seguinte à consoante perdida’ foi considerada a segunda mais significativa no trabalho de COLLISCHONN (2002), sendo de as fricativas não sibilantes (como [v]) foram as que mais apareceram após epêntese. A autora concluiu que a fricativa não sibilante se mostrou a mais favorecedora de epêntese porque ela estava, em 100% dos dados analisados, em uma posição pretônica (posição apontada no trabalho como a mais favorável para a inserção vocálica). As nasais foram o segundo contexto seguinte que mais apareceram após epêntese. As sibilantes [s] e [z], por sua vez, foram as menos favorecedoras do fenômeno, fato que motivou Collischonn a admitir a formação de africadas com a junção dessas sibilantes com as consoantes oclusivas que as antecedem.

Tabela 8: Tipos de contextos seguintes analisados

Descrição do contexto seguinte		Exemplo em português	Exemplo em inglês
oclusiva bilabial	[p]	<i>sobpõem</i>	<i>archpriests</i>
	[b]	<i>magbás</i>	<i>rugby</i>
oclusiva alveolar	[t]	<i>aftas</i>	<i>posttest</i>
	[d]	<i>lambda</i>	<i>postdate</i>
oclusiva velar	[k]	<i>vodca</i>	<i>napkin</i>
	[g]	<i>sobgraves</i>	<i>dodgasted</i>
nasal	[m]	<i>pigmeus</i>	<i>enmity</i>
	[n]	<i>hidnáceas</i>	<i>amnesia</i>
fricativa labiodental	[f]	<i>estagflação</i>	<i>archfiend</i>
	[v]	<i>obversão</i>	<i>postvocalic</i>
fricativa alveolar	[s]	<i>absorveu</i>	<i>archsee</i>
	[z]	<i>eczemas</i>	<i>sabzi</i>
fricativa alveopalatal	[ʃ]	<i>subchefe</i>	<i>outshone</i>
	[ʒ]	<i>Djalma</i>	<i>luxurious [gʒ]</i>
africada alveopalatal	[tʃ]	<i>éctipo</i>	<i>subchief</i>
	[dʒ]	<i>écdico</i>	<i>outjockey</i>
fricativa interdental desvozeada	[θ]	-	<i>diphthongs</i>
lateral alveolar vozeada	[l]	<i>adligação</i>	<i>ashlar</i>
retroflexa alveolar vozeada	[ɽ]	-	<i>postrecord</i>
vibrante alveolar vozeada* ⁴¹	[r̃]	<i>sub-raças</i>	-
tepe alveolar vozeado*	[r]	-	<i>outrun</i>
fricativa velar desvozeada	[x]	<i>ad-rogar</i>	-
fricativa glotal desvozeada*	[h]	<i>ad-rogar</i>	-

Nos dados do inglês de PEREYRON (2008), a fricativa [v] apresentou peso relativo neutro. A fricativa [s], por sua vez, teve comportamento diverso nas análises acústica e perceptual, mostrando-se inibidora na primeira e favorecedora na última dessas análises. A nasal [n] foi a única que apresentou comportamento semelhante ao estudo de COLLISCHONN (2002), com peso relativo favorecedor para a epêntese.

⁴¹ Os tipos de contexto seguinte marcados com um asterisco foram incluídos após a transcrição dos dados coletados e correspondem, portanto, a contextos inicialmente não esperados (a vibrante alveolar vozeada foi a menos esperada, principalmente em razão dos resultados de MONARETTO, 1997, que demonstram a gradativa substituição do /r/ alveolar pela fricativa velar). Para efeito das análises estatísticas, fez-se a amalgamação da vibrante alveolar vozeada [r̃], do tepe alveolar vozeado [r] e da retroflexa alveolar vozeada [ɽ], já que compartilham o ponto de articulação (alveolar). Também foram amalgamados os casos de realização de fricativa glotal desvozeada [h] e de fricativa velar desvozeada [x], ambas articuladas na parte posterior do trato vocal.

Espera-se, assim, verificar se a vogal epentética em dados do PB e do inglês de fato apresenta comportamento diverso no que se refere à sua relação com a consoante imediatamente posterior.

2.2.2.1.5 Vozeamento do contexto seguinte

A consoante que vem após a consoante perdida ou a vogal epentética pode ser desvozeada ou vozeada. Alguns exemplos dos vocábulos analisados que apresentam contexto seguinte desvozeado são *af[t]as*, *ob[f]irmam*, *cap[tj]ure* e *nap[k]in*. Exemplos que apresentam contexto seguinte vozeado são *sub-[x]aças*, *pig[m]eus*, *post[g]raduate* e *ash[l]ar*.

Assim como previsto para a variável ‘vozeamento da consoante perdida’, acredita-se que o traço [+ vozeado] da segunda consoante do *cluster* motive a realização da vogal, que também possui marcação positiva para [vozeado].

2.2.2.1.6 Acento

Esta variável diz respeito ao contexto prosódico no qual se localiza a consoante perdida dentro do vocábulo e se subdivide em duas possibilidades: pretônica e postônica.

Se a consoante aparece antes do núcleo da sílaba que detém o acento principal da palavra, é classificada como pretônica. Exemplos de vocábulos com consoantes perdidas pretônicas são [abdʒɪ'kɔw] (*abdicou*) e [æm'niʒə] (*amnesia*), em que o [b] e o [m], respectivamente, precedem o início da sílaba que detém o acento principal da palavra (identificado pelo diacrítico “’”).

Quando a consoante perdida é pronunciada após o acento principal da palavra, é classificada como postônica. Nas palavras [lã^mbdə] (*lambda*) e [ræbdəm] (*rhabdom*), [b] corresponde a uma consoante perdida localizada na pauta acentual postônica.

Vale observar que os dados dessa categoria foram codificados de acordo com a forma como o falante produziu o dado no momento da gravação. Assim, se uma palavra como *éctipo*, por exemplo, foi lida como [ek'tjipɔ], ou seja, foi produzida com a transferência do acento principal da palavra para outra sílaba (nesse caso, provocando a mudança do contexto da consoante perdida de postônico para pretônico), sua codificação foi registrada com a alteração feita pelo falante.

Enquanto o trabalho de COLLISCHONN (2002) apresenta essa variável como a mais significativa na realização da epêntese (sendo o fator ‘postônico’ visto como o maior inibidor do fenômeno), PEREYRON (2008) sequer discute seu comportamento nos dados da língua inglesa, já que suas rodadas estatísticas não a selecionaram durante o processo de *step-up*.

COLLISCHONN (2002) argumentou que o fator postônico se mostrou inibidor do fenômeno porque a inserção da vogal após o acento implicaria o afastamento deste do final da palavra (de [et.ni.ku] para ?[e.t[ɪ.ni.ku] *ético*, por exemplo). Esse afastamento não é lícito no PB, em que se admite acento até a terceira sílaba antes do fim de palavra.

Observe-se que, em inglês, é possível que o acento esteja após a terceira sílaba (a contar da direita para a esquerda), como é o caso de [kʌm-fər-tə-blɪ] *comfortably*. Com base nessa informação, não surpreende o fato de que essa variável não tenha sido selecionada na análise de PEREYRON (2008), uma vez que o deslocamento do acento em inglês não implica necessariamente a produção de uma forma não licenciada na pauta acentual do inglês.

Espera-se, pois, confirmar a hipótese de que a variável ‘acento’ tenha relevância apenas na análise dos dados de fala do PB, para a qual a posição postônica deve inibir a aplicação da epêntese. Nos dados da língua inglesa, a inserção de uma vogal após o acento da palavra não deve ter razão de ser evitada, já que isso não resulta em um dado estranho à pauta acentual dessa língua.

2.2.2.1.7 Tipo de *cluster*

Este fator linguístico indica o idioma em que ocorrem os encontros consonantais presentes nos dados de fala analisados nas amostras do PB e do inglês. Inicialmente foram propostas três categorias de classificação dos dados para essa variável.

A primeira categoria incluiu todos os vocábulos que apresentam *clusters* consonantais que possam favorecer a epêntese e que existem em PB, mas não existem em inglês. A Tabela 9 (cf. próxima página) apresenta quais são esses *clusters* e quais foram os vocábulos selecionados nesta pesquisa.

Tabela 9 - *Clusters* relacionados à epêntese vocálica em PB inexistentes em inglês

Tipo de <i>cluster</i> consonantal	Exemplo
[bʒ]	<i>objeto</i>
[bx]	<i>sub-raças</i>
[bř]* ⁴²	<i>sub-raças</i>
[bh]*	<i>sub-raças</i>
[dʒ]	<i>ad-digital</i>
[dx]	<i>ad-rogar</i>
[dř]*	<i>ad-rogar</i>
[dh]*	<i>ad-rogar</i>
[kp]	<i>ecpiemas</i>
[kz]	<i>eczemas</i>
[kdʒ]	<i>écdico</i>
[gs]	<i>UFRGS</i>

A segunda categoria, por sua vez, inclui todos os encontros consonantais existentes na língua inglesa e ausentes em PB, e que, na fala de aprendizes brasileiros de inglês, podem servir de contexto de aplicação da epêntese. Os encontros analisados nesta pesquisa para essa subcategoria encontram-se na Tabela 10 (cf. próxima página).

⁴² Os *clusters* indicados com um asterisco foram incluídos após a audição dos dados coletados. Inicialmente, não se esperava a ocorrência desses encontros, uma vez que se partiu do pressuposto de que, em Porto Alegre, a forma mais usual de se pronunciar o ‘r’ em início de sílaba pós-prefixal é através da fricativa alveolar desvozeada [x].

Tabela 10 - *Clusters* relacionados à epêntese vocálica em inglês inexistentes em PB

Tipo de <i>cluster</i> consonantal	Exemplo
[pk]	<i>napkin</i>
[pm]	<i>chipmunk</i>
[pf]	<i>options</i>
[pθ]	<i>naphtha</i>
[tp]	<i>outpassed</i>
[tb]	<i>outbreathing</i>
[tt]	<i>posttest</i>
[td]	<i>postdate</i>
[tʃ]	<i>outshone</i>
[tdʒ]	<i>outjockey</i>
[tθ]	<i>outthink</i>
[tr]	<i>postrecord</i>
[kʃ]	<i>auction</i>
[gz]	<i>eczema</i>
[gʒ]	<i>luxurious</i>
[gdʒ]	<i>suggested</i>
[nm]	<i>enmity</i>
[fθ]	<i>diphthongs</i>
[ʃl]	<i>ashlar</i>
[ʒm]	<i>cashmere</i>
[tʃp]	<i>archpriests</i>
[tʃb]	<i>archbishops</i>
[tʃk]	<i>archconfraternity</i>
[tʃf]	<i>archfiend</i>
[tʃs]	<i>archsee</i>
[dʒp]	<i>hodgepodge</i>
[dʒl]	<i>fledglings</i>
[θb]	<i>athbash</i>
[θm]	<i>arithmetic</i>
[θn]	<i>ethnic</i>
[θl]	<i>athletes</i>
[ðm]	<i>rhythmic</i>

A terceira categoria englobou os demais grupos consonantais que podem favorecer a epêntese vocálica medial e que existem tanto em português brasileiro quanto em inglês. Esses grupos podem ser verificados na Tabela 11 (cf. próxima página).

Tabela 11 - *Clusters* relacionados à epêntese vocálica existentes em PB e em inglês

Tipo de <i>cluster</i>	Exemplo ⁴³
[pt]	<i>optou</i>
[pn]	<i>apneias</i>
[ps]	<i>opção</i>
[ptʃ]	<i>capture</i>
[bp]	<i>sobpõem</i>
[bt]	<i>obturação</i>
[bd]	<i>lambda, subdivided</i>
[bk]	<i>subconsciously</i>
[bg]	<i>sobgraves</i>
[bf]	<i>obfirmam</i>
[bv]	<i>obversão</i>
[bs]	<i>absinto</i>
[bz]	<i>subzone</i>
[bʃ]	<i>subchefe</i>
[btʃ]	<i>subchief</i>
[bdʒ]	<i>abdicou, objects</i>
[bl]	<i>sublicense</i>
[tg]	<i>postgraduate</i>
[tf]	<i>portfolio</i>
[tm]	<i>ritmo, postmortal</i>
[tn]	<i>outnumbered</i>
[tv]	<i>bodisatvas, postvocalic</i>
[ts]	<i>postsurgical</i>
[tz]	<i>quartzo</i>
[dk]	<i>vodka</i>
[dg]	<i>dodgasted</i>
[dz]	<i>adze</i>
[dʒ]	<i>adjunto</i>
[dn]	<i>hidnáceas</i>
[dl]	<i>adligação</i>
[kb]	<i>écbase</i>
[kd]	<i>ecdêmicas</i>
[kg]	<i>marcgraviáceas</i>
[ktʃ]	<i>actinites, pictures</i>
[gp]	<i>Nagpur</i>
[gb]	<i>magbás, rugby</i>
[gm]	<i>pigmeus</i>
[mn]	<i>amnésias, amnesia</i>
[gf]	<i>estagflação</i>
[ft]	<i>aftas, fifty</i>
[fn]	<i>hafnium</i>
[fm]	<i>marshmallows</i>

⁴³ Nesta tabela foram exemplificados apenas os vocábulos utilizados na coleta de dados de fala desta pesquisa, que não chegou a incluir uma palavra de cada língua para cada tipo de *cluster* existente, uma vez que isso tornaria o instrumento de coleta muito extenso. Palavras como *aptitude*, *capsule*, *obséquo* e *futsal* são apenas alguns dos exemplos que provam a ocorrência desses encontros consonantais em ambas as línguas. A lista completa e exemplificada de *clusters* relacionados à epêntese em PB em inglês pode ser conferida no Anexo IV.

Acrescentou-se, após a audição dos dados, uma quarta categoria a esse grupo de fatores, na qual foram enquadrados os casos de *clusters* produzidos pelos falantes que inexistem na língua em que foram pronunciados ou, ainda, que inexistem em ambas as línguas. Um exemplo de encontro consonantal incluído nessa categoria é [vt], pronunciado para o dado *Latvian* (note-se que, nesse caso, o que ocorreu foi a inversão da posição da consoante perdida com a consoante seguinte).

COLLISCHONN (2002) investigou em seu estudo a variável ‘origem do vocábulo’, para a qual considerou os fatores ‘estrangeiro’ (*partner*) e ‘nativo’ (*absurdo*). A autora formulou a hipótese de que a epêntese seria menos aplicada em vocábulos estrangeiros, sob o argumento de que os falantes têm mais cuidado na pronúncia de vocábulos não nativos de sua língua. As rodadas estatísticas de seu estudo, no entanto, não consideraram a origem do vocábulo como estatisticamente relevante, e, em razão disso, COLLISCHONN (2002) não chegou a discutir os resultados obtidos para essa variável.

PEREYRON (2008), com base na variável ‘origem do vocábulo’ de COLLISCHONN (2002), admitiu em seu estudo a variável ‘tipo de *cluster*’, que foi subdividida nos fatores ‘encontros consonantais comuns ao português e ao inglês’ (como [kn], presente tanto em *acne* como em *acknowledge*) e ‘encontros consonantais que pertencem somente ao inglês’ (como [pk] em *napkin*). A hipótese apresentada para essa variável foi a de que os encontros existentes somente em inglês seriam mais observados pelos falantes e apresentariam menos epêntese em razão disso. PEREYRON (2008) confirmou sua hipótese em suas rodadas estatísticas, que apontaram pesos relativos baixos para a epêntese em *clusters* existentes somente em inglês.

Nesta pesquisa, além dos fatores analisados em PEREYRON (2008) para essa variável, incluíram-se os casos de *clusters* que existem em português, mas não existem no inglês (cf. Tabela 9). A partir da inclusão desse fator, são propostas as seguintes hipóteses para a variável ‘tipo de *cluster*’:

(i) *clusters* existentes em inglês e inexistentes em PB devem apresentar taxas menores de inserção vocálica do que *clusters* existentes em ambas as línguas ou existentes em PB e inexistentes em inglês;

(ii) *clusters* existentes em PB e inexistentes em inglês devem apresentar taxas maiores de inserção vocálica do que *clusters* existentes em ambas as línguas.

Para propor a hipótese (i), assume-se que a ‘consciência’ dos falantes sobre o fato de que a epêntese não corresponde a um fenômeno linguístico comum no inglês se manifeste mais em vocábulos que apresentem padrões silábicos exclusivos⁴⁴ daquele idioma. Por analogia, a hipótese (ii) foi formulada para verificar se *clusters* existentes em PB e inexistentes em inglês motivam mais a epêntese, já que correspondem a encontros consonantais exclusivos da língua em que a inserção vocálica é lícita.

2.2.2.2 Variáveis independentes extralinguísticas

Nos próximos subitens são apresentados os fatores de ordem extralinguística considerados de relevância para a inserção vocálica ora investigada. As células sociais, que apresentam a distribuição exata dos informantes selecionados com base nesses fatores, podem ser visualizadas no Quadro 2 (cf. seção 2.1.3).

Novamente, as principais hipóteses apresentadas partem dos resultados já apontados em outros estudos sobre a epêntese – especialmente os de COLLISCHONN (2002), para os dados do PB, e os PEREYRON (2008), para os dados do inglês.

2.2.2.2.1 Sexo

Esta variável contempla as categorias ‘feminino’ e ‘masculino’ e busca mostrar, assim como grande parte dos trabalhos que seguem a metodologia sociolinguística, possíveis diferenças entre os informantes de cada sexo no que diz respeito aos percentuais de aplicação do fenômeno sob investigação.

Trabalhos como os de LABOV (1972) indicam que as mulheres buscam produzir as formas que são prestigiadas em sua língua. Na maioria das vezes, essas formas prestigiadas são aquelas que respeitam a norma culta de sua língua.

CAMARA JR. (1970) observa que existe uma tendência a reduzir a emissão da vogal epentética no registro formal da língua culta. Admitindo-se que essa redução da emissão da vogal seja a forma mais prestigiada entre os falantes, presume-se que o fator ‘feminino’ seja inibidor do processo de epêntese em PB, já que as mulheres pressupostamente respeitam mais

⁴⁴ A noção de ‘exclusivo’ aqui é usada para dizer que, no par inglês/português, o encontro consonantal ocorre em somente um dos idiomas. No entanto, nada impede que esse encontro esteja presente em outras línguas.

o registro formal da língua, evitando, portanto, a produção de formas epentéticas. Por analogia, tal hipótese também é feita para os dados do inglês.

2.2.2.2.2 Idade

Neste estudo, o fator ‘idade’ foi igualmente considerado para os informantes que procederam à leitura dos dados que compõem a Amostra de Fala 1 e a Amostra de Fala 2. Foram estipulados dois grupos de faixa etária: um que incluiu oito participantes com idades entre 16 e 30 anos, e outro que contemplou oito informantes com idades entre 38 e 61 anos.

O estudo de COLLISCHONN (2002) evidencia que o fator idade não tem grande relevância na variação da epêntese em PB (essa variável sequer foi selecionada pelas rodadas estatísticas). Em razão disso, não há uma hipótese específica sobre a relação existente entre o fator ‘idade’ e a epêntese nos dados de PB investigados nesta pesquisa.

Com relação à epêntese em inglês, PEREYRON (2008) demonstrou que, nos dados analisados, a aplicação da epêntese é mais frequente entre os aprendizes mais velhos. Neste trabalho, hipotetiza-se, por analogia, que os falantes da faixa etária mais elevada, i.é, os que têm idade entre 38 e 61 anos, apliquem mais a epêntese em inglês do que os falantes com idade entre 16 e 30 anos.

2.2.2.2.3 Informante

O grupo ‘Informante’ considera individualmente cada um dos dezesseis participantes que procederam à leitura dos dados referentes à Amostra de Fala 1 e à Amostra de Fala 2.

COLLISCHONN (2002) não chegou a incluir essa variável em seu trabalho. PEREYRON (2008), contrariamente, investigou o comportamento da epêntese em inglês para cada indivíduo inglês, mas essa variável não foi considerada significativa em suas rodadas estatísticas.

Muito embora não existam hipóteses precisas para esta variável, acredita-se que, por questões externas à língua, alguns informantes possam aplicar mais a epêntese do que outros, tanto nos dados do português como nos do inglês como LE. Nesse sentido, as informações contidas nos formulários sobre informações pessoais dos participantes (cf. modelo no Anexo VII) podem auxiliar na explicação das possíveis diferenças entre as taxas de aplicação da vogal epentética para cada informante.

2.2.2.2.4 Proficiência em inglês

O fator ‘proficiência em inglês’ corresponde, neste trabalho, à quantidade de tempo, em anos, a que o informante já havia se dedicado ao estudo da língua inglesa à época em que ocorreu a coleta dos dados.⁴⁵ A partir dessa definição, os participantes foram divididos em dois grupos. No primeiro, foram incluídos aqueles que possuíam até dois anos de aprendizagem da língua inglesa. No segundo, foram distribuídos os participantes com quatro anos ou mais de instrução na língua estrangeira.

Cabe ressaltar que, para efeitos dos resultados referentes à Amostra de Fala 1 (cf. 3.3), tal variável não foi levada em consideração, uma vez que os dados dessa amostra correspondem a palavras do PB, configurando uma situação em que o tempo de estudo da língua estrangeira não tem relevância. Por analogia, a variável foi igualmente desconsiderada na análise que contrasta os resultados da Amostra de Fala 2 com os resultados da Amostra de Fala 1 (cf. seção 3.5).

Os resultados da análise perceptual de PEREYRON (2008) indicaram, ainda que de maneira tímida, o desfavorecimento da epêntese pelo fator ‘avançado’ (mais de 4 anos de estudo de inglês): peso relativo de 0,47. A hipótese de que falantes com mais tempo de instrução em inglês apliquem menos a epêntese é igualmente sustentada neste trabalho, uma vez que se presume que indivíduos em contato com a língua inglesa há mais tempo estão mais cientes de que, nessa língua, a inserção da vogal não é usual.

2.3 Instrumentos para a interpretação dos dados coletados

Para executar a interpretação dos dados coletados na Amostra de Fala 1 e na Amostra de Fala 2 desta pesquisa, contou-se com o auxílio de dois programas computacionais: Praat, v. 5.1.07 (BOERSMA e WEENINK, 2009), e GoldVarb X, v. 3.0b3 (SANKOFF, TAGLIAMONTE e SMITH, 2005).

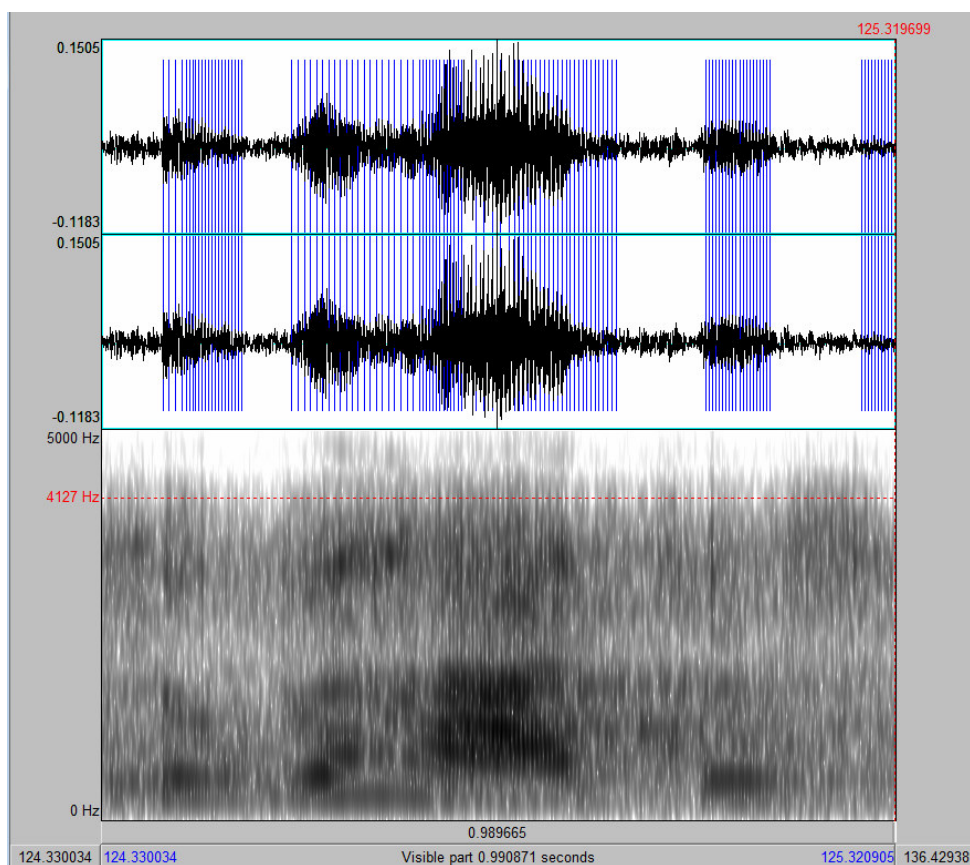
Praat é um *software* comumente utilizado na síntese e na análise de dados de fala. A partir de arquivos de áudio de diferentes extensões (no caso desta pesquisa, .mp3), o programa gera um desenho contínuo das ondas sonoras captadas nas gravações e a partir delas cria

⁴⁵ O tempo de instrução na língua estrangeira foi verificado a partir do nível (isto é, o semestre) do curso de idiomas em que o participante estava matriculado e a partir da declaração feita pelo próprio informante no Formulário sobre Informações Pessoais (cf. Anexo VII).

espectrogramas, i.é, representações visuais dos sons contidos nesses arquivos. Através desses espectrogramas, o usuário do programa consegue literalmente enxergar os sons que foram gravados.

Nesta pesquisa, o som que se buscou enxergar foi aquele produzido entre os *clusters* consonantais presentes nos dados de fala coletados, i.é, a epêntese vocálica. No presente estudo, considerou-se caso de epêntese somente segmentos com duração $\geq 0,03$ segundos.

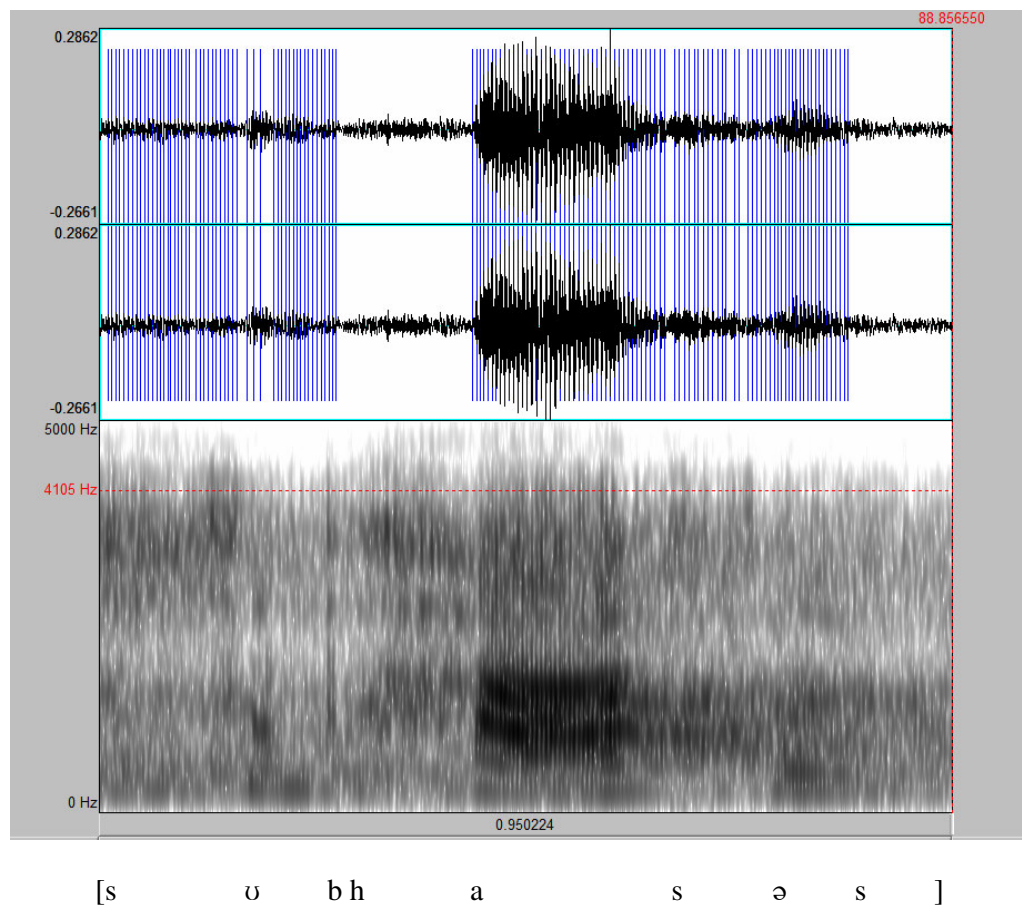
A parte inferior da imagem a seguir ilustra um exemplo de espectrograma em que ocorre a inserção da vogal [ɪ] no vocábulo *sub-raças*.



[s u b ɪ x a s ə s]

Observe-se que, neste exemplo, as vogais podem ser identificadas sem dificuldades através das áreas mais escuras do espectrograma que em geral se concentram na parte inferior do mesmo. Com relação à vogal epentética [ɪ], sua produção pode ser identificada pela pequena esfera de tom mais escuro que aparece logo acima de sua representação fonética, na

parte de baixo do espectrograma. Com o auxílio da próxima imagem, que apresenta um caso de ausência de epêntese, é possível entender melhor essa forma de identificação da vogal.



Nessa imagem, percebe-se que, para o mesmo vocábulo *sub-raças* (agora produzido por outro informante), a área mais escura antes presente entre [b] (aqui bem identificado pela barra vertical que representa a explosão dessa oclusiva) a fricativa subsequente agora desapareceu.

A identificação das vogais epentéticas através da simples visualização das áreas mais escuras, no entanto, nem sempre é eficaz, já que essas vogais geralmente têm duração curta na fala e podem, em razão disso gerar espectros que se confundem com os das consoantes circundantes. Soma-se a isso o fato de que o barulho captado no local da gravação tende a escurecer as imagens dos espectrogramas, o que pode causar confusão na hora de definir o que é ruído e o que é vibração das cordas vocais. Em razão disso, utilizou-se adicionalmente o recurso de visualização de pulsos glotais oferecido pelo programa.

Os pulsos glotais – representados pelas barras verticais azuis que cortam as ondas que aparecem na parte superior das ilustrações acima – permitem identificar os pontos em que houve vibração das cordas vocais, ou seja, os pontos em que houve produção de segmentos vozeados (como as vogais epentéticas). Neste estudo, à semelhança de PEREYRON (2008), considerou-se a ocorrência de no mínimo dois pulsos glotais vocálicos entre a consoante perdida e a consoante seguinte para que se admitisse um caso de epêntese.

O outro programa utilizado nesta pesquisa, GoldVarb X, é um aplicativo computacional que executa a análise multivariacional de dados. Desenvolvido na Universidade de York, EUA, o GoldVarb X corresponde a uma versão atualizada do Pacote VARBRUL 2S e tem como uma das principais vantagens o fato de que permite o acesso a todos os programas do pacote se sua versão anterior em um único arquivo executável (no Pacote VARBRUL 2S, os programas deviam ser executados individualmente). BRESCANCINI (2002) apresenta uma descrição detalhada sobre os programas que compõem o Pacote VARBRUL.

Os dados obtidos com o auxílio do programa Praat foram codificados e, em seguida, submetidos à análise do aplicativo GoldVarb X. Após a correção dos erros de codificação detectados e a criação dos arquivos de condições para cada uma das três análises executadas (Amostra de Fala 1, Amostra de Fala 2 e Amostra de Fala 1 + Amostra de Fala 2), foram gerados os arquivos de células, que apresentaram os percentuais de aplicação da epêntese para cada fator.

Através do arquivo de células, foi possível identificar os casos de *knockout*, i.é, situações em que houve 0% ou 100% de aplicação da epêntese para um determinado fator.⁴⁶ Uma vez eliminados os *knockouts*, os arquivos de células corrigidos serviram de *input* para a execução da análise binomial, que calculou o peso relativo de cada um dos fatores relacionados à inserção vocálica.

O peso relativo, valor numérico que varia entre 0,00 e 1,00, indica, através de sua comparação com o ponto neutro (= 0,50), se um determinado fator favorece (peso relativo > 0,50) ou inibe (peso relativo < 0,50) a ocorrência do fenômeno. BRESCANCINI (2002) observa, no entanto, que o ponto de referência não deve ser adotado como o único critério para verificar se um fator de uma determinada variável contribui ou não para a realização do fenômeno. Nesse sentido, a autora argumenta que o peso relativo de um fator também deve

⁴⁶ Os *knockouts* não são admitidos pelo programa na etapa de execução da análise binomial e, por isso, precisam ser corrigidos, seja pela eliminação do dado, seja pela sua recodificação.

ser interpretado de modo relativo, ou seja, “(...) considerando a ordenação relativa entre os pesos dos fatores do grupo.” (BRESCANCINI, 2002:48).

Ainda com base no arquivo de células, a realização de rodadas-teste entre grupos de fatores permitiu que se detectasse um problema de ordem teórico-metodológica: a baixa ortogonalidade de alguns desses grupos. GUY (1998) argumenta que os grupos de fatores devem ser ‘ortogonais’, ou quase ‘ortogonais’, ou seja, “(...) eles devem coocorrer livremente, e não ser sub- ou supercategorias uns dos outros.” (GUY, 1998:31).

Um exemplo de baixa ortogonalidade verificado neste estudo está no grupo ‘vozeamento da consoante perdida’. Esse grupo apresenta a informação sobre o tipo de marcação (positiva ou negativa) para o traço [vozeado]; essa informação, no entanto, já está contida nos fatores que compõem o grupo ‘tipo de consoante perdida’ (para uma consoante [b], por exemplo, já se sabe que a marcação de [vozeado] é positiva).

Em razão dos casos de baixa ortogonalidade detectados em rodadas gerais (i.é, rodadas que continham todos os grupos de fatores), optou-se pela realização de duas rodadas para cada uma das duas análises executadas. Assim, a distribuição dos grupos de fatores em cada uma das rodadas ficou de acordo com o que está no Quadro 4, na próxima página. A rodada 1 excluiu os grupos que causavam problemas de ortogonalidade (‘vozeamento da consoante perdida’, ‘vozeamento do contexto seguinte’, ‘sexo’, ‘idade’ e ‘proficiência em inglês’⁴⁷), enquanto a rodada 2 acrescentou esses grupos, em substituição às variáveis com as quais criavam conflitos.

⁴⁷ Os grupos ‘sexo’, ‘idade’ e ‘proficiência em inglês’ têm problemas de ortogonalidade em relação ao grupo ‘informante’, uma vez que, para um dado indivíduo, as informações sobre gênero, idade e grau de instrução na LE são únicas.

Quadro 4 - Distribuição das variáveis por rodada e por análise realizada

		Análise realizada	
		Amostra de Fala 1 (PB)	Amostra de Fala 2 (inglês)
Rodada	1	Contexto morfológico Tipo de consoante perdida Tipo de contexto seguinte Acento Tipo de <i>cluster</i> Informante	Contexto morfológico Tipo de consoante perdida Tipo de contexto seguinte Acento Tipo de <i>cluster</i> Informante
	2	Contexto morfológico Vozeamento da consoante perdida Vozeamento do contexto seguinte Acento Tipo de <i>cluster</i> Sexo Idade	Contexto morfológico Vozeamento da consoante perdida Vozeamento do contexto seguinte Acento Tipo de <i>cluster</i> Sexo Idade Proficiência em inglês

A partir da interpretação comparativa entre os pesos relativos de cada fator nas rodadas realizadas, foi possível analisar as relações de favorecimento e de desfavorecimento para a aplicação do fenômeno da epêntese, possibilitando-se, assim, a verificação das hipóteses formuladas para os grupos de fatores considerados para a análise da inserção vocálica nos dados da Amostra de Fala 1 e da Amostra de Fala 2. A apresentação e a discussão dos resultados obtidos nessas amostras, assim como nos testes de verificação de transparência de prefixos em PB e em inglês como LE, são o foco do próximo capítulo deste trabalho.

3 ANÁLISE E DISCUSSÃO DOS RESULTADOS

No presente capítulo, são apresentadas, inicialmente, as respostas obtidas em cada um dos instrumentos aplicados com vistas à verificação da transparência de prefixos e à obtenção de dados de fala em português brasileiro (PB) e em inglês para investigar os fatores relacionados à epêntese medial em ambas as línguas (cf. seções 2.1.1 a 2.1.4). A partir dessas respostas, procede-se à interpretação e à discussão das mesmas, de acordo com as hipóteses e com os objetivos apresentados para cada um dos instrumentos.

A análise e a discussão dos resultados estão distribuídas da seguinte maneira neste capítulo: nas seções 3.1 e 3.2, são abordados os dados referentes ao instrumento de verificação de transparência para o PB e para o inglês, respectivamente; na seção 3.3, são discutidos os resultados da Amostra de Fala 1, referente à análise de inserção da epêntese medial em PB; na seção 3.4, procede-se à interpretação dos dados coletados na Amostra de Fala 2, em que se objetivou verificar a aplicação da epêntese medial em dados de fala do inglês; por fim, na seção 3.5, são equiparados alguns resultados dos dados de fala da língua portuguesa com alguns dos resultados referentes à língua inglesa.

3.1 Teste de verificação de transparência dos prefixos em PB

Após a elaboração e aplicação deste primeiro instrumento de coleta de dados (cf. seção 2.1.1), os resultados foram tabulados de acordo com o que está representado no Quadro 5, a seguir. Dado o número elevado de informações constantes desse quadro, optou-se por interpretá-lo de forma segmentada, de acordo com o que está exposto nos próximos parágrafos.

Inicialmente, faz-se necessário esclarecer o modo como se deu a classificação das respostas obtidas pelos informantes. Como cada um dos entrevistados devia circular, em cada uma das 46 palavras, a parte que julgasse corresponder ao núcleo do significado das mesmas, já era esperado que houvesse grande variabilidade no número de respostas para o exercício proposto. De fato, a diversidade encontrada nos dados informados tornou inviável a tarefa de discriminar cada um dos segmentos de palavras julgados como ‘detentores do núcleo do significado’. Uma vez que o enfoque deste instrumento foi analisar exclusivamente o isolamento ou não de determinados elementos prefixais nas palavras listadas, optou-se pela separação dos dados em três categorias: (i) Prefixo isolado (falso ou verdadeiro), (ii) Outro

prefixo isolado e (iii) Prefixo não isolado. Todos e quaisquer elementos isolados pelo informante sob a forma de sufixo (ou seja, todas as partes não circuladas após os segmentos identificados como bases dos vocábulos) foram ignorados nessa classificação.

Quadro 5 - Resultado do teste de transparência de prefixos em PB

Vocábulo	Frequência / nº de informantes			Vocábulo	Frequência / nº de informantes		
	Prefixo isolado	Outro prefixo isolado	Prefixo não isolado		Prefixo isolado	Outro prefixo isolado	Prefixo não isolado
01 <i>obedecer*</i>	2/20	4/20	14/20	24 <i>eczema</i>	1/20	1/20	18/20
02 <i>ablação</i>	3/20	2/20	15/20	25 <i>subtração*</i>	8/20	1/20	11/20
03 <i>adaptar</i>	3/20	1/20	16/20	26 <i>adjunto</i>	12/20	0/20	8/20
04 <i>suburitano*</i>	3/20	2/20	15/20	27 <i>sobgrave</i>	16/20	0/20	4/20
05 <i>obviar</i>	3/20	0/20	17/20	28 <i>obediente*</i>	0/20	4/20	16/20
06 <i>subsolo</i>	16/20	0/20	4/20	29 <i>sublocar</i>	17/20	0/20	3/20
07 <i>abacate*</i>	0/20	5/20	15/20	30 <i>eclético</i>	0/20	2/20	18/20
08 <i>eclipse</i>	2/20	3/20	15/20	31 <i>adrenalina*</i>	0/20	6/20	14/20
09 <i>sobalçou</i>	7/20	3/20	10/20	32 <i>sobdominante</i>	17/20	0/20	3/20
10 <i>adotado*</i>	0/20	11/20	9/20	33 <i>abnegar</i>	17/20	1/20	2/20
11 <i>abaxial</i>	13/20	3/20	4/20	34 <i>sobejo*</i>	3/20	5/20	12/20
12 <i>subarquado</i>	17/20	0/20	3/20	35 <i>obstruir</i>	3/20	1/20	16/20
13 <i>sobrado*</i>	0/20	3/20	17/20	36 <i>administração</i>	9/20	0/20	11/20
14 <i>adligação</i>	13/20	1/20	6/20	37 <i>ecumênico*</i>	0/20	4/20	16/20
15 <i>obturar*</i>	3/20	0/20	17/20	38 <i>abstenção</i>	8/20	0/20	12/20
16 <i>eclesiástico</i>	0/20	5/20	15/20	39 <i>subnutrido</i>	16/20	1/20	3/20
17 <i>sobescavar</i>	16/20	0/20	4/20	40 <i>adolescente*</i>	1/20	4/20	15/20
18 <i>ecdêmico</i>	4/20	3/20	13/20	41 <i>obnubilar</i>	14/20	1/20	5/20
19 <i>abestado*</i>	1/20	14/20	5/20	42 <i>ecmnésia</i>	7/20	5/20	8/20
20 <i>subarrendar</i>	15/20	2/20	3/20	43 <i>obrigou*</i>	0/20	4/20	16/20
21 <i>obumbrar</i>	7/20	3/20	10/20	44 <i>advogada</i>	5/20	2/20	13/20
22 <i>economia*</i>	0/20	2/20	18/20	45 <i>sobestar</i>	12/20	1/20	7/20
23 <i>abdicar</i>	6/20	1/20	13/20	46 <i>ablitense*</i>	3/20	1/20	16/20

O * indica vocábulo não prefixado ou com prefixo diferente daqueles investigados.

Na categoria 'Prefixo isolado (falso ou verdadeiro)', identificado no Quadro 5 sob a forma reduzida 'Prefixo isolado', foram registrados os casos em que o informante separou, à esquerda das bases, as formas gráficas idênticas aos seis prefixos analisados, inclusive nos vocábulos em que *ab-*, *ad-*, *ec-*, *ob-*, *sob-* e *sub-* não correspondem a prefixos (esses vocábulos estão marcados com um asterisco). Note-se que, para 17 *sobescavar* e 20 *subarrendar*, que possuem tanto as bases *escavar* e *arrendar*, como *cavar* e *renda*,

respectivamente, foram admitidas, na categoria ‘Prefixo isolado (falso ou verdadeiro)’, tanto a separação pelo informante de *sob-* e *sub-*, quanto a de *sobes-* e de *subar-* (ou *suba-*). Admitiu-se, pois, que o informante, ao ser capaz de depreender as bases *cavar* e *renda* a partir de *sobescavar* e *subarrendar*, também teria condições de depreender as bases *escavar* e *arrendar* desses mesmos vocábulos, separando os prefixos *sob-* e *sub-* ora investigados.

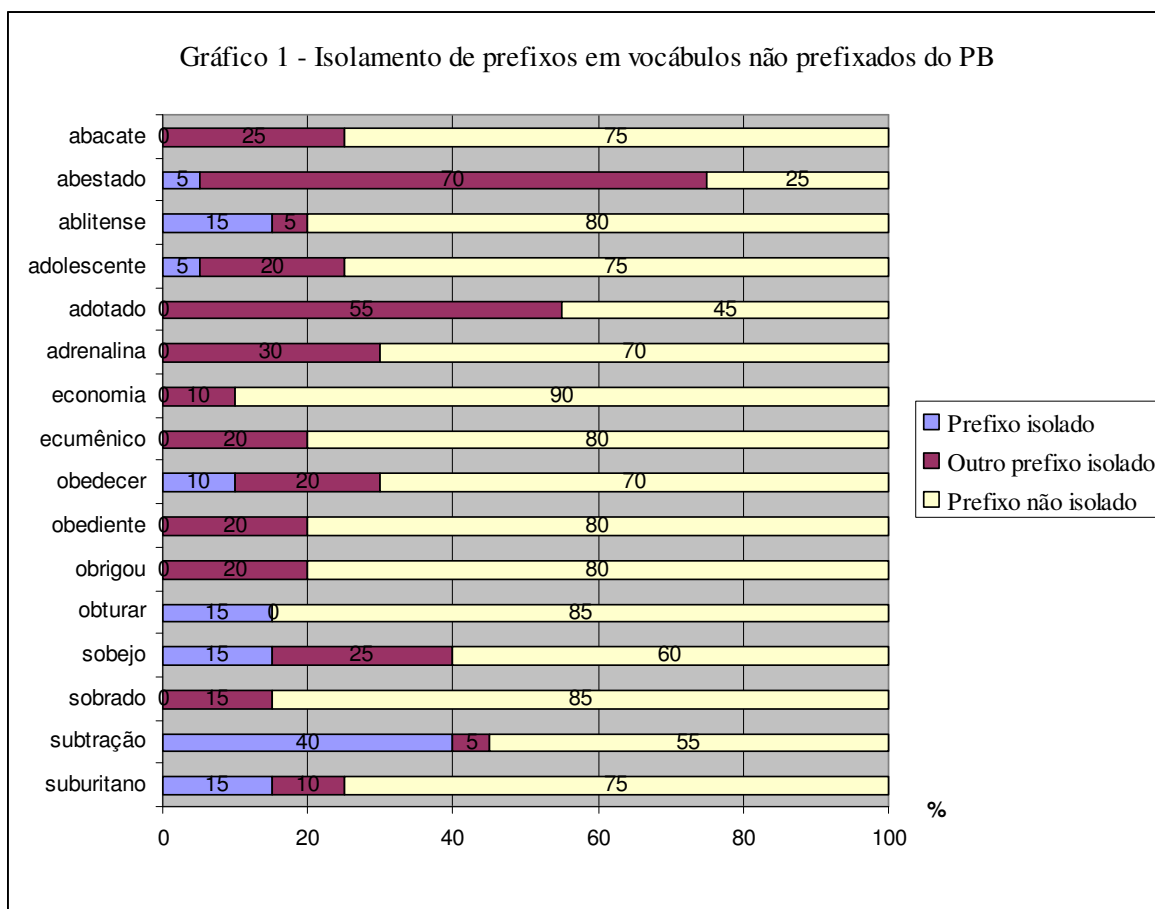
Na segunda categoria, ‘Outro prefixo isolado’, incluíram-se os casos em que o respondente havia deixado fragmentos à esquerda das partes circuladas que diferiam das formas gráficas equivalentes aos prefixos analisados. Esses fragmentos correspondiam tanto a prefixos diferentes dos que estavam sendo investigados (por exemplo, *a-*, em 10 *adotado* e em 19 *abestado*) quanto a partes sem significação precisa (por exemplo, *su-*, em 04 *suburitano*, *aba-*, em 07 *abacate*, *obu-*, em 21 *obumbrar* e *adre-*, em 31 *adrenalina*).⁴⁸

Para a categoria ‘Prefixo não isolado’, finalmente, foram considerados os casos em que o respondente circulou toda a parte inicial da palavra, ou seja, não identificou a existência de qualquer prefixo no vocábulo. Note-se que nessa coluna também foram incluídos os casos em que o informante considerou o prefixo (ou a forma gráfica equivalente ao prefixo) como sendo o detentor do núcleo do significado da palavra (por exemplo, casos em que, numa palavra prefixada como *abstenção*, o informante circulou apenas *ab-*). Vale registrar que, toda vez que esse tipo de dado era atestado, verificava-se se o informante havia circulado os afixos igualmente nas demais palavras para confirmar se ele havia compreendido o exercício. Nenhum caso como esse foi constatado e, além disso, foram poucos os casos em que somente o prefixo (ou o falso prefixo) foi circulado na palavra.

Numa análise prévia, verificamos nos resultados duas possibilidades de interferência na execução do teste (cf. 2.1.1), quais sejam: (i) os informantes poderiam se basear nos critérios de separação silábica no momento de separar as bases dos prefixos e (ii) os informantes poderiam se basear nas semelhanças gráficas entre as palavras prefixadas e as não prefixadas (e, a partir disso, separar as bases dos prefixos indiscriminadamente, mesmo nos casos em que não havia prefixos). Em relação à situação (i), a simples análise visual dos segmentos que foram circulados pelos falantes no instrumento foi suficiente para confirmar que a ordem do exercício havia sido entendida por todos (i.é., não houve nenhum caso em que o entrevistado isolou categoricamente as bases dos afixos pelo critério de separação silábica).

⁴⁸ Os resultados mostram que o isolamento de partes sem significação foi reduzido.

Para discutir a situação (ii), elaborou-se o Gráfico 1 abaixo, que apresenta a distribuição dos percentuais para as categorias ‘prefixo isolado’, ‘outro prefixo isolado’ e ‘prefixo não isolado’ para cada um dos 16 vocábulos não prefixados que foram incluídos no primeiro instrumento.

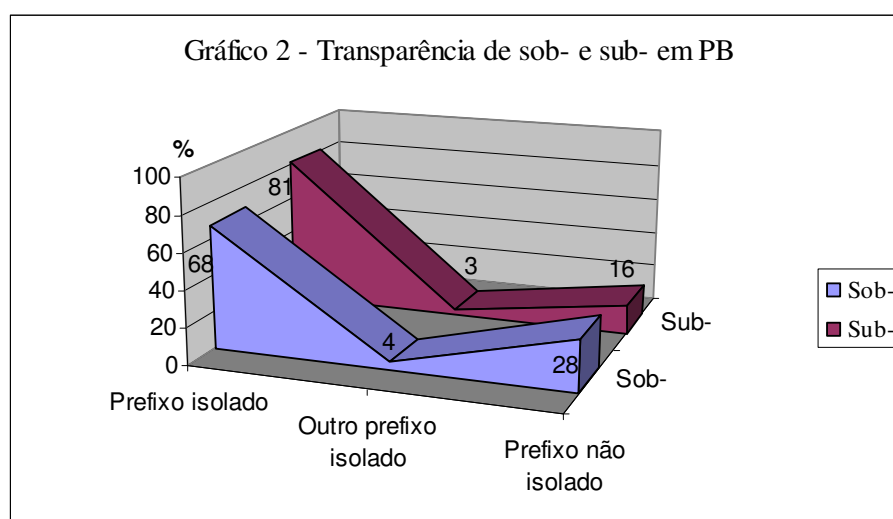


Como se pode observar através dos segmentos de barra na cor azul desse gráfico, os percentuais de isolamento de formas idênticas aos prefixos *ad-*, *ad-*, *ec-*, *ob-*, *sob-* e *sub-* nas palavras não prefixadas é mínimo, não ultrapassando, à exceção do vocábulo *subtração*,⁴⁹ a faixa de 15%. Através das barras em amarelo, nota-se que a grande maioria dos informantes ‘acertadamente’ deixou de isolar prefixos nessas palavras. Apenas as palavras *abestado* e *adotado* apresentaram índices inferiores a 50% de não isolamento, mas isso é justificado pelo

⁴⁹ No caso do vocábulo *subtração*, é provável que o principal critério adotado para isolar ‘sub’ de ‘tração’ foi o fato de que essa última parte exista na língua como forma livre, e não porque ‘sub’ tenha sido reconhecido como prefixo pelos informantes. Prova disso está no fato de que a palavra *suburitano*, igualmente iniciada com ‘sub’, teve esse ‘falso prefixo’ isolado em apenas 15% dos casos.

fato de que os informantes identificaram nelas o prefixo *a-*, (como mostram os percentuais elevados para a categoria 'outro prefixo isolado'). Isso posto, é seguro dizer que os informantes realizaram a tarefa sem serem influenciados pelos dados do próprio exercício, o que garante a fiabilidade do mesmo.

A primeira constatação de relevância sobre os resultados obtidos trata da transparência dos prefixos *sob-* e *sub-*. Os percentuais de isolamento, distribuídos conforme as três categorias apresentadas no Quadro 5 podem ser visualizados no Gráfico 2, abaixo.



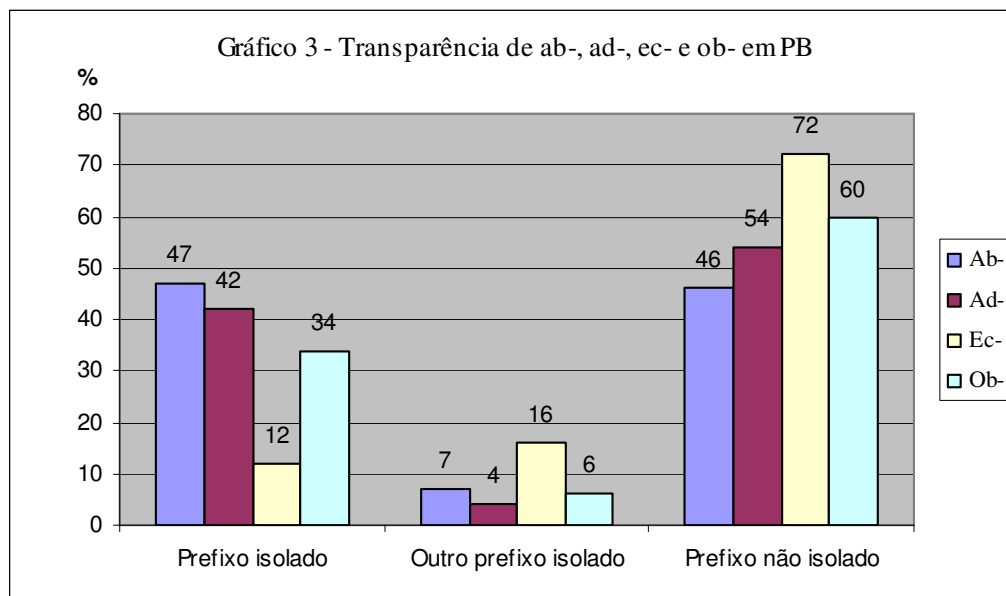
Os prefixos *sob-* e *sub-* apresentaram os maiores índices de isolamento nos resultados obtidos para o teste de verificação de transparência em PB. Esses índices, convertidos em taxas percentuais no Gráfico 2, evidenciam que os informantes entrevistados identificam os vocábulos que detêm esses afixos através da estrutura [prefixo [base]].

Uma das prováveis razões para o fato de o prefixo *sub-* ser facilmente isolado está relacionada à recorrência de palavras com esse afixo na língua portuguesa. Consultando-se o Dicionário Aurélio Eletrônico (1999), por exemplo, verifica-se a existência de aproximadamente 400 vocábulos com o prefixo *sub-*. A ocorrência elevada desse afixo propicia a memorização do mesmo por parte do falante.

Para o caso de *sob-*, que aparece em uma lista menor de vocábulos (menos de 50 palavras), cremos que sua semelhança com a grafia e com o significado de *sub-* pode estar relacionada aos índices elevados de isolamento (e ao conseqüente *status* de prefixo

transparente). Observe-se que, contrastando-se os resultados para os dois afixos (cf. Gráfico 2), o percentual de não isolamento de *sob-* equivale a quase o dobro do percentual de não isolamento de *sub-* (28% contra 16%). Em contraste com os demais prefixos, como veremos a seguir, observa-se, no entanto, que a taxa de isolamento de *sob-* é bem maior.

O Gráfico 3, abaixo, apresenta os percentuais de isolamento para os prefixos *ab-*, *ad-*, *ec-* e *ob-* no português brasileiro.



Comparando-se o gráfico acima com o anterior, é possível ver com clareza a distinção que os falantes de PB fazem no reconhecimento dos prefixos *sob-* e *sub-* e dos prefixos *ab-*, *ad-*, *ec-* e *ob-*. O percentual de isolamento para todos os prefixos desse último grupo é bem menor.

Analisando os resultados apresentados no Gráfico 3, percebe-se que o tratamento dado aos vocábulos com prefixo *ab-* e com prefixo *ad-* é semelhante. Há uma pequena diferença percentual entre os dois (5%) mostrando que *ab-* foi isolado mais vezes que *ad-*. Essa diferença aumenta se forem somadas as duas primeiras categorias ('prefixo isolado' e 'outro prefixo isolado'). Juntas, seus resultados mostram que os falantes enxergam algum prefixo em 54% dos vocábulos com *ab-* e em 48% dos vocábulos com o prefixo *ad-*.

No entanto, ainda que somadas as taxas das duas primeiras colunas para cada um dos dois prefixos, os percentuais não atingem um valor que permita classificá-los como transparentes para os falantes de PB. As taxas de não isolamento de *ab-* e de *ad-* (46% e 54%, respectivamente) atestam, se pensarmos na existência de um *continuum* de transparência (cf. 1.2), a posição medial desses afixos nessa escala.

Vale ainda observar que os vocábulos 11 *abaxial*, 33 *abnegar*, 14 *adligação* e 26 *adjunto* (cf. Quadro 5) contribuíram para a elevação do percentual de reconhecimento dos prefixos *ab-* e *ad-*. Essas palavras apresentaram índices individuais de isolamento do afixo correspondentes a, no mínimo, 60%. A justificativa para isso possivelmente se encontra no fato de que, nessas palavras, a identificação dos afixos é facilitada porque as bases a que são adjungidos correspondem a formas livres na língua portuguesa (*axial*, *negar*, *ligação* e *junto*, respectivamente).

Com relação aos prefixos *ab-* e *ad-*, portanto, não é possível dizer que são transparentes, nem que são opacos, já que os percentuais de isolamento dos mesmos se encontram numa faixa medial do *continuum* de transparência (entre 40% e 60%), o que impossibilita fazer qualquer outra afirmação sobre seu *status* sincrônico na língua. A situação é diferente no que diz respeito aos dois prefixos restantes, quais sejam, *ec-* e *ob-*.

Verificando-se a taxa de isolamento do prefixo *ec-*, fica claro que o grau de transparência do mesmo é mínimo (seu índice de isolamento foi de apenas 12%), o que indica fortemente que esse afixo corresponde a uma forma já opaca na língua. Chama-se também atenção para o percentual elevado (se contrastado com os demais prefixos) de isolamento de outra forma prefixal nos vocábulos com *ec-* (16%, valor superior à coluna ‘prefixo isolado’). Os informantes, mesmo nos casos em que identificaram algum afixo no início dos vocábulos com *ec-*, não foram capazes de identificar com clareza qual era esse prefixo e acabaram isolando outra forma (*ecd-*, *ecze-*, *ecl-*, dentre outras). Essa informação, somada ao fato de que não houve isolamento de *ec-* em 72% dos casos em que ele apareceu, permite afirmar que tal forma está próxima do extremo opaco do *continuum* de transparência por nós admitido (cf. seção 1.2).

Para o prefixo *ob-*, o percentual de isolamento também é considerado baixo: 34% dos vocábulos apresentaram o afixo separado de suas bases. Esse percentual é ainda mais

reduzido (22%) se for desconsiderado na contagem o vocábulo 41 *obnubilar*, que teve o afixo identificado por 70% dos entrevistados.⁵⁰

Em 60% das palavras iniciadas com o prefixo *ob-* não houve isolamento de prefixo e em 6% delas ocorreu o isolamento de algum outro afixo à esquerda de suas bases. Juntos, esses resultados reforçam a afirmação de que o prefixo *ob-* parece não ser mais considerado transparente no português brasileiro.

A partir dos resultados obtidos nesse primeiro instrumento, sobre a análise da transparência de prefixos que podem motivar a epêntese vocálica em PB, tivemos a confirmação da hipótese, apresentada em 2.1.1, de que os índices de isolamento dos prefixos *sob-* e *sub-* de suas bases (equivalentes a seu grau de transparência na língua) são maiores do que aqueles observados para os prefixos *ab-*, *ad-*, *ec-* e *ob-*. Acredita-se que a razão disso não está apenas no fato de que *sob-* e *sub-* são prefixos mais frequentes no léxico do PB, mas também está no fato de que eles se configuram como prefixos adjuntos, i.é., sílabas que aparecem à esquerda de palavras fonológicas para formarem, juntas, uma palavra fonológica recursiva (cf. 1.2).

Para os demais quatro prefixos, por sua vez, os baixos índices de isolamento são justificados não apenas porque esses afixos são menos recorrentes na língua, mas também porque eles geralmente estão ligados a bases não existentes como formas livres na língua, frequentemente eruditas (especialmente no caso do prefixo *ec-*). Assume-se que, para os casos de palavras eruditas prefixadas, a distância temporal entre sua origem (a língua clássica) e seu uso corrente compromete a identificação precisa de sua estrutura interna por parte dos usuários da língua moderna.

A hipótese principal que podemos extrair com base nos resultados deste teste de verificação de transparência de prefixos relacionados à epêntese em PB é a seguinte: acredita-se que vocábulos com os prefixos *sob-* e *sub-*, que têm graus elevados de transparência na língua, apresentem percentuais elevados de aplicação de epêntese vocálica se comparados com os vocábulos prefixados por *ab-*, *ad-*, *ec-* e *ob-*, ou que possuam *clusters* consonantais no interior da base. A resposta para tal hipótese foi obtida com a análise dos dados coletados na Amostra de Fala 1, que é apresentada na seção 3.3.

⁵⁰ É possível admitir que esse percentual elevado de isolamento de *ob-* em *obnubilar* seja decorrente da consciência dos indivíduos de que, em PB, o encontro *bn* nunca ocorre no interior de uma base, mas apenas na fronteira entre um prefixo e uma base (como em 36 *subnutrido*).

3.2 Teste de verificação de transparência dos prefixos em inglês

Este segundo teste de verificação de transparência foi feito com o objetivo de analisar a competência lexical de aprendizes brasileiros de inglês no que diz respeito ao reconhecimento da estrutura interna de vocábulos prefixados por *arch-*, *out-*, *post-* e *sub-* em inglês (cf. 2.1.2). O Quadro 6 abaixo apresenta o resultado desse teste.

Quadro 6 - Resultado do teste de transparência de prefixos em inglês

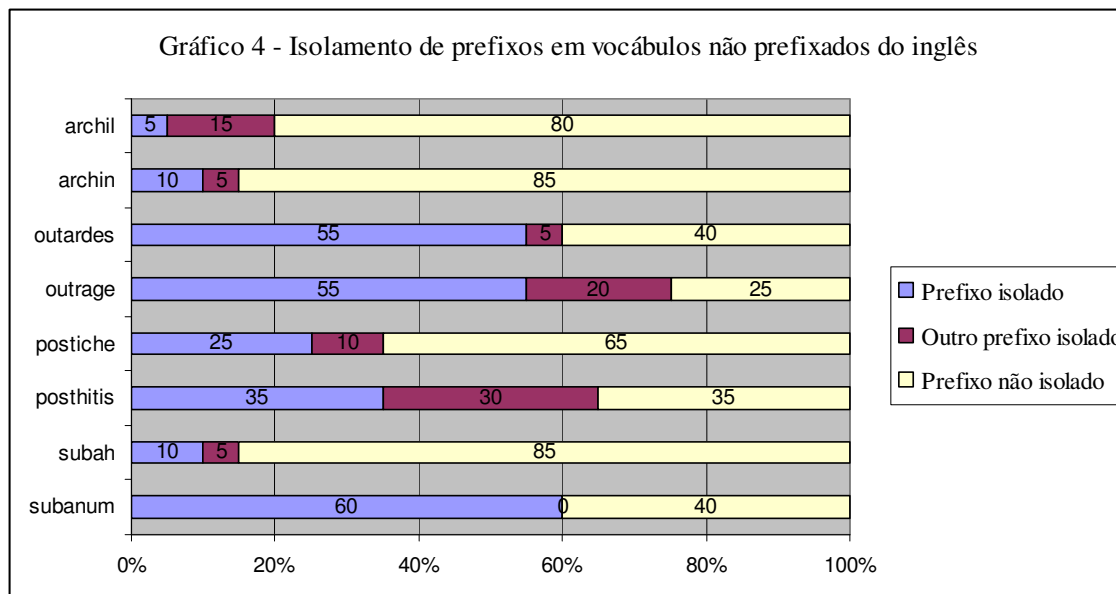
Vocábulo	Frequência / n° de informantes			Vocábulo	Frequência / n° de informantes		
	Prefixo isolado	Outro prefixo isolado	Prefixo não isolado		Prefixo isolado	Outro prefixo isolado	Prefixo não isolado
01 <i>outrage*</i>	11/20	4/20	5/20	15 <i>subah*</i>	2/20	1/20	17/20
02 <i>archbishop</i>	13/20	1/20	6/20	16 <i>postconsonantal</i>	18/20	1/20	1/20
03 <i>outdated</i>	17/20	1/20	2/20	17 <i>outrun</i>	15/20	1/20	4/20
04 <i>postocular</i>	18/20	1/20	1/20	18 <i>subnitrate</i>	20/20	0/20	0/20
05 <i>archil*</i>	1/20	3/20	16/20	19 <i>archin*</i>	2/20	1/20	17/20
06 <i>subcutaneous</i>	18/20	0/20	2/20	20 <i>postdate</i>	18/20	0/20	2/20
07 <i>outfighting</i>	19/20	0/20	1/20	21 <i>archfiend</i>	11/20	0/20	9/20
08 <i>posthitis*</i>	7/20	6/20	7/20	22 <i>subanum*</i>	12/20	0/20	8/20
09 <i>archduke</i>	17/20	0/20	3/20	23 <i>outmaneuver</i>	18/20	0/20	2/20
10 <i>sublet</i>	14/20	0/20	6/20	24 <i>subarid</i>	19/20	0/20	1/20
11 <i>outbreathe</i>	19/20	0/20	1/20	25 <i>postwar</i>	18/20	0/20	2/20
12 <i>postiche*</i>	5/20	2/20	13/20	26 <i>outardes*</i>	11/20	1/20	8/20
13 <i>archpriest</i>	19/20	0/20	1/20	27 <i>archdeacon</i>	17/20	0/20	3/20
14 <i>postfix</i>	17/20	1/20	2/20	28 <i>subdominant</i>	19/20	0/20	1/20

O * indica vocábulo não prefixado.

A tabulação do Quadro 6 obedece aos mesmos padrões de classificação dos resultados para o teste de verificação dos prefixos relacionados à epêntese em PB e, portanto, o detalhamento sobre a subdivisão dos resultado em ‘Prefixo isolado’, ‘Outro prefixo isolado’ e ‘Prefixo não isolado’ pode ser conferido na seção 3.1.

A primeira tarefa que deve ser realizada na análise dos resultados do Quadro 6 é verificar se os informantes respondentes se basearam nas semelhanças de grafia existentes entre os vocábulos da lista e consequentemente isolaram formas que originalmente não correspondem aos prefixos *arch-*, *out-*, *post-* e *sub-*. Tal verificação pode ser feita a partir do Gráfico 4 (cf. próxima página).

Com base na interpretação dos resultados do Gráfico 4, percebe-se que o percentual de isolamento de formas homógrafas aos prefixos analisados em inglês (cf. barras azuis) não pode ser desprezado em alguns dos vocábulos, especialmente em *outardes* (55%), *outrage* (55%), *posthitis* (35%) e *subanum* (60%). Pelo menos duas observações podem ser feitas a respeito desses resultados.

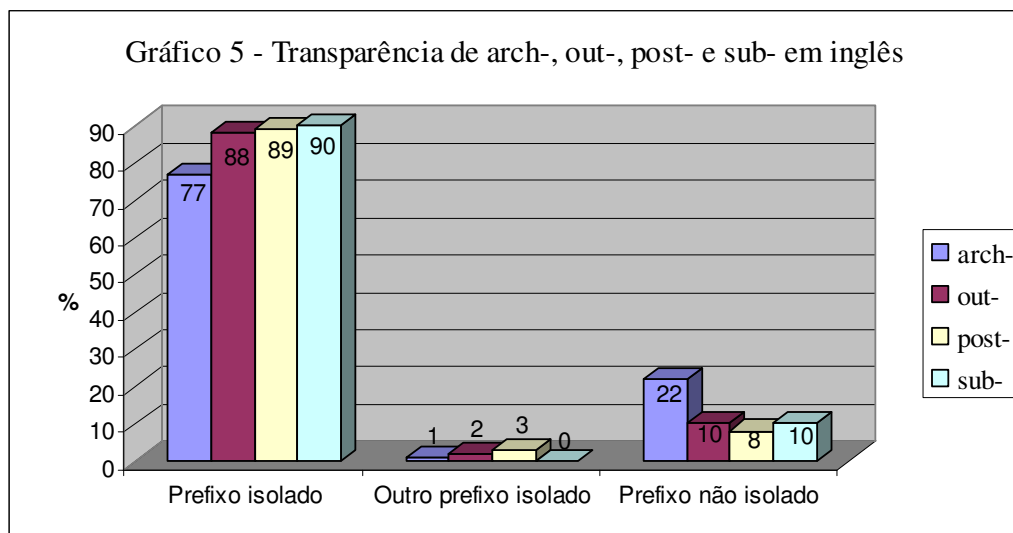


A primeira observação está relacionada ao instrumento de coleta como um todo. Como mostra o Quadro 6 supra, o teste constituiu-se de um conjunto de apenas 28 vocábulos, todos iniciados por apenas 4 possibilidades de grafia: *arch-*, *out-*, *post-* ou *sub-*. É possível, pois, que a recorrência dessas formas em tão poucas palavras tenha se sobressaído em alguns momentos na execução do exercício, provocando índices consideráveis de isolamento de prefixos em palavras nas quais não estão presentes.

A outra observação, talvez mais importante, está relacionada ao fato de que as formas *outardes*, *outrage*, *posthitis* e *subanum*, que têm maiores índices de isolamento de um ‘falso prefixo’, correspondem graficamente a formas mais longas que *archil*, *archin* e *subah*, cujos índices de isolamento de um ‘falso prefixo’ são bem menores (5%, 10% e 10%, respectivamente). Parece, pois, que há alguma imposição com relação à extensão mínima de

um vocábulo para que um falante admita a existência de estruturas morfológicas dentro dele, ainda que essas estruturas efetivamente não existam.⁵¹

Os dados apresentados no Gráfico 4, no entanto, não devem ser considerados prejudiciais para efeitos da análise de transparência dos quatro prefixos sob investigação. Prova disso está nos resultados que foram obtidos para os vocábulos do instrumento que de fato contêm os prefixos *arch-*, *out-*, *post-* e *sub-*, conforme o Gráfico 5 abaixo.



Através da primeira sequência de colunas do Gráfico 5, é possível verificar que os índices de isolamento de *arch-*, *out-*, *post-* e *sub-* em vocábulos com esses afixos são bem mais altos do que nas palavras que apresentam grafia inicial idêntica a prefixos, apresentadas no Gráfico 4. Isso mostra que, se, por um lado, os aprendizes de inglês hesitam no momento de depreender prefixos onde sequer existem, por outro, eles não apresentam quase nenhuma dificuldade em localizar esses afixos onde eles de fato estão presentes.

Os resultados do Gráfico 5, além de demonstrarem um percentual elevado de isolamento para os quatro prefixos investigados, revelam uma diferença que merece ser discutida, ainda que não influencie na inegável constatação de que estamos diante de quatro prefixos transparentes. Trata-se pois, dos percentuais verificados para *arch-* nas colunas 'Prefixo isolado' (na qual *arch-* apresenta percentual inferior aos demais, de 77%) e 'Prefixo

⁵¹ Observe-se que, em *archin*, o fato de *in* existir isoladamente na língua inglesa não é razão suficiente para que os informantes façam a separação *arch - in* (registrada em só 10% dos casos).

não isolado' (na qual o prefixo é enquadrado em 22% dos casos, mais do que o dobro de vezes se comparado aos demais afixos nessa categoria).

Pela análise do Quadro 6, é possível constatar que *arch-* é menos identificado no vocábulo 02 *archbishop* (isolamento em 13 de 20 ocorrências) e no vocábulo 21 *archfiend* (isolamento de 11 em 20 ocorrências). A partir dessa informação, fez-se necessário analisar os testes de cada participante individualmente para se tentasse identificar alguma informação que justificasse tal fato. O que se encontrou nesses dados foi um número elevado de isolamento de *arch-* não como prefixo, mas como detentor exclusivo do núcleo do significado da palavra. Isso é justificável na medida em que se admite que os informantes possam ter se baseado na existência de *arch* ('arco') como forma livre na língua inglesa para atribuírem *arch-* (prefixo) o núcleo do significado dessas palavras.

A conclusão que se faz a partir dos resultados do teste de verificação de transparência de prefixos do inglês é que, como previsto inicialmente, os prefixos *arch-*, *out-*, *post-* e *sub-* apresentam grau elevado no *continuum* de transparência admitido para o inglês. Isso pode, em princípio, ser motivado pela semelhança que os prefixos têm com formas livres da língua inglesa ou com alguns prefixos da língua portuguesa. É mais interessante pensar, no entanto, que estamos diante de quatro prefixos *adjuntos*, que se anexam a palavras fonológicas existentes como formas livres no léxico do inglês e que, por isso, são mais facilmente identificados.

Uma vez que *arch-*, *out-*, *post-* e *sub-* estão próximos ao extremo 'transparente' do *continuum* de transparência, resta-nos admitir que essas formas devem implicar taxas elevadas de epêntese vocálica nas palavras em que estão presentes (cf. 1.2). A verificação disso é discutida na seção 3.4.

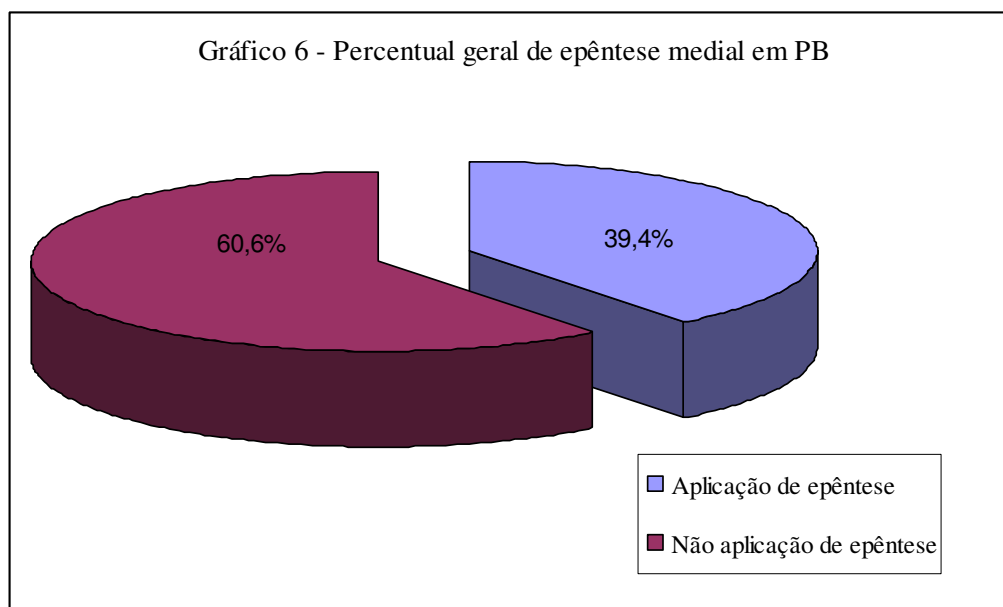
3.3 Amostra de Fala 1 (PB)

Nesta seção, são apresentados e discutidos os resultados referentes aos dados que compõem a Amostra de Fala 1 (cf. 2.1.3), que serviu para a análise da epêntese vocálica em PB. A análise dos resultados obedece à seguinte ordem: primeiramente, em 3.3.1, é mostrado e discutido o percentual geral da epêntese que ocorre em contexto medial em PB; num segundo momento, em 3.3.2 e subseções, são elencados os grupos de fatores selecionados pelas duas rodadas estatísticas realizadas na análise da Amostra de Fala 1 (cf. 2.3), bem como são discutidos os resultados específicos dos fatores que constituem esses grupos selecionados.

3.3.1 Percentual geral de epêntese medial em PB

O percentual geral de aplicação da epêntese em PB corresponde ao número total de vezes em que a inserção vocálica ocorreu dividido pelo número total de contextos nos quais ela poderia ocorrer. Nesse cálculo foram incluídos, portanto, os casos em que o contexto para aplicação de epêntese foi eliminado (através da não produção da consoante perdida ou do contexto seguinte, o que ocorreu em 6 casos), bem como as ocorrências que foram posteriormente excluídas ou recodificadas porque geravam *knockouts* nas rodadas estatísticas.

O Gráfico 6 abaixo apresenta o percentual geral de epêntese medial em PB.



Num total de 624 ocorrências de vocábulos com contexto para aplicação, a epêntese medial foi atestada 246 vezes, o que corresponde a um percentual igual a 39,4 pontos. Esse valor comprova a recorrência da epêntese medial na língua portuguesa falada por porto-alegrenses e indica, sob a perspectiva variacionista, que a inserção vocálica é uma característica presente na fala da capital gaúcha, o que parece não ser registrado no dialeto de

outras (em Florianópolis, por exemplo, a epêntese parece ser desfavorecida, cf. COLLISCHONN, 2002).

Cabe aqui comparar o percentual de aplicação de epêntese de nosso trabalho com o de COLLISCHONN (2002). Nos resultados de COLLISCHONN (2002), 58% dos dados produzidos por falantes porto-alegrenses apresentaram epêntese no interior do vocábulo (aplicação em 107 de um total de 183 ocorrências). Essa diferença razoável entre os percentuais dos dois estudos (58% em COLLISCHONN, 2002, e 39,4% em nosso trabalho) não deve ser vista como uma tendência à redução da epêntese na fala do português em Porto Alegre. É preferível, pois, acreditar que existam diferenças de ordem metodológica entre os dois estudos que expliquem esses resultados diferentes.

Os dados de COLLISCHONN (2002) foram analisados por meio da audição de fitas magnéticas de áudio que compõem o banco de dados do projeto VARSUL, ao passo que o presente trabalho envolveu a análise de arquivos de áudio digitais submetidos a um programa de análise acústica. A partir dessa informação, é correto entender que a análise perceptual dos dados (no caso, a audição de fitas cassete) deva implicar a detecção de um número maior de casos de epêntese do que a análise acústica.⁵² Isso já foi confirmado em estudos como o de PEREYRON (2008), que investigou a epêntese em inglês tanto pela simples audição de arquivos de áudio (33% de aplicação) quanto pela interpretação acústica desses arquivos (8% de aplicação). Não se deve desconsiderar, portanto, que as diferenças resultantes das especificidades de cada estudo, ainda que sejam sobre um mesmo fenômeno, podem implicar resultados finais diferentes.

3.3.2 Resultados das análises estatísticas para dados do PB

Em um momento prévio da análise dos dados de fala do PB, foram executadas algumas rodadas-teste com as codificações feitas, o que permitiu detectar um número total de 11 *knockouts*, referentes a casos em que houve: (i) eliminação da consoante perdida (como em [epi'ezməs] *ecpiesmas*); (ii) eliminação do contexto seguinte (como em [adʒigla'sãw] *adligação*); (iii) realização única de *cluster* inexistente na língua ([pʒ] em [epzi'eməs] *ecpiesmas*); e (iv) realizações de vibrante alveolar no contexto seguinte (como em [adʁo'gar] *ad-rogar*). Os casos (i), (ii) e (iii) foram eliminados da análise estatística. No caso (iv), a

⁵² Soma-se a isso o fato de que parte dos dados de COLLISCHONN (2002) foi obtida de entrevistas com informantes de escolaridade equivalente ao ensino fundamental (nível inferior ao de nossos informantes).

vibrante alveolar (que apareceu somente duas vezes nos dados), foi recodificada como fricativa velar [x].⁵³

A relação dos grupos de fatores analisados por rodada estatística e dos grupos selecionados em cada uma das rodadas são apresentadas no Quadro 7 a seguir. Algumas observações gerais devem ser feitas sobre os resultados expostos nele.

Primeiramente, nota-se que, à exceção de ‘acento’, todos os grupos submetidos a análise na primeira rodada foram selecionados; na segunda rodada, foram eliminados os fatores ‘sexo’ e ‘idade’ apenas (e ‘acento’ agora foi selecionado, ainda que na última colocação). É possível perceber, também, que existe uma relação direta entre os dois primeiros colocados da rodada 2 (‘vozeamento da consoante perdida’ e ‘vozeamento do contexto seguinte’) e o 1º e terceiro colocados da rodada 1 (‘tipo de consoante perdida’ e ‘tipo de contexto seguinte’). Nota-se, ainda, que ‘contexto morfológico’ e ‘tipo de *cluster*’ são selecionados nas duas rodadas, muito embora não fique claro qual grupo tem precedência na seleção (entre uma rodada e outra, a ordem de colocação entre eles de inverte). Por fim, chama-se atenção para o fato de que ‘informante’ foi o segundo grupo de maior relevância na rodada em que foi analisado.

Quadro 7 - Grupos analisados e grupos selecionados por rodada (PB)

	Rodada	
	1	2
Grupos analisados	Contexto morfológico Tipo de consoante perdida Tipo de contexto seguinte Acento Tipo de <i>cluster</i> Informante	Contexto morfológico Vozeamento da consoante perdida Vozeamento do contexto seguinte Acento Tipo de <i>cluster</i> Sexo Idade
Grupos selecionados	1º Tipo de consoante perdida 2º Informante 3º Tipo de contexto seguinte 4º Tipo de <i>cluster</i> 5º Contexto morfológico	1º Vozeamento da consoante perdida 2º Vozeamento do contexto seguinte 3º Contexto morfológico 4º Tipo de <i>cluster</i> 5º Acento

Optamos, aqui, por analisar apenas os grupos de fatores selecionados pelas rodadas. Os grupos ‘sexo’ e ‘idade’, portanto, não estão contemplados na análise. Além disso, com

⁵³ Achamos mais interessante recodificar os dados a excluí-los das rodadas (muito embora isso nem sempre seja possível). No caso específico da vibrante, pareceu-nos mais sensato recodificá-la como [x], já que estudos como o de Monaretto (1997) apontam para o processo de mudança da posição da vibrante de anterior para posterior.

base nas observações gerais feitas para o Quadro 7, propomos segmentar a discussão dos resultados de acordo com o que segue: em 3.3.2.1, são analisados, em conjunto, os resultados referentes aos grupos ‘vozeamento da consoante perdida’ e ‘tipo de consoante perdida’; em 3.3.2.2, procedemos à averiguação conjunta dos resultados para ‘vozeamento do contexto seguinte’ e ‘tipo de contexto seguinte’; em 3.3.2.3, faz-se a análise dos resultados para o grupo ‘informante’; em 3.3.2.4, discutem-se os resultados para ‘contexto morfológico e ‘tipo de *cluster*’; e, por fim, em 3.3.2.5, discute-se o papel do grupo ‘acento’ nas rodadas.

3.3.2.1 Vozeamento da consoante perdida e Tipo de consoante perdida

Os grupos ‘vozeamento da consoante perdida’ e ‘tipo de consoante perdida’ foram os primeiros selecionados nas rodadas estatísticas. Em razão disso e do fato de que tais grupos possuem relação direta entre si (um é uma supercategoria do outro), seus resultados são analisados aqui conjuntamente.

Como ponto de partida, verificamos as taxas de frequência de epêntese vocálica e os pesos relativos atribuídos aos fatores do grupo ‘vozeamento da consoante perdida’. Os resultados para esse grupo podem ser verificados na Tabela 12 abaixo.

Tabela 12 - Resultado para Vozeamento da consoante perdida (PB)

	Apl./Total	%	Peso relativo
desvozeado (<i>aftas</i>)	76/240	31	0,40
vozeado (<i>absorveu</i>)	169/375	45	0,56
TOTAL	245/615	39	

Input: 0,38

Significância: 0,014

Os resultados percentuais da Tabela 12 mostram que a epêntese é mais frequente em dados que apresentam consoante perdida marcada pelo traço [+vozeado]. Esse resultado confirma, pois, a hipótese inicial de que a inserção vocálica ocorreria com maior frequência após consoantes vozeadas. Os resultados encontrados no trabalho de KINZEL (2000), que investigou justamente o papel da sonoridade em contextos onde a epêntese pode ser aplicada em PB, já apontava para essa direção, i.é, para o fato de que contextos vozeados (como em

pigmeu) favorecem a epêntese. Nossa investigação reforça, portanto, a tese de que consoantes perdidas sonoras, por compartilharem com as vogais a marcação positiva do traço [vozeado], estão diretamente relacionadas ao fenômeno de inserção vocálica.

Pode-se interpretar, por outro viés, que o fato de as consoantes perdidas desvozeadas apresentarem peso relativo desfavorecedor (0,40) tem prevalência sobre a posição de favorecimento de ‘vozeado’, que é representada na Tabela 12 por um peso relativo próximo do ponto neutro: 0,56. Essa parece, pois, ser uma constatação igualmente correta a se fazer, ainda que não dispusemos de mais informações para esclarecer essa questão. A análise dos resultados para os tipos específicos de consoantes perdidas pode ser útil nesse sentido.

Antes de apresentarmos os resultados de aplicação de epêntese por tipo de consoante perdida, deve-se mencionar que, no que diz respeito à consoante perdida /d/, seu registro fonético na transcrição e na codificação dos dados foi feito de duas formas, de modo que se permitisse distinguir os casos de realização dessa obstruinte como [d] das situações em que foi produzida como [dʒ]. Ainda que a africada [dʒ] seja considerada um alofone de /d/, que ocorre quando este fonema está diante de [i] e variantes (cf. SILVA, 2002:132), o fato observado nos dados coletados é que /d/ também é realizado como africada mesmo sem que efetivamente esteja diante de uma vogal, como atestam os exemplos [ɪdʒ'nasjəs] *hidnáceas* e ['vɔdʒkə] *vodka*.

É importante observar que o gatilho da palatalização de /d/ em [ɪdʒ'nasjəs] e ['vɔdʒkə] é justamente a vogal epentética, que não é realizada. Numa abordagem teórica como a Fonologia Lexical, que admite a existência de regras e processos no léxico, é possível justificar a ocorrência de tais produções assumindo-se que a estrutura subespecificada para a epêntese na subjacência motiva é suficiente para motivar a realização de [dʒ] na superfície, ainda que a posição vazia para a epêntese permaneça sem ser preenchida. Dado o número razoável de casos (19 no total) em que /d/ foi produzido como [dʒ] sem que a vogal epentética fosse emitida, considerou-se [d] e [dʒ] como duas consoantes perdidas distintas para efeitos da análise feita.

Isso posto, apresentamos, na Tabela 13 (cf. próxima página), os resultados para a variável ‘tipo de consoante perdida’, que foram listados por ordem decrescente de peso relativo por fator.

Tabela 13 - Resultado para Tipo de consoante perdida (PB)

	Apl./Total	%	Peso relativo
[g] (<i>magbás</i>)	38/61	62	0,82
[dʒ] (<i>vodca</i>)	34/53	64	0,77
[k] (<i>actinities</i>)	50/126	39	0,65
[m] (<i>amnésias</i>)	8/16	50	0,45
[b] (<i>obfirmam</i>)	82/192	42	0,42
/t/ (<i>bodisatva</i>)	13/48	27	0,40
[p] (<i>opção</i>)	12/49	24	0,30
[d] (<i>hidnáceas</i>)	7/54	12	0,21
[f] (<i>afta</i>)	1/16	6	0,04
TOTAL	245/615	39	

Input: 0,33

Significância: 0,023

A primeira observação a ser feita é que, como esperado, o peso relativo para [dʒ] foi bem mais elevado do que para [d] (0,77 contra 0,21), isso porque a realização de [dʒ] para o fonema /d/ é, por regra, motivada justamente pela presença de uma vogal após essa consoante (em nosso estudo, uma vogal epentética). Note-se, no entanto, que a superficialização da vogal epentética após [dʒ] está um tanto longe de ser categórica: a vogal é realizada em 65% dos casos.

Outra informação que deve ser observada é que as consoantes perdidas desvozeadas - [k,t,p,f] - apresentam as menores taxas percentuais de aplicação de epêntese (somente a vozeada [d] aparece com taxa reduzida, isso porque contemplou casos pouco frequentes de epêntese após /d/ sem palatalização, como *a[də]ligação* e *a[dɪ]digital*).⁵⁴ Observe-se ainda que, após a fricativa [f], a epêntese ocorreu apenas uma vez, resultando em um peso relativo quase insignificante para essa consoante (0,04). Essas informações tornam mais claros os resultados da Tabela 12, referentes à variável ‘vozeamento da consoante perdida’, pois

⁵⁴ Uma ressalva deve ser feita com relação à consoante perdida [k]. Embora o percentual de ocorrência de epêntese após essa consoante seja menor do que para [m] e [b] (39% contra 50% e 42%, respectivamente), o peso relativo para ela acabou sendo maior do que para essas duas labiais (0,65 contra 0,45 e 0,42). No nível 1 da rodada, [k] de fato apresentou peso inferior a [m] e [b] (0,51 contra 0,61 e 0,54); essa relação se inverteu, no entanto, com a seleção da variável ‘tipo de consoante perdida’ a partir do nível 4 (a seleção dessa variável também provocou alteração no peso relativo de fatores de outros grupos). Através dos resultados para ‘contexto morfológico’ (cf. seção 3.3.2.4), é possível entender que o peso de [k] na Tabela 13 extrapola sua real significância, já que os contextos morfológicos em que aparece (ou dentro de uma base, ou no prefixo *ec-*) têm pesos relativos mais próximos do ponto neutro.

indicam que todas as consoantes perdidas desvozeadas, em especial [f], tendem a desfavorecer a epêntese.

A oclusiva velar [g] foi a consoante que mais favoreceu a epêntese. A inserção vocálica ocorreu em 62% das vezes em que ela foi realizada, o que resultou num peso relativo igual a 0,82. De acordo com o que mostra o Quadro 1, são quatro os vocábulos analisados com esse contexto: *magbás*, *pigmeus*, *estagflação* e *UFRGS*. Através da análise individual da ocorrência de epêntese nesses vocábulos, foram constatadas as seguintes informações: *magbás* - aplicação em 8 de 16 casos; *pigmeus* - aplicação em 13 de 16 casos; *estagflação* - aplicação em 9 de 16 casos; e *UFRGS* - aplicação em 9 de 16 casos. O percentual de epêntese em cada vocábulo é de, no mínimo, 50%; é no vocábulo *pigmeus*, no entanto, que se registra a maior taxa: 81,25 %. Como veremos nas próximas seções, esse vocábulo contém quase todos os fatores linguísticos considerados favorecedores da epêntese em PB (quais sejam, consoante perdida vozeada velar, contexto seguinte vozeado nasal e consoante perdida na posição pretônica).

Voltando a atenção para a coluna com as taxas percentuais da Tabela 13, constata-se que a frequência de epêntese é superior à média geral (39%) não só após [g] e [dʒ], mas também após [m] (50%) e [b] (42%). A consoante perdida [m] ocorre apenas em um vocábulo, *amnésia* (cf. Quadro 1), enquanto [b] aparece em 12 vocábulos, dos quais 9 (75%) são prefixados.

O vocábulo *amnésia*, o único com consoante perdida [m], apresenta, assim como *pigmeu*, contexto seguinte nasal (vozeado) e consoante perdida em posição pretônica. Como veremos nas próximas seções, essas características tendem a favorecer a aplicação de epêntese medial. Os vocábulos com consoante perdida [b], por sua vez, além de conterem em sua maior parte essa consoante na pauta pretônica, são majoritariamente formados por prefixos (*ab-*, *ob-*, *sob-* e *sub-*) adjungidos a bases (*sorveu*, *firmam*, *põem*, *chefe* etc). Veremos na seção 3.3.2.4 que alguns desses prefixos estão intimamente relacionados à epêntese, o que ajuda a explicar, neste momento, a frequência elevada de inserção de vogal após [b].

No intuito de estabelecer um paralelo com a análise do comportamento das consoantes perdidas nos dados de COLLISCHONN (2002), realizamos a amalgamação das consoantes perdidas em três supercategorias: *labial*, *alveolar* e *velar*. Para isso, no entanto, eliminamos da análise a fricativa [f], já que seu peso relativo muito baixo criava nas rodadas-teste uma

aparente distorção dos pesos em outros fatores. O resultado desse amálgama é observado na Tabela 14.

Tabela 14 - Resultado de amálgama para Tipo de consoante perdida (PB)

	Apl./Total	%	Peso relativo
velar [k,g] (<i>éctipo, magbás</i>)	88/187	47	0,72
alveolar /t/, [d,d ₃] (<i>ritmo, vodca</i>)	54/155	34	0,42
labial [p,b,m] (<i>opção, absinto, amnésias</i>)	102/257	39	0,37
TOTAL	244/615	39	

Input: 0,35

Significância: 0,004

Ainda que a tabela acima não apresente um resultado estatisticamente ideal (dada a inversão de ordem que existe entre os percentuais e os pesos relativos das categorias *alveolar* e *labial*), ela serve principalmente para mostrar a relação que existe entre as velares e a inserção vocálica: com peso relativo igual a 0,72, [k] e [g] parecem favorecer esse fenômeno, ainda mais quando seu peso é contrastado com os das demais categorias (0,42 para alveolares e 0,37 para labiais). Esse resultado vai de encontro aos resultados de COLLISCHONN (2002) e à hipótese por nós sustentada, que previa o favorecimento de epêntese após /t/ e /d/.

Quanto a essas consoantes alveolares, observe-se que, na Tabela 13, /d/ de fato apresenta, através do alofone [d₃], peso relativo elevado (0,77). Os pesos relativos de /t/ e de [d], no entanto, são baixos de acordo com essa tabela, e é isso que provoca a diminuição do peso relativo para a supercategoria *alveolares* na amalgamação dos dados apresentada na Tabela 14.

Essas constatações sobre /t/ e /d/, entretanto, não desvalorizam o peso relativo elevado atribuído às velares em nosso estudo. Através da análise dos quatro vocábulos do Quadro 1 que apresentam consoante perdida [g] (a velar que de fato mais favoreceu a epêntese), percebe-se que três deles (75%) possuem contexto seguinte labial, [b], [m] ou [f], que como veremos na próxima seção, correspondem a contextos em que a epêntese ocorre acima da sua frequência média. Admite-se, aqui, que existam motivações de ordem articulatória envolvidas na inserção de uma vogal epentética entre uma velar como [g] e consoantes labiais. Parece-

nos, pois, que distância existente entre o ponto em que uma velar como [g] é produzida no trato oral (parte posterior) e o ponto extremo oposto em que uma labial é produzida (i.é, os lábios) pode causar certa dificuldade ao falante na articulação desses sons em sequência, o que pode resultar na inserção de um elemento vocálico entre essas consoantes, como ocorre em *mag[ə]bás* e *estag[i]flação*.

3.3.2.2 Vozeamento do contexto seguinte e Tipo de contexto seguinte

A variável ‘vozeamento da consoante perdida’ foi a segunda selecionada na rodada 2, ao passo que a variável ‘tipo de contexto seguinte’ foi a terceira escolhida na rodada 1. O resultado para a primeira dessas variáveis é verificado na Tabela 15.

Tabela 15 - Resultado para Vozeamento do contexto seguinte (PB)

	Apl./Total	%	Peso relativo
vozeado (<i>lambda</i>)	153/335	45	0,58
desvozeado (<i>optou</i>)	92/280	32	0,40
TOTAL	245/615	39	

Input: 0,38

Significância: 0,014

Os valores da tabela acima mostram que a epêntese medial é mais frequente em vocábulos com contexto seguinte vozeado. Nossa hipótese que de a marcação positiva do traço [vozeado] para a consoante subsequente à vogal epentética favoreceria esse fenômeno é, portanto, confirmada. Mais uma vez, as constatações feitas por KINZEL (2000) são ratificadas, o que permite concluir que *clusters* consonantais favorecem, de fato, a inserção de um elemento epentético.

Passemos agora à análise dos resultados obtidos para a variável ‘tipo de contexto seguinte’ em PB. Esses resultados estão apresentados na Tabela 16 (cf. próxima página).

Tabela 16 - Resultado para Tipo de contexto seguinte (PB)⁵⁵

	Apl./Total	%	Peso relativo
[l] (<i>adli</i> gação)	7/14	50	0,87
[m] (<i>pigm</i> eus) ou [n] (<i>amn</i> ésias)	43/80	53	0,80
[t] (<i>opt</i> ou) ou [d] (<i>ecd</i> êmicas)	30/81	37	0,76
[f] (<i>obf</i> irmam) ou [v] (<i>obv</i> ersão)	25/61	41	0,58
[k] (<i>vodc</i> a) ou [g] (<i>sobg</i> raues)	24/48	50	0,56
[p] (<i>subp</i> õem) ou [b] (<i>magb</i> ás)	30/63	47	0,43
[ʃ] (<i>subch</i> efe) ou [ʒ] (<i>obj</i> eto)	25/64	39	0,41
[s] (<i>opç</i> ão) ou [z] (<i>quartz</i> o)	25/97	25	0,27
[tʃ] (<i>actin</i> ites) ou [dʒ] (<i>écd</i> ico)	18/75	24	0,23
[x], [h] ou [ʁ] (<i>ad-r</i> ogar)	18/32	56	0,20
TOTAL	245/615	39	

Input: 0,33

Significância: 0,023

Com base nos pesos relativos e nos percentuais de aplicação de epêntese apresentados na tabela acima, propomos a divisão da análise dos resultados em três grupos: o primeiro é formado por [x,h,ʁ], [l] e [m,n], que têm os pesos relativos ou percentuais de aplicação de epêntese mais elevados; o segundo contempla [t,d], [f,v], [k,g], [p,b] e [ʃ,ʒ], que têm pesos e percentuais médios; e o terceiro inclui [s,z] e [tʃ,dʒ], que têm peso relativo e percentual de aplicação baixos. Passemos à análise do primeiro grupo.

A taxa percentual mais alta de aplicação de epêntese foi registrada para os casos em que a vogal inserida estava diante de *r* produzido como fricativa (velar ou glotal) ou como vibrante alveolar (56%). De acordo com a lista de vocábulos analisados (cf. Quadro 1), essa

⁵⁵ Conforme já observado, a variável ‘tipo de contexto seguinte’ gerou alterações nos pesos relativos de alguns fatores dos grupos analisados na rodada 1, e essas alterações ocorreram também para fatores da própria variável (cf. [t,d], [f,v] e, principalmente, [x,h,ʁ] nesta Tabela). No intuito de encontrar uma resposta para isso, foram realizadas algumas rodadas-teste com os fatores da rodada 1. Em cada rodada, excluiu-se uma das variáveis potencialmente conflitantes com ‘tipo de contexto seguinte’. A conclusão a que se chegou com isso, no entanto, foi de que a rodada 1 original era a que apresentava os pesos mais fidedignos em relação às taxas de frequência para cada tipo de consoante seguinte após a epêntese. São estes os pesos que estão representados nesta Tabela.

consoante só aparece em *sub-raças* e em *ad-rogar* como contexto seguinte. Para *ad-rogar*, a epêntese foi registrada em 7 vezes num total de 16 ocorrências (quase 50%). Note-se que tal palavra tem /d/ como consoante perdida, consoante essa que tende a favorecer a inserção da vogal (cf. [d₃] na Tabela 13). Foi no vocábulo *sub-raças*, no entanto, que ocorreu maior frequência de epêntese: em 16 ocorrências desse vocábulo, a inserção vocálica foi registrada 11 vezes (quase 70%). Como veremos na seção 3.3.2.4, o fato de *sub-raças* conter o prefixo *sub-* parece atribuir a essa palavra uma probabilidade maior de ser produzida com epêntese. De momento, cabe observar que o fato de *r* ser produzido como fricativa ou vibrante (e não como tepe alveolar) após *ad-* e *sub-* nos dois vocábulos referidos parece mostrar que existe uma fronteira bem definida entre esses afixos e as bases a que se adjungem.

A lateral [l] foi a consoante seguinte que recebeu o maior peso relativo (0,87) de acordo com os resultados apresentados na Tabela 16. Esse peso condiz com a frequência elevada de epêntese registrada antes dessa consoante (50%). Vale observar que [l] aparece como contexto seguinte apenas na palavra *adligação*, que é formada por um prefixo que tem /d/ como consoante perdida, consoante que, como já sabemos, tende a favorecer a epêntese. A essa informação, soma-se o fato de que, de acordo com a análise das transcrições fonéticas das ocorrências desse vocábulo, os informantes buscam evitar a produção de *dl* não só com epêntese, mas também com metátese (*'adiglação'*) ou com pausa (*'ad#ligação'*). Isso parece indicar que os falantes têm consciência de que *adligação* não pode ser produzido como (*a.dli.ga.ção*), i.é, eles sabem que [dl] não ocorre em *onset* silábico no PB.

As consoantes nasais [m,n] são as últimas do grupo com pesos relativos e percentuais de aplicação de epêntese elevados. Elas aparecem como contexto seguinte em 80 ocorrências, e a epêntese ocorre em 43 dessas ocorrências (53%). Foram analisados, no total, cinco vocábulos com nasal no contexto seguinte: *apneias*, *ritmo*, *hidnáceas*, *pigmeus* e *amnésia*. Como se observa, em todos esses vocábulos, a consoante nasal é antecedida ou por uma oclusiva, ou por outra consoante nasal, i.é, a sonoridade da nasal é muito próxima ou idêntica à sonoridade da consoante que a antecede. De acordo com o que vimos sobre a sílaba do PB (cf. 1.3.2), sequências consonantais de sonoridade vizinha ou idêntica são rejeitadas na língua em respeito ao Princípio de Sonoridade Sequencial. A atuação desse princípio parece ser confirmada nesses dados, já que a epêntese se aplica nesses contextos com frequência maior do que a média para desfazer tais sequências.

No que diz respeito ao segundo grupo por nós proposto para a análise da Tabela 16, entendemos que [t,d], [f,v], [k,g], [p,b] e [ʃ,ʒ] correspondem a contextos seguintes aos quais

são atribuídos pesos relativos próximos ao ponto neutro.⁵⁶ Isso significa, pois, que não há como trazer aqui evidências sobre do papel que esses contextos têm na realização ou não da epêntese. Em razão disso, preferimos nos abster de fazer comentários mais aprofundados sobre esse grupo.

O terceiro e último grupo de contextos seguintes, formado pelas consoantes [s,z] e [tʃ,dʒ], apresenta os pesos relativos e percentuais de aplicação de epêntese mais reduzidos. Para as consoantes [s] e [z], que foram antecedidas por vogal epentética em 25% das vezes em que ocorreram, o peso relativo ficou em 0,27, ao passo que, para [tʃ] e [dʒ], antecedidas por epêntese em 24% das vezes, o peso relativo atribuído foi de 0,23.

As sibilantes [s] e [z] aparecem como contexto seguinte em seis dos vocábulos que analisamos: *opção*, *quartzo*, *eczemas*, *absinto*, *absorveu* e *UFRGS*. Note-se que, em dois desses vocábulos (*eczemas* e *UFRGS*), a consoante perdida é uma velar (consoante que, como foi visto, tende a favorecer a epêntese). São justamente esses dois vocábulos os que mais têm epêntese no grupo das seis palavras com contexto seguinte [s] ou [z]. Das 25 ocorrências de epêntese registradas para esse grupo, 16 delas (64%) são encontradas em *eczemas* e *UFRGS* (oito em cada vocábulo). As nove ocorrências restantes aparecem nos encontros [ps], [tz~ts], [bs] dos vocábulos *opção*, *quartzo*, *absinto* e *absorveu*, que, juntos, somam 64 dados. Nesses encontros, portanto, a taxa de epêntese é bem reduzida ($9/64 = 14\%$).

Os resultados do estudo de COLLISCHONN (2002) mostraram que a epêntese era pouco frequente após sibilantes [s,z]. A partir desses resultados, hipotetizamos para nosso estudo que as taxas de aplicação de epêntese antes de sibilantes seriam igualmente reduzidas, e isso acabou se confirmando. Podemos, nesse sentido, acatar a proposição feita por COLLISCHONN (2002) de que as fricativas sibilantes tendem a formar com a oclusiva precedente uma espécie de africada, como em [k^war.tʃu] *quartzo*, com a consoante perdida posicionada no ataque da sílaba (situação em que a epêntese não se aplica). Em razão dos resultados específicos de nosso estudo, que mostram taxas de aplicação de epêntese bem menores entre oclusivas labiais ou alveolares e sibilantes (*opção*, *quartzo*) do que entre oclusivas velares e sibilantes (*eczemas*), pode-se acrescentar que a oclusiva que forma com a sibilante uma africada é preferencialmente uma consoante labial ou alveolar.

⁵⁶ Excetua-se se aí o contexto [t,d], ao qual foi atribuído peso relativo igual a 0,76, valor não condizente com a taxa percentual de epêntese encontrada antes dessas consoantes (37%). Optou-se por entender, nesse caso, que o peso relativo real para esse contexto está mais próximo do ponto de referência (0,50).

Voltando à análise dos resultados da Tabela 16, observa-se que o contexto seguinte que menos está relacionado com o fenômeno de epêntese é formado por [tʃ] ou [dʒ]. Em 75 ocorrências de vocábulos com esse contexto, a inserção da vogal foi registrada apenas 18 vezes (24%). A esse percentual, foi atribuído um peso relativo igual a 0,23.

De acordo com às transcrições fonéticas dos dados coletados, [tʃ] ou [dʒ] foram produzidos como contexto seguinte em cinco vocábulos: *abdicou*, *actinites*, *éctipo*, *écdico* e *ad-digital*. A frequência de aplicação de epêntese nos casos em que [tʃ] e [dʒ] foram produzidos (e não seus alofones [t] e [d]) é a seguinte: *abdicou* - 4/16 (25%); *actinites* - 2/15 (13%), *éctipo* - 0/15 (0%), *écdico* - 8/13 (61%); e *ad-digital* - 4/16 (25%). Há pelo menos um motivo que ajuda a explicar essa aparente irregularidade entre os percentuais de aplicação de epêntese dos vocábulos que compartilham o mesmo tipo de contexto seguinte.

Observe-se que, nesses vocábulos, a taxa de aplicação de epêntese mais elevada foi registrada em *écdico* (61%). Ao comparar as características morfofonológicas desse vocábulo com as de *éctipo* (0% de epêntese), percebe-se que uma das poucas diferenças existentes é o tipo de vozeamento do contexto seguinte (vozeado em *écdico* e desvozeado em *éctipo*). Como vimos no início desta seção, o traço [+vozeado] do contexto seguinte tende a favorecer a aplicação da epêntese. Isso parece justificar, nos resultados referentes ao contexto seguinte [tʃ,dʒ], o fato de que a frequência de epêntese em *ab[dʒ]icou*, *éc[dʒ]ico* e *ad-[dʒ]igital* é maior do que em *ac[tʃ]inities* e *éc[tʃ]ipo* (que têm contexto seguinte desvozeado).

3.3.2.3 Informante

A variável ‘informante’ foi a segunda selecionada durante o *stepping-up* da rodada 1 e foi a única variável extralinguística selecionada nas duas rodadas. Os resultados obtidos para essa variável estão na Tabela 17 (cf. próxima página).

Os resultados apresentados na Tabela 17 permitem constatar a variação considerável que existe na aplicação da epêntese vocálica entre os falantes nativos de PB inquiridos. Ainda que não tenhamos formulado hipóteses a respeito do comportamento dessa variável, os resultados atestados para ela foram bastante surpreendentes e por isso merecem ser discutidos.

Tabela 17 - Resultado para Informante (PB)

	Apl./Total	%	Peso relativo
N	29/39	74	0,87
H	24/39	61	0,78
K	24/39	61	0,75
I	23/39	58	0,72
C	22/38	57	0,72
O	20/37	54	0,68
A	19/39	48	0,66
F	18/38	47	0,65
D	15/38	39	0,49
B	10/38	26	0,36
P	10/38	26	0,34
E	9/38	23	0,32
L	8/39	20	0,26
G	6/38	15	0,23
M	6/39	15	0,21
J	2/39	5	0,05
TOTAL	245/615	39	

Input: 0,33

Significância: 0,023

Seis informantes produziram epêntese em mais de 50% dos vocábulos analisados, enquanto sete outros produziram epêntese em menos de 30% desses vocábulos. Dentre os que mais produziram epêntese, destaca-se o informante N, que inseriu uma vogal para desfazer os *clusters* analisados em 74% das vezes (13% a mais do que os próximos dois colocados). Dentre os que menos produziram epêntese, destaca-se o informante J, que realizou epêntese em somente dois dos 39 vocábulos produzidos (i.é, 5% das vezes).

Estudos como o de LABOV (1972) apontam para diversos aspectos sociais que podem estar relacionadas à variação de um fenômeno linguístico entre os indivíduos de uma comunidade de fala. Alguns desses aspectos são: a existência de diferentes registros de fala de acordo com o nível socioeconômico, o prestígio atribuído às formas produzidas nesses diferentes registros, o modo como a coleta de dados é feita, o grau de formalidade do contexto de fala, etc.

Para a presente investigação especificamente, baseando-se na metodologia adotada (que envolveu a seleção dos informantes pelo método aleatório estratificado) e nas fichas sociais preenchidas por cada informante (cf. modelo no Anexo VII), as seguintes considerações podem ser feitas:

a) A variação na aplicação de epêntese que ocorre entre cada um dos 16 informantes pode ser atribuída tanto ao fato de que esses informantes podem pertencer a diferentes contextos socioeconômicos (sendo que a epêntese é mais prestigiada em uns e menos prestigiada em outros desses contextos) quanto ao fato de que os informantes podem ter se comportado de maneira diversa diante do fato de que estavam sendo gravados (alguns poderiam se sentir mais inibidos com isso, por exemplo, o que poderia causar a produção mais ‘cuidada’ dos vocábulos);

b) A frequência bastante reduzida para o informante J pode estar relacionada a uma espécie de hipercorreção. Esse informante, que é descendente de japoneses, informou em sua ficha social que seus pais são falantes de japonês e de português. Pode-se cogitar que, como os pais desse informante podem produzir, por influência do japonês, um número considerável de vocábulos com epêntese em PB,⁵⁷ esse informante entende que tais formas epentetizadas na verdade não são tão frequentes e tende a evitar produzi-las em sua fala.

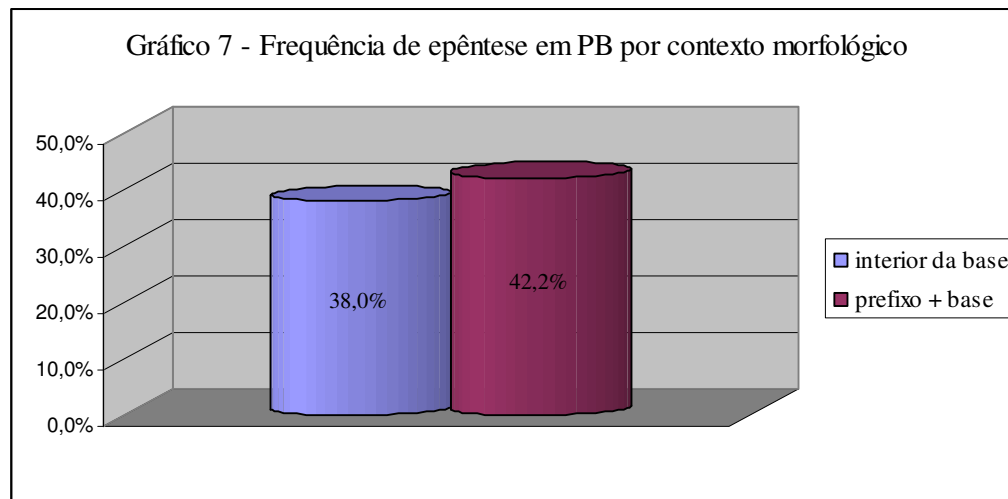
Dado o fato de que a variável ‘informante’ foi também selecionada na análise estatística com os dados de língua inglesa, deixaremos a discussão de outros aspectos sobre a mesma para um momento posterior (cf. seção 3.4.2.3).

3.3.2.4 Contexto morfológico e Tipo de *cluster*

As variáveis ‘contexto morfológico’ e ‘tipo de *cluster*’ foram consideradas tanto na rodada 1 quanto na rodada 2. Na rodada 1, a variável ‘tipo de *cluster*’ foi selecionada um nível antes de ‘contexto morfológico’. Na rodada 2, isso se inverteu: ‘contexto morfológico’ foi selecionado um nível antes de ‘tipo de *cluster*’. Na tentativa de explicar o porquê dessa aparente relação entre as duas variáveis, decidimos analisá-las numa mesma seção. Começaremos por ‘contexto morfológico’.

Apresentamos, inicialmente, os percentuais de aplicação de epêntese verificados no grupo de vocábulos com *cluster* no interior da base e no grupo de vocábulos com *cluster* na fronteira prefixo + base. Esses resultados estão no Gráfico 7 (cf. próxima página).

⁵⁷ De acordo com ITÔ (1986) as sílabas do japonês terminam em vogal ou semivogal (embora possam terminar também em uma nasal ou consoante geminada, como em *nip.pon*, cf. ITÔ, 1986:144).



Em 347 ocorrências de vocábulos com *cluster* no interior da base, a epêntese foi registrada 132 vezes (38%), ao passo que, em 268 dados com *cluster* na fronteira entre prefixo e base, a inserção da vogal ocorreu 113 vezes (42,2%). Ainda que a diferença percentual seja tímida entre esses dois grupos, pode-se notar, através do gráfico acima, que o contexto morfológico entre um prefixo e uma base parece ser mais propício para a realização de epêntese, o que confirma nossa hipótese inicial sobre essa questão (cf. 2.2.2.1.1).

Cabe averiguar, agora, o comportamento individual de cada um dos fatores da variável ‘contexto morfológico’ em PB, que são: *interior da base*, *ab- + base*, *ad- + base*, *ec- + base*, *ob- + base*, *sob- + base* e *sub- + base*. Os resultados para esses fatores podem ser visualizados na Tabela 18.

Tabela 18 - Resultado para Contexto morfológico (PB)

	Apl./ Total	%	Peso relativo Rodada 1	Peso relativo Rodada 2
sub- (<i>subchefe</i>)	23/37	62	0,87	0,68
sob- (<i>sobpõem</i>)	15/27	55	0,83	0,61
ec- (<i>eczemas</i>)	24/63	38	0,39	0,54
interior da base (<i>ritmo</i>)	132/347	38	0,42	0,53
ob- (<i>obversão</i>)	21/48	43	0,58	0,41
ad- (<i>adjunto</i>)	22/61	36	0,39	0,29
ab- (<i>absorveu</i>)	8/32	25	0,71	0,28
TOTAL	245/615	39		

Input: 0,33

Significância: 0,023

Input: 0,38

Significância: 0,014

No que diz respeito aos pesos relativos apresentados na Tabela 18, vamos dar atenção somente aos valores da rodada 2.⁵⁸ Analisando os pesos relativos da rodada 2 em conjunto com os percentuais de frequência de epêntese para cada um dos fatores, é possível classificar os fatores em três grupos: o grupo 1, formado por *sob-* e por *sub-*, que apresentam peso relativo e frequência de epêntese elevados; o grupo 2, formado por *ec-*, *interior da base*, *ob-* e *ad-*, que apresentam valores médios na análise conjunta entre peso relativo e frequência de epêntese; e o grupo 3, formado pelo fator *ab-* apenas, que apresentou peso relativo e frequência de epêntese reduzidos. Considerando-se a neutralidade atestada para os fatores do grupo 2, analisaremos apenas os grupos 1 e 3.⁵⁹

O grupo 1, como foi dito, é formado pelos fatores que apresentaram os pesos relativos mais elevados na rodada 2: *sob-* (peso igual a 0,61) e *sub-* (peso igual a 0,68). Os percentuais de frequência de epêntese nos vocábulos com os prefixos *sob-* e *sub-* foram também os mais elevados (55% e 62%, respectivamente) e se distanciam dos percentuais referentes aos demais contextos morfológicos analisados. Esses resultados confirmam, portanto, nossa hipótese, formulada com base na suspeita levantada por COLLISCHONN (2002), de que vocábulos com *sub-* (e, por extensão, com *sob-*) deveriam apresentar maiores índices de inserção vocálica.

Os resultados encontrados para *sob-* e *sub-* parecem apontar, ainda, para o fato de que estamos tratando, nesses casos, de dois *prefixos adjuntos* (cf. 1.2). A essa afirmação, acrescentamos nossa crença de que os prefixos adjuntos *sob-* e *sub-* possam ser interpretados não apenas como sílabas átonas à esquerda de uma palavra fonológica, mas, também, como sílabas com acento próprio e que correspondem, nesse caso, a palavras fonológicas (que podem inclusive aparecer isoladas, especialmente no caso de *sub-*⁶⁰). Note-se que, sob essa perspectiva, não estaríamos tratando de casos de epêntese medial, mas, sim, de epêntese final (ou seja, a epêntese em *sob*[I]*graves* e em *sub*[I]*chefe* estaria ocorrendo num contexto morfológico semelhante ao da epêntese atestada em *PUC*[I], *BIG*[I] etc.).

⁵⁸ Como já foi observado, na rodada 1, a seleção da variável ‘tipo de contexto seguinte’ causou alterações nos pesos relativos de alguns fatores. Isso se refletiu sobre o grupo ‘contexto morfológico’, especialmente sobre o fator ‘ab- + base’, para o qual foi atribuído peso relativo inverso ao percentual de frequência (0,71 vs. 25%).

⁵⁹ A neutralidade atribuída aos fatores *ec-*, *ob-* e *ad-* permite concluir que os prefixos aos quais esses fatores se referem são considerados, sob o ponto de vista prosódico, sílabas átonas com características semelhantes às sílabas que estão no interior de vocábulos. Prova disso está na semelhança entre a frequência de aplicação de epêntese após esses prefixos e a frequência de aplicação da epêntese no interior da base (cf. Tabela 18).

⁶⁰ SCHWINDT (2000:109) observa, nesse sentido, que o prefixo *sub-* “(...) só aparecerá isolado quando se constatar na sentença um relação de oposição/substituição, do tipo *Se o gerente não está, falarei com o sub*, uma vez que, em princípio, não se compreenderia algo como *O sub está?* (isolado de contexto).”

Resta-nos analisar o grupo 3 proposto para a classificação dos resultados da Tabela 18. Esse grupo é formado apenas pelo fator *ab-*, que apresentou tanto peso relativo quanto percentual de aplicação de epêntese reduzido (0,28 e 25%, respectivamente). Note-se, como ponto de partida, o fato de que o prefixo *ab-* apresenta a mesma consoante perdida que os prefixos *sob-* e *sub-*, o que mostra que o papel da consoante perdida aqui não é suficiente para explicar as diferenças entre esses dois grupos de prefixos.

São dois os vocábulos analisados que contêm o prefixo *ab-*: *absorveu* e *abdicou*. Para cada vocábulo, a epêntese foi registrada exatamente 4 vezes de um total de 16 ocorrências. Contata-se que, nesses vocábulos, os contextos seguintes são, respectivamente, [s] e [d₃]. Como vimos na seção 3.3.2.2, esses são os contextos seguintes que menos estão relacionados com a epêntese vocálica. Parece, pois, que, nesses dados, a informação sobre o contexto seguinte do *cluster* tem prevalência sobre a informação do contexto morfológico em que esse *cluster* se encontra. Sob esse ponto de vista, é possível considerar que o prefixo *ab-* se comporta de modo semelhante aos contextos morfológicos contemplados no grupo 2 de nossa classificação.

Passando-se à análise da variável ‘tipo de *cluster*’, temos como resultados o que está expresso na Tabela 19.

Tabela 19 - Resultado para Tipo de *cluster* (PB)

	Apl./ Total	%	Peso relativo Rodada 1	Peso relativo Rodada 2
existente em PB e em inglês (<i>magbás, rugby</i>)	189/504	37	0,43	0,46
existente só em português (<i>sub-raças</i>)	56/111	50	0,77	0,63
TOTAL	245/615	39		

Input: 0,33 Input: 0,38
Significância: 0,023 Significância: 0,014

A variável ‘tipo de *cluster*’ foi originalmente admitida neste estudo para que se estabelecesse uma comparação entre *clusters* existentes em PB e *clusters* existentes em inglês, e vice-versa (cf. 2.2.2.1.7). Para que isso acontecesse, no entanto, pressupunha-se que a

variável ‘tipo de *cluster*’ deveria ser selecionada não só em PB, mas também em inglês, o que não aconteceu (cf. 3.4.2).

Ainda que essa comparação não possa ser feita, os resultados obtidos para a variável ‘tipo de *cluster*’ em PB, que mostram o favorecimento de epêntese em *clusters* existentes em PB e inexistentes em inglês (cf. pesos relativos de 0,77 e 0,63 da Tabela 19), servem, em parte, para corroborar os resultados obtidos para ‘contexto morfológico’. A partir da visualização da Tabela 9, que contém os oito vocábulos analisados com *clusters* existentes em PB e inexistentes em inglês, nota-se a presença de três vocábulos prefixados: *ad-digital*, *ad-rogar* e *sub-raças*. As informações morfofonológicas desses vocábulos, especialmente as de *ad-rogar* e de *sub-raças* - quais sejam, a presença do prefixo *sub-* em um deles e a presença de uma fricativa/vibrante como contexto seguinte em ambos -, colaboram parcialmente para explicar a seleção da variável ‘tipo de *cluster*’ em PB.⁶¹

3.3.2.5 Acento

A variável ‘acento’ foi submetida a análise estatística tanto na rodada 1 quanto na rodada 2 dos dados do PB. Foi apenas na segunda rodada, no entanto, que ela foi selecionada. Os resultados para essa variável estão expressos na Tabela 20.

Tabela 20 - Resultado para Acento (PB)

	Apl./Total	%	Peso Relativo
pretônico (<i>ob’jeto</i>)	203/468	43	0,53
postônico (<i>’ritmo</i>)	42/147	28	0,39
TOTAL	245/615		

Input: 0,38

Significância: 0,014

Indiferente do fato de ter sido selecionada em apenas uma das duas rodadas em que foi analisada, a variável ‘acento’ parece mostrar, através dos resultados da Tabela 20, papel

⁶¹ Outras razões que podem ter motivado a seleção dessa variável são passíveis de ser inferidas a partir das informações linguísticas dos oito vocábulos que representam o fator favorecedor (cf. Tabela 9). São algumas delas: a presença de velar como consoante perdida em quatro vocábulos; a presença de consoante perdida vozeada em cinco vocábulos; e a presença de consoante perdida na pauta acentual pretônica em seis vocábulos.

importante na aplicação do fenômeno de epêntese em PB. Os resultados dessa tabela apontam, pois, para uma tendência que parece existir no sentido de se evitar a realização de epêntese em vocábulos nos quais a vogal epentética é produzida na pauta postônica (cf. peso relativo 0,39), já que, nesses casos, a epêntese implica o deslocamento do acento principal da palavra (na direção direita → esquerda), como ocorre no exemplo ?[ɛ.tʃi.ni.ku] *ét.ni.co*.

Fala-se aqui em uma provável tendência porque a variável ‘acento’ foi a primeira a ser selecionada nas rodadas estatísticas do estudo de COLLISCHONN (2002) e os resultados para essa variável, nesse estudo, igualmente apontaram para o papel desfavorável que o contexto prosódico ‘postônico’ representa para a aplicação da epêntese (peso relativo 0,24 contra peso relativo 0,68 para contexto pretônico).

Dadas as semelhanças existentes entre os resultados de COLLISCHONN (2002) (cf. 1.3.4.1.1) e os de nossa investigação para a variável ‘acento’, resta-nos reforçar aqui a conclusão a que esse estudo anterior ao nosso já havia chegado:

“(...) a explicação para o comportamento da epêntese em relação ao acento baseia-se nas restrições de acento em português, que claramente favorecem a colocação do acento na penúltima sílaba da palavra, sendo evitada qualquer inserção de sílaba à direita, para impedir o deslocamento de acento.”

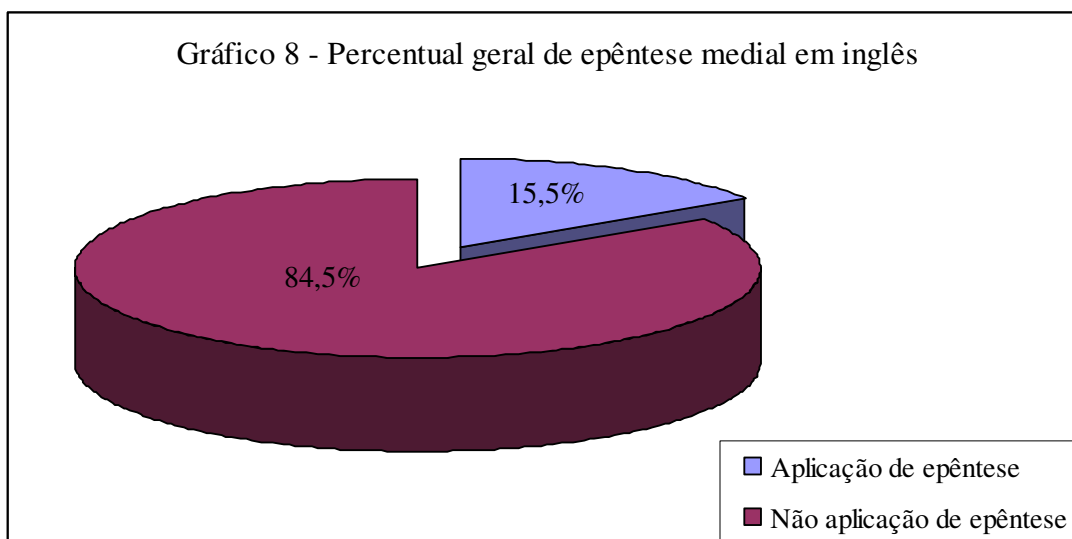
COLLISCHONN (2002:228)

3.4 Amostra de Fala 2 (inglês)

Na presente seção, apresentamos e discutimos os resultados obtidos a partir da coleta de dados que compõem a Amostra de Fala 2, através da qual realizamos a análise da epêntese vocálica que ocorre no inglês aprendido por brasileiros. Subdividimos a análise desses resultados da seguinte forma: em 3.4.1, discutimos aspectos relacionados ao percentual geral da aplicação de epêntese nos dados dessa Amostra; em 3.4.2 e subitens, apresentamos e discutimos os resultados obtidos para as rodadas estatísticas realizadas com os dados de língua inglesa.

3.4.1 Percentual geral de epêntese medial em inglês

O percentual geral de aplicação de epêntese medial que apresentamos para os dados do inglês foi calculado com base no número total de ocorrências para os 61 vocábulos analisados e no número de vezes em que a epêntese foi empregada nesse total de ocorrências. Foram incluídos nesse cálculo, portanto, as ocorrências em que os falantes buscaram outras formas de evitar a realização do *cluster* consonantal (como apagamento da consoante perdida ou apagamento do contexto seguinte). O percentual geral de epêntese para os dados do inglês podem ser conferido no Gráfico 8 abaixo.

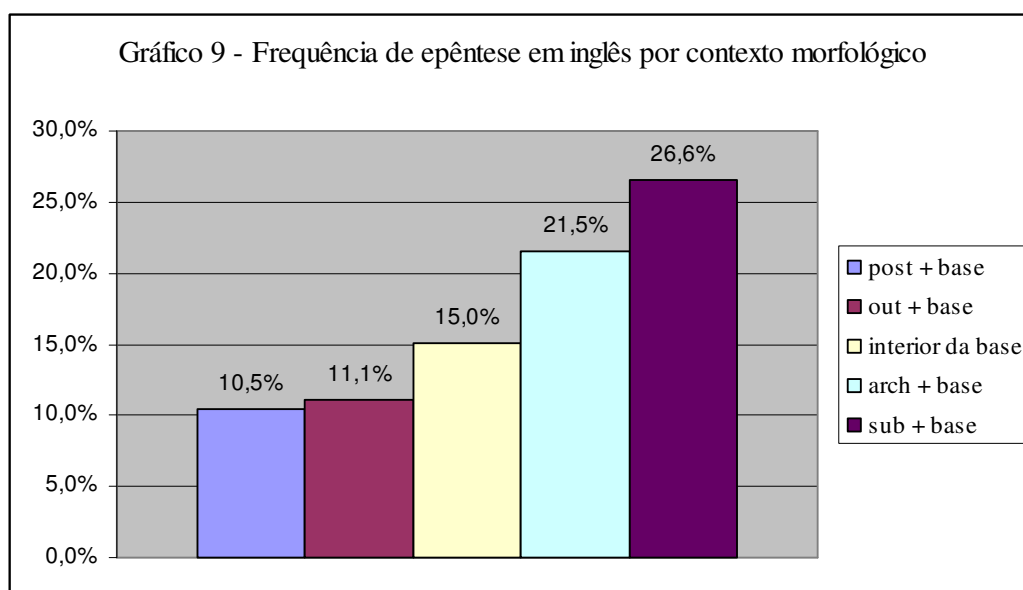


Das 976 ocorrências dos vocábulos analisados na Amostra de Fala 2, apenas 151 apresentaram vogal epentética: 15,5%, cf. Gráfico 8. Esse percentual é significativamente menor (menos da metade) do que o registrado nos vocábulos do PB (39,4 %, cf. Gráfico 6). Esse percentual era de fato esperado, já que se considerava que alguns dos informantes teriam maior consciência sobre o fato de que em inglês as sílabas podem terminar em oclusivas e produziram, portanto, menos epêntese.

A comparação do percentual de 15,5% de aplicação de epêntese com o percentual de aplicação de epêntese obtido por PEREYRON (2008) em sua análise acústica (8%) permite constatar a existência de uma diferença considerável entre os resultados desses dois estudos (ainda que eles se assemelhem sob o ponto de vista metodológico). Parece-nos, pois, que a justificativa para isso se encontra nas diferenças que existem entre os fatores contextuais

envolvidos no desenvolvimento de cada um desses estudos, o que inclui o ambiente em que a coleta dos dados foi feita, as características individuais dos participantes envolvidos, entre outros.

Decidimos ainda apresentar, aqui, o percentual geral de epêntese obtido para a variável ‘contexto morfológico’ em inglês. Como veremos na próxima seção (3.4.2), essa variável não foi selecionada nas rodadas estatísticas para os dados do inglês e, em razão disso, acabou não sendo discutida pormenorizadamente neste estudo. Ainda assim, julgamos interessante apresentar os percentuais de epêntese atribuídos a cada um dos fatores da variável (*interior da base*, *arch-*, *out-*, *post-* e *sub-*). Esses percentuais podem ser visualizados no Gráfico 9 abaixo.



Os resultados do gráfico acima comprovam que, por mais que as rodadas estatísticas realizadas com os dados de inglês não tenham selecionado a variável ‘contexto morfológico’, existem diferenças consideráveis entre as frequências de epêntese dos diferentes contextos morfológicos analisados. Destacam-se aí as frequências registradas para os fatores *arch-* (21,5%) e, principalmente, *sub-* (26,6%), pois ambas estão acima da média geral de frequência de epêntese em inglês (15,5%).

Admite-se, no que concerne ao percentual de epêntese após os prefixos *arch-* e *sub-*, que os falantes brasileiros de inglês possam estar sujeitos a uma espécie de transferência linguística que envolve o nível morfológico da L1. Sob esse ponto de vista, as ocorrências em

PB dos prefixos *arqui-* e *sub-* podem motivar a aplicação de epêntese após *arch-* e *sub-* na interlíngua (no primeiro caso, pelo próprio fato de que *arqui-* já contém a vogal epentética *i*, e, no segundo caso, porque, como já vimos, *sub-* em PB é um prefixo favorecedor de epêntese). Não é nosso objetivo pormenorizar essa discussão aqui, no entanto. Cabe concluir, no que diz respeito aos percentuais gerais apresentados para a variável ‘contexto morfológico’ em inglês (cf. Gráfico 9), que a fronteira existente entre prefixo e base pode ter, sim, um papel relevante na investigação da interlíngua entre o português e o inglês.

3.4.2 Resultados das análises estatísticas para dados do inglês

A análise dos resultados das rodadas estatísticas referentes aos dados de língua inglesa foi antecedida pela exclusão de algumas ocorrências inesperadas na produção oral dos vocábulos coletados nessa língua. Tal exclusão foi realizada porque determinados vocábulos foram produzidos sem o contexto apropriado para a aplicação de epêntese vocálica (i.é, sem a presença de um *cluster* consonantal), gerando-se, em consequência disso, 0% de aplicação de epêntese nesse conjunto.

Excluíram-se da análise, pois, as ocorrências referentes a vocábulos produzidos sem a presença da consoante perdida (como [rʌbɪ] em vez de [rʌgbɪ] *rugby*) ou sem a presença do contexto seguinte (como [næpɪŋ] em vez de [næpkɪn] *napkin*). Também foram excluídos ‘outros casos’, que incluíram, por exemplo, ocorrências em que o informante, em vez de produzir um determinado vocábulo, acabou produzindo outro (como, por exemplo, *opinions* em vez de *options*). Do total de 54 exclusões realizadas, a grande maioria (40) se referia a casos de consoantes perdidas omitidas. Oito das exclusões se referiam a contextos seguintes omitidos e as seis exclusões restantes se referiam a ‘outros casos’.

Note-se que os casos de omissão de consoante perdida e de omissão de contexto seguinte juntos contemplam 48 ocorrências, valor que representa 4,9% do total de ocorrências analisadas (976). Esse percentual, ainda que seja pequeno, parece indicar que, na fala dos aprendizes brasileiros de inglês, não só a epêntese como também o processo de apagamento é utilizado como estratégia de reparo a sílabas malformadas.

As rodadas estatísticas com os dados codificados foram ainda precedidas por algumas rodadas-teste, nas quais foram detectados quatro grupos com *knockout*: (i) 0% de aplicação de epêntese em oito ocorrências com consoante perdida [θ]; (ii) 0% de aplicação de epêntese em 11 ocorrências com contexto seguinte [θ,ð]; (iii) ocorrência única de [ð] como consoante

perdida; e (iv) ocorrência única de [ʒ] como consoante perdida. As 21 ocorrências que geravam *knockout* foram recodificadas atribuindo-se o código da consoante que mais aparecia em substituição à consoante em situação de *knockout* (por exemplo, [ð] em *rhythmic*, foi recodificada como [t], que foi a consoante mais recorrente na pronúncia do *th* nesse vocábulo).

Realizados esse procedimentos, procedeu-se à execução da rodada 1 e da rodada 2 (cf. Quadro 4) pelo programa computacional GoldVarb X. Os resultados para essas rodadas podem ser verificados no Quadro 8.

Quadro 8 - Grupos analisados e grupos selecionados por rodada (inglês)

	Rodada	
	1	2
Grupos analisados	Contexto morfológico Tipo de consoante perdida Tipo de contexto seguinte Acento Tipo de <i>cluster</i> Informante	Contexto morfológico Vozeamento da consoante perdida Vozeamento do contexto seguinte Acento Tipo de <i>cluster</i> Sexo Idade Proficiência em inglês
Grupos selecionados	1º Tipo de consoante perdida 2º Acento 3º Informante	1º Vozeamento da consoante perdida 2º Acento

Como podemos ver no quadro acima, foram quatro os grupos selecionados nas rodadas com dados da língua inglesa: ‘tipo de consoante perdida’, ‘acento’ (selecionada tanto na primeira quanto na segunda rodada), ‘vozeamento da consoante perdida’ e ‘informante’.

À semelhança do procedimento adotado na análise dos resultados da Amostra de Fala 1, analisaremos aqui apenas os grupos que foram selecionados pelas rodadas estatísticas (excluem-se da análise, portanto, ‘contexto morfológico’, ‘tipo de contexto seguinte’, ‘tipo de *cluster*’, ‘vozeamento do contexto seguinte’, ‘sexo’, ‘idade’ e ‘proficiência em inglês’). Propomos, ainda, a subdivisão da análise de acordo com o seguinte: em 3.4.2.1, analisamos os resultados para ‘vozeamento da consoante perdida’ e ‘tipo de consoante perdida’ conjuntamente; em 3.4.2.2, tratamos dos resultados para ‘acento’; e em 3.4.2.3, finalmente, discutimos os resultados para o grupo ‘informante’.

3.4.2.1 Vozeamento da consoante perdida e Tipo de consoante perdida

O primeiro grupo de fatores a ser selecionado na rodada 1 foi ‘tipo de consoante perdida’ e o primeiro grupo a ser selecionado na rodada 2 foi ‘vozeamento da consoante perdida’. Dada a relação estreita que existe entre tais grupos, realizamos, nesta seção, a análise conjunta dos mesmos.

Iniciamos pela análise dos resultados para o grupo ‘vozeamento da consoante perdida’. Esses resultados estão expressos na Tabela 21 abaixo.

Tabela 21 - Resultado para Vozeamento da consoante perdida (inglês)

	Apl./Total	%	Peso relativo
vozeado (<i>sabzi</i>)	77/182	29	0,70
desvozeado (<i>outpassed</i>)	74/663	11	0,41
TOTAL	151/922	16	

Input: 0,14

Significância: 0,003

Os resultados apresentados na Tabela 21 mostram que a epêntese medial é registrada em 29% das ocorrências em que a consoante perdida é vozeada, ao passo que ela é registrada em apenas 11% das ocorrências em que a consoante perdida é desvozeada. Na comparação dos percentuais dos fatores ‘vozeado’ e ‘desvozeado’ com a média geral de 16% de epêntese, nota-se que é o fator ‘vozeado’ que está em posição de destaque, já que está bem mais distante da média geral do que o fator ‘desvozeado’.

Em estudos que investigaram quantitativamente a epêntese inicial e final em inglês como L2/LE, como REBELLO (1997), CAGLIARI (2001), CORNELIAN JR. (2003), BETTONI-TECHIO (2005) e PEREYRON (2008), a epêntese foi igualmente atestada com maior frequência em ambiente vozeado. Parece-nos, pois, que a epêntese vocálica definitivamente possui uma relação estreita com esse tipo de ambiente, o que pode ser justificado pelo fato de que a vogal epentética compartilha com a consoante próxima a ela a mesma marcação (+) para o traço [vozeado].

Passemos à análise do grupo ‘tipo de consoante perdida’. Cabe observar, de início, que muitas das produções de fala analisadas para o inglês apresentaram algum tipo de alteração

fonética, isso quando comparadas com as produções efetivamente realizadas por um falante nativo desse idioma. Dentre esses tipos de alterações, inclui-se a substituição do segmento referente à consoante perdida na forma-alvo por um segmento com ponto e modo de articulação semelhante (por exemplo, a substituição de [θ] por [t] em *athbash*). Optamos, nesses casos, por transcrever e codificar os dados de acordo com as produções realizadas pelos falantes.

Isso dito, passemos aos resultados obtidos para o grupo ‘tipo de consoante’. Esses resultados estão dispostos na Tabela 22.

Tabela 22 - Resultado para Tipo de consoante perdida (inglês)

	Apl./Total	%	Peso relativo
[dʒ] (<i>fledglings</i>)	13/20	65	0,94
[g] (<i>rugby</i>)	14/34	41	0,87
[tʃ] (<i>archbishop</i>)	21/70	30	0,74
[b] (<i>sublicense</i>)	36/127	28	0,72
[m] (<i>amnesia</i>)	5/17	29	0,67
[p] (<i>napkin</i>)	9/65	13	0,56
[k] (<i>auction</i>)	17/113	15	0,55
[d] (<i>adze</i>)	7/44	15	0,54
[ʃ] (<i>marshmallows</i>)	6/55	10	0,42
[n] (<i>enmity</i>)	1/16	6	0,35
[t] (<i>portfolio</i>)	20/296	6	0,29
[f] (<i>hafnium</i>)	2/65	3	0,21
TOTAL	151/922	16	

Input: 0,10

Significância: 0,005

A primeira observação a ser feita nos resultados da Tabela 22 é que, à exceção de [d] e [n], os fatores que representam consoantes vozeadas estão concentrados na parte superior dessa tabela, que é onde estão os fatores para os quais é registrada maior frequência de epêntese. Essa observação nos ajuda a entender um pouco mais sobre os resultados da tabela anterior a esta (Tabela 21), na medida em que permite entender que a frequência de epêntese é alta em qualquer ambiente vozeado, indiferente do tipo de consoante vozeada presente nesse ambiente.

As consoantes perdidas [dʒ] e [g] são, de acordo com os percentuais da Tabela 22, aquelas que mais são seguidas de uma vogal epentética. No caso específico de [dʒ],

verificamos nas transcrições fonéticas dos dados coletados que a grande maioria dos registros de epêntese após essa africada foi atestada para o vocábulo *hodgepodge*. Existe uma vogal *e* na forma gráfica desse vocábulo, que, na pronúncia de um falante nativo de inglês, não é produzida, o que justificou sua inclusão em nossa análise. Por influência da forma escrita desse vocábulo, no entanto, os aprendizes brasileiros de inglês acabaram inserindo uma vogal entre o *cluster* [dʒp]. Isso justifica o porquê de a epêntese ter sido aplicada com grande frequência após essa africada em nossos dados.

Em razão do exposto, parece-nos que a consoante perdida que genuinamente é acompanhada mais vezes por epêntese é a velar [g]. Essa constatação vai ao encontro dos resultados apresentados em PEREYRON (2008), que mostraram que [g] é a consoante perdida mais intimamente relacionada à epêntese medial em inglês (peso relativo 0,84 na análise acústica).

Depois do fator [g], aparecem ainda [tʃ], [b] e [m] com percentuais elevados de epêntese (cf. Tabela 22). Com relação aos resultados para [tʃ] e [b], acreditamos que as frequências elevadas de epêntese para esses fatores estejam diretamente relacionadas às frequências (igualmente elevadas) de epêntese encontradas para os vocábulos com os prefixos *arch-* e *sub-* (cf. Gráfico 9). Já em relação aos resultados para [m], fator admitido exclusivamente em razão do vocábulo *amnesia*, parece que a realização do platô de sonoridade que envolve nasal labial + nasal alveolar ([mn]) não é algo bem aceito na interlíngua dos informantes que inquirimos, que preferem inserir uma vogal para desfazer um encontro consonantal de sonoridade idêntica.

No outro extremo da Tabela 22, encontram-se os fatores [n], [t] e [f], para os quais registraram-se percentuais de aplicação de epêntese bem inferiores à média geral. No que diz respeito a [n], vale observar, primeiramente, que tal fator foi registrado apenas nas ocorrências do vocábulo *enmity*. Observe-se que, nesse vocábulo, ocorre um platô de sonoridade com consoantes nasais ([nm]) semelhante ao platô nasal atestado em *amnesia* (para o qual a epêntese se mostrou frequente). O fato de ter havido baixa frequência de epêntese em *enmity* parece mostrar que o platô nasal alveolar + nasal labial parece ser mais aceitável na interlíngua dos participantes. Essa maior aceitabilidade parece ser resultante do fato de que a produção do [n] em coda na interlíngua não parece ser considerada tão ‘difícil’ quanto a produção de [m] em coda para os falantes brasileiros de inglês, que acabam dispensando a produção de epêntese em vocábulos como *enmity*.

Finalmente, no que concerne aos resultados para [t] e [f], queremos acreditar que as duas consoantes perdidas a que esses fatores se referem estejam entre as primeiras consoantes obstruintes a serem admitidas na coda silábica do inglês aprendido por brasileiros (o que explicaria a dispensa do uso de epêntese nesses casos). Essa nossa crença parece ser confirmada através dos resultados apresentados em PEREYRON (2008), por exemplo. Nesse estudo, houve o registro de apenas uma aplicação de epêntese em 109 ocorrências com a consoante perdida [t] (PEREYRON, 2008:117). No caso de [f], não houve registro algum de epêntese no total de 65 ocorrências de vocábulos com essa consoante perdida (o que inclusive motivou a exclusão dessas ocorrências nas rodadas estatísticas, cf. PEREYRON, 2008:116).

3.4.2.2 Acento

O grupo de fatores ‘acento’ foi incluído nas duas rodadas estatísticas realizadas com os dados de língua inglesa. Tanto na rodada 1 quanto na rodada 2, esse grupo foi o segundo a ser selecionado como estatisticamente relevante.

Os resultados obtidos para ‘acento’ estão expressos na Tabela 23 abaixo.

Tabela 23 - Resultado para Acento (inglês)

	Apl./ Total	%	Peso relativo Rodada 1	Peso relativo Rodada 2
pretônico (<i>am'nesia</i>)	99/495	20	0,59	0,56
postônico (<i>'rhabdom</i>)	52/427	12	0,39	0,42
TOTAL	151/922	16		

Input: 0,10 Input: 0,14
Significância: 0,005 Significância: 0,003

Através dos resultados da tabela acima, é possível perceber que a epêntese registrada na pauta postônica dos vocábulos analisados teve menor frequência (12% de aplicação) do que a epêntese que ocorreu na pauta pretônica desses vocábulos (20% de aplicação).

A hipótese que fizemos a respeito do comportamento de ‘acento’ foi a de que não deveria haver frequência reduzida de aplicação de epêntese no contexto ‘postônico’ em dados do inglês, uma vez que pressupúnhamos que os falantes da interlíngua teriam consciência de que o deslocamento do acento para depois da penúltima sílaba (considerando-se a

direcionalidade direita → esquerda) é algo admissível em inglês, como comprova a formação da palavra *comfortable* a partir de *comfort*: [kʌm-fərt] *ˈcomfort* → [kʌm-fər-tə-bl] *ˈcomfortable*. Nossa pressuposição, no entanto, não foi confirmada nos resultados obtidos.

Cabe antecipar aqui parte da comparação que propomos fazer entre os resultados da Amostra de Fala 1 e os da Amostra de Fala 2 (cf. 3.5). Na comparação dos resultados obtidos para ‘acento’ em inglês (cf. Tabela 23) com os resultados referentes à variável ‘acento’ em PB (cf. Tabela 20), podemos constatar que o comportamento atestado para a epêntese no contexto postônico na interlíngua é semelhante ao comportamento atestado para a epêntese na L1 (cf. pesos relativos 0,39 e 0,42 em inglês e 0,39 em PB para o fator ‘postônico’ nas tabelas referidas). Isso parece indicar, pois, que o padrão acentual do PB exerce algum tipo de influência na interlíngua.

Acreditamos que a semelhança existente entre os resultados para ‘postônico’ nas duas línguas investigadas indique que a transferência linguística que admitimos operar da L1 para a interlíngua (cf. 1.1) inclui a transposição de informações sobre a pauta acentual da língua materna para a língua aprendida. Isso parece que se confirma através análise das transcrições fonéticas dos dados produzidos pelos falantes.

Nessas transcrições, percebe-se um número razoável de ocorrências em que houve alteração da posição do acento principal da palavra. Alguns exemplos de alteração de acento atestados são: *arithˈmetic* (produzido em vez de *aˈrithmetic*); *aˈthletes* (produzido em vez de *ˈathletes*) e *enˈmity* (produzido em vez de *ˈenmity*).

Pode-se concluir, a partir das constatações que fizemos, que o grau de relevância atribuído a ‘acento’ nos dados do inglês (considerando-se que esse grupo foi selecionado em todas as rodadas em que foi analisado) é reflexo do papel que a variável ‘acento’ desempenha nos dados do PB. Admite-se, nesse sentido, que a transferência linguística que ocorre da L1 para a L2 opere sobre a pauta acentual dessas duas línguas.

3.4.2.3 Informante

O grupo de fatores ‘informante’ foi o terceiro e último selecionado na rodada 1 dos dados de língua inglesa. Os resultados obtidos para esse grupo são apresentados na Tabela 24.

Note-se que o ordenamento dos fatores nessa tabela obedece à ordem decrescente dos percentuais de aplicação de epêntese atestados para esses fatores.⁶²

Tabela 24 - Resultado para Informante (inglês)

	Apl./Total	%	Peso relativo
H	15/57	26	0,70
F	15/57	26	0,68
D	15/61	24	0,72
K	14/57	24	0,72
C	13/54	24	0,55
N	12/54	22	0,68
E	12/61	19	0,67
I	11/59	18	0,46
M	10/57	17	0,57
P	7/57	12	0,51
O	7/56	12	0,44
L	6/59	10	0,36
A	6/57	10	0,34
B	3/59	5	0,27
G	3/58	5	0,21
J	2/59	3	0,14
TOTAL	151/922	16	

Input: 0,10

Significância: 0,005

Através da análise dos percentuais de aplicação de epêntese apresentados na tabela acima, percebe-se que houve seis informantes (representados pelos fatores H, F, D, K, C e N nessa tabela) que realizaram epêntese em mais de 20% dos dados produzidos. Por outro lado, houve três informantes (identificados pelos fatores B, G e J na tabela) que realizaram epêntese em menos de 6% dos dados produzidos oralmente.

Esse distanciamento considerável que existe entre os percentuais de produção de epêntese por informante parece nos indicar que a realização de epêntese não depende única e exclusivamente do contexto linguístico em que ela ocorre, mas também depende do

⁶² À semelhança do que aconteceu na rodada 1 do PB, a rodada 1 do inglês apresentou resultados em que a relação entre o percentual de aplicação de epêntese e o peso relativo dos fatores de um mesmo grupo não ficasse muito clara, o que é constatado nas pequenas diferenças de ordenamento nas duas últimas colunas da Tabela 24. Consideramos mais importante, na análise dos dados em inglês, considerar a ordem do percentual de aplicação de epêntese.

informante que a produz. Essa nossa suspeita pode ser confirmada em estudos como o de CAGLIARI (2001).

CAGLIARI (2001) investigou a produção dos encontros consonantais /sC/ do inglês por falantes nativos de PB. Ao analisar os casos em que houve epêntese na produção desses encontros /sC/, a referida autora constatou que, enquanto um de seus informantes havia produzido 0% de aplicação de epêntese, outro aplicou epêntese em 90% dos dados produzidos (esses participantes possuíam o mesmo tempo de instrução na LE). A discrepância entre os percentuais de aplicação de epêntese entre os informantes do estudo de CAGLIARI (2001) e a discrepância apontada na aplicação desse fenômeno entre os informantes de nosso estudo parecem confirmar que a aplicação da epêntese está também condicionada a características individuais do falante que a produz.

No que diz respeito à aplicação da epêntese na língua inglesa, conclui-se que a diferença registrada entre os informantes de nosso estudo pode ser explicada com base na *variação linguística intra-aprendiz* que assumimos existir na interlíngua de um indivíduo (cf. 1.1). Nesse sentido, fatores envolvidos no processo de aquisição de segunda língua, como aptidão, motivação, autonomia, etc. podem explicar o fato de que alguns de nossos informantes aplicaram mais epêntese do que outros.

3.5 Comparações entre a Amostra de Fala 1 e a Amostra de Fala 2

Esta seção tem por objetivo comparar os resultados obtidos para os grupos de fatores selecionados nas análises estatísticas referentes à Amostra de Fala 1 (PB) com os resultados obtidos para os grupos de fatores selecionados nas análises estatísticas referentes à Amostra de Fala 2 (inglês). Com isso, queremos ver até que ponto pode atuar a transferência linguística que admitimos existir entre a L1 e a interlíngua.

Apresentamos, no Quadro 9 (cf. próxima página), os grupos de fatores selecionados em PB e em inglês em cada rodada realizada. Podemos constatar, através da análise do desse quadro, que todos os grupos de fatores selecionados nas rodadas com dados do inglês foram também selecionados nas rodadas do PB, sem exceções. São estes os grupos: ‘tipo de consoante perdida’, ‘vozeamento da consoante perdida’, ‘acento’ e ‘informante’. Devemos observar que a comparação dos resultados para o grupo ‘acento’ já foi realizada na seção 3.4.2.2 e que, por isso, não a vamos rerepresentá-la aqui. Cabe aqui apenas lembrar que, tanto

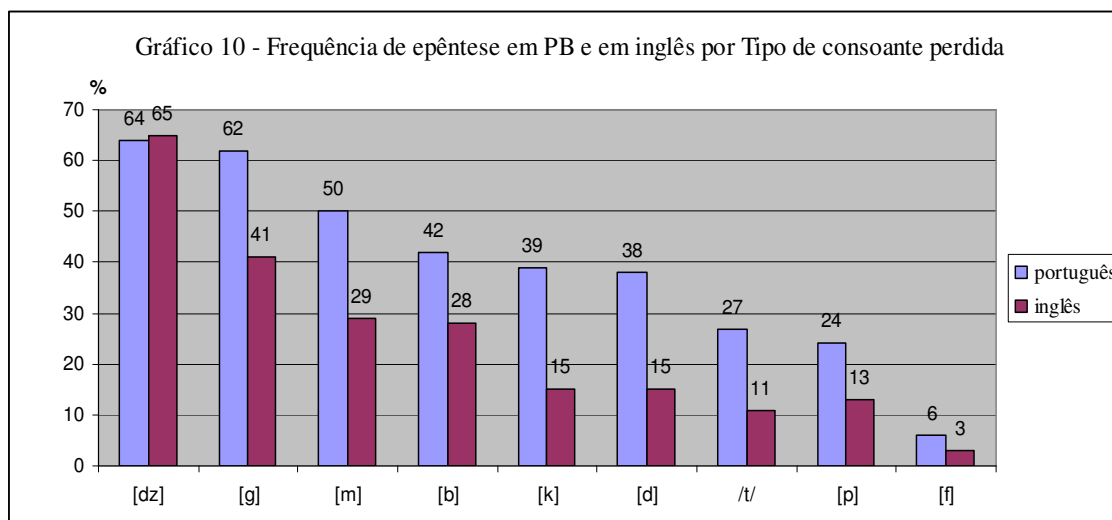
no PB quanto na interlândia, os dados com consoante perdida no contexto acentual postônico apresentaram frequências reduzidas de aplicação de epêntese.

Quadro 9 - Grupos selecionados por rodada em PB e em inglês

	Rodada	
	1	2
Grupos selecionados em PB	1º Tipo de consoante perdida 2º Informante 3º Tipo de contexto seguinte 4º Tipo de <i>cluster</i> 5º Contexto morfológico	1º Vozeamento da consoante perdida 2º Vozeamento do contexto seguinte 3º Contexto morfológico 4º Tipo de <i>cluster</i> 5º Acento
Grupos selecionados em inglês	1º Tipo de consoante perdida 2º Acento 3º Informante	1º Vozeamento da consoante perdida 2º Acento

O grupo ‘vozeamento da consoante perdida’, em PB, apresentou o fator ‘vozeado’ como favorecedor de aplicação de epêntese (peso relativo 0,56). Em inglês, o mesmo aconteceu (inclusive de forma mais evidente, cf. peso relativo 0,70 para ‘vozeado’). Essas informações permitem inferir que a marcação de traço [+vozeado] em consoantes perdidas, por ser idêntica à marcação atestada em vogais, crie um ambiente propício para a realização de uma vogal epentética, seja em PB ou em inglês como LE.

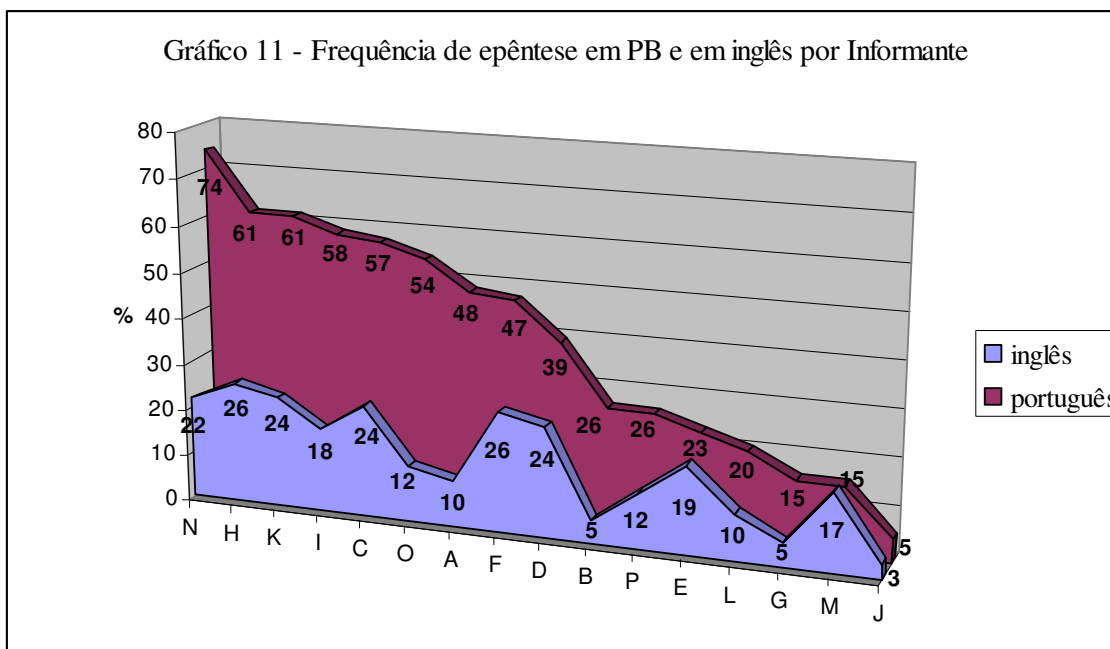
Passemos à análise do grupo ‘tipo de consoante perdida’. O Gráfico 10 apresenta a frequência de epêntese para as consoantes perdidas produzidas tanto em PB quanto em inglês.



A primeira informação que extraímos do Gráfico 10 é que a epêntese é sempre mais frequente após consoantes perdidas produzidas em vocábulos do PB (exceto após [dz], representada no gráfico por [dz]). Outra informação que podemos obter desse gráfico é que os percentuais de aplicação de epêntese para as consoantes perdidas do inglês acompanham a ordem decrescente apresentada para os percentuais de aplicação de epêntese para as consoantes perdidas em português (exceto o percentual de [p], que foi ligeiramente superior ao percentual de /t/).

Tais informações parecem mostrar que, em nosso estudo, existe uma correspondência entre a frequência de epêntese que se aplica em português e a que se aplica em inglês após um determinado tipo de consoante perdida. Em outras palavras, as consoantes perdidas que tendem a favorecer a aplicação de epêntese em PB são também as consoantes que estão mais frequentemente associadas à aplicação de epêntese na interlíngua. Essa constatação nos permite admitir que, em nossos dados, a transferência linguística que ocorre da L1 para a interlíngua é bastante operante.

Cabe, por fim, comparar as frequências de aplicação de epêntese em PB e em inglês entre cada um dos 16 informantes que participaram de nossa pesquisa. Para realizarmos essa comparação, valemo-nos do Gráfico 11 abaixo.



Através das informações do Gráfico 11, podemos primeiramente constatar que a aplicação de epêntese por informante é sempre mais frequente nas produções de fala em PB, exceto pelo informante M. Outra constatação que pode ser feita é que, dos cinco informantes que apresentam as maiores frequências de epêntese em português (N, H, K, I e C), todos se incluem entre os que aplicaram epêntese em inglês acima da média geral de aplicação nesse idioma (16%).

Ao olharmos, no Gráfico 11, para os resultados sobre os informantes B, P, E, L, G, M e J, constatamos que tais informantes produziram epêntese abaixo da média atestada para o PB (39%). Percebe-se, pois, que esses informantes estão também entre os que produziram epêntese abaixo da média nos dados do inglês (16%), à exceção de E e de M. A análise dos resultados para esses informantes nos permite ainda constatar que o informante J foi o que menos produziu epêntese, tanto em português quanto em inglês.⁶³

Através das constatações feitas, concluímos que existe certa correlação entre a frequência com que um falante nativo de PB aplica epêntese nessa língua e a frequência com que esse falante aplica epêntese na interlíngua. Em outras palavras, acreditamos que, quanto mais um falante produz epêntese na L1, maiores são as chances de ele produzir epêntese na interlíngua. Essa nossa crença se baseia principalmente no princípio da transferência linguística que admitimos operar entre a L1 e a interlíngua. Note-se, no entanto, que essa correlação nem sempre é confirmada em nossos resultados: de acordo com o Gráfico 11, houve informantes, como O e A, por exemplo, que aplicaram epêntese em PB com frequência alta (acima da média); esses mesmos informantes, entretanto, ao produzirem dados em inglês, aplicaram epêntese com frequência abaixo da média. Para esses casos, admitimos que o percentual reduzido de produção de epêntese em inglês se justifique pela atuação de fatores de natureza intra-aprendiz, como motivação, autonomia, proficiência na língua, etc. que estão acima do poder exercido pela transferência linguística.

⁶³ Na seção 3.3.2.3, apresentamos uma razão que tenta justificar a baixa frequência de epêntese atestada para esse informante.

4 CONCLUSÕES

No presente estudo, realizamos uma análise quantitativa da aplicação do fenômeno de epêntese vocálica medial na produção oral de vocábulos do PB e de vocábulos do inglês. Com base na suspeita de COLLISCHONN (2002) de que a inserção de um elemento vocálico após o prefixo *sub-* parece ser mais frequente do que após uma consoante que esteja no interior de uma palavra, investigamos nos dados coletados, dentre outros aspectos, a possível existência de diferenças entre os percentuais de aplicação de epêntese para *clusters* localizados no interior de uma base e os percentuais de aplicação de epêntese para *clusters* localizados na fronteira entre prefixos e bases.

Partimos do pressuposto de que a inserção de uma vogal em *clusters* de vocábulos do inglês seja reflexo da epêntese vocálica que ocorre em PB como estratégia de reparo a uma sílaba malformada. Em outras palavras, pressupomos a existência e a atuação de um princípio de *transferência linguística* entre a L1 e a interlíngua de um aprendiz, que faz com que a estratégia de reparo silábico utilizada em PB seja igualmente utilizada na interlíngua.

Outro pressuposto admitido em nosso estudo foi o de que, entre os prefixos do PB e do inglês que podem motivar epêntese, existem prefixos mais transparentes e prefixos menos transparentes. Os prefixos mais transparentes, em nosso entendimento, estariam relacionados a frequências mais altas de aplicação de epêntese em comparação com os prefixos menos transparentes. Para verificar o grau de transparência dos prefixos relacionados à epêntese, aplicamos um teste de transparência com dados do PB e um teste de transparência com dados do inglês. Os resultados do teste de transparência com dados do PB indicaram que *sob-* e *sub-* são os mais transparentes, enquanto os resultados do teste de transparência com dados do inglês indicaram que tanto *arch-* como *out-*, *post-* e *sub-* são transparentes em inglês. Com base nesses resultados, hipotetizamos que o percentual de aplicação de epêntese em vocábulos que contêm esses prefixos seria maior do que o percentual de aplicação de epêntese em vocábulos com outros prefixos ou sem prefixo algum.

Antes de realizar a análise dos resultados que obtivemos a partir da coleta de dados em português e em inglês, levantamos uma série de hipóteses sobre os fatores linguísticos e os fatores extralinguísticos que poderiam estar relacionados à aplicação de epêntese em cada uma dessas línguas. Para efetuarmos o levantamento dessas hipóteses, baseamo-nos nos resultados de outros estudos quantitativos sobre a inserção vocálica em PB e em inglês, em especial os estudos de COLLISCHONN (2002) e de PEREYRON (2008).

Os resultados referentes à coleta de dados em português mostram que a epêntese ocorreu em 39,4% dos contextos analisados. Os fatores linguísticos apontados como significativos para a aplicação de epêntese foram ‘vozeamento da consoante perdida’, ‘tipo de consoante perdida’, ‘vozeamento do contexto seguinte’, ‘tipo de contexto seguinte’, ‘contexto morfológico’, ‘tipo de *cluster*’ e ‘acento’.

Para o grupo ‘vozeamento da consoante perdida’, o fator ‘vozeado’ foi o que mais favoreceu a inserção vocálica (peso relativo igual a 0,56). Para o grupo ‘tipo de consoante perdida’, o fator ‘[g]’ (ou seja, a consoante velar vozeada) foi o que mais favoreceu a aplicação de epêntese (peso relativo 0,82). Outras consoantes perdidas que demonstraram favorecer a epêntese em PB foram [dʒ], [m] e [b]. Por outro lado, a consoante perdida [f] foi a que menos favoreceu a aplicação de epêntese. Num total de 16 ocorrências de [f] em coda de sílaba, a epêntese foi aplicada apenas uma vez. Através da amalgamação das consoantes perdidas analisadas em três supercategorias (‘velar’, ‘alveolar’ e ‘labial’), pudemos constatar que, em nossos dados, as velares são as que mais favorecem epêntese em PB. Esse resultado foi de encontro aos resultados obtidos por COLLISCHONN (2002), que indicaram maior frequência de epêntese após consoantes perdidas alveolares.

Na análise do grupo ‘vozeamento do contexto seguinte’, observou-se que o contexto vozeado foi o que mais favoreceu a inserção de uma vogal (peso relativo 0,58). Para o grupo ‘tipo de contexto seguinte’, os fatores [x,h,ʁ], [l] e [m,n] foram os que apresentaram os maiores pesos relativos (ou as maiores frequências de aplicação de epêntese). A análise dos vocábulos que contêm esses tipos de contexto seguinte em nossa amostra de fala permitiu entender o porquê da frequência elevada de epêntese antes desses contextos. Ainda para o grupo ‘tipo de contexto seguinte’, as fricativas [s] e [z] e as africadas [tʃ] e [dʒ] foram as que menos estiveram relacionadas à aplicação de epêntese. No que diz respeito aos resultados para [s] e [z] especificamente, acreditamos ter confirmado em nossos dados a hipótese (cf. COLLISCHONN, 2002) de que essas duas sibilantes podem formar uma espécie de africada com as oclusivas que as antecedem (como se pode atestar em [k^war.tʃu] *quartzo*).

A variável ‘informante’ foi a única variável extralinguística a ser selecionada nas rodadas estatísticas com dados do PB. Os resultados obtidos para essa variável permitiram constatar que a aplicação de epêntese varia muito entre um indivíduo e outro (N aplicou epêntese em 74% dos dados produzidos, enquanto J aplicou epêntese em apenas 5% dos dados que produziu). Esses resultados nos fazem admitir que fatores não linguísticos também podem estar relacionados à aplicação de epêntese.

Os resultados para o grupo ‘contexto morfológico’ nos permitiram constatar, primeiramente, que a frequência de epêntese é ligeiramente maior para o contexto morfológico ‘prefixo + base’ (42,2%) do que para o contexto morfológico ‘interior da base’ (38%). Outra constatação que pudemos fazer através dos resultados para esse grupo foi a de que as frequências de epêntese mais elevadas são atestadas para os fatores ‘*sub-*’ e ‘*sob-*’ (62% e 55%, respectivamente). Através da análise conjunta dos resultados do teste de transparência de prefixos em PB e dos resultados obtidos para a variável ‘contexto morfológico’, concluímos que *sob-* e, principalmente, *sub-*, são prefixos que apresentam configuração morfofonológica diferenciada quando comparados com os demais prefixos que analisamos.

Através dos resultados obtidos para a variável ‘tipo de *cluster*’, percebemos que a epêntese é mais frequentemente aplicada em *clusters* que só ocorrem em PB (ou melhor, ocorrem em PB, mas não ocorrem em inglês). Nosso objetivo era, em princípio, comparar os resultados obtidos para ‘tipo de *cluster*’ em PB com os resultados obtidos para esse grupo em inglês. Essa comparação acabou sendo impossibilitada pelo fato de que, em inglês, o grupo ‘tipo de *cluster*’ não foi considerado relevante nas rodadas estatísticas executadas.

A última variável selecionada nas rodadas estatísticas para os dados do PB foi ‘acento’. Os resultados individuais para os fatores dessa variável nos permitiram constatar que a frequência de aplicação de epêntese no contexto postônico (*‘ritmo*) é menor do que a frequência média de aplicação de epêntese (peso relativo 0,39). Esses resultados se assemelham bastante aos resultados obtidos por COLLISCHONN (2002), uma vez que, nesse estudo, a aplicação de epêntese na pauta acentual postônica também foi pouco frequente. Dadas as semelhanças entre os dois estudos quanto a esse aspecto, optamos por reforçar a conclusão de COLLISCHONN (2002), qual seja: a explicação para o comportamento da epêntese se baseia nas restrições de acento em PB, que favorecem a colocação do acento na penúltima sílaba da palavra, sendo evitada a inserção de sílaba à direita, para impedir o deslocamento de acento.

Na análise dos resultados referentes aos dados de língua inglesa, vimos que a epêntese foi aplicada em 15,5% do total de ocorrências de contexto para aplicação. Os grupos selecionados nas rodadas estatísticas foram ‘vozeamento da consoante perdida’, ‘tipo de consoante perdida’, ‘acento’ e ‘informante’. Esperávamos que o grupo ‘contexto morfológico’ fosse igualmente selecionado nas análises estatísticas com dados do inglês, mas isso não aconteceu. Ainda que esse grupo não tenha sido selecionado nas rodadas estatísticas, decidimos apresentar os resultados referentes à frequência de aplicação de epêntese em cada

contexto morfológico admitido. A análise desses resultados nos permitiu constatar que, nos contextos morfológicos analisados, a aplicação de epêntese é mais frequente em vocábulos que contêm os prefixos *arch-* e *sub-*.

Os resultados referentes ao grupo ‘vozeamento da consoante perdida’ nos permitiram concluir que, em inglês, a recorrência de epêntese é maior após consoantes perdidas vozeadas. O percentual médio de aplicação de epêntese após consoante vozeadas foi de 29%, valor equivalente a quase o dobro do percentual médio de aplicação de epêntese em inglês. De acordo com os resultados obtidos para o grupo ‘tipo de consoante perdida’, as consoantes [dʒ], [g], [tʃ], [b] e [m] estão relacionadas a frequências elevadas de aplicação de epêntese. Interpretamos que a frequência de epêntese após [dʒ] só foi elevada porque os falantes, ao lerem a palavra *hodgepodge*, inseriram uma vogal epentética por influência do *e* expresso na forma gráfica dessa palavra. Em razão dessa interpretação, consideramos que [g] foi o fator que genuinamente apresentou a maior frequência de epêntese no grupo ‘tipo de consoante perdida’. Para esse grupo ainda, os fatores [n], [t] e [f] foram os que menos estiveram relacionados à aplicação de epêntese. No que concerne especificamente aos resultados para [t] e [f], admitimos que as consoantes a que esses fatores se referem estejam entre as primeiras consoantes obstruintes a serem admitidas na coda silábica do inglês aprendido por brasileiros. Essa nossa crença parece que se confirma em outros estudos (cf. PEREYRON, 2008), que também mostraram que as consoantes perdidas [t] e [f] estão relacionadas a percentuais reduzidos de epêntese.

Os resultados para o grupo ‘acento’ em inglês nos permitiram constatar que a epêntese foi menos recorrente em vocábulos nos quais a consoante perdida é registrada numa posição postônica. Essa constatação constituiu uma das razões que nos levou à conclusão de que a *transferência linguística* que admitimos ocorrer da L1 para a interlíngua opera sobre a pauta acentual dessas línguas. Isso quer dizer que, se a realização de epêntese é evitada na pauta postônica em PB, ela também poderá ser evitada na pauta postônica da interlíngua do aprendiz.

Finalmente, na análise dos resultados obtidos para o grupo ‘informante’, pudemos constatar que os percentuais de aplicação de epêntese entre um participante e outro variam consideravelmente. Concluímos, a partir desses resultados, que o fato de um aprendiz produzir mais epêntese do que outro na interlíngua esteja relacionado aos aspectos intra-aprendiz, que incluem motivação, personalidade, autonomia, etc.

Na parte final da análise e discussão de nossos resultados, comparamos os resultados obtidos para os grupos de fatores selecionados nas análises estatísticas referentes aos dados do

PB com os resultados obtidos para os grupos de fatores selecionados nas análises estatísticas referentes aos dados do inglês. Constatamos que todos os grupos selecionados na análise dos dados do inglês haviam sido igualmente selecionados na análise dos dados do PB. Através dessa constatação, suspeitamos que pudesse haver alguma correlação entre os fatores que favorecem a epêntese em PB e os fatores que estão relacionados a frequências elevadas de aplicação de epêntese em inglês.

As comparações que fizemos nos permitiram confirmar que existe uma relação muito próxima entre os fatores que podem condicionar a aplicação da epêntese em PB e os fatores que estão relacionados a percentuais elevados de aplicação de epêntese em inglês. Concluimos que essa relação estreita seja resultado da transferência linguística que opera na direção L1 → interlíngua.

REFERÊNCIAS BIBLIOGRÁFICAS⁶⁴

ALVES, U. K. *O papel da instrução explícita na aquisição fonológica do inglês como L2: evidências fornecidas pela Teoria da Otimidade*. Dissertação de Mestrado. Pelotas: Universidade Católica de Pelotas, 2004.

_____. *A aquisição das seqüências finais de obstruintes do inglês (L2) por falantes do Sul do Brasil: Análise via Teoria da Otimidade*. Tese de Doutorado. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul, 2008.

ARONOFF, M. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press, 1976.

BASILIO, M. *Estruturas lexicais do Português*. Petrópolis: Vozes, 1980.

BECHARA, E. *Moderna gramática portuguesa*. 37. ed., revista e ampliada. Rio de Janeiro: Lucerna, 2006.

BEEBE, L. M. Five sociolinguistic approaches to second language acquisition. In: BEEBE, L. M. (ed.) *Issues in second language acquisition. Multiple perspectives*. Rowley, MA: Newbury House, 1988. pp. 43-77

BETTONI-TECHIO, M. *Production of final alveolar stops in Brazilian Portuguese/English interphonology*. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 2005.

BISOL, L. A sílaba e seus constituintes. In: Neves, M. H. M. (org.). *Gramática do Português Falado*. v.VII. Campinas: Editora da UNICAMP, 1999. pp. 701-742

BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer* (v. 5.1.07) [programa computacional]. 2009. Disponível em <http://www.praat.org/>.

⁶⁴ Incluímos aqui todos os trabalhos que foram referidos ao longo do texto, ainda que alguns deles não tenham sido diretamente consultados.

BOOIJ, G. E. Cliticization as prosodic integration: the case of Dutch. In: *The Linguistic Review*. nº 13. Berlim, 1996. pp. 219-242

BRESCANCINI, C. R. A análise de regra variável e o programa VARBRUL 2S. In: BISOL, L.; BRESCANCINI, C. R. (orgs.). *Fonologia e variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, 2002. pp. 13-76

BROWN, D. H. *Principles of language learning and teaching*. Englewood Cliffs: Prentice Hall Regents, 1987.

CAGLIARI, L. C. *Elementos de fonética do português brasileiro*. Tese de livre docência. Campinas: Universidade de Campinas, 1981.

CAGLIARI, A. *A produção dos encontros consonantais sC do inglês por falantes nativos do português brasileiro*. Dissertação de Mestrado. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2001.

CAMARA JR., J. M. *Problemas de Linguística Descritiva*. Petrópolis: Editora Vozes, 1969.

_____. *Estrutura da língua portuguesa*. Petrópolis: Editora Vozes, 1970.

_____. *Dicionário de linguística e gramática: referente à língua portuguesa*. 17ª ed. Petrópolis: Editora Vozes, 1996.

CAMBRIDGE Advanced Learner's Dictionary. Versão disponível na Internet <http://dictionary.cambridge.org/>. (acessado em 10/01/2009)

CARDOSO, W. The variable development of English word-final stops by Brazilian Portuguese speakers: a stochastic optimality theoretic account. In: *Language Variation and Change*. v. 19. Cambridge, Massachusetts: Cambridge University Press, 2007.

CHOMSKY, N.; HALLE, M. *The Sound Pattern of English*. New York: Harper & Row, 1968.

CLEMENTS, G. N. The role of sonority cycle in core syllabification. In: KINGSTON, J.; BECKMAN, M. (orgs.). *Papers in laboratory phonology 1: between the Grammar and Physics of Speech*. New York: CUP, 1990. pp. 283-333

CLEMENTS, G. N.; KEYSER, S. J. *CV phonology: a generative theory of the syllable*. Cambridge, Massachusetts: MIT Press, 1983.

COLLISCHONN, G. *Análise prosódica da sílaba em português*. Tese de Doutorado. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul, 1997.

_____. A epêntese vocálica no português do Sul do Brasil: análise variacionista e tratamento pela Teoria da Otimalidade. In: *Letras de Hoje*. v. 35, nº 1. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul, março de 2000. pp. 285-318

_____. A epêntese vocálica no português do Sul do Brasil. In: BISOL, L.; BRESCANCINI, C. R. (orgs.). *Fonologia e variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, 2002. pp. 205-230

_____. Epêntese vocálica no português do Sul do Brasil: variáveis extralinguísticas. In: *Revista Letras*. n. 61, especial. Curitiba: Editora UFPR, 2003. pp. 285-297

_____. Epêntese vocálica e restrições de acento no português do Sul do Brasil. In: *Signum: Estudos da Linguagem*. n. 7/1. Londrina, junho de 2004. pp. 61-78

COLLISCHONN, G.; SCHWINDT, L. C. Considerações sobre a sequência /sC/ inicial em português brasileiro. In: *Revista Língua(gem)*. v. 2, nº 2. Macapá, 2005. pp. 249-266

CORNELIAN JR., D. *Brazilian learners' production of initial /s/ clusters: phonological structure and environment*. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 2003.

CRYSTAL, D. *A Dictionary of Linguistics and Phonetics*. 5th edition. Oxford: Blackwell Publishing, 2003.

CUNHA, C.; CINTRA, L. F. L. *Nova gramática do português contemporâneo*. 3. ed. Rio de Janeiro: Nova Fronteira, 2001.

DICIONÁRIO Eletrônico Michaelis. v. 5.0. DTS Software Ltda., 1998.

ELLIS, R. *The study of Second Language Acquisition*. Oxford: Oxford University Press, 1994.

GOLDSMITH, J. *Autosegmental Phonology*. Tese de Doutorado (PhD). Cambridge, Massachusetts: MIT Press, 1976.

GUY, G. VARBRUL: análise avançada. In: *Cadernos de Tradução*. n. 1. Tradução de Ana Maria Stahl Zilles. Porto Alegre: Universidade Federal do Rio Grande do Sul, janeiro de 1998. pp. 27-50

HAMMOND, M. *The phonology of English: a prosodic optimality-theoretic approach*. Oxford: Oxford University Press, 1999.

ITÔ, J. *Syllable theory in Prosodic Phonology*. Tese de Doutorado. University of Massachusetts at Amherst: MIT Press, 1986.

KAHN, D. *Syllable based generalizations in English phonology*. Tese de Doutorado. University of Massachusetts: MIT Press, 1976.

KATAMBA, F.; STONHAM, J. T. *Morphology*. 2. ed. New York: Palgrave Macmillan, 2006.

KINZEL, G. R. *O papel da sonoridade na epêntese do português do sul do Brasil*. Monografia de conclusão do curso de graduação em Letras. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2000.

KOERICH, R. D. *Perception and production of word-final vowel epenthesis by Brazilian EFL students*. Tese de Doutorado. Florianópolis: Universidade Federal de Santa Catarina, 2002.

- LABOV, W. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, 1972.
- LADO, R. *Linguistics across cultures: applied linguistics for teachers*. Ann Arbor: University of Michigan Press, 1957.
- LENNEBERG, E. *Biological Foundations of Language*. New York: Wiley, 1967.
- LONGMAN Dicionário Escolar. v. 1.0.0.1. Pearson Education Ltd., 2004.
- MCCARTHY, J.; PRINCE, A. *Generalized alignment*. 1993. Disponível na Internet em <http://roa.rutgers.edu>.
- MONARETTO, V. N. de O. *Um reestudo da vibrante: análise variacionista e fonológica*. Tese de Doutorado. Porto Alegre, Pontifícia Universidade Católica do Rio Grande do Sul, 1997.
- MERRIAM-WEBSTER Unabridged Dictionary. v. 2.5. Merriam-Webster Incorporated, 2000.
- PEREYRON, L. *Epêntese vocálica em encontros consonantais mediais por falantes porto-alegrenses de inglês como língua estrangeira*. Dissertação de Mestrado. Porto Alegre, Pontifícia Universidade Católica do Rio Grande do Sul, 2008.
- PERSEGONA, M. *A fossilização no processo de aquisição de segunda língua*. Dissertação de Mestrado em Linguística. Curitiba: Universidade Federal do Paraná, 2005.
- PIGOTT, G. Epenthesis and syllable weight. *Natural Language and Linguistic Theory*. V. 13, p. 283-326, 1995.
- PRINCE, A.; SMOLENSKY, P. *Optimality Theory*. Rutgers University, 1993. Ms.
- QUIRK, R.; GREENBAUM, S. *A university grammar of English*. London: Longman, 1973. pp. 431-435

QUIRK, R.; GREENBAUM, S.; LEECH, G.; SVARTVIK, J. *A comprehensive grammar of the English language*. Harlow: Longman, 1985. pp. 1540-1546

REBELLO, J. T. *The acquisition of initial /s/ clusters by Brazilian EFL learners*. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 1997.

ROACH, P. *English Phonetics and Phonology - a practical course*. 3. ed. Cambridge: Cambridge University Press, 2000.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. *GoldVarb: a multivariate analysis application* (v. 3.0b3) [programa computacional]. 2005. Disponível em <http://individual.utoronto.ca/tagliamonte/goldvarb.htm>

SCHWINDT, L. C. *O prefixo no português brasileiro: análise morfofonológica*. Tese de Doutorado. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul, 2000.

_____. Revisitando o estatuto prosódico e morfológico de palavras prefixadas do PB em uma perspectiva de restrições. In: *Alfa*. 52 (2). São Paulo, 2008. pp. 391-404

SCOVEL, T. Foreign accents, language acquisition, and cerebral dominance. In: *Language learning*. v. 19. 1969. pp. 245-254

SELKIRK, E. The syllable. In: HULST, H. van der; SMITH, N. (ed.). *The structure of phonological representations*. Dordrecht: Foris. pp. 337-383

SILVA, T. C. *Fonética e Fonologia do Português*. 6.ed. São Paulo: Contexto, 2002.

VIGÁRIO, M. C. *The prosodic word in European Portuguese*. Tese de Doutorado. Lisboa: Universidade de Lisboa, 2001.

ANEXOS

ANEXO I

Leia antes de responder. O presente exercício constitui parte de um projeto vinculado ao Programa de Pós-Graduação em Letras da UFRGS - Área de Concentração Estudos da Linguagem - e tem por objetivo analisar a intuição dos falantes de português brasileiro com relação a aspectos linguísticos de vocábulos dessa língua. As informações prestadas serão mantidas sob sigilo. Por favor, assine este documento como forma de consentimento de sua participação nesse projeto e registre seu e-mail para eventual contato. Muito obrigado!

Nome: _____ Assinatura: _____

E-mail: _____

1. Circule a parte das palavras em itálico abaixo que, **em sua opinião**, possui o núcleo do significado das mesmas. Obs.: Algumas das palavras possuem sua definição entre parênteses.

Exemplos: mar; bananeira; presentimento; comprava.

- | | |
|--|--|
| 01. <i>obedecer</i> | 24. <i>eczema</i> (dermatose inflamatória) |
| 02. <i>ablação</i> (ação de tirar, de arrancar) | 25. <i>subtração</i> |
| 03. <i>adaptar</i> | 26. <i>adjunto</i> |
| 04. <i>suburitano</i> (proveniente de Subura, Itália) | 27. <i>sobgrave</i> (inferior ao som mais lento) |
| 05. <i>obviar</i> (evitar, prevenir a ocorrência de) | 28. <i>obediente</i> |
| 06. <i>subsolo</i> | 29. <i>sublocar</i> (alugar para outrem) |
| 07. <i>abacate</i> | 30. <i>eclético</i> |
| 08. <i>eclipse</i> | 31. <i>adrenalina</i> |
| 09. <i>sobalçou</i> (exaltou, aclamou) | 32. <i>sobdominante</i> |
| 10. <i>adotado</i> | 33. <i>abnegar</i> (renunciar a) |
| 11. <i>abaxial</i> (que está afastado do eixo) | 34. <i>sobejo</i> (enorme, excessivo) |
| 12. <i>subarquado</i> (com pequeno arqueamento) | 35. <i>obstruir</i> |
| 13. <i>sobrado</i> | 36. <i>administração</i> |
| 14. <i>adligação</i> (fixação de uma planta em outra) | 37. <i>ecumênico</i> |
| 15. <i>obturar</i> (fechar, impedir a passagem) | 38. <i>abstenção</i> |
| 16. <i>eclesiástico</i> (relativo à Igreja) | 39. <i>subnutrido</i> |
| 17. <i>sobescavar</i> (abrir espaço por baixo) | 40. <i>adolescente</i> |
| 18. <i>ecdêmico</i> (relativo a mal de alcance restrito) | 41. <i>obnubilar</i> (escurecer) |
| 19. <i>abestado</i> (que se embruteceu) | 42. <i>ecmnésia</i> (certo distúrbio de memória) |
| 20. <i>subarrendar</i> (transferir a terceiro direitos de) | 43. <i>obrigou</i> |
| 21. <i>obumbrar</i> (tornar sombrio, escuro) | 44. <i>advogada</i> |
| 22. <i>economia</i> | 45. <i>sobestar</i> (ser inferior a) |
| 23. <i>abdicar</i> (renunciar, desistir de) | 46. <i>ablitense</i> (relativo a Ablitas, Espanha) |

ANEXO II

Leia antes de responder. O presente exercício constitui parte de um projeto vinculado ao Programa de Pós-Graduação em Letras da UFRGS - Área de Concentração Estudos da Linguagem - e tem por objetivo analisar a intuição de brasileiros aprendizes de inglês com relação a aspectos linguísticos de vocábulos dessa língua. As informações prestadas serão mantidas sob sigilo. Por favor, preencha os dados solicitados abaixo e assine este documento como forma de consentimento de sua participação nesse projeto. Muito obrigado!

Nome: _____ Assinatura: _____

E-mail: _____

É (ou foi) estudante de Letras? () Sim () Não

É falante nativo de português? () Sim () Não

Há quanto tempo estuda inglês (em anos e meses)? _____

1. Circule a parte das palavras em itálico abaixo que, **em sua opinião**, possui o núcleo do significado das mesmas. Obs.: O significado das palavras está entre parênteses.

Exemplos: sea; worker; preplanned; cloudy

- | | |
|--|---|
| 01. <i>outrage</i> (an act of violence) | 15. <i>subah</i> (a province of the Mogul Empire) |
| 02. <i>archbishop</i> (a chief clergyman) | 16. <i>postconsonantal</i> (following a consonant) |
| 03. <i>outdated</i> (antiquated) | 17. <i>outrun</i> (to go beyond a particular limit) |
| 04. <i>postocular</i> (located behind the eye) | 18. <i>subnitrate</i> (a basic salt of nitric acid) |
| 05. <i>archil</i> (a violet dye obtained from lichens) | 19. <i>archin</i> (a Russian unit of length) |
| 06. <i>subcutaneous</i> (situated beneath the skin) | 20. <i>postdate</i> (to assign an event to a future day) |
| 07. <i>outfighting</i> (to be defeating in a fight) | 21. <i>archfiend</i> (a chief enemy, especially Satan) |
| 08. <i>posthitis</i> (inflammation of the prepuce) | 22. <i>subanum</i> (a pagan people of Philippines) |
| 09. <i>archduke</i> (a sovereign prince) | 23. <i>outmaneuver</i> (to surpass in maneuverability) |
| 10. <i>sublet</i> (to rent a rented property to sb. else) | 24. <i>subarid</i> (moderately or slightly dry) |
| 11. <i>outbreathe</i> (to exhale) | 25. <i>postwar</i> (occurring after a war) |
| 12. <i>postiche</i> (false hair) | 26. <i>outardes</i> (wild geese found in Canada) |
| 13. <i>archpriest</i> (the highest title of honor given
a member of the secular clergy) | 27. <i>archdeacon</i> (a chief ecclesiastical
dignitary next in rank below a bishop) |
| 14. <i>postfix</i> (to affix at the end) | 28. <i>subdominant</i> (incompletely commanding) |

ANEXO III

LISTA DE FRASES EM PORTUGUÊS

Por favor, proceda à leitura das seguintes frases:

1. Eduardo optou por estudar Biologia. Mal sabia ele que, com essa opção, teria de estudar sobre a adligação em plantas briófitas e epífitas, além de aprender todos os detalhes sobre a vida das hidnáceas, das rosáceas e das marcgraviáceas.

2. Luiz Carlos abdicou de vários cursos na universidade antes de se formar em Odontologia. Hoje ele é professor adjunto na área de obturação em São Paulo, além de ser subchefe do Departamento de sua Unidade.

3. Os doutores Mário e Júlio são irmãos e ambos são médicos. Este é especializado no tratamento de aftas, eczemas, micoses, actinites e outras dermatoses, enquanto que aquele é cirurgião e atua há anos na correção de ecpiesmas cranianos.

4. Djalma estava decidido a se formar em História, mas desistiu ao descobrir que o ritmo de estudos seria muito acelerado. Após algumas noitadas regadas a absinto, uísque e vodca, absorveu a ideia de que o melhor seria aprender sobre a vida de duendes, magbás, pigmeus, bodisatvas e outras sub-raças, existentes ou não.

5. Floriano se formou em Ciências Sociais. Com o dom da oratória, vive dizendo que sabe o significado da palavra obversão e que condena aqueles “que obfirmam no equívoco de jamais criticar os governos que sobpõem o particular ao público”.

6. Apenas um geólogo é capaz de distinguir uma pedra de uma rocha e um quartzo de um cristal de rocha, assim como apenas um artista plástico sabe diferenciar um protótipo de um éctipo.

7. Muitas pessoas vão em busca do Doutor Sigmund na esperança de encontrarem uma cura para doenças graves como amnésias, transtornos obsessivos compulsivos e até mesmo apneias.

8. Adriana Souza, economista de renome, advertiu que a crise poderá provocar a chamada estagflação em nosso País, ocasionando o aumento concomitante de desempregos e de preços de produtos.

9. Alguns músicos possuem o que é chamado de ouvido absoluto. Essas pessoas sabem identificar o tom dos sons que ouvem e têm a capacidade de distinguir sons sobgraves de sons graves, por exemplo.

10. Joana, que estudava Letras, era a única do grupo que sabia que delta, epsilon, lambda e omega eram letras do alfabeto grego.

11. Etimologistas ainda hoje buscam descobrir as origens das palavras ‘édico’, ‘sádico’, ‘ébase’ e ‘perífrase’.

12. Com a nova ortografia, ninguém mais sabe se a forma correta de grafar as palavras ‘ad-rogar’ e ‘ad-digital’ leva hífen ou não.

13. A UFRGS deverá lançar nos próximos anos um programa de estudos cujo objeto de investigação é o tratamento de doenças ecdêmicas na região noroeste do estado.

ANEXO IV – Lista de encontros consonantais que favorecem a epêntese medial em PB e na aquisição de inglês como LE por brasileiros *

Contexto	Exemplo em português	Exemplo em inglês	
/p/ +	t	<i>captou</i>	<i>neptune</i>
	k	-	<i>napkin</i>
	m	-	<i>chipmunk</i>
	n	<i>apneia</i> **	<i>grapnel</i>
	θ	-	<i>naphtha</i>
	s	<i>psicose</i>	<i>capsule</i>
	ʃ	-	<i>option</i>
/b/ +	tʃ	<i>réptil</i>	<i>capture</i>
	p	<i>subproduto</i>	<i>subpopulation</i>
	b	<i>sub-base</i>	<i>subbasement</i>
	t	<i>obter</i>	<i>obtain</i>
	d	<i>abduzir</i>	<i>rhabdom</i>
	k	<i>subconsciente</i>	<i>babka</i>
	g	<i>sobgrave</i>	<i>subgroup</i>
	m	<i>submarino</i>	<i>submerse</i>
	n	<i>abnegado</i>	<i>obnoxious</i>
	f	<i>subfaturar</i>	<i>obfuscate</i>
	v	<i>óbvio</i>	<i>obvious</i>
	s	<i>absoluto</i>	<i>lobster</i>
	z	<i>obséquio</i>	<i>sabzi</i>
	ʃ	<i>subchefe</i>	<i>subshining</i>
	ʒ	<i>objeto</i>	-
tʃ	<i>subtítulo</i>	<i>subchief</i>	
dʒ	<i>abdicar</i>	<i>object</i>	
l	<i>sublocação</i>	<i>sublicense</i>	
x	<i>sub-reptício</i>	-	
/t/ +	p	-	<i>outpass</i>
	b	-	<i>outbreathe</i>
	t	-	<i>posttest</i>
	d	-	<i>postdate</i>
	k	<i>nutca</i>	<i>catkin</i>
	g	<i>wittgensteiniano</i>	<i>postgraduate</i>
	m	<i>ritmo</i>	<i>atmosphere</i>
	n	<i>etnia</i>	<i>chutney</i>
	f	<i>portfólio</i>	<i>outfight</i>
	v	<i>bodisatva</i>	<i>latvian</i>
	θ	-	<i>outthink</i>
	s	<i>futsal</i>	<i>flotsam</i>
	z	<i>quartzo</i>	<i>postzygapophysis</i>
ʃ	-	<i>outshine</i>	

* Lista elaborada com base na revisão das listas apresentadas em CAGLIARI (1981) *apud* COLLISCHONN (2002:206) e em PEREYRON (2008:53), feita a partir de consulta ao *Dicionário Eletrônico Houaiss da Língua Portuguesa* (2002, v. 1.0.5) e ao *Merriam-Webster Unabridged Dictionary* (2000, v. 2.5).

** Em negrito estão os *clusters* favorecedores de epêntese não apontados em estudos anteriores.

		tʃ	-	-
		dʒ	-	<i>outjockey</i>
		ɾ	-	<i>outrun</i>
/d/	+	p	<i>adpresso</i>	<i>adpromission</i>
		b	<i>strindberguiano</i>	<i>tidbit</i>
		k	<i>vodca</i>	<i>vodka</i>
		g	<i>Edgar</i>	<i>dodgasted</i>
		m	<i>admirar</i>	<i>admire</i>
		n	<i>adnominal</i>	<i>kidney</i>
		v	<i>advogado</i>	<i>advantage</i>
		s	<i>adstringente</i>	<i>adsorb</i>
		z	<i>adzâneni</i>	<i>adze</i>
		ʃ	<i>mandchu</i>	<i>hardship</i>
		ʒ	<i>adjetivo</i>	<i>adjective</i>
		dʒ	<i>ad-digital</i>	-
		l	<i>adligar-se</i>	<i>bedlam</i>
x	<i>ad-rogar</i>	-		
/k/	+	p	<i>ecpiesma</i>	-
		b	<i>écbase</i>	<i>ecbolics</i>
		t	<i>compacto</i>	<i>cactus</i>
		d	<i>ecdêmico</i>	<i>ekdemite</i>
		g	<i>marcgraviáceas</i>	<i>marcgravia</i>
		m	<i>ecmnésia</i>	<i>acme</i>
		n	<i>técnica</i>	<i>acne</i>
		θ	-	<i>ichthyology</i>
		s	<i>fixe</i>	<i>accent</i>
		z	<i>eczema</i>	-
		ʃ	-	<i>auction</i>
tʃ	<i>actinite</i>	<i>picture</i>		
dʒ	<i>édico</i>	-		
/g/	+	p	<i>Nagpur</i>	<i>Nagpur</i>
		b	<i>magbá</i>	<i>rugby</i>
		d	<i>Bagdá</i>	<i>magdalen</i>
		m	<i>pigmeu</i>	<i>enigma</i>
		n	<i>ignorância</i>	<i>magnet</i>
		f	<i>estagflação</i>	<i>mugful</i>
		s	<i>UFRGS</i>	-
		z	-	<i>eczema</i>
		ʒ	-	<i>luxurious</i>
dʒ	-	<i>suggest</i>		
/m/	+	n	<i>amnésia</i>	<i>amnesia</i>
/n/	+	m	-	<i>enmity</i>
/f/	+	t	<i>afta</i>	<i>nifty</i>
		n	<i>dafnomancia</i>	<i>hafnium</i>
		θ	-	<i>diphthong</i>
/v/	+	n	<i>ovni</i>	-
/θ/	+	b	-	<i>athbash</i>
		m	-	<i>arithmetic</i>

		n	-	<i>ethnic</i>
		l	-	<i>athlete</i>
/ð/	+	m	-	<i>rhythmic</i>
/ʃ/	+	m	<i>leishmaniose</i>	<i>marshmallow</i>
		l	-	<i>ashlar</i>
/ʒ/	+	m	-	<i>cashmere</i>
/tʃ/	+	p	-	<i>hotchpotch</i>
		b	-	<i>archbishop</i>
		d	-	<i>archduke</i>
		k	-	<i>archconfraternity</i>
		f	-	<i>archfiend</i>
		s	-	<i>archsee</i>
/dʒ/	+	p	-	<i>hodgepodge</i>
		l	-	<i>fledgling</i>

ANEXO V

LISTA DE FRASES EM INGLÊS

Por favor, proceda à leitura dos seguintes parágrafos:

1. An archconfraternity met in a subzone of Nagpur last week. The quantity of archpriests and archbishops outnumbered the allowed capacity of the archsee, resulting in a hodgepodge of ideas and opinions. Some people thought that the archfiend Satan was taking control of them. When their subchief realized that situation, he suggested they were subdivided in groups of fifty, according to their options and subjects of interest. They subconsciously said ‘dodgasted!’ to themselves and started to move in a rhythmic and strange way. When the hotchpotch was finally over, peace reigned among them and enmity not even had the chance to rise.

2. A group of postgraduate students decided to meet up some years after their graduation, so they set up a rugby match to celebrate the occasion. The ones who had become athletes at the university managed to outrun their rivals and outpassed them, winning the game. After the match, they decided to light a fire by the woods to cook and eat marshmallows. One of the boys went into the forest with an adze in hand to cut some wood for the fire and then he saw a nest with two fledglings waiting for their mother bird, and a chipmunk eating nuts near there. He first wanted to capture the animals, but then he decided just to take pictures of them. By the end of the day they were all outbreathing loud laughs of happiness and their smiles outshone the fire light.

3. Yesterday I went to a cool restaurant built with ashlar stones and saw a Latvian couple eating a soup with sabzi vegetables. The girl had a band-aid on her nose and was wearing an athbash pendant and a luxurious scarf made of cashmere. She was telling her boyfriend that the postsurgical pain she felt was terrible and started to cry. The man dried her eyes with a napkin and then said he would have to postdate their wedding because he had got a sublicense from his boss to go to an auction to buy some naphtha and hafnium and some flotsam objects for the chemical industry where he worked. She suddenly passed out and hit her head on the floor, getting an eczema and amnesia with the accident. The desperate boy could not outthink his actions and killed himself. Now he still visits his forgetful girl in his postmortal life.

4. After the class of arithmetic, the teacher of English came into class and decided to give the students a posttest, since they did not get good grades in the first test. There were lots of difficult questions, and the students could not notice the difference between the words ‘postvocalic’, ‘outjockey’, and ‘postrecord’ and the words ‘portfolio’, ‘ethnic’, and ‘rhabdom’. The teacher had to explain that none of these words had postfixes, but the first three of them had prefixes. He also explained to the group that the word ‘outjockey’ had two diphthongs.

ANEXO VI

TERMO DE CONSENTIMENTO

Prezado aluno:

Através da assinatura do presente Termo de Consentimento, você estará contribuindo para o desenvolvimento de um projeto de mestrado vinculado ao Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul - Área de Concentração Estudos da Linguagem - que tem por objetivo analisar aspectos linguísticos referentes à língua portuguesa e à língua inglesa falada por aprendizes brasileiros.

Sua tarefa como colaborador desta pesquisa consiste em cumprir as etapas abaixo descritas:

- a) preenchimento e assinatura da Declaração deste Termo de Consentimento;
- b) preenchimento do formulário sobre informações pessoais (anexo);
- c) pronúncia de uma sequência de frases em português e em inglês.

Observações:

- a) Todas as informações pessoais serão mantidas sob sigilo absoluto e serão utilizadas apenas pelo pesquisador, como contribuição para o desenvolvimento de seu estudo;
- b) A leitura da sequência de frases será gravada por equipamento de áudio.

DECLARAÇÃO

Eu, _____, declaro que estou de acordo com as informações deste Termo de Consentimento e que consinto participar como colaborador da pesquisa acima referida.

Assinatura

Em ____ / ____ / ____.

ANEXO VII

FORMULÁRIO SOBRE INFORMAÇÕES PESSOAIS DO ALUNO
COLABORADOR

1. Nome completo: _____

2. E-mail: _____

3. Sexo: () masculino () feminino

4. Idade: _____

5. É falante nativo de português? () Sim. () Não.

6. Em que município residiu até os 10 anos de idade? Caso haja mais de um, discrimine os nomes e informe quanto tempo viveu em cada local. (Ex.: Itu-SP - 0 aos 6; Canoas - 6 aos 10).

7. Atualmente reside em qual município? _____

8. Há quantos anos reside nesse município? _____

9. Qual é seu nível de escolaridade? _____

10. É (ou foi) estudante de graduação em Letras? () Sim. () Não.

11. Por quantos anos teve aula de inglês na educação básica? _____

12. Há quantos anos estuda inglês nesta escola de idiomas? _____

13. Por quanto tempo estudou em outras escolas de idiomas? _____

14. Já morou em país de língua inglesa? Em caso afirmativo, relacione o(s) país(es) e o tempo de permanência.

15. Já teve de usar o inglês para se comunicar com alguém verbalmente? Relate.

16. Você fala ou tem contato com pessoas que falam outros idiomas? Quais? Relate.

ANEXO VIII – Transcrição fonética e significado dos vocábulos analisados na Amostra de Fala 2

Vocábulo em inglês	Transcrição fonética*	Significado**	Vocábulo em inglês	Transcrição fonética*	Significado**
<i>adze</i>	[ˈædz]	enxó	<i>napkin</i>	[ˈnæpkɪn]	guardanapo
<i>amnesia</i>	[æmˈni:zə]	amnésia	<i>object</i>	[ˈɒbdʒɪkt]	objeto
<i>archbishop</i>	[ɑ:rtʃˈbɪʃəp]	arcebispo	<i>option</i>	[ˈɒpʃən]	opção
<i>archconfraternity</i>	[ɑ:rtʃ kənfrəˈtɜ:nəti]	um tipo de irmandade	<i>outbreathe</i>	[aʊtˈbri:ð]	exalar
<i>archfiend</i>	[ɑ:rtʃˈfi:nd]	arqui-inimigo	<i>outjockey</i>	[aʊtˈdʒɔ:kɪ]	obter vantagem
<i>archpriest</i>	[ɑ:rtʃˈpri:st]	arcipreste	<i>outnumber</i>	[aʊtˈnʌmbər]	exceder em número
<i>archsee</i>	[ɑ:rtʃˈsi:]	arquidiocese	<i>outpass</i>	[aʊtˈpæs]	ir além de
<i>arithmetic</i>	[əˈrɪθmətɪk]	aritmética	<i>outrun</i>	[aʊtˈrʌn]	correr mais que
<i>ashlar</i>	[ˈæʃlə]	pedra silhar	<i>outshine</i>	[aʊtˈʃaɪn]	brilhar mais que
<i>athbash</i>	[æθˈbæʃ]	criptografia do hebraico	<i>outthink</i>	[aʊtˈθɪŋk]	pensar mais que
<i>athlete</i>	[ˈæθli:t]	atleta	<i>picture</i>	[ˈpɪktʃər]	fotografia
<i>auction</i>	[ˈɔ:kʃən]	leilão	<i>portfolio</i>	[pɔ:tfəʊliəʊ]	portfólio
<i>capture</i>	[ˈkæptʃər]	capturar	<i>postdate</i>	[pəʊstˈdeɪt]	pós-datar
<i>cashmere</i>	[ˈkæʒmɪər]	caxemira	<i>postfix</i>	[ˈpəʊstfɪks]	posfixo
<i>chipmunk</i>	[ˈtʃɪpmʌŋk]	tâmia	<i>postgraduate</i>	[pəʊstˈgrædʒuət]	pós-graduando
<i>diphthong</i>	[ˈdɪfθɔ:ŋ]	ditongo	<i>postmortal</i>	[pəʊstˈmɔ:rtəl]	pós-morte
<i>dodgasted</i>	[ˈdɒdgæstɪd]	eufemismo de <i>God blasted</i> (maldito Deus)	<i>postrecord</i>	[pəʊstrɪˈkɔ:rd]	pós-gravar
<i>eczema</i>	[ˈeksɪmə]	eczema	<i>postsurgical</i>	[pəʊstˈsɜ:rdʒɪkəl]	pós-operatório
<i>enmity</i>	[ˈenmɪti]	inimizade	<i>posttest</i>	[ˈpəʊsttest]	pós-teste
<i>ethnic</i>	[ˈeθnɪk]	étnico	<i>postvocalic</i>	[pəʊstvəˈkælɪk]	pós-vocálico
<i>fifty</i>	[ˈfɪftɪ]	cinquenta	<i>rhabdom</i>	[ˈræbdəm]	parte de um olho composto
<i>fledgling</i>	[ˈfledʒlɪŋ]	filhote de ave	<i>rhythmic</i>	[ˈrɪðmɪk]	rítmico
<i>flotsam</i>	[ˈflɔ:tsəm]	fragmentos de naufrágio	<i>rugby</i>	[ˈrʌgbɪ]	rúgbi
<i>hafnium</i>	[ˈhæfnɪəm]	háfnio	<i>sabzi</i>	[ˈsæbzɪ]	tipo de legume
<i>hodgepodge</i>	[ˈhɔ:dʒpɔ:dʒ]	confusão	<i>subchief</i>	[ˈsʌbtʃi:f]	subchefe
<i>hotchpotch</i>	[ˈhɔ:tʃpɔ:tʃ]	confusão	<i>subconscious</i>	[sʌbˈkɔ:nʃəs]	subconsciente
<i>Latvian</i>	[ˈlætvi:ən]	letão	<i>subdivide</i>	[sʌbdɪˈvaɪd]	subdividir
<i>luxurious</i>	[lʌˈdʒʊəriəs]	luxuoso	<i>sublicense</i>	[sʌbˈlaɪsənts]	sublicença
<i>marshmallow</i>	[ˈmɑ:ʀ[mæləʊ]	<i>marshmallow</i>	<i>subzone</i>	[ˈsʌbzəʊn]	sub-região
<i>Nagpur</i>	[ˈnɑɡpʊr]	Nagpur (cidade)	<i>suggest</i>	[səˈdʒest]	sugerir
<i>naphtha</i>	[ˈnæpθə]	nafta	-	-	-

* A transcrição fonética baseou-se na pronúncia norte-americana e foi feita com o auxílio do International Phonetic Alphabet e da versão eletrônica do *Cambridge Advanced Learner's Dictionary*. O Dicionário está disponível em <http://dictionary.cambridge.org/>.

** As traduções e as explicações do significado das palavras (nos casos em que não há tradução direta do inglês para o português) foram feitas com o auxílio do *Dicionário Eletrônico Michaelis v. 5.0*.

