

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE COMUNICAÇÃO E EXPRESSÃO  
PÓS-GRADUAÇÃO EM LINGÜÍSTICA

*Estudo do VOT no Português  
Brasileiro*

Simone Klein

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Cláudia Borges De Faveri

Dissertação apresentada ao Curso de Pós-Graduação em Língua Portuguesa da Universidade Federal de Santa Catarina, como requisito parcial para a obtenção do título de Mestre em Língua Portuguesa.

Florianópolis - SC  
1999

## Agradecimentos

A meus pais, Vit3lio e Medy; 3s “meninas”, Monica, Cris e Lilian; 3 Sarah e ao Victor.

3 Cl3udia, pela orienta33o.

3 Sandra, pela amizade, for3a, companheirismo, hist3rias e “tratamentos de choque”... durante todos os anos de “conv3ncia”.

3 Izabel, a quem nem todas as palavras de agradecimento do mundo fariam justi3a!  
Valeu Iza! O trabalho est3 3!

3 Tati, pelas noites e noites afora de discuss3es acad3micas e existenciais.

3 Sanir.

3s amigas Ina e Sandra Mara, pela amizade, for3a, caf3s, etc...

Ao Ronaldo.

Aos amigos Du, Iza, Ronaldo e Sandra, pela presteza e paci3ncia com que “suportaram” a 3rdua tarefa de servirem como informantes para o presente estudo.

Ao Prof. Ap3stolo, pela disponibilidade e bom humor com que sempre atendeu 3s minhas d3vidas referentes ao ingl3s.

3 Claudia, Julio e Cris, pelos empurr3es, tanto na gradua33o quanto na p3s.

3 C3lia, por tudo!

Ao CNPq e 3 CAPES, pelo suporte financeiro.

*Uma voz nunca se torna parte de você como um ruído.  
O mar, você acaba não ouvindo mais.  
Uma voz não, uma voz é impossível não escutar.\**

---

\* “Além das nuvens”, filme de Michelangelo Antonioni e Wim Wenders.

# Sumário

Lista de figuras	viii
Lista de gráficos	ix
Lista de tabelas	x
Resumo .....	xiv
Abstract .....	xv
Introdução .....	1

## Capítulo 1: Referencial teórico

1.1. Oclusivas .....	6
1.2. <i>Voice onset time</i> .....	10
1.3. Revisão da literatura .....	16
1.3.1. O VOT na produção da fala .....	17
1.3.2. O VOT na percepção da fala .....	31

## Capítulo 2: Metodologia

2.1. Seleção de informantes .....	39
2.2. <i>Corpus</i> .....	40
2.3. Coleta de dados .....	42
2.4. Registro e análise dos dados .....	43
2.5. Parâmetros estatísticos utilizados na análise .....	44

## Capítulo 3: Análise e discussão dos resultados

3.1. Introdução .....	48
3.2. Pré-sonorização .....	49

3.2.1. Oclusivas bilabiais .....	52
3.2.2. Oclusivas dentais .....	53
3.2.3. Oclusivas velares .....	53
3.2.4. O VOT e o contexto vocálico .....	54
3.2.5. O VOT e a tonicidade da sílaba .....	56
3.2.6. O VOT em diferentes posições na sentença .....	61
3.3. Retardo curto .....	63
3.3.1. Oclusivas bilabiais .....	65
3.3.2. Oclusivas dentais .....	66
3.3.3. Oclusivas velares .....	67
3.3.4. O VOT e o contexto vocálico .....	67
3.3.5. O VOT e a tonicidade da sílaba .....	70
3.3.6. O VOT em diferentes posições na sentença .....	74
3.4. Considerações finais .....	75
Conclusão .....	87
Referencial bibliográfico .....	90
Bibliografia consultada .....	94
Apêndice 1 .....	97
Apêndice 2 .....	100

## Lista de Figuras

Figura 1.1 - Formas de onda e espectrogramas da área de retardo longo da oclusiva bilabial aspirada [p <sup>h</sup> ] (quadros A e D), da área de retardo curto da oclusiva bilabial [p] (quadros B e E) e da área de pré-sonorização da oclusiva bilabial sonora [b] (quadros C e F) .....	11
Figura 2.1 - Formas de onda e espectrogramas da área de pré-sonorização e retardo curto das consoantes [d] e [t] nos segmentos [do] e [to] .....	44
Figura 3.2.1 - Forma de onda e espectrograma da área de pré-sonorização de [g] no segmento [ga] .....	49
Figura 3.3.1 - Forma de onda e espectrograma da área de retardo curto da consoante oclusiva [k] no segmento [ka] .....	63
Figura 3.4.1 - Formas de onda e espectrogramas da área de retardo curto (quadros A e B), da “tendência ao retardo longo” (quadros C e D) e da área de retardo longo (quadros E e F) da oclusiva bilabial surda do PB .....	84
Figura 3.4.2 - Formas de onda e espectrogramas da área de retardo curto (quadros A e B), e de retardo longo (quadros C, D e E, F) da oclusiva velar surda do PB .....	85

## Lista de Gráficos

Gráfico 3.2.1 – Comparação da área de pré-sonorização do PB em Istre em e nossos dados .....	50
Gráfico 3.2.2 – Comparação da área de pré-sonorização do PB e em outras línguas (Lisker e Abramson, 1964) .....	51
Gráfico 3.3.1 – Comparação da área de retardo curto do PB em Istre e em nossos dados .....	64
Gráfico 3.3.2 – Comparação da área de retardo curto do PB e em outras línguas (Lisker e Abramson, 1964) .....	65
Gráfico 3.4.1 – Distribuição do VOT das oclusivas bilabiais do PB em logatomas produzidos isoladamente .....	77
Gráfico 3.4.2 – Distribuição do VOT das oclusivas bilabiais do PB em logatomas inseridos em frases-veículo .....	77
Gráfico 3.4.3 – Distribuição do VOT das oclusivas dentais do PB em logatomas produzidos isoladamente .....	79
Gráfico 3.4.4 – Distribuição do VOT das oclusivas dentais do PB em logatomas inseridos em frases-veículo .....	79
Gráfico 3.4.5 – Distribuição do VOT das oclusivas velares do PB em logatomas produzidos isoladamente .....	81
Gráfico 3.4.6 – Distribuição do VOT das oclusivas velares do PB em logatomas inseridos em frases-veículo .....	81

## Lista de Tabelas

Tabela 3.2.1 – Média geral de VOT e área observada em ms para a oclusiva [b] nos logatomas produzidos isoladamente e nos logatomas inseridos em frases-veículo .....	52
Tabela 3.2.2 – Média geral de VOT e área observada em ms para a oclusiva [d] nos logatomas produzidos isoladamente e nos logatomas inseridos em frases-veículo .....	53
Tabela 3.2.3 – Média geral de VOT e área observada em ms para a oclusiva [g] nos logatomas produzidos isoladamente e nos logatomas inseridos em frases-veículo .....	54
Tabela 3.2.4 – Média de VOT e área observada em ms para a oclusiva [b] diante das vogais [a], [e] e [o], para os logatomas produzidos isoladamente .....	54
Tabela 3.2.5 – Média de VOT e área observada em ms para a oclusiva [b] diante das vogais [a], [e] e [o], para os logatomas inseridos em frases-veículo .....	54
Tabela 3.2.6 – Média de VOT e área observada em ms para a oclusiva [d] diante das vogais [a], [e] e [o], para os logatomas produzidos isoladamente .....	55
Tabela 3.2.7 – Média de VOT e área observada em ms para a oclusiva [d] diante das vogais [a], [e] e [o], para os logatomas inseridos em frases-veículo .....	55
Tabela 3.2.8 – Média de VOT e área observada em ms para a oclusiva [g] diante das vogais [a], [e] e [o], para os logatomas produzidos isoladamente .....	55
Tabela 3.2.9 – Média de VOT e área observada em ms para a oclusiva [g] diante das vogais [a], [e] e [o], para os logatomas inseridos em frases-veículo .....	56
Tabela 3.2.10 – Média geral de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico .....	57
Tabela 3.2.11 – Média geral de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico .....	57
Tabela 3.2.12 – Média geral de VOT e área observada em ms para a oclusiva [g] nos contextos pretônico, tônico e postônico .....	58
Tabela 3.2.13 – Média de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico, diante da vogal [a] .....	59



Tabela 3.2.14 – Média de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico, diante da vogal [e] .....	59
Tabela 3.2.15 – Média de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico, diante da vogal [o] .....	59
Tabela 3.2.16 – Média de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico, diante da vogal [a] .....	60
Tabela 3.2.17 – Média de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico, diante da vogal [e] .....	60
Tabela 3.2.18 – Média de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico, diante da vogal [o] .....	60
Tabela 3.2.19 – Média de VOT e área observada em ms para a oclusiva [g] nos contextos pretônico, tônico e postônico, diante da vogal [a] .....	61
Tabela 3.2.20 – Média de VOT e área observada em ms para a oclusiva [g] nos contextos pretônico, tônico e postônico, diante da vogal [e] .....	61
Tabela 3.2.21 – Média de VOT e área observada em ms para a oclusiva [g] nos contextos pretônico, tônico e postônico, diante da vogal [o] .....	61
Tabela 3.2.22 – Média de VOT e área observada em ms para a oclusiva [b] em diferentes posições em sentenças .....	62
Tabela 3.2.23 – Média de VOT e área observada em ms para a oclusiva [d] em diferentes posições em sentenças .....	62
Tabela 3.2.24 – Média de VOT e área observada em ms para a oclusiva [g] em diferentes posições em sentenças .....	62
Tabela 3.3.1 – Média geral de VOT e área observada em ms para a oclusiva [p] inserida em logatomas produzidos isoladamente e inseridos em frases-veículo .....	66
Tabela 3.3.2 – Média geral de VOT e área observada em ms para a oclusiva [t] inserida em logatomas produzidos isoladamente e inseridos em frases-veículo .....	66
Tabela 3.3.3 – Média geral de VOT e área observada em ms para a oclusiva [k] inserida em logatomas produzidos isoladamente e inseridos em frases-veículo .....	67

Tabela 3.3.4 – Média de VOT e área observada em ms para a oclusiva [p] diante das vogais [a], [e] e [o], para logatomas produzidos isoladamente .....	68
Tabela 3.3.5 – Média de VOT e área observada em ms para a oclusiva [p] diante das vogais [a], [e] e [o], para logatomas inseridos em frases-veículo .....	68
Tabela 3.3.6 – Média de VOT e área observada em ms para a oclusiva [t] diante das vogais [a], [e] e [o], para logatomas produzidos isoladamente .....	68
Tabela 3.3.7 – Média de VOT e área observada em ms para a oclusiva [t] diante das vogais [a], [e] e [o], para logatomas inseridos em frases-veículo .....	69
Tabela 3.3.8 – Média de VOT e área observada em ms para a oclusiva [k] diante das vogais [a], [e] e [o], para logatomas produzidos isoladamente .....	69
Tabela 3.3.9 – Média de VOT e área observada em ms para a oclusiva [k] diante das vogais [a], [e] e [o], para logatomas inseridos em frases-veículo .....	69
Tabela 3.3.10 – Média geral de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico .....	70
Tabela 3.3.11 – Média geral de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico .....	70
Tabela 3.3.12 – Média geral de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico .....	70
Tabela 3.3.13 – Média de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico, diante da vogal [a] .....	72
Tabela 3.3.14 – Média de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico, diante da vogal [e] .....	72
Tabela 3.3.15 – Média de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico, diante da vogal [o] .....	72
Tabela 3.3.16 – Média de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico, diante da vogal [a] .....	72
Tabela 3.3.17 – Média de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico, diante da vogal [e] .....	73

Tabela 3.3.18 – Média de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico, diante da vogal [o] .....	73
Tabela 3.3.19 – Média de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico, diante da vogal [a] .....	73
Tabela 3.3.20 – Média de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico, diante da vogal [e] .....	73
Tabela 3.3.21 – Média de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico, diante da vogal [o] .....	74
Tabela 3.3.22 – Média de VOT e área observada em ms para a oclusiva [p] em diferentes posições em sentenças .....	74
Tabela 3.3.23 – Média de VOT e área observada em ms para a oclusiva [t] em diferentes posições em sentenças .....	75
Tabela 3.3.24 – Média de VOT e área observada em ms para a oclusiva [k] em diferentes posições em sentenças .....	75

## Resumo

Há diversas pistas que podem ser utilizadas na distinção entre segmentos sonoros e surdos de uma língua. No caso das consoantes oclusivas, um dos parâmetros mais eficazes é o *Voice onset time (VOT)* - a relação temporal entre a soltura da oclusão e o início da excitação glotal. O VOT apresenta três categorias que irão caracterizá-lo, a saber, pré-sonorização, retardo curto e retardo longo, segundo o início da sonorização seja anterior à soltura da oclusão; simultânea a ou imediatamente posterior à soltura (10 ms a 35 ms) ou, posterior à soltura em cerca de 35 ms a 100 ms, respectivamente. Todas as línguas que fazem o contraste sonoro *vs* surdo irão apresentar ou as três categorias, ou apenas duas. O português brasileiro, que distingue entre oclusivas sonoras e surdas, apresenta as categorias de pré-sonorização e retardo curto. Nesta pesquisa, analisamos como se constituem essas duas áreas, bem como a influência de determinados contextos sobre o VOT, tais como contexto vocálico, contexto de tonicidade e diferentes posições em sentenças. Os resultados revelaram que o português brasileiro caracteriza-se por apresentar precedência longa de sonorização para as oclusivas sonoras, e retardo curto para as oclusivas surdas, embora as últimas tenham apresentado valores encaminhando-se para a área de retardo longo, ou seja, apresentando certa aspiração. Os resultados revelaram ainda que o VOT é sensível aos contextos vocálico, de tonicidade e a diferentes posições em sentenças.

## Abstract

There are several acoustic cues which serve to make the distinction between voiced and voiceless segments in a language. In the case of stop consonants, one of the most efficient parameters is the **voice onset time (VOT)** – the timing relation between voice onset and the release of the occlusion. VOT is defined in terms of three categories: voicing lead – glottal pulsing begins prior to release of occlusion; short voicing lag – onset of vocal folds vibration begins simultaneously with or slightly after (10 – 35 ms) release of occlusion; and long voicing lag – a longer interval ( 35 – 100 ms) exists between voicing onset and release of occlusion. Those languages which have voicing contrast have either two or three voicing categories. Brazilian Portuguese, which distinguishes between voiced and voiceless stops, has voicing lead and short lag categories. This study investigates how these two categories are organized and what the influence of determined context on VOT, such as vowel context, stressed context and different positions in sentences. The results have shown that Brazilian Portuguese has long voicing lead and short voicing lag for voiced and voiceless stops, respectively, although the latter have values shift toward long lag range, that is, they have slight aspiration. The results have also shown that VOT is sensitive to vowel and stressed contexts and different positions in sentences.

# Introdução

Investigações das propriedades acústicas que correspondem às dimensões fonéticas da fala têm sugerido que para qualquer categoria fonética dada há sempre uma propriedade ou grupo de parâmetros acústicos que definem aquela categoria fonética (Stevens e Blumstein, 1981; Stevens, 1989, *apud* Kessinger e Blumstein, 1997), e parece haver também, para cada uma dessas categorias, uma área de valores que irá caracterizar a propriedade acústica em questão. Por exemplo, uma das propriedades acústicas responsáveis pelo aspecto sonoro é a presença de energia espectral de baixa frequência ou periodicidade que se deve à vibração das pregas vocais. A ausência dessa energia caracteriza os segmentos surdos.

Todavia, esta não é a única pista acústica que pode contribuir para o contraste de sonorização dos segmentos de determinada língua. Há várias outras, como por exemplo:

- frequência do primeiro formante no início da sonorização;
- natureza das mudanças na frequência fundamental ( $F_0$ ) que seguem imediatamente a soltura da consoante;
- frequência da explosão, segundo formante e formantes mais altos;
- intensidade do ruído de aspiração com relação à vogal;

- duração de uma vogal precedente.

Dentre as línguas do mundo, a maioria - 88,9% - distingue entre segmentos sonoros e não sonoros (ou surdos) (cf. Kessinger e Blumstein, 1997). Em português, por exemplo, as consoantes oclusivas /b/, /d/, /g/ - sonoras - opõem-se às consoantes oclusivas /p/, /t/, /k/ - surdas.

Uma das pistas acústicas associadas à distinção sonoro *vs.* surdo de consoantes oclusivas homorgânicas é o *Voice onset time (VOT)* - relação temporal entre o início de sonorização e a soltura da oclusão (cf. Lisker e Abramson, 1964) - sendo um parâmetro usado exclusivamente no caso de consoantes oclusivas, haja vista serem estas as únicas consoantes a realizar uma oclusão total no trato vocal.

O VOT tem sido estudado em inúmeras línguas e tem se mostrado um parâmetro bastante eficiente em áreas de pesquisa relacionadas à:

- *patologia*, na investigação de distúrbios da fala resultantes de danos neurológicos, tal como afasia, por comparar o VOT de sujeitos normais com o de sujeitos afásicos;
- *aquisição da linguagem*, comparando o desenvolvimento de crianças em processo normal de aquisição com crianças que apresentam problemas nesse processo;
- *dialetologia*, mais especificamente o sotaque, na tentativa de explicar o porquê de, muitas vezes, falantes, principalmente falantes bilíngües, pronunciarem oclusivas sonoras como surdas e vice-versa;
- *identificação de falantes*, na possibilidade de utilização deste parâmetro como pista da identidade do falante.

De acordo com Lisker e Abramson (1964), o VOT caracteriza-se por apresentar três áreas de valores que irão especificá-lo:

- *pré-sonorização*: o início da sonorização é anterior à soltura da oclusão (VOT negativo);

- *retardo curto*: o início da sonorização e a soltura da oclusão ou são simultâneos (VOT zero), ou a sonorização segue imediatamente a soltura da oclusão (10 a 35 ms);
- *retardo longo*: o início da sonorização começa cerca de 35 ms a 100 ms após a soltura da oclusão.

Segundo Istre (s/d), o português brasileiro, que apresenta um contraste entre oclusivas sonoras e surdas, caracteriza-se por apresentar precedência média de sonorização e retardo curto.

Nosso interesse, neste trabalho, centra-se, então, no estudo do VOT no português brasileiro, tendo em vista sua importância como parâmetro acústico para a distinção de sonorização de oclusivas e os poucos trabalhos existentes para essa língua.

O objetivo geral desta pesquisa será, portanto, estudar o VOT das consoantes oclusivas sonoras e surdas do português brasileiro.

Como objetivos específicos, pretende-se:

- a) verificar como se constituem as áreas de pré-sonorização e retardo curto;
- b) observar o comportamento do VOT entre oclusivas inseridas em logatomas produzidos isoladamente e em logatomas inseridos em contextos maiores;
- c) observar a influência do contexto vocálico sobre o VOT;
- d) observar a influência da tonicidade sobre o VOT;
- e) verificar se diferentes posições em sentenças provocarão mudanças significativas nos valores de VOT das oclusivas no português brasileiro.

Tendo como suporte teórico as leituras realizadas para a elaboração desta pesquisa (cf. Capítulo 1 – Referencial teórico), bem como alguns experimentos realizados previamente, levantamos as seguintes hipóteses:



- a) as oclusivas sonoras do português, ao contrário do que mostraram os resultados obtidos por Istre (s/d), caracterizar-se-iam por uma precedência longa de sonorização, e não por uma precedência média;
- b) as oclusivas surdas do português apresentariam certa aspiração;
- c) o VOT das oclusivas inseridas em logatomas do tipo CV produzidos isoladamente apresentariam valores médios maiores do que as oclusivas dos mesmos logatomas inseridos em frases-veículo;
- d) o VOT seria influenciado em função do contexto vocálico;
- e) o VOT de oclusivas em contexto tônico apresentaria valores médios maiores do que oclusivas em contexto átono;
- f) oclusivas inseridas em palavras em posição final de sentença apresentariam valores de VOT maiores do que oclusivas em palavras em posição inicial ou medial de sentença.

Utilizando fundamentos estatísticos de tomada de decisão, como o teste *t-de-Student*, verificaremos se a variação dos valores de VOT são relevantes nos diferentes contextos vocálicos, de tonicidade e em diferentes posições em sentenças.

Quanto à organização da dissertação, no primeiro capítulo serão apresentados diversos estudos sobre o parâmetro VOT e sua aplicação nas mais diversas áreas, tanto em produção quanto em percepção da fala. Neste capítulo também será revisado o trabalho de Lisker e Abramson (1964), pioneiro no tratamento desse parâmetro acústico, apresentando as diferenças concernentes às diferentes línguas do mundo.

No segundo capítulo, serão descritos os métodos e técnicas empregados na obtenção do material sonoro utilizado nas análises, bem como os procedimentos adotados no processo de registro dos dados, juntamente com a apresentação de todo aparato estatístico utilizado para a validação dos resultados alcançados.

No terceiro capítulo, serão apresentadas as análises e a discussão dos dados obtidos a partir das análises acústicas para as duas categorias de VOT, a pré-sonorização e o retardo curto, das oclusivas bilabiais, dentais e velares,

observando também, com a ajuda dos testes estatísticos, a influência que o contexto vocálico, tonicidade da sílaba e posição da palavra na sentença exercem sobre o VOT.

Por fim, o quarto e último capítulo, apresentará as conclusões finais desta pesquisa e sugestões para trabalhos futuros.

# *Referencial Teórico*

## **1.1 Oclusivas**

As consoantes oclusivas, também chamadas plosivas, são, do ponto de vista articulatorio, aqueles sons da fala caracterizados por uma obstrução total e momentânea da corrente de ar iniciada nos pulmões, em algum ponto da cavidade oral. Tanto sons nasais quanto sons orais podem ser classificados como oclusivos, embora o termo seja usualmente reservado para os sons orais. A obstrução pode, em português, ocorrer em três áreas - lábios, dentes e véu palatino - resultando, respectivamente, em oclusivas bilabiais, dentais e velares, a saber, /p, b/, /t, d/ e /k, g/<sup>1</sup>. Entretanto, diferentes línguas fazem uso de uma variedade maior de oclusivas, utilizando diferentes pontos de articulação, mecanismos de produção de soltura e ruídos que as acompanham (Johnson, 1997). Há línguas, por exemplo, que apresentam um tipo de oclusão não muito usual - a oclusão glotal - produto da obstrução do fluxo de ar na glote, como o próprio nome diz. Segundo Cagliari (1981),

---

<sup>1</sup> À série de oclusivas velares, Mattoso Câmara Jr. dá o nome de “oclusivas velopalatais”, pois se articulam no véu palatino diante de /a/ ou de vogal posterior, ou, no extremo final do palato duro ou pós-palato, diante de vogal anterior.

em português brasileiro (doravante PB), um exemplo de oclusiva glotal pode ser encontrado na expressão 'aha', pronunciada com uma pequena pausa entre os dois 'as'. Essa pequena pausa é a oclusiva glotal.

Da mesma forma com que toda consoante oclusiva envolve uma fase na qual o fluxo de ar é completamente obstruído em qualquer ponto entre a glote e os lábios, *fase de obstrução ou intervalo de fechamento*, essa consoante envolve também uma fase antecedente e uma subsequente a essa obstrução. Na fase antecedente, *fase de aproximação ou de fechamento*, os articuladores entram em contato, provocando assim uma parada no fluxo do ar. Na fase subsequente, *fase de soltura ou explosão*, o ar encerrado pela obstrução escapa. Esse ar é empurrado para fora da boca em uma velocidade bastante rápida e produz um impulso de pressão que dura somente 2 ms ou 3 ms.

No entanto, na articulação de uma oclusiva, tanto a *fase de aproximação ou fechamento* quanto a *fase de soltura* podem estar ausentes ou, potencialmente ausentes. Por exemplo, quando se pronuncia a palavra *pá*, começando a fazê-lo de uma posição de repouso, com os lábios fechados, não há uma aproximação observável dos articuladores, embora alguns eventos preparatórios estejam, provavelmente, tomando lugar. Eventos tais como o levantamento do véu palatino para fechar a entrada para a cavidade nasal e alguma tensão dos músculos dos lábios. Quanto à soltura da oclusão, em uma língua como o inglês americano por exemplo, em que ocorrem palavras como *stop* e *cat*, as oclusivas /p/ e /t/ finais podem realizar-se sem uma soltura audível do fechamento. Dessa forma, a única das três fases de uma oclusiva que é absolutamente essencial é, de fato, o *intervalo de fechamento ou oclusão*, que deve ter sempre uma duração perceptível (Cf. Asher, 1994).

Do ponto de vista acústico, a oclusiva é caracterizada, em seu período de oclusão, por um breve silêncio, embora alguma energia de sonorização no caso de oclusivas sonoras possa se estender por parte ou todo o período. Quando essa sonorização está presente, é manifestada por energia de baixa frequência nos harmônicos mais baixos, especialmente o primeiro harmônico ou frequência fundamental.

No que tange à soltura, esta é caracterizada acusticamente por um segmento de ruído muito breve - a *explosão*, que é um dos eventos mais curtos, se não o mais curto, comumente analisado em fala. Se a oclusiva for seguida por um som vocálico, a explosão será seguida por outro intervalo acústico - a *transição*. No caso de oclusiva sonora, esse intervalo de transição é caracterizado por uma mudança rápida no padrão de formantes. O espectro da explosão da oclusiva irá variar de acordo com o ponto de articulação da mesma. Essa variação é atribuída a propriedades de ressonância do trato vocal. Dessa forma, conforme Kent e Read (1992), as oclusivas bilabiais tendem a ser caracterizadas por uma explosão em faixas de baixa frequência, as alveolares em faixas de alta frequência e as velares, por uma explosão em faixas de média frequência. Segundo Johnson (1997), no caso das bilabiais, essa predominância de energia de explosão em faixas de baixa frequência pode ser dada em função da ausência de uma cavidade de ressonância em frente aos lábios, de forma que o espectro dessa consoante seja determinado pelas propriedades acústicas da fonte sonora. Já as oclusivas dentais e alveolares possuem uma cavidade anterior pequena, conseqüentemente, picos de energia espectral de alta-frequência. Oclusivas velares têm uma cavidade anterior maior, por conseguinte, picos de frequência mais baixa no espectro e, geralmente, sua estrutura aproxima-se mais da estrutura de formantes. De acordo com pesquisas realizadas por Halle, Hughes e Radley (1957), as bilabiais /p/ e /b/ apresentam maior concentração de energia de explosão na faixa entre 500 e 1500 Hz, as alveolares /t/ e /d/ acima de 4000 Hz e as velares /k/ e /g/, entre 1500 e 4000 Hz. Ainda, com relação à soltura, as oclusivas são classificadas segundo sejam aspiradas ou não aspiradas. Uma consoante aspirada é caracterizada por um ruído surdo, ou sopro, gerado à medida que o ar passa através das pregas vocais parcialmente fechadas e faringe, entre a explosão da consoante e a vogal seguinte, principalmente se a vogal for acentuada. Na articulação das oclusivas aspiradas, a glote encontra-se em sua posição aberta e, antes que as pregas vocais entrem em vibração para a vogal seguinte através de seu acoplamento, passa-se algum tempo, durante o qual o ar flui, produzindo o ruído do sopro. No caso de oclusiva não aspirada, durante a oclusão, a glote encontra-se em sua posição fechada.

As pregas vocais, portanto, podem entrar em vibração no momento da explosão, pois já encontram-se na posição exigida para a fonação, e a vogal começa imediatamente após a explosão (cf. Malmberg, 1954).

No que diz respeito à sonorização, as oclusivas são classificadas em sonoras e surdas. No PB, /b, d, g/ são segmentos oclusivos sonoros, enquanto /p, t, k/ são surdos. Como já mencionado anteriormente, a propriedade acústica responsável pelo aspecto sonoro é a presença de energia espectral de baixa frequência ou periodicidade que se deve à vibração das pregas vocais. A ausência, pois, dessa energia, caracteriza os segmentos surdos. De acordo com Stevens e Blumstein (1981 *apud* Lieberman e Blumstein, 1996):

*The presence of this low frequency periodicity occurs for consonant sounds over a time interval of 20-30 milliseconds in the vicinity of the acoustic discontinuity that precedes or follows the consonant constriction interval (Lieberman e Blumstein, 1996: 195).*

Contudo, há várias outras pistas acústicas que contribuem para o contraste sonoro *vs.* surdo dos segmentos de determinada língua, a saber, frequência do primeiro formante no início da sonorização; natureza das mudanças na frequência fundamental ( $F_0$ ) que seguem imediatamente a soltura da consoante; frequência da explosão, segundo formante e formantes mais altos; intensidade do ruído de aspiração com relação à vogal; duração de uma vogal precedente.

Há que se ressaltar, no entanto, que estas múltiplas pistas acústicas que especificam o aspecto fonético da sonorização dependem das manobras articulatórias do trato vocal que iniciam o processo. Desse modo, os aspectos fonéticos da sonorização possuem tanto uma base articulatória - à medida que todos os segmentos da fala são produzidos pelo trato vocal, envolvendo, portanto, laringe, cavidades oral e nasal, lábios, língua, etc. - quanto uma base acústica - à medida que esses aspectos fonéticos produzem atributos perceptuais relevantes para distinguir os contrastes de sons que ocorrem na língua.

No que diz respeito à base articulatória, a atividade muscular da laringe independe da atividade muscular que faz com que os lábios se abram ou a língua se mova, quando, por exemplo, se produz uma consoante oclusiva como [p] ou [b] ou

[d]. Da mesma forma, a atividade articulatória do trato vocal supralaríngeo é independente da laringe e é idêntica para as consoantes oclusivas bilabiais [b] e [p], por exemplo, produzidas em contextos similares, como nas palavras 'bato' e 'pato'. O que difere é a atividade dos músculos laríngeos em relação à abertura abrupta dos lábios que marca o começo do som inicial dessas duas palavras. Tanto em 'bato' quanto em 'pato', o sinal acústico audível começa no instante em que o falante abre os lábios. No caso de 'bato', no instante em que o falante abre seus lábios, as pregas vocais ou estão em sua posição fechada ou praticamente fechada. A fonação, aqui, pode começar em cerca de 100 ou 120 milissegundos antes do fluxo de ar escapar pelos lábios. No caso de 'pato', no instante em que o falante abre seus lábios, as pregas vocais encontram-se em sua posição aberta, o que faz com que o som inicial seja surdo - uma vez que não há vibração das pregas. O fato de as pregas vocais estarem abertas permite que haja um fluxo de ar relativamente grande e o som inicial seja gerado pela turbulência do ar quando da abertura dos lábios. O falante, quando produz a palavra 'pato', começa a fechar suas pregas vocais depois da soltura do [p].

A distinção sonoro-surdo de consoantes homorgânicas pode então repousar na variação do tempo entre a soltura da oclusão e o início de sonorização, ou seja, no *Voice onset time* (VOT). Em outras palavras, o VOT é a relação temporal entre a soltura do fechamento articulatório e o início de vibração das pregas vocais em consoantes oclusivas pré-vocálicas (cf. Lisker e Abramson, 1964), sendo um parâmetro acústico utilizado exclusivamente no estudo das oclusivas, já que essas são as únicas consoantes a envolver oclusão total da cavidade oral.

## 1.2 *Voice Onset Time*

O VOT - *voice onset time* - definido por Lisker e Abramson (1964) como o intervalo de tempo entre a soltura do fechamento da oclusiva e o início de excitação glotal (ou sonorização), é um dos parâmetros acústicos mais eficientes para a distinção de consoantes oclusivas homorgânicas sonoras e surdas em todas as línguas que possuem contraste de sonorização. Nas palavras dos próprios autores:

*This measure is both easy to make and at the same time most promising as providing the single best basis for the physical discrimination of stop manner categories. (Lisker & Abramson, 1964: 387)*

Segundo os autores, o VOT caracteriza-se por apresentar três categorias que irão especificá-lo: (1) *pré-sonorização*, em que o início de sonorização é anterior à soltura da oclusão, sendo, portanto, os valores representados pelo sinal negativo; (2) *retardo curto*, que apresenta início de sonorização e soltura da oclusão simultâneos (VOT=zero) ou início de sonorização imediatamente posterior à soltura, cujos valores distribuem-se em uma área de cerca de 10 ms a 35 ms (cf. Smith, 1978; Istre, s/d) e (3) *retardo longo*, em que o início de sonorização é posterior à soltura, distribuindo-se de 35 ms a 100 ms (cf. Smith, 1978; Istre, s/d).

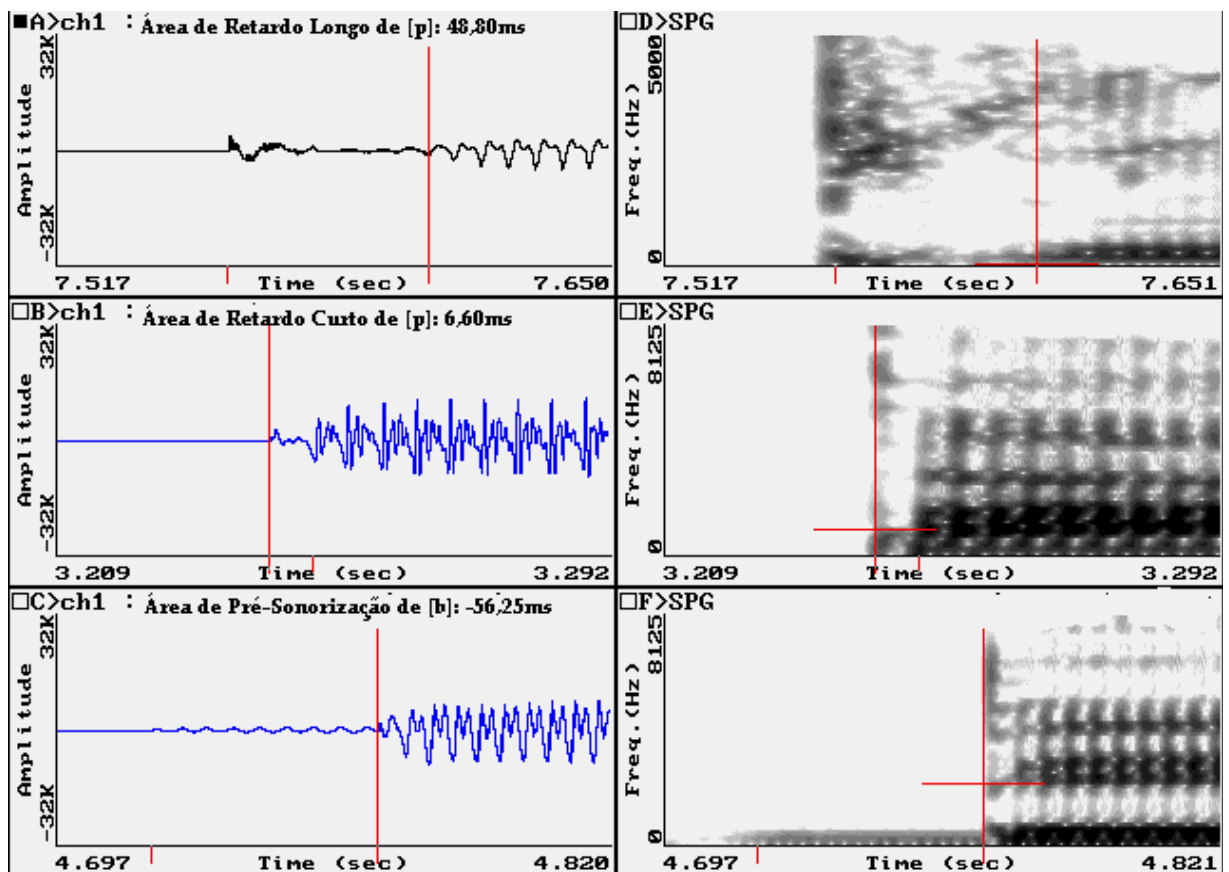


Figura 1.1 – Formas de onda e espectrogramas da área de retardo longo da oclusiva bilabial aspirada [p<sup>h</sup>] (quadros A e D), da área de retardo curto da oclusiva bilabial surda [p] (quadros B e E) e da área de pré-sonorização da oclusiva bilabial sonora [b] (quadros C e F).



Em um trabalho pioneiro, Lisker e Abramson (1964) estudaram onze línguas - inglês americano, cantonês, holandês, húngaro, espanhol porto-riquenho, tâmil, coreano, armênio oriental, tai, hindi e marati - e descobriram que, embora diferentes línguas possam apresentar categorias de VOT em comum, há variabilidade suficiente entre elas, fato que permite se pensar em ajustes específicos às línguas.

Todas as línguas que fazem o contraste sonoro *vs.* surdo irão apresentar ou as três categorias, ou apenas duas (Maddieson, 1984, *apud* Kessinger e Blumstein, 1997) e esse contraste se dará ou entre as categorias de pré-sonorização e retardo curto ou entre as categorias de retardo curto e retardo longo (Kessinger e Blumstein, 1997).

Lisker e Abramson (1964) gravaram 17 falantes que produziram uma série de palavras reais com todas as oclusivas pré-vocálicas encontradas em suas línguas, e duas sentenças para cada palavra a fim de mostrar o uso dessas oclusivas nas posições iniciais e não iniciais, com a fluência e a naturalidade de uma conversação normal. O VOT foi medido através de espectrogramas de banda larga, demarcando-se o intervalo entre a soltura da oclusão e o início da sonorização. O instante da soltura da oclusão foi encontrado marcando-se o ponto em que o padrão mostra uma mudança abrupta no espectro, e o ponto de início da sonorização foi determinado localizando-se a primeira das estriações verticais regularmente espaçadas que indica a pulsação glotal.

Os resultados revelaram que as onze línguas examinadas podem ser classificadas em três grupos, de acordo com o número de categorias de oclusivas em cada ponto de articulação, a saber, línguas de duas, três e quatro categorias.

As línguas de duas categorias caracterizam-se por apresentar, nos pontos de articulação labial, dental e velar, ou um contraste entre oclusivas sonoras e surdas, como o inglês americano (/b/-/p/, /d/-/t/ e /g/-/k/), o húngaro (/b/-/p/, /d/-/t/ e /g/-/k/), o espanhol porto-riquenho (/b/-/p/, /d/-/t/ e /g/-/k/), o tâmil (/b/-/p/, /d/-/t/ e /g/-/k/) e o holandês, que não apresenta o contraste sonoro *vs.* surdo no ponto de articulação velar, exibindo nesse caso somente a oclusiva surda (/b/-/p/, /d/-/t/ e /k/) ou, o contraste entre oclusivas surdas não aspiradas e oclusivas surdas aspiradas (/p/-/p<sup>h</sup>/, /t/-/t<sup>h</sup>/ e /k/-/k<sup>h</sup>/), como o cantonês.

Das línguas de três categorias, o armênio oriental (/b/-/p/-/p<sup>h</sup>/, /d/-/t/-/t<sup>h</sup>/ e /g/-/k/-/k<sup>h</sup>/) e o tai (/b/-/p/-/p<sup>h</sup>/, /d/-/t/-/t<sup>h</sup>/ e /k/-/k<sup>h</sup>/) caracterizam-se por apresentar oclusivas sonoras, surdas não aspiradas e surdas aspiradas nos mesmos pontos de articulação das línguas de duas categorias. O tai, como o holandês, não apresenta a oclusiva velar sonora em seu inventário fonológico. O coreano apresenta três pares de oclusivas surdas (/p/-/p<sup>c</sup>/-/p<sup>h</sup>/, /t/-/t<sup>c</sup>/-/t<sup>h</sup>/ e /k/-/k<sup>c</sup>/-/k<sup>h</sup>/)<sup>2</sup>, a saber, oclusivas surdas não aspiradas, surdas levemente aspiradas e surdas fortemente aspiradas.

As línguas de quatro categorias - hindi (/b/-/b<sup>h</sup>/-/p/-/p<sup>h</sup>/, /d/-/d<sup>h</sup>/-/t/-/t<sup>h</sup>/, /ḍ/-/ḍ<sup>h</sup>/-/ṭ/-/ṭ<sup>h</sup>/ e /g/-/g<sup>h</sup>/-/k/-/k<sup>h</sup>/) e marati (/b/-/b<sup>h</sup>/-/p/-/p<sup>h</sup>/, /d/-/d<sup>h</sup>/-/t/-/t<sup>h</sup>/, /ḍ/-/ḍ<sup>h</sup>/-/ṭ/-/ṭ<sup>h</sup>/ e /g/-/g<sup>h</sup>/-/k/-/k<sup>h</sup>/) - apresentam as oclusivas sonoras não aspiradas, sonoras aspiradas, surdas não aspiradas e surdas aspiradas, nos pontos de articulação labial, dental, domal e velar.

Das línguas encontradas no grupo (1), pode-se dizer que o holandês, o espanhol, o húngaro e o tâmil são similares, no sentido de que suas oclusivas caracterizam-se por apresentar as categorias de pré-sonorização e retardo curto, caracterizando, respectivamente as oclusivas sonoras e surdas; enquanto o cantonês caracteriza-se pelas categorias de retardo curto e retardo longo, especificando as oclusivas sonoras e surdas, respectivamente. O inglês americano assemelha-se ao cantonês, ou seja, apresenta as categorias de retardo curto e retardo longo, exceto por um falante, que produziu as oclusivas /b, d, g/ caracterizadas pela categoria de pré-sonorização: destas oclusivas produzidas com precedência de sonorização, 95% foram produzidas por um único falante. Os 5% restantes foram produzidas, todas, por outro falante. A partir dos dados do inglês pode-se, então, concluir que os falantes não produzem aleatoriamente oclusivas com valores positivos e negativos de tempo de início relativo mas, cada falante produz, pelo menos no caso de palavras isoladas, sempre ou quase sempre, um único tipo de /b, d, g/ (cf. Lisker e Abramson, 1964). Apesar de algumas diferenças em magnitude de médias e faixas de valores de

<sup>2</sup> A escolha dos símbolos para representar as oclusivas do coreano não foi baseada nas práticas de lingüistas especializados nessa língua, mas representa uma transcrição simples baseada nos

língua para língua, a total similaridade entre essas seis línguas é impressionante, pois parece que as categorias de oclusivas caem, todas, geralmente, em três faixas - uma de cerca de -125 ms a -75 ms, uma de 0 (zero) a +25 ms, e uma terceira de cerca de +60 ms a +100 ms. Os valores medianos de VOT para essas áreas são -100 ms, +10 ms e +75 ms, respectivamente. Dessa forma, o holandês, o espanhol, o húngaro e o tâmil mostram essencialmente a mesma distribuição de valores, ocupando as faixas de cerca de -100 ms e +10 ms, o cantonês localiza suas categorias na faixa e +10 ms e +75 ms e o inglês mostra uma distribuição como a do cantonês, exceto para os valores de pré-sonorização apresentados por um dos falantes (cf. Lisker e Abramson, 1964).

Das línguas do grupo (2) ou, de três categorias, o armênio oriental e o tai têm suas oclusivas sonoras, surdas não aspiradas e surdas aspiradas caracterizadas, respectivamente, pelas áreas de pré-sonorização, retardo curto e retardo longo, valores negativos e positivos de VOT, ao passo que o coreano, peculiarmente, caracteriza-se pelo fato de seus três pares de consoantes oclusivas exibirem valores positivos de VOT (oclusivas surdas não aspiradas, levemente aspiradas e fortemente aspiradas). Embora os valores de VOT dos pares de oclusivas não aspiradas e levemente aspiradas do coreano sejam muito similares, Lisker e Abramson não acreditam que a medida do VOT não dê conta de separar essas duas classes de oclusivas:

*“...while the distribution of values is thus somewhat anomalous, we cannot say with reasonable assurance that our measure of voice onset time fails to separate the three categories of Korean stops; it will certainly suffice to distinguish the aspirated set from the other two and it may still well be the single most important measure for separating the latter.” (Lisker e Abramson, 1964: 403)*

Para as línguas de quatro categorias - hindi e marati -, contudo, Lisker e Abramson reconhecem que o VOT parece ser insuficiente para distinguir entre todas as categorias, no sentido de que:

*“... the voiced unaspirated and voiced aspirated stops show differences in average values that are almost systematic; nevertheless they occupy ranges that are nearly coextensive. It seems very likely that voiced aspirates are distinguished*

*from the other voiced category by the presence of low amplitude buzz mixed with noise in the interval following release of the stop.” (Lisker e Abramson, 1964: 403)*

As descobertas de Lisker e Abramson mostram também que o VOT exibe, nas várias línguas analisadas, variações intrínsecas como uma função de ponto de articulação - os valores de VOT serão tanto maiores quanto mais posterior for o ponto de articulação, pelo menos quando se tratar de valores positivos de VOT, ou seja, no caso de oclusivas que apresentam as categorias de retardo curto e retardo longo. Dessa forma, as oclusivas bilabiais terão valores menores que as oclusivas alveolares, que terão valores menores que as oclusivas velares.

Os autores chegaram ainda à conclusão, comparando palavras isoladas com palavras inseridas em sentenças, de que para categorias com precedência de sonorização, em posição inicial, a sonorização se dá, usualmente, de forma contínua, através de um ambiente sonoro precedente durante a oclusão; qualquer interrupção de vibração glotal depende dos sons que ocorrem antes da oclusão. Esse mesmo efeito é observado também com relação às oclusivas /b, d, g/ do inglês e /p<sup>c</sup>, t<sup>c</sup>, k<sup>c</sup>/ (oclusivas surdas levemente aspiradas) do coreano. Outro efeito observado pelos pesquisadores, como consequência da inserção de palavras em sentenças, é o de que os valores de VOT, tanto os que definem a área de pré-sonorização quanto os que definem as áreas de retardo curto e retardo longo, tendem a ser reduzidos nas palavras inseridas em sentenças. Nos vários casos em que ocorre essa redução, há naturalmente uma diminuição do intervalo que separa as categorias de pré-sonorização, retardo curto e retardo longo umas das outras ao longo da dimensão de VOT. Lisker e Abramson mostram que, em consequência de tal redução, poderá haver sobreposição entre as categorias, que de outra forma seriam distintas.

O trabalho destes autores foi o primeiro de uma série na investigação do parâmetro acústico VOT. Depois dele, vários estudos vêm sendo realizados, tanto em produção quanto em percepção da fala, em diferentes línguas.

No PB, o VOT foi investigado por Istre (s/d) que, baseado nos resultados obtidos por Lisker e Abramson (1964) para as diversas línguas, principalmente nos dados do espanhol porto-riquenho, resolveu testar duas hipóteses com relação ao português.

O espanhol porto-riquenho, assim como o PB, realizam um contraste entre oclusivas sonoras e oclusivas surdas não aspiradas. Os valores médios de VOT encontrados por Lisker e Abramson (1964), para as sonoras do espanhol, foram de -138 ms para a bilabial /b/, -110 ms para a dental /d/ e -108 ms para a velar /g/. Os valores médios de VOT para as oclusivas surdas foram de 4 ms, 9 ms e 29 ms para a labial /p/, a dental /t/ e a velar /k/, respectivamente.

Das hipóteses levantadas por Istre para o PB, a primeira foi de que as oclusivas sonoras do PB apresentariam uma precedência longa de sonorização e a segunda, de que as oclusivas surdas mostrariam retardo curto.

Istre investigou o VOT de quatro falantes masculinos, monolíngües, do PB, naturais do Estado de Santa Catarina, que realizaram segmentos monossilábicos do tipo CV, apresentando as seis oclusivas - /b, d, g/ e /p, t, k/ - e as sete vogais - /i, e, ε, a, o, u/.

Os valores médios de VOT obtidos para as oclusivas sonoras /b, d, g/ foram de -39,45 ms, -60,59 ms e -53,69 ms, respectivamente. Os valores médios para as oclusivas surdas /p, t, k/ foram, respectivamente, de 11,95 ms, 18,48 ms e 38,53 ms. Esses resultados revelaram que as oclusivas do PB, comparadas a oclusivas de outras línguas, podem ser classificadas como apresentando retardo curto para as oclusivas surdas e uma precedência média de sonorização para as oclusivas sonoras. Desta forma, a primeira hipótese levantada por Istre, de que as oclusivas sonoras apresentariam uma área de precedência longa de sonorização, não foi comprovada pelos dados.

### 1.3 Revisão da literatura

#### 1.3.1. O VOT na produção da fala

Weismer (1979) empreendeu um estudo sobre o inglês que visava determinar se o VOT da oclusiva inicial de palavras monossilábicas CVC sofreria influências sistemáticas através de uma distinção de vogal tensa/relaxada, assim como confirmar e definir mais acuradamente o efeito do caráter sonoro/surdo da consoante de coda sobre o VOT da consoante de *onset* dessas mesmas palavras. Analisou, então, as oclusivas /p, t, k/ iniciais de palavra, realizando tantas combinações quanto possível entre vogais tensas e relaxadas e oclusivas sonoras e surdas em posição de coda. Como resultado, obteve as quatro combinações seguintes: tenso/sonoro, tenso/surdo, relaxado/sonoro e relaxado/surdo. Para a determinação das vogais tensas e relaxadas, seguiu-se a classificação de Ladefoged (1975, *apud* Weismer, 1976), incluindo ainda as tensas /i/, /ei/ e /u/ e as relaxadas /ɪ/, /ɛ/ e /æ/.

Os resultados obtidos por Weismer indicaram que, (1) o valor médio de VOT é maior quando a vogal na palavra-teste é tensa do que quando é relaxada, independentemente da característica de sonorização da consoante de coda; (2) o VOT é maior quando a consoante final é sonora do que quando é surda e (3), como confirmação das descobertas (1) e (2), as palavras em que a vogal é tensa e a consoante final sonora têm um valor de VOT maior para a consoante de *onset* do que palavras em que a vogal é relaxada e a consoante final surda.

Weismer chegou, então, à conclusão de que o VOT sofre influências tanto da vogal seguinte quanto da consoante de coda, pelo menos em se tratando de palavras monossilábicas do tipo CVC. Contudo, essa influência não se dá de forma sistemática, uma vez que o efeito da oposição tensa/relaxada da vogal sobre o VOT é levemente maior do que o efeito do contraste sonoro/surdo da consoante final.

Bond, Eddey e Bermejo (1980) desenvolveram uma pesquisa visando investigar os efeitos do aprendizado de uma segunda língua nos valores de VOT de oclusivas iniciais de palavras na fala infantil. Foram analisadas e comparadas as produções de palavras - três pares mínimos do espanhol e do inglês que iniciavam com as oclusivas bilabiais, dentais e velares - de dois irmãos, falantes nativos do espanhol. Uma das crianças apresentava atraso na aquisição de sua língua materna e as duas estavam aprendendo inglês como segunda língua.

Foram levantadas duas hipóteses pelos autores. A primeira, de que, após seis meses de exposição à língua inglesa, as crianças começariam, espontaneamente, a diferenciar os valores de VOT do espanhol e do inglês, e a segunda, de que haveria diferenças entre as produções das crianças devido aos problemas de linguagem de uma delas.

Sabe-se que falantes nativos adultos do espanhol apresentam as áreas de pré-sonorização e retardo curto caracterizando, respectivamente, as oclusivas sonoras e surdas e que falantes nativos adultos do inglês apresentam a área de retardo curto, especificando as oclusivas sonoras, e a área de retardo longo, as oclusivas surdas.

Os resultados obtidos pelos autores levaram-nos a verificar que a criança com desenvolvimento normal da linguagem seguia um padrão de desenvolvimento, estendendo os valores de VOT da área de retardo curto em direção às áreas de pré-sonorização e retardo longo, aproximando-se dos valores que caracterizam o VOT dos adultos, ou seja, os valores de VOT que caracterizam as línguas espanhola e inglesa. Esses resultados parecem estar de acordo com os de outras pesquisas relacionadas ao desenvolvimento da distinção de sonorização em crianças. Zlatin e Koenigsknecht (1976 *apud* Bond *et al.*, 1980) mostraram que mudanças de desenvolvimento na produção de oclusivas iniciais de palavras em todos os pontos de articulação ocorreram entre crianças de dois e seis anos de idade. Descobriram que, até a idade de três anos, as produções das crianças eram unimodais, concentradas na área de retardo curto do *continuum* de VOT. Após essa idade, começaram a surgir distribuições bimodais (de VOT), isto é, os valores começaram a encaminhar-se para as áreas de pré-sonorização e retardo longo. Kewly-Port e

Preston (1974 *apud* Bond *et al.*, 1980) mostraram que crianças entre dois e quatro anos e meio de idade produziam tanto oclusivas apicais sonoras quanto surdas na área de retardo curto. Sugeriram, então, que os ajustes necessários para produzir uma oclusiva na área de retardo curto são menos complexos que o controle articulatorio necessário para produzir oclusivas com precedência ou retardo longo de sonorização.

A criança que apresentava problemas no processo de aquisição, analisada por Bond *et al.*, entretanto, estava atrasada em desenvolver a distribuição bimodal do VOT, uma vez que os valores de VOT para suas produções ficaram confinados à área de retardo curto tanto para o espanhol quanto para o inglês.

Dos resultados obtidos por Bond *et al.* pode-se verificar que suas duas hipóteses iniciais se confirmaram. A primeira, contudo, confirmou-se somente em parte, já que apenas a criança com desenvolvimento normal da linguagem tinha começado a diferenciar, após alguns meses de exposição ao inglês, os valores de VOT das duas línguas. Os autores sugeriram que talvez o atraso no desenvolvimento em alcançar os valores de VOT característicos da fala dos adultos, da criança com problemas no processo de aquisição da linguagem, tanto para o espanhol quanto para o inglês, seja uma função de seu perfil global de desenvolvimento tardio da linguagem. Essa relação já havia sido mostrada por Menyuk e Looney (1972 *apud* Bond *et al.*, 1980), que descobriram que crianças com dificuldades para produzir estruturas sintáticas específicas, apresentavam dificuldades também para produzir estruturas fonológicas.

Seguindo os passos de Bond *et al.* (1980), Konefal e Fokes (1981) examinaram os valores de VOT para seis oclusivas iniciais de palavras do espanhol e do inglês produzidas por três irmãs - uma de quatro anos, uma de sete e outra de dez - falantes nativas do espanhol e expostas à língua inglesa há três anos. As duas crianças mais novas não apresentavam quaisquer problemas no desenvolvimento da linguagem, enquanto sua irmã mais velha havia desenvolvido tardiamente o processo de aquisição.



Os resultados revelaram que, para as duas crianças com desenvolvimento normal da linguagem, tanto as categorias de pré-sonorização e retardo curto do espanhol quanto as categorias de retardo curto e retardo longo do inglês apresentaram diferenças significativas. A criança de quatro anos manteve o contraste de sonorização na língua inglesa com a maior parte dos valores de VOT para as oclusivas sonoras na área de retardo curto. Suas oclusivas surdas apresentaram estabilidade, com os valores de VOT agrupando-se em torno de uma área de 80 ms a 120 ms, retardo longo, portanto. As oclusivas sonoras do espanhol, para essa mesma criança, fixaram-se na área de retardo curto, enquanto as oclusivas surdas agruparam-se entre 20 ms e 80 ms, parte na área de retardo curto e parte na área de retardo longo. Relembrando o que Kewly-Port e Preston (1974 *apud* Bond *et al.* e Konefal e Fokes, 1981) sugeriram, os ajustes articulatórios necessários para produzir uma oclusiva na área de retardo curto são menos complexos do que o controle articulatório necessário para produzir oclusivas com pré-sonorização e retardo longo. A criança de sete anos também manteve o contraste sonoro-surdo do inglês, com as oclusivas sonoras produzidas na área de retardo curto e as oclusivas surdas agrupadas na área de retardo longo - em cerca de 40 ms a 60 ms - bastante reduzida em relação à mesma área apresentada por sua irmã. As oclusivas sonoras do espanhol para essa criança caracterizaram-se pela área de pré-sonorização, e todas as suas oclusivas surdas foram produzidas na área de retardo curto. Novamente, em contraste com sua irmã, essa criança apresentou os valores de VOT nas três categorias que o especificam nas duas línguas, a saber, pré-sonorização, retardo curto e retardo longo. Ela adquiriu, portanto, os contrastes sonoro-surdo tanto em espanhol quanto em inglês.

A criança com distúrbios de linguagem, no entanto, não manteve uma distinção entre nenhuma das categorias que especificam as oclusivas sonoras e surdas, nem no espanhol, nem no inglês. Parece ter havido uma sobreposição das categorias de VOT que definem as oclusivas sonoras e surdas produzidas por essa criança em ambas as línguas. Parte da área de retardo curto - que no inglês caracteriza as oclusivas sonoras - estendeu-se sobre a área de retardo longo,

localizando-se em uma faixa de 40 ms a 70 ms. A área de retardo longo, que caracteriza as oclusivas surdas, agrupou-se entre 40 ms e 80 ms. Quanto ao espanhol, sua língua materna, ela apresentou alguma pré-sonorização para as oclusivas sonoras, mas a maior parte dos valores de VOT caiu na área de retardo curto. A categoria de retardo curto estendeu-se sobre a categoria de retardo longo. Neste caso, também, não parece ter havido distinção clara entre as categorias sonora e surda do espanhol. O fato de as áreas de VOT que caracterizam suas produções terem ficado predominantemente limitadas à área de retardo curto, tanto no espanhol quanto no inglês, aponta para a questão de que essa criança possa estar atrasada em desenvolver uma distribuição bimodal de VOT. Esse atraso pode ser uma manifestação de seu distúrbio em linguagem.

Konefal e Fokes concluíram, assim, que deve haver sim, um padrão de desenvolvimento de valores de VOT, que vai de uma área de retardo curto a uma área de retardo longo e uma área de pré-sonorização para falantes bilíngües de espanhol e inglês que tenham estado em contato suficiente com as duas línguas. Concluíram, ainda, que uma criança com problemas de linguagem pode, em conseqüência, manifestar um atraso em parâmetros acústicos da fala específicos à língua.

Barton e Macken (1980) desenvolveram um estudo no qual investigaram o VOT de oclusivas iniciais de palavras produzidas por crianças de quatro anos de idade, monolíngües, falantes nativas do inglês americano. A pesquisa envolveu comparações com dados de estudos anteriores sobre a fala de crianças de dois anos de idade (Macken e Barton, 1980 *apud* Barton e Macken, 1980) e de adultos (Lisker e Abramson, 1964). Para tal investigação, Barton e Macken realizaram dois experimentos. O primeiro envolveu a análise de uma amostra de fala espontânea a fim de se obter uma visão geral do VOT das oclusivas das crianças de quatro anos. O segundo, a análise de sinais emitidos isoladamente comparados a sinais inseridos em sentenças.

Sabe-se que o contraste de sonorização das oclusivas do inglês se dá em três pontos de articulação - bilabial, alveolar e velar - e que, no caso das oclusivas sonoras /b, d, g/, o início de pulsação glotal se dá, usualmente, simultâneo a ou imediatamente após a soltura da oclusão (área de retardo curto), embora alguns falantes demonstrem tendência a produzir oclusivas sonoras com vibração glotal anterior à soltura (área de pré-sonorização). Já as oclusivas surdas são produzidas com um intervalo de tempo mais longo entre a soltura do fechamento da oclusiva e o início de vibração das pregas vocais (área de retardo longo).

Em seu primeiro trabalho sobre o VOT, Lisker e Abramson (1964) mostraram que, dos quatro sujeitos estudados, adultos, falantes nativos do inglês americano, um apresentou a área de pré-sonorização caracterizando as oclusivas sonoras, enquanto os outros três falantes apresentaram as mesmas oclusivas caracterizadas pela área de retardo curto. Quando as palavras eram faladas isoladamente, o VOT das oclusivas apresentava valores maiores do que quando essas mesmas palavras eram inseridas em sentenças. Esse trabalho mostrou também que não houve sobreposição entre as áreas de valores de VOT em nenhum dos pontos de articulação para as palavras faladas isoladamente. Quando as palavras eram inseridas em sentenças, os valores de VOT para as oclusivas surdas eram consideravelmente menores do que aqueles constatados nas palavras isoladas e a distinção entre as categorias de VOT era reduzida, podendo haver, nesse caso, sobreposição entre as áreas de valores que caracterizam as oclusivas sonoras e surdas.

Macken e Barton (1980 *apud* Barton e Macken, 1980), em estudo anterior, ao investigar o VOT de crianças com dois anos de idade, falantes nativas do inglês americano, encontraram evidências para a existência de três estágios na aquisição do contraste de sonorização. No primeiro estágio, as crianças não produzem qualquer distinção em sua fala, em termos de VOT, em oclusivas que são fonemicamente sonoras e surdas na forma adulta - muitos dos sinais tanto para as oclusivas sonoras quanto para as oclusivas surdas localizam-se na área de 0 (zero) a 30 ms, não havendo diferença nos valores médios obtidos para as duas categorias. No segundo estágio, as distribuições para oclusivas sonoras e surdas, em qualquer dos pontos de

articulação, apresentam sobreposição considerável - muitos dos sinais, tanto para as oclusivas sonoras quanto para as surdas, encontram-se na área de retardo curto, mas as médias para as oclusivas surdas são significativamente maiores do que as médias para as oclusivas sonoras. Esse é o primeiro estágio em que há evidência de uma distinção entre as oclusivas sonoras e surdas nas produções das crianças. No terceiro estágio, as crianças, primeiro, 'ultrapassam' os valores dos adultos, produzindo oclusivas surdas que são consideravelmente maiores, aproximando-se, depois, gradualmente dos valores de VOT dos adultos.

Os resultados obtidos por Barton e Macken revelaram que a fala espontânea das crianças de quatro anos mostrou que elas apresentavam distribuições distintas para as oclusivas sonoras e surdas nos três pontos de articulação. Mostraram ainda que houve alguma sobreposição entre cognatos sonoros e surdos. Quanto ao ponto de articulação, as crianças de quatro anos seguiram a mesma tendência dos adultos, apresentando valores de VOT maiores para as oclusivas velares, e menores um pouco para as alveolares e bilabiais, respectivamente. Quanto ao contexto, essas crianças apresentaram, como os adultos do estudo de Lisker e Abramson (1964), alguma sobreposição quando os segmentos eram inseridos em sentenças nos três pontos de articulação. Entretanto, os valores médios de VOT apresentados pelas crianças eram maiores do que os valores médios apresentados pelos adultos, tanto para os sinais produzidos isoladamente quanto para aqueles inseridos em sentenças.

No que diz respeito às crianças de dois anos de idade, cada uma encontrava-se em um estágio diferente ao final da pesquisa (Macken e Barton, 1980 *apud* Barton e Macken, 1980). Uma delas não apresentava um contraste claro entre as oclusivas sonoras e surdas e produzia muitos desses sinais na área de retardo curto. As outras três tinham 'ultrapassado' os valores dos adultos para as oclusivas surdas. Dessas, uma já se aproximava das produções dos adultos e outra, embora apresentasse maior variabilidade, parecia estar se aproximando dos valores apresentados pelos adultos. Comparando os dados apresentados pelas crianças de dois anos de idade com as de quatro, observou-se a mesma tendência de diminuição nos valores de VOT de segmentos inseridos em sentenças.

Os resultados do estudo levaram à conclusão de que, ao produzirem oclusivas surdas, no que diz respeito ao VOT, as crianças de quatro anos, numa primeira fase, ‘ultrapassam’ os valores dos adultos e, então, gradualmente, os diminuem em direção aos valores de VOT destes.

Flege e Port (1981) realizaram um estudo no qual compararam a implementação fonética do contraste de sonorização de oclusivas produzidas no árabe por falantes nativos do dialeto falado na Arábia Saudita e no inglês, por falantes nativos de inglês e do mesmo dialeto árabe. Os autores examinaram várias dimensões acústicas que são correlatos fonéticos do contraste fonológico entre oclusivas sonoras e surdas, dentre elas, o VOT, nas duas línguas em questão.

A escolha do árabe como contraponto ao inglês deu-se porque o contraste entre oclusivas sonoras e surdas no árabe parece diferir daquele do inglês (Yeni-Komshian, Caramazza e Preston, 1977; Port, Al-ani e Maeda, 1980 *apud* Flege e Port, 1981) e porque, no árabe, falta uma das oclusivas encontradas no inglês - a oclusiva bilabial surda /p/ (Al-ani, 1970 *apud* Flege e Port, 1981). Essas diferenças entre as duas línguas ofereceram a oportunidade de se avaliar como diferenças no inventário fonológico e na implementação fonética de um contraste fonológico afetariam a produção dos sons da fala de uma língua estrangeira por aprendizes adultos.

Os resultados revelaram que as oclusivas do inglês produzidas por árabes manifestaram valores de VOT similares aos valores de VOT encontrados nas oclusivas de sua língua materna (o árabe), ou seja, as oclusivas do inglês realizadas por árabes tiveram valores de VOT menores do que as mesmas oclusivas realizadas por americanos.

Em termos de percepção, a despeito desse tipo de interferência fonética do árabe sobre o inglês, ouvintes americanos praticamente não tiveram dificuldade em identificar as oclusivas do inglês produzidas por árabes, com exceção do /p/. Esse fonema, que não existe no árabe, foi freqüentemente produzido com pulsação glotal

durante o intervalo de fechamento da oclusiva, sendo percebido, portanto, como /b/.

Fokes, Bond e Steinberg (1985) desenvolveram um estudo no qual investigaram a aquisição do contraste de sonorização das consoantes oclusivas em posição inicial de sílaba do inglês, dado pelo VOT, na fala de crianças árabes.

Sabe-se que o árabe e o inglês possuem oclusivas em comum. Todavia, essas oclusivas diferem tanto nos detalhes fonéticos que especificam o contraste sonoro-surdo, quanto no inventário total de oclusivas. No que diz respeito ao árabe, parece haver algumas diferenças entre suas variedades. Al-Ani (1970 *apud* Fokes *et al.*, 1985) registra as surdas /t, k/ e as sonoras /b, d/ como as oclusivas encontradas no árabe *standard* falado no Iraque. Flege e Port (1981) registram, como oclusivas do árabe falado na Arábia Saudita, as surdas /t, k/ e as sonoras /b, d, g/. Em ambas as classificações, em posição inicial, as oclusivas surdas são aspiradas e as sonoras caracterizadas pela presença de vibração glotal durante a fase de oclusão. O VOT das oclusivas /t/ e /k/ do árabe, no entanto, é menor do que o VOT das mesmas oclusivas do inglês. Além disso, o contraste de sonorização, no inglês, se dá nos três pontos de articulação, tanto para as oclusivas sonoras quanto para as surdas, a saber, bilabiais (/b/ e /p/), alveolares (/d/ e /t/) e velares (/g/ e /k/).

Apesar de os inventários das oclusivas diferirem nas duas línguas, uma das características distintivas primárias das oclusivas sonoras e surdas em posição inicial parece ser a aspiração.

Como conseqüência destas diferenças entre as oclusivas das duas línguas, os falantes árabes em processo de aquisição do inglês deparam-se com dois problemas: (1) têm de incorporar uma oclusiva bilabial surda a seu inventário fonêmico e (2) precisam modificar os correlatos das oclusivas surdas do árabe para a norma do inglês, empregando valores de VOT mais longos para esses segmentos iniciais de sílaba. Flege e Port (1981) mostraram que adultos falantes nativos do árabe fazem uso dos correlatos acústicos do contraste de sonorização de sua língua materna ao produzirem oclusivas sonoras e surdas do inglês, uma vez que os valores de VOT

que caracterizam a língua árabe são tipicamente mais curtos do que aqueles que caracterizam o inglês.

Quanto às crianças, como já demonstraram Bond *et al.* (1980) e Konefal e Fokes (1981), o controle de desenvolvimento do VOT segue uma seqüência nitidamente determinada, no sentido de que as crianças produzem primeiro as oclusivas com sonorização na área de retardo curto, ou seja, seguindo imediatamente a soltura do fechamento da oclusiva. Nesse estágio, os valores de VOT formam uma distribuição unimodal, pois concentram-se em uma única área de VOT. Depois, as crianças começam a desenvolver uma distribuição bimodal dos valores de VOT, atrasando ou antecipando consideravelmente o início de sonorização com relação ao momento de soltura da oclusão, produzindo, respectivamente, oclusivas nas áreas de retardo longo e de pré-sonorização.

Fokes *et al.* (1985), ao empreenderem seu experimento para investigar a aquisição de correlatos acústicos do contraste de sonorização na fala de crianças árabes, procuraram descobrir:

1. se, a despeito das diferenças em VOT entre o inglês e o árabe para oclusivas iniciais de palavras, crianças árabes que estivessem aprendendo inglês como segunda língua iriam apresentar características da língua árabe na produção de oclusivas do inglês ou, se estariam aptas a adquirir os contrastes do inglês;
2. caso adquirissem os contrastes do inglês, se haveria uma tendência similar na aquisição da distinção entre oclusivas labiais e alveolares iniciais;
3. caso existisse esta tendência, se a aquisição seria mais evidente como um resultado da idade ou da exposição ao inglês;
4. se as tendências de aquisição para as crianças árabes seriam similares àquelas mostradas por crianças falantes nativas do inglês em fase de aquisição de sua língua materna.

Fokes *et al.* gravaram então doze crianças de dois a onze anos de idade, agrupadas de acordo com a idade e com a experiência com o inglês. Os resultados mostraram que, das crianças com menos experiência com o inglês, somente duas

(com apenas dois meses de exposição ao inglês) apresentaram sobreposição extensiva dos valores de VOT para /p/ e /b/. Após praticamente um ano de experiência com o inglês, nenhuma das crianças mostrou sobreposição extensiva das categorias sonora/surda, o que indica algum controle na produção do VOT. Várias crianças apresentaram valores de pré-sonorização e de retardo muito altos. O que parece acontecer é que as crianças são capazes de usar contrastivamente o VOT, mas o fazem com grande variabilidade fonética, assim como fazem as crianças falantes nativas do inglês.

Dos resultados obtidos, os autores puderam concluir, respondendo à primeira questão, que após alguma experiência com a língua inglesa, as crianças árabes parecem ter adquirido o contraste de sonorização do inglês. No entanto, os valores de VOT para essas crianças foram consideravelmente exagerados em relação aos valores de VOT dos adultos falantes nativos do inglês, mostrados na pesquisa de Lisker e Abramson (1964).

Com relação à segunda questão, parece ter havido uma certa consistência na forma com que as crianças usaram o VOT em diferentes pontos de articulação: ocorrendo diferenças para oclusivas alveolares, distinções similares ocorriam para oclusivas bilabiais, mesmo que o contraste de sonorização não existisse, no árabe, nesse ponto de articulação. Essa descoberta sugere que as crianças árabes devem simplesmente adicionar um contraste fonêmico ao aprenderem inglês.

Os resultados mostraram também, no que tange à terceira questão, que as crianças com menos experiência com o inglês tiveram um desempenho muito similar ao desempenho das crianças com mais experiência.

Finalmente, quanto à quarta questão, Fokes *et al.* concluíram que crianças árabes em processo de aquisição do inglês apresentam valores de VOT mais próximos daqueles produzidos por falantes adultos nativos do inglês do que os falantes adultos do árabe.



Andrade e Mascarenhas (1994)<sup>3</sup>, em um estudo sobre os mecanismos de produção subjacentes ao desvozeamento das vogais altas não acentuadas do português europeu (doravante PE), acabaram por estudar o VOT da oclusiva coronal /t/ diante da vogal alta /i/ nos contextos pretônico, tônico e postônico, em palavras que ocorreram em diferentes posições nas frases.

Em um estudo piloto, analisando o segmento /ti/, Andrade (1984 *apud* Andrade e Mascarenhas, 1994) encontrou valores de VOT maiores em sílaba não acentuada do que em sílaba acentuada, no PE. As diferenças foram pequenas mas estatisticamente significativas, permitindo que se levantasse a hipótese de o ataque da vogal em sílaba não acentuada tender a ser mais tardio do que em sílaba acentuada. Andrade e Mascarenhas, em seu experimento, utilizaram diversas frases que foram lidas várias vezes em ordem aleatória e da forma mais natural possível pelos sujeitos, em uma velocidade considerada normal pelos falantes. A entonação foi descendente, sem marcações prosódicas notáveis.

Em sua maioria, os trabalhos sobre o VOT no PE dizem respeito a sílabas isoladas ou a sílabas tônicas no caso de mais de uma sílaba. Em um estudo sobre a sonorização e a africacão de oclusivas no PE, Viana (1984 *apud* Andrade e Mascarenhas, 1994) demonstrou, para segmentos produzidos isoladamente, valores médios de VOT positivo que são maiores no caso de vogais altas do que no caso de vogais não altas. Andrade (1980 *apud* Andrade e Mascarenhas, 1994) analisou o VOT de oclusivas diante de vogal não alta e seus resultados foram convergentes com os de Viana (1984).

Quanto ao trabalho de Andrade e Mascarenhas (1994), seus resultados mostraram pequenas diferenças com relação à sílaba /ti/ em posição tônica e pretônica. Todavia, essas diferenças foram de encontro à hipótese inicialmente levantada, de que oclusivas em sílabas não acentuadas apresentariam valores de VOT maiores do que oclusivas em sílabas acentuadas. Na verdade, ocorreu o contrário: “a sílaba não acentuada foi associada às médias de VOT mais baixas, e não

---

<sup>3</sup> Trabalho que se insere no projeto “Estudo experimental de processos de lenição vocálica em Português Europeu”, do Centro de Linguística da Universidade de Lisboa.

às mais altas” (Andrade e Mascarenhas, 1994). Mas a influência da tonicidade manifestou-se na duração média da oclusão da coronal: *“o período de silêncio da consoante é, em média, mais longo em posição acentuada do que em posição pré-acentual, podendo a variação de VOT acompanhar esta diferença, embora mais fracamente e de forma não sistemática.”*<sup>4</sup> (Andrade e Mascarenhas, 1994: s/p). No que diz respeito à posição postônica, os resultados, como no caso das oclusivas pretônicas, também não apoiaram a hipótese inicial, isto é, as médias de VOT mais baixas foram associadas à sílaba não acentuada, e não à sílaba acentuada.

Quanto à posição na frase, por considerar uma possível influência dessa variável nos valores de VOT, as autoras analisaram as realizações de /ti/ de duas das palavras analisadas (‘tinha’ e ‘Saqueti’) em posição final e não final na frase. Observaram que dois dos três falantes que participaram do experimento manifestaram o mesmo tipo de comportamento, produzindo valores de VOT maiores em posição final de frase. O outro falante, contudo, surpreendeu pelo fato de apresentar valores bastante elevados de VOT na sílaba tônica da palavra ‘tinha’ (que ora funciona como verbo principal, ora como verbo auxiliar) não final de frase, em contraste com os valores relativamente baixos em final de frase. Esse falante parece ter adotado uma estratégia de compensação com relação ao VOT e à duração do período de oclusão da coronal: em ‘tinha’ final, as oclusões foram longas e o VOT breve, ao passo que em ‘tinha’ não final (auxiliar) ocorreu o inverso. Quanto à sílaba não acentuada da palavra ‘Saqueti’, observou-se somente uma diferença muito pequena entre as duas posições na frase para este falante.

As conclusões alcançadas por Andrade e Mascarenhas (1994) mostraram, primeiramente, que os valores de VOT mais elevados não estão associados à sílaba não acentuada e sim, à sílaba acentuada. Observaram, ainda, que os valores elevados de VOT refletem a africacão da oclusiva. *“Este fenômeno, característico do português do Brasil”* (Andrade e Mascarenhas, 1997: s/p), ocorre somente nas realizações não finais da palavra ‘tinha’ (funcionando como verbo auxiliar) na pesquisa em questão.

---

<sup>4</sup> Grifo no original.

Voltado para a área de Identificação de Falantes, Figueiredo (1994), ao lado de Inspeção Visual de Espectrogramas e mais cinco parâmetros acústicos - Formantes Vocálicos, Frequência Fundamental, Espectro de Longo Termo, Velocidade de Fala e Consoantes Nasais - estudou o VOT da consoante oclusiva dental surda do PB.

Figueiredo (1994) realizou sua pesquisa com o intuito de verificar se o VOT serviria como parâmetro para identificar falantes e observar a variação inter-falantes desse parâmetro.

Wolf (1972 *apud* Figueiredo, 1994) já enfatizava, ao definir os principais requisitos na escolha de parâmetros acústicos eficientes para identificar falantes, “aquelas medidas relacionadas a aspectos da fala cujo mecanismo articulatório subjacente seja transparente para o falante, ou seja, ações articulatórias sobre as quais o falante não possa exercer controle consciente”. (Figueiredo, 1994: 277). Esse tipo de parâmetro, segundo Wolf, seria mais resistente a tentativas de disfarce e imitação. Como exemplo, o autor cita a duração da pré-sonorização em uma oclusiva sonora precedida por um segmento surdo. No caso do inglês, a distinção entre oclusivas sonoras e surdas em termos de sinal positivo ou negativo do VOT é válida somente para a posição medial, uma vez que em posição inicial, tanto as oclusivas sonoras quanto as surdas são produzidas com valores positivos de VOT, ou seja, são caracterizadas pelas áreas de retardo curto (sonoras) e retardo longo (surdas). Lisker e Abramson (1964) estudaram o VOT em diferentes línguas e os resultados indicaram que, na grande maioria, a variação intra-falante do VOT ocorreu com as oclusivas sonoras, mesmo que as duas categorias de oclusivas - sonoras e surdas - ocorressem no mesmo ambiente fonético.

“Esses resultados sugerem que as plosivas não sonoras podem fornecer informação mais consistente do que as sonoras no que diz respeito à identidade do falante, ao contrário do que sugerira Wolf. (Figueiredo, 1994: 278)

Baseado na maior variação dos segmentos sonoros, Figueiredo analisou o VOT da consoante oclusiva dental /t/, diante do segmento vocálico /ɔ/, na palavra *laboratório*. A palavra-alvo ocupou diferentes posições em frases, que foram lidas na

velocidade considerada normal pelos falantes. Um dos falantes participou de duas sessões de gravação, com uma defasagem de alguns meses, visto que por ocasião da primeira sessão, encontrava-se sob forte estado gripal, apresentando, provavelmente, inflamação na laringe.

“Assim, os resultados para esse falante poderão servir como indicadores para a estabilidade do VOT em amostras não contemporâneas e, acumulativamente, em função da presença de alterações orgânicas transientes”. (Figueiredo, 1994: 279).

Os resultados obtidos por Figueiredo mostraram que a faixa total de variação inter-falante foi de 12,30 ms. Apenas dois dos falantes, um com a média de VOT mínima e o outro com a média máxima do grupo, foram significativamente diferentes de todos os demais falantes. O falante que por ocasião da primeira gravação encontrava-se sob forte estado gripal apresentou valores de VOT semelhantes nas duas sessões de gravação, o que sugere que o parâmetro em questão é relativamente estável, pelo menos no mesmo contexto fonético. As conseqüências da inflamação da laringe - variação de massa e consistência dos tecidos na região da glote e/ou diminuição da pressão trans-glotal em função da capacidade pulmonar - parecem não influenciar o VOT, o que vai contra a afirmação de Löfqvist (1992 *apud* Figueiredo, 1994), que atribui a alguns aspectos relacionados à glote (área glotal, pressão trans-glotal e tensão) uma função importante no comportamento do VOT.

O VOT, segundo Figueiredo, mostrou ser um parâmetro ‘potencialmente útil’ na área de Identificação de Falantes. Em graus de comparação com a Inspeção Visual de Espectrogramas e os outros cinco parâmetros acústicos estudados por ele, o VOT foi o segundo em eficácia. No entanto, não é propriamente um “determinante de identidade”, mas sim, um “indicador genérico”.

“Informações dessa natureza não deixam de ser válidas, especialmente na aplicação forense, onde a evidência de identificação negativa (exclusão de suspeitos) também cumpre um papel relevante” (Figueiredo, 1994: 286).

Os trabalhos expostos acima têm, todos, relação com a produção da fala. A seguir, serão expostos alguns estudos com relação à percepção.

### 1.3.2. O VOT na percepção da fala

Uselding (1977) empreendeu dois experimentos cujo objetivo era discriminar segmentos de fala sintética que apresentavam variação em VOT. Os segmentos incluíram oclusivas bilabiais e velares.

Uma importante descoberta que tem resultado do uso de segmentos sinteticamente produzidos é que, embora o VOT seja um *continuum*, a percepção de adultos ao longo desse *continuum* é categórica, isto é, sílabas CV que variam quanto ao VOT são “perceptualmente agrupadas”<sup>5</sup> em categorias que têm fronteiras claras. Enquanto essas fronteiras podem variar para diferentes procedimentos de construção na síntese de segmentos CV, descobriu-se que a fronteira fixa-se em cerca +25 ms para oclusivas bilabiais, +30 ms a +40 ms para oclusivas velares e +35 ms para oclusivas apicais em posição inicial, associadas à vogal /a/ (Lisker e Abramson, 1970 *apud* Uselding, 1977). O “agrupamento perceptual” ao longo do *continuum* de VOT é visível nos dados de estudos de discriminação. Lisker e Abramson (1970 *apud* Uselding, 1977), por exemplo, descobriram que adultos podem discriminar facilmente entre duas sílabas CV que se localizam sobre uma fronteira de fonema (por exemplo, a fronteira entre /ba/ e /pa/), mas realizam apenas casualmente a tarefa de discriminação entre sílabas CV que representam variações acústicas na mesma categoria fonêmica (por exemplo, /ba/ com VOT de 0 (zero) ms *vs.* /ba/ com VOT de 10 ms).

Uma outra forma de se trabalhar com os mecanismos de percepção de VOT, que fornece informações bastante relevantes, é através da adaptação seletiva de detectores de aspectos de VOT (*VOT features detectors*). A adaptação seletiva é o processo que envolve a repetição de um segmento a fim de produzir uma mudança

---

<sup>5</sup> Aspas no original.

na fronteira fonêmica de uma série de segmentos que variam continuamente ao longo de algum parâmetro. Embora haja divergências nas explicações sobre o fenômeno, um grande número de estudos têm, consistentemente, concordado que há uma mudança em fronteira após vários períodos de adaptação (Sharf e Ohde, 1981).

Eimas e Corbit (1973 *apud* Uselding, 1977) fizeram uso de um procedimento de adaptação seletiva a fim de obter evidências para apoiar a noção de que existe um mecanismo de detecção de aspectos que é sensível a mudanças em VOT em segmentos de fala - basicamente um mecanismo dual. Parece haver dois detectores de aspecto, um para segmentos sonoros e outro para segmentos surdos, que são organizados de maneira antagônica: na ausência de um, o outro domina temporariamente. Utilizando segmentos CV, sinteticamente produzidos de oclusivas bilabiais e apicais em posição inicial, Eimas e Corbit mostraram uma mudança na fronteira de fonemas seguindo repetidas apresentações de cada série de oclusivas, sonoras ou surdas. Por exemplo, após adaptação com um segmento surdo, um sujeito tenderia a identificar segmentos próximos à fronteira do fonema como sonoro.

Posteriormente, Eimas, Cooper e Corbit (1973 *apud* Uselding, 1977) demonstraram que a susceptibilidade aos efeitos de adaptação não é igual para os detectores sonoros e surdos: "*A greater adaptation effect is found for the detector mediating perception of the voiceless stops*" (Eimas, Cooper e Corbit, 1973 *apud* Uselding, 1977: 368).

Baseado em trabalhos anteriores é que Uselding desenvolveu sua pesquisa, cujo objetivo era avaliar a existência de assimetria na detecção de mudanças de VOT por parte de adultos e, no caso de haver, qual seria a direção dessa assimetria.

No primeiro experimento, foram produzidos segmentos com valores de VOT de 0 (zero) e 20 ms, identificados como /ba/ e segmentos com valores de VOT de 40 ms, 60 ms e 80 ms, identificados como /pa/. Os sujeitos que participaram do experimento deveriam ouvir algumas séries de sons e detectar qualquer mudança nos sons em cada série. Os dados obtidos indicaram que os efeitos de diminuição e aumento nos valores de VOT foram assimétricos, pois mudanças eram mais corretamente detectadas quando havia uma diminuição de VOT através da fronteira

de +30 ms do que quando havia aumentos acima dessa fronteira. Essa assimetria já havia sido observada nos dados de Eimas, Cooper e Corbit (1973 *apud* Uselding, 1977) e apóia a teoria de que existe um mecanismo detector dual mediador para a percepção categórica do VOT, um para segmentos sonoros e outro para segmentos surdos.

O segundo experimento foi realizado para verificar se esse efeito assimétrico é geral ou limita-se apenas às oclusivas bilabiais. Foram empregadas então as oclusivas velares /g/ e /k/. Os segmentos apresentaram VOT de 0 (zero), 20 ms, 40 ms e 60 ms.

Os dados confirmaram os resultados já obtidos no Experimento 1, de que a percepção de VOT dos adultos é categórica e assimétrica, isto é, se dá mais facilmente entre dois segmentos que se localizam sobre uma fronteira de fonema do que entre dois segmentos que variam dentro da mesma categoria fonêmica. Mostraram também, como no Experimento 1, mais mudanças de segmentos surdos a sonoros do que mudanças de segmentos sonoros a surdos. Isso reflete uma maior susceptibilidade à adaptação para o detector de aspecto surdo, como já haviam mostrado Eimas *et al.* (1973).

Jamieson e Cheesman (1987) desenvolveram uma pesquisa cujo objetivo foi estudar o processo de adaptação seletiva na produção do VOT. Para tal, empreenderam três experimentos.

Como o contexto em que um som da fala é ouvido pode produzir mudanças significativas na percepção daquele som, após ouvir uma série de sons com VOT longo, como por exemplo [p<sup>h</sup>a] ou [t<sup>h</sup>a], os sons da fala com valores intermediários são ouvidos mais provavelmente como sonoros (cf. Eimas, Cooper e Corbit, 1973; Ohde, 1982 *apud* Jamieson e Cheesman, 1987).

Em trabalhos anteriores, Cooper, juntamente com vários pesquisadores, mostraram que ouvir repetidamente um som da fala pode alterar o VOT na produção de consoantes oclusivas (efeito de adaptação percepto-motora). Cooper (1974 *apud* Jamieson e Cheesman, 1987) mostrou que a média de VOT das produções de sílabas

[p<sup>hi</sup>] foi reduzida ao seguir adaptação perceptual com segmentos [p<sup>hi</sup>], ao passo que o VOT para a produção dos mesmos segmentos não foi afetado pela adaptação perceptual dos sons [i] ou [bi]. Em contrapartida, o VOT para as produções de [b] não foi afetado pelo uso dos segmentos adaptadores [p<sup>hi</sup>], [bi] ou [i].

Com o intuito de confirmar os resultados de trabalhos anteriores é que Jamieson e Cheesman realizaram seus experimentos. O primeiro deles envolveu a participação de falantes nativos do inglês canadense. Cada sujeito participou de três sessões, uma com cada um dos adaptadores - [p<sup>h</sup>a], [ba] ou [a]. Cada sessão consistiu da apresentação de 40 repetições do adaptador em questão. Os sujeitos deveriam produzir a sílaba [p<sup>h</sup>a] tão logo cada seqüência de adaptação fosse concluída.

Os resultados mostraram que o valor médio de VOT para a produção de [p<sup>h</sup>a] após a apresentação dos sons [a] e [ba] foi maior que o valor médio de VOT de [p<sup>h</sup>a] após a apresentação dos sons [p<sup>h</sup>a]. Esses resultados confirmaram a conclusão de Cooper (1979 *apud* Jamieson e Cheesman, 1987) de que ouvir consoantes oclusivas surdas pode levar a mudanças no VOT no que diz respeito à produção de oclusivas surdas. Ouvir oclusivas sonoras, no entanto, não influencia na produção de VOT de oclusivas surdas. Segundo Cooper (1979), ainda, a adaptação perceptual não altera o VOT na produção de oclusivas sonoras. Esses efeitos limitam-se à produção de oclusivas surdas.

No Experimento 2, após a conclusão de cada série de adaptação, os sujeitos deveriam pronunciar a sílaba [ba]. Os resultados revelaram que as médias de VOT para [ba] foram menores para as emissões que seguiram os sons [ba] do que para aquelas que seguiram os sons [a]. Posteriormente, testes de variância indicaram, entretanto, que o VOT não foi alterado na produção de [ba], nem após a apresentação do segmento [p<sup>h</sup>a], em comparação com o segmento [a], nem após a apresentação de [ba], em comparação com o segmento [a]. Esses resultados confirmaram, novamente, a conclusão de Cooper (1979), de que o VOT das oclusivas sonoras é resistente à alteração por adaptação.



No Experimento 3, foram empregados os sons [a] e [p<sup>h</sup>a]. Foram apresentadas 40 repetições de cada som. Os resultados revelaram que o valor médio de VOT para o som [p<sup>h</sup>a] produzido imediatamente após o segmento [p<sup>h</sup>a] foi menor do que o valor médio de VOT para [p<sup>h</sup>a], medido logo após a apresentação dos segmentos [a]. Essa diminuição dos valores médios de VOT subseqüentes a uma sessão de testes perceptuais, utilizando adaptadores surdos, confirma os resultados do Experimento 1 e os de Cooper (1974, 1979 *apud* Jamieson e Cheesman, 1987). A grande maioria dos sujeitos mostrou esse efeito, embora, como no Experimento 1, tivesse havido variação considerável em ambas as áreas de VOT e na magnitude da mudança de VOT que acompanha a adaptação de [p<sup>h</sup>a]. O efeito de adaptação na produção parece ter “vida curta”, uma vez que os valores de VOT para os sinais [p<sup>h</sup>a] emitidos 30 segundos após a adaptação não são estatisticamente distintos daqueles emitidos sem adaptação.

Carney, Gandour, Petty *et al.* (1988) investigaram a produção do VOT de oclusivas iniciais de palavras de um falante do tai que apresentava um quadro de perda acidental profunda da audição sensorio-neural.

Segundo Monsen (1976; 1978 *apud* Carney *et al.*, 1988), a produção do VOT é um importante correlato de inteligibilidade da fala na surdez.

Estudos têm mostrado que sujeitos que perderam a audição acidentalmente, ou seja, que apresentavam audição normal durante a fase adulta e em consequência de doença, ferimento ou ingestão de medicamentos tiveram sua audição profundamente prejudicada, representam uma prova interessante da necessidade de *feedback* auditivo para a manutenção da fala normal (cf. Carney *et al.*, 1988).

Zimmerman e Rettaliata (1981 *apud* Carney *et al.*, 1988), por exemplo, compararam as produções de um sujeito adulto, surdo acidentalmente, com as produções de um falante adulto com audição normal. Os resultados mostraram diferenças articulatórias consideráveis entre as produções dos dois falantes. Esses resultados sugeriram que diferenças fisiológicas nos movimentos articulatórios e no

*timing* nem sempre resultam em mudanças perceptuais. Os autores concluíram que o *feedback* auditivo pode funcionar como um instrumento de ajuste, reduzindo a variabilidade na produção quando esta excede um nível crítico.

Carney *et al.*, para sua investigação, gravaram um sujeito do sexo masculino, falante nativo do tai, de 43 anos de idade, que havia perdido a audição aos 18 anos em consequência de ingestão de medicamentos. O sujeito produziu uma série de palavras monossilábicas reais que incluíram as três distinções de sonorização para as oclusivas bilabiais (/b, p, p<sup>h</sup>/) e alveolares (/d, t, p<sup>h</sup>/) e as duas distinções de sonorização para as oclusivas velares (/k, k<sup>h</sup>/), características do tai. As produções deste sujeito foram comparadas às produções de sujeitos sem quaisquer problemas auditivos analisados em trabalhos anteriores (Gandour e Dardarananda, 1984; Gandour *et al.*, 1987 *apud* Carney *et al.*, 1988). Dez falantes nativos do tai com audição normal participaram dos testes de percepção.

Os resultados mostraram que as produções dos falantes normais das oito oclusivas do tai foram corretamente identificadas em 96% dos casos (Gandour *et al.*, 1987 *apud* Carney *et al.*, 1988). As produções de /p, p<sup>h</sup>, d, t, t<sup>h</sup>, k, k<sup>h</sup>/ do sujeito surdo foram corretamente identificadas e as produções de /b/ foram identificadas em 85% dos casos.

As distribuições dos valores de VOT associados à produção das oclusivas bilabiais, alveolares e velares para os sujeitos normais não apresentaram qualquer sobreposição entre as categorias, seja entre pré-sonorização e retardo curto, seja entre retardo curto e retardo longo. Os valores de VOT para todas as oclusivas agruparam-se em regiões bem definidas ao longo do *continuum* de VOT - os valores para /b, d/ localizaram-se na região de pré-sonorização, para /p, t, k/ na região de retardo curto e para /p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/ na região de retardo longo.

No que diz respeito às produções do sujeito surdo, embora a área total do *continuum* de VOT fosse um pouco mais estreita do que a área total dos sujeitos com audição normal, os valores de VOT para as oito consoantes oclusivas localizaram-se, em grande parte, em regiões distintas, ou seja, praticamente não apresentaram

sobreposição, com a exceção de /b/ e /p/ e /d/ e /t/. No caso das bilabiais, os sinais com valores de VOT atípicos foram aqueles que levaram a identificações incorretas. O mesmo não ocorreu com as alveolares, que foram corretamente identificadas.

Estes resultados levaram Carney *et al.* (1988) a concluírem que as produções das oclusivas do tai do sujeito com problemas auditivos indicam que houve somente pequenas perturbações articulatórias nas consoantes oclusivas iniciais de palavras, de forma que a percepção dos sujeitos com audição normal não fosse seriamente afetada. Contudo, há pequenas alterações na produção do VOT, e essas alterações são consistentes com as descobertas articulatório/fisiológicas de Zimmerman e Rettaliata (1981). Essa descoberta, porém, necessita ser confirmada através de estudos de um maior número de sujeitos com perda de audição na fase adulta a fim de determinar se é particular a um pequeno subgrupo de pacientes ou se é característico de uma população de pessoas acidentalmente surdas como um todo.

## Capítulo 2

# *Metodologia*

Ao longo desta seção serão descritos os métodos e técnicas adotados na obtenção do material sonoro empregado nas análises. Serão apresentados também os procedimentos aplicados no processo de registro sonoro dos dados.

### **2.1. Seleção de informantes**

Foram selecionados, para o presente estudo, quatro sujeitos, dois homens e duas mulheres, falantes nativos do português brasileiro (PB), com idades entre 20 e 40 anos, denominados SK, RL, IS e ES. É interessante ressaltar que, aparentemente, nenhum dos falantes apresenta qualquer tipo de problema articulatorio ou auditivo. Todos possuem nível superior e três deles trabalham na área de fonética.

## 2.2. Corpus

Para a montagem do *corpus*, foram empregados logatomas e palavras reais. Os segmentos seguiram, todos, a estrutura CV (consoante + vogal) e consistiram das oclusivas /b, d, g/ e /p, t, k/ e das vogais /e, a, o/ do PB.

O objetivo, ao se empregar três vogais diferentes, foi verificar se o contexto vocálico influenciaria os valores de VOT, uma vez que há fortes evidências de que a informação acústica que especifica um determinado segmento fonético varia extensivamente com uma série de fatores contextuais, devido às implementações articulatória e fonética da fala serem relativamente imprecisas. Esses fatores podem ser falante, velocidade de fala (Miller, 1994; Kessinger e Blumstein, 1997), posição silábica, tonicidade e contexto vocálico (Lisker e Abramson, 1967 e Klatt, 1975 *apud* Weismer, 1979). No que diz respeito ao efeito do contexto vocálico sobre o VOT, no entanto, há divergências nos estudos de Lisker e Abramson (1967) e Klatt (1975) (*apud* Weismer, 1979). Lisker e Abramson, ao medirem as oclusivas /b, d, g/ e /p, t, k/ do inglês, em combinação com doze núcleos silábicos, concluíram que a vogal não exerce influência sistemática sobre o VOT. Klatt, porém, mostrou uma diferença estatisticamente significativa nos valores de VOT para as oclusivas surdas, dependendo de a oclusiva ser seguida por vogal alta ou média. Sugeriu, então, que a não detecção de influência vocálica sobre o VOT em Lisker e Abramson (*apud* Weismer, 1979) pudesse estar relacionada ao escopo limitado de dados de seu experimento.

Em concordância com o trabalho de Klatt estão os resultados de Weismer (1979) que, estudando as oclusivas surdas do inglês seguidas por vogais tensas e relaxadas, demonstrou, também, que o VOT é modificado em função da vogal seguinte: o valor médio de VOT é maior quando a vogal é tensa do que quando é relaxada.

A escolha das vogais médias fechadas anterior e posterior e central baixa, e não das vogais altas anterior e posterior (como a central baixa, extremos do trapézio

vocálico), deu-se, principalmente, pelo fato de que, como resultado da combinação da vogal /i/ com as oclusivas velares /d/ e /t/, obteríamos, na maioria dos casos, as africadas /dz/ e /tʃ/. Da mesma forma, as vogais médias abertas anterior e posterior foram descartadas por não aparecerem em contexto átono, isto no que se refere às palavras.

Nossa pesquisa, então, dividiu-se em dois estudos-piloto. O primeiro foi constituído de logatomas, todos monossilábicos, produzidos de duas formas: isoladamente e inseridos em frases-veículo do tipo *Digo \_\_\_\_\_ pra ela*. Cabe observar que o objetivo de se inserir os logatomas em frases-veículo foi confirmar, para o PB, as descobertas de Lisker e Abramson (1964) de que palavras inseridas em sentenças terão seu VOT reduzido em relação a segmentos produzidos isoladamente. Por outro lado, decidiu-se por frases-veículo e não sentenças, por terem as primeiras a vantagem de reduzir consideravelmente alguns efeitos de natureza prosódica, como por exemplo, entonação, ou ainda, o efeito conhecido como “forma de citação”. Com esse primeiro experimento verificaremos também a influência do contexto vocálico sobre o VOT.

O segundo estudo-piloto, constituído de palavras produzidas isoladamente e inseridas em diferentes posições em sentenças, tem como objetivo verificar a influência da tonicidade da sílaba e da palavra e de diferentes posições da palavra na sentença sobre o VOT.

No que tange às palavras, foram elaborados testes estatísticos (*t-de-Student*), que não mostraram influência do número de sílabas sobre o VOT. Dessa forma, foram empregadas tanto palavras dissilábicas quanto trissilábicas e polissilábicas. Procurou-se estabelecer uma lista de vocábulos bastante comuns, embora, algumas vezes, tenha-se tornado necessário empregar palavras não usuais ao vocabulário cotidiano.

O contexto vocálico levado em consideração para a análise do VOT das oclusivas nas palavras foi o mesmo dos logatomas. Controlou-se, além da vogal, a tonicidade da sílaba, ou seja, as posições pretônica, tônica e postônica.

O passo seguinte consistiu em inserir as palavras em sentenças, cujas estruturas variaram de acordo com a posição das palavras dentro das mesmas. Nesse caso, foram analisados somente os segmentos tônicos. Tentou-se, com isso, verificar uma possível influência da posição da palavra na frase sobre o VOT. As sentenças foram elaboradas de forma a apresentarem, todas, 14 sílabas e, a palavra-alvo, sempre trissílaba, ocupou as seguintes posições:

- a) posição inicial zero (0), por exemplo, *Tapada pega aquela mulher do seu José Carlos.;*
- b) posição inicial (1) (seguindo um determinante - os artigos definidos feminino ou masculino), por exemplo, *A tapada pega na mulher do seu José Carlos.;*
- c) posição medial, em que o segmento-alvo é intercalado na sentença após a 8ª sílaba, por exemplo, *Aquela mulher era tapada demais da conta.;*
- d) posição final de sentença, por exemplo, *O senhor José Carlos tinha uma mulher tapada.*

### 2.3. Coleta dos dados

Os dados foram organizados em fichas e apresentados aleatoriamente a cada informante. Em uma primeira fase, foram apresentados os logatomas isolados, a seguir, os logatomas inseridos em frases-veículo, seguidos, então, pelas palavras produzidas isoladamente e, por fim, das palavras nas sentenças.

Os logatomas isolados, os logatomas inseridos em frases-veículo e as palavras isoladas foram lidos dez vezes por cada informante e as palavras inseridas nas sentenças, cinco vezes. A leitura dos segmentos deveria proceder da forma mais natural e o mais próximo possível da fala normal de cada sujeito.

## 2.4. Registro e análise dos dados

As gravações dos dados e análises posteriores foram realizadas em um microcomputador com dispositivo de análise e síntese de fala - o Computer Speech Lab (CSL Modelo 4300 B da Kay Elemetrics), pertencente ao Laboratório de Fonética Acústica do Departamento de Língua e Literatura Vernáculas da Universidade Federal de Santa Catarina. A sala de gravação não apresentou nenhum tratamento acústico. Mesmo assim, não foram observados ecos, reverberações ou ruídos de fundo. O microfone utilizado é unidirecional, modelo *Shure SM 48I*, e os informantes receberam instruções de mantê-lo a uma distância de pelo menos 15 cm da boca.

Os valores de VOT foram obtidos medindo-se o intervalo entre a soltura da oclusão e o início de vibração das pregas vocais ou, sonorização. Utilizando-se a forma de onda, vinculada ao espectrograma, mediu-se os valores de VOT para as oclusivas com retardo curto, em milissegundos (ms), do início da explosão da oclusiva até o início de sonorização, valores positivos de VOT. Para as oclusivas com pré-sonorização, o VOT foi medido do início de sonorização até o início da explosão da oclusiva, daí os valores negativos de VOT. O ponto de soltura ou, explosão, pode ser localizado, na forma de onda, no caso de oclusiva surda (Figura 2.1 à direita), no ponto em que a energia (praticamente zero durante a oclusão) (Istre, s/d) mostra uma mudança significativa em amplitude, enquanto o início de sonorização se dá no momento em que começam os pulsos glotais para a realização da vogal seguinte. No caso de oclusiva sonora (Figura 2.1 à esquerda), a medição se dá de maneira diferente, ou seja, é feita a partir da sonorização já existente durante o intervalo de fechamento da oclusiva, manifestada por energia de baixa-freqüência até a soltura que, nesse caso, é posterior à sonorização.



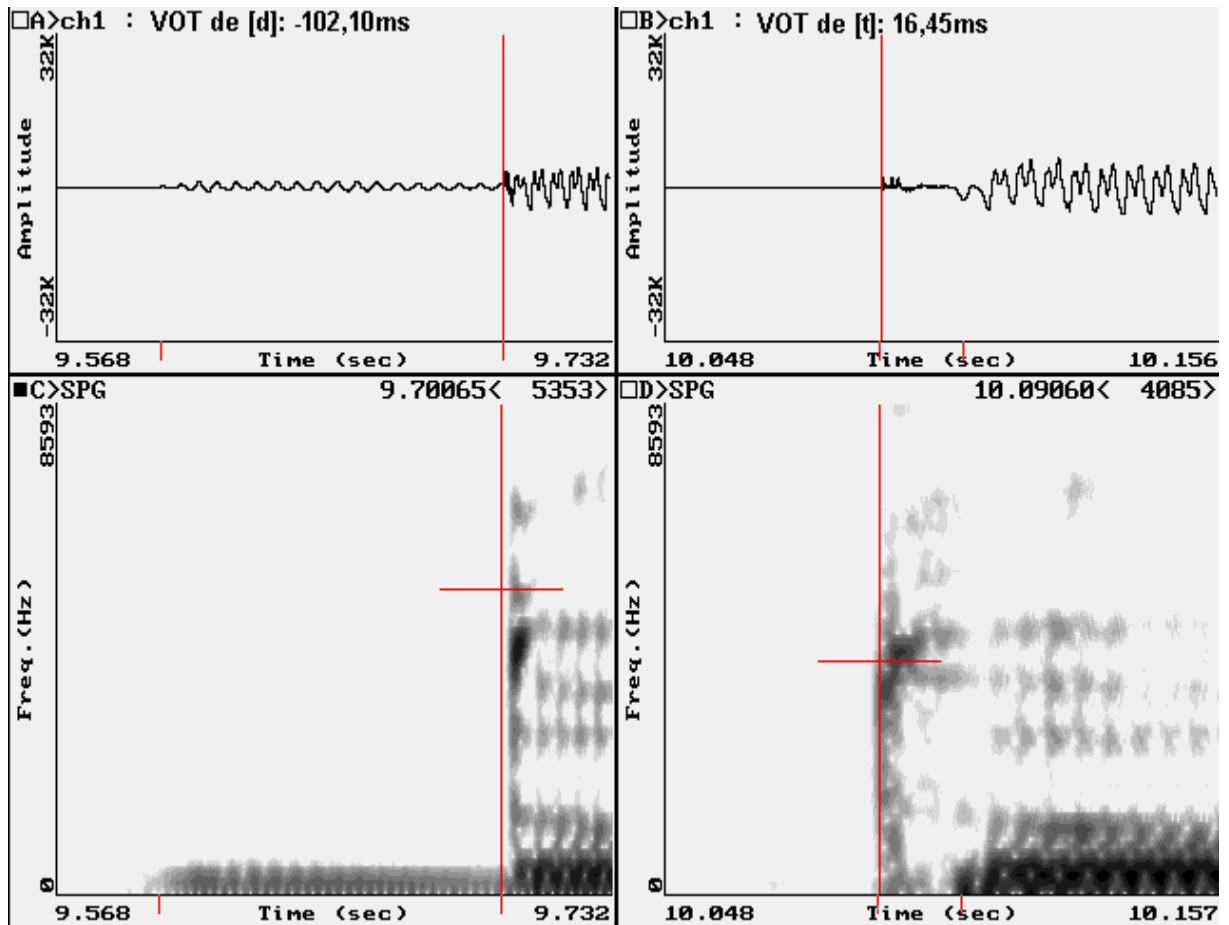


Figura 2.1 – Formas de onda e espectrogramas das áreas de pré-sonorização e retardo curto das consoantes oclusivas [d] e [t] nos segmentos [do] e [to].

## 2.5. Parâmetros estatísticos utilizados na análise

Nesta pesquisa, utilizamos testes de influência estatística como a média, o desvio-padrão e o teste *t-de-Student*. Média é o ponto de qualquer distribuição em torno do qual se equilibram diferenças positivas e negativas. Desvio-padrão é uma medida de dispersão que serve para verificar o quanto os dados coletados variam em torno da média, ou seja, representa a “variabilidade média” de uma distribuição. Quanto maior a variabilidade em torno da média de uma distribuição, maior o desvio-padrão.

Em análises estatísticas, geralmente se estabelece duas hipóteses: a *hipótese nula* e a *hipótese experimental*. A *hipótese nula* afirma “terem duas amostras sido extraídas

da mesma população”<sup>1</sup> (Levin, 1987: 145), ou seja, uma diferença entre duas médias amostrais nesse caso não apresentaria uma verdadeira diferença entre as médias populacionais. Esta é estabelecida com a esperança de ser rejeitada, já que o que se busca são relações entre variáveis e o que se espera são provas muito mais de que elas existam do que não existam. Rejeitando-se a hipótese nula, aceitamos a outra hipótese, a *experimental*, que afirma existir uma verdadeira diferença de valores, e diz que essa diferença obtida da comparação entre médias amostrais é grande demais para ser explicada apenas como erro amostral. Quando se rejeita a hipótese nula significa que os dados observados não dão evidência suficiente para afirmar que existem diferenças entre os valores analisados. Ou seja, isso não quer dizer que não haja realmente diferenças entre esses valores. Já, quando se aceita a hipótese experimental, significa que os dados observados dão evidência suficiente para dizer que os valores, por exemplo, de VOT, têm diferenças que são determinadas pelos variados contextos em que a oclusiva se insere.

Para se decidir se esta diferença amostral obtida é estatisticamente significativa, estabelece-se um *nível de confiança* (também chamado de *nível de significância*), que representa a probabilidade com que a hipótese nula pode ser rejeitada com segurança, ou seja, a probabilidade com que a hipótese experimental pode ser aceita com confiança.

Para a presente pesquisa, optamos por um nível de significância de 0,05, isto é, rejeitamos a hipótese nula se a diferença amostral obtida ocorrer somente cinco vezes ou menos em 100, 5% no máximo. Como a maior parte de nossas amostras continha em torno de 30 dados, optamos por utilizar o teste *t-de-Student*, teste de tomada de decisões habitualmente empregado para pequenas amostras, já que esse teste, para esse número de dados, ainda permanece válido mesmo que a variável não tenha uma distribuição normal (Barbetta, 1998).

Para a interpretação do teste de tomada de decisão, necessitamos ainda do *grau de liberdade* (liberdade de variação em um conjunto de escores) que é dado por  $gl=n-1$ , onde

---

<sup>1</sup> Grifo no original.

$gl$  é o grau de liberdade e,  
 $n$  é o número de dados da amostra.

Assim, com a ajuda de um programa de análise estatística, determinamos a média aritmética, o desvio padrão e o teste *t-de-Student* com nível de significância de 0,05. Nosso objetivo era verificar se as diferenças entre os valores de VOT em diferentes contextos vocálicos, diferentes contextos de tonicidade e diferentes posições de uma palavra inserida em sentenças seriam estatisticamente significantes, ou seja, que existiria uma verdadeira diferença entre os valores médios de VOT nessas diferentes circunstâncias que não eram devido a erros amostrais, mas sim, a diferenças circunstanciais. Nesse caso, pretende-se rejeitar a hipótese nula com grau de confiança de 95%, ou seja, 95% de probabilidade de se estar acertando.

Os testes estatísticos apresentaram, então, como diferenças estatisticamente significativas entre os valores médios de VOT:

1) o contexto vocálico para a área de pré-sonorização das consoantes oclusivas bilabiais, dentais e velares em logatomas inseridos em frases-veículo e em sílabas tônicas inseridas em palavras; e somente no caso de inserção em logatomas isolados para as oclusivas dentais e velares; para a área de retardo curto das consoantes oclusivas bilabiais, dentais e velares, tanto em logatomas produzidos isoladamente, em logatomas inseridos em frases-veículo, quanto em sílabas tônicas inseridas em palavras;

2) o contexto de tonicidade para a área de pré-sonorização e retardo curto das consoantes oclusivas bilabiais, dentais e velares;

3) o veículo que contém o estímulo para a área de pré-sonorização das oclusivas dentais e velares entre logatomas isolados e logatomas inseridos em frases-veículo e entre logatomas produzidos isoladamente e sílaba tônica inserida em palavra; para a bilabial sonora somente entre logatomas isolados e logatomas inseridos em frases-veículo; para a área de retardo curto da consoante oclusiva bilabial entre logatomas isolados e logatomas inseridos em frases-veículo e entre logatomas isolados e sílaba tônica inserida em palavra, e, para a velar, somente entre logatomas produzidos isoladamente e logatomas inseridos em frases-veículo;

4) as diferentes posições nas sentenças para a área de pré-sonorização da oclusiva bilabial entre a posição medial e final; para a área de retardo curto das oclusivas dentais e velares também entre a posição medial e final e das oclusivas bilabiais entre as posições inicial zero e 1 em relação à medial.

## *Análise e discussão dos dados*

### 3.1. Introdução

O português brasileiro, assim como o espanhol, o holandês, o húngaro e o tâmil, é uma língua de duas categorias, pois faz distinção entre consoantes oclusivas sonoras e surdas. Similarmente a essas línguas, suas oclusivas sonoras são caracterizadas pela área de pré-sonorização e as surdas, pela área de retardo curto.

A distinção entre oclusivas sonoras e surdas pode ser representada como segue:

Labial	b	p
Dental	d	t
Velar	g	k

Para uma melhor análise dos dados de nossos experimentos, apresentaremos os resultados divididos em duas seções, uma referente à área de pré-sonorização e a outra, referente à área de retardo curto.

### 3.2. Pré-sonorização

Como definiram Lisker e Abramson (1964), na categoria de pré-sonorização - que caracteriza as oclusivas sonoras na maioria das línguas analisadas pelos autores - o início da voz se dá anteriormente à soltura da oclusão, resultando, por conseguinte, em valores negativos de VOT, conforme Figura 3.2.1, que ilustra a área de pré-sonorização da oclusiva sonora /g/.

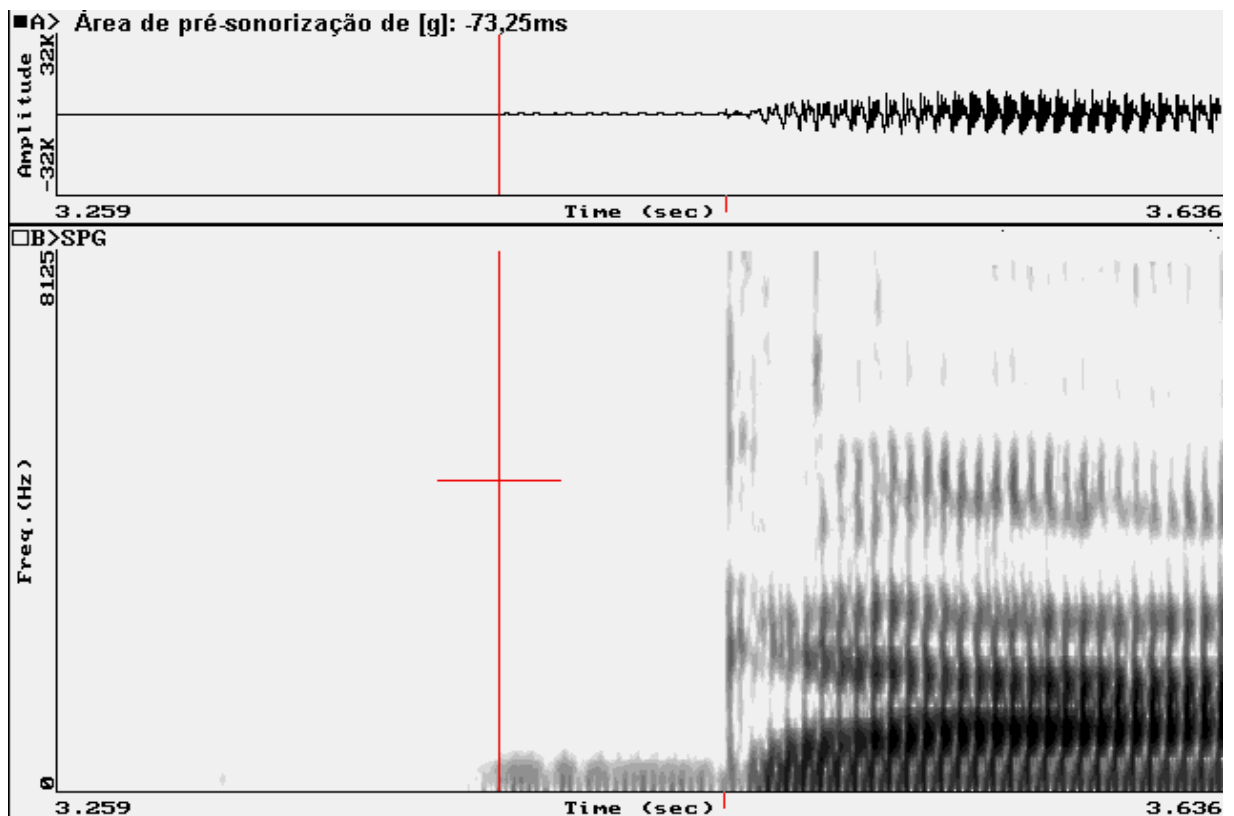


Figura 3.2.1 – Forma de onda e espectrograma da área de pré-sonorização de [g] no segmento [ga]

Em uma análise mais geral, os valores de VOT para a área de pré-sonorização obtidos para as oclusivas sonoras do PB variaram, para a oclusiva bilabial [b], de

-162,75 a -26,70 ms, obtendo-se um valor médio de -92,27 ms; para a oclusiva dental [d], os valores distribuíram-se em uma área de -167,95 a -31,90 ms, obtendo-se um valor médio de -92,07 ms; para a oclusiva velar [g], os valores de VOT com precedência de sonorização variaram em uma faixa de -172,50 a -22,75 ms, obtendo-se como valor médio, -78,20 ms.

Comparando nossos dados aos de Istre (s/d), parece mais razoável classificar as consoantes oclusivas do português brasileiro como tendo uma precedência longa de sonorização, concordando com a hipótese inicial desse autor, que foi, no entanto, rejeitada por seus resultados, mostrando as oclusivas sonoras do PB com uma precedência média de sonorização (cf. Gráfico 3.2.1). Istre apresenta um valor médio de VOT de -39,45 ms para a oclusiva bilabial, -60,59 ms para a dental e -53,69 ms para a velar. As médias de VOT encontradas em nossa pesquisa para as oclusivas bilabial e dental mostram que não há, praticamente, diferenças nos valores entre as duas classes de oclusivas, haja vista a área de pré-sonorização de ambas encontrar-se na mesma faixa, em torno de -92,00 ms.

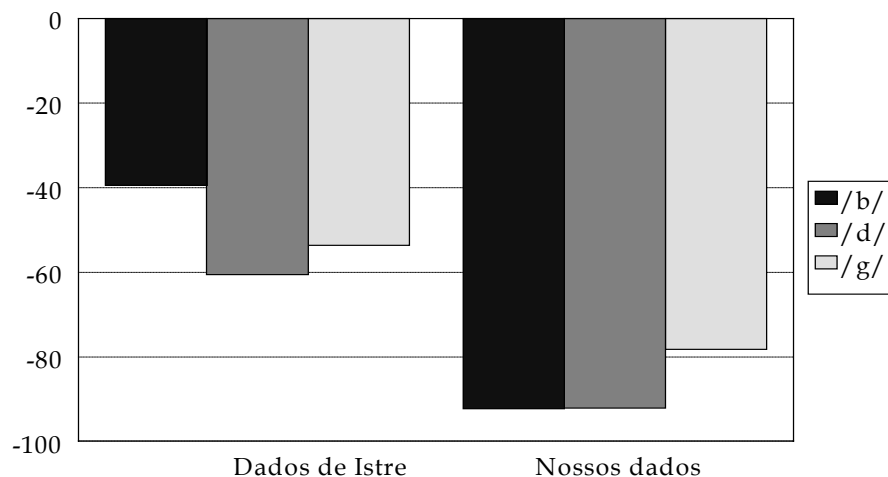


Gráfico 3.2.1 - Comparação da área de pré-sonorização do PB em Istre e em nossos dados.

A despeito da similaridade apresentada pela área de pré-sonorização que caracteriza as oclusivas bilabial e dental, o PB apresenta dados compatíveis aos apresentados pelas diversas línguas analisadas, conforme Gráfico 3.2.2. Fica claro, a

partir desse gráfico, que, embora as diferentes línguas apresentem a categoria de pré-sonorização em comum, há grande variabilidade entre elas, o que sugere, como já dito anteriormente, ocorrerem ajustes específicos às línguas. Por exemplo, o espanhol difere do PB pelo fato de a área de pré-sonorização do primeiro ser consideravelmente maior do que a mesma área para o segundo. No entanto, o PB assemelha-se muito ao holandês e ao húngaro quanto à extensão da área de pré-sonorização das oclusivas bilabial e dental.

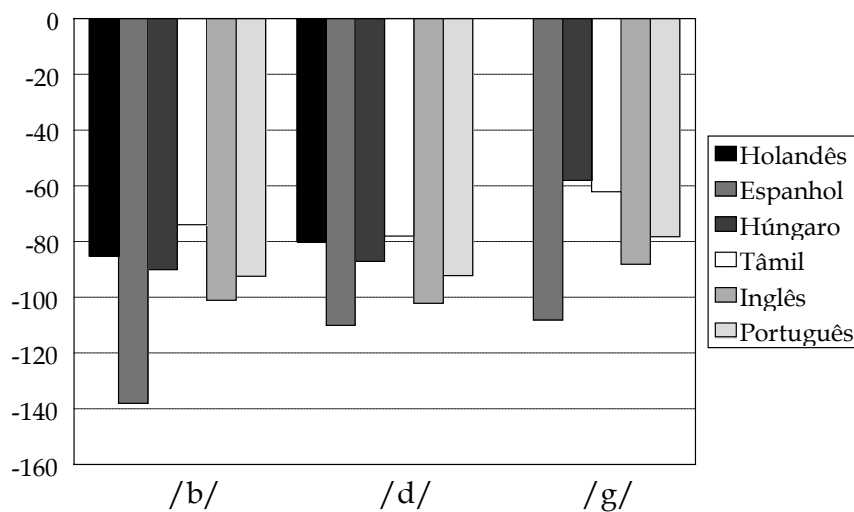


Gráfico 3.2.2 - Comparação da área de pré-sonorização em PB e outras línguas (Lisker e Abramson, 1964)

Conforme citado no Capítulo 2 - Metodologia, nossa pesquisa foi dividida em dois estudos-piloto. O primeiro buscando analisar o comportamento do VOT diante de diferentes contextos vocálicos e em inserções de oclusivas em logatomas produzidos isoladamente e inseridos em frases-veículo e, o segundo, buscando observar se o VOT variaria significativamente em diferentes contextos de tonicidade e diferentes posições em sentenças, verificando esse comportamento em sílabas tônicas e átonas em palavras isoladas, e em sílaba tônica de palavras inseridas em diferentes posições em sentenças.



## Resultados referentes ao primeiro estudo-piloto

### 3.2.1 Oclusivas bilabiais

A área de pré-sonorização para a oclusiva bilabial do PB apresentou variação significativa nos diferentes contextos testados em nosso estudo, ou seja, os valores de VOT dessa oclusiva variaram segundo estivessem inseridas em logatomas produzidos isoladamente ou em logatomas inseridos em frases-veículo.

A Tabela 3.2.1 nos mostra que a média de VOT desta consoante inserida em logatomas isolados foi de -86,57 ms e em logatomas inseridos em frases-veículo, de -104,98 ms. A faixa de valores apresentada por [b] inserido em logatomas isolados variou de -162,65 ms a -26,70 ms e de logatomas inseridos em frases-veículo, de -162,30 ms a -76,95 ms.

Comparando os segmentos inseridos em logatomas isolados e em logatomas inseridos em frases-veículo, podemos verificar que os valores médios de VOT medidos em logatomas produzidos isoladamente são, ao contrário do que supúnhamos, significativamente menores do que os valores em logatomas inseridos em frases-veículo (Tabela 1 do apêndice). Isso vai de encontro à descoberta de Lisker e Abramson (1964), de que, tanto oclusivas com precedência de sonorização quanto aquelas com retardo - seja curto ou longo -, inseridas em contextos maiores, tendem a apresentar valores de VOT reduzidos em comparação aos valores de segmentos medidos em palavras produzidas isoladamente.

	<b>logatoma</b>	<b>logatoma/FV</b>
<b>Média</b>	-86,57	-104,98
<b>Área observada</b>	-162,75 : -26,70	-162,30 : -76,95
<b>Nº de sinais</b>	96	96

Tabela 3.2.1 - Média geral de VOT e área observada em ms da oclusiva [b] em logatomas produzidos isoladamente e logatomas inseridos em frases-veículo

### 3.2.2. Oclusivas Dentais

A categoria de pré-sonorização da oclusiva dental [d], conforme mostra a Tabela 2 do apêndice, apresentou reduções significativas como consequência de sua inserção em logatomas inseridos em frases-veículo, comparados à oclusiva inserida em logatomas produzidos isoladamente. Enquanto a média de VOT para essa consoante inserida em logatomas isolados foi de -98,84 ms, o valor médio de VOT para a mesma oclusiva em logatomas inseridos em frases-veículo foi de -93,31 ms, conforme Tabela 3.2.2.

Assim, no caso da oclusiva dental, os resultados corroboraram a descoberta de Lisker e Abramson (1964), bem como nossa hipótese inicial.

	<b>logatoma</b>	<b>logatoma/FV</b>
<b>Média</b>	-98,84	-93,31
<b>Área observada</b>	-167,95 : -39,85	-138,45 : -52,75
<b>Nº de sinais</b>	96	93

Tabela 3.2.2 - Média geral de VOT e área observada em ms da oclusiva [d] em logatomas produzidos isoladamente e logatomas inseridos em frases-veículo

### 3.2.3. Oclusivas Velares

A oclusiva velar sonora [g] apresentou o mesmo comportamento da oclusiva dental, mostrando redução significativa no VOT à medida que o segmento era inserido em contextos mais longos (Tabela 3 no apêndice). O valor médio da área de pré-sonorização para essa consoante em logatomas inseridos em frases-veículo foi significativamente menor do que aquele para a oclusiva inserida em logatomas produzidos isoladamente, -76,93 ms e -91,16 ms, respectivamente.

A variação na área de pré-sonorização para a oclusiva velar foi de -172,50 a -22,75 ms para o segmento inserido em logatomas produzidos isoladamente e de -130,90 a -35,45 ms para aquele em logatomas inseridos em frases-veículo, como podemos observar na Tabela 3.2.3.

	<b>logatoma</b>	<b>logatoma/FV</b>
<b>Média</b>	-91,16	-76,93
<b>Área observada</b>	-172,50 : -22,75	-130,90 : -35,45
<b>Nº de sinais</b>	92	92

Tabela 3.2.3 - Média geral de VOT e área observada em ms da oclusiva [g] em logatomas produzidos isoladamente e logatomas inseridos em frases-veículo

### 3.2.4. O VOT e o contexto vocálico

Como mostra a Tabela 25, no apêndice, a oclusiva bilabial, quando inserida em logatomas produzidos isoladamente (Tabela 3.2.4), não apresentou diferenças significativas na área de pré-sonorização para nenhuma das vogais que a seguiam. Sua média de VOT diante de [a] foi de -91,84 ms, diante de [e] foi de -81,43 ms e diante de [o], de -87,91 ms.

Diferentemente do exposto acima, quando a oclusiva [b] foi inserida em logatomas produzidos em frases-veículo (Tabela 3.2.5), houve diferenças significativas em função da vogal seguinte: a média de VOT de [b] em contexto vocálico [o] foi significativamente menor do que as médias de VOT diante da vogal baixa [a] e da vogal média anterior [e], conforme Tabela 31, apêndice.

	<b>[a]</b>	<b>[e]</b>	<b>[o]</b>
<b>Média</b>	-91,84	-81,43	-87,92
<b>Área observada</b>	-149,65 : -40,75	-129,50 : -44,90	-150,65 : -26,70
<b>Nº de sinais</b>	31	32	32

Tabela 3.2.4 - Média de VOT e área observada em ms para a oclusiva [b] diante das vogais [a], [e] e [o] para os logatomas produzidos isoladamente

	<b>[a]</b>	<b>[e]</b>	<b>[o]</b>
<b>Média</b>	-106,88	-108,80	-99,25
<b>Área observada</b>	-137,65 : -82,95	-162,30 : -78,55	-122,10 : -76,95
<b>Nº de sinais</b>	32	32	32

Tabela 3.2.5 - Média de VOT e área observada em ms para a oclusiva [b] diante das vogais [a], [e] e [o] para os logatomas inseridos em frases-veículo

A oclusiva dental [d], inserida em logatomas produzidos isoladamente (Tabela 3.2.6), apresentou valor médio de VOT significativamente maior diante da vogal média anterior [e] em relação à vogal média posterior [o] e à baixa [a]. Quando essa oclusiva era inserida em logatomas produzidos em frases-veículo (Tabela 3.2.7), foram as vogais médias anterior e posterior que apresentaram VOT significativamente maior em relação à vogal baixa (Tabelas 26 e 32 no apêndice).

	[a]	[e]	[o]
<b>Média</b>	-89,15	-110,51	-96,85
<b>Área observada</b>	-155,05 : -50,20	-167,95 : -43,20	-131,55 : -39,85
<b>Nº de sinais</b>	32	32	32

Tabela 3.2.6 - Média de VOT e área observada em ms para a oclusiva [d] diante das vogais [a], [e] e [o] para os logatomas produzidos isoladamente

	[a]	[e]	[o]
<b>Média</b>	-85,15	-100,42	-95,02
<b>Área observada</b>	-123,30 : -54,15	-135,00 : -51,35	-138,45 : -54,85
<b>Nº de sinais</b>	32	29	32

Tabela 3.2.7 - Média de VOT e área observada em ms para a oclusiva [d] diante das vogais [a], [e] e [o] para os logatomas inseridos em frases-veículo

A oclusiva velar, para os segmentos inseridos em logatomas produzidos isoladamente (Tabela 3.2.8), mostrou diferenças estatisticamente significativas da vogal média anterior [e] em relação à vogal posterior [o], conforme Tabela 27 no apêndice. No caso de logatomas inseridos em frases-veículo, as alterações ocorreram da vogal baixa [a] em relação às vogais médias, conforme Tabela 33, apêndice.

	[a]	[e]	[o]
<b>Média</b>	-88,46	-99,14	-82,38
<b>Área observada</b>	-172,50 : -22,75	-167,25 : -40,15	-143,15 : -35,00
<b>Nº de sinais</b>	31	32	32

Tabela 3.2.8 - Média de VOT e área observada em ms para a oclusiva [g] diante das vogais [a], [e] e [o] para os logatomas produzidos isoladamente

	[a]	[e]	[o]
<b>Média</b>	-68,11	-81,73	-79,84
<b>Área observada</b>	-120,65 : -35,45	-130,90 : -53,60	-129,50 : -40,20
<b>Nº de sinais</b>	28	32	32

Tabela 3.2.9 - Média de VOT e área observada em ms para a oclusiva [g] diante das vogais [a], [e] e [o] para os logatomas inseridos em frases-veículo

Pelos dados acima descritos, podemos concluir então, que a área de pré-sonorização das oclusivas sonoras do PB sofre influências do contexto vocálico, porém, essas alterações são assimétricas através dos diferentes pontos de articulação.

Em relação à colocação de Lisker e Abramson (1964), de que os valores de VOT apresentariam redução quando inseridos em contextos maiores foi confirmada para a oclusiva bilabial diante das vogais [a], [e] e [o] (Tabelas 7, 8 e 9 no apêndice), para a velar diante de [a] e [e] (Tabelas 13 e 14 do apêndice), e para a dental somente diante de [e] (Tabela 11, apêndice).

## Resultados referentes ao segundo estudo-piloto

### 3.2.5. O VOT e a tonicidade da sílaba

Segundo a Tabela 3.2.10, no caso da oclusiva bilabial, a média de VOT foi significativamente maior para a oclusiva inserida em contexto tônico - 85,30 ms - comparada aos contextos átonos, que apresentaram valores médios de -72,27 ms e -59,01 ms para as posições pós e pretônicas, respectivamente. Pela Tabela 37 do apêndice, podemos verificar que as diferenças foram estatisticamente significativas para todos os contextos de tonicidade.

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-85,30	-59,01	-72,27
<b>Área observada</b>	-126,45 : -28,05	-120,15 : -21,50	-103,65 : -46,00
<b>Nº de sinais</b>	91	77 <sup>1</sup>	91

Tabela 3.2.10 - Média geral de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico

Podemos observar o mesmo comportamento para a oclusiva dental [d] (conforme Tabela 3.2.11), que, em contexto tônico, apresentou valor médio de VOT de -85,99 ms; em contexto postônico, de -74,20 ms e, em contexto pretônico, de -54,58 ms. As alterações foram, como no caso da oclusiva bilabial, estatisticamente significativas entre todos os contextos de tonicidade.

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-85,99	-54,58	-74,20
<b>Área observada</b>	-137,85 : -31,90	-103,90 : -24,95	-106,55 : -46,45
<b>Nº de sinais</b>	91	91	84 <sup>2</sup>

Tabela 3.2.11 - Média geral de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico

No caso da oclusiva velar, o valor médio de VOT foi significativamente mais alto na posição tônica, -66,81 ms. Com relação às posições pré e postônicas, não houve alterações significativas nos valores médios entre as duas posições. O valor médio de VOT para a velar na posição pretônica foi de -51,37 ms e na postônica, de -51,12 ms (conforme Tabela 39 no apêndice). A faixa de variação dos valores de VOT pode ser vista na Tabela 3.2.12:

<sup>1</sup> Em posição pretônica, muitos dos sinais, principalmente aqueles em contexto vocálico [a] e [e], não apresentaram explosão.

<sup>2</sup> O número de sinais [do] em contexto postônico é bastante reduzido em relação aos contextos tônico e pretônico porque, nesse caso, alguns falantes alternaram produção de [do] e [du].

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-66,81	-51,37	-51,12
<b>Área observada</b>	-132,00 : -25,25	-110,00 : -20,95	-83,95 : -23,45
<b>Nº de sinais</b>	85	85	83

Tabela 3.2.12 - Média geral de VOT e área observada em ms para a oclusiva [g]<sup>3</sup> nos contextos pretônico, tônico e postônico

Nossos dados revelaram uma simetria no comportamento do VOT com relação à tonicidade da sílaba, haja vista as três oclusivas sonoras terem apresentado valores médios de VOT maiores no contexto tônico. No entanto, quanto aos valores concernentes às oclusivas nas posições átonas, as oclusivas bilabial e dental apresentaram valores de VOT para a área de pré-sonorização significativamente maiores na posição postônica do que na pretônica.

Analisando nossos dados segundo tonicidade e contexto vocálico, verificamos que a área de pré-sonorização, para a oclusiva bilabial, apresentou alterações significativas nos valores de VOT para as três vogais, sendo que na posição tônica os valores foram significativamente maiores para as vogais [a] e [o] (Tabelas 43 e no apêndice). No caso da vogal baixa [a] e da vogal média posterior [o], houve alterações significativas entre as três posições em questão, ou seja, entre as posições tônica, pré e postônicas. Para a bilabial diante da vogal média anterior, essa relação não é mais consistente, já que se mantém a diferença de valores médios de VOT somente entre as posições tônica e postônica em relação à posição pretônica, conforme Tabela 44 do apêndice.

<sup>3</sup> O número de sinais [g] no caso das palavras é consideravelmente menor em comparação ao número apresentado pelas outras oclusivas pelo fato de inúmeras realizações não terem apresentado explosão, além de ter havido, em alguns desses casos, sonorização contínua do som sonoro precedente. Isso já havia sido observado nos dados de Lisker e Abramson (1964), porém, quando da inserção de palavras em sentenças maiores: “...for categories with voicing-lead, we find that in non-initial position voicing usually proceeds unbroken from a preceding voiced environment into the closure interval; any interruption of glottal buzz depends on what sounds occur before the stop.” (Lisker e Abramson, 1964: 413). A maioria dos sinais em nosso trabalho ocorre em posição não inicial, mesmo quando se tratando das posições átonas. Câmara Jr. declara que “as oclusivas intervocálicas, em português, apresentam uma articulação um tanto enfraquecida pelo ambiente vocálico em cujo meio se acham.” (Câmara Jr., 1998: 49)

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-95,55	-56,87	-71,25
<b>Área observada</b>	-121,50 : -67,85	-96,90 : -23,55	-94,80 : -50,00
<b>Nº de sinais</b>	31	21 <sup>4</sup>	31

Tabela 3.2.13 - Média de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico, diante da vogal [a]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-74,76	-59,41	-74,41
<b>Área observada</b>	-111,05 : -28,05	-120,15 : -26,25	-103,65 : -58,20
<b>Nº de sinais</b>	30	25	30

Tabela 3.2.14 - Média geral de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico, diante da vogal [e]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-85,71	-60,70	-72,23
<b>Área observada</b>	-126,45 : -44,05	-100,55 : -21,50	-102,00 : -46,00
<b>Nº de sinais</b>	29	30	30

Tabela 3.2.15 - Média geral de VOT e área observada em ms para a oclusiva [b] nos contextos pretônico, tônico e postônico, diante da vogal [o]

Para as oclusivas dentais, ocorrem também alterações significativas para todas as vogais, conforme Tabelas 46, 47 e 48 no apêndice. Nesse caso, os valores de VOT são significativamente maiores na posição tônica para a vogal baixa [a] e para a vogal média anterior [e]. No caso da vogal média posterior [o], o valor médio de VOT foi significativamente maior para a posição postônica em relação à pretônica, embora, em termos estatísticos, não tivesse havido diferenças significativas entre as posições postônica e tônica. No caso de [o], essas diferenças estatisticamente significativas ocorreram das posições postônica e tônica em relação à posição pretônica. Entre as posições átonas, os maiores valores de VOT foram apresentados pela oclusiva em posição postônica.

<sup>4</sup> Conforme nota de rodapé (1), página 57.



	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-80,81	-50,02	-75,93
<b>Área observada</b>	-105,00 : -39,20	-62,15 : -37,05	-91,85 : -58,75
<b>Nº de sinais</b>	31	27 <sup>5</sup>	31

Tabela 3.2.16 - Média de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico, diante da vogal [a]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-101,56	-58,43	-70,22
<b>Área observada</b>	-129,55 : -68,40	-103,90 : -24,95	-94,65 : -46,45
<b>Nº de sinais</b>	32	32	32

Tabela 3.2.17 - Média de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico, diante da vogal [e]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-71,14	-57,49	-77,68
<b>Área observada</b>	-119,00 : -31,90	-69,10 : -36,45	-106,55 : -56,10
<b>Nº de sinais</b>	32	32	21 <sup>6</sup>

Tabela 3.2.18 - Média de VOT e área observada em ms para a oclusiva [d] nos contextos pretônico, tônico e postônico, diante da vogal [o]

A oclusiva velar, pelas Tabelas 3.2.19, 3.2.20 e 3.2.21, apresentou valores médios de VOT maiores na posição tônica no caso das três vogais - [a], [e] e [o]. Para essa oclusiva diante de [a], as diferenças foram estatisticamente significativas das posições tônica e postônica em relação à posição pretônica (Tabela 49, apêndice). Diante das vogais médias [e] e [o], os valores médios de VOT foram significativamente maiores da posição tônica em relação às posições átonas, conforme Tabelas 50 e 51 no apêndice.

<sup>5</sup> Conforme nota de rodapé (1), página 57.

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-57,86	-51,12	-55,20
<b>Área observada</b>	-89,45 : -25,25	-73,75 : -20,95	-79,65 : -28,40
<b>Nº de sinais</b>	28	28	28

Tabela 3.2.19 - Média de VOT e área observada em ms para a oclusiva [g] nos contextos pretônico, tônico e postônico, diante da vogal [a]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-67,97	-56,64	-53,25
<b>Área observada</b>	-132,00 : -25,95	-110,20 : -27,45	-67,90 : -36,50
<b>Nº de sinais</b>	31	31	30

Tabela 3.2.20 - Média de VOT e área observada em ms para a oclusiva [g] nos contextos pretônico, tônico e postônico, diante da vogal [e]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	-76,08	-45,00	-45,13
<b>Área observada</b>	-93,70 : -42,85	-57,20 : -29,55	-62,35 : -23,45
<b>Nº de sinais</b>	25	25	22

Tabela 3.2.21 - Média de VOT e área observada em ms para a oclusiva [g] nos contextos pretônico, tônico e postônico, diante da vogal [o]

### 3.2.6. O VOT em diferentes posições na sentença

Como vários contextos em que um som da fala é encontrado podem vir a influenciar os valores desse som, tais como posição silábica, contexto vocálico, velocidade de fala, entre outros, acreditamos que diferentes posições em sentenças ocasionarão alguma mudança nos valores de VOT.

Queremos lembrar aqui, que os segmentos-alvo, ou seja, as consoantes oclusivas, encontram-se em posição tônica de palavras trissílabas, todas em contexto vocálico [a], e que as diferentes posições se referem a:

- a) inicial (0), por exemplo, *Cabala lida pelo rabino será sagrada.*
- b) inicial (1), seguindo um determinante - os artigos definidos feminino ou masculino - por exemplo, *A cabala lida pelos rabinos é sagrada.*

<sup>6</sup> Conforme nota de rodapé (2), página 57.

- c) medial, em que o segmento-alvo é intercalado na sentença após a 8ª sílaba, por exemplo, *O rabino lia na cabala que é sagrada*.
- d) final de sentença, por exemplo, *O rabino tava lendo na sagrada cabala*.

As Tabelas 61, 62 e 63, no apêndice, nos mostram que a única oclusiva a apresentar diferenças significativas nos valores de VOT em função da posição na sentença foi a bilabial, entre as posições medial e final. Na posição final, diferentemente do que esperávamos, o valor médio de VOT foi menor. As oclusivas dental e velar não mostraram diferenças significativas em nenhuma das posições. As Tabelas 3.2.22, 3.2.23 e 3.2.24 apresentam os valores médios de VOT para as oclusivas bilabial, dental e velar, assim como a área de valores observada.

	<b>inicial (0)</b>	<b>inicial (1)</b>	<b>medial</b>	<b>final</b>
<b>Média</b>	-84,00	-79,20	-87,92	-77,89
<b>Área observada</b>	-101,70 : -64,10	-91,70 : -68,15	-107,15 : -67,95	-86,35 : -70,90
<b>Nº de sinais</b>	12	11	12	10

Tabela 3.2.22 - Média de VOT e área observada em ms para a oclusiva [b] em diferentes posições em sentenças

	<b>inicial (0)</b>	<b>inicial (1)</b>	<b>medial</b>	<b>final</b>
<b>Média</b>	-67,89	-69,47	-68,79	-65,79
<b>Área observada</b>	-88,45 : -51,30	-101,70 : -52,85	-78,50 : -62,20	-79,80 : -53,35
<b>Nº de sinais</b>	12	12	12	12

Tabela 3.2.23 - Média de VOT e área observada em ms para a oclusiva [d] em diferentes posições em sentenças

	<b>inicial (0)</b>	<b>inicial (1)</b>	<b>medial</b>	<b>final</b>
<b>Média</b>	-47,34	-43,94	-48,24	-49,85
<b>Área observada</b>	-57,85 : -34,10	-58,45 : -26,25	-57,80 : -35,95	-58,00 : -41,45
<b>Nº de sinais</b>	8	10	9	10

Tabela 3.2.24 - Média de VOT e área observada em ms para a oclusiva [g] em diferentes posições em sentenças

### 3.3. Retardo Curto

A área de retardo curto, que, segundo Lisker e Abramson (1964), especifica as oclusivas surdas não aspiradas na maioria das línguas estudadas, caracteriza-se por apresentar soltura da oclusão e início da pulsação glotal simultâneos (VOT zero) ou, início da pulsação glotal seguindo imediatamente a soltura, em cerca de 10 ms a 35 ms. A Figura 3.3.1 ilustra a área de retardo curto da oclusiva velar [k] no PB.

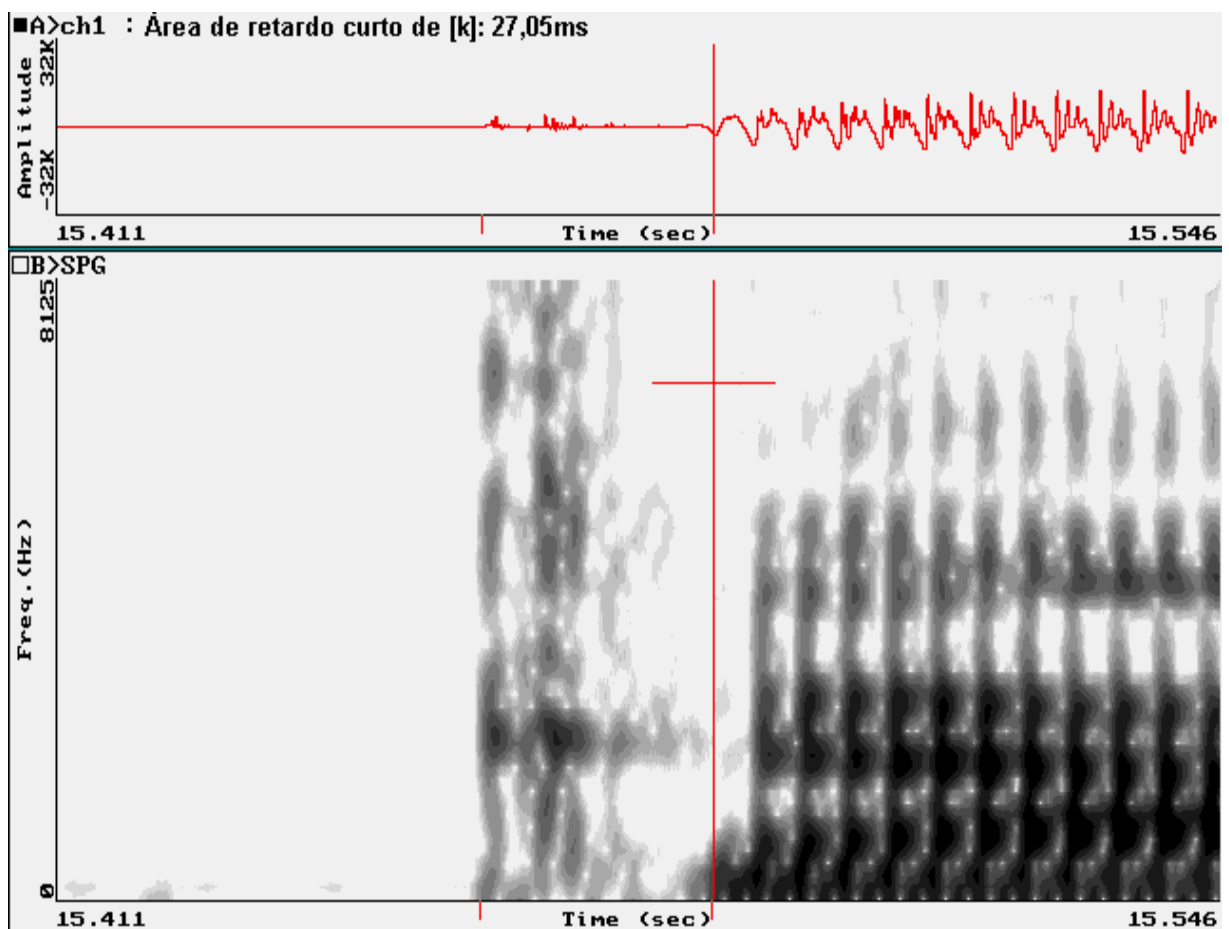


Figura 3.3.1 - Forma de onda e espectrograma da área de retardo curto da consoante oclusiva [k] no segmento [ka]

Em uma análise mais geral, os valores de VOT para a área de retardo curto das oclusivas surdas do PB distribuíram-se, no caso da bilabial [p], em uma faixa de 5,90 ms a 37,00 ms, obtendo-se um valor médio de 15,58 ms; no caso da dental [t], a distribuição dos valores deu-se de 6,95 ms a 38,00 ms, obtendo-se um valor médio de

16,69 ms; no caso da velar, os valores distribuíram-se em uma faixa de 16,65 ms a 66,20 ms, sendo a média de VOT de 36,36 ms.

Comparando nossos dados aos de Istre (s/d), como mostra o Gráfico 3.3.1, podemos observar uma certa similaridade entre as oclusivas surdas, com exceção da oclusiva bilabial [p], que, diferentemente do que esperávamos, apresentou valores de VOT bastante altos, aproximando-se dos valores que caracterizam a oclusiva dental. Enquanto o valor médio apresentado por Istre para a oclusiva [p] foi de 11,95 ms, o valor médio dessa mesma consoante em nossos dados foi de 15,58 ms. As oclusivas dental e velar apresentaram como valores médios, nos dados de Istre, 18,48 ms e 38,53 ms, respectivamente, e, em nossos dados, 16,69 ms e 36,36 ms, respectivamente. Apesar da aproximação dos valores das oclusivas bilabial e dental, a relação descoberta por Lisker e Abramson (1964), de que, quanto mais posterior o ponto de articulação, maiores os valores de VOT, foi observada também em nossos dados.

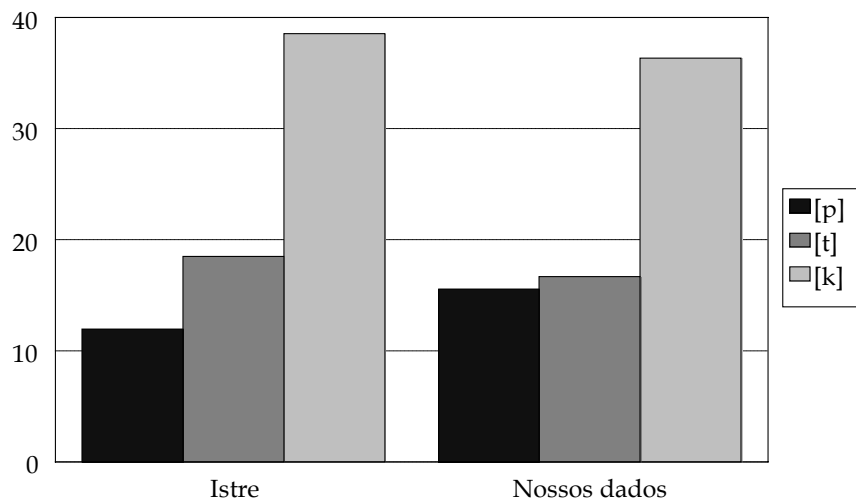


Gráfico 3.3.1 - Comparação da área de retardo curto do PB em Istre e nossos dados

Nossos resultados mostraram ainda que, com relação à categoria de retardo curto, o PB, em comparação com o holandês, o espanhol, o húngaro e o tâmil, apresentou os maiores valores de VOT para a oclusiva bilabial [p]; assemelhou-se ao holandês e ao húngaro com relação à oclusiva dental [t], e apresentou valores de VOT maiores que qualquer uma das línguas para a velar [k], conforme Gráfico 3.3.2.

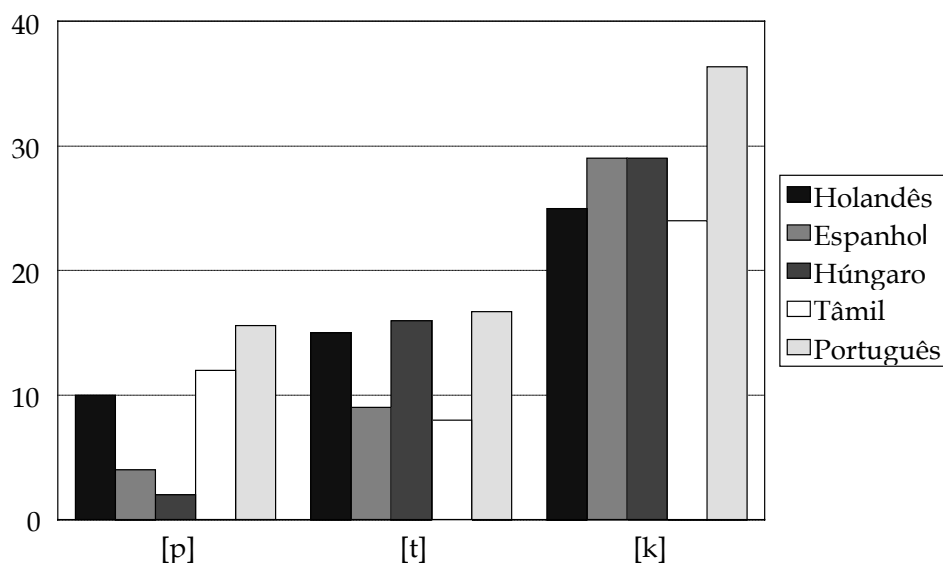


Gráfico 3.3.2 - Comparação da área de retardo curto em PB e outras línguas (Lisker e Abramson, 1964)

Nas seções seguintes, apresentaremos separadamente os resultados do primeiro estudo-piloto que dará conta dos dados referentes às inserções de oclusivas em logatomas (produzidos isoladamente e inseridos em frases-veículo) e aos diversos contextos vocálicos. E, os do segundo estudo, referentes à tonicidade da sílaba em palavras produzidas isoladamente, e de palavras em diferentes posições na sentença.

## Resultados do primeiro estudo-piloto

### 3.3.1 Oclusivas bilabiais

A área de retardo curto para as oclusivas bilabiais apresentou variação significativa nos diferentes contextos testados em nosso estudo. Isto é, essa oclusiva inserida em logatomas isolados ou logatomas em frases-veículo apresentou valores de VOT diferenciados.

Pela Tabela 3.3.1, vemos que a média de VOT desta consoante inserida em logatomas produzidos isoladamente foi de 17,70 ms e em logatomas inseridos em frases-veículo foi de 15,49 ms. Comparando esses dados, observamos que os valores de VOT medidos em produções isoladas são significativamente maiores do que aqueles inseridos em frases-veículo (Tabela 4 no apêndice). Essa ocorrência já é enfatizada por diversos estudiosos, dentre eles, Lisker e Abramson (1964), que mostraram ser o VOT mais reduzido à medida que se aumenta o “veículo” que contém o segmento sob análise.

	<b>logatoma</b>	<b>Logatoma/FV</b>
<b>Média</b>	17,70	15,49
<b>Área observada</b>	5,90 : 37,00	6,90 : 28,30
<b>Nº de sinais</b>	96	96

Tabela 3.3.1 - Média geral de VOT e área observada em ms da oclusiva [p] inserida em logatomas produzidos isoladamente e inseridos em frases-veículo.

### 3.3.2. Oclusivas Dentais

	<b>logatoma</b>	<b>logatoma/FV</b>
<b>Média</b>	16,78	17,13
<b>Área observada</b>	6,95 : 30,45	8,20 : 28,50
<b>Nº de sinais</b>	96	96

Tabela 3.3.2 - Média geral de VOT e área observada em ms da oclusiva [t] inserida em logatomas produzidos isoladamente e inseridos em frases-veículo.

O valor médio de VOT para a oclusiva [t] inserida em produções isoladas foi de 16,78 ms e inserida em frases-veículo, de 17,13 ms (Tabela 3.3.2). Nesse caso, para o PB, nossos dados relativos à oclusiva dental divergiram daqueles apresentados para a oclusiva bilabial, conforme seção 3.3.1, mostrando um comportamento diferente daquele que esperávamos. A categoria de retardo curto para a oclusiva dental, não apresentou, no entanto, alterações significativas em função da inserção em diferentes contextos. Quer dizer, não houve redução considerável do valor médio de VOT quando da inclusão do segmento em logatomas inseridos em frases-veículo

em relação aos logatomas produzidos isoladamente, conforme Tabela 5 mostradas no apêndice.

### 3.3.3. Oclusivas Velares

Pela Tabela 3.3.3, o valor médio de VOT para as oclusivas velares inseridas em logatomas produzidos isoladamente foi de 37,85 ms e em logatomas inseridos em frases-veículo, de 33,90 ms.

A área de retardo curto para estas oclusivas apresentou alterações significativas no VOT entre a oclusiva inserida em produções isoladas e a inserida em frases-veículo, conforme Tabela 6 do apêndice, corroborando a colocação de Lisker e Abramson (1964), de que a inserção do segmento-alvo em sentenças maiores tende a diminuir os valores de VOT do segmento.

	<b>logatoma</b>	<b>logatoma/FV</b>
<b>Média</b>	37,85	33,90
<b>Área observada</b>	20,60 : 66,20	16,65 : 54,90
<b>N° de sinais</b>	96	93

Tabela 3.3.3 - Média geral de VOT e área observada em ms da oclusiva [k] inserida em logatomas produzidos isoladamente e inseridos em frases-veículo.

### 3.3.4. O VOT e o contexto vocálico

O contexto vocálico provocou alterações significativas na área de retardo curto das oclusivas bilabial, dental e velar, porém, essas alterações não foram similares para as diferentes oclusivas.

Nos segmentos produzidos isoladamente, as consoantes bilabial, velar e dental apresentaram comportamentos similares, isto é, mostraram alterações significativas dos valores médios de VOT da vogal baixa em relação às médias [e] e [o], sendo que os valores de VOT das vogais médias foram maiores do que os valores da vogal



baixa (Tabelas 28, 29 e 30 do apêndice). A bilabial, no entanto, foi sensível a todos os contextos vocálicos, porém, quando comparando as vogais médias entre si, percebemos que os maiores valores de VOT são apresentados pela bilabial diante da vogal média posterior.

Para segmentos inseridos em frases-veículo, as áreas de retardo curto das oclusivas bilabial e dental mostraram-se sensíveis a todos os contextos vocálicos, no entanto, comparando os valores de VOT entre as vogais médias, enquanto para a bilabial é a vogal média posterior a de maior valor, para a dental é a vogal média anterior. Já os valores médios de VOT da oclusiva velar apresentaram diferenças significativas somente entre a vogal baixa e as vogais médias, preservando, entretanto, as relações anteriores, ou seja, a área de retardo curto da oclusiva velar diante das vogais médias é maior do que diante da vogal baixa (Tabelas 34, 35 e 36 do apêndice).

	<b>[a]</b>	<b>[e]</b>	<b>[o]</b>
<b>Média</b>	11,34	13,98	27,77
<b>Área observada</b>	6,65 : 21,20	5,90 : 26,35	18,85 : 37,00
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.4 - Média de VOT e área observada em ms para a oclusiva [p] diante das vogais [a], [e] e [o] para os logotomas produzidos isoladamente

	<b>[a]</b>	<b>[e]</b>	<b>[o]</b>
<b>Média</b>	12,51	14,11	19,84
<b>Área observada</b>	6,90 : 21,90	8,40 : 22,15	10,35 : 28,30
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.5 - Média de VOT e área observada em ms para a oclusiva [p] diante das vogais [a], [e] e [o] para os logotomas inseridos em frases-veículo

	<b>[a]</b>	<b>[e]</b>	<b>[o]</b>
<b>Média</b>	14,03	18,64	17,67
<b>Área observada</b>	6,95 : 21,70	9,45 : 30,45	11,95 : 23,40
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.6 - Média de VOT e área observada em ms para a oclusiva [t] diante das vogais [a], [e] e [o] para os logotomas produzidos isoladamente

	[a]	[e]	[o]
<b>Média</b>	15,07	19,19	17,13
<b>Área observada</b>	8,20 : 22,70	12,95 : 28,50	10,20 : 29,55
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.7 - Média de VOT e área observada em ms para a oclusiva [t] diante das vogais [a], [e] e [o] para os logotomas inseridos em frases-veículo

	[a]	[e]	[o]
<b>Média</b>	33,84	40,67	39,04
<b>Área observada</b>	20,60 : 62,90	25,35 : 66,20	29,55 : 48,65
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.8 - Média de VOT e área observada em ms para a oclusiva [k] diante das vogais [a], [e] e [o] para os logotomas produzidos isoladamente

	[a]	[e]	[o]
<b>Média</b>	29,68	36,37	35,83
<b>Área observada</b>	16,65 : 36,55	26,65 : 54,90	28,90 : 50,55
<b>Nº de sinais</b>	32	32	29

Tabela 3.3.9 - Média de VOT e área observada em ms para a oclusiva [k] diante das vogais [a], [e] e [o] para os logotomas inseridos em frases-veículo

Viana (1984 *apud* Andrade e Mascarenhas, 1994), para o português europeu, demonstrou que os valores médios de VOT positivo são maiores para as vogais altas do que para as não altas. Nossos dados não incluem as vogais altas, porém, talvez pudéssemos, a partir desses resultados, inferir que, quanto mais alta a vogal, maior o VOT, já que, nos diferentes contextos analisados, tivemos sempre essa relação.

Retomando os resultados apresentados por Lisker e Abramson (1964), com relação à redução dos valores de VOT quando os segmentos são inseridos em contextos maiores, nossos dados ratificam essa colocação para as oclusivas velares diante das vogais médias [e], [o] e da baixa [a] (Tabelas 22, 23 e 24 do apêndice) e para a bilabial diante da vogal média posterior [o] (conforme Tabela 18, apêndice). Nos outros contextos não houve alterações significativas.

## Resultados referentes ao segundo estudo-piloto

### 3.3.5. O VOT e a tonicidade da sílaba

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	13,56	12,98	17,05
<b>Área observada</b>	6,80 : 28,00	6,70 : 22,50	7,25 : 31,65
<b>Nº de sinais</b>	96	96	96

Tabela 3.3.10 - Média geral de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	16,16	16,09	20,11
<b>Área observada</b>	7,45 : 38,00	6,90 : 35,05	6,95 : 41,95
<b>Nº de sinais</b>	96	96	82 <sup>7</sup>

Tabela 3.3.11 - Média geral de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	37,34	33,53	33,10
<b>Área observada</b>	18,85 : 58,50	16,70 : 54,70	17,25 : 63,15
<b>Nº de sinais</b>	96	95	96

Tabela 3.3.12- Média geral de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico

De forma geral, nossos dados mostraram que a área de retardo curto das oclusivas [p], [t] e [k] do PB, como esperávamos, sofreu influências em função da tonicidade da sílaba.

Todavia, observando as Tabelas 3.3.10, 3.3.11 e 3.3.12 e as Tabelas 40, 41 e 42 apresentadas no apêndice, percebemos que as alterações não foram sistemáticas, pois, enquanto as oclusivas bilabial e dental apresentaram valores médios de VOT

<sup>7</sup> O número de sinais [to] em contexto postônico é bastante reduzido em relação aos contextos tônico e pretônico porque, nesse caso, alguns falantes alternaram produção de [to] e [tu].

maiores para a posição postônica - 17,05 ms e 20,11 ms, respectivamente -, a oclusiva velar apresentou média de VOT mais alta para a posição tônica - 33,53 ms. Dessa forma, nossa hipótese inicial, de que o VOT apresentaria maior valor médio em contexto tônico, foi apenas parcialmente confirmada, ou seja, somente no caso da oclusiva velar.

Andrade (1984 *apud* Andrade e Mascarenhas, 1994), estudando o VOT da oclusiva [t] inserida na sílaba *ti*, encontrou, para o PE, valores de VOT maiores em sílabas não-acentuadas do que em sílabas acentuadas. Embora as diferenças tivessem sido pequenas, mostraram-se significativas.

Andrade e Mascarenhas (1994), analisando o VOT da mesma oclusiva, no mesmo contexto vocálico, também encontraram diferenças significativas com relação às posições tônica, pré e postônicas, contrárias, porém, a sua hipótese inicial e à descoberta de Andrade (1984): seus resultados revelaram valores de VOT mais altos associados às sílabas acentuadas, e não às não-acentuadas.

Com relação à oclusiva dental [t], comparando nossos resultados aos obtidos por Andrade (1984) e Andrade e Mascarenhas (1994), observamos que nossos resultados concordam com os de Andrade (1984), no sentido de que os maiores valores médios de VOT são apresentados pela oclusiva inserida na sílaba não-acentuada, no caso de nossa pesquisa, mais precisamente, na posição postônica.

Analisando nossos dados segundo tonicidade e contexto vocálico, verificamos que a área de retardo curto apresentou valor médio de VOT, para a oclusiva bilabial, maior em posição postônica diante da vogal baixa [a] e da vogal média posterior [o], como podemos observar pelas Tabelas 3.3.13, 3.3.15 e 52 e 54 mostradas no apêndice. No caso da vogal média anterior [e] (Tabela 3.3.14), os valores de VOT fixaram-se na faixa de 12,00 ms, sem alterações significativas, portanto, como mostra a Tabela 53 do apêndice.

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	11,50	10,78	17,18
<b>Área observada</b>	6,80 : 16,95	6,75 : 15,35	8,65 : 27,35
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.13 - Média de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico, diante da vogal [a]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	12,26	12,67	12,41
<b>Área observada</b>	8,15 : 16,75	7,05 : 20,15	7,25 : 17,95
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.14 - Média de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico, diante da vogal [e]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	16,91	15,49	21,56
<b>Área observada</b>	8,10 : 35,80	6,70 : 21,05	15,50 : 31,65
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.15 - Média de VOT e área observada em ms para a oclusiva [p] nos contextos pretônico, tônico e postônico, diante da vogal [o]

As oclusivas dentais em contexto vocálico [a] e [e] (Tabelas 3.3.16 e 3.3.17) apresentaram alterações para todos os contextos de tonicidade, sendo que os valores médios de VOT foram significativamente maiores na posição postônica, conforme mostram as Tabelas 55 e 56 do apêndice. Diante da vogal [o], os valores de VOT não tiveram variação estatisticamente significativa entre os contextos de tonicidade, fixando-se em uma faixa de 15,00 a 16,00 ms, conforme Tabela 3.3.18 e Tabela 57 (apêndice).

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	14,23	16,39	19,12
<b>Área observada</b>	7,75 : 28,60	10,45 : 24,50	8,00 : 32,90
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.16 - Média de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico, diante da vogal [a]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	18,43	15,98	23,35
<b>Área observada</b>	7,55 : 38,00	6,90 : 26,90	11,50 : 41,95
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.17 - Média de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico, diante da vogal [e]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	15,81	16,21	16,11
<b>Área observada</b>	7,95 : 35,05	7,45 : 25,40	6,95 : 28,60
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.18 - Média de VOT e área observada em ms para a oclusiva [t] nos contextos pretônico, tônico e postônico, diante da vogal [o]

Tomados juntos, contexto de tonicidade e contexto vocálico, para a oclusiva velar, como mostram as Tabelas 3.3.19 a 3.3.21, diante da vogal [a], o valor médio de VOT da oclusiva é significativamente maior na posição postônica; diante de [e], é significativamente menor nessa posição e, diante de [o], a média de VOT é significativamente maior na posição tônica.

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	30,22	30,69	34,75
<b>Área observada</b>	18,85 : 45,55	20,80 : 39,20	17,25 : 49,65
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.19 - Média de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico, diante da vogal [a]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	37,63	34,69	31,02
<b>Área observada</b>	16,70 : 48,80	22,85 : 53,00	18,95 : 63,15
<b>Nº de sinais</b>	32	31	32

Tabela 3.3.20 - Média de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico, diante da vogal [e]

	<b>tônico</b>	<b>pretônico</b>	<b>postônico</b>
<b>Média</b>	44,16	35,24	33,54
<b>Área observada</b>	27,20 : 58,50	23,20 : 55,75	20,90 : 51,00
<b>Nº de sinais</b>	32	32	32

Tabela 3.3.21 - Média de VOT e área observada em ms para a oclusiva [k] nos contextos pretônico, tônico e postônico, diante da vogal [o]

Concluindo, cruzando contexto de tonicidade e contexto vocálico através do teste *t-de-Student* para todas as oclusivas surdas do PB, obtivemos os seguintes resultados:

- em contexto vocálico [a], o VOT das consoantes oclusivas bilabial [p], dental [t] e velar [k] foi significativamente maior para a posição postônica, conforme Tabelas 52, 55 e 58, no apêndice;
- em contexto vocálico [e], a oclusiva [t] apresentou média de VOT significativamente maior na posição postônica (Tabela 56, apêndice), enquanto a oclusiva [k] apresentou redução significativa do valor médio de VOT nessa mesma posição (Tabela 59, apêndice).
- em contexto vocálico [o], a oclusiva bilabial apresentou VOT estatisticamente maior na posição postônica e a oclusiva velar, na posição tônica (Tabelas 54 e 60 do apêndice).

### 3.3.6. O VOT em diferentes posições na sentença

	<b>inicial (0)</b>	<b>inicial (1)</b>	<b>medial</b>	<b>final</b>
<b>Média</b>	11,59	11,20	13,12	12,00
<b>Área observada</b>	7,10 : 14,55	6,65 : 14,50	7,00 : 19,40	7,65 : 17,20
<b>Nº de sinais</b>	12	12	12	12

Tabela 3.3.22 - Média de VOT e área observada em ms para a oclusiva [p] em diferentes posições em sentenças

	<b>inicial (0)</b>	<b>inicial (1)</b>	<b>medial</b>	<b>final</b>
<b>Média</b>	12,54	12,98	11,95	13,72
<b>Área observada</b>	7,30 : 20,50	6,60 : 20,40	7,00 : 16,95	7,95 : 22,75
<b>N° de sinais</b>	12	12	12	12

Tabela 3.3.23 - Média de VOT e área observada em ms para a oclusiva [t] em diferentes posições em sentenças

	<b>inicial (0)</b>	<b>inicial (1)</b>	<b>medial</b>	<b>final</b>
<b>Média</b>	28,68	28,30	27,78	30,93
<b>Área observada</b>	20,65 : 32,10	18,15 : 40,65	18,50 : 35,85	21,10 : 37,95
<b>N° de sinais</b>	12	12	12	12

Tabela 3.3.24 - Média de VOT e área observada em ms para a oclusiva [k] em diferentes posições em sentenças

Aplicando os testes estatísticos, pudemos verificar alterações significativas nos valores de VOT em função das diferentes posições na sentença. No caso da oclusiva bilabial, houve alteração significativa da área de retardo curto da posição medial em relação às posições inicial (0) e inicial (1) (Tabela 64 do apêndice). No caso das oclusivas dental e velar, diferenças estatisticamente significativas foram observadas somente entre as posições medial e final, conforme Tabelas 65 e 66 apresentadas no apêndice.

Conforme podemos observar pela Tabela 3.3.22 e pela Tabela 64 do apêndice, o maior valor médio de VOT foi apresentado pela oclusiva bilabial em posição medial de sentença, a saber, 13,12 ms. Para as oclusivas dental e velar, as maiores médias de VOT ocorreram em posição final (13,27 ms para a dental e 30,93 ms para a velar), o que corrobora nossa hipótese inicial (Tabelas 3.3.23 e 3.3.24 e Tabelas 65 e 66, apêndice).

### 3.4. Considerações finais

A similaridade existente entre as diferentes línguas de duas categorias (holandês, húngaro, espanhol, tâmil, cantonês e inglês), observada por Lisker e



Abramson (1964), é bastante surpreendente, pois as categorias de oclusivas tendem a cair, todas, geralmente em três faixas - a) de -125 ms a -75 ms, b) de 0 (zero) a +25 ms e c) de +60 ms a +100 ms - sendo os valores medianos para essas faixas de -100 ms, +10 ms e +75 ms, respectivamente. O holandês, o húngaro, o espanhol e o tâmil mostram essencialmente a mesma distribuição de valores, ocupando as faixas em torno de -100 ms e +10 ms; o cantonês e o inglês (exceto para os valores de pré-sonorização apresentados por um dos falantes) apresentam uma distribuição dos valores de retardo curto e retardo longo em torno das faixas de +10 ms e +75 ms.

A fim de melhor visualizar como o VOT separa as categorias de oclusivas no PB segundo sua sonoridade, apresentaremos a seguir os gráficos de frequência de distribuição para cada série de oclusivas. Esse tipo de gráfico permite que se veja com mais clareza se os valores determinados para itens pertencentes a alguma categoria em particular tendem a agrupar-se próximos a algum valor dito “modal”. Cada categoria de oclusiva é representada por um conjunto de linhas verticais ao longo do eixo horizontal, em intervalos de 5 ms. A altura de cada linha indica a percentagem de itens pertencentes à categoria.

O PB distribui seus valores de VOT em torno de -165 ms a -20 ms e +5 ms a +60 ms, apresentando valores medianos para essas faixas de -95 ms e +30 ms.

Considerando primeiramente os gráficos de distribuição das oclusivas inseridas em logatomas produzidos isoladamente, passaremos, então, à análise da distribuição das oclusivas bilabiais [b] e [p] (Gráficos 3.4.1 e 3.4.2). Se lembrarmos da média de VOT apresentada pelo segmento surdo produzido isoladamente - 17,70 ms - e se observarmos a distribuição de seus valores ao longo do eixo horizontal, veremos que esses valores, que têm maior percentagem de ocorrência em 5 ms e 10 ms (30 e 40%, respectivamente), parecem encaminhar-se para a área de retardo longo, uma vez que, embora em percentagem bem menor, esses valores cheguem a alcançar 35 ms, o que parece indicar que as oclusivas bilabiais surdas apresentam certa aspiração, confirmando nossa hipótese inicial. Já a área de pré-sonorização tem seus valores distribuídos numa área bastante extensa, de -162,75 ms a -26,70 ms, apresentando uma média de -86,57 ms, evidenciando assim a razão de se considerar

esse VOT como sendo de precedência longa de sonorização. Os valores para essa área concentraram-se entre -110 ms e -75 ms.

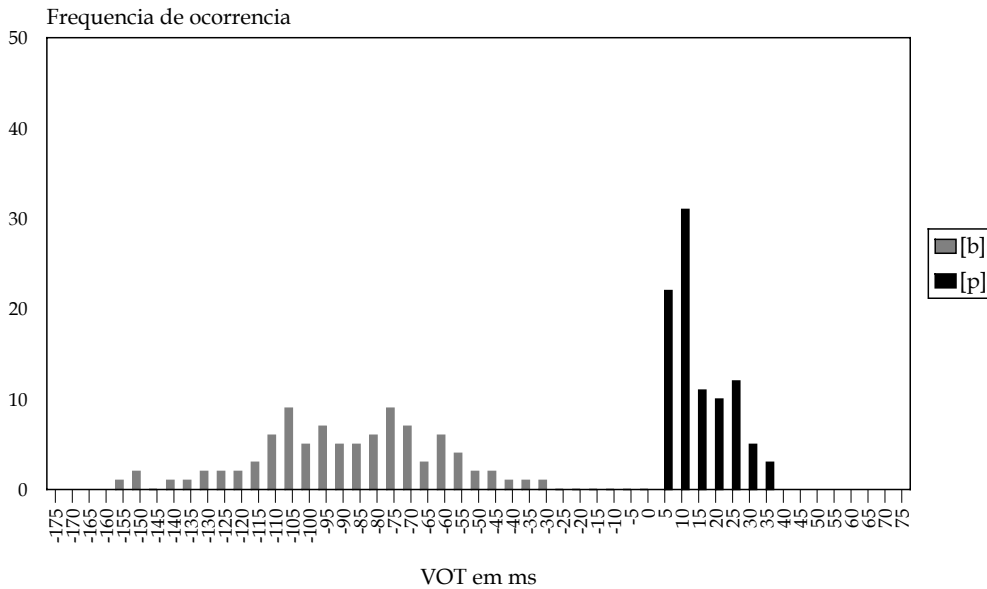


Gráfico 3.4.1 - Distribuição do VOT das oclusivas bilabiais do PB inseridas em logatomas produzidos isoladamente

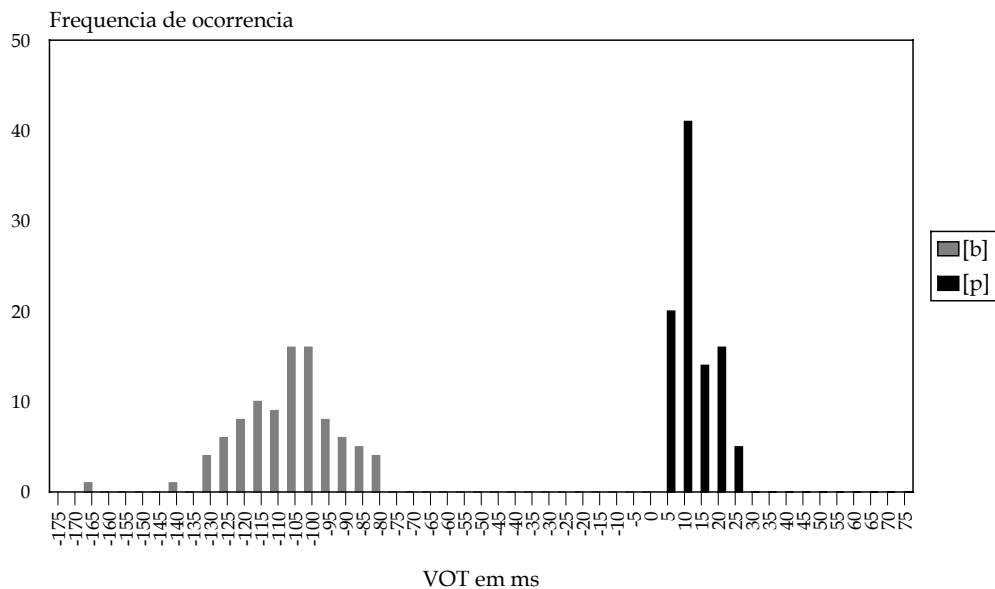


Gráfico 3.4.2 - Distribuição do VOT das oclusivas bilabiais do PB em logatomas inseridos em frases-veículo

No caso de inserção em logatomas em frases-veículo, a oclusiva [p] (Gráfico 3.4.2), cujo valor médio foi de 15,49 ms, significativamente menor do que o valor médio para a mesma oclusiva produzida em logatomas isolados, teve seus valores distribuídos de 5 ms a 25 ms, indicando, portanto, pequena redução em sua distribuição. A área de maior concentração de sinais fixou-se, ainda nesse caso, entre 5 ms e 10 ms, continuando com seu pico máximo de frequência (pouco mais de 40%) em 10 ms. Para a oclusiva [b], a redução na distribuição dos valores de VOT, no entanto, foi consideravelmente maior. Seus valores, que na produção de logatomas isolados vão de -160 ms a -30 ms, no caso de inserção em frases-veículo reduziu-se para uma área de -130 ms a -80 ms, com uma pequena dispersão em -165 ms e -140 ms. Os valores de [b] em logatomas produzidos em frases-veículo tiveram maior frequência de ocorrência em -105 ms e -100 ms. A redução, no entanto, aconteceu apenas na distribuição de valores ao longo do *continuum* de VOT, pois em termos de valor médio, a oclusiva [b] inserida em logatomas produzidos em frases-veículo foi maior do que o valor médio para a consoante inserida em logatomas produzidos isoladamente, -104,98 ms e -86,57 ms, respectivamente.

As oclusivas dentais [d] e [t] têm a distribuição de seus valores apresentada nos Gráficos 3.4.3 e 3.4.4.

A sonora [d], inserida em logatomas produzidos isoladamente, apresentou maior frequência de ocorrência, para sua área de pré-sonorização, em torno de -120 ms e -100 ms, sendo que seus valores se distribuíram ao longo de uma área de -155 ms e -40 ms, com um valor médio de -98,84 ms, precedência longa de sonorização, como previa nossa hipótese inicial. Essa oclusiva, inserida em logatomas produzidos em frases-veículo, apresentou redução na distribuição de seus valores, localizando-os entre -140 ms e -55 ms, com maior frequência de ocorrência entre -95 ms e -85 ms. Como o valor médio dessa oclusiva inserida em logatomas produzidos em frases-veículo é de -93,31 ms, vemos que em se tratando tanto de valores médios quanto de distribuição de valores ao longo do *continuum* de VOT, o comportamento é o mesmo, ou seja, redução no caso de inserção em contextos maiores.

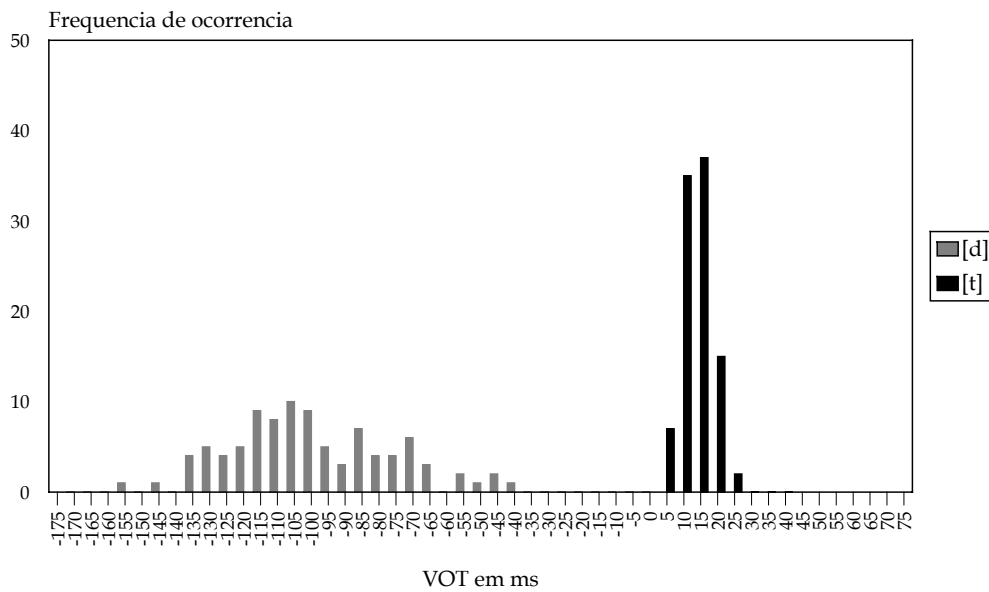


Gráfico 3.4.3 - Distribuição do VOT das oclusivas dentais do PB inseridas em logatomas produzidos isoladamente

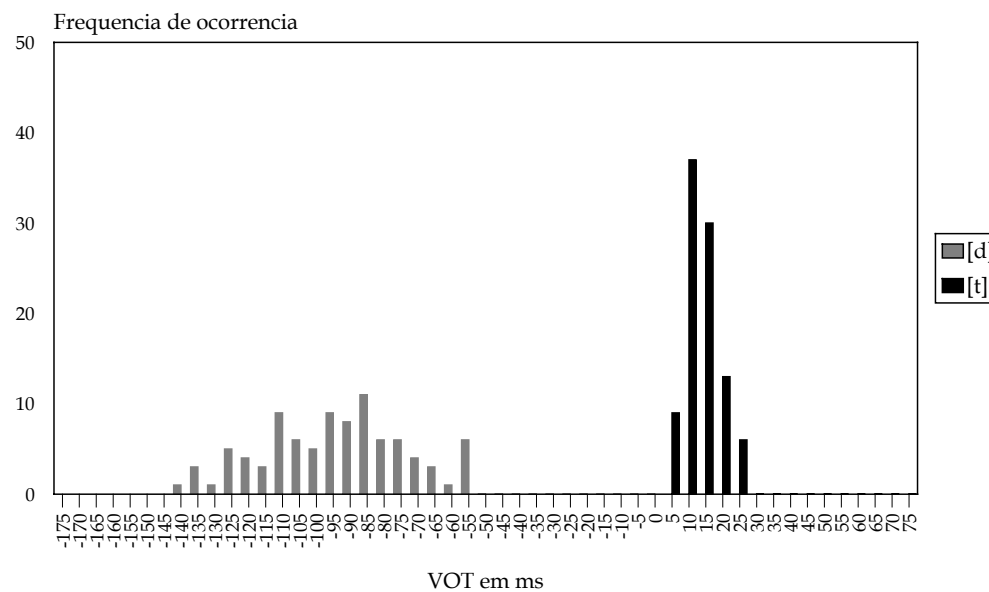


Gráfico 3.4.4 - Distribuição do VOT das oclusivas dentais do PB em logatomas inseridos em frases-veículo

No caso da dental surda [t], seus valores distribuíram-se em uma área de 5 ms a 25 ms, com maior frequência de ocorrência em 10 ms e 15 ms (em torno de 40%)

tanto para a oclusiva inserida em logatomas produzidos isoladamente, quanto para aquelas inseridas em logatomas produzidos em frases-veículo. Sua distribuição, diferentemente do que aconteceu com [p], não parece se encaminhar para a área de retardo longo, haja vista os valores médios e a distribuição apresentados.

Nos Gráficos 3.4.5 e 3.4.6, é apresentada a distribuição dos valores de VOT das oclusivas velares [g] e [k].

Os valores de VOT para a sonora [g], no caso de logatomas produzidos isoladamente, distribuíram-se em cerca de -165 ms a -20 ms, sendo que houve maior concentração de sinais na área entre -110 ms e -80 ms. A distribuição, juntamente com o valor médio obtido para a oclusiva velar sonora - -78,20 ms - comprovam a precedência longa de sonorização que esperávamos. Quando da inserção dos logatomas em frases-veículo, observou-se uma redução da área de pré-sonorização, cujos valores distribuíram-se de -130 ms a -40 ms.

Os valores de VOT para a área de retardo curto da oclusiva [k] distribuíram-se de 20 ms a 60 ms, “ultrapassando o limite” da área de retardo curto, ou seja, os cerca de 35 ms que a caracterizam. A área de maior concentração de sinais localizou-se entre 20 ms e 45 ms, tendo seu pico máximo de frequência de ocorrência em 35 ms (30%), o que corrobora nossa hipótese inicial, de que as oclusivas surdas do PB apresentariam certa aspiração, embora essa aspiração não seja distintiva, como no cantonês ou no inglês (para oclusivas em posição inicial) ou nas línguas de três categorias (tai, hindi, marati, coreano). A média geral dessa oclusiva (37,85 ms) já mostra uma certa tendência a encaminhar-se para a área caracterizada como retardo longo. A distribuição dos valores de VOT para essa oclusiva inserida em logatomas produzidos em frases-veículo apresentou uma pequena redução, indo de 15 ms a 55 ms.

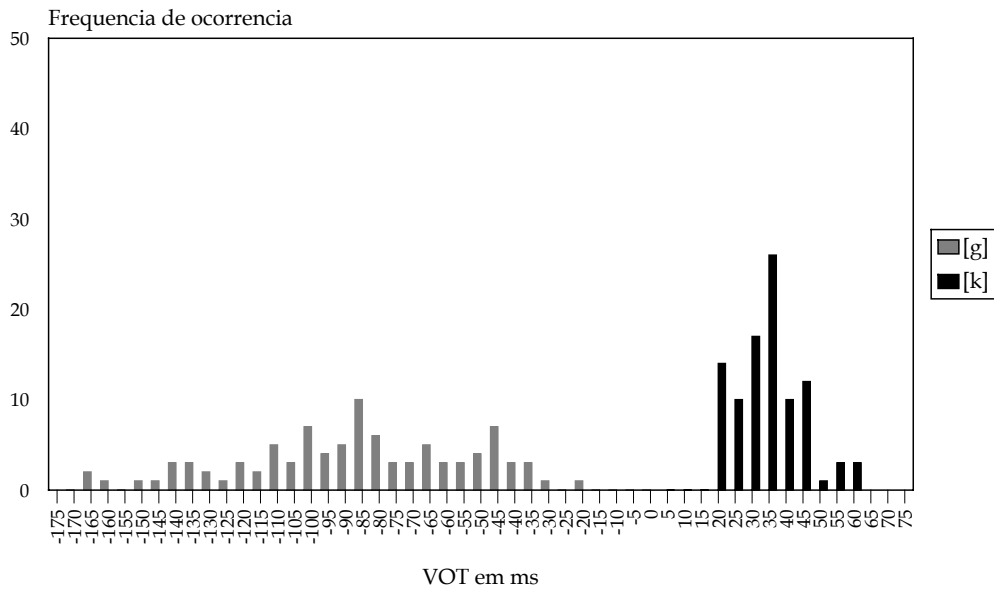


Gráfico 3.4.5 - Distribuição do VOT das oclusivas velares do PB inseridas em logotomas produzidos isoladamente

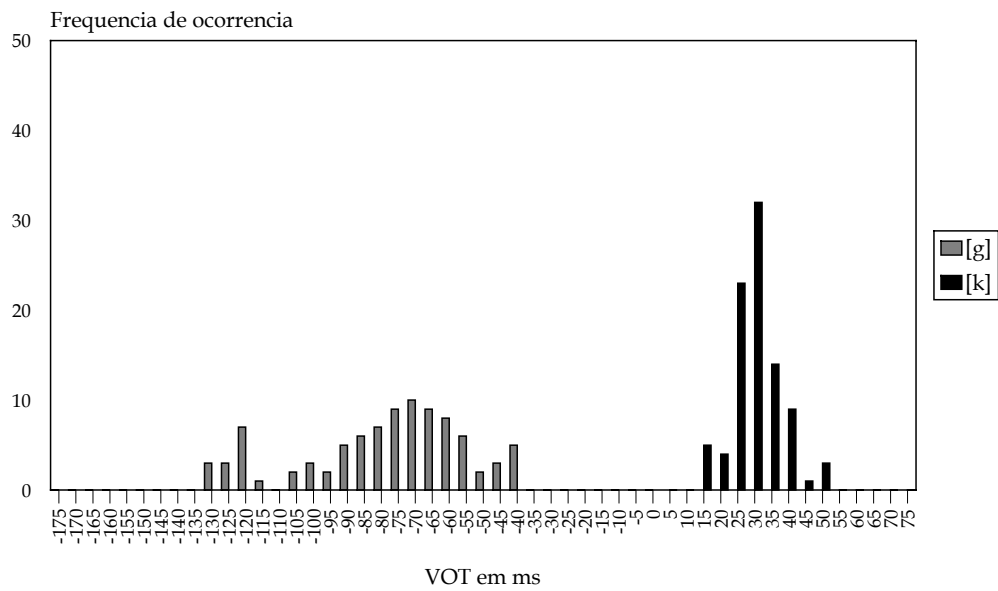


Gráfico 3.4.6 - Distribuição do VOT das oclusivas velares do PB em logotomas inseridos em frases-veículo

Conforme podemos observar pelos gráficos anteriormente apresentados, as categorias de sonorização foram diferentemente afetadas, ou seja, enquanto a área de pré-sonorização teve as distribuições de seus valores consideravelmente reduzidas em função da inserção dos logatomas em contextos maiores, a distribuição da área de retardo curto praticamente não foi alterada. Kessinger & Blumstein (1997), estudando a influência da velocidade de fala sobre o VOT, observaram essa mesma assimetria e levantaram a hipótese da necessidade de se manter um contraste fonético na língua, ou seja, enquanto houver alguma variabilidade inerente à produção de VOT como uma função de velocidade de fala, é essencial que o contraste de sonorização seja mantido. Como consequência, uma das categorias necessita ser o mínimo afetada pela velocidade de fala, para que esse contraste possa ser mantido. Julgamos que a mesma justificativa possa ser empregada no caso da inserção de segmentos em contextos maiores.

Lisker e Abramson (1964) já haviam observado que a inserção de segmentos em contextos maiores resultava na redução tanto da média de VOT quanto da distribuição dos valores de VOT ao longo do *continuum* e, nos casos em que ocorre essa redução na distribuição, ocorre também, naturalmente, uma diminuição do intervalo que separa as categorias de pré-sonorização, retardo curto e retardo longo. Como consequência de tal redução, poderá ocorrer sobreposição entre categorias, que de outra forma seriam distintas.

Sobreposição foi observada também no estudo de Kessinger & Blumstein (1977), em que as autoras investigaram a influência da velocidade de fala sobre as oclusivas bilabiais e dentais em três línguas com diferentes categorias fonéticas de sonorização - tai, inglês e francês. Enquanto o tai apresenta as áreas de pré-sonorização, retardo curto e retardo longo, caracterizando, respectivamente, as oclusivas sonoras, surdas não-aspiradas e surdas aspiradas, o inglês apresenta a área de retardo curto caracterizando as oclusivas sonoras e de retardo longo, as oclusivas surdas. Já o francês apresenta as áreas de pré-sonorização e retardo curto distinguindo as oclusivas sonoras das surdas. As autoras investigaram três

condições: segmentos produzidos isoladamente e inseridos em sentenças em velocidade lenta e rápida.

Em termos de distribuição dos valores de VOT, o tai, o francês e o inglês, tanto para as oclusivas bilabiais quanto para as dentais, quando comparando as condições isolada e de velocidade lenta de fala, apresentaram uma distribuição virtualmente idêntica. Entretanto, na condição de fala rápida, ambas as línguas apresentaram alterações nas distribuições dos valores ao longo do *continuum* de VOT. Para o tai, menos sinais foram produzidos com retardo muito longo para as oclusivas aspiradas ou com precedência muito longa para as oclusivas sonoras. Parece ter havido algum movimento dos valores dessas categorias em direção à categoria de retardo curto. Esse movimento implicou uma sobreposição mínima, porém, somente entre as categorias de retardo curto e retardo longo, a saber, [p] e [p<sup>h</sup>] e [t] e [t<sup>h</sup>]. Para o francês e o inglês, contudo, apesar do movimento dos valores de VOT das áreas de pré-sonorização (francês) e retardo longo (inglês) em direção à área de retardo curto, não houve sobreposição entre as áreas.

Podemos comparar o comportamento do francês para a variável “velocidade de fala” ao comportamento do PB para a variável “inserção de segmentos em contextos maiores” – já que o controle de sonorização para essas duas línguas se dá entre as mesmas categorias (pré-sonorização) para as oclusivas sonoras e retardo curto para as surdas) – e verificar que são similares, ou seja, nem o francês, nem o PB apresentaram sobreposição. Para o PB, curiosamente, e indo contra a descoberta de Lisker e Abramson (1964) já mencionada anteriormente, ao invés de o intervalo entre as categorias sofrer redução em função da inserção dos segmentos em contextos maiores, ocorreu o contrário. Os intervalos ou “espaços em branco” que separam as áreas de VOT aumentaram, diminuindo ainda mais qualquer possibilidade de sobreposição.

O que se pode observar para os dados do PB e que também já havia sido verificado por Lisker e Abramson (1964) para os dados do inglês, é uma sobreposição entre os valores de oclusivas de diferentes pontos de articulação. No caso do inglês, esses autores constataram sobreposição na área de +20 ms e +30 ms, não entre [b] e



[p] ou [d] e [t] ou [g] e [k], mas entre [g] e [p]. Nossos dados também mostraram sobreposição de valores para diferentes pontos de articulação, na área de +15 ms para as oclusivas [p] e [t] e +30 ms para [p] e [k]. A sobreposição dos valores de VOT das oclusivas [p] e [k] em 30 ms parece reforçar, mais uma vez, nossa hipótese inicial de que o PB apresentaria certa aspiração.

A fim de ilustrar essa aspiração das oclusivas surdas do PB, exibimos a seguir, as formas de onda e os espectrogramas de [p] e [k] e seus alofones aspirados.

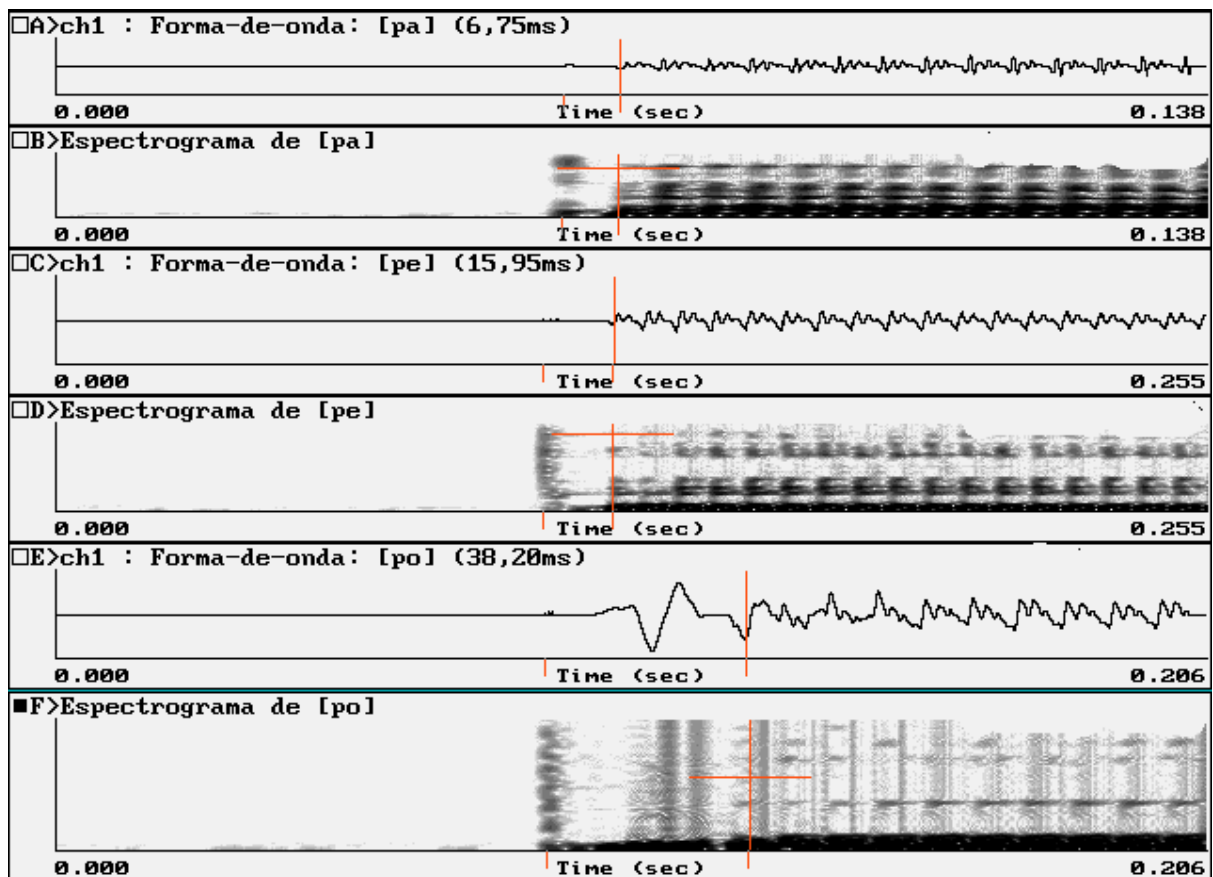


Figura 3.4.1 - Formas de onda e espectrogramas da área de retardo curto (quadros A e B), da “tendência ao retardo longo” (quadros C e D) e da área de retardo longo (quadros E e F) da oclusiva bilabial surda do PB

Os quadros A e B ilustram a oclusiva bilabial surda [p] em contexto vocálico [a]. Podemos verificar, tanto pela forma de onda quanto pelo espectrograma que, nesse contexto, a oclusiva não apresenta qualquer aspiração, pois o início de sonorização ou o primeiro pulso glotal, referente à vogal seguinte, começa imediatamente após a soltura da oclusão.

Nos quadros C e D, que apresentam a oclusiva bilabial surda em contexto vocálico [e], podemos observar um certo atraso no início da sonorização com relação à soltura da oclusão. Percebemos, aqui, que ocorre a sobreposição entre valores de VOT de oclusivas com pontos de articulação diferentes já mencionada por Lisker e Abramson (1964). Isto é, o VOT de [p] entra na área de valores que caracteriza a oclusiva [t].

Os quadros E e F mostram claramente uma oclusiva bilabial surda aspirada, uma vez que o intervalo de tempo entre a soltura da oclusão e o início de sonorização é maior do que 35 ms, valor limite entre as áreas de retardo curto e retardo longo. Esta figura parece ilustrar bem a relação observada no Capítulo 3, seção 3.3.4, de que a oclusiva bilabial surda apresenta o maior valor médio diante da vogal média posterior. Novamente aqui podemos verificar a sobreposição entre valores de VOT de diferentes pontos de articulação, no caso, VOT de [p] “invadindo” a área de valores característicos de [k].

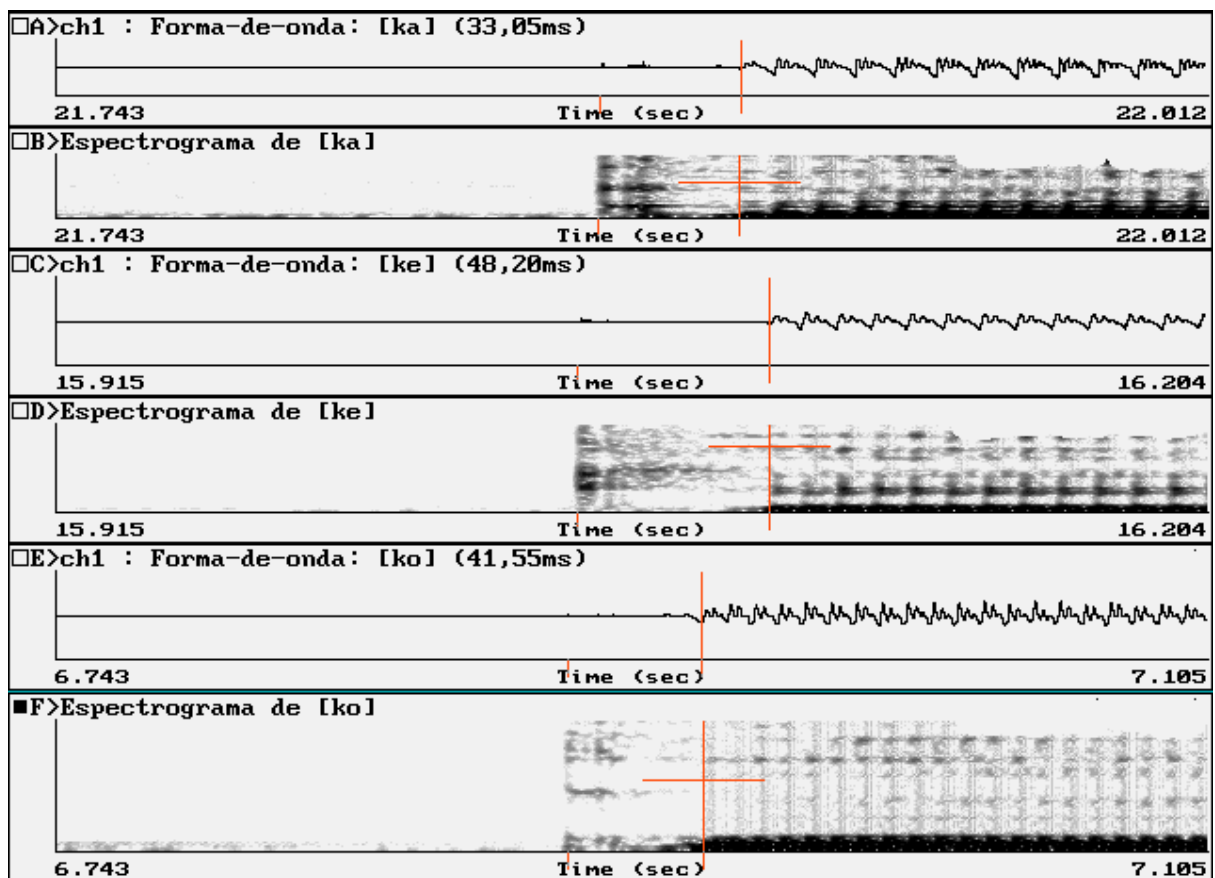


Figura 3.4.2 – Formas de ondas e espectrogramas da área de retardo curto (quadros A e B) e de retardo longo (quadros C, D e E, F) da oclusiva velar

Os quadros A e B nos mostram o VOT da oclusiva [k] em contexto vocálico [a]. Podemos verificar, tanto pela forma de onda quanto pelo espectrograma que, nesse contexto, a oclusiva não parece apresentar qualquer aspiração, pois o início de sonorização, referente à vogal seguinte, começa imediatamente após a soltura da oclusão, fixando-se na área limite de até 35 ms, característica da categoria de retardo curto.

As oclusivas [k] ilustradas nos quadros C, D e E e F, seguidas pelas vogais [e] e [o], respectivamente, parecem mostrar certa aspiração, pois seus valores de VOT ultrapassam o limite de fronteira de categoria, ou seja, 35 ms. Diante de [e] (quadros C e D), a oclusiva [k] apresenta um valor de 48,20 ms, diante de [o] (quadros E e F), essa oclusiva apresenta um valor de 41,55 ms.

# Conclusão

Os resultados das análises acústicas demonstraram, primeiramente, que o VOT do PB, em concordância com nossa hipótese inicial, caracteriza-se por apresentar uma precedência longa de sonorização especificando as oclusivas sonoras e uma área de retardo curto especificando as oclusivas surdas, cujos valores parecem encaminhar-se para a área de retardo longo, o que pode sugerir alguma aspiração dessas últimas.

Nossos dados revelaram também que os valores médios de VOT diminuem em consequência da inserção dos segmentos-alvo em contextos maiores. Isso foi verificado para as oclusivas sonoras [d] e [g] e para as surdas [p] e [k], corroborando nossa hipótese inicial e a descoberta de Lisker e Abramson (1964). O mesmo comportamento não foi observado para as oclusivas bilabial sonora e dental surda.

De uma forma geral, como esperávamos, os valores de VOT mostraram-se sensíveis ao contexto vocálico, tanto no caso de logatomas produzidos isoladamente quanto no caso de logatomas inseridos em frases-veículo. Para as oclusivas [p], [t] e [k] inseridas em logatomas produzidos isoladamente e para as oclusivas [d], [g], [p], [t] e [k] inseridas em logatomas produzidos em frases-veículo, os valores médios de

VOT diante das vogais médias foram sempre maiores do que os valores médios diante da vogal baixa. Entre as vogais médias, para as oclusivas [t] e [k] inseridas em logatomas produzidos isoladamente e para [d], [g] e [k] inseridas em logatomas produzidos em frases-veículo, os maiores valores médios de VOT foram apresentados pela oclusiva diante da vogal média anterior. Para a oclusiva [p] em logatomas isolados e inseridos em frases-veículo, o VOT foi significativamente maior diante da vogal média posterior. Viana (1984 *apud* Andrade e Mascarenhas, 1994) já havia mostrado, para o PE, que os valores médios de VOT positivo são maiores para as vogais altas do que para as não-altas. Embora nossos dados não incluam as vogais altas, eles parecem demonstrar que a média de VOT será tanto maior quanto mais alta for a vogal.

Quanto ao contexto de tonicidade, as oclusivas [b], [d], [g] e [k] apresentaram valores médios de VOT maiores na posição tônica, como previa nossa hipótese inicial e como já havia sido mostrado por Andrade e Mascarenhas (1994) para a oclusiva [t]. Para as oclusivas [p] e [t], os maiores valores médios de VOT foram apresentados na posição postônica, o que vai de encontro às descobertas dessas autoras, mas concorda com Andrade (1984), que apresenta valores médios de VOT maiores para oclusivas em sílabas não-acentuadas.

Tomados juntos, o contexto de tonicidade e o contexto vocálico revelaram que as oclusivas [b] e [d] diante da vogal baixa [a]; [b], [g] e [k] diante da vogal média posterior [o] e [d] e [g] diante da vogal média anterior [e], apresentaram valores médios de VOT maiores na posição tônica. Para as oclusivas [b] e [d] diante de [a]; [b] diante de [o] e [d] diante de [e], a relação entre as posições átonas foi similar, no sentido de que o maior valor de VOT foi apresentado pela oclusiva na posição postônica.

Com respeito às diferentes posições nas sentenças, nossos dados demonstraram que o VOT mostrou-se sensível a essa variável. Todavia, esses dados não corroboraram de todo nossa hipótese inicial que previa maiores valores de VOT em oclusivas inseridas em palavras em posição final de sentença, haja vista isso ter sido observado somente no caso das oclusivas [t] e [k]. Para [b] e [p], a maior média

de VOT foi apresentada em posição medial de sentença. [d] e [g] não mostraram alterações significativas para nenhuma das posições dentro da sentença.

Diversas questões ainda merecem análises mais aprofundadas, tais como estudos relacionados à percepção da fala e influências de outras variáveis tais como velocidade de fala.

Com relação à influência da velocidade de fala sobre o VOT, vários estudos (Miller e Volaitis, 1989; Volaitis e Miller, 1992; Kessinger e Blumstein, 1997; Klein, 1999) já exibiram a variação desse parâmetro em diferentes velocidades de fala.

Testes perceptuais, partindo dos dados apresentados nesta pesquisa, poderão ser elaborados com a finalidade de se verificar a questão da percepção categórica para as oclusivas do PB.

## Referencial Bibliográfico

ANDRADE, A.; MASCARENHAS, I. **Para um estudo do vozeamento em início de vogal diante de consoante oclusiva**. Centro de Lingüística da Universidade de Lisboa, CIP-Lisboa, 1994.

ASHER, R. E. (Ed.) **The encyclopedia of language and linguistics**. Vol. 6, Pergamon Press, 1994 .

BARTON, D.; MACKEN, M. A. An instrumental analysis of the voicing contrast in word-initial stops in the speech of four-year-old English-speaking children. **Language and Speech**, v. 23, p. 159-169, 1980.

BOND, Z. S.; EDDEY, J. E.; BERMEJO, J. J. VOT del Español to English: comparison of a language-disordered and normal child. **Journal of Phonetics**, v. 8, p. 287-291, 1980.

CAGLIARI, L. C. **Elementos de fonética do português brasileiro**. Campinas: UNICAMP, 1981. Tese (Livre Docência).

CÂMARA Jr., J. M. **Estrutura da língua portuguesa**. 28. ed. Petrópolis: Vozes, 1998.

CARNEY, A. E.; GANDOUR, J.; PETTY, S. H.; ROBBINS, A. M., MYRES, W.; MIYAMOTO, R. The effect of adventitious deafness on the perception and production of voice onset time in Thai: a case study. **Language and Speech**, v. 31, p. 273-282, 1988.

FIGUEIREDO, R. M. **Identificação de falantes: aspectos teóricos e metodológicos.** Campinas: UNICAMP, 1994. Tese (Doutorado).

FLEGE, J.E.; PORT, R. Cross-language phonetic interference: Arabic to English. **Language and Speech**, v. 24, p. 125-146, 1981.

FOKES, J.; BOND, Z. S. & STEINBERG, M. Acquisition of the English voicing contrast by Arab children. **Language and Speech**, v. 23, p. 159-169, 1985.

HALLE, M.; HUGHES, G. W.; RADLEY, J. P. Acoustic properties of stops consonants. In: BAKEN, R.J.; DANILOFF, R. (Ed.) **Readings in clinical spectrography of speech.** San Diego/New Jersey: Singular Publishing Group, Inc & Kay Elemetrics Corp, 1991, p. 216-225.

ISTRE, G. L. (s/d) **Um estudo do VOT em monolíngües brasileiros.** (trabalho não publicado).

JAMIESON, D. G.; CHEESMAN, M. F. The adaptation of produced voice-onset time, v. 15, p. 15-27, 1987

JOHNSON, K. **Acoustic & Auditory Phonetics.** Cambridge: Blackwell Publishers, 1997.

KENT, R. D.; READ, C. **The acoustic analysis of speech.** San Diego: Singular Publishing Group, Inc., 1992.

KESSINGER, R. H.; BLUMSTEIN, S. E. Effects of speaking rate on voice-onset time in Thai, French, and English. **Journal of Phonetics**, v. 25, p. 143-168, 1997.



KONEFAL, J. A. & FOKES, J. Voice onset time: the development of Spanish/English distinction in normal and language disordered children. **Journal of Phonetics**, v. 9, p. 437-444, 1981.

LEVIN, J. **Estatística aplicada a ciências humanas**. 2. ed. São Paulo: Harbra, 1987.

LIEBERMAN, P.; BLUMSTEIN, S. E. **Speech physiology, speech perception, and acoustic phonetics**. Melbourne: Cambridge University Press, 1996.

LISKER, L. & ABRAMSON, A. A cross-language study of voicing in initial stops: Acoustical measurements. **Word**, v. 20, p. 384-422, 1964.

MALMBERG, B. **A fonética**. Lisboa: Livros do Brasil, 1954.

MILLER, J. L.; VOLAITIS, L. E. Effect of speaking rate on the perceptual structure of phonetic category. **Perception & Psychophysics**, v.46, p. 505-512, 1989.

MILLER, J. L. On the internal structure of phonetic categories: a progress report. **Cognition**, v. 50, p. 271-285, 1994.

SHARF, D. J. & OHDE, R. N. Recovery from adaptation to stimuli varying in voice onset time. **Journal of Phonetics**, v. 9, p. 79-87, 1981.

SMITH, B. L. Temporal aspects of English speech production: A developmental perspective. **Journal of Phonetics**, v. 6, p. 37-67, 1978.

USELDING, D. K. A temporal order effect in voice onset time discrimination. **Language and Speech**, v. 20, p. 366-376, 1977.

VOLAITIS, L. E. & MILLER, J. L. Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. **Journal of the Acoustical Society of America**, v. 92, p. 723-735, 1992.

WEISMER, G. Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. **Journal of Phonetics**, v. 7, 1p. 97-204, 1979.

## **Bibliografia Consultada**

BARBETTA, P. A. **Estatística aplicada às ciências humanas**. 2. ed. ver. Florianópolis: Editora da UFSC, 1998.

CHANG, S. & BLUMSTEIN, S. The role of onsets in perception of stop place of articulation: effects of spectral and temporal discontinuity. **Journal of Acoustic Society of America**, v. 70, p. 39-44, 1981.

CLARK, J.; YALLOP, C. **An introduction to phonetics and phonology**. Cambridge/Massachussets: Blasil Balckwell, 1990.

CRYSTAL, D. **A dictionary of linguistics and Phonetic**. 4. ed. Updated and enlarged. Oxford/Massachussets: Balckwell Publishers, 1997.

DUBOIS, J.; GIACOMO, M.; GUESPIN, L. *et al.* **Dicionário de lingüística**. São Paulo: Cultrix, 1993.

FISCHER-JØRGENSEN, E. Acoustic analysis of stops consonants. *Miscellenea Phonetica*, Vol. II, *In*: LEHISTE, I. (ed.). **Readings in acoustic phonetics**. 2. ed. Cambridge, Massachussets: M.I.T. Press, 1996. p. 137-154.

FLEGE, J. E. Laryngeal timing and phonation onset in utterance-initial English stops. **Journal of Phonetics**, v. 10, p. 177-192, 1982.

FLEGE, J. E. The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance. **Journal of Phonetics**, v. 10, p. 335-345, 1982.

FRY, D.B. **The physics of speech**. Cambridge: Cambridge University Press, 1994.

GARMAN, M. **Psycholinguistics**. Cambridge: Cambridge University Press, 1990.

HEFFNER, R. M. S. **General phonetics**. Madison: University of Wisconsin Press, 1964.

KLATT, D. Voice onset time, frication, and aspiration in word-initial consonant clusters. In: BAKEN, R.; DANILOFF, R. G. (ed.) **Readings in clinical spectrography of speech**. San Diego/New Jersey: Singular Publishing Group, Inc. & Kay Elemetrics Corp, 1991.

LAVER, T. **Principles of phonetics**. New York: Cambridge University Press, 1994.

LISKER, L.; LIBERMAN, A. M. & ERICKSON, D. M. On publishing the voice-onset-time (VOT) boundary about. **Language and Speech**, v. 20, p. 209-216, 1977.

LÖFQVIST, A. Acoustic and aerodynamic effects of interarticulator timing in voiceless consonants. **Language and Speech**, v. 35, p. 15-28, 1992.

MILLER, J. L.; CONNINE, C. M.; SCHERMER, T. M. & KLUENDER, K. R. A possible auditory basis for internal structure of phonetic categories. **Journal of Acoustic Society of America**, v. 73, p. 2124-2133, 1983.

SAERENS, M; SERNICLAES, W. & BEECKMANS, R. Acoustic versus contextual factors in stops voicing perception in spontaneous French. **Language and Speech**, v. 32, p. 291-314, 1989.

SCHIAVETTI, N.; WHITEHEAD, R. L.; METZ, D. E.; WHITEHEAD, B. & MIGNEREY, M. Voice onset time in speech produced during simultaneous communication. **Journal of Speech and Hearing Research**, v. 39, p. 565-572, 1996.

SOCK, R.; LÖFQVIST, A. Some timing constraints in the production of bilabial stops. **Journal of Phonetics**, v. 23, p. 129-138, 1995.

## Apêndice 1

*Corpus* utilizado para a obtenção do material sonoro:

### Logatomas

/ba/	/be/	/bo/
/pa/	/pe/	/po/
/da/	/de/	/do/
/ta/	/te/	/to/
/ga/	/ge/	/go/
/ka/	/ke/	/ko/

### Frases-veículo

Digo <i>ba</i> pra ela.	Digo <i>be</i> pra ela.	Digo <i>bo</i> pra ela.
Digo <i>pa</i> pra ela.	Digo <i>pe</i> pra ela.	Digo <i>po</i> pra ela.
Digo <i>da</i> pra ela.	Digo <i>de</i> pra ela.	Digo <i>do</i> pra ela.
Digo <i>ta</i> pra ela.	Digo <i>te</i> pra ela.	Digo <i>to</i> pra ela.
Digo <i>ga</i> pra ela.	Digo <i>ge</i> pra ela.	Digo <i>go</i> pra ela.
Digo <i>ka</i> pra ela.	Digo <i>ke</i> pra ela.	Digo <i>ko</i> pra ela.

### Palavras

Contexto pré-tônico

/ba/ - <i>baguete</i>	/be/ - <i>Bebeto</i>	/bo/ - <i>bodoque</i>
/pa/ - <i>pataca</i>	/pe/ - <i>aperiódico</i>	/po/ - <i>hipopótamo</i>

/da/ - predador	/de/ - decoro	/do/ - adoçante
/ta/ - tacape	/te/ - tequila	/to/ - topete
/ga/ - gadelha	/ge/ - guedelha	/go/ - logotipo
/ka/ - cabala	/ke/ - esqueleto	/ko/ - cobiça

## Contexto tônico

/ba/ - cabala	/be/ - bêbedo	/bo/ - boba
/pa/ - tapada	/pe/ - tapete	/po/ - popa
/da/ - didata	/de/ - guedelha	/do/ - doce
/ta/ - pataca	/te/ - Atenas	/to/ - atômico
/ga/ - legado	/ge/ - gueto	/go/ - agogô
/ka/ - tacape	/ke/ - etiqueta	/ko/ - decoro

## Contexto pós-tônico

/ba/ - bêbado	/be/ - bêbedo	/bo/ - abóbora
/pa/ - popa	/pe/ - bípede	/po/ - dípede
/da/ - tapada	/de/ - nádega	/do/ - ídolo
/ta/ - didata	/te/ - apótema	/to/ - cátodo
/ga/ - cágado	/ge/ - Águeda	/go/ - dígono
/ka/ - pataca	/ke/ - tíquete	/ko/ - cócoras

**Palavras inseridas em sentenças:**

Tapada pega aquela mulher do seu José Carlos.

A tapada pega na mulher do seu José Carlos.

Aquela mulher era tapada demais da conta.

O senhor José Carlos tinha uma mulher tapada.

Cabala lida pelo rabino será sagrada.  
A *cabala* lido pelos rabinos é sagrada.  
O rabino lia na *cabala* que é sagrada.  
O rabino tava lendo na sagrada *cabala*.

*Pataca* foi uma moeda de prata antiga.  
A *pataca* foi a moeda de prata antiga.  
Moeda de prata foi *pataca* que é antiga.  
A moeda de prata antiga foi a *pataca*.

*Didata* foi autor deste livro maravilhoso.  
O *didata* foi autor do livro maravilhoso.  
O carisma daquele *didata* conquistou todos.  
Este livro chocante foi trabalho de *didata*.

*Tacape* seria daquela forma particular.  
O *tacape* será daquela forma particular.  
Era dessa forma o *tacape* resplandecente.  
Era dessa forma o resplandecente *tacape*.

Legado deixado pra você vai ser nosso saber.  
O legado deixado por nós vai ser nosso saber.  
Deixaremos como seu legado nosso dinheiro.  
Deixaremos nosso dinheiro como nosso legado.



## Apêndice 2

### Tabelas Estatísticas

Oclusiva bilabial [b]						
Variável	Média	DP	N	t	gl	p <sup>1</sup>
Log.isol.	-86.57	26.67	96	6.14576	95	.000000
Log. fv	-104.98	14.01				

Tabela 1 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora inserida em logatomas isolados e em logatomas produzidos em frases-veículo

Oclusiva dental [d]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-99.20	25.47	93	-2.17017	92	.032569
Log. fv	-93.31	12.34				

Tabela 2 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora inserida em logatomas isolados e em logatomas produzidos em frases-veículo

Oclusiva velar [g]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-91.16	33.59	92	-3.83356	91	.000232
Log. fv	-76.93	24.18				

Tabela 3 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora inserida em logatomas isolados e em logatomas produzidos em frases-veículo

<sup>1</sup> Os dados são significativos quando  $p < 0,05$  e, para uma melhor visualização, estes serão marcados pela cor cinza.

Oclusiva bilabial [p]						
Variável	Média	DP	N	t	gl	p
Log.isol.	17.70	8.67	96	3.58965	95	.000526
Log. fv	15.49	5.56				

Tabela 4 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda inserida em logatomas isolados e em logatomas produzidos em frases-veículo

Oclusiva dental [t]						
Variável	Média	DP	N	t	gl	p
Log.isol.	16.78	4.54	96	-.95057	95	.344233
Log. fv	17.13	4.90				

Tabela 5 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda inserida em logatomas isolados e em logatomas produzidos em frases-veículo

Oclusiva velar [k]						
Variável	Média	DP	N	t	gl	p
Log.isol.	37.82	10.04	93	4.38522	92	.000031
Log. fv	33.90	7.48				

Tabela 6 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[ba]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-91.84	28.60	31	3.04664	30	.004792
Log. fv	-106.74	12.71				

Tabela 7 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, em contexto vocálico [a], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[be]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-81.43	22.31	32	5.27695	31	.000010
Log. fv	-108.80	15.80				

Tabela 8 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, em contexto vocálico [e], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[bo]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-87.92	27.87	32	2.42188	31	.021482
Log. fv	-99.25	11.92				

Tabela 9 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, em contexto vocálico [o], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[da]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-89.15	24.51	32	-1.12233	31	.270340
Log. fv	-85.15	15.05				

Tabela 10 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, em contexto vocálico [a], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[de]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-111.71	24.92	29	-2.34179	28	.026534
Log. fv	-100.42	25.43				

Tabela 11 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, em contexto vocálico [e], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[do]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-96.85	21.93	32	-.44357	31	.660436
Log. fv	-95.02	20.53				

Tabela 12 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, em contexto vocálico [o], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[ga]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-91.50	35.48	28	-2.78261	27	.009720
Log. fv	-68.11	26.11				

Tabela 13 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, em contexto vocálico [a], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[ge]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-99.14	33.56	32	-2.95757	31	.005887
Log. fv	-81.73	21.11				

Tabela 14 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, em contexto vocálico [e], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[go]						
Variável	Média	DP	N	t	gl	p
Log.isol.	-82.38	30.97	32	-.53336	31	.597588
Log. fv	-79.84	23.97				

Tabela 15 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, em contexto vocálico [o], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[pa]						
Variável	Média	DP	N	t	gl	p
Log.isol.	11.34	3.67	32	-1.99904	31	.054436
Log. fv	12.51	4.09				

Tabela 16 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, em contexto vocálico [a], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[pe]						
Variável	Média	DP	N	t	gl	p
Log.isol.	13.98	5.30	32	-.15799	31	.875491
Log. fv	14.11	3.94				

Tabela 17 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, em contexto vocálico [e], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[po]						
Variável	Média	DP	N	t	gl	p
Log.isol.	27.77	5.32	32	8.45496	31	.000000
Log. fv	19.84	5.64				

Tabela 18 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, em contexto vocálico [o], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[ta]						
Variável	Média	DP	N	t	gl	p
Log.isol.	14.03	3.59	32	-1.82083	31	.078298
Log. fv	15.07	4.03				

Tabela 19 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, em contexto vocálico [a], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[te]						
Variável	Média	DP	N	t	gl	p
Log.isol.	18.64	5.04	32	-.883113	31	.383973
Log. fv	19.19	4.78				

Tabela 20 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, em contexto vocálico [e], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[to]						
Variável	Média	DP	N	t	gl	p
Log.isol.	17.67	3.54	32	.77986	31	.441389
Log. fv	17.13	50.6				

Tabela 21 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, em contexto vocálico [o], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[ka]						
Variável	Média	DP	N	t	gl	p
Log.isol.	33.84	12.23	32	2.248528	31	.031794
Log. fv	29.68	6.34				

Tabela 22 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, em contexto vocálico [a], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[ke]						
Variável	Média	DP	N	t	gl	p
Log.isol.	40.67	10.05	32	2.784514	31	.009057
Log. fv	36.37	8.54				

Tabela 23 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, em contexto vocálico [e], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

[ko]						
Variável	Média	DP	N	t	gl	p
Log. isol.	39.07	4.97	29	2.84276	28	.008256
Log. fv	35.83	5.23				

Tabela 24 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, em contexto vocálico [o], inserida em logatomas isolados e em logatomas produzidos em frases-veículo

Oclusiva bilabial [b]						
Variável	Média	DP	N	t	gl	p
[ba]	-91.84	28.60	31	-1.81092	30	.080178.
[be]	-82.61	21.64				
[ba]	-91.84	28.60	31	-.62707	30	.535355
[bo]	-88.01	28.33				
[be]	-81.43	22.31	32	1.33430	31	.191824
[bo]	-87.92	27.87				

Tabela 25 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, em diferentes contextos vocálicos, inserida em logatomas produzidos isoladamente

Oclusiva dental [d]						
Variável	Média	DP	N	t	gl	p
[da]	-89.15	24.51	32	4.59728	31	.000068
[de]	-110.51	24.94				
[da]	-89.15	24.51	32	1.4533	31	.157011
[do]	-96.85	21.93				
[de]	-110.51	24.94	32	-3.55132	31	.001248
[do]	-96.85	21.93				

Tabela 26 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, em diferentes contextos vocálicos, inserida em logatomas produzidos isoladamente

Oclusiva velar [g]						
Variável	Média	DP	N	t	gl	p
[ga]	-88.46	35.37	31	1.71991	30	.095752
[ge]	-100.40	33.34				
[ga]	-88.46	35.37	31	-.77008	30	.447275
[go]	-83.62	30.66				
[ge]	-99.14	33.56	32	-3.08430	31	.004266
[go]	-82.38	30.97				

Tabela 27 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, em diferentes contextos vocálicos, inserida em logatomas produzidos isoladamente

Oclusiva bilabial [p]						
Variável	Média	DP	N	t	gl	p
[pa]	11.34	3.67	32	-3.9394	31	.000432
[pe]	13.98	5.30				
[pa]	11.34	3.67	32	-14.6469	31	.000000
[po]	27.77	5.32				
[pe]	13.98	5.30	32	-10.3554	31	.000000
[po]	27.77	5.32				

Tabela 28 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, em diferentes contextos vocálicos, inserida em logatomas produzidos isoladamente

Oclusiva dental [t]						
Variável	Média	DP	N	t	gl	p
[ta]	14.03	3.59	32	-6.98816	31	.000000
[te]	18.64	5.04				
[ta]	14.03	3.59	32	-6.39258	31	.000000
[to]	17.67	3.54				
[te]	18.64	5.04	32	1.28184	31	.209402
[to]	17.67	3.54				

Tabela 29 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, em diferentes contextos vocálicos, inserida em logatomas produzidos isoladamente



Oclusiva velar [k]						
Variável	Média	DP	N	t	gl	p
[ka]	33.84	12.23	32	-3.45423	31	.001620
[ke]	40.67	10.05				
[ka]	33.84	12.23	32	-2.40138	31	.22518
[ko]	39.04	4.88				
[ke]	40.67	10.05	32	.89087	31	.379862
[ko]	39.04	4.88				

Tabela 30 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, em diferentes contextos vocálicos, inserida em logatomas produzidos isoladamente

Oclusiva bilabial [b]						
Variável	Média	DP	N	t	gl	p
[ba]	-106,88	12.53	32	.75178	31	.457854
[be]	-108,80	15.80				
[ba]	-106.88	12.53	32	-2.68562	31	.011524
[bo]	-99.25	11.92				
[be]	-108.80	15.80	32	-3.14807	31	.003621
[bo]	-99.25	11.92				

Tabela 31 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, em diferentes contextos vocálicos, inserida em logatomas produzidos em frases-veículo

Oclusiva dental [d]						
Variável	Média	DP	N	t	gl	p
[da]	-86.24	15.41	29	3.49793	28	.001585
[de]	-100.42	25.43				
[da]	-85.15	15.05	32	3.94916	31	.000421
[do]	-95.02	20.53				
[de]	-100.42	25.43	29	1.17036	28	.251723
[do]	-95.81	21.36				

Tabela 32 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, em diferentes contextos vocálicos, inserida em logatomas produzidos em frases-veículo

Oclusiva velar [g]						
Variável	Média	DP	N	t	gl	p
[ga]	-68.11	26.11	28	5.989432	27	.000002
[ge]	-83.64	21.63				
[ga]	-68.11	26.11	28	4.551233	27	.000102
[go]	-83.08	23.77				
[ge]	-81.73	21.11	32	-.631376	31	.532423
[go]	-79.84	23.97				

Tabela 33 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, em diferentes contextos vocálicos, inserida em logotomas produzidos em frases-veículo

Oclusiva bilabial [p]						
Variável	Média	DP	N	t	gl	p
[pa]	12.51	4.09	32	-2.45173	31	.020051
[pe]	14.11	3.94				
[pa]	12.51	4.09	32	-7.43420	31	.000000
[po]	19.84	5.64				
[pe]	14.11	3.94	32	-6.95492	31	.000000
[po]	19.84	5.64				

Tabela 34 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, em diferentes contextos vocálicos, inserida em logotomas produzidos em frases-veículo

Oclusiva dental [t]						
Variável	Média	DP	N	t	gl	p
[ta]	15.07	4.03	32	-5.54513	31	.000004
[te]	19.19	4.78				
[ta]	15.07	4.03	32	-2.88197	31	.007116
[to]	17.13	5.06				
[te]	19.19	4.78	32	3.76971	31	.000690
[to]	17.13	5.06				

Tabela 35 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, em diferentes contextos vocálicos, inserida em logotomas produzidos em frases-veículo

Oclusiva velar [k]						
Variável	Média	DP	N	t	gl	p
[ka]	29.68	6.34	32	-4.87952	31	.000030
[ke]	36.37	8.54				
[ka]	29.62	6.53	29	-5.56697	28	.000006
[ko]	35.83	5.23				
[ke]	36.57	8.96	29	.49335	28	.625611
[ko]	35.83	5.23				

Tabela 36 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, em diferentes contextos vocálicos, inserida em logatomas produzidos em frases-veículo

Oclusiva bilabial [b]						
Variável	Média	DP	N	t	gl	p
pretônico	-59.01	20.65	77	8.17257	76	.000000
tônico	-85.36	19.84				
pretônico	-59.01	20.65	77	5.93855	76	.000000
postônico	-73.50	12.04				
tônico	-85.30	19.68	91	-5.35715	90	.000001
postônico	-72.27	11.89				

Tabela 37 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora em diferentes contextos de tonicidade

Oclusiva dental [d]						
Variável	Média	DP	N	t	gl	p
pretônico	-54.58	18.55	91	10.0127	90	.000000
tônico	-85.99	21.53				
pretônico	-55.41	18.95	84	8.0785	83	.000000
postônico	-74.20	11.53				
tônico	-86.36	22.05	84	-4.4743	83	.000024
postônico	-74.20	11.53				

Tabela 38 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora em diferentes contextos de tonicidade

Oclusiva velar [g]						
Variável	Média	DP	N	t	gl	p
pretônico	-51.37	16.02	85	7.59056	84	.000000
tônico	-66.81	18.84				
pretônico	-51.39	16.17	83	-.15095	82	.880389
postônico	-51.12	11.43				
tônico	-66.91	18.94	83	-6.80230	82	.000000
postônico	-51.12	11.43				

Tabela 39 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora em diferentes contextos de tonicidade

Oclusiva bilabial [p]						
Variável	Média	DP	N	t	gl	p
pretônico	12.98	3.88	96	-1.19727	95	.234181
tônico	13.56	4.62				
pretônico	12.98	3.88	96	-7.83814	95	.000000
postônico	17.05	5.81				
tônico	13.56	4.62	96	-6.45877	95	.000000
postônico	17.05	5.81				

Tabela 40 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda em diferentes contextos de tonicidade

Oclusiva dental [t]						
Variável	Média	DP	N	t	gl	p
pretônico	16.09	5.07	96	-.12361	95	.901886
tônico	16.16	5.96				
pretônico	15.57	4.70	82	-6.40668	81	.000000
postônico	20.11	7.90				
tônico	15.78	6.25	82	-6.78747	81	.000000
postônico	20.11	7.90				

Tabela 41 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda em diferentes contextos de tonicidade

Oclusiva velar [k]						
Variável	Média	DP	N	t	gl	p
pretônico	33.53	7.90	95	-3.59328	94	.000522
tônico	37.24	9.29				
pretônico	33.53	7.90	95	.58102	94	.562617
postônico	33.00	9.09				
tônico	37.34	9.30	96	3.46008	95	.000810
postônico	33.10	9.10				

Tabela 42 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda em diferentes contextos de tonicidade

[ba]						
Variável	Média	DP	N	t	gl	p
pretônico	-56.87	18.34	21	10.9927	20	.000000
tônico	-99.75	6.52				
pretônico	-56.87	18.34	21	3.2744	20	.003793
postônico	-72.95	9.86				
tônico	-95.55	11.79	31	-11.9976	30	.000000
postônico	-71.25	9.57				

Tabela 43 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, diante da vogal baixa [a], em diferentes contextos de tonicidade

[be]						
Variável	Média	DP	N	t	gl	p
pretônico	-59.41	24.46	25	2.82234	24	.009426
tônico	-76.26	20.33				
pretônico	-59.41	24.46	25	2.84551	24	.008930
postônico	-76.28	10.81				
tônico	-74.76	19.12	30	-.09008	29	.928841
postônico	-74.41	10.82				

Tabela 44 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, diante da vogal média anterior [e], em diferentes contextos de tonicidade

[bo]						
Variável	Média	DP	N	t	gl	p
pretônico	-61.23	19.49	29	5.06182	28	.000023
tônico	-85.71	21.80				
pretônico	-60.70	19.37	30	2.76815	29	.009720
postônico	-72.23	14.53				
tônico	-85.71	21.80	29	-2.82544	28	.008610
postônico	-72.02	14.74				

Tabela 45 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, diante da vogal média posterior [o], em diferentes contextos de tonicidade

[da]						
Variável	Média	DP	N	t	gl	p
pretônico	-50.02	7.92	27	18.44	26	.000000
tônico	-83.36	9.84				
pretônico	-50.02	7.92	27	15.37	26	.000000
postônico	-77.08	9.21				
tônico	-80.81	11.76	31	-2.1549	30	.039321
postônico	-75.93	9.30				

Tabela 46 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, diante da vogal baixa [a], em diferentes contextos de tonicidade

[de]						
Variável	Média	DP	N	t	gl	p
pretônico	-58.43	22.00	32	13.29	31	.000000
tônico	-101.56	19.14				
pretônico	-58.43	22.00	32	3.5669	31	.001197
postônico	-70.22	11.13				
tônico	-101.56	19.14	32	-11.79	31	.000000
postônico	-70.22	11.13				

Tabela 47 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, diante da vogal média anterior [e], em diferentes contextos de tonicidade

[do]						
Variável	Média	DP	N	t	gl	p
pretônico	-57.49	9.58	32	3.43073	31	.001724
tônico	-71.14	22.51				
pretônico	-61.28	6.32	21	5.12807	20	.000051
postônico	-77.68	13.69				
tônico	-73.99	23.57	21	.54878	20	.589231
postônico	-77.68	13.69				

Tabela 48 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, diante da vogal média posterior [o], em diferentes contextos de tonicidade

[ga]						
Variável	Média	DP	N	t	gl	p
pretônico	-51.12	16.35	28	2.57105	27	.015968
tônico	-57.86	15.05				
pretônico	-51.12	16.35	28	1.67554	27	.105373
postônico	-55.20	11.78				
tônico	-57.86	15.05	28	-1.32481	27	.196338
postônico	-55.20	11.78				

Tabela 49 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, diante da vogal baixa [a], em diferentes contextos de tonicidade

[ge]						
Variável	Média	DP	N	t	gl	p
pretônico	-56.64	19.17	31	3.03704	30	.004909
tônico	-67.97	23.98				
pretônico	-57.32	19.11	30	-1.18991	29	.243732
postônico	-53.25	7.37				
tônico	-67.96	24.39	30	-3.41601	29	.001900
postônico	-53.25	7.37				

Tabela 50 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, diante da vogal média anterior [e], em diferentes contextos de tonicidade

[go]						
Variável	Média	DP	N	t	gl	p
pretônico	-45.00	7.68	25	15.9860	24	.000000
tônico	-76.08	8.43				
pretônico	-44.68	7.72	22	.1235	21	.902886
postônico	-45.13	12.30				
tônico	-76.28	8.81	22	-9.4088	21	.000000
postônico	-45.13	12.30				

Tabela 51 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, diante da vogal média posterior [o], em diferentes contextos de tonicidade

[pa]						
Variável	Média	DP	N	t	gl	p
pretônico	10.78	2.52	32	-1.51841	31	.139044
tônico	11.50	2.75				
pretônico	10.78	2.52	32	-8.05083	31	.000000
postônico	17.18	5.43				
tônico	11.50	2.75	32	-6.88152	31	.000000
postônico	17.18	5.43				

Tabela 52 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, diante da vogal baixa [a], em diferentes contextos de tonicidade

[pe]						
Variável	Média	DP	N	t	gl	p
pretônico	12.67	3.11	32	.873721	31	.388989
tônico	12.26	1.87				
pretônico	12.67	3.11	32	.635795	31	.529576
postônico	12.41	3.20				
tônico	12.26	1.87	32	-.293238	31	.771295
postônico	12.41	3.20				

Tabela 53 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, diante da vogal média anterior [e], em diferentes contextos de tonicidade



[po]						
Variável	Média	DP	N	t	gl	p
pretônico	15.49	4.31	32	-1.11695	31	.272598
tônico	16.91	6.05				
pretônico	15.49	4.31	32	-6.95108	31	.000000
postônico	21.56	4.52				
tônico	16.91	6.05	32	-4.26496	31	.000174
postônico	21.56	4.52				

Tabela 54 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, diante da vogal média posterior [o], em diferentes contextos de tonicidade

[ta]						
Variável	Média	DP	N	t	gl	p
pretônico	16.39	3.62	32	2.64583	31	.012684
tônico	14.23	5.48				
pretônico	16.39	3.62	32	-2.53818	31	.016385
postônico	19.12	6.75				
tônico	14.23	5.48	32	-5.91787	31	.000002
postônico	19.12	6.75				

Tabela 55 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, diante da vogal baixa [a], em diferentes contextos de tonicidade

[te]						
Variável	Média	DP	N	t	gl	p
pretônico	15.98	6.24	32	-2.45760	31	.019780
tônico	18.43	7.03				
pretônico	15.98	6.24	32	-8.07933	31	.000000
postônico	23.35	8.49				
tônico	18.43	7.03	32	-5.02354	31	.000020
postônico	23.35	8.49				

Tabela 56 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, diante da vogal média anterior [e], em diferentes contextos de tonicidade

[to]						
Variável	Média	DP	N	t	gl	p
pretônico	16.21	5.40	32	.43714	31	.665040
tônico	15.81	4.49				
pretônico	13.94	3.38	18	-1.50148	17	.151577
postônico	16.11	6.64				
tônico	13.83	4.42	18	-1.27643	17	.218965
postônico	16.11	6.64				

Tabela 57 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, diante da vogal média posterior [o], em diferentes contextos de tonicidade

[ka]						
Variável	Média	DP	N	t	gl	p
pretônico	30.69	5.58	32	.43400	31	.667293
tônico	30.22	6.56				
pretônico	30.69	5.58	32	-3.50330	31	.004120
postônico	34.75	9.56				
tônico	30.22	6.56	32	-2.68950	31	.011417
postônico	34.75	9.56				

Tabela 58 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, diante da vogal baixa [a], em diferentes contextos de tonicidade

[ke]						
Variável	Média	DP	N	t	gl	p
pretônico	34.69	9.76	31	-1.43844	30	.160661
tônico	37.40	7.43				
pretônico	34.69	9.76	31	2.12912	30	.041567
postônico	30.75	9.88				
tônico	37.63	7.42	32	3.32202	31	.002300
postônico	31.02	9.84				

Tabela 59 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, diante da vogal média anterior [e], em diferentes contextos de tonicidade

[ko]						
Variável	Média	DP	N	t	gl	p
pretônico	35.24	7.31	32	-4.56874	31	.000074
tônico	44.16	8.16				
pretônico	35.24	7.31	32	1.15018	31	.258865
postônico	33.54	7.61				
tônico	44.16	8.16	32	6.14789	31	.000001
postônico	33.54	7.61				

Tabela 60 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, diante da vogal média posterior [o], em diferentes contextos de tonicidade

[ba]						
Variável	Média	DP	N	t	gl	p
inicial (0)	-84.98	12.30	11	-1.34077	10	.209656
inicial (1)	-79.20	8.57				
inicial (0)	-84.00	12.21	12	1.16821	11	.267418
medial (1)	-87.92	10.21				
inicial (0)	-86.62	11.63	10	-2.07646	9	.067652
final	-77.89	4.61				
inicial (1)	-79.20	8.57	11	1.93230	10	.082121
medial	-88.42	10.55				
inicial (1)	-80.31	8.16	10	-1.16710	9	.273160
final	-77.89	4.61				
medial	-88.50	11.12	10	-2.57753	9	.029817
final	-77.89	4.61				

Tabela 61 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial sonora, em contexto vocálico [a], em diferentes posições em sentenças

[da]						
Variável	Média	DP	N	t	gl	p
inicial (0)	-67.89	12.01	12	.48737	11	.635562
inicial (1)	-69.47	14.18				
inicial (0)	-67.89	12.01	12	.27912	11	.785333
medial (1)	-68.79	6.69				
inicial (0)	-67.89	12.01	12	-.83640	11	.420728
final	-65.79	10.92				
inicial (1)	-69.47	14.18	12	-.19886	11	.846000
medial	-68.79	6.69				
inicial (1)	-69.47	14.18	12	-1.20092	11	.255005
final	-65.79	10.92				
medial	-68.79	6.69	12	-1.10089	11	.294450
final	-65.79	10.92				

Tabela 62 - Parâmetros estatísticos referentes ao VOT da oclusiva dental sonora, em contexto vocálico [a], em diferentes posições em sentenças

[ga]						
Variável	Média	DP	N	t	gl	p
inicial (0)	-47.34	8.96	8	-.52726	7	.614304
inicial (1)	-44.58	11.59				
inicial (0)	-47.34	8.96	8	.01235	7	.990490
medial (1)	-47.39	7.41				
inicial (0)	-47.34	8.96	8	.52876	7	.613320
final	-49.58	5.94				
inicial (1)	-44.61	10.84	9	1.01600	8	.339369
medial	-48.24	7.39				
inicial (1)	-43.94	10.44	10	1.88107	9	.092648
final	-49.85	5.86				
medial	-48.24	7.39	9	.32743	8	.751744
final	-49.13	5.72				

Tabela 63 - Parâmetros estatísticos referentes ao VOT da oclusiva velar sonora, em contexto vocálico [a], em diferentes posições em sentenças

[pa]						
Variável	Média	DP	N	t	gl	p
inicial (0)	11.59	2.98	12	1.36617	11	.199163
inicial (1)	11.20	2.82				
inicial (0)	11.59	2.98	12	-2.49933	11	.029541
medial (1)	13.12	4.16				
inicial (0)	11.59	2.98	12	-.48281	11	.638699
final	12.00	3.32				
inicial (1)	11.20	2.82	12	-3.12592	11	.009648
medial	13.12	4.16				
inicial (1)	11.20	2.82	12	-1.20232	11	.254484
final	12.00	3.32				
medial	13.12	4.16	12	1.07960	11	.303422
final	12.00	3.32				

Tabela 64 - Parâmetros estatísticos referentes ao VOT da oclusiva bilabial surda, em contexto vocálico [a], em diferentes posições em sentenças

[ta]						
Variável	Média	DP	N	t	gl	p
inicial (0)	12.54	4.57	12	-.50127	11	.626060
inicial (1)	12.98	4.61				
inicial (0)	12.54	4.57	12	.85286	11	.411923
medial (1)	11.95	4.21				
inicial (0)	12.54	4.57	12	-1.16047	11	.270422
final	13.72	5.47				
inicial (1)	12.98	4.61	12	.98927	11	.343794
medial	11.95	4.21				
inicial (1)	12.98	4.61	12	-.49693	11	.629023
final	13.72	5.47				
medial	11.95	4.21	12	-2.61113	11	.024208
final	13.72	5.48				

Tabela 65 - Parâmetros estatísticos referentes ao VOT da oclusiva dental surda, em contexto vocálico [a], em diferentes posições em sentenças

[ka]						
Variável	Média	DP	N	t	gl	p
inicial (0)	28.68	4.00	12	.30032	11	.769533
inicial (1)	28.30	6.66				
inicial (0)	28.68	4.00	12	1.07154	11	.306874
medial (1)	27.78	5.42				
inicial (0)	28.68	4.00	12	-2.01471	11	.069040
final	30.93	5.65				
inicial (1)	28.30	6.66	12	.33499	11	.743937
medial	27.78	5.42				
inicial (1)	28.30	6.66	12	-2.00965	11	.069642
final	30.93	5.65				
medial	27.78	5.42	12	-3.25803	11	.007626
final	30.93	5.65				

Tabela 66 - Parâmetros estatísticos referentes ao VOT da oclusiva velar surda, em contexto vocálico [a], em diferentes posições em sentenças