

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE COMUNICAÇÃO E EXPRESSÃO
DEPARTAMENTO ACADÊMICO DE LÍNGUAS ESTRANGEIRAS MODERNAS
CURSO DE LICENCIATURA EM LETRAS PORTUGUÊS/INGLÊS

CAMILLE FERNANDA CORTIANO NAGAKURA

**COMPARAÇÃO DE VOZ A PARTIR DO TEMPO DE ATAQUE DE
VOZEAMENTO (VOT)**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA

2014

CAMILLE FERNANDA CORTIANO NAGAKURA

**COMPARAÇÃO DE VOZ A PARTIR DO TEMPO DE ATAQUE DE
VOZEAMENTO (VOT)**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção de título de licenciado em Licenciatura em Letras Português/Inglês da Universidade Tecnológica Federal do Paraná.

Professor: Prof. Dr. Roberlei Bertucci.

Orientadora: Prof^a. Dra. Maria Lúcia de Castro Gomes

CURITIBA

2014

RESUMO

NAGAKURA, Camille Fernanda C. **Comparação de voz a partir do tempo de ataque de vozeamento (VOT)**. 2015. págs. Trabalho de Conclusão de Curso de Licenciatura em Letras Português/Inglês - Universidade Tecnológica Federal do Paraná.

A fonética trabalha com a fala em três áreas (KENT; READ, 2002) – como ela é produzida (fonética articulatória ou fisiológica), como ela é transmitida acusticamente (fonética acústica) e como ela é percebida (fonética perceptiva). Fonética Forense, área norteadora deste trabalho, analisa a produção e a percepção dos sons da fala, fazendo uso da tecnologia com uma perspectiva interdisciplinar. No presente projeto, a perspectiva focada será a acústica, analisando o segmento da plosiva bilabial não vozeada, explorando também a perspectiva perceptiva, comparando voz normal e voz disfarçada. Primeiramente será feita uma análise teórica das consoantes plosivas, e do Tempo de Ataque de Vozeamento (VOT, na sigla em inglês). Considerando isso, o trabalho tem como principal objetivo analisar o VOT da plosiva /p/ em início de palavras, que pode ser um parâmetro importante para análises forenses. Também visou-se identificar a diferença em voz normal e voz disfarçada comparando vozes femininas e masculinas. A partir dos resultados, foi possível relacionar a estratégia de disfarce com as alterações nos valores de VOT. A estratégia de voz áspera, por exemplo, resultou em diminuição de VOT, enquanto que a imitação de sotaque americano ocasionou aumento de VOT.

Palavras-chave: Acústica. Fonética. Fonética Forense. VOT. Consoantes plosivas. Disfarce. Comparação de voz.

ABSTRACT

NAGAKURA, Camille Fernanda C. **Voice Onset Time (VOT) as a Parameter for Voice Comparison**. 2015. págs. Trabalho de Conclusão de Curso de Licenciatura em Letras Português/Inglês - Universidade Tecnológica Federal do Paraná.

The three areas that are embraced by phonetics are the acoustic phonetics, the articulatory phonetics and the auditory phonetics. Forensic Phonetics, one of the guiding areas of this work analyses the production and the perception of the sounds that are produced during the act of speaking, working with a multidisciplinary perspective. This project focuses in the acoustic area, analysing the stop unvoiced consonants. firstly, was presented a theoretical analysis of the stop consonants, comparing the disguised and the normal voice. Secondly, an important cue in speech recognition, the Voice Onset Time, was defined. Taking this under consideration, the main goal of this project was to verify a pattern for voice recognition by analysing the VOT of the stop consonant /p/ at initial position. After that, the differences between feminine and masculine undisguised and disguised voices were compared. The software Praat (BOERSMA; WEENINK, 2014) was used to define the data. The final analysis was made taking the kind of disguise under consideration.

Keywords: *Phonetics. Forensic Phonetics. VOT. Stop consonants. Disguise. Voice comparison.*

LISTA DE TABELAS

Tabela 1 - Médias de VOT do Grupo Controle

Tabela 2 - Médias de VOT do Grupo Mulheres

Tabela 3 - Médias de VOT do Grupo Homens

Tabela 4 – Disfarces estabelecidos a partir do BP - VPAS

Tabela 5 - Medidas de VOT do Grupo Controle

Tabela 6 - Medidas de VOT do Grupo Mulheres

Tabela 7 - Medidas de VOT do Grupo Homens

SUMÁRIO

1.0 Introdução	8
2.0 Fonética Forense	10
2.1 A Fonética Forense e a Identificação de Locutor	10
2.2 Disfarce	13
3.0 <i>Voice Onset Time</i> (VOT)	14
3.1 Consoantes Oclusivas e o VOT	15
3.2 O VOT no Português Brasileiro (PB)	17
4.0 Metodologia	19
4.1 Informantes	19
4.2 Coleta de Dados.....	19
4.3 Estratégias de Disfarce.....	20
4.4 Procedimentos de medidas	22
4.5 Resultados	24
5.0 Considerações sobre os dados	26
6.0 Conclusão.....	30
Referências.....	31

1.0 Introdução

No capítulo introdutório do livro *Acoustic Analysis of Speech* (KENT; READ, 2002) os autores respondem à pergunta “What is speech?”. De acordo com eles a fala (*speech*) é uma modalidade da linguagem. A comunicação por meio da voz é comum para quase todos os seres humanos (exceto aos surdos). Kent e Read definem o sinal acústico como o produto final da fala.

Dentro da fonética, a fala é trabalhada em três áreas: como ela é produzida (fonética articulatória ou fisiológica), como é transmitida acusticamente (fonética acústica) e, por fim, como a fala é percebida/recebida (fonética perceptiva) (KENT; READ, 2002).

Como ressalta Gillier (2011), tentou-se por vários anos separar a análise perceptiva da análise acústica, mas atualmente usam-se as abordagens de maneira conjunta. Os parâmetros acústicos são de extrema importância, uma vez que a Fonética Acústica permite fazer uma análise visual, quantitativa e precisa dos sinais que se percebem auditivamente e que caracterizam uma voz.

No presente trabalho, o foco será na área acústica da fonética. O sinal acústico da fala é “o evento físico transmitido em telecomunicações ou gravado em diferentes mídias” (KENT; READ, 2002)¹. Tal sinal contém a mensagem da fala, ao passo que o ouvinte pode reconhecer a mensagem ao ouvi-la. Quando se relaciona a identidade de um indivíduo a uma voz, entra-se em áreas como a Fonética Forense, por exemplo (ERIKSSON, 2010).

O foneticista forense se preocupa com todos os aspectos do discurso, visando usá-los como evidências (COULTHARD; JOHNSON, 2007). A perícia em Fonética Acústica Forense tem por base gravações realizadas principalmente através de escutas telefônicas (MARTINS *et al*, 2014). A importância da perícia não é questionável, a segurança que ela passa, entretanto, depende da competência dos peritos e da confiabilidade dos resultados da análise. Para que isso aconteça é preciso que os conhecimentos das propriedades da linguagem sejam ampliados, sendo

¹ “*The acoustic signal of speech is physical event that is transmitted in telecommunications or is recorded in magnetic tape, laser disc, or other mediums*” (Kent; Read, 2002, p.9).

necessários parâmetros objetivos para caracterizar e identificar o falante. De acordo com Martins *et al* (2014), a Fonética Acústica Forense assume atualmente uma importância crescente, uma vez que os meios comunicativos estão cada vez mais sofisticados, fato que reitera a busca por parâmetros estatisticamente mais seguros no que diz respeito à identificação ou eliminação do falante.

A necessidade de parâmetros objetivos que permitam o reconhecimento do falante tem-se mostrado muitíssimo importante em termos judiciais. Porém, a metodologia habitualmente utilizada não tem se apresentado efetiva. A avaliação é feita de acordo com uma escala de base qualitativa. Por ter um caráter qualitativo, a escala levanta problemas e, por isso, parâmetros mais objetivos, que possibilitem a formulação de conclusões de base quantitativa (a partir de análise estatística), se fazem necessários (MARTINS *et al*, 2014).

Um dos parâmetros mais importantes em Fonética Forense é a frequência fundamental (f_0), considerada por muitos autores um parâmetro robusto, podendo ser extraído mesmo em gravações de menor qualidade (ROSE, 2002, *apud* GILLIER, 2011). Outro parâmetro é o padrão formântico das vogais.

Neste trabalho, no entanto, em lugar de vogais, consoantes serão o foco, e o parâmetro, por tratar-se de consoantes, será o VOT (Voice Onset Time). O segmento analisado será o da plosiva bilabial não-vozeada em posição de ataque. O VOT é considerado uma ferramenta essencial para o reconhecimento de voz, tomando por base suas características como instrumento eficaz na percepção e produção de plosivas em início de palavras (WHITESIDE; MARSHALL, 2000).

Assim, esta pesquisa tem por objetivo principal avaliar a importância do parâmetro VOT para comparação de voz. Busca-se também, verificar a existência de um padrão para reconhecimento de voz a partir da análise do VOT da plosiva /p/ em início de palavras para comparação de falantes. Tais constatações serão realizadas através da análise dos dados extraídos por meio do programa *Praat*, versão 5.4.02 liberada em 26 de novembro de 2014, de autoria dos holandeses Paul Boersma e David Weenink, ofertado de maneira gratuita na Internet. Através do programa *Praat* é possível medir o tempo de ataque de vozeamento (VOT) em milissegundos. O trabalho

também busca identificar a diferença intrafalante em voz normal e disfarçada, trabalhando também com comparação de vozes femininas e masculinas.

Com esses objetivos, primeiramente foi feito um levantamento teórico sobre Fonética Forense, disfarce, e sobre o objeto de estudo desta pesquisa, o VOT. Foram analisados alguns trabalhos que tratam dos temas previamente mencionados. Em sequência, será apresentada a metodologia utilizada na pesquisa, os resultados encontrados e a análise destes dados.

2.0 Fonética Forense

De acordo com Eriksson (2010), a Fonética Forense trata de linguagem e identidade, ainda que em uma área bem específica. Dentro da área são utilizados conhecimentos, teorias e métodos fonéticos com finalidades judiciais (GILLIER, 2011). O principal objeto de estudo da Fonética Forense é o reconhecimento de falantes, mas outras atividades, como autenticação de gravações ou a identificação de conteúdo de gravações, entre outras, também são realizadas (ROSE, 2002 *apud* GILLIER, 2011).

2.1 A Fonética Forense e a Identificação de Locutor

A Fonética Forense tem por objetivo buscar traços individuais na voz de um falante, em comparação com outro(s) falante(s). A perícia em Fonética Acústica Forense tem por base gravações realizadas principalmente através de escutas telefônicas (MARTINS *et al*, 2014). O reconhecimento de falantes pode ser definido como “qualquer processo de decisão que usa algumas características do sinal de fala para determinar se uma pessoa em particular é o falante de determinado enunciado” (ROSE, 2002 *apud* GILLIER, 2011).

De acordo com Nolan (1999) o reconhecimento de um falante pode ser definido como “qualquer atividade em que uma amostra de discurso é atribuída a uma pessoa com base em suas propriedades fonético-acústicas”. A partir desse conceito mais abrangente, o autor estabelece duas divisões para o reconhecimento de falantes: a verificação e a identificação. Para o autor, a verificação trata da comparação de uma amostra de fala de um indivíduo com outras amostras do mesmo sujeito, com a

finalidade de observar se existe correspondência entre as amostras. Nesse caso temos alguém que quer ser identificado. Já no caso da identificação tem-se um enunciado de um sujeito desconhecido que precisa ser identificado em uma população de possíveis amostras de referência (NOLAN, 1999). Segundo o autor, esse é o contexto mais comum da área forense.

Hollien (2002) também trabalha com tais termos e estabelece uma diferenciação entre o reconhecimento de falante, da mesma forma feita por Nolan: verificação e identificação de falante. O primeiro termo - reconhecimento - é o mais amplo, um conceito geral que abrange os outros dois. Para a verificação de falante, tem-se alguém que quer ser identificado, em uma ação cooperativa, como por exemplo, o acionamento de dispositivos eletrônicos por comando de voz. Já na identificação, o contexto é aquele mais comum da forense (um suspeito de um crime, por exemplo), uma amostra de fala que pode ter sofrido alguma distorção, tornando mais difícil a análise da fala. Segundo esses dois autores, então, Hollien (2002) e Nolan(1999), o termo apropriado para o reconhecimento de falante na área forense é 'identificação'.

No entanto, de acordo com Gomes e Carneiro (2014), o termo identificação não é aplicado no Brasil dentro do contexto forense, o termo mais comumente usado tem sido 'verificação de locutor'. Braid (2003, *apud* GOMES; CARNEIRO, 2014) afirma que "a verificação de locutor é a atividade parcial dentro da Fonética Forense capaz de determinar se duas falas foram produzidas por um mesmo falante". Mais recentemente, Valente (2012, *apud* GOMES; CARNEIRO, 2014) elegeu "comparação de locutor" como o termo adotado pela perícia criminal oficial brasileira. O autor argumenta que pesquisadores da área têm preferido evitar nomenclaturas com uma ideia de individualização. Gomes *et al* (2012) também discordam do uso do termo 'identificação' uma vez que "a natureza da fala humana e os atuais procedimentos de análise disponíveis não permitem individualizar um falante a partir de um amplo banco de dados, com a precisão que se obtém com o emprego da genética molecular e através de impressões digitais". Foulkes e French (2012, *apud* GOMES; CARNEIRO, 2014), no entanto, argumentam que a análise de características vocais, embora não

determine a identidade do falante, consegue fornecer importantes informações sobre o falante em questão.

Enfim, no Brasil, o termo primeiramente escolhido foi ‘verificação de locutor’, que vem sendo substituído por ‘comparação forense de locutor’ ou, simplesmente, ‘comparação de locutor’.

Partindo da terminologia para os métodos de análise, atualmente as perícias em Fonética Forense têm base na metodologia do American Board of Recorded Evidence (ABRE) e da International Association of Forensic Phonetics and Acoustics (IAFPA), com os resultados seguindo uma escala de base qualitativa (MARTINS *et al*, 2014).



Imagem 1: Níveis de confiança nas conclusões em Fonética Forense (ABRE *apud* MARTINS *et al*, 2014)

A classificação de acordo com a escala depende de julgamentos pessoais dos peritos, por isso pode receber uma valoração inferior àquela possível a partir de estudos quantitativos estatisticamente significativos (MARTINS *et al*, 2014). A escala feita dentro deste formato remete ao teste diagnóstico (TD) feito em estudos médicos para estabelecer a probabilidade de um indivíduo ter uma doença para a qual ele é positivo, porém, tal fato não pode ser tido como certeza. A razão de probabilidades é definida como “probabilidade do resultado de um dado teste entre pessoas com uma doença dividida pela probabilidade daquele teste em pessoas sem a doença” (CÂMARA, 2009).

Gold e French (2011, *apud* GOMES; CARNEIRO, 2014) realizaram pesquisa sobre os métodos de análise mais comuns em vários países e que a perícia no Brasil

adota “como metodologia de trabalho a análise auditiva combinada com a análise acústica”.

Um fator importante a ser apontado dentro da identificação de falante é a diferenciação entre o sexo do falante, algo relativamente acessível (GILLIER, 2011), uma vez que a separação entre o sexo feminino e o masculino é acentuada em termos de frequência fundamental (f_0), geralmente mais elevada nas mulheres (200Hz, em média (STEVENS, 2000 *apud* GILLIER, 2011)) em relação aos homens (120Hz em média).

Muitas vezes a metodologia de análise pode ser definida pela condição do material questionado. Uma das possibilidades é o uso de disfarce de voz, condição escolhida para análise neste trabalho.

2.2 Disfarce

Eriksson (2010) destaca que a existência de disfarce em uma amostra de voz é um dos fatores que mais dificulta a tarefa do perito forense. Porém, de acordo com o autor, a maior parte dos casos não faz uso de recursos sofisticados, o que complica menos a tarefa do perito.

Gillier (2011) define disfarce como “a ação deliberada de um falante que altera sua voz, discurso ou língua com o propósito de esconder sua identidade”. Ela diferencia o disfarce em deliberado versus não-deliberado e entre eletrônico versus não-eletrônico. O disfarce deliberado-eletrônico acontece quando é feito uso de aparelhos que promovem alteração da voz, o não deliberado-eletrônico corresponde às distorções provocadas pelos canais de transmissão, como defeitos em uma ligação telefônica, por exemplo. O disfarce não deliberado- não eletrônico está relacionado a fatores patológicos (já previamente mencionados) ou ao consumo de álcool ou drogas, enquanto que, o disfarce deliberado-não eletrônico diz respeito às alterações voluntárias como alterações no trato vocálico, produção de falsetes ou obstrução nas fossas nasais.

No caso dos disfarces não deliberados, vários fatores podem promover modificações no trato vocal, como por exemplo: estados patológicos específicos

(gripes, variações de estado psicológico, entre outros). Já dentro de disfarces deliberados temos: hipernasalidade, redução ou exagero na variação do *pitch* (GILLIER, 2011).

A presença de disfarce na voz representa uma grande dificuldade para o perito Forense, uma vez que pode distorcer a voz e produzir uma enorme variação em parâmetros como: frequência fundamental (f_0), frequência dos formantes, banda formântica e duração dos segmentos (GILLIER, 2011).

O estudioso alemão Hermann Joseph Kunzel (2000) detalhou que, dos casos atendidos pela sessão de identificação de falantes do *Bundeskriminalamt* (Polícia Federal Alemã) nas últimas duas décadas, entre 20 e 15% ao ano faziam uso de algum tipo de disfarce. As formas mais comuns de disfarce descritas por ele são: *falsetto*, voz vibrante produzida pela laringe com pouca soltura de ar, voz sussurrada (substituindo a respiração por fonação), fingir um sotaque estrangeiro e o bloqueio das narinas, de forma a anasalar a voz. Dentro de sua pesquisa, Kunzel concluiu que quando os disfarces diminuem a frequência fundamental ou anasalam a voz, a frequência fundamental original do falante pode ser apurada, com uma certa margem de erro. Quando se trata de aumento de f_0 , os dados não admitem muita interferência da frequência original (KUNZEL, 2000).

A pesquisa de Kunzel teve foco na frequência fundamental, porém, outros parâmetros podem ser alterados com o uso de disfarce na fala, como, por exemplo, o *Voice Onset Time*.

3.0 *Voice Onset Time* (VOT)

O Tempo de Ataque de Vozeamento (VOT) é classificado como um parâmetro para identificar auditivamente as consoantes plosivas surdas e sonoras, (GREGIO; CAMARGO, 2005). No Português Brasileiro (PB), o VOT é considerado o parâmetro mais robusto em relação ao contraste de vozeamento (SMITHS, 2004 *apud* BARZAGHI et al, 2007).

3.1 Consoantes Oclusivas e o VOT

Dentro da acústica temos as características das vogais e das consoantes. Como já mencionado, no presente trabalho, o foco será nas consoantes. As características acústicas das consoantes são mais complicadas que as das vogais (KENT; READ, 2002). As vogais possuem, essencialmente, as mesmas características acústicas, como, por exemplo, duração ou padrão dos formantes. Já as consoantes se diferenciam de maneira significativa no que diz respeito às propriedades acústicas. Aqui trataremos das consoantes oclusivas.

O português brasileiro apresenta três consoantes oclusivas não vozeadas (/p/, /t/ e /k/) e três vozeadas (/b/, /d/ e /g/) (CELESTE; TEIXEIRA, 2009). O vozeamento, de acordo com Cagliari (2007, *apud* BANDEIRA; ZIMMER, 2011) é a dimensão fonética primária usada para diferenciar os pares mínimos, assim, as consoantes surdas /p, t, k/ possuem em /b, d, g/ suas contrapartes sonoras. Essas consoantes possuem três correlatos acústicos: tempo de silêncio (oclusão total do trato vocal pelos articuladores), barra de explosão (liberação da pressão pelos articuladores) e *voice onset time* (VOT) (HERLICH; ACKERMAN, 2007; LOUSADA, 2006; ROCCA, 2003 *apud* CELESTE; TEIXEIRA, 2009).

De acordo com Prestes (2012), as oclusivas surdas e sonoras podem ser diferenciadas por diferentes aspectos, como por exemplo: barra de sonoridade (identificada na porção inferior do espectrograma durante a fase de oclusão) (KENT; READ, 1992, *apud* PRESTES, 2012), duração da consoante (a consoante tende a ser maior nas oclusivas) (VELOSO, 1995 *apud* PRESTES, 2012), e a amplitude do *burst* que tende a ser levemente mais intensa para os seguimentos surdos (LADEFAGED; MADDISON, 1997 *apud* PRESTES, 2012).

Além de todos os parâmetros mencionados, é possível usar um outro parâmetro para caracterizar as consoantes oclusivas: o VOT (*Voice Onset Time*). Prestes (2012) afirma que a primeira descrição deste aspecto encontra-se em um trabalho de Lisker e Abramson (1964) “no qual os autores buscaram identificar um parâmetro único para distinguir oclusivas surdas e sonoras”. Os autores (LISKER; ABRAMSON, 1964 *apud*

BANDEIRA; ZIMMER, 2011) consideram o VOT “como o meio mais efetivo de percepção e produção de palavras com plosivas iniciais”.

Dentro das consoantes plosivas, o VOT é definido como o intervalo entre a soltura da pausa (*stop burst*) e o aparecimento de uma modulação (vozeamento) para o próximo som (KENT; READ, 2002). As oclusivas não-vozeadas da Língua Inglesa são produzidas com VOT longo e aspirado. Já no português esta medida é mais curta e não há aspiração (CELESTE; TEIXEIRA, 2009). O tempo medido é, geralmente, expresso em milissegundos (ms), sendo decisivo na percepção acurada das plosivas não-vozeadas /p, t, k/ e das vozeadas /b, d, g/ (BANDEIRA; ZIMMER, 2011).

O VOT pode assumir valores positivos, negativos ou iguais a zero. Ao medir-se o VOT, o momento da soltura é condicionado como “ponto zero” (PRESTES, 2012). Quando o vozeamento acontece antes da soltura, o VOT possui valor negativo, típico de uma consoante sonora. Caso o vozeamento seja posterior à soltura, o VOT é positivo, e a consoante é uma oclusiva surda. Em alguns casos, o vozeamento coincide com o momento da soltura, e diz-se que o VOT é igual a zero (PRESTES, 2012). Esses três domínios foram assim descritos por Rocca (2003):

- oclusivas sonoras, caracterizadas por barra de sonoridade em que o vozeamento é anterior à liberação da oclusão (VOT negativo de -125 a -75msec.);
- oclusivas surdas sem aspiração, caracterizadas por intervalo curto ou simultâneo, em que o vozeamento é simultâneo ou dá-se logo após a liberação da oclusão (VOT zero de 0 a +25msec.);
- oclusivas surdas com aspiração, caracterizadas por intervalo longo, em que o espaço de tempo entre a liberação da oclusão e o início da vogal é superior a 60msec.(VOT positivo de +60 a +100msec.) (Rocca, 2003).

Para o Inglês, os valores típicos de VOT ficam em torno de 55ms para /p/, 70ms para /t/ e 80ms para /k/ (KENT; READ, 1992 *apud* BANDEIRA; ZIMMER, 2011).

Estudos têm mostrado que a diferença de idade pode interferir tanto na percepção quanto na produção do VOT (CELESTE; TEIXEIRA, 2009). Além da idade, o gênero do falante também interfere nas medidas de VOT. Uma pesquisa feita por Whiteside e Marshall (2001, *apud* CELESTE; TEIXEIRA, 2009) com adolescentes entre 9 e 11, falantes de inglês britânico, mostrou que os indivíduos do sexo feminino apresentavam medidas de VOT próximas dos valores dos adultos, o que não acontecia

com os indivíduos do sexo masculino. Celeste e Teixeira (2009) apontam que o estudo das oclusivas em Língua Portuguesa, na variante do português brasileiro, é escasso, principalmente no que diz respeito à variação de grupos.

3.2 O VOT no Português Brasileiro (PB)

Rocca (2003) realizou uma pesquisa de comparação entre falantes bilíngues de português brasileiro e inglês, com foco no VOT de consoantes plosivas surdas, com objetivo de detectar se as diferenças sutis na realização fonética de um contraste fonológico que estão presentes na produção dos monolíngues são reproduzidas na fala desses bilíngues. Tal pesquisa é uma importante fonte de informação relacionada ao VOT no PB (português brasileiro). A pesquisadora cita Sancier e Fowler (1997), que, ao estudarem as oclusivas surdas em um falante nativo de PB passando um período nos Estados Unidos, chegaram à seguinte conclusão: apesar de as medidas de VOT do sujeito serem sempre menores no português do que as apresentadas no inglês, os VOTs das oclusivas apresentavam-se menores em ambas as línguas quando o sujeito passava alguns meses no Brasil.

Bandeira e Zimmer (2011) realizaram uma pesquisa comparativa entre crianças monolíngues e multilíngues, visando investigar as diferenças relativas aos padrões de VOT em plosivas surdas. Os participantes foram divididos em dois grupos, um com 20 crianças monolíngues, e o outro com 20 crianças bilíngues. O grupo “monolíngue” falava PB e o multilíngue falava pomerano (língua original da região da Pomerânia, extinta na Europa, porém ainda falada no Brasil), alemão e PB. Ambos os grupos estavam aprendendo Inglês. O estudo concluiu que os bilíngues produziram valores médios de aspiração muito elevados, inclusive para língua portuguesa (/p/ 50ms, /t/ 59ms e /k/ 67ms). O fato mostra que mesmo o PB não apresentando aspiração nas oclusivas surdas em posição inicial, ela é realizada pelos falantes bilíngues.

Outro trabalho usado como referência para o estudo de VOT dentro do PB na presente pesquisa é o trabalho de Reis (2009). A estudiosa menciona a pesquisa de Istre (1985), que definiu resultados médios para os valores de VOT no português: em torno de 38 ms para o /k/, 18 ms para o /t/, e 12 ms para o /p/. Vale ressaltar, que, de

acordo com Reis (2009), os valores de VOT não são entidades isoladas em nenhum sistema fonológico: “estes valores recebem influência de fatores no âmbito segmental e suprasegmental, como a prosódia, contextos fonéticos, a tonicidade de palavra e de frase, a idade do falante e a velocidade de fala” (KLEIN, 1999; COHEN, 2004 *apud* REIS, 2009). A autora reitera, ainda, que o fator de maior influência nas medidas de VOT é a velocidade de fala: quanto mais rápida a fala, menor tende a ser o valor de VOT (SCHMIDT; FLEGE, 1996 *apud* REIS, 2009). Outro ponto a ser observado em relação à velocidade de fala é que em uma velocidade rápida de fala os valores de VOT curto e longo tendem a se sobrepor, fazendo com que a identificação entre segmentos vozeados e não-vozeados se torne mais problemática (COHEN, 2004 *apud* REIS, 2009).

As diferentes pesquisas que apresentaram valores médios para o VOT de plosivas não-vozeadas no PB foram mostradas por Prestes (2013):

Português brasileiro			
	[p]	[t]	[k]
Istre (1985)	11,95	18,48	38,53
Major (1992)	11	15	35
Klein (1999)	15,58	16,69	36,36
Reis e Nobre-Oliveira (2007)	17,27	23,55	46,55
Alves e Dias (2010)	37,67/31,38	36,52	47,83/52,95
França (2011)	19,56	21,66	47,20
Schwartzhaupt (2012)	15,13	17,87	58,05

Tabela 4: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas do português brasileiro em diferentes estudos

Imagem 2 - Valores médios de VOT no Português Brasileiro (PRESTES, 2013)

Na tabela temos os valores da plosiva /p/ variando, em sua maioria, entre 11 e 15 milissegundos. A autora menciona que na pesquisa de Alves e Dias (2010) os valores ficaram mais altos devido à posição das palavras-alvo (posição medial ou final na sentença), confirmando a influência do contexto na medida de VOT.

4.0 Metodologia

Os dados a serem utilizados nesta pesquisa foram coletados a partir do projeto do Grupo de Pesquisa Estudos dos Sons da Fala². Os procedimentos necessários foram: escolha dos sujeitos, manuseio dos aparelhos de gravação e etiquetagem dos dados no Software. A metodologia escolhida teve foco comparativo, objetivando analisar as diferenças do VOT entre os grupos escolhidos e entre voz normal e disfarçada.

4.1 Informantes

A pesquisa foi composta por 5 grupos de 10 integrantes cada, separados por gênero e idade, denominados: grupo Controle (25-55 anos), Mulheres (30-55 anos), Garotas (19-25 anos), Homens (30-55 anos) e Garotos (19-25 anos). O presente trabalho irá focar nos grupos Mulheres, Homens e Controle, totalizando 30 indivíduos.. Como o trabalho objetivou analisar a diferença de gênero, e não de idade, tais grupos foram selecionados, enquanto os dados dos grupos Garotas e Garotos não foram considerados. O Grupo Controle, apesar de ter uma informante na faixa etária do grupo Garotas, contava com as demais informantes na faixa etária dos grupos Mulheres e Homens.

4.2 Coleta de Dados

Após aceitarem participar do projeto, foi verificada com os participantes a disponibilidade de horário de cada um, bem como do Estúdio de Coleta e Análise de Dados (ECAD) da UTFPR para que as gravações fossem realizadas. A coleta de dados foi realizada em laboratório com isolamento acústico com o seguinte equipamento: Computador Pentium Dual Core 5.300 2.60 GHz, 1.99 Gb RAM Processador XP 2002 Service Pack 3, placa de som externa M-Audio Fast Track Pro 4x4, Microfone AKG C

² O grupo foi criado em 2010 e possui diversos trabalhos publicados em periódicos e anais de eventos nacionais e internacionais (GOMES, M. L.; RICHERT, L. C.; MALAKOSKI, J., 2013; GUILHERME, 2013; KREMER, 2013; GUILHERME, 2013; GOMES, M. L.C.; CARNEIRO, D. O., 2014; KREMER, R.; GOMES, M. L. C., 2014). Trabalhos de Iniciação Científica (KREMER, 2014; GUILHERME, 2014) também foram desenvolvidos, e Trabalhos de Conclusão de Curso estão ainda em andamento (dois TCC2 a serem defendidos neste final de semestre, 2/2014, e um TCC1 sendo desenvolvido para o primeiro semestre 2015).

3000 B. O texto elaborado simulava o pedido de resgate em uma situação de sequestro. O programa utilizado para as gravações foi o *Audacity*, um software gratuito de edição de áudio desenvolvido por voluntários, versão 2.0.3.

Em cada coleta, o registro de áudio foi capturado com as seguintes configurações: formato *wave monocal*, frequência de amostragem de 44,1 KHz e 16 bits. Os arquivos foram nominados de acordo com o grupo e o número sequencial do informante, para facilitar os procedimentos seguintes.

Os participantes gravaram em laboratório o seguinte texto, composto por 75 palavras:

“Alô?! Quero falar com a dona Teca. Dona Teca, aqui fala o Capeta, estamos com o pato do teu marido na toca e vamos matar ele, picar e tacar dentro de uma garrafa de coca. Quer salvar ele, então não me deixa puto. Quero um milhão, coloca tudo num pacote perto da pipa do cateto e se pica. Não faz caca, se não vai levar toco. Não chama a polícia, se não te cutuco e espeto teu coco.”
(Texto elaborado pelo Grupo de Pesquisa de Fonética Forense da UTFPR)

O texto foi criado para simular uma situação de sequestro. Foi pedido aos informantes que lessem o texto quatro vezes: duas com voz normal e duas com algum tipo de disfarce (da escolha de cada um). O número de gravações escolhido foi esse por motivos de segurança, caso ocorresse algum problema com as gravações. Foi indicado que tentassem usar algum disfarce que dificultasse ao máximo o reconhecimento por parte do ouvinte.

Esse banco de dados foi coletado para um projeto de análise de vogais, mas tem servido para vários outros projetos. Neste trabalho, em particular, que tem como foco a análise de VOT, mas especificamente na consoante plosiva bilabial não vozeada, as palavras escolhidas para análise foram três: pato, puto, pipa. Tais palavras foram escolhidas pois possuem o /p/ precedido e seguido de vogais distintas, dentro do contexto do texto. Assim, o resultado foi 360 palavras/sons a serem analisados (30 indivíduos x 4 repetições x 3 palavras).

4.3 Estratégias de Disfarce

Um grupo de seis pessoas (juízes) analisou as vozes, com a orientação de uma fonoaudióloga, para definir algumas estratégias de disfarce. Para tanto foi feito uso do

BP-VPAS (CAMARGO; MADUREIRA, 2008)³. De acordo com a professora Maria Lúcia Gomes, coordenadora do Grupo de Pesquisa, o protocolo não foi utilizado em toda sua extensão e com todos os seus detalhes. Também não foram utilizados os níveis de graduação, uma vez que o objetivo era apenas detectar o ajuste escolhido pelo participante para disfarçar sua voz. As vozes foram ouvidas várias vezes, com os juízes preenchendo seus formulários individualmente. Após o preenchimento dos formulários, as respostas eram comparadas entre os juízes e, voltando a ouvir quantas vezes fosse necessário, uma resposta final era negociada.

Abaixo, está apresentada uma tabela contendo os informantes e o tipo de disfarce estabelecido pelos juízes, a partir do protocolo. O nome do disfarce está aqui apresentado seguindo o modelo proposto por Laver (2007 *apud* CAMARGO; MADUREIRA, 2008).

³ Essa fase foi feita pelo Grupo de Pesquisa e está servindo de base para as pesquisas individuais dos membros do grupo, como este trabalho que aqui se apresenta.

Tabela 4 – Disfarces estabelecidos a partir do BP – VPAS

Informante	Disfarce	Informante	Disface	Informante	Disfarce
GC1	Backed tongue body Tense vocal tract Foreign accent	GM1	High pitch	GH1	Fronted tongue body Tense larynx Harsh voice
GC2	Harsh voice	GM2	Nasal voice	GH2	Open jaw High pitch High pitch
GC3	Falsetto	GM3	Tense larynx Open jaw	GH3	Soothing voice
GC4	Backed and lowered tongue body	GM4	Lip protrusion Low pitch	GH4	Open jaw Accent
GC5	High pitch	GM5	Raised and tense larynx	GH5	High pitch Nasal
GC6	Open jaw	GM6	Low pitch	GH6	Open jaw High pitch Nasal
GC7	Lip protrusion Low pitch	GM7	Pinch noise	GH7	High pitch Nasal Accent
GC8	Tense and raised larynx Nasal voice	GM8	Low pitch	GH8	Harsh voice
GC9	High pitch (mean or variability?) Nasal voice	GM9	High pitch Regional accent	GH9	Open jaw High pitch Nasal
GC10	Backed tongue body High pitch (mean or variability?)	GM10	Backed tongue body Foreign accent	GH10	Tense larynx Harsh voice

4.4 Procedimentos de medidas

A partir das informações obtidas com o grupo de pesquisa, deu-se início à análise do VOT a partir do software *Praat*, versão 5.4.02, de 2014.

Ao abrir a janela de edição do software é possível observar o espectrograma e a forma da onda referentes à produção de determinada palavra. O espectrograma é uma representação espectro-temporal do som (SOUZA, 2012). A direção horizontal representa o tempo, enquanto que a vertical representa a frequência. Utilizando o *Praat* pode-se identificar visualmente e auditivamente a duração da aspiração, um período que se inicia no momento seguinte à aspiração do ar e termina no início da barra de sonoridade.

Para medir o VOT usando o software basta estabelecer dois pontos no espectrograma: a soltura da pausa e o início do vozeamento. Como citado por Styler (2011), esta é a sequência de passos para medição:

1. Encontrar a soltura da pausa
2. Encontrar o início do vozeamento
3. Selecionar o segmento entre os dois pontos
4. Identificar a duração do espaço selecionado (em segundos), delimitada na barra de duração localizada na parte inferior da janela de Edição.
5. Utilizar os comandos *Editor* → *Query* → *Get selection length* e ler a seleção na janela de informação
6. Se o início do vozeamento acontecer antes da soltura da pausa, o VOT será negativo, se não, será positivo.

Prestes (2013), ao analisar o VOT com o software *Praat*, menciona a possibilidade do cálculo da duração relativa do VOT, tomando como base a palavra-alvo, porém, a maior parte dos trabalhos que embasam esta tese usou os valores absolutos para medida de VOT. Assim, neste trabalho, o VOT será medido a partir da duração absoluta (ver imagem 2). O padrão seguido para a extração das medidas foi baseado no parâmetro seguido por Prestes (2013) em relação às oclusivas surdas, coletando a duração em milissegundos entre o *burst* e o primeiro pulso regular da vogal seguinte.

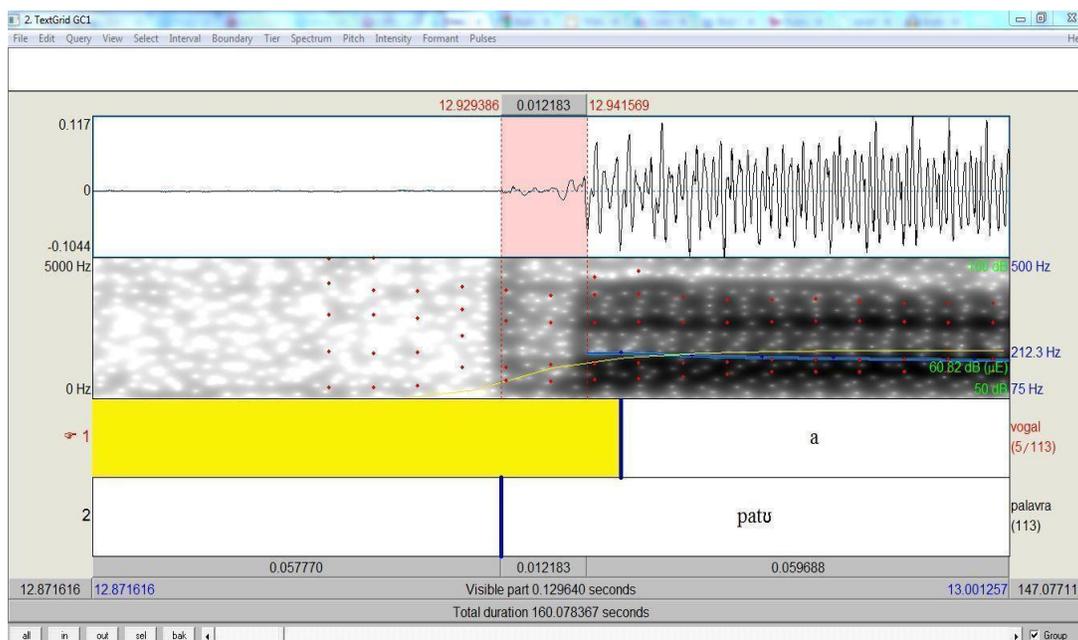


Imagem 3: medida de VOT da palavra “pato”. Fonte: Software Praat

O valor do VOT no Software é medido em segundos. Para facilitar o estudo e a análise, as medidas foram apresentadas nas tabelas em “msec”. As tabelas estão numeradas e intituladas.

4.5 Resultados

Foram formadas seis tabelas com os dados retirados do software *Praat*. As três primeiras (Anexo 1) com os dados das quatro vezes em que os informantes leram o texto. As colunas foram divididas entre as palavras-foco: PA para “pato”, PU para “puto” e PI para “pipa”. As letras “n” e “d” foram usadas para definir “voz normal” e “voz disfarçada” respectivamente.

As três tabelas seguintes trazem as médias aritméticas das medidas. Foram feitas as médias entre o valor de VOT da palavra na primeira e na segunda leitura com voz normal, e, depois, com voz disfarçada.

Tabela 1 - Médias de VOT do Grupo Controle

INFORMANTE	Average PAn	Average PUn	Average PIn	Average PAd	Average PUd	Average PId
GC1	13.71	17.36	17.44	73.62	63.05	53.13
GC2	12.10	23.8	24.03	11.61	15.41	15.1
GC3	18.93	24.14	29.65	16.14	39.79	20.3
GC4	21.69	37.28	16.76	14.67	22.41	21.81
GC5	16.42	18.62	15.55	27.96	25.74	19.72
GC6	39.05	64.69	41.85	85.39	41.69	24.64
GC7	14.12	25.95	15.12	19.37	22.01	17.75
GC8	14.75	30.14	26.72	20.11	30.09	16.03
GC9	16.68	27.07	26.57	17.53	21.33	18.99
GC10	13.05	17.96	16.22	11.18	17.09	12.36

Tabela 2 - Médias de VOT do Grupo Mulheres

INFORMANTE	Average PAn	Average PUn	Average PIn	Average PAd	Average PUd	Average PId
GM1	22.22	37.5	23.53	25.09	31	29.11
GM2	21.45	32.31	30.82	17.28	18.08	19.72
GM3	27.87	46.26	21.2	19.27	22.98	14.3
GM4	18.07	24.04	24.88	12.52	14.25	19.31
GM5	11.78	18.51	29.24	15.07	23.44	23.64
GM6	17.85	23.02	13.85	17.35	22.77	20.17
GM7	20.19	23.64	15.28	13.97	27.9	13.71
GM8	15.55	29.09	28.95	14	25.02	18.45
GM9	25.76	32.56	19.84	18.91	25.36	16.06
GM10	18.2	30.15	25.06	45.79	41.76	32.31

Tabela 3 - Médias de VOT do Grupo Homens

INFORMANTE	Average PAn	Average PUn	Average PIn	Average PAd	Average PUd	Average PId
GH1	15.42	23.11	19.59	17.36	13.81	13.02
GH2	18.6	25.75	29.82	13.13	26.56	8.13
GH3	21.22	14.9	18.57	15.17	28.51	26.66
GH4	13.67	12.73	22.49	10.25	24.51	11.14
GH5	9.97	12.31	7.98	9.97	20.78	8.84
GH6	12.31	18.46	20.08	8.28	9.51	8.17
GH7	16.99	28.95	9.34	14.37	16.48	16.5
GH8	12.85	20.26	22.59	16.08	17.01	16.21
GH9	26.03	22.68	23.75	19.39	16.78	10.13
GH10	19.3	27.87	24.49	13.54	24.44	19.17

Para a análise dos dados foram usadas as medidas encontradas nas tabelas 1, 2 e 3.

5.0 Considerações sobre os dados

A partir dos dados levantados e apresentados, no Anexo 1 em sua totalidade, e nas tabelas acima em suas médias, vários pontos merecem consideração e podem ser aproveitados para a perícia em fonética forense. Nesta seção, vou apresentar essas considerações, procurando relacionar os fatos verificados com a literatura consultada.

Paulina Rocca (2003) apontou que, diferentemente do Inglês, no Português, as oclusivas /p, t, k/ são produzidas com um intervalo de VOT curto e com ausência de aspiração. Já no Inglês, as mesmas oclusivas são produzidas com um intervalo de VOT longo, seguido de aspiração, quando em sílaba tônica e em posição inicial na palavra, assim, o tempo de ataque do vozeamento é superior a 25 msec.

Pode-se perceber, na presente pesquisa, medidas de VOT com valores próximos aos da Língua Inglesa, tanto em voz normal quanto em voz disfarçada, se considerarmos os grupos Mulheres e Controle. É importante ressaltar o fato de muitas informantes serem professoras de Inglês (informantes 1, 8, 9 e 10 no GC e informantes 2, 6, 7, 8, 9 e 10 no GM).

No Grupo Controle a informante GC1 usou o sotaque americano (*foreign accent*) como estratégia de disfarce. Tal disfarce teve reflexos no valor de VOT para todas as consoantes. O valor em voz disfarçada ficou elevado, tendo uma medida próxima aos valores de Língua Inglesa encontrados por Rocca (2003). O mesmo aconteceu para a informante 10 do Grupo Mulheres.

O uso do VOT com disfarce simulando sotaque americano se mostrou bastante interessante, uma vez que o parâmetro possui características específicas para Língua Portuguesa e Língua Inglesa.

Outra questão observada foi o posicionamento da palavra na sentença. As palavras estão em posição medial (“pato” e “pipa”) e final (“puto”) na sentença, o que poderia acarretar valores maiores de VOT quando comparados com os valores base de pesquisas anteriores. De acordo com Prestes (2013), as vogais altas ([i] e [u]) possuem, em média, um VOT 7% mais longo que vogais baixas ([a]). As plosivas medidas também estão em posição tônica nas palavras analisadas, o que acarreta um VOT geralmente mais longo (LISKER; ABRAMSON, 1967; GIMSON, 1980 *apud* PRESTES, 2013).

Nas análises, buscou-se também, relacionar a estratégia de disfarce com os valores encontrados. Os homens mostraram, em média, um parâmetro de VOT mais baixo em relação aos outros grupos, tanto em voz normal quanto em voz disfarçada. Os disfarces mais comuns neste grupo foram a elevação do *pitch* (*high pitch*) e a voz mais áspera (*harsh voice*). O disfarce com voz áspera (usado pelos informantes 1, 7, 8 e 10 do Grupo Homens) mostrou diminuição nos valores de VOT ao compararmos com a voz normal do participante. A estratégia de voz áspera também mostrou diminuição de valor de VOT com a única informante do Grupo Controle (GC2) a usar esta estratégia. Tal diminuição pode ser devido ao aumento no movimento da língua, como sugerido por Cho e Ladefoged (1999).

A nasalização (*nasal voice*) mostrou uma diminuição dos valores de VOT, na relação voz normal para voz disfarçada. Tal mudança ocorreu tanto para as vozes femininas (GM), quanto para as vozes masculinas (GH). Os informantes que usaram

esse tipo de disfarce foram os informantes 2 e 7 no Grupo Mulheres, e 5, 6 e 9 no Grupo Homens.

Em relação ao abaixamento do *pitch* como estratégia de disfarce, o ocorrido foi a diminuição dos valores de VOT. Tal resultado já era esperado. Em um estudo, Narayan e Bowden (2013) afirmam que o VOT e a frequência fundamental (f_0) estão ligados, na forma da perturbação do *pitch*, a um fenômeno de depressão de f_0 que segue plosivas vozeadas (HAGGARD et al, 1970 *apud* NARAYAN; BOWDEN, 2013). Alguns modelos sugerem que o aumento de f_0 vem acompanhado de um VOT mais longo. Os autores ressaltam que, embora existam muitos estudos sobre a perturbação de *pitch*, o efeito da f_0 no VOT é muito pouco entendido.

Já em relação ao efeito no VOT dos níveis de f_0 em vozes masculinas, McCrea e Morris (2005, *apud* NARAYAN; BOWDEN, 2013) encontraram que consoantes plosivas desvozeadas com um f_0 alto produziram um VOT mais curto, quando comparadas com produções de f_0 mais baixo ou mediano.

Para facilitar a observação do leitor, no Anexo II podem ser vistos alguns gráficos e a seguir, vê-se um resumo das principais informações que tiramos dos resultados da presente pesquisa:

No Grupo Controle, alguns pontos foram observados.

- O valor mais alto de VOT foi encontrado com a palavra “pato” em voz disfarçada do informante 6 do Grupo Controle, disfarce com mandíbula aberta (*open jaw*)
- Outro fato observado foi o aumento do valor de VOT na palavra “pato” para a maior parte dos informantes do Grupo Controle, a exceção dos informantes 2, 3 e 4, que usaram voz áspera, falsete e língua recuada (*backed and lowered tongue body*) como disfarce, respectivamente.
- Já palavra “puto” mostrou alteração entre voz normal e disfarçada no GC, aumentando seu valor nos informantes 1, 3 e 5 (disfarce com aumento de *pitch*) e diminuindo para o restante. O valor de VOT na palavra “pipa” aumentou para os informantes 1, 4, 5 e 7.

Dentro do Grupo Mulheres, os fatos observados por meio do gráfico foram:

- aumento do valor de VOT na palavra “pato” para os informantes 1, 5 e 10, este que usou sotaque americano como disfarce, mostrando o mesmo padrão que o informante GC1
- geral diminuição do VOT da palavra “puto”, a exceção dos informantes 5 e 7
- VOT maior em GM1, 6 e 10 para a palavra “pipa”. O uso do disfarce do informante 10 mais uma vez mostrou padrão similar ao do informante GC1. O disfarce usado pelo informante 6 - diminuição do *pitch* - também foi usado por GC7, também com aumento do VOT.

No Grupo Homens o gráfico mostrou:

- VOT com valor menor em “pato”, menos para GH3, quando o valor se manteve o mesmo. Os disfarces concomitantes são a voz mais áspera, e a mandíbula aberta
- em “puto” houve uma predominante diminuição no VOT, fato que não ocorreu em 2, 3, 4 e 5. Três informantes (GH2, GH3 e GH5) aumentaram o *pitch* como forma de disfarce, de maneira semelhante ao GC5. A nasalização também se repetiu em GH5 e GM7
- para a palavra “pipa” foi observado aumento de VOT em 3, 5 e 7. Todos os informantes usaram *high pitch* como disfarce.

Em relação à comparação de voz feminina e masculina, vale ressaltar que os valores do VOT para o Grupo Mulheres foram superiores. O GH, mesmo quando usou o disfarce de elevação do *pitch*, não chegou a ter medidas com valores similares.

A comparação em voz normal e voz disfarçada a partir do VOT se mostrou válida na medida em que existe diferença entre os valores médios de um mesmo falante entre os dois tipos de voz, na maior parte dos participantes. Para pesquisas posteriores vale ressaltar que seria um tanto mais eficaz o estabelecimento de um tipo específico de disfarce, visando uma análise de dados mais objetiva.

6.0 Conclusão

Esta pesquisa teve como principal objetivo avaliar a importância do parâmetro VOT para comparação de voz. O trabalho se justificou tendo em vista a escassa produção de trabalhos com base no VOT dentro da área de Fonética Forense, como ressalta Martins *et al* (2014): "a necessidade de parâmetros objetivos que permitam caracterizar e identificar o falante (duas das tarefas mais comuns da Fonética Forense) tem-se feito sentir em sede judicial".

O VOT como parâmetro para comparação de voz se mostrou bastante válido, pois diversas diferenças foram observadas, principalmente entre voz normal e voz disfarçada, área que carece de estudos.

Há muito que ser pesquisado na área, e trabalhos posteriores podem se dedicar a outras oclusivas, fazendo uso do mesmo texto, ou até produzir um banco de dados com o estabelecimento de um padrão no que diz respeito a estratégias de disfarce. Os dados encontrados na pesquisa podem ser usados para investigar posteriormente os disfarces mencionados neste trabalho.

Estudar Fonética Forense com o Português Brasileiro ainda representa um desafio a ser superado, mas, com as pesquisas desenvolvidas pelo Grupo de Estudos da UTFPR, percebe-se que o campo está cada vez menos desconhecido.

Referências

- BRAID, A. C. M. **Fonética Forense**. 2.ed. Campinas, SP: Millenium, 2003
- CÂMARA, Fernando Portela. **Psiquiatria e Estatística VI: Fundamentos dos Testes Diagnósticos - Parte I**. Psychiatry on-line Brazil. Vol. 14, n. 5. Rio de Janeiro, 2009. Disponível em <<http://www.polbr.med.br/ano09/cpc0509.php>>. Acesso em novembro de 2014.
- BARZAGHI L. *et al.* **Deficiência de audição e contraste de vozeamento em oclusivas do Português Brasileiro: análise acústica e perceptiva**. *Disturb Comun* vol. 19, n. 3. São Paulo, 2007.
- BOERSMA, P; WEENINK, D. **Praat: doing phonetics by computer**. Versão 5.4.02. Amsterdam: University of Amsterdam, 2014.
- CELESTE, L.; TEIXEIRA, E. **Efeito do sexo e idade na produção do VOT**. *Revista de Letras da Universidade Católica de Brasília*, vol.2 n.1. Brasília, 2009.
- CHO, Taehong; LADEFOGED, Peter. **Variation and Universals in VOT: Evidences in 18 Languages**. *Journal of Phonetics*. Los Angeles, 1999.
- DAS, Sharmistha; HANSEN, John. **Detection of Voice Onset Time (VOT) for Unvoiced Stops (/p/, /t/, /k/) Using the Teager Energy Operator (TEO) for Automatic Detection of Accented English**. *Proceedings of the 6th Nordic Signal Processing Symposium*. Espoo, 2004.
- ERIKSSON, Anders. **The Disguised Voice: Imitating Accents or Speech Styles and Impersonating Individuals**. In: LLAMAS, C.; WATT, D. *Language and Identities*. Edinburgh: Edinburgh University Press, 2010.
- GILLIER, Raïssa. **O Disfarce da Voz em Fonética Forense**. *Dissertação de Mestrado em Linguística*. Lisboa: Universidade de Lisboa, 2011.
- GOMES, M. L. C.; CARNEIRO, D. O., **A fonética forense no Brasil: cenários e atores**. *Language and Law/Linguagem e Direito*, Vol. 1(1), 2014, p. 22-36.
- GOMES, M. L.; RICHERT, L. C. ; MALAKOSKI, J. **Identificação de Locutor na Área Forense: A Importância da Pesquisa Interdisciplinar**. *Anais do X Encontro do*

CELSUL. UNIOESTE, Cascavel, 2012. Disponível em <http://celsul.org.br/Encontros/10_index.htm#M>. Acesso em abril de 2013.

GREGIO, F; CAMARGO, Z. **Dados de tempo de início do vozeamento (VOT) na avaliação do sinal vocal de indivíduos com paralisia unilateral de prega vocal.** Distúrbios da Comunicação vol. 17, n.3. São Paulo, 2005.

GUILHERME, M. L. F. **A variação da intensidade e da qualidade de vocal como estratégia de disfarce de voz.** Anais do XXI Seminário do Centro de Estudos Linguísticos e Literários do Paraná, 2013, p. 348 a 357.

_____. **Intensidade e qualidade vocal como estratégias de disfarce de voz.** XIX Seminário de Iniciação Científica e Tecnológica da UTFPR, 2014.

HOLLIEN, H. **Forensic Voice Identification.** San Diego: Academic Press, 2002.

KENT, Raymond D.; READ, Charles. **Acoustic Analyses of Speech.** Albany: Thomson Learning, 2002.

KREMER, R. L. **A eficiência do disfarce na voz: uma análise da frequência fundamental.** Anais do XXI Seminário do Centro de Estudos Linguísticos do Paraná. Paranaguá, 2013.

_____. **A eficiência do disfarce em vozes femininas: uma análise da frequência fundamental.** XIX Seminário de Iniciação Científica e Tecnológica da UTFPR, 2014.

KREMER, R. L.; GOMES, M. L. C.. **A eficiência do disfarce em vozes femininas: uma análise da frequência fundamental.** *ReVEL*, vol. 12, n. 23, 2014. Disponível em <www.revel.inf.br>.

KUNZEL, Hermann J. **Effects of Voice Disguise on Speaking Fundamental Frequency.** Forensic Linguistics 7. Birmingham: University of Birmingham Press, 2000.

MARTINS, F. *et al* **A Fonética Forense na produção de prova do ordenamento jurídico do Português: o parâmetro do pré-vozeamento.** *ReVel*, vol.12, n.23. Lisboa, 2014.

NARAYAN, Chandan; BOWDEN, Mary. **Pitch affects voice onset time (VOT): A cross-linguistic study.** 21st International Congress on Acoustics. Montreal, 2013.

NOLAN, F. Speaker Recognition and Forensic Phonetics. In: LAVER, J.; HARDCASTLE, W.J. **The Handbook of Phonetic Sciences**. Oxford: Blackwell Publishers Ltd, 1999.

PRESTES, Suzana P. **Produção de Consoantes Oclusivas Iniciais do Inglês por Falantes Nativos de PB**. Dissertação de Mestrado em Estudos Linguísticos do Curso de Pós-Graduação em Letras. Curitiba: Universidade Federal do Paraná, 2013.

REIS, Mara. **Efeitos de Treinamento Perceptual na Percepção e Produção de Plosivas não-vozeadas do Inglês**. 3º Colóquio de Estudos Linguísticos e Literários. Maringá, 2007,

ROCCA, Paulina D. A. **O desempenho de falantes bilíngues: evidências advindas da investigação do VOT de oclusivas surdas do inglês e do português**. Revista DELTA: Documentação de Estudos em Linguística Teórica e Aplicada. Vol 19, n. 2. São Paulo, 2003.

SOUZA, Aline de. **A Produção das oclusivas desvozeadas do Inglês por aprendizes brasileiros: uma análise acústica**. Trabalho de Conclusão de Curso de Licenciatura em Letras Português-Inglês. Curitiba: Universidade Tecnológica Federal do Paraná, 2012.

STYLER, Will. **Using Praat for Linguistic Research**. Versão 1.4.5. University of Colorado, 2014.

WHITESIDE, S.P.; MARSHALL, J. **Developmental Trends In Voice Onset Time: Some Evidence for Sex Differences**. Sheffield: Karger, 2000.

ZIMMER, Márcia; BANDEIRA, Marta. **The dynamics of interlinguistic transfer of VOT patterns in multilingual children**. Linguagem & Ensino, Vol 15, n. 2. Pelotas, 2012.

Anexos

Anexo I – Tabelas de Medidas de VOT

Tabela 5 - Medidas de VOT do Grupo Controle

INFO	PA		PU		PI		PA		PU		PI	
	n1	n2	n1	n2	n1	n2	d1	d2	d1	d2	d1	d2
	0.012	0.015	0.013	0.021	0.014	0.019	0.068	0.078	0.067	0.058	0.031	0.074
GC1	25	17	05	67	92	96	29	96	21	9	44	83
	0.011	0.012	0.020	0.026	0.021	0.026	0.013	0.009	0.015	0.015	0.015	0.014
GC2	82	39	81	8	34	72	38	84	7	13	78	43
	0.018	0.019	0.022	0.026	0.030	0.029	0.023	0.008	0.049	0.030	0.012	0.028
GC3	55	31	03	25	08	23	8	497	1	48	45	16
	0.024	0.018	0.037	0.036	0.012	0.020	0.020	0.009	0.021	0.022	0.022	0.021
GC4	45	93	69	87	99	54	13	21	91	92	42	21
	0.019	0.013	0.010	0.026	0.014	0.016	0.023	0.032	0.019	0.032	0.019	0.020
GC5	03	81	64	61	51	6	46	46	43	06	2	25
	0.044	0.033	0.060	0.068	0.047	0.036	0.151	0.018	0.036	0.046	0.022	0.027
GC6	17	94	91	47	62	08	92	86	7	69	09	2
	0.010	0.017	0.030	0.021	0.013	0.017	0.014	0.023	0.022	0.021	0.017	0.018
GC7	71	53	03	88	23	01	76	99	11	92	33	18
	0.017	0.011	0.029	0.030	0.029	0.023	0.025	0.015	0.033	0.026	0.013	0.019
GC8	53	97	52	76	52	92	12	1	46	72	02	05
	0.015	0.018	0.025	0.028	0.033	0.019	0.011	0.023	0.012	0.030	0.016	0.021
GC9	32	04	3	84	69	45	36	7	24	42	72	27
	0.015	0.010	0.015	0.019	0.015	0.016	0.011	0.010	0.019	0.014	0.012	0.012
GC10	63	47	97	96	52	93	91	45	84	35	55	17

Tabela 6 - Medidas de VOT do Grupo Mulheres

INFO	PA		PU		PI		PA		PU		PI	
	n1	n2	n1	n2	n1	n2	d1	d2	d1	d2	d1	d2
GM1	0.014	0.030	0.035	0.039	0.022	0.024	0.021	0.028	0.030	0.031	0.027	0.030
	1	35	15	86	33	74	44	75	07	93	33	9
GM2	0.021	0.020	0.034	0.030	0.024	0.037	0.018	0.016	0.012	0.023	0.027	0.011
	97	94	47	16	08	57	01	55	77	4	85	6
GM3	0.027	0.028	0.042	0.049	0.021	0.021	0.018	0.019	0.020	0.025	0.018	0.010
	54	21	8	72	32	09	67	88	79	17	49	12
GM4	0.016	0.019	0.027	0.021	0.024	0.025	0.014	0.010	0.015	0.012	0.018	0.019
	98	17	05	04	58	19	8	24	92	59	88	75
GM5	0.011	0.011	0.016	0.020	0.023	0.034	0.018	0.011	0.022	0.024	0.019	0.027
	64	92	95	07	96	52	57	57	43	46	81	47
GM6	0.024	0.011	0.019	0.026	0.011	0.016	0.013	0.021	0.019	0.025	0.014	0.025
	21	5	85	19	51	19	5	21	94	61	61	73
GM7	0.025	0.014	0.016	0.030	0.017	0.013	0.014	0.013	0.033	0.022	0.015	0.012
	73	66	7	58	15	42	8	14	11	7	35	08
GM8	0.011	0.019	0.027	0.030	0.029	0.028	0.016	0.011	0.023	0.026	0.013	0.023
	96	15	81	37	17	74	93	07	97	08	57	33
GM9	0.025	0.025	0.027	0.037	0.018	0.021	0.019	0.018	0.025	0.024	0.016	0.015
	78	74	88	25	46	22	52	3	74	99	16	97
GM10	0.017	0.019	0.026	0.034	0.029	0.020	0.028	0.063	0.033	0.049	0.033	0.031
	03	37	14	16	19	93	43	15	77	75	1	52

Tabela 7 - Medidas de VOT do Grupo Homens

INFO	PA	PA	PU	PU			PA	PA	PU	PU		
	n1	n2	n1	n2	PI n1	PI n2	d1	d2	d1	d2	PI d1	PI d2
	0.015	0.015	0.016	0.029	0.017	0.022	0.018	0.016	0.013	0.013	0.007	0.018
GHI	65	18	54	69	16	02	51	22	88	75	5	55
	0.019	0.017	0.027	0.024	0.034	0.025	0.008	0.018	0.007	0.045	0.009	0.006
GH2	32	87	45	06	03	61	24	03	33	79	57	69
	0.022	0.019	0.013	0.016	0.018	0.019	0.016	0.014	0.027	0.029	0.025	0.027
GH3	51	92	48	33	13	01	01	34	74	28	94	39
	0.013	0.013	0.012	0.012	0.029	0.015	0.010	0.009	0.027	0.021	0.014	0.008
GH4	59	74	89	57	38	61	7	81	68	34	19	09
	0.012	0.007	0.015	0.008	0.007	0.007	0.010	0.009	0.022	0.018	0.006	0.011
GH5	17	77	95	67	97	99	16	79	97	6	25	44
	0.013	0.011	0.013	0.023	0.019	0.020	0.009	0.007	0.009	0.009	0.008	0.008
GH6	21	41	45	47	7	47	01	55	31	72	14	21
	0.021	0.012	0.032	0.025	0.006	0.012	0.011	0.017	0.014	0.018	0.017	0.015
GH7	89	08	82	09	45	23	23	52	76	21	68	32
	0.014	0.010	0.025	0.015	0.023	0.021	0.020	0.011	0.020	0.013	0.018	0.013
GH8	70	99	01	51	61	58	97	19	09	93	6	83
	0.025	0.026	0.024	0.020	0.026	0.020	0.015	0.023	0.015	0.018	0.005	0.015
GH9	43	63	6	77	6	9	09	7	46	11	04	22
	0.019	0.019	0.027	0.027	0.025	0.023	0.013	0.013	0.024	0.024	0.026	0.011
GH10	10	5	86	88	49	49	49	59	09	8	98	36

Anexo II – Gráficos

Gráfico 1 – Grupo Controle

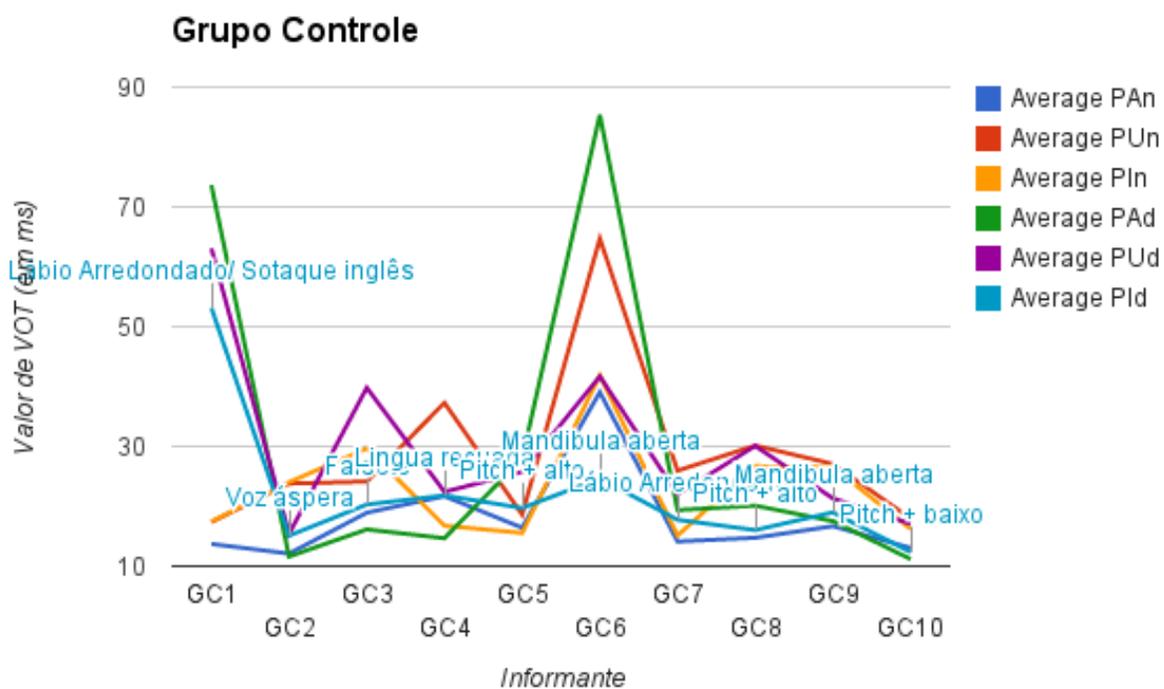


Gráfico 2 – Grupo Mulheres

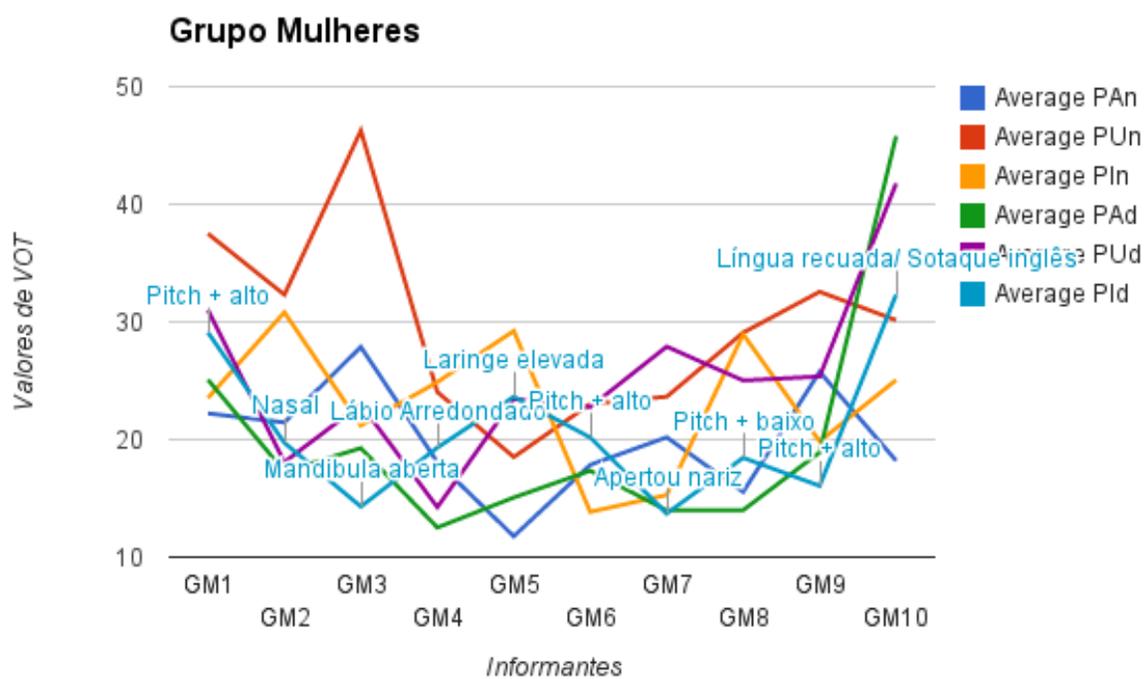


Gráfico 3 – Grupo Homens

