

CINTIA SCHIVINSCKI GONÇALVES

**TAXA DE ELOCUÇÃO E DE ARTICULAÇÃO
EM *CORPUS* FORENSE DO PORTUGUÊS BRASILEIRO**

Tese apresentada como requisito
para obtenção do grau de Doutor,
pelo Programa de Pós-Graduação em Letras
da Faculdade de Letras da Pontifícia Universidade
Católica do Rio Grande do Sul.

Orientador: Profa. Dra. Cláudia Regina Brescancini

Porto Alegre

2013

Dados Internacionais de Catalogação na Publicação (CIP)

G635t Gonçalves, Cintia Schivinski

Taxa de elocução e de articulação em *corpus* forense do português brasileiro / Cintia Schivinski Gonçalves. – Porto Alegre, 2013.

192 f.

Tese (Doutorado) – Fac. de Letras, PUCRS.

Orientador: Profa. Dra. Cláudia Regina Brescancini.

1. Linguagem. 2. Fonética Forense. 3. Taxa de Elocução. 4. Taxa de Articulação. 5. Tempo de Fala. 6. Prosódia. I. Brescancini, Cláudia Regina. II. Título.

CDD 469.15

Bibliotecária Responsável: Dênira Remedi – CRB 10/1779

CINTIA SCHIVINSCKI GONÇALVES

**TAXA DE ELOCUÇÃO E DE ARTICULAÇÃO EM *CORPUS* FORENSE
DO PORTUGUÊS BRASILEIRO**

Tese apresentada como requisito para
obtenção do grau de Doutor, pelo
Programa de Pós-Graduação em
Letras da Faculdade de Letras da
Pontifícia Universidade Católica do Rio
Grande do Sul.

Aprovada em 25 de março de 2013

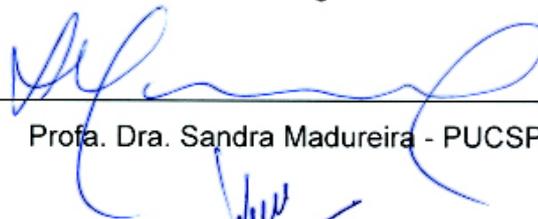
BANCA EXAMINADORA:



Prof. Dra. Claudia Regina Brescancini - PUCRS



Prof. Dr. Alexandre Rodrigues Meireles - UFES



Prof. Dra. Sandra Madureira - PUCSP



Prof. Dr. Denis Fernandes - PUCRS



Prof. Dr. Rodrigo Ghiringhelli de Azevedo - PUCRS

DEDICATÓRIA

Dedico este trabalho às minhas filhas, Lana e June, desejando que elas tenham em suas vidas: a coragem de empreender, a obstinação necessária à defesa de ideias e a perseverança para vencer os obstáculos que surgem no caminho.

AGRADECIMENTO

Agradeço aos que se empenharam na qualificação deste trabalho, especialmente: à profa. Cláudia Regina Brescancini, que me orientou com rigor e destacada competência; aos profs. Sandra Madureira e Alexsandro Rodrigues Meireles, pela valiosa interação e análise que conduziu à elaboração desta versão final de tese; aos demais membros da banca examinadora, pelo tempo e esforço despendido.

Agradeço ao meu esposo e aos amigos generosos, particularmente os da Seção de Perícias em Áudio e Imagens do Departamento de Criminalística do RS, por terem suportado minhas limitações durante esse intenso processo de construção.

Agradeço, ainda, a outros bons espíritos, que nos momento mais difíceis me inspiraram positivamente, com ideias e renovação de ânimo.

“Nem tudo que se enfrenta pode ser modificado,
mas nada pode ser modificado até que seja enfrentado.”

Albert Einstein

RESUMO

Este estudo tem por tema a taxa de elocução (TE) e de articulação (TA) em fala espontânea do português brasileiro (PB), angariada em gravação desavisada (interceptações telefônicas judicialmente autorizadas, realizadas sem a ciência dos locutores) e avisada (entrevista semidirigida sabida e consentida), contraponto situacional comumente encontrado em Fonética Forense, especificamente na perícia de Comparação de Locutor (CL). Objetivou-se através dele estabelecer o potencial individualizante da TE e da TA, visando à incorporação das taxas temporais ao conjunto de parâmetros técnico-comparativos utilizados na perícia de CL, assim como verificar a relação existente entre tais taxas e as variáveis independentes idade, sexo, escolaridade, *gap* temporal (entre as gravações desavisada e avisada), tipo de gravação e tamanho do intervalo de fala. Investigou-se a TE e a TA na fala de sete sujeitos (cinco do sexo masculino e dois do sexo feminino), estabelecidos no Rio Grande do Sul (Brasil), com o PB como língua materna e dialetos indefinidos devido à ocasional(is) aprisionamento(s). Os sujeitos integram o banco de dados do Instituto Geral de Perícias (IGP), órgão da Secretaria de Segurança Pública (SSP) do referido estado, nele figurando como alvo da perícia de CL, com resultado positivo para o confronto de perfil de voz e fala então efetuado. Mensurou-se os tempos de fala global e localmente, analisando-se 539 turnos de fala e 748 intervalos interpausais, considerados no cálculo, respectivamente, das TEs e das TAs. Os resultados apontam a existência de diferença não significativa entre os tipos de taxa (mostrando-se a TA menos variável do que a TE, especialmente na mensuração local), mas diferença significativa entre as formas de mensuração na TE. Quanto ao potencial individualizante das taxas, somente na TA a variabilidade intersujeitos foi superior à intrassujeito, obtendo-se, nesse caso, um coeficiente de correlação intraclasse indicativo de satisfatório poder discriminatório de falante. A análise da variabilidade por sexo e por tipo de gravação apontou como significativo somente o tipo de gravação na TE. Observou-se que os sujeitos de ambos os sexos tendem a diminuir as taxas quando têm ciência da gravação (diminuição maior no sexo masculino) e que na fala casual (gravação desavisada) são prevalentemente os homens os falantes com as maiores taxas. Encontrou-se diferença significativa entre os fatores da variável tipo de gravação na TE e correlação significativa entre a TE e a variável tamanho do intervalo de fala e entre a TA e a variável *gap* temporal. Considerando-se as 1.287 taxas locais, são significativos preditores do aumento da TE e da TA o fator masculino da variável sexo e o avanço na escolaridade e no *gap* temporal, enquanto que significativos preditores da diminuição da TE e da TA o avanço na idade e a ciência de gravação. Conclui-se pela indicação da incorporação da TA local média ao rol de parâmetros técnico-comparativos utilizados na perícia de CL, resguardada a máxima contemporaneidade entre as gravações confrontadas, e pela necessidade de adoção de providências que visem minimizar o impacto da ciência da gravação e de diferenças decorrentes tanto do estilo de fala próprio a cada um dos tipos de gravação quanto de eventuais incrementos na escolarização, ocorridos no *gap* temporal existente entre os áudios do cotejo.

Palavras-chave:

taxa de elocução – taxa de articulação – tempo de fala – organização temporal da fala – prosódia – Comparação de Locutor – Fonética Forense

ABSTRACT

This study deals with the speech rate (SR) and the articulation rate (AR) in Brazilian Portuguese (BP) spontaneous speech obtained from unaware audio recordings (legally authorized telephone interceptions) and known ones (semistructured interview), situations commonly counterposed in the Forensic Phonetics area, specifically in Speaker Comparison (SC). The objective was to establish the speaker-discriminating power of SR and AR, contributing to the discussion about the effectiveness of their use as a comparative parameter in SC. In addition, the relationship between the rates and the independent variables of study (age, sex, formal education, temporal gap between the recordings, type of recording, length of speech stretch) was investigated. The SR and AR in the speech of seven subjects (five males and two females) was analyzed. The subjects live in the State of Rio Grande do Sul, Brazil, have BP as their L1 and undefined dialects due to occasional detention(s). The recordings belong to the database of Instituto Geral de Perícias, a state department of security. All subjects present a positive result for the SC. Measurement of the rates was carried out globally and locally. SRs calculation involved 539 turn talks, and ARs calculation involved 748 interpause stretches. Statistical analysis showed that there is no significant difference between the rate types (SR and AR) and between forms of measurement in SR. The AR was less variable than the SR, especially in the local measurement. As far as the speaker-discriminating power of the rates, only the AR showed higher between-speaker variability than within-speaker variability, reaching a satisfactory intraclass correlation coefficient. The variability analysis by sex and type of recording was significant only for type of recording in SR. As a trend, it was observed that the subjects of both sexes decrease their rates when they are aware of the recording (reduction higher in males), and that in naturalistic speech (unaware recording) males predominantly presented the highest rates. Significant difference was found between the factors of the variable type of recording in SR and significant correlations between SR and length of speech stretch, and between AR and temporal gap. Based on the multiple local stretches, the significant predictors of SR and AR increase found in the analysis were sex (male factor), educational advancement, and higher temporal gaps. The significant predictors of decrease in SR and AR found in the analysis are advancing age and awareness of recording. It is possible to conclude that it is indicated to incorporate the AR local mean to the list of parameters used in the SC, provided there is a minimum temporal gap between contrasted recordings. Caution is recommended to minimize the impact of awareness of recording and of differences related to inherent speech style, as well as differences associated to educational advancement that might occur during the temporal gap between the confronted recordings.

Keywords:

speech rate – articulation rate – speech tempo – temporal organization of speech – prosody – Speaker Comparison – Forensic Phonetics

LISTA DE ILUSTRAÇÕES

Figura 1 – Configuração comumente encontrada na Comparação de Locutor no Brasil	27
Figura 2 – Recursos de análise comumente utilizados no exame acústico da Comparação de Locutor	30
Figura 3 – Esquema da Teoria fonte-filtro para vogais	34
Figura 4 – Espectrogramas em banda estreita e larga de frequência	34
Figura 5 – Participação das fontes sonoras na constituição das classes	35
Figura 6 – Proveniência dos formantes e antiformantes dos contoides ressoantes de segmentos	36
Figura 7 – Periodicidade tipicamente associada à enunciação vocálica	37
Figura 8 – Exemplar de ditongo decrescente e crescente	38
Figura 9 – Pré-plosão, soltura e zona de ruído típicos	39
Figura 10 – VOTs tipicamente observados na enunciação de plosivas	40
Figura 11 – Localização dos vestígios tipicamente associados às fricativas [f, v, s, z, ʃ, ʒ]	40
Figura 12 – Regularidade oscilográfica e antirressonância típicas das nasais ...	41
Figura 13 – Disposição do tempo de fala em relação aos demais elementos da produção da fala	45
Figura 14 – Variação estilística: principais abordagens e respectivas fontes de Variação	54
Figura 15 – Terminologia e conceitos relativos às duas principais medidas do tempo de fala	61
Figura 16 – Variáveis e fatores de variáveis admitidos no estudo	86
Figura 17 – Exemplar da etiquetagem utilizada no procedimento do estudo	92
Figura 18 – Exemplar de turno de fala com mais de um intervalo interpausal e da aplicação dos critérios de notação “IL”, “T” seguido de n ^o de ordem e “FS”	94
Figura 19 – Exemplos de turnos de fala notados como “INI”, “VOC” e “PEQ” .	95
Figura 20 – Exemplar de turno de fala notado como “INT”	96
Figura 21 – Exemplar de apontamentos adicionais e da transcrição fonética local empregada	99
Figura 22 – Exemplar de alongamento consonantal, na emissão de [ʒ]	100
Figura 23 – Exemplar de alongamento vocálico em situação de ênfase	101

Figura 24 – Exemplo de alongamento vocálico em pausa preenchida (“ã...”) ...	101
Figura 25 – Exemplo de descarte de turno de fala com aproveitamento de intervalos interpousais	103
Figura 26 – Esquema da forma de cálculo global e local médio das taxas do estudo	104
Figura 27 – Gráficos Q-Q dos dados relativos à TE global e local média	111
Figura 28 – Gráficos Q-Q dos dados relativos à TA global e local média	111
Figura 29 – Médias de TE global e local média por tipo de gravação e sexo	116
Figura 30 – Médias de TA global e local média por tipo de gravação e sexo	116
Figura 31 – Gráficos P-P dos resíduos relativos a TE global e local média	124
Figura 32 – Gráficos P-P dos resíduos relativos a TA global e local média	124
Figura 33 – Histograma e gráficos Q-Q dos resíduos relativos às TE locais	129
Figura 34 – Histograma e gráficos Q-Q dos resíduos relativos às TAs locais	130

LISTA DE TABELAS

Tabela 1 – Dados dos sete sujeitos selecionados para participação no estudo .	90
Tabela 2 – Quantidade e duração dos turnos de fala úteis dos sujeitos do estudo	97
Tabela 3 – Quantidade e duração total dos intervalos interpausais úteis dos sujeitos do estudo	103
Tabela 4 – Resumo estatístico relativo às quatro variáveis de resposta do estudo	109
Tabela 5 – Correlação entre as formas de mensuração das taxas pesquisadas	112
Tabela 6 – Teste t para amostras pareadas quanto à forma de mensuração	113
Tabela 7 – Estatísticas descritivas e Teste t para amostras independentes por sexo das quatro variáveis de resposta	114
Tabela 8 – Estatísticas descritivas e Teste t para amostras pareadas por tipo de gravação das quatro variáveis de resposta	114
Tabela 9 – Análise de variância da TE global e local média quanto ao sexo e tipo de gravação	115
Tabela 10 – Análise de variância da TA global e local média quanto ao sexo e tipo de gravação	115
Tabela 11 – TE e TA global e local média dos sujeitos por tipo de gravação	118
Tabela 12 – Teste de Brown-Forsythe para igualdade de médias entre sujeitos	118
Tabela 13 – Estimativas da variância e CCI para as quatro variáveis de resposta	120
Tabela 14 – Correlação entre as taxas pesquisadas e as variáveis idade, escolaridade e <i>gap</i> temporal	121
Tabela 15 – Resumo estatístico da regressão linear com cinco variáveis independentes	122
Tabela 16 – Coeficientes da regressão linear relativos a TE global e local média	122
Tabela 17 – Coeficientes da regressão linear relativos a TA global e local média	123
Tabela 18 – Resumo estatístico da regressão linear com seis variáveis independentes	127
Tabela 19 – Coeficientes da regressão linear relativos às TEs locais	128
Tabela 20 – Coeficientes da regressão linear relativos às TAs locais	128
Tabela 21 – Propriedades da amostra de fala do sujeito identificado como “S1”	142
Tabela 22 – Correlação entre as variáveis idade e <i>gap</i> temporal	147

SUMÁRIO

INTRODUÇÃO	14
1 FUNDAMENTOS CONSIDERADOS	22
1.1 FONÉTICA FORENSE	22
1.1.1 A perícia de Comparação de Locutor	25
1.2 FONÉTICA SEGMENTAL E SUPRASSEGMENTAL	33
1.2.1 A Teoria Fonte-Filtro	33
1.2.2 Acústica dos segmentos	36
1.2.3 O nível suprasegmental da fala: prosódia e tempo (taxa) de fala	42
1.3 SOCIOFONÉTICA	46
1.3.1 A entrevista sociolinguística	50
1.3.2 Variações estilísticas	52
1.3.3 Comunidade de fala, redes sociais e comunidade de prática	57
2 SOBRE O TEMPO DE FALA	60
2.1 ASPECTOS TEMPORAIS DA FALA	60
2.2 VARIABILIDADE, POTENCIAL INDIVIDUALIZANTE E REFERÊNCIA AO USO FORENSE DA TE E DA TA	73
2.3 POSSÍVEIS CONDICIONADORES DA TE E DA TA	76
2.3.1 Extralinguísticos	76
2.3.2 Linguísticos	81
3 METODOLOGIA	83
3.1 DELINEAMENTO DO ESTUDO	83
3.1.1 Tipo de pesquisa	83
3.1.2 Variáveis	83
3.1.3 Objetivos	86
3.1.4 Hipóteses	88
3.2 AMOSTRA	89
3.2.1 Autorizações indispensáveis ao desenvolvimento do trabalho	89
3.2.2 Seleção dos sujeitos	89
3.3 ORGANIZAÇÃO DOS DADOS	90
3.3.1 Equipamentos e <i>softwares</i> utilizados	90
3.3.2 Montagem dos <i>corpora</i>	91
3.3.3 Compilação das gravações desavisadas	91
3.3.4 Segmentação e etiquetagem dos áudios	92
3.4 INSTRUMENTOS DE ANÁLISE	105
3.4.1 Procedimento de verificação acústica	105
3.4.2 Tratamento estatístico dos dados	105
4 RESULTADOS	108
4.1 APRESENTAÇÃO DOS RESULTADOS OBTIDOS	108
4.1.1 Análises baseadas nas TEs e TAs globais e TEs e TAs locais médias	109

4.1.2 Análises baseadas nas TEs e TAs globais e TEs e TAs locais múltiplas	125
4.2 DISCUSSÃO DOS RESULTADOS	130
4.2.1 A pertinência das decisões teórico-metodológicas	130
4.2.2 Resposta aos objetivos estabelecidos	135
4.2.3 Justificativas para os comportamentos não previstos	140
CONCLUSÃO	151
REFERÊNCIAS BIBLIOGRÁFICAS	156
GLOSSÁRIO	171
APÊNDICES	175
ANEXOS	184

INTRODUÇÃO

A Comparação de Locutor^{1,2} é um dos tipos de perícia realizado na esfera criminal. Atinente à área de Fonética Forense, é caracterizada pelo confronto que estabelece entre duas amostras de fala, visando definir se essas foram produzidas pelo mesmo indivíduo.

Os métodos de análise empregados no desenvolvimento da Comparação de Locutor no Brasil estão em relativa consonância com o apregoado pela comunidade científica internacional (segundo panorama apresentado por Gold e French, 2011), tendo-se como concordante também o entendimento acerca da necessidade de se aumentar o número de parâmetros quantitativos considerados (KÜNZEL, 1994), a fim de diminuir o ainda prevalente subjetivismo.

Atualmente já se tem consolidada a percepção de que no confronto forense de voz e fala não existe parâmetro que possa, isoladamente, ser utilizado como referência individualizante indelével, de forma que as conclusões sobre a autoria das emissões orais consideram não apenas um, mas um conjunto de parâmetros técnico-comparativos, sendo o comportamento vocal e linguístico dos locutores do cotejo escrutinado em suas características gerais e particularizantes.

Nesse contexto, entende-se como muito bem-vindo todo e qualquer elemento que contribua para a definição confiável acerca da compatibilidade (ou não) entre as amostras de fala confrontadas, especialmente aqueles que propiciem a observação quantitativa.

Apesar de ser um elemento tão variável na fala, especialmente na fala espontânea, e a despeito da ocorrência de modificações intencionais, as medidas de emprego do tempo na fala têm indicação de uso na perícia de Comparação de Locutor (CAO e WANG, 2011; ERIKSSON, 2012; FRENCH e HARRISON, 2007;

¹ Neste estudo será utilizada a expressão “Comparação (Forense) de Locutor”, em detrimento de similares, em razão de ser essa a forma que efetivamente representa (cf. ERIKSSON, 2012; FRENCH e HARRISON, 2007; GOLD e FRENCH, 2011; WATT, 2010) o tipo de reconhecimento técnico-pericial aqui considerado (discussão sobre a questão é apresentada na Subseção 1.1, Capítulo 1).

² A acadêmica-autora deste trabalho é Perita Criminal Oficial do Estado do Rio Grande do Sul, lotada desde o ano de 2005 na Seção de Perícias em Áudio e Imagens (SEPAI) do Departamento de Criminalística (DC), onde realiza como rotina, entre outras, a perícia de Comparação de Locutor referida.

FRENCH et al., 2010; GOLD e FRENCH, 2011; HOLLIEN, 2002; JESSEN, 2007; KÜNZEL, 1997; ROSE, 2002), o que se atribui ao relatado fato de que alguns indivíduos falam habitualmente mais rápido ou mais lentamente do que outros (TSAO et al., 2006; TSAO e WEISMER, 1997). Para Jacewicz et al. (2010), tais diferenças individuais refletem o caráter individualizante da velocidade dos movimentos articulatórios, assim como do uso singular da prosódia e das pausas. Esses autores, em investigação a partir do inglês americano, concluem que o tempo usado na fala é uma característica própria do falante.

Diagnóstico similar é feito por Meireles e Barbosa (2009, p.25), segundo os quais:

Cada falante possui um tempo intrínseco próprio que faz com que não haja uma taxa de fala única para cada dialeto. Essa taxa depende de cada um dos falantes. Há falantes, que, por natureza, falam mais rápido, e falantes que falam mais lento.

No Brasil, os peritos em Fonética Forense que incluem em suas análises a verificação do tempo de fala³, avaliam-no, normalmente, de forma impressionística, expressando qualitativamente o que de oitiva constata. Os resultados por eles apresentados consideram uma escala que classifica globalmente as produções de um dado sujeito⁴, através do julgamento subjetivo de parença com o que seria uma taxa normal de fala, uma taxa rápida (fala acelerada) ou uma taxa lenta (fala lentificada). A conduta denota não só a carência de cientificismo ao empregar-se o parâmetro no mencionado tipo de perícia, como também a importância de se investir (os peritos em registros de áudio, ao longo de suas formações) em estudo e pesquisa que permitam aprofundar e desenvolver conhecimento em linguística.

A fim de contribuir para a diminuição de tal déficit, este estudo investiga o comportamento das duas medidas temporais de fala mais difundidas, a taxa de elocução (doravante referida simplesmente como “TE”) e a taxa de articulação (doravante referida simplesmente como “TA”), empregadas em situação de fala

³ “Tempo de fala” é o termo genérico normalmente utilizado quando não é oportuna a especificação acerca de um tipo de taxa em especial, sendo esse empregado ao longo do presente texto, sempre que for conveniente. A adoção da expressão genérica também é feita por Künzel (1997) e Dancovičová (1997), que empregam, respectivamente, “*Speaking tempo*” (forma não analisada da relação existente entre a fala e o tempo de sua produção) e “*rate of speaking*” (forma indistinta de taxa de produção de fala).

⁴ Como a apresentada por Hollien (2002, p.84).

espontânea por sete falantes vinculados à prática delituosa, pertencentes ao banco de dados VOXDC (caracterizado na Seção 3.2, Capítulo 3), que têm o português brasileiro (PB) como língua materna. Grosso modo, admite-se a TE como sendo o número de unidades linguísticas constantes em um turno de fala dividido pela duração do mesmo (TE local) ou o resultado da soma das unidades linguísticas constantes nos diversos turnos de fala do locutor no diálogo dividido pelo resultado da soma das respectivas durações (TE global). Já a TA, neste estudo, é entendida como sendo o número de unidades linguísticas constantes em um intervalo de fala antecedido e sucedido por pausas silenciosas dividido pela duração do mesmo (TA local) ou o resultado da soma das unidades linguísticas constantes nos diversos intervalos interpausais do locutor no diálogo dividido pelo resultado da soma das respectivas durações (TA global).

Destaca-se que a diferença paradigmática entre os tipos de taxa reside no fato de que a primeira, por conceito, é pausa-includente e a segunda pausa-excludente. A TE, contudo, não representa o maior desafio neste estudo, uma vez que, conforme Verhoeven et al. (2004), implica um número bem menor de decisões teórico-metodológicas se comparado ao necessário para cálculo da TA.

De maneira geral objetiva-se contribuir para a discussão acerca da efetividade da incorporação das medidas temporais (especificamente da TE e da TA) ao rol de parâmetro técnico-comparativos utilizados na realização da perícia de Comparação de Locutor, através da investigação da variabilidade da TE e da TA em *corpus* do PB, valendo-se, para tanto, de um método científico de mensuração que seja dirigido à rotina em Fonética Forense.

Na tarefa, será investigada a existência de diferença significativa entre as taxas obtidas a partir da mensuração global em relação às obtidas por média das mensurações locais, assim como se buscará a identificação do tipo de taxa que apresenta maior potencial discriminante de locutor. Ainda, pretende-se identificar possíveis condicionadores das taxas estudadas, através da verificação do impacto gerado pelas variáveis independentes admitidas (idade, sexo, escolaridade, *gap* temporal, tipo de gravação e tamanho do intervalo de fala) nos valores de TE e de TA empregados pelos sujeitos da pesquisa.

Tendo-se por base os resultados de trabalhos afins já disponibilizados na literatura, parte-se da hipótese de que a TA seja a medida mais indicada à aplicação forense, mostrando-se em algum grau capaz de distinguir falantes. Ambas as taxas

(TE e TA), infere-se, são afetadas pela idade (com sujeitos mais velhos empregando menores taxas), pelo sexo (com os homens sendo os falantes com taxas mais altas) e, especialmente, pelo tipo de gravação e pelo tamanho do intervalo de fala.

Quanto aos dois últimos possíveis condicionadores prevê-se, com certo destaque, que a ciência de estar sendo gravado, existente em somente um dos dois tipos de gravação considerados nesta pesquisa, interfere significativamente na taxa temporal empregada (reduzindo-a) e, ainda, que intervalos de fala compostos de um número maior de sílabas são veiculados a taxas mais elevadas, acarretando a diminuição da duração média das sílabas.

Cao e Wang (2011), em estudo sobre a variação da TA no chinês, indicam que a TA global e a TA média (derivada de múltiplas observações locais) podem ser utilizadas como um fator discriminatório na Comparação Forense de Locutor. Também Künzel (1997), pesquisando falantes do alemão, aponta a TA como um promissor parâmetro individualizante. No entanto, muito há a se indagar e descobrir acerca da variabilidade das taxas estudadas (TE e TA) e, por consequência, quanto ao caráter individualizante das mesmas. Encontra-se na literatura desde autores que exacerbam seus potenciais discriminatórios de indivíduo (GOLDMAN-EISLER, 1968) quanto os que concluem pela substancial variabilidade intrassujeito das mesmas (GROSJEAN E DESCHAMPS, 1975). Algumas das divergências devem-se ao equívoco que se estabelece ao se comparar resultados provindos de estudos com diferentes assunções teóricas e/ou metodológicas.

No Brasil, a TE e a TA são prevalentemente abordadas como intervenientes, ou seja, como variáveis independentes. Pode-se citar a observação da TE e/ou TA, por exemplo:

- a) na composição de um dado modelo, como na proposta do Modelo Dinâmico do Ritmo da Fala (BARBOSA, 2006), onde é observado, entre outros, se e de que forma a TE promove a reestruturação rítmica dos enunciados;
- b) no estudo da geração de novos padrões rítmicos estáveis e da redução silábica provocada pela modificação fonética decorrente da variação na TE (MEIRELES, 2007);
- c) na observação acerca dos processos fônicos prosódicos no dialeto capixaba e a consequente reorganização rítmica gerada pela manipulação da TE (SILVA, 2010);

- d) no estudo da interferência da TE na redução vocálica postônica e do grau de coarticulação das consoantes postônicas (MEIRELES e BARBOSA, 2008);
- e) na verificação da aplicação de processos de sândi externo (especificamente elisão, degeminação e ditongação) em fala espontânea produzida em diálogo e narrativa (COSTA, 2008).

As publicações que contemplam a TE e/ou a TA como variável dependente são em maioria produções estrangeiras, cuja leitura revela a dimensão do que cabe ainda ser pesquisado sobre o uso do tempo de fala, especialmente na produção espontânea, sendo a falta de uniformidade entre os métodos das pesquisas, reitera-se, um dos obstáculos à comparação dos resultados já existentes.

No manuseio do conteúdo percebe-se que a maioria dos achados acerca das taxas temporais considera amostras linguísticas geradas em ambiente laboratorial, a partir de leitura oral, alguns com promoção artificial de variação da TE e/ou da TA (BUTCHER, 1981; CRYSTAL E HOUSE, 1990; GRUNWALD, 1983; KOREMAN, 2005; ZELLNER, 1998), ou seja, em condições que se contrapõem às da fala naturalística, fato que salienta a importância deste estudo e a possibilidade de que os resultados aqui alcançados contribuam para um melhor entendimento do tema abordado.

A necessidade de se investir no desenvolvimento de pesquisas que contemplem a fala como naturalmente produzida é destacada por Rischel (1992), sendo tal convicção concordante com a de inúmeros operadores forenses.

Nesse sentido contrapõem-se, na configuração do estudo em tela, os dados de fala advindos de dois tipos de gravação espontânea⁵: a gravação desavisada (que compreende um ou mais registros de interceptação telefônica judicialmente autorizada) e a gravação avisada (registro ambiental único de entrevista semidirigida, colhido para fins periciais), sendo tais gravações compulsoriamente não contemporâneas. A díade situacional mencionada compreende o delineamento

⁵ A viabilidade legal para utilização acadêmica de dados de natureza criminal, neste caso sob a guarda da Secretaria de Segurança Pública do Estado do Rio Grande do Sul (SSP-RS), baseou-se no fato de ser precipuo à entidade pericial o interesse no desenvolvimento científico e na qualificação dos profissionais que nela atuam. A autorização formal para tal uso foi obtida através do Ofício n.º 1815/2009-DG/IGP, datado de 07 de julho de 2009, assinado pelo então Diretor-Geral do Instituto Geral de Perícias (IGP) do RS, sendo uma cópia do mesmo apresentada em anexo (ANEXO A). Destaca-se que no pleito que deu origem ao mencionado ofício consta o comprometimento, por parte da acadêmica-autora, com a exclusão de todo trecho de áudio que possa vir a revelar a identificação das partes envolvidas no inquérito ou processo.

de confronto mais comumente encontrado na rotina de realização da perícia de Comparação de Locutor no Brasil.

Contrastando os mencionados tipos de gravação tem-se, em essência, o meio em que se estabelece o diálogo espontâneo (se ao telefone, no caso da gravação via interceptação, ou presencialmente, no caso da gravação ambiental) e a ciência (ou não) dos locutores acerca da feitura da gravação. Podem ser citados como aspectos divergentes entre as gravações desavisada e avisada, ainda, o grau de familiaridade entre os locutores do diálogo (na gravação desavisada esses são ao menos parceiros em atividade ilícita, enquanto que na gravação avisada são desconhecidos, com certa hierarquia a desfavor do locutor-alvo⁶), assim como particularidades decorrentes do sistema de coleta de áudio, entre elas, a localização do microfone em relação à fonte sonora.

Destaca-se que, no levantamento bibliográfico realizado, não foi encontrado qualquer estudo que tivesse contraposto, como o em tela, dados da configuração de trabalho mais comum na rotina brasileira de Comparação de Locutor: gravação desavisada *versus* gravação avisada e consentida. Dessa forma, tem-se a verificação do comportamento da TE e da TA realizada neste estudo como singular, sendo os achados maximamente aplicados à prática forense brasileira. Entende-se que sua principal contribuição é empírica e metodológica, não se desprezando, contudo, a possibilidade de que os dados apresentados e a discussão consequente agreguem ao constructo teórico em desenvolvimento.

Neste estudo, não se pretendeu a verificação do comportamento das taxas temporais em uma determinada comunidade de fala, tampouco se objetivou mostrar padrões regionais ou sociais específicos, isso em razão de não se ter disponível informações confiáveis no que se refere ao histórico sociolinguístico e ao tipo de exposição linguística ocorrida no *gap* temporal existente entre uma e outra gravação. Ainda assim, admite-se sua filiação à Sociofonética em razão da adoção da fala espontânea como material de análise, da investigação sobre o condicionamento provocado por variáveis sociais (se a TE e a TA são por elas influenciadas e em que grau), da consideração de fala com características estilísticas próprias e do emprego de procedimentos de verificação acústica.

⁶ Neste estudo o locutor-alvo corresponde, dos sujeitos que participam do diálogo contido no áudio examinado, ao admitido como participante da pesquisa.

Destacando-se pelo caráter aplicado (Linguística Forense), tem-se que esta investigação quantitativa se debruça sobre a fala efetivamente produzida, trazendo à luz dados, indicadores e regularidades generalizáveis, o que, conforme Serapioni (2000, p.191), reproduz “a magnitude e a causa do fenômeno social”.

A carência de estudos que verifiquem as taxas estudadas em material de fala espontânea do PB permite estimar que os resultados alcançados podem vir a colaborar para o aprimoramento da técnica empregada na perícia de Comparação de Locutor, qualificando os peritos e conseqüentemente seus Laudos Periciais, o que reverterá na melhora do nível do atendimento prestado à solução de demandas da sociedade relativas à materialização de delitos durante a persecução penal.

Com o objetivo de facilitar a concepção geral do trabalho aqui desenvolvido, organizou-se o conteúdo em quatro capítulos, os quais se subdividem logicamente em seções e essas em subseções.

No primeiro capítulo têm-se os preceitos teórico-conceituais que sustentam a configuração do estudo e que fundamentam a discussão dos resultados, desenvolvendo-se conteúdo referente à Fonética Forense, à Fonética Segmental e Suprasegmental e à Sociofonética.

A revisão de literatura relativa à TE e à TA é tema do segundo capítulo, sendo convenientemente abordados os conflitos teórico-metodológicos que permeiam a concepção de um estudo sobre os parâmetros temporais estudados (entre eles o tratamento a ser dado à pausa e à fala disfluente), a questão da variabilidade das referidas taxas nos sujeitos e entre os sujeitos pesquisados e a relação existente entre as taxas estudadas e cada uma das variáveis admitidas como independentes.

No terceiro capítulo é exposta a metodologia utilizada no desenvolvimento da tese, com foco no delineamento de pesquisa adotado, os critérios de composição do *corpus*, a forma de organização dos dados dos sete sujeitos selecionados como participantes do estudo, assim como o procedimento de verificação acústica e o tratamento estatístico utilizado.

Os resultados obtidos, bem como a análise e a discussão a partir desses elaborada, são apresentados no quarto capítulo, tendo-se considerado o aparato teórico e as informações da revisão de literatura na avaliação da pertinência das decisões metodológicas assumidas, na resposta frontal aos objetivos formulados e na proposição de justificativas fundamentadas para os resultados encontrados, especialmente os não previstos.

Após o último capítulo seguem a conclusão a que se chegou a partir dos resultados obtidos, apontando-se questões para pesquisas futuras, as referências relativas às obras mencionadas no corpo do texto (e nesse não explicitadas), um glossário com expressões técnicas específicas e os apêndices e anexos.

1 FUNDAMENTOS CONSIDERADOS

O tema desta tese (tempo de fala) insere-se no campo de estudo da fonética, sendo admitido, em concordância com Laver (1994), como um dos elementos suprasegmentais da organização temporal da fala. O presente estudo compreende um trabalho aplicado, de caráter prevalentemente quantitativo, que considera em seu desenvolvimento a utilização da análise acústica e, devido ao método empregado, premissas teóricas relativas à Fonética Segmental.

Ressalva-se que o tempo de fala, como já objetivamente referido na literatura (DANCOVIČOVÁ, 2001), carece de aparato teórico próprio, que satisfatoriamente explique o fenômeno *per se* (e não sua participação como interveniente no comportamento de outras variáveis). Em decorrência, a fundamentação teórica aqui apresentada tenciona, basicamente, sustentar as decisões metodológicas tomadas ao longo da implementação do trabalho e subsidiar a interpretação crítica dos resultados obtidos. Adicionalmente, contudo, conduzir-se-á a discussão no sentido de que a mesma possa convenientemente contribuir para o necessário desenvolvimento teórico acerca das medidas temporais de fala.

Neste capítulo, a primeira subseção oferecida compreende um panorama sobre Fonética Forense, com destaque para a perícia de Comparação de Locutor, ao que seguem os pressupostos atinentes à perspectiva fonética de segmentos e de suprasegmentos e uma subseção dedicada à exposição das premissas conceituais e metodológicas inerentes à área da Sociofonética.

1.1 FONÉTICA FORENSE

Nesta subseção será contextualizada a perícia de Comparação de Locutor, *locus* de inserção almejado para as taxas em estudo, com a apresentação de um panorama da atividade pericial relacionada à análise de materiais linguísticos.

Cabe ressaltar de antemão que, no Brasil, diferentemente do que ocorre em outros países (entre eles o Reino Unido, profícuo em publicações na área de Fonética Forense), a perícia criminal está a cargo exclusivamente do Estado, cabendo à iniciativa privada somente a participação autônoma como Perito Judicial ou Assistente Técnico da parte.

Na definição da área de aplicação, tem-se a Linguística Forense como o “estudo científico da linguagem dirigido aos objetivos e contextos forenses” (MCMENAMIN, 2002, p.86, tradução nossa). Tal conceituação permite a percepção de que a Linguística Forense abrange uma ampla gama de tipos de perícias relacionadas à linguagem oral (atendidas nos órgãos periciais brasileiros pelas seções de perícias em registros de áudio ou correlatas) ou escrita (atendidas nos órgãos periciais brasileiros pelas seções de documentoscopia ou correlatas). Independente do meio de manifestação, as perícias em produção de linguagem podem objetivar:

- a) a definição da origem de produção(ões) oral(is) ou escrita(s), ou seja, determinar o autor (locutor ou escritor) responsável pela feitura de uma determinada amostra linguística (essa intitulada, conforme Rose, 2002, amostra questionada);
- b) a verificação da autenticidade do conteúdo gravado ou escrito, assim como da mídia que porta, respectivamente, as falas e as escritas que se suspeita terem sido vitimadas por manipulação fraudulenta;
- c) a apreensão do texto originalmente falado (quando o áudio mostra-se ininteligível) ou escrito (nos casos de sobretraçados de adulteração). A transposição da oralidade para o meio gráfico é costumeiramente referida no Brasil como Degravação ou Transcrição Fonográfica (BRAID, 2003), requerendo, quando se observa baixo nível de inteligibilidade, a aplicação prévia de recursos de amplificação e filtragem para melhoramento da qualidade do sinal e consequente discriminação do proferido.

Especificamente quanto às análises em registros de áudio, de competência da intitulada área de Fonética Forense, são atualmente realizadas no Brasil: a Análise de Conteúdo (não havendo consenso quanto à pertinência de seu enquadramento como perícia) e as perícias de Verificação de Edição e de Comparação de Locutor (MORISSON et al., 2009), ressalvadas as variações na nomenclatura decorrentes de particularidades administrativas (dos Departamentos/ Institutos de Criminalística das Unidades da Federação) ou da formação e rigor técnico dos peritos responsáveis pelo atendimento das perícias em áudio.

A intitulada Análise de Conteúdo aproxima-se da tradicional rotina de degravação, porém, difere dessa essencialmente por não ser literal e por ter como

premissa a impossibilidade de correspondência fidedigna entre a versão oral e a versão gráfica (transcrita) da fala. Admitindo-se a impossibilidade de fazer tal transposição sem que ocorra a emissão de juízo⁷, procede-se à segmentação do sinal de áudio, sendo o conteúdo de cada trecho demarcado vinculado ao respectivo arquivo digital e apresentando em forma de sinopse, com localização das temáticas tratadas nos diálogos mantidos.

Na Verificação de Edição são procurados elementos indicativos de alteração no conteúdo original da gravação, especificamente os que possam levar à compreensão diversa da realidade, modificando o sentido das declarações, imputando ou eliminando culpa. Assim, procuram os peritos por vestígios de inserções (introdução de trechos oriundos de outras gravações), supressões (eliminação de trechos), superposições (mistura de ruídos ou de sons ao registro original ou apagamento) e remanejamentos (mudança na ordem dos eventos).

A metodologia utilizada na Verificação de Edição contempla as análises perceptivo-auditiva e acústica e prevê a localização de discontinuidades e anormalidades do sinal de áudio, a verificação de compatibilidade entre a banda de frequência do sinal e a mídia de armazenamento, a detecção de mudanças nos padrões de ruído de fundo que sejam incoerentes com a sequência de eventos, a identificação de pontos de ausência de sinal e de sinais transitórios que possam estar relacionados a operações analógicas de edição, além da examinação do processo de produção das falas registradas no sinal sonoro (coarticulação, prosódia, nexos discursivos, coerência contextual, coerência entre as intensidades do áudio e a posição relativa das fontes sonoras, entre outros).

O reconhecimento de um indivíduo a partir de sua voz e fala é visto na literatura sob o título de “Reconhecimento de Locutor” (HOLLIEN, 2002; NOLAN, 1983; ROSE, 2002). Segundo os autores, no Reconhecimento de Locutor pode-se almejar a identificação ou a verificação, envolvendo, ambas as tarefas, (NOLAN, 1983, p.6, tradução nossa) “[...] um processo de decisão que confirma ou nega que duas amostras de voz foram produzidas pelo mesmo aparato vocal”.

Nesse sentido BRAID (2003, p.6) afirma:

⁷ Não é atribuição do Perito Criminal a emissão de julgamento, entre eles o relativo ao significado comunicativo implícito (por exemplo, se uma dada sentença denota ironia, raiva, desprezo, etc).

Um exame de Verificação de Locutor também é capaz de desvincular o envolvimento de um inocente num crime que lhe possa estar sendo imputado, o que talvez seja até mais importante do que incriminar um culpado.

Na Identificação de Locutor procede-se à comparação da amostra de fala de um indivíduo desconhecido com um grupo de amostras de fala pertencentes a locutores de identidade sabida, montado para fins de realização de confronto⁸, ou que integram um determinado banco de dados de produções orais (ainda inexistente de forma consistente no Brasil).

Atualmente, as perícias em registro de áudio em âmbito nacional mostram-se imersas em um contexto no qual prevalece a tarefa de verificação, sendo confrontadas propriedades de, via de regra, duas amostras, tendo sido uma delas, a relativa ao locutor que se deseja saber a autoria, obtida através de interceptação telefônica⁹ e, a outra, a relativa ao locutor de identidade conhecida, recolhida pelos próprios peritos em procedimento de coleta técnica de padrão vocal¹⁰. O referido contraponto situacional compreende a intitulada Comparação de Locutor.

1.1.1 A perícia de Comparação de Locutor

Como exposto na subseção precedente, na Comparação de Locutor tem-se a verificação contrastiva de amostras de fala, com vistas ao reconhecimento de um locutor. Especificamente no Brasil, são possíveis solicitantes desse tipo de perícia a Autoridade Policial (Delegado(a)), o Policial Militar (Oficial responsável por Inquéritos

⁸ Na literatura internacional, encontra-se referência ao trabalho com “*voice line-ups*”, uma espécie de alinhamento de vozes (podendo o padrão do locutor-suspeito estar ou não entre elas), utilizado para reconhecimento por testemunha, profissional da área ou sistema automático (HOLLIEN, 2002).

⁹ Tem se mostrado cada vez mais raro o envio aos Departamentos/ Institutos de Criminalística de áudios armazenados em mídias analógicas (fita cassete, microcassete, VHS, etc). Já os dispositivos móveis de armazenamento de dados por meio eletrônico (cartões de memória, *pen-drives*, etc), quando encaminhados, acondicionam arquivos digitais de áudio normalmente compactados, cuja tendência é a inadequação para a perícia de Comparação de Locutor devido à inerente degradação acústica.

¹⁰ Procedimento que compreende a apresentação dos presentes, prólogo sobre o procedimento a ser realizado, teste de gravação para verificação da qualidade do sinal (com os locutores frente a frente, estando o microfone em pedestal de mesa localizado a pouca distância do fornecedor do padrão), gravação de fala espontânea sobre tema de interesse do locutor-alvo (sendo comuns relatos sobre história de vida pessoal e sobre as circunstâncias de eventuais aprisionamentos), atentando-se para que haja mínima interferência dos peritos. A isso segue registro de repetição de sentenças pré-definidas, desconsiderado no presente estudo.

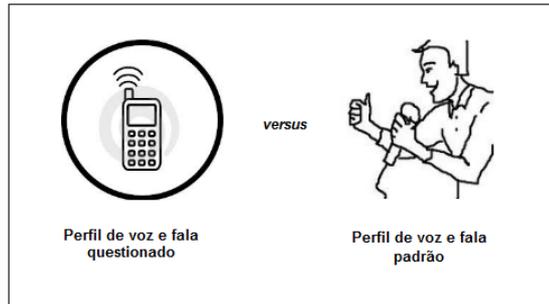
Policiais Militares), a Autoridade Judiciária (Juiz(a) de Direito) e o Ministério Público (Promotor de Justiça).

Por conceito, a Comparação de Locutor envolve “a comparação de uma amostra de fala teste [amostra questionada] com uma amostra de referência [amostra padrão] de um único locutor” (Nolan, 1983, p.8, tradução nossa). Na prática, contudo, observa-se, por vezes, a ocorrência de pequenas variações em tal delineamento de confronto, encontrando-se, por exemplo, a indicação de um locutor de identidade sabida e uma gravação (ou mais) com duas vozes passíveis de lhe serem atribuídas (leia-se: um locutor-padrão *versus* dois locutores-questionados). Ainda, pode-se encontrar a indicação de uma amostra de fala de autoria desconhecida e dois ou três suspeitos de a terem produzido (leia-se: um locutor-questionado *versus* dois ou mais locutores-padrão, esses, porém, não pertencentes a alinhamento pré-estabelecido ou banco de dados, mas apontados como suspeitos pelo solicitante da perícia).

Na perícia de Comparação de Locutor objetiva-se definir a autoria de falas armazenadas em uma determinada mídia, avaliando se essas, de fato, foram produzidas (ou não) pelo aparelho fonador de um determinado indivíduo (suspeito, indiciado ou réu). Segundo Nolan (1997, p.745, tradução nossa): “na maior parte das vezes a tarefa forense consiste em comparar duas amostras, [...] aplicando-se um critério para decidir se as amostras são suficientemente similares a ponto de procederem de um mesmo locutor”.

Para tanto, são analisadas as amostras apresentadas para confronto, tendo-se de um lado o áudio questionado fornecido pelo requerente (interceptações telefônicas legalmente autorizadas, radiodifusão ou gravações ambientais) e, de outro, o áudio padrão, preferivelmente recolhido por peritos da área em procedimento técnico de colheita de material. Os perfis de voz e de linguagem expressiva oral do locutor questionado e do locutor de identidade sabida são então cotejados (conforme Figura 1), identificando-se quais parâmetros, dos elencados, são indicativos de convergência e de divergência entre as amostras.

Figura 1 – Configuração comumente encontrada na Comparação de Locutor no Brasil



Fonte: A autora (2013).

Os elementos investigados podem ser qualitativos ou quantitativos. Não há “um conjunto pré-determinado de parâmetros, sendo a escolha ditada pela realidade das circunstâncias” (ROSE, 2002, p.33, tradução nossa).

French et al. (2010) apontam como comumente admitidos na Comparação de Locutor os seguintes parâmetros:

- a) configuração vocal (LAVIER, 1980, 1994), qualidade vocal e *pitch* (o último obtido através da média e da variação da frequência fundamental, doravante f_0);
- b) taxa de articulação;
- c) entonação e traços rítmicos;
- d) processos da fala encadeada (como padrões de assimilação e de elisão);
- e) traços consonantais (por exemplo, o *locus* de energia das fricativas; a soltura das plosivas; a duração de nasais, das líquidas e das fricativas em contexto fonológico específico; o tempo de início de vozeamento (VOT) das plosivas; a presença ou a ausência de pré-vozeamento em plosivas átonas e variáveis sociolinguísticas discretas);
- f) traços vocálicos (incluindo configuração formântica, frequência central, densidade, largura de banda e qualidade auditiva de variáveis (sic) sociolinguísticas);
- g) informações linguísticas de níveis mais altos (uso e padrão de marcadores discursivos, escolhas lexicais, variantes morfológicas e sintáticas, comportamento pragmático como os encontrados na tomada de turno de fala e no atendimento de ligações telefônicas, comportamentos multilinguais (como o *codeswitching*));

h) evidências de comprometimento de fala, de patologia de voz ou de fala e de traços não linguísticos característicos do falante (por exemplo, respiração audível, limpezas de garganta, cliques linguais e elementos marcadores de hesitação).

Ao inquirirem seus entrevistados quanto a que parâmetro pode ser considerado distintivo de locutor, Gold e French (2011) obtêm como resposta: qualidade vocal, variantes dialetais, formantes vocálicos, tempo de fala, f_0 , ritmo, escolhas lexicais e gramaticais, realizações vocálicas e consonantais típicas, processos fonológicos da fala encadeada e fluência da fala.

Segundo Nolan (1983), os parâmetros selecionados devem preferencialmente ter alta variabilidade interfalantes e baixa variabilidade intrafalante; serem resistentes à tentativa de disfarce; serem usualmente observados (mesmo em pequenas amostras); serem robustos a diferenças na transmissão (ou seja, não variar suas propriedades se advindos, por exemplo, de gravação telefônica ou de gravação ambiental) e serem facilmente mensuráveis.

Na problemática que perpassa a realização da perícia de Comparação de Locutor destaca-se a dificuldade em se ter definido o quanto cada um dos vários parâmetros técnico-comparativos utilizados varia no indivíduo (variabilidade intrassujeito) e entre os indivíduos da população (variabilidade intersujeito), já que é a razão entre as variâncias que representam essas variabilidades que explicita o potencial discriminante de um parâmetro, assim como a má qualidade acústica e a ingerência sobre o material sonoro típicos das amostras questionadas (ROSE, 2002; HOLLIEN, 2002).

As mencionadas variabilidades refletem diferenças sonoras e acústicas resultantes de desigualdades orgânicas (associadas à estrutura e funcionalidade dos elementos que integram o aparelho fonador¹¹), linguísticas (decorrentes, entre outros, da variedade de língua adotada pelo indivíduo), além de paralinguísticas¹².

Quanto à depreciação do sinal sonoro e à impossibilidade de controle das condições de gravação, cabe comentar que no Brasil, em razão de se primar pela coleta técnica de padrão vocal, esses tendem a se restringir à amostra questionada.

¹¹ O aparelho fonador compreende as pregas vocais, laringe supraglótica, faringe, cavidades oral e nasal, seios paranasais e lábios.

¹² Para Laver (1994) enquadra-se como paralinguístico o comportamento comunicativo não linguístico e não verbal projetado para alcançar duas metas na interação conversacional: comunicar o estado afetivo, atitudinal e emocional corrente do indivíduo e regular o tempo compartilhado na conversação.

Ressalva-se, contudo, que o padrão tecnicamente colhido, a despeito de contar com equipamentos e metodologia adequados, é comumente produzido em ambiente físico com propriedades não desejáveis, como salas de Presídios, de Delegacias ou de Fóruns, encarregando-se os peritos, nesses casos, de minimizar a interferência de efeitos sonoros adversos.

Quanto ao método empregado na Comparação de Locutor, cabe referir que em maio de 2007, no Brasil, teve início a implementação da Capacitação Nacional para Peritos Criminais em Fonética Forense, promovida pela Secretaria Nacional de Segurança Pública (SENASP), sob condução do Instituto Nacional de Criminalística (órgão central de criminalística da Polícia Federal), que apregoou e pulverizou (visto que contava com a participação de Peritos Criminais Oficiais estaduais, do Distrito Federal e da Polícia Federal de quase todas as Unidades da Federação) a procedência da utilização do método combinado, constituído das análises perceptivo-auditiva e acústica.

O panorama apresentado por Gold e French (2011), baseado na participação de 36 peritos de 13 países distintos (entre eles o Brasil), informa que também internacionalmente, considerando-se os respondentes, é prevalente o uso do método combinado, de emprego já referido por Byrne e Foulkes (2004); Kuwabara e Sagisaka (1995); McDougall (2005); Nolan (2001); Rodman et al. (2002); Romito e Galatà (2004); Rose (2002); Watt (2010), entre outros.

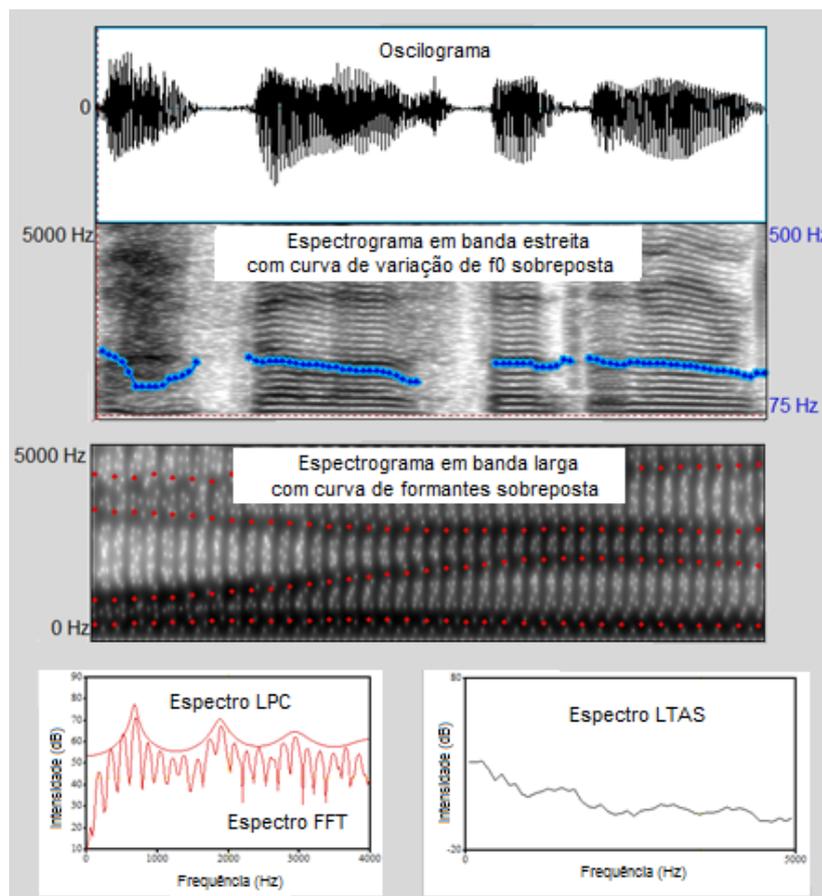
Desta feita, na Comparação de Locutor é utilizada, basicamente, a conjunção de duas análises, a perceptivo-auditiva e a acústica. Complementarmente, a perícia oficial de alguns estados brasileiros considera, ainda, resultados provenientes de sistemas de reconhecimento automático de locutor.

Amiúde, na análise perceptivo-auditiva (por vezes referida como “perceptual”) são investigados elementos vocais de caracterização geral do indivíduo, assim como os relativos à variedade linguística dos locutores confrontados. São observados fatores indicativos do sexo, da fase do ciclo de vida (se aparente infância, adolescência, fase adulta ou velhice), do estado de saúde dos órgãos fonoarticulatórios e da provável condição sociocultural e intelectual; características referentes à qualidade vocal e aos ajustes musculares utilizados na fonação; hábitos vocais típicos (pigarro, estalos, cliques, etc); forma de articulação; presença de desvios fonéticos (distorções) e/ou fonológicos (problemas no sistema de contrastes

da língua); alterações temporais, de ritmo ou fluência da fala; padrão entonacional empregado; coordenação pneumofonoarticulatória; idioleto e dialeto¹³; entre outros.

Já na análise acústica (por vezes referida como “instrumental”) são extraídas medidas físicas que documentam a condição e o comportamento de fatores segmentais e suprasegmentais, resultantes de configurações específicas do aparelho fonador, objetivando-se corroborar ou refutar os achados perceptuais. As informações são obtidas com a aplicação de recursos de análise disponibilizados em *softwares* de edição de áudio, sendo comumente utilizados oscilogramas (formas de onda), espectrogramas (em banda larga e estreita de frequência), curvas de formantes e de variação da f_0 e espectros FFT (*Fast Fourier Transform*), LPC (*Linear Predictive Code*) e LTAS (*Long-term Average Spectrum*). Os recursos de análise recém-mencionados são ilustrados na Figura 2, a seguir.

Figura 2 – Recursos de análise comumente utilizados no exame acústico da Comparação de Locutor



Fonte: A autora (2013).

¹³ Neste estudo assume-se a concepção de dialeto apresentada por Llamas et al. (2007, p.211), segundo a qual esse compreende “a pronúncia, o léxico e a gramática de uma variedade de língua, associada a uma área geográfica ou grupo social particular” (tradução nossa).

No cotejo entre as amostras podem ser consideradas sentenças, palavras, sílabas ou segmentos (fones), atentando-se para que os segmentos confrontados sejam pares quanto ao acento (ao menos lexical) e imersos em ambiente fonético antecedente e seguinte maximamente análogos.

Os resultados obtidos com as análises realizadas são apresentados tanto qualitativamente (especialmente no que se refere ao comportamento linguístico manifesto) quanto quantitativamente, por meio de estatística descritiva, restrita normalmente à exposição das medidas extraídas durante o exame acústico¹⁴ e à caracterização da diferença percentual existente entre as amostras questionada e padrão.

Ressalta-se que, até o momento, como oportunamente exposto por Bonastre et al. (2003), não há um processo científico capaz de, uma vez aplicado, gerar resultados que permitam a identificação com absoluta certeza de um indivíduo a partir de sua voz e/ou fala. Assim, é aconselhável que o perito considere o maior número possível de parâmetros técnico-comparativos, sendo obviamente preferidos os de maior potencial individualizante.

Na interpretação dos resultados da examinação, há referência na literatura do uso de escalas verbais tanto para indicação direta da conclusão (a exemplo a escala de nove pontos apresentada por Eriksson, 2012, p.60, a qual varia entre os extremos “os resultados suportam a hipótese [de unicidade das amostras] com quase certeza” e “os resultados contrariam a hipótese com quase certeza” (tradução nossa)) quanto para enquadramento *a posteriori* de resultados quantitativos (como no caso da razão de verossimilhança, cuja classificação verbal tem associada uma escala de escores baseada no grau de suporte da hipótese assumida).

No Brasil existe certa concordância entre os Peritos Criminais Oficiais quanto à necessidade de se considerar, na expressão das conclusões da perícia de Comparação de Locutor, tanto o número de vestígios a favor ou contra a hipótese de unicidade das fontes emissoras das falas cotejadas quanto o potencial individualizante dos parâmetros técnico-comparativos utilizados, o que vai, de certa

¹⁴ De uso mais recente, a razão de verossimilhança (*Likelihood Ratio*) constitui um novo paradigma na confecção e exposição dos resultados da perícia de Comparação de Locutor (MORRISON, 2009a; ROSE, 2002, 2006). Considerando os objetivos do estudo em tela, no entanto, não se entende como apropriado o aprofundamento acerca de tal recurso metodológico.

forma, ao encontro da proposta difundida internacionalmente como “UK Position Statement”¹⁵ (FRENCH e HARRISON, 2007).

Tal proposta torna oficial e público o posicionamento técnico de nove signatários e 16 cossignatários (destacados profissionais da área de Fonética Forense), que apregoam uma nova rotina de exposição de conclusões, a qual considera duas etapas. Na primeira etapa é definido se as amostras analisadas são consistentes entre si (tendo-se as opções, conforme os autores, p.141, tradução nossa, “Consistente”, “Não consistente” e “Nenhuma decisão”). Já a segunda etapa, restrita às amostras compatíveis entre si, é reservada ao estabelecimento do grau de distintividade do conjunto de parâmetros identificados como convergentes entre as amostras. Para tanto, é proposta uma escala verbal de cinco pontos, que evolui do “Não distintivo” ao “Excepcionalmente distintivo – a possibilidade dessa combinação de traços ser compartilhada por outro locutor é considerada remota” (FRENCH e HARRISON, 2007, p.141, tradução nossa).

Na prática observa-se que o corpo probatório construído no desenvolvimento do trabalho pericial naturalmente indica a in/compatibilidade entre as amostras, permitindo que o perito formule um juízo de convencimento tendente à unicidade ou não das fontes emissoras das falas cotejadas.

Gold e French (2011, p.303, tradução nossa) ressaltam:

Curiosamente, embora não surpreendentemente, ampla maioria dos participantes referiu que, a despeito da significância de algum parâmetro isolado, é a combinação do conjunto de traços que eles consideram crucial para a discriminação de falantes.

Dessa forma, fica claro que o resultado na Comparação de Locutor não considera apenas um, mas um conjunto de parâmetros (o mais extenso e robusto possível), no qual constam, segundo Cao e Wang (2011), Eriksson (2012), French e Harrison (2007), French et al. (2010), Gold e French (2011), Hollien (2002), Jessen (2007), Künzel (1997) e Rose (2002), as medidas do tempo de fala.

¹⁵ O documento oficial com o mencionado posicionamento é disponibilizado no sítio <http://www.forensic-speech-science.info/docs/position-statement.pdf>. Acesso em: 15 dez. 2012.

1.2 FONÉTICA SEGMENTAL E SUPRASSEGMENTAL

Após breve explanação introdutória sobre a Teoria Fonte-Filtro (FANT, 1960), será apresentada, nesta seção, a perspectiva fonética de ambos os níveis de análise linguística, segmental e suprasegmental, estabelecendo-se os fundamentos que norteiam as relações acústico-articulatórias consideradas especialmente durante o procedimento instrumental empregado no estudo em tela.

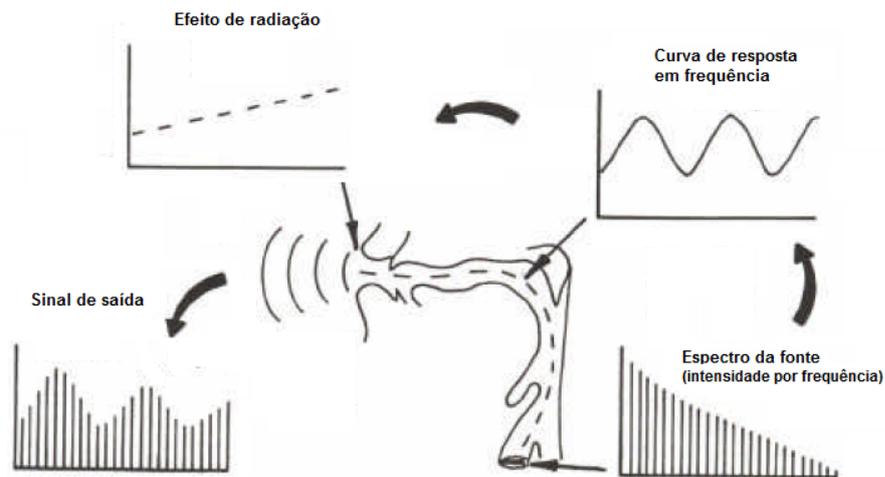
1.2.1 A Teoria Fonte-Filtro

Antes da exposição sobre fonética acústica segmental faz-se necessária breve explicitação acerca da difundida Teoria Acústica da Produção de Fala, também referida como Teoria Fonte-Filtro (FANT, 1960), concepção que desencadeou uma série de outras produções sobre a matéria, entre elas, Johnson (1997), Kent e Read (1992) e Ladefoged (2001, 2003).

De acordo com a mencionada teoria, o aparato fonador (trato vocal) é representado por um tubo, fechado em uma extremidade (a que corresponde ao nível glótico) e aberto na outra (a que corresponde ao limite dos lábios). O modelo, originalmente aplicado à produção de vogais, propõe que a energia de saída é resultante da combinação dos efeitos gerados pela fonte de energia (laríngea) e pelo ressonador (filtro).

Na Figura 3 é apresentado um esquema (elaborado a partir de Kent e Read, 1992) que ilustra a proposição feita pela Teoria Fonte-Filtro. Nele, o espectro da fonte (exibido na base da figura, à direita) tem energia que decresce à medida que aumenta a frequência do sinal, sendo filtrado pela função de transferência do filtro (curva de resposta em frequência exibida no alto da figura, à direita), decorrente das propriedades ressonantes do trato vocal. O sinal é submetido então ao efeito de radiação, que se comporta como uma espécie de filtro passa-alta (mais restritivo à passagem de energia em baixas frequências, como é possível observar no espectro exibido no alto da figura, à esquerda), resultando no sinal de saída ou *output* cujo espectro é apresentado na base da figura, à esquerda.

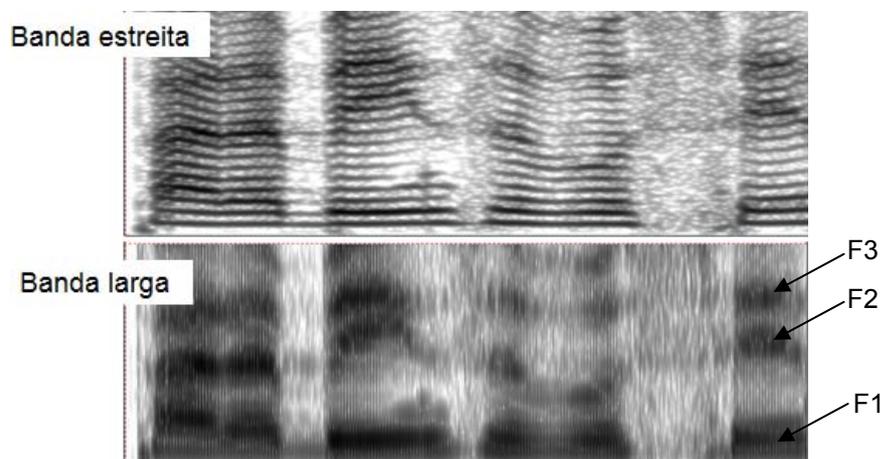
Figura 3 – Esquema da Teoria Fonte-Filtro para vogais



Fonte: Adaptado de Kent e Read (1992, p.19).

O espectro da fonte revela os harmônicos do som, também visualizados como linhas horizontais no espectrograma em banda estreita apresentado na Figura 4 (quadro superior), sendo os mesmos, por definição, múltiplos inteiros da f_0 . A filtragem ocorre nas porções do trato vocal subsequentes às pregas vocais (supralaringe, faringe, cavidade oral e nasal), onde existem zonas específicas de incremento de energia conhecidas como formantes. Os formantes são ressonâncias do trato vocal (KENT e READ, 1992) correspondentes na imagem dos espectrogramas em banda larga, conforme Figura 4 (quadro inferior), às barras largas horizontais formadas pelos tons mais escuros da escala de cinza, sendo classificados sequencialmente (F_1 , F_2 , $F_3...$), de acordo com o aumento de seus valores em frequência.

Figura 4 – Espectrogramas em banda estreita e larga de frequência



Fonte: A autora (2013).

Nota: No eixo horizontal consta a escala temporal e, no eixo vertical, a escala em frequência.

Cabe destacar que, apesar da energia dos sons da fala se estender sobre uma largura de banda (de frequência) de mais de 10.000 Hz, em transmissões telefônicas, um dos tipos de gravação contemplado nesta pesquisa, a banda passante restringe-se à faixa de 300 Hz a 3400 Hz (BYRNE e FOULKES, 2004).

Nessa perspectiva (FANT, 1960), consideram-se os segmentos em termos de suas fontes sonoras constituintes, agrupando-os em classes naturais de sons, essas dispostas em dois grandes grupos: o das obstruintes (plosivas, fricativas e africadas) e o das ressoantes (líquidas e nasais, além de vogais e semivogais), conforme ilustrado na Figura 5.

Figura 5 – Participação das fontes sonoras na constituição das classes de segmentos

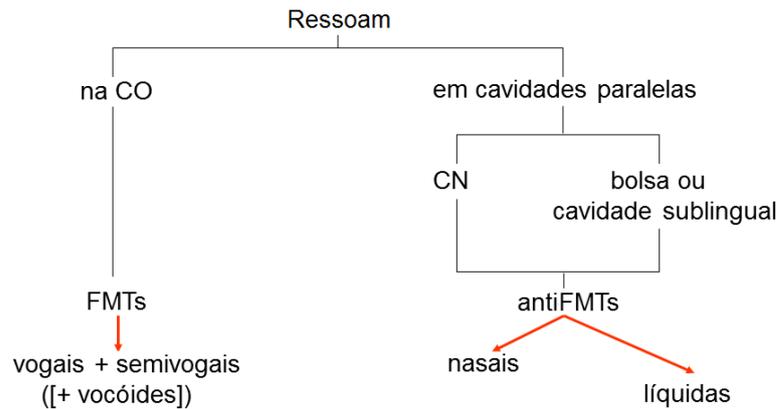
		FRT	FRC	FSON
Obstruintes	Plosiva surda	X		
	Plosiva sonora	X		X
	Fricativa surda		X	
	Fricativa sonora		X	X
	Africada surda	X	X	
	Africada sonora	X	X	X
Ressoantes	Líquidas, nasais, vogais e semivogais			X

Legenda: FRT= fonte de ruído transiente, FRC= fonte de ruído contínuo, FSON= fonte de sonoridade.
Fonte: A autora (2013).

Os segmentos ressoantes são assim definidos por apresentarem uma configuração articulatória propícia à ocorrência de ressonâncias no trato vocal, sendo, por essa razão, preferidos na inspeção formântica. Os formantes tendem a ser mais visíveis nos casos em que a ressonância efetiva-se na cavidade oral (como na produção de vogais e de semivogais, conforme consta na Figura 6) do que quando essa conta, majoritária ou secundariamente, com a participação de cavidades paralelas como a cavidade nasal ou a cavidade formada com a elevação da ponta da língua própria dos segmentos alveolares¹⁶, respectivamente presentes na produção das nasais e das líquidas (segundo exposto na Figura 6).

¹⁶ Nos segmentos alveolares a ponta da língua toca a região dos alvéolos, localizada atrás dos dentes incisivos superiores.

Figura 6 – Proveniência dos formantes e antiformantes dos contoides ressoantes



Legenda: CO= cavidade oral, FMTs= formantes, CN= cavidade nasal.
Fonte: A autora (2013).

Finalizado a apresentação relativa à Teoria Fonte-Filtro, aonde se procurou estabelecer conceitos acústicos elementares, prosseguir-se-á com a caracterização fonético-acústica dos segmentos da fala.

1.2.2 Acústica dos segmentos

No desenvolvimento deste estudo fez-se necessário recorrentemente associar as informações articulatórias aos achados acústicos correspondentes, tendo sido tal conhecimento especialmente importante durante a delimitação precisa dos intervalos de fala a serem examinados, na caracterização das pausas do discurso (quando se faz discriminação entre pausa silenciosa e pausa articulatória¹⁷) e na definição tanto dos segmentos efetivamente produzidos quanto do número de sílabas fonéticas (definidas na Subseção 2.1, Capítulo 2) constantes em cada um dos intervalos de fala identificados como passíveis de submissão ao cálculo das taxas em estudo.

Por essa razão apresenta-se, a seguir, a caracterização acústica das principais classes de sons linguísticos, ressaltando-se que a descrição dos vestígios associados aos vocóides e contoides aqui exibida compreende uma reunião do publicado em Fujimura e Erickson (1997), Kent e Read (1992), Ladefoged (2001, 2003), Ladefoged e Maddieson (1996) e Martins (1988).

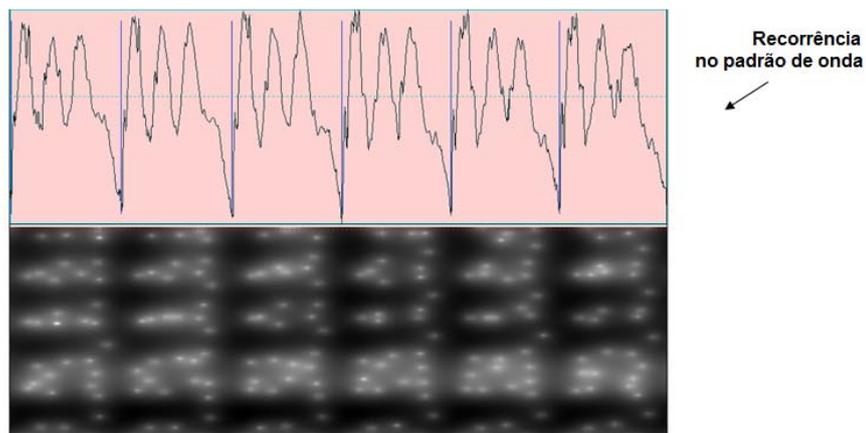
¹⁷ A pausa silenciosa compreende o tempo de não emissão usado linguisticamente ou para reposição do suporte aéreo enquanto que a pausa articulatória refere-se ao tempo de não emissão necessário para reposicionamento dos articuladores, questão expandida na Subseção 2.1 do Capítulo 2.

1.2.2.1 Vocoides

Os vocoides, que têm por representantes as vogais e as semivogais, são produzidos sem constrição significativa no aparelho fonador, mas com modificações nos articuladores que produzem padrões específicos de ressonância, visualizados especialmente através da configuração dos dois primeiros formantes. Quanto mais alto é o ponto articulatório da vogal ou da semivogal, menor tende a ser o valor em frequência de seu F1 (formante associado ao posicionamento vertical de língua) e quanto mais anterior é o ponto de articulação da vogal, maior tende a ser o valor de seu F2 (formante associado ao posicionamento ântero-posterior da língua). A fonte de energia nessa classe de segmentos é a sonoridade. Os achados acústicos representativos especificamente das vogais indicam que essas:

- a) são quase periódicas, como é possível observar no oscilograma e no espectrograma da Figura 7 (respectivamente exibidos no quadro superior e inferior), o que justifica serem classicamente utilizadas para visualização de harmônicos e formantes;

Figura 7 – Periodicidade tipicamente associada à enunciação vocálica



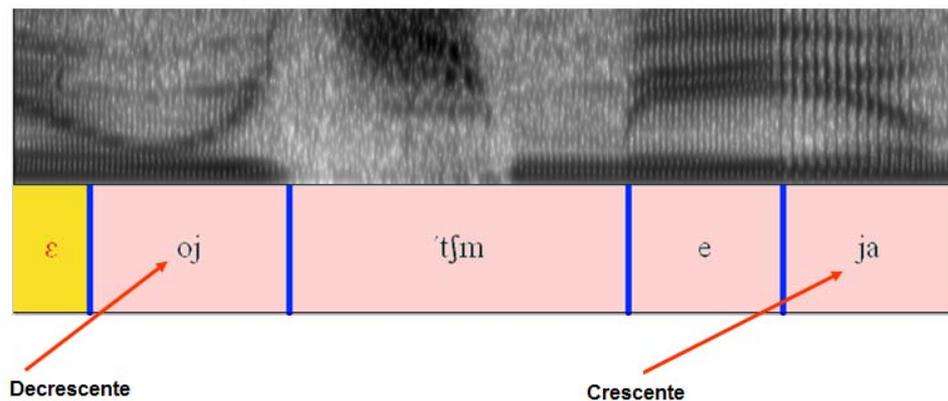
Fonte: A autora (2013).

- b) são intensas (por não terem fonte de ruído conjugada) e apresentam a maior parte da energia até 4000 Hz;
- c) possuem duração de, aproximadamente, 50 a 400ms;
- d) se nasais ou nasalizadas, são mais fracas do que as vogais orais, têm formantes com maior largura de banda e formante nasal de baixa frequência.

As semivogais apresentam propriedades acústicas muito similares às das vogais, podendo-se citar como diferencial o fato de que as semivogais:

- a) são mais curtas e mais fracas do que as vogais (pois implicam constrição mais estreita entre os articuladores), contendo menor porção estável, essa inclusive inexistente no segundo formante, conforme Padgett (2008);
- b) participam na formação dos ditongos decrescentes (vogal seguida de semivogal) e crescentes (semivogal seguida de vogal), respectivamente ilustrados à esquerda e à direita na Figura 8, com valores correspondentes aos monotongos [ɪ, ʊ]. São descritas normalmente quanto aos seus valores de entrada e saída.

Figura 8 – Exemplo de ditongo decrescente e crescente



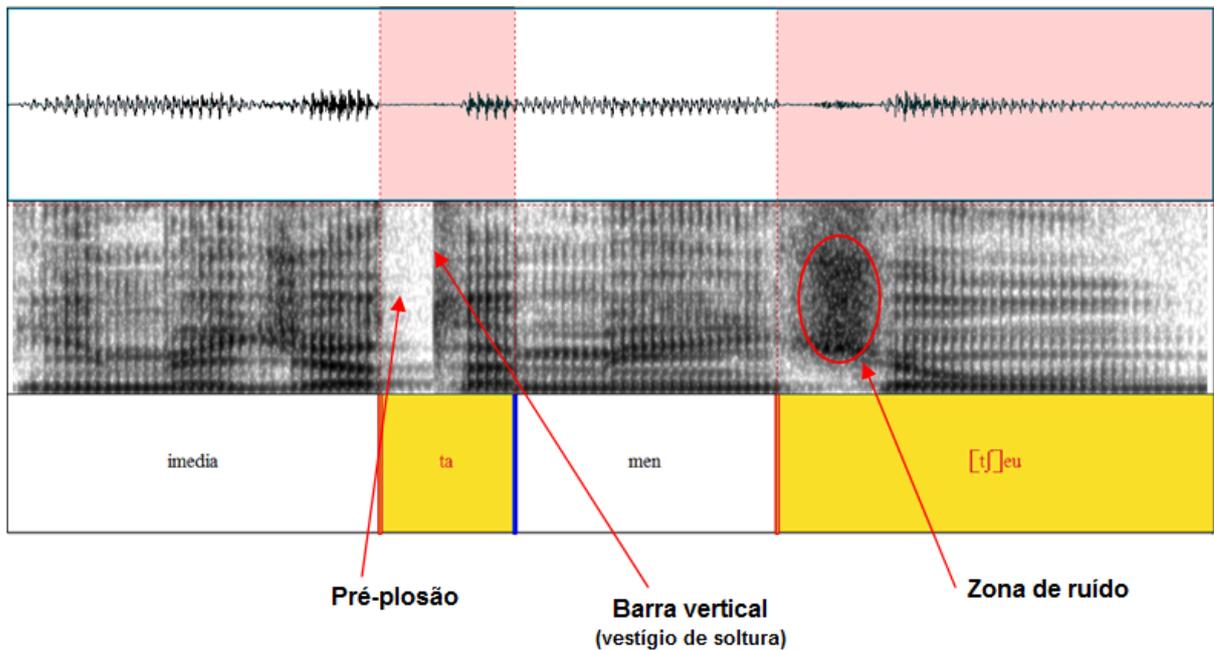
Fonte: A autora (2013).

1.2.2.2 Contoides

Os contoides caracterizam-se pela presença de obstrução da corrente aérea durante a articulação, sendo representantes desse grupo as plosivas, as fricativas e as africadas, adiante caracterizadas.

As plosivas e africadas preveem em sua produção um tempo inicial de contato absoluto entre os articuladores, ao que segue a soltura (explosão) da corrente aérea expiratória tipicamente encontrada nessas classes de segmento e, no caso das africadas, após a mencionada soltura, uma zona de ruído associada ao período fricativo do segmento. Os mencionados vestígios são ilustrados na Figura 9.

Figura 9 – Pré-plosão, soltura e zona de ruído típicos

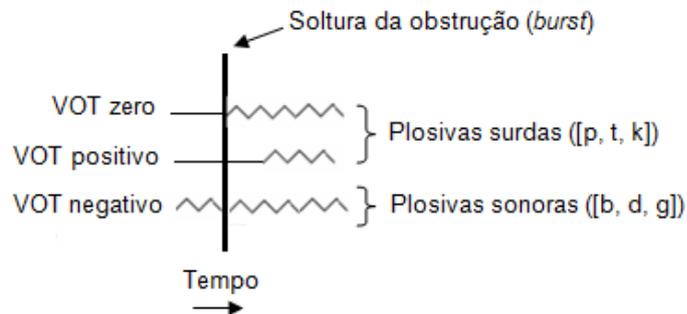


Fonte: A autora (2013).

São achados acústicos típicos das plosivas:

- o tempo pré-plosão, visualizado como intervalo de silêncio ou de baixa energia no sinal, apontado como menor nos exemplares sonoros;
- a barra vertical com aproximadamente 10 a 30ms, mais tênue nas manifestações sonoras e cuja concentração de energia varia de acordo com o ponto articatório da plosiva (bilabiais em torno de 500 a 1500 Hz, dorsais de 1500 a 2500 Hz e alveolares entre 2500 e 4000 Hz, podendo haver, no caso das plosivas velares, mais de uma barra vertical e, ainda, no caso do [k], ruído após a soltura;
- o VOT (*Voice Onset Time*), correspondente ao intervalo (em milissegundos) entre a soltura da plosão e o início do vozeamento da vogal seguinte (no caso das plosivas desvozeadas) ou da sonorização do próprio segmento (no caso das plosivas vozeadas), como exibido na Figura 10 a seguir. É um recurso utilizado para diferenciar as plosivas surdas de suas contrapartes sonoras (tendo as últimas o VOT negativo).

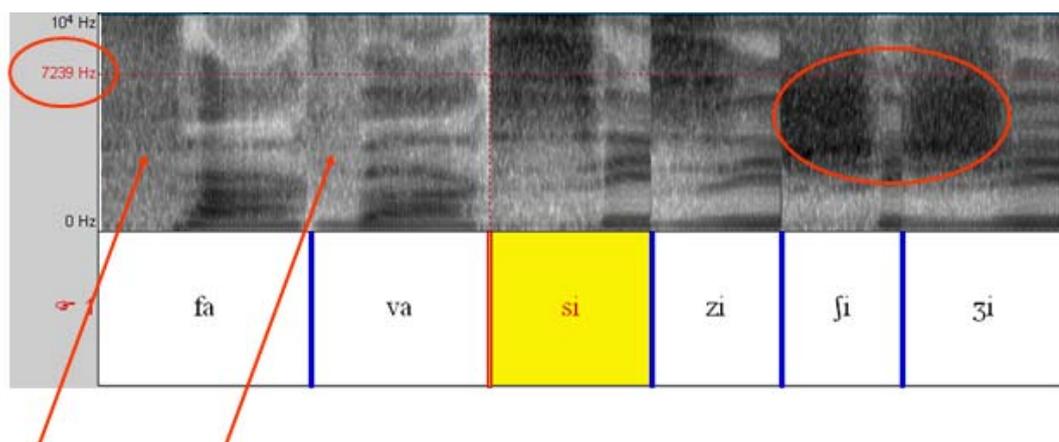
Figura 10 – VOTs tipicamente observados na enunciação de plosivas



Fonte: A autora (2013).

As fricativas são produzidas com constrição que gera um ruído turbulento, considerado a fonte de energia típica dessa classe de segmentos (embora não necessariamente a única, dada a existência de fricativas sonoras). Apresentam como referência acústica a zona de ruído espectral em altas frequências, essa mais longa nos exemplares desvozeados ou posteriores e com valor de frequência associado ao tamanho da cavidade anterior ao local da articulação (acima dos 4000 Hz em [s, z] e ao redor dos 3000 Hz em [ʃ, ʒ], sendo o espectro das labiodentais [f, v] fraco, plano e difuso). A imagem espectrográfica exibida na Figura 11, a seguir, ilustra as propriedades mencionadas.

Figura 11 – Localização dos vestígios tipicamente associados às fricativas [f, v, s, z, ʃ, ʒ]



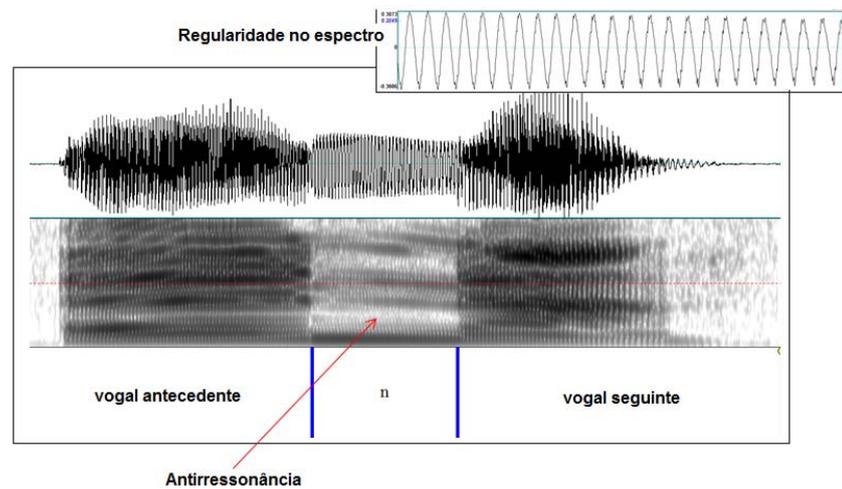
zona de ruído espectral difusa

Fonte: A autora (2013).

As nasais têm o vozeamento (ativação das pregas vocais) como fonte de energia e a presença de antiformantes. São achados acústicos típicos dessa classe de segmentos:

- a) o “murmúrio nasal”¹⁸, equivalente à energia acústica periódica (conforme detalhe sobreposto, na Figura 12), de baixa intensidade, com formantes fracos e elevada largura de banda em decorrência do amortecimento gerado pela rápida absorção de intensidade nas paredes da cavidade nasal (identificada como “Antirressonância” na Figura 12);

Figura 12 – Regularidade oscilográfica e antirressonância típicas das nasais



Fonte: A autora (2013).

- b) o “formante nasal”, equivalente ao primeiro formante, que se apresenta tipicamente baixo, entre 200 e 300 Hz;
- c) os vales nos espectros FFTs e LPCs que variam conforme o ponto de articulação da nasal (quanto mais posterior o ponto articulatório da nasal, maior o valor da frequência dos antiformantes).

Finalmente, as líquidas, com características acústicas gerais semelhantes às das nasais, destacam-se por exibirem no sinal sonoro:

- a) primeiro formante baixo, como é típico nos segmentos soantes;
- b) porção estável (*steady-state portion*);
- c) antiformantes em frequência bem superior (em torno de 2000 Hz) ao encontrado no caso das nasais.

¹⁸ Conforme Kent e Read (1992, p.130, tradução nossa), o murmúrio “é o segmento acústico associado à radiação exclusivamente nasal do som”.

1.2.3 O nível suprasegmental da fala: prosódia e tempo (taxa) de fala

Nesta subseção será discutido o fato da fala poder ser analisada ao nível do segmento ou acima desse (nível suprasegmental), o último referido na literatura como tendo associado traços suprasegmentais ou prosódicos. Por extensão, será apresentada a concepção fonética e fonológica de prosódia e o que fundamenta o posicionamento assumido neste estudo, de disposição do tempo de fala (taxa), em concordância com Laver (1994), em organização à parte da prosódica.

Como recém-mencionado, a análise linguística da fala pode ater-se aos segmentos, que em sucessão a compõem, ou pode extrapolar o nível desses, observando unidades maiores, comumente referidas na literatura como pertencentes ao nível suprasegmental. O nível suprasegmental de descrição da fala considera, segundo Laver (1994), todos os fatores que podem ser prolongados acima do domínio do segmento. Tal concepção é concorde com a apresentada por Nolan (1983, p.122, tradução nossa):

A cadeia suprasegmental da fala envolve [...] um domínio potencialmente maior que o do segmento, sendo tomada, tradicionalmente, como a modulação das dimensões fisicamente mensuráveis de f₀, amplitude e duração.

Uma série de autores, entre eles Kent e Read (1992) e McQueen e Cutler (1997), rotulam de prosódia as variações dos mencionados parâmetros físicos (f₀, amplitude e duração), assim como dos respectivos correlatos perceptivos (*pitch*, *loudness*¹⁹ e alongamento), ocorridas ao longo do enunciado.

O vocábulo “prosódia” vem do grego e tem por significado “melodia que acompanha o discurso”. Fujisaki (1997) se refere à prosódia como a organização sistemática de várias unidades linguísticas em um enunciado ou um grupo coerente de enunciados no processo de produção de fala.

O papel da prosódia no discurso é o de salientar ou diminuir o valor de algo no texto (CAGLIARI, 1997), exercendo, os elementos ou traços prosódicos, funções linguísticas e não linguísticas. A dinâmica da fala tem em si propriedades prosódicas que, de maneira geral, objetivam facilitar a compreensão do material linguístico

¹⁹ *Pitch* e *loudness* são medidas subjetivas relacionadas, respectivamente, à sensação psicoacústica de frequência e de intensidade.

veiculado. Tais propriedades são intrincadas e relacionais, mostrando-se difíceis de serem isoladas em seus mecanismos fisiológicos, uma vez que se estabelecem sempre em relação a outro elemento (ou seja, comparativamente).

A partir do levantamento bibliográfico realizado foi possível observar que o termo “prosódia” é utilizado com diferentes acepções. Cutler et al. (1997) concordam que o termo prosódia é de fato usado diferentemente pelos pesquisadores. Segundos os autores, em um extremo estão aqueles que mantêm uma definição abstrata (fonológica) de prosódia, tendo-a como a estrutura que organiza os sons, e no outro, aqueles que se referem a ela como a própria realização (perspectiva fonética de prosódia), havendo ainda os que, por conveniência, conjugam componentes de ambas as concepções.

Nooteboom (1997, p.640, tradução nossa) promove um acerto nessa caracterização quando expõe:

Na fonética moderna a palavra “prosódia” e a sua forma adjetival “prosódico” é frequentemente utilizada para referir aquelas propriedades que não podem ser derivadas das sequências segmentais dos fonemas subjacentes aos enunciados humanos. [...] Na Fonologia Gerativa Moderna, a palavra “prosódia” tem um significado diferente, sendo referida como um aspecto não segmental da estrutura linguística abstrata.

Na perspectiva fonética os traços prosódicos têm função exclusivamente descritiva e explanatória (LADEFOGED, 2001), enquanto que na perspectiva fonológica esses, isoladamente ou de maneira combinada (compondo um traço prosódico derivado), assumem função contrastiva, de oposição de significado (FIRTH, 1951; NESPOR e VOGEL, 1986).

Para Mira-Mateus (2004), embora os traços prosódicos possuam inicialmente um caráter fonético, podem ser utilizados fonologicamente para marcar os limites das unidades (por exemplo, o início ou o fim de uma palavra), além de para opor significados lexicais (por exemplo, em línguas onde o tom é distintivo) e frasais (diferenciando, por exemplo, frases interrogativas e afirmativas).

Foram também encontradas divergências quanto à quais seriam exatamente os traços prosódicos, o que fica visível ao serem confrontadas, por exemplo, as definições de prosódia apresentadas por Crystal (1985, p.249, tradução nossa), a saber, “termo usado na fonética suprasegmental e fonologia para referir

coletivamente as variações no *pitch* e na *loudness*, no tempo e no ritmo” (visão que mistura elementos perceptivos, de organização temporal e fonológico), e por Ladefoged (2001, p.276, tradução nossa), a saber, “são traços fonéticos como acento, duração, tom e entonação, não sendo essas propriedades de consoantes ou vogais isoladas e sim de unidades maiores”.

Considerando a área de aplicação deste estudo, Fonética Forense, Watt (2010), descrevendo os traços normalmente observados nas análises da Comparação de Locutor, elenca como características prosódicas a entonação, o ritmo, o tempo e a qualidade vocal.

Já Laver (1994) refere como sendo traços suprasegmentais os apontados por Crystal (1985), acrescentando, contudo, os padrões de variação relativos à continuidade da fala²⁰. Clark et al. (2007) pactua com Ladefoged (2001) ao admitir como suprasegmentos o acento, o tom e a entonação, abstendo-se, no entanto, no que se refere à duração.

A despeito da não concordância exata acerca de quais seriam os traços prosódicos ou suprasegmentais, Fletcher (2010, p.523, tradução nossa) afirma que tais traços são, “para muitos foneticistas e cientistas da fala, sinônimo de variação em parâmetros como f_0 , intensidade e duração”, derivando da combinação desses os demais fenômenos suprasegmentais, como o acento, a entonação e o ritmo.

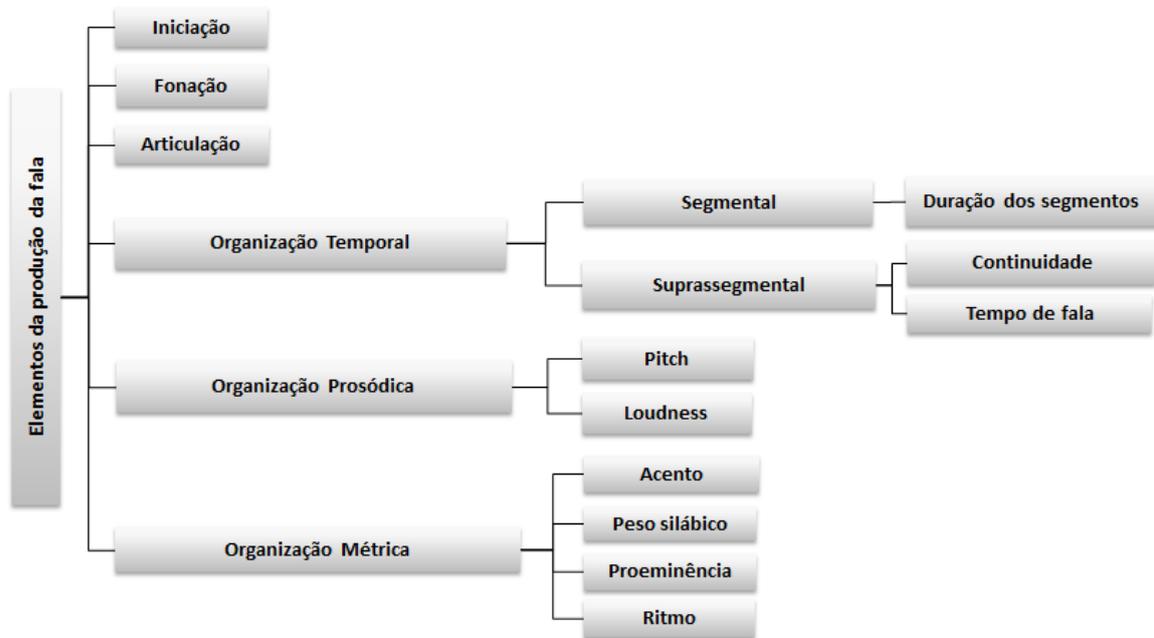
Quanto aos demais traços prosódicos mencionados, encontra-se na literatura referência à taxa de articulação e à qualidade vocal (a última representada, entre outros, pelo *pitch*), como sendo (CLARK et al., 2007, p.327, tradução nossa) “outras configurações e ajustes de longo termo”, e ao tempo de fala e à continuidade como sendo (LAVÉR, 1994) fatores não linguísticos, utilizados para sinalizar informações paralinguísticas (relacionadas ao estado atitudinal e emocional do locutor) e extralinguísticas (indicativas da identidade e da personalidade do locutor).

Laver (1994) propõe a análise da produção da fala através de três módulos organizacionais: o de organização temporal (módulo que abrange a duração segmental, a continuidade e a taxa ou tempo de fala), o de organização prosódica (restrita ao *pitch* e à *loudness*) e o de organização métrica (constando nesse último grupo o acento, o peso silábico, a proeminência e o ritmo). A partir da concepção de Laver (1994, Parte VI, p.431 a 546) acerca do contexto de inserção do tempo de fala

²⁰Segundo Laver (1994), a continuidade relaciona-se à incidência das pausas na cadeia da fala.

no conjunto de elementos constituintes da produção da fala, formulou-se o esquema apresentado na Figura 13, abaixo.

Figura 13 – Disposição do tempo de fala em relação aos demais elementos da produção da fala



Fonte: Elaborado a partir de LAYER (1994, p.431-546).

Nota: Conforme o autor, “Iniciação” refere-se ao mecanismo respiratório, “Fonação”, à atividade laríngea e “Articulação”, às mudanças rápidas nas ações dos órgãos do trato vocal para criação dos padrões fonéticos associados aos segmentos linguísticos.

Segundo o autor (LAYER, 1994), o tempo de fala compreende um fenômeno do nível suprasegmental da produção da fala, atinente ao módulo de organização temporal (conforme indica a Figura 13 recém-apresentada). No presente estudo, reitera-se, assumir-se-á tal concepção, a despeito da revisão acerca do termo prosódica (justificada pelo fato desse ser utilizado, não raro, como sinônimo de suprasegmento), bem como de suas funções e *status*. Entende-se como oportuna e mais apropriada, considerando-se que as taxas temporais aqui estudadas constam como variáveis dependentes e não como intervenientes, a adoção de um referencial teórico onde o tempo de fala é disposto em uma organização didaticamente apartada da prosódica.

1.3 SOCIOFONÉTICA

Com o evoluir dos recursos computacionais e tecnológicos e a incorporação da utilização desses nos métodos de pesquisa científica tem-se a Sociofonética como um recente (e promissor) campo de investigação sociolinguística.

A Sociolinguística é o ramo da linguística que estuda a relação entre língua e sociedade, a associação que há entre manifestações linguísticas e atitudes sociais. A Sociolinguística Variacionista ou Teoria da Variação (LABOV, 1966, 1972, 1982, 1994) compreende a perspectiva teórica na qual a análise da língua “envolve a adoção de um conjunto de fatores não apenas linguísticos, mas também extralinguísticos como os de natureza social e estilística” (CAMACHO, 2010, p.149). Nela, observa-se a língua “tal como usada na vida diária por membros da ordem social” (LABOV, 2008 [1972], p.13, tradução nossa), sendo o comportamento linguístico verificado nas comunidades onde a fala efetivamente transcorre, com o intento de desvendar o sistema dentro da variação.

Na assunção sociolinguística assume-se a língua como um fenômeno social, atendo-se a empregos linguísticos concretos e a questões como a identidade linguística de grupos sociais, as atitudes sociais relativas à linguagem, as formas padrão e não padrão de linguagem, as variedades e níveis de linguagem, entre outros (CRYSTAL, 1985).

Alguns dos preceitos desenvolvidos sob a concepção sociovariacionista de linguagem são de especial interesse para o presente estudo, entre eles, o que promulga a interferência do observador e de variações estilísticas nas manifestações de fala.

A Sociofonética representa a interface entre a Sociolinguística e a Fonética Acústica, excluindo a tradicional transcrição fonética impressionística (THOMAS, 2011). De acordo com os preceitos variacionistas, essa abordagem observa o falante no contexto social de uso da língua e nela a variação linguística, a despeito de motivações biológicas, é vista como socialmente estruturada. Nessa perspectiva, a coleta de dados dá-se *in locuo*, constituindo um material não controlado, que se opõe às falas produzidas em laboratório ou obtidas a partir de leitura oral.

Intrinsecamente empírica, a Sociofonética é definida por Trudgill (2003) como sendo o estudo sociolinguístico de traços fonéticos e/ou o uso de técnicas fonéticas

na condução de trabalhos sociolinguísticos. Definição mais ampla, no entanto, é encontrada em Swann et al. (2004, p.288, tradução nossa):

A Sociofonética envolve a aplicação da fonética aos estudos sociolinguísticos. Por exemplo, uma pesquisa pode valer-se da mensuração acústica dos segmentos da fala para investigar aspectos da variação regional, da variação social, da variação estilística ou de mudança linguística.

Foulkes e Docherty (2006) argumentam que apesar dessas serem definições atuais, são limitadas e carentes de adequação. Os autores formularam, então, uma proposta na qual a Sociofonética atem-se ao estudo sociolinguístico de traços fonéticos e fonológicos e/ou ao uso de técnicas relativas a esses domínios com o propósito de identificar e descrever as fontes de variação, assim como os lugares e os parâmetros da variação socialmente estruturada.

Tal noção fundamenta-se na premissa de que “os falantes ajustam-se aos contextos [sociais] através do ajuste de suas produções orais” (THOMAS, 2011, p.2, tradução nossa), ou seja, a língua é inerentemente instável, de forma a permitir que o falante, frente às inúmeras situações de interação comunicativa a que se expõe, constantemente adapte-se e acomode-se. Foulkes e Docherty (2006) ressaltam, no entanto, que a variabilidade na fala não é aleatória, tampouco caótica, mas sim determinada por um número específico de fontes de variação, que conduzem à realização de padrões sistemáticos.

Para Foulkes et al. (2010, p.704, tradução nossa), a Sociofonética refere-se à “como a variação socialmente estruturada [...] é aprendida, armazenada cognitivamente e processada em atos de compreensão e expressão linguística”. Para os autores, a Sociofonética subsidia e desenvolve uma gama de disciplinas aplicadas, entre elas, a de terapia de fala, a de educação, a de tecnologia da fala e a de ciências da fala em âmbito forense.

A aplicação forense da Sociofonética caracteriza-se, segundo Foulkes e Docherty (2006), pela descrição de parâmetros de variação individuais ou de grupo, assim como pela identificação das fontes adicionais de variação, entre elas, o meio de transmissão sonora e as influências externas, como estresse e drogas. Nesse sentido, Foulkes et al. (2010, p.737, tradução nossa) afirmam que a “[...]”

sociofonética tem um papel central no crescimento da área de Fonética Forense, pois a compreensão da variação intrassujeito e intersujeitos é essencial para a Comparação de Locutor [...].”

Nolan (2001, p.17, tradução nossa) apoia a aplicação do conhecimento sociovariacionista e do instrumental fonético-acústico no trabalho de detecção de formas próprias de falar, alvo da investigação forense. Segundo o autor:

[...] a linguagem e a fala formam um imenso e complexo sistema plástico, e o entendimento dos meios através dos quais a identidade do falante mostra-se na fala requer uma consistente fundamentação ao menos em linguística, dialetologia, sociolinguística, fonética e acústica.

Assim, entende-se que o presente estudo vai ao encontro de premissas assumidas pelo modelo sociofonético pois, embora conte com um número restrito de participantes (sendo as razões devidamente expostas no Capítulo 3), compreende uma investigação foneticamente orientada, que contempla a fala maximamente naturalística (referência à gravação de diálogo feita sem a consciência dos locutores nele participantes) e as variáveis sociais normalmente investigadas em pesquisas sociovariacionistas.

Nesse sentido, Foulkes et al. (2010) afirmam que tanto o número de sujeitos considerados quanto a quantidade e extensão dos dados coletados decorrem, em um estudo sociofonético, das perguntas que se pretende responder, ou seja, dos objetivos da investigação. Os autores referem que os estudos que se debruçam amiúde sobre um fenômeno de língua muito específico, utilizando uma análise custosa e complexa, inevitavelmente têm de admitir um número reduzido de sujeitos.

Di Paolo e Yaeger-Dror (2011), discutindo a questão do número de sujeitos da amostra de estudos sociolinguísticos, ressaltam apenas que os trabalhos com um número pequeno de sujeitos, como o em tela, devem contemplar um grande número de *tokens*²¹ por falante, para que seja possível atender aos requisitos dos testes estatísticos.

Especificamente quanto à variação dos suprasegmentos, conforme Foulkes et al. (2010), essa tem sido menos estudada à luz da Sociofonética em razão da dificuldade de se estabelecer a equivalência funcional entre as formas linguísticas

²¹ O termo “*token*” refere-se a uma das realizações de uma palavra em uma amostra de fala.

alternativas em materiais não controlados de fala. Algumas das variáveis passíveis de estudo mostram-se de complexa análise na fala naturalística. Os autores citam, a exemplo, a capacidade de interferência do tempo de fala empregado durante a interação espontânea, entre outros, no VOT.

Thomas (2011, p.185, tradução nossa), que classifica o tempo de fala como um “fator global” da prosódia, refere a subexploração da variação prosódica quando afirma: “além do ritmo prosódico e da entonação, alguns outros poucos tópicos, como a taxa de fala e a análise das pausas, tem sido esporadicamente estudados”.

Segundo Kendall (2009), algumas pesquisas já perscrutam as taxas temporais de fala como um fenômeno sociolinguístico, tendo alcançado resultados que indicam que esse parâmetro pode ser uma profícua área para os estudos variacionistas, por exemplo, no que se refere ao comportamento dos tempos de fala em diferentes dialetos. A exemplo, podem-se citar estudos sobre diferenças dialetais nos quais é verificado o tempo de fala espontaneamente manifesto:

- a) Byrd (1994), que observou os processos de redução e a TE associada em oito dialetos do inglês americano;
- b) Quené (2008), que aprofundou e sofisticou a análise da variabilidade intra e interindivíduo, analisando a duração média das sílabas no holandês falado na Holanda e na Bélgica;
- c) Robb et al. (2004), que investigaram a TE e a TA nas variedades americana e neozelandesa do inglês;
- d) Verhoeven et al. (2004), que observaram o comportamento da TE e da TA em quatro dialetos do holandês da Holanda e em quatro dialetos do holandês falado na Bélgica;
- e) Hewlett e Rendall (1998), que contrapuseram a TE e a TA obtida a partir do inglês urbano de Edimburgo (Escócia) com a obtida a partir do inglês rural de Orkney (Escócia);
- f) Jacewicz et al. (2009, 2010), que enfatizaram a análise da variabilidade intra e interindividual da TA no inglês de Wisconsin (EUA) e da Carolina do Norte (EUA);
- g) Syrdal (1996), que observou a TE em oito dialetos do inglês americano, tendo por base a localização geográfica dos dez primeiros anos de vida dos falantes.

Ressalta-se, contudo, que no presente estudo a investigação de questões dialetais (variações regionais) é prejudicada em razão da impossibilidade de acesso a histórico sociolinguístico confiável dos locutores cujas falas são analisadas, uma vez que muitos deles possuem exposição linguística peculiar, decorrente de eventual(is) temporada(s) de privação de liberdade.

Ademais, embora os resultados dos trabalhos de Ramsaran (1978) e Siptár (1991), respectivamente sobre o inglês e o húngaro, apontem a existência de independência entre o grau de formalidade do discurso e a TE empregada pelos sujeitos pesquisados, Thomas (2011) destaca como evidente a interferência de variações estilísticas no tempo de fala, razão pela qual o tema é nas próximas seções abordado.

1.3.1 A entrevista sociolinguística

A entrevista sociolinguística é basicamente uma sequência dialógica, uma vez que envolve ao menos dois locutores, que se alternam na produção de turnos conversacionais. Esse tipo de entrevista é projetado para angariar grande quantidade de dados de falantes em condição tão casual e natural quanto possível (embora apresente certa estruturação), obtendo do informante a produção oral na forma vernacular, definida por Labov (2008 [1972], p. 244, tradução nossa) como sendo aquela em que “o mínimo de atenção é dada ao monitoramento da fala”.

Em uma entrevista sociolinguística invariavelmente tem-se tanto a fala monitorada, que conforme Labov (2008 [1972], p.63, tradução nossa) compreende “uma fala [...] controlada em resposta à presença de um observador externo”, quanto a fala espontânea, equivalente à fala casual, contudo, de ocorrência reservada aos contextos formais, a partir do momento em que se têm superado os constrangimentos próprios do formalismo.

Neste estudo assume-se que a fala obtida através da gravação desavisada (interceptação telefônica feita sem a ciência dos locutores) é prevalentemente casual. Já a gravação avisada, provinda da coleta técnica de padrão vocal para fins de perícia, é admitida (em razão de compreender uma entrevista semidirigida, contexto enquadrado por Labov, 1972, como formal, e de se ter exercitado os dois ou

três primeiros minutos iniciais da conversação) como sendo prevalentemente espontânea²².

Segundo Görski (2011, p.9), no tipo de entrevista sociolinguística as perguntas do entrevistador “funcionam como gatilhos para o surgimento de diferentes sequências textuais nas respostas do entrevistado”. Durante o procedimento de entrevista incentivam-se as narrativas de experiência pessoal, gênero discursivo em que os informantes, por estarem emocionalmente envolvidos com a elaboração do relato, dispõem menos atenção à fala (TARALLO, 1986).

Llisterri (1992), em levantamento feito a partir dos trabalhos apresentados no *workshop* sobre “Speaking styles in speech research”, observa que as pesquisas que integraram o evento consideram materiais basicamente produzidos em situação de leitura oral, na realização de tarefas e em situações menos pré-definidas (como entrevistas dirigidas ou semidirigidas), abrangendo tanto materiais produzidos a partir de registro ambiental clássico (*in locuo*) como em ambiente laboratorial. As entrevistas utilizadas para colheita de dados são normalmente monólogos curtos elaborados em resposta a perguntas genéricas do entrevistador que se esforça para interferir o menos possível nos relatos, fazendo-o somente quando há exaustão do tópico em curso, a fim de formular novo questionamento.

A despeito de ser rara a interferência do interlocutor, Llisterri (1992) destaca que a entrevista, mesmo semidirigida, não corresponde à conversação, pois nela há violação, ao menos, das tomadas espontâneas de turno. Segundo o autor, o mais próximo de uma conversação natural seria o diálogo mantido entre o falante em observação e alguém conhecido, contexto encontrado em um dos tipos de gravação considerados na pesquisa aqui desenvolvida (na gravação desavisada).

Neste estudo são extraídos tempos de fala (TEs e TAs) a partir de dois tipos de gravação (e de situações comunicativas) distintos, um deles, recém-mencionado, referente a interceptação telefônica e o outro, à entrevista semidirigida. Conduzida por peritos em registros de áudio, a gravação da entrevista com fins de confronto de voz e fala não difere, em essência, da descrita entrevista sociolinguística. Preocupam-se os peritos, da mesma forma, em extrair do falante a fala mais espontânea (menos monitorada) possível, propondo-lhe tópicos de conversação emocionalmente fortes, como os relacionados ao seu histórico familiar, experiência

²² Ao longo deste texto sempre que for irrelevante a diferenciação entre fala casual e espontânea será utilizada, genericamente, a expressão “fala espontânea”.

pessoal pregressa, envolvimento no crime, condição carcerária a que está exposto (se recluso em unidade prisional), perspectiva de vida futura, entre outros.

1.3.2 Variações estilísticas

Variações estilísticas referem-se às “alternâncias através das quais um falante adapta sua linguagem ao contexto imediato do ato de fala” (LABOV, 2008 [1972], p. 313, tradução nossa). Conforme o autor, na pesquisa de campo destacam-se cinco princípios fundamentais relativos ao(s) estilo(s) de fala, a saber:

- a) alternância de estilo – o falante altera o estilo de sua fala durante a produção oral, ou seja, não há falante de estilo único;
- b) atenção – os estilos de fala variam em um *continuum* que se estende da fala casual (vernacular), “fala cotidiana usada em situações informais” (LABOV, 2008 [1972], p.111, tradução nossa), à fala padrão (monitorada), sendo esse escalonado pelo grau de atenção dispensado pelo falante à própria fala, o que é associado à formalidade da situação comunicativa;
- c) vernáculo – o vernáculo é um estilo de fala que se destaca pela sistematicidade, sendo considerado o mais regular quanto à estrutura e na relação com a evolução da língua;
- d) formalidade – a observação sistemática do falante é um contexto formal, onde mais que um mínimo de atenção é dispensado à fala;
- e) bons dados – a entrevista individual é uma observação sistemática óbvia, que garante a obtenção de dados de fala em quantidade suficiente.

Os princípios elencados conduzem a um importante paradoxo metodológico (intitulado “Paradoxo do Observador”), segundo o qual, para obter-se o dado mais importante para a teoria linguística (provindo da fala casual), deve-se analisar como as pessoas falam quando não estão sendo observadas. Nesse caso, tem-se um problema aparentemente insolúvel: como sistematicamente observar os falantes em momento de interação comunicativa natural sem a introdução da formalidade tipicamente associada à presença do entrevistador?

No que se refere à caracterização dos estilos de fala, Labov (1970) afirma que existem mais estilos e dimensões estilísticas do que um analista pode constatar,

sendo possível, no entanto, avaliar o estilo de fala em razão do grau de atenção prestado à mesma.

O problema na investigação da variação estilística em Sociolinguística é, segundo Labov (1972), controlar adequadamente os contextos e definir os estilos que ocorrem em cada um deles. O autor aponta a existência de ao menos cinco níveis estilísticos que afetam diferentemente a fala, a saber: fala casual, fala monitorada, leitura oral, lista de palavras e pares mínimos.

Foulkes et al. (2010) entendem que a classificação laboviana de níveis estilísticos (baseada no grau de atenção dispensado à fala) é um tanto simplista, pois os falantes fazem ajustes fonéticos não só em razão do automonitoramento, mas também em decorrência de fatores externos, como condição física, tópico conversacional e audiência (os dois últimos também referidos por Bell, 1984).

Schilling-Estes (2002) aponta como limitações da concepção de variação estilística baseada na atenção dispensada à fala as dificuldades encontradas quando se tenta separar, em um registro de entrevista, a fala casual da monitorada e as relativas à quantificação do nível de atenção prestado à fala, assim como o fato de que há estilos que não se ajustam à proposta de condicionamento em razão da formalidade e/ou da atenção. A autora ressalva ainda que, sendo admitida a existência de variações estilísticas intencionais, o nível de atenção dispensado à fala não estaria rigorosamente associado à formalidade da situação comunicativa.

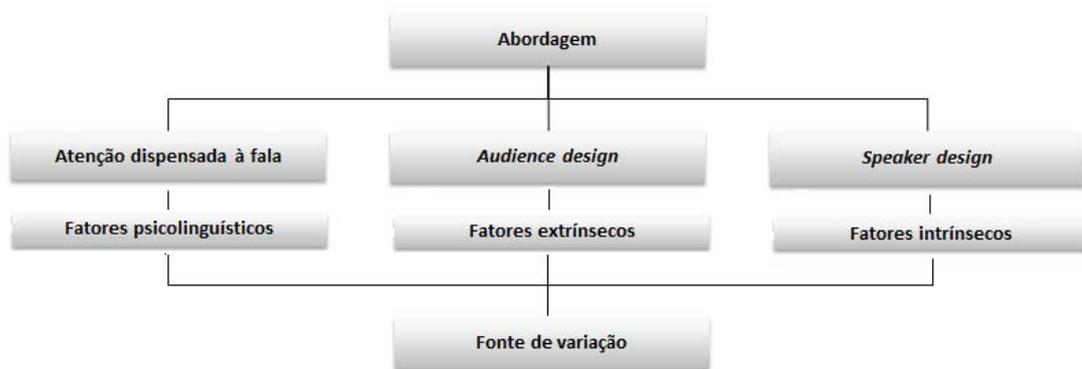
A alternância de estilo é discutida por Schilling-Estes (2002, 2004), que a considera uma variação de fala de ordem intrassujeito, passível de acometer a forma de uso de língua de um grupo (dialeto)²³, a forma prevista para uma situação de uso de língua em particular (registro) ou um gênero linguístico específico²⁴. Nessa perspectiva, os estilos de fala variam tanto na fala de um mesmo indivíduo (variação intrassujeito) quanto entre grupos de falantes (variação intersujeitos). Para a autora, o falante pode incorrer na alteração do estilo durante a emissão oral tanto de forma voluntária quanto inconsciente, por período temporalmente curto ou extenso e em uma escala que vai do estilo conversacional ao estilo formal.

²³ Tal acometimento pode se dar em qualquer um dos níveis de organização da língua (fonológico, morfossintático, lexical, semântico, pragmático ou discursivo).

²⁴ Gêneros discursivos são definidos pela autora como “variedades rotineiras, altamente ritualizadas, usualmente associadas com algum tipo de *performance* ou apresentação artística” (SCHILLING-ESTES, 2002, p.375, tradução nossa).

A variação estilística tem sido caracterizada a partir de três abordagens sociolinguísticas principais: a primeira, já mencionada, concernente ao grau de atenção dispensado à fala (LABOV, 1972, 2001); a segunda, de cunho interacional, que privilegia o papel da audiência e considera as relações interpessoais envolvidas (intitulada “*audience design*”, conforme Bell, 1984, 2001) e a terceira, referente à questão da identidade social (intitulada “*speaker design*”, conforme Coupland, 1996), paradigmas ilustrados na Figura 14.

Figura 14 – Variação estilística: principais abordagens e respectivas fontes de variação



Fonte: A autora (2013).

Através do esquema apresentado na figura percebe-se que cada abordagem de variação estilística tem associada uma fonte de variação particular, de natureza, psicolinguística (relativos especialmente à função cognitiva atenção), extrínseca (entre eles, a audiência, o tópico conversacional e o contexto situacional) ou intrínseca de outra ordem (entre eles, a intenção comunicativa).

A diferenciação entre as duas primeiras abordagens (a baseada na atenção dispensada à fala e a projetada na audiência) decorre essencialmente das mencionadas motivações. Enquanto para Labov (1972) a variação no estilo é resultado da alteração no grau de automonitoramento da fala, para Bell (1984) ela refere-se à adequação do estilo de fala ao interlocutor, visando o falante a aproximação ou o distanciamento em relação aos membros da audiência.

Bell (1984, 2001) atribui as variações estilísticas a adaptações do locutor à composição e características da audiência. Segundo o autor, os falantes ajustam sua fala na intenção de expressar solidariedade (visando parecerça com a(s) pessoa(s) com quem está falando) e intimidade em relação ao interlocutor (presente ou não), assim como, sendo a intenção, o distanciamento. Em tal proposta, aplicável

não só à situação de entrevista, mas também a outras formas de interação conversacional (por exemplo, à conversa naturalística entre pares e colegas), o impacto dos membros constituintes da audiência sobre o resultado da fala é proporcional ao nível de consciência que o falante tem da presença desses no contexto da fala.

Já entre as duas primeiras abordagens e a terceira (projetada no falante) há, conforme Schilling-Estes (2002), uma mudança conceitual envolvendo o papel do falante na variação estilística em curso (admitido como passivo nas duas primeiras²⁵ e como ativo na última) e o fato daquelas serem unidimensionais (por considerarem basicamente, nessa ordem, ou o grau de atenção dispensado à fala ou a composição/ propriedades da audiência) enquanto a última é multidimensional (contemplando um maior número de fatores como possíveis fontes de variação, entre eles, os relacionados aos níveis de organização da língua, ao tópico conversacional, ao contexto situacional, ao humor, ao meio de transmissão da linguagem, além de fatores paralinguísticos e não linguísticos). Para a autora há dois pontos principais na abordagem projetada no falante:

(1) os falantes não alternam o estilo somente, ou principalmente, como reação a elementos da situação de fala (quer seja a formalidade ou a audiência); antes disso, são ativos e criativos nos usos de recursos estilísticos [...] e (2) [os falantes] usam suas falas para moldar e remoldar as situações externas (em contexto interacional imediato ou interações sociais mais amplas), os relacionamentos interpessoais e, especialmente, suas identidades pessoais (SCHILLING-ESTES, 2002, p.378, tradução nossa).

Na abordagem projetada no falante destaca-se a questão da intencionalidade (“agency”) e da prática social (ECKERT, 2000). Tal perspectiva, mais condizente com o proposto no presente estudo, prevê que a variação estilística não é determinada pela reatividade ao meio externo e sim provocada pelo próprio falante, não só para marcar afiliação a grupo social específico ou o atributo social desejado, mas também para significar diferentes propósitos conversacionais na interação em curso (SCHILLING-ESTES, 2002).

Com o exposto, depreende-se que os estudos variacionistas, no que se refere ao entendimento acerca das variações no estilo de fala, evoluíram da ideia de

²⁵ Embora Bell (2001) defenda que a responsividade ao ambiente não é uma ação passiva.

adaptação de variáveis individuais à formalidade da situação comunicativa à noção de uso de modos distintos de fala, vinculados à identidade social e à intencionalidade do discurso (CAMACHO, 2010).

Neste estudo aborda-se a questão dos estilos de fala em razão de serem consideradas gravações realizadas em situações comunicativas distintas, a saber: conversação ao telefone celular entre um locutor-alvo e algum indivíduo conhecido, sem que ambos tivessem, à ocasião, ciência de que estavam sendo gravados, e diálogo presencial com interlocutor até então desconhecido do locutor-alvo, que transcorre com o consentimento dos presentes, em ambiente onde constam visíveis os equipamentos utilizados na gravação.

A caracterização recém-feita permite a percepção de que há graus de formalidade particulares e audiência própria para cada um dos tipos de gravação considerados na presente pesquisa. Ressalta-se, ainda, que não se objetivou a verificação dos estilos empregados em um e outro tipo de gravação, no entanto, a exploração quantitativa perpetrada disponibilizará informações que contribuirão para a avaliação da procedência da difundida concepção (apontada, entre outros por Abaurre-Gnerre, 1981) de que a fala classificada por Labov (1972) como casual (encontrada na gravação desavisada, provinda de interceptação telefônica feita sem a ciência dos locutores) é incondicionalmente rápida e de que a fala por ele classificada como espontânea (encontrada na gravação avisada, obtida em colheita técnica de padrão vocal para fins de perícia) é compulsoriamente lenta.

No delineamento de pesquisa da Comparação Forense de Locutor é prevalente o confronto de amostras de fala produzidas em situações comunicativas distintas, sendo comuns em Laudos Periciais ressalvas acerca da possível interferência, entre outros, das propriedades do contexto situacional e do estado emocional do locutor em relação ao tópico de conversação desenvolvido. Assim, entende-se como mais adequada à prática forense a proposta de Schilling-Estes (2007, p.984)²⁶, concernente à intencionalidade e à multidimensionalidade, do que as propostas baseadas na atenção dispensada à fala e projetadas na audiência, que apregoam a reatividade e a unidimensionalidade.

²⁶ Disponível em: <<http://www.um.es/lacell/aeisla/contenido/pdf/9/schilling.pdf>>. Acesso em: 12 nov. 2012.

1.3.3 Comunidade de fala, redes sociais e comunidade de prática

Neste estudo é considerado um grupo particular de falantes: indivíduos por alguma razão associados à prática delituosa, especificamente a relacionada ao tráfico de entorpecente e/ou a (tentativa de) homicídio(s), apresentando todos eles histórico de aprisionamento, seja em unidade de reabilitação socioeducativa²⁷ ou em unidade prisional convencional, em razão de cumprimento de pena ou de prisão preventiva.

Tais indivíduos, apesar de nunca terem interagido comunicativamente entre si, solidarizam-se no emprego de comportamentos linguísticos tipicamente associados ao contexto de execução de crimes, entre eles, o referente ao uso de expressões cujo significado é de entendimento restrito ao grupo de usuários de drogas e/ou traficantes e a quem os investiga ou pericia, a exemplo: “partir o cara” significando assassinar alguém, “tá baixado” significando estar foragido, “fazer uma caminhada” significando fazer algo determinado por um superior (por exemplo uma entrega de drogas ou uma execução), “escolher um bicho” significando escolher um veículo, que será furtado ou roubado para cometimento de outros delitos, “escama” significando maconha, expressões rotineiramente observadas em gravações de diálogos mantidos entre criminosos.

Ao se considerar o critério de definição de comunidade apontado por Saville-Troike (2003), qual seja, o de compartilhamento de alguma significativa dimensão da experiência, pode-se aceitar a proposição de que, mesmo advindos de facções ou grupos criminosos distintos, os traficantes e, por vezes, consequentes homicidas, em razão da função em exercício (que requer compreensão rápida e maximamente limitada aos membros do grupo), desenvolvem uma linguagem própria, adaptada ao contexto laboral.

Na perspectiva sociolinguística tem-se difundida a noção de comunidade de fala, a qual, segundo Saville-Troike (2003, p.15, tradução nossa), está relacionada “ao modo com que os membros do grupo usam, valoram e interpretam a linguagem”.

Por conceito (GUMPERZ apud ROMAINE, 1980, p.13, tradução nossa), a comunidade de fala compreende “um grupo de falantes que compartilham um

²⁷ Dois dos sete sujeitos participantes desta pesquisa eram menores de idade e encontravam-se detidos na Fundação de Atendimento Sócio-Educativo (FASE) do Estado do RS à época da gravação desavisada.

conjunto de normas e de regras referentes ao uso da linguagem”. Refere-se ao uso convergente de determinadas variáveis por uma dada comunidade, em razão de motivação social comum, essa considerada na organização dos significados linguísticos.

Para Romaine (1994, p.22, tradução nossa), uma comunidade de fala é:

Um grupo de pessoas que não necessariamente compartilham a mesma língua, mas que compartilham um conjunto de normas e regras a serem consideradas no emprego da linguagem. As fronteiras entre as comunidades de fala são essencialmente sociais e não linguísticas.

A composição das comunidades de fala considera fatores demográficos como gênero, idade, raça, etnia, classe social, local de trabalho (entre outros) ou observa uma população geograficamente definida (local de residência, região, etc). No entanto, há ao menos duas outras importantes abordagens de comunidades de falantes: as redes sociais e as comunidades de prática (MULLANY, 2007).

A abordagem de comunidade conhecida como redes sociais (MILROY, 1987) foca as relações sociais que falantes específicos mantêm entre si, examinando como tais relações afetam o comportamento linguístico desses locutores. A unidade de referência nessa abordagem é a força da rede, avaliada como densa ou leve (em decorrência da sucessão de contatos interativos entre os membros do grupo) e como única ou múltipla (em decorrência do número de ambientes de interação existente entre os pares).

No entanto, é a comunidade que se constitui a partir do exercício de uma atividade regular conjunta, a intitulada “comunidade de prática”, que melhor exprime a relação sociolinguística existente entre os participantes do presente estudo.

Segundo Eckert (2006, p.1)²⁸, a comunidade de prática é “[...] um grupo de pessoas que recorrentemente se envolvem em algum empreendimento comum”. Para a autora, as comunidades de prática formam-se em razão de interesses ou posições sociais comuns aos falantes que a compõem, desenvolvendo esses “jeitos de fazer coisas, visões, valores, relações de poder, jeitos de falar” próprios.

Nessa abordagem, a prática linguística reflete a identidade que o falante constrói enquanto membro do grupo, identidade essa que guia o comportamento

²⁸ Disponível em <<http://www.stanford.edu/~eckert/PDF/eckert2006.pdf>>. Acesso em: 12 nov. 2012.

linguístico manifesto e que faz com que seja reconhecido como um falante típico, como o falante participante desta pesquisa, cujo padrão de fala é associado, no âmbito da criminalística, como sendo, entre outros, do traficante e/ou do homicida, o que, para Caldeira (2000) corresponde à “fala do crime”.

Da Hora (2008), ao analisar a problemática social da droga, destaca a autenticidade e os modos peculiares de construção da realidade encontrados nesse contexto social e aponta, ao discutir as interações locais socialmente construídas, a linguagem como um dos elementos marcadores do grupo.

Feitas as considerações teóricas julgadas relevantes à compreensão do desenvolvimento deste trabalho, assim como da discussão encaminhada em item próprio, segue, no próximo capítulo, a revisão de literatura relativa aos aspectos temporais da fala; à variabilidade, potencial individualizante e aplicação forense das taxas em estudo e à relação dessas com os possíveis fatores condicionadores.

2 SOBRE O TEMPO DE FALA

Neste capítulo será apresentada uma revisão de literatura sobre o tempo de fala. Na Seção 2.1 aborda-se a conceituação, classificação e contextualização das medidas comumente utilizadas para representá-lo e a problemática que permeia a sua investigação. A Seção 2.2 contém informações referentes à variabilidade das taxas temporais estudadas (TE e TA), ao respectivo potencial individualizante (de indivíduo) e à aplicação forense de tais medidas.

2.1 ASPECTOS TEMPORAIS DA FALA

O uso do tempo na fala é interpretado por Nespor e Vogel (1986) como sendo um fator de desempenho, posição também assumida por Laver (1994) quando o classifica como um fator não linguístico que compreende a duração individual dos segmentos, a taxa ou tempo de fala e a continuidade da fala.

Os mencionados elementos temporais, contudo, tem seus conceitos utilizados diferentemente pelos autores que pesquisam o tema, o que conduz à constatação da existência de inúmeras divergências metodológicas quando do confronto dos trabalhos. As disparidades assumem tal proporção e tem tanto poder de afetar a comparação de resultados, que os textos produzidos tradicionalmente dedicam espaço em prólogo para discussão das controvérsias vigentes e para a exposição do posicionamento assumido quanto à categorização e quantificação do tempo de fala.

A problemática é vista, no entanto, como inerente à natureza do objeto, por ser ele um fenômeno dinâmico (logo, previsivelmente variável), que se manifesta na fala em uso (essa suscetível a adaptações às diferentes situações comunicativas) e que perpassa níveis (segmental e suprasegmental) e domínios gramaticais distintos da língua.

Quanto à terminologia, como já mencionado, normalmente é utilizada uma expressão genérica, como forma indistinta de taxa de produção de fala (neste estudo, “tempo de fala”²⁹), quando se quer evitar a especificação de um tipo particular de mensuração quantitativa (DANCOVIČOVÁ, 1997; KENDALL, 2009).

²⁹ Em inglês *rate of speaking*, *rate of speech* ou *speech tempo*.

A Figura 15, apresentada a seguir, contém uma exposição terminológica que exprime a concepção prevalente acerca das duas medidas temporais de fala mais difundidas, a TE e TA, com a ilustração dos respectivos pontos controversos. Na coluna da esquerda, no caso da “Taxa de elocução”, é feita menção às formas utilizadas no inglês. As referências bibliográficas apontadas na figura dão a procedência de determinado aspecto do conceito ou denotam o posicionamento assumido pelo respectivo autor, sendo possível perceber que a TA é a medida mais conflitiva em termos teórico-metodológicos.

Figura 15 – Terminologia e conceitos relativos às duas principais medidas do tempo de fala

<p>Taxa de elocução (“<i>speech rate</i>”, “<i>speaking rate</i>”, “<i>syllable rate</i>”)</p>	<p>Número de unidades linguísticas constantes em um intervalo de fala dividido pela duração desse intervalo.</p> <p>Considera todo material vocal, linguístico e não linguístico, além das pausas (silenciosas e preenchidas) e de outras interrupções na fluência (JESSEN, 2007; LAVER, 1994).</p> <p>É medida tanto global (no texto oral integral) como localmente (em um intervalo de fala menor que o texto oral integral).</p>
<p>Taxa de articulação</p>	<p>Número de unidades linguísticas constantes em um intervalo de fala pausa-excludente dividido pela duração desse intervalo.</p> <p>Há consenso quanto à exclusão das pausas silenciosas, mas controvérsia quanto à consideração (LAVER, 1994; MILLER et al., 1984; ROBB et al., 2004; TSAO e WEISMER, 1997; VERHOEVEN et al., 2004) ou não (DANCOVIČOVÁ, 1997, 2001; JACEWICZ et al., 2009; JESSEN, 2007; KÜNZEL, 1997) das pausas preenchidas e da fala disfluente.</p> <p>Também pode ser medida global e localmente.</p> <p>É alternativamente definida com base na duração média das sílabas, obtida através da divisão da duração do intervalo de fala pelo número total de sílabas (CRYSTAL e HOUSE, 1990; MILLER et al., 1984; QUENÉ, 2008).</p>

Fonte: A autora (2013).

Quanto às premissas que perpassam inevitavelmente o desenvolvimento de uma pesquisa sobre taxa temporal de fala, cada uma delas implicando laboriosa reflexão e posicionamento teórico coerente, reincide na literatura da área especialmente os questionamentos acerca de qual unidade linguística utilizar no cálculo, do *status* a ser atribuído à unidade selecionada e da abrangência da análise. Sendo o caso, ainda, debate-se o tipo de intervalo de fala a ser submetido ao exame local.

Com relação ao primeiro questionamento, o que se refere a qual unidade linguística utilizar no cálculo das taxas, foi encontrada, no levantamento bibliográfico realizado, referência ao uso de ao menos seis tipos de unidades. São eles: sentença (RAMUS, 2002); palavra (BINNENPOORTE, 2005; KRAUSE e BRAIDA, 2002; SYRDAL, 1996; TAUROZA e ALLISON, 1990; YUAN et al., 2006); sílaba (BYRD, 1994; DANCOVIČOVÁ, 1997, 2001; GOLDMAN-EISLER, 1968; GROSJEAN e DESCHAMPS, 1975; JACEWICZ et al., 2010; JESSEN, 2007; KÜNZEL, 1997; MALÉCOT et al., 1972; ROBB et al., 2004; TSAO e WEISMER, 1997; TSAO et al., 2006; VERHOEVEN et al., 2004); unidade VV³⁰ (BARBOSA, 2006; MEIRELES, 2007; MEIRELES e BARBOSA, 2008, 2009); mora (KAKITA, 1998; KOISO et al., 1998) e segmento (BINNENPOORTE, 2005; FONAGY e MAGDICS, 1960; MIRGHAFORI et al., 1996; PFITZINGER, 1998; TROUVAIN et al. 2001; VERHASSELT e MARTENS, 1996).

Na seleção da unidade linguística a ser utilizada na contagem, Trouvain (2004) estabelece alguns critérios, sendo eles, o grau de popularidade (quão difundido é o uso da unidade), a comparabilidade entre os estudos (se é uma unidade relativamente estável na língua e entre as línguas), a facilidade na contagem (se é uma unidade mais ou menos custosa à inspeção acústica), a clareza na definição (se é uma unidade cujo conceito é consolidado ou controverso) e o potencial de evidenciação da variância temporal (se é uma unidade que permite observar com certa facilidade a variância das taxas).

Andrade et al. (2003, p.82, tradução nossa) apontam como determinante a escolha da unidade de contagem, uma vez que cada uma traz consigo uma informação específica: “o número de palavras por minuto é o índice da produção de informação, enquanto que o número de sílabas por minuto indica [...] a habilidade de transição motora”.

Com respeito à sílaba, Crystal e House (1990), mostrando-se preocupados com a possível interferência da variação estrutural e de complexidade da mesma nos resultados da TA, propõem que o cálculo da taxa seja feito via duração média da sílaba (*Average Syllable Duration – ASD*), método também adotado por Goldman-Eisler (1956), Miller et al. (1984) e Quené (2008). Cabe apontar que a

³⁰ A unidade VV é “[...] constituída por uma vogal e todos os segmentos assilábicos que a seguem, independentemente da fronteira silábica, até o *onset* da vogal seguinte [...]” (BARBOSA, 2006, p.30).

decisão de se obter a TA através do tamanho médio de sílaba ou da divisão do número de sílabas pela duração do intervalo de fala não é exatamente impactante, uma vez que há possibilidade de versão dos resultados de uma para outra forma de cálculo. Contudo, percebe-se como difundido entre os estudos consultados, sendo por essa razão aqui adotado, o uso do número de unidades linguísticas como numerador da razão utilizada no cálculo da TA.

Embora seja generalizadamente empregado na literatura nacional e internacional o número de sílabas por segundo, pesquisas interlinguísticas sugerem que a unidade de medida mais apropriada, devido às diferenças rítmicas decorrentes da estrutura da sílaba, não seria a mesma em todas as línguas. Nesse sentido, Ramus (2002) aponta como quase ilusória a proposição de uma medição que seja igualmente eficiente em todas as línguas.

No Brasil, entre os estudos sobre entonação, ritmo e processos de redução que procederam ao cálculo da TE é profícua a utilização da unidade VV, sendo essa, no entanto, por conceito, pausa-includente, o que inviabiliza sua adoção em concepção original como unidade de contagem no estabelecimento da TA. Dessa forma, por ser mais razoável a consideração de um único tipo de unidade de contagem para ambas as taxas investigadas, neste estudo, optou-se pelo uso da sílaba.

O segundo questionamento, relacionado ao *status* da unidade selecionada, refere-se à perspectiva de observação da unidade de contagem, o que equivale, em outras palavras, ao nível de fala considerado na análise. Em se tratando de sílaba, por exemplo, é necessário definir se serão computadas, no procedimento, as sílabas ortográficas, fonológicas ou fonéticas ou auditivamente percebidas.

Conceitualmente menos controversa, a sílaba ortográfica é também referida como gramatical. Já a noção fonológica de sílaba, segundo o paradigma não linear autossegmental proposto por Selkirk (1982), apregoa que as sílabas possuem uma organização hierárquica interna, composta dos elementos ataque (*onset*) e rima, a última subdividida em núcleo e coda, sendo o núcleo da sílaba o único elemento obrigatório da estrutura.

Quanto às sílabas percebidas, Tauroza e Allison (1990) afirmam que as mesmas correspondem às ocorrências de sílabas ouvidas. Bento et. al. (1998, p.91), no entanto, vinculam-nas à inteligibilidade, uma vez que, segundo os autores,

implicam a “recepção e discriminação dos sons, [...], processo que envolve a atividade sensorial e cortical”.

De fato problemática é a concepção acerca da sílaba fonética. Stetson (apud LAVER, 1994) explica-a a partir do mecanismo de corrente de ar pulmonar, correspondendo as sílabas, nessa perspectiva, a impulsos expiratórios (constituídos de contração dos músculos respiratórios e dos consequentes jatos de ar).

Ladefoged (2001, p.227-229, tradução nossa) sugere uma definição de sílaba fonética baseada na sonoridade inerente a cada som, entendendo por sonoridade “a sensação psicoacústica de intensidade de um som em relação a outro com igual duração, acento e tom”. A proposta prevê que os picos de sonoridade coincidem com os picos de silabidade (núcleos silábicos), argumento que, segundo o próprio autor falha em uma série de casos. Complementarmente, refere que na identificação das sílabas fonéticas podem-se considerar os picos de proeminência, dados pela duração, acento e tom próprios dos segmentos, mas ressalta que o parâmetro integrado é subjetivo e não mensurável.

Destaca-se que, assim como quanto ao tipo de unidade de contagem, a concepção assumida acerca do *status* da unidade selecionada produz resultados e interpretações próprios. Ainda, ressalva-se que, na dinâmica da fala, o número de sílabas ortográficas ou fonológicas tende a ser reduzido com a aplicação de processos típicos do estilo coloquial, razão pela qual um dado material de fala apresentará taxas temporais menores em cálculos realizados com base nas sílabas efetivamente produzidas (fonéticas) ou percebidas do que os realizados com base nas sílabas ortográficas ou fonológicas.

Nesse sentido Trouvain (2004, p.44, tradução nossa) esclarece:

As sílabas podem ser subjacentes, derivadas da forma lexical da palavra, ou realizadas. Para sílabas subjacentes, o número de sílabas parece ser sempre claro. Em contraste às sílabas fonêmicas, a presença de uma sílaba fonética é por vezes difícil de detectar. Sílabas podem ser apagadas [...] e em muitos casos é difícil decidir quando uma sílaba está ou não presente.

Jessen (2007) refere como mais difundida no meio forense a utilização da sílaba fonética, o que pode dever-se ao fato de que essa propicia o emprego de métodos automáticos de detecção dos picos de energia no sinal (correspondentes

aos núcleos silábicos, constituídos de vogais) e ao fato da forma subjacente nem sempre ser clara.

Künzel (1997) sugere que a escolha da versão de fala ou *status* da unidade seja condizente com o objetivo da investigação, ou seja, se o desejo for verificar a quantia de informação lexical transmitida em um intervalo de tempo, a contagem deve ter por base a sílaba subjacente, já se a articulação for o foco, a perspectiva fonética de sílaba deve ser a admitida.

Considerando que a maioria dos estudos sobre TE e TA expressa seus resultados em sílabas por segundo e que é difundido o uso da sílaba fonética (CAO e WANG, 2011; DANCOVIČOVÁ, 1997, 2001; HEWLETT e RENDALL, 1998; JACEWICZ et al., 2009, 2010; JESSEN, 2007; KENDAL, 2009; ROBB et al., 2004; TSAO e WEISMER, 1997), assumiu-se no cálculo das taxas aqui examinadas a perspectiva fonética de sílaba.

Considerando a proposição de coincidência entre picos de sonoridade e de silabidade feita por Ladefoged (2001), neste estudo, as sílabas fonéticas tiveram sua presença definida em razão da existência no sinal de, ao menos, um dos vestígios acústicos associados às vogais plenas (silábicas) ou de “pistas remanescentes do gesto vocálico” (MENESES, 2012), a saber: regularidade na forma de onda, barra de sonoridade no espectrograma e valores formânticos compatíveis com os da vogal auditivamente percebida³¹.

Com relação ao terceiro questionamento, referente à abrangência da análise e ao tipo de intervalo de fala a ser submetido ao exame local, tem-se como possível a obtenção de taxas a partir da duração integral de um determinado material de fala, realizando-se um único cálculo de taxa (o que se intitula mensuração global) ou a obtenção de taxas a partir da avaliação de diversos trechos, tomadas parciais de um registro integral, o que implica múltiplos cálculos de taxa (configurando a intitulada mensuração local).

Jessen (2007) explica que a mensuração global das taxas caracteriza o tempo habitual do sujeito, enquanto que a observação local evidencia as mudanças no uso do tempo durante a produção da fala. Quanto à obtenção de medidas globais e locais de TA, aplicando-se parcialmente o raciocínio à TE, o autor refere (p.54, tradução nossa):

³¹ Para tanto, adotou-se como referência os escores relativos ao PB apresentados por Escudero et al. (2009, p.1383), reproduzidos no ANEXO B.

A TA global é capturada através da TA média e a TA local pelos resultados dos intervalos [...] individuais e pelo desvio padrão da TA, que é um índice da variabilidade da TA ao longo das diferentes partes da gravação.

A mensuração local, menos investigada do que a global, além de indicar as variações de taxa naturalmente ocorridas ao longo do discurso e a variabilidade intrassujeito, permite a detecção de padrões de comportamento e do domínio de variação (investigados, entre outros, por Dancovičová, 1997, 2001). A abordagem local requer, no entanto, que se defina um tipo de intervalo de análise.

No que se refere ao tipo de intervalo de fala a ser submetido ao exame local, a literatura sobre o assunto relata a utilização de turnos de fala (BORSEL e MAESSCHALCK, 2008; GOLDMAN-EISLER, 1956), orações (DANCOVIČOVÁ, 1997, 2001), intervalos interpausais³² (CRYSTAL e HOUSE, 1990; DANCOVIČOVÁ, 1997, 2001; JACEWICZ et al., 2010; KOREMAN, 2005; MALÉCOT et al., 1972; MILLER et al., 1984; TSAO e WEISMER, 1997) e frases entonacionais (DANCOVIČOVÁ, 1997, 2001; KOREMAN, 2005; TROUVAIN, 2004).

Jessen (2007) aponta argumentos desfavoráveis à utilização dos dois últimos (intervalos interpausais e frases entonacionais), destacando que o intervalo interpausal tem tamanho variável em razão da frequência das pausas no discurso e que a frase entonacional, embora mais uniforme em comprimento, tem delimitação que envolve conhecimentos linguísticos de aplicação, com exceções, subjetiva. O autor propõe, então, um tipo de intervalo de fala diferenciado, o trecho de memória (correspondente à porção de fala fluente com um número de sílabas que pode ser facilmente retido na memória de curto prazo), posteriormente adotado por Cao e Wang (2011).

Koreman (2005), em um de seus experimentos, escolhe o intervalo interpausal em detrimento da frase entonacional sob o argumento de que o primeiro é mais fácil de ser determinado, ressalvando, contudo, que tais tipos de intervalo de maneira geral equivalem-se.

Em razão de figurar entre os objetivos do presente estudo a verificação da existência de diferença significativa entre as formas de mensuração (global e local) e

³² Já definido na Introdução deste documento como sendo o intervalo de fala limitado por duas pausas silenciosas consecutivas, de duração maior ou igual a 130ms.

entre os tipos de taxa (TE e TA), considerando-se, ainda, que nos diálogos a maior unidade linguística produzida pelo locutor-alvo é o turno de fala e que o tipo de intervalo mais comumente utilizado no cálculo da TA é o intervalo interpausal, previu-se que o método a ser empregado admitiria os dois mencionados escopos de observação e, como intervalo de fala a ser submetido ao cálculo da TE e da TA, respectivamente, o turno de fala e o intervalo interpausal.

Apresentados os três tópicos de conflito mais discutidos na literatura, são necessárias, ainda, considerações acerca de outras questões de menor destaque, porém não menos importantes, posto que capazes de, em algum grau, influenciar os resultados (taxas temporais obtidas):

- a) o tipo de material de fala examinado (se espontâneo ou não), sua forma de eliciação e condições de gravação;
- b) o tratamento dado às pausas (silenciosas e preenchidas) e à fala disfluente;
- c) o tamanho do intervalo de fala admitido para análise.

Quanto à primeira questão, considerando o levantamento bibliográfico realizado, identificou-se como preponderante a utilização da fala não espontânea gravada em ambiente laboratorial, sendo especialmente profusa a utilização da tarefa de leitura oralizada na confecção dos *corpora*. Entre as razões para que tal técnica de tomada de fala e ambiente de coleta seja preferida está o fato de que nessas circunstâncias há oportunidade de maior controle sobre o conteúdo emitido (levando à obtenção de um material texto-dependente, pouco afeito à ocorrência de manifestações adversas como, por exemplo, a de alongamentos segmentais associados à hesitação) e a geração de um sinal sonoro com suficiente qualidade acústica (condição pouco frequente em gravações produzidas em ambientes onde não há controle dos sons adversos).

Não obstante, há relativo consenso quanto à existência de diferença entre os tempos de fala obtidos a partir da leitura oral em relação àqueles obtidos a partir de fala espontânea, encontrando-se registro de maiores valores de taxa em material lido em Butcher (1981), Crystal e House (1990), Jessen (2007), Künzel (1997), Mixdorff e Pfitzinger (2005), Trouvain (2004) e Tsao e Weismer (1997).

Braun (1995) argumenta que as diferenças relacionadas às taxas temporais existentes entre os dois tipos de material de fala (o lido oralmente e o espontâneo) deve também advir do fato de que a leitura requer processamento cognitivo e

fonológico diferenciados, envolvendo as habilidades do indivíduo como leitor. Além disso, o autor justifica que os valores mais altos de TE na leitura podem estar associados às características desse tipo de material de fala, constituído de um menor número de pausas e de intervalos mais longos sem pausa quando comparado à fala casual ou espontânea (definidas na Subseção 1.3.1, Capítulo 1). As pausas fazem aumentar a duração do intervalo de fala em examinação, levando a um quociente (valor de taxa) conseqüentemente menor.

Foram encontrados, ainda, trabalhos com taxas geradas a partir de leitura oral, nos quais houve indução intencional para produção de diferentes taxas de fala, o que foi feito através do uso de instrumentos como metrônomo ou de recursos oferecidos em *softwares* de edição de áudio (a exemplo, respectivamente, Meireles, 2001, 2007; Tillman e Pfitzinger, 2003) ou, ainda, de orientações verbais orais impressionísticas, como por exemplo: “fale como se estivesse falando com um amigo”, “fale o mais rápido que puder sem que sua fala seja ininteligível”, “fale o mais rápido que puder sem apagar sons ou distorcê-los” e “fale o mais devagar que puder sem introduzir pausas”, referidas por Meireles (2007) e Tsao e Weismer (1997).

Deve-se registrar que foram encontrados trabalhos cujas metodologias contemplam a verificação da TE e/ou da TA em fala produzida ao telefone, a saber: Binnenpoort et al. (2005), Cao e Wang (2011), Jessen (2007), Künzel (1997), Syrdal (1996) e Yuan et al. (2006). Desses, há os que contrapõem, e por essa razão são de especial interesse neste estudo, o comportamento da TE e/ou da TA observado na fala presencial com o observado em fala manifesta em conversação telefônica, a saber, Jessen (2007), Künzel (1997) e Binnenpoort et al. (2005), tendo sido utilizada como unidades de contagem no último, contudo, a palavra e o fone.

Dessa forma, logra-se que os trabalhos com delineamento mais compatível com o do estudo em tela são Jessen (2007) e Künzel (1997), ressaltando-se que nesses, diferentemente do que se tem no presente estudo e na realidade forense, os locutores dos diálogos telefônicos tinham ciência de estarem sendo gravados.

Llisterri (1992), advertindo quanto aos cuidados necessários à comparação de resultados provindos de configurações de pesquisa díspares, especificamente quanto à questão dos diferentes estilos de fala, afirma que a comparação significativa entre resultados de diferentes pesquisas somente é possível se as amostras de fala forem obtidas em condições similares, o que justifica o empenho

perpetrado no sentido de assegurar que a presente investigação considerasse a configuração forense real (fala casual produzida ao telefone e fala espontânea produzida em situação de entrevista semidirigida, gravada ambientalmente).

Adentrando à segunda questão, a que diz respeito às pausas e às disfluências, cabe salientar que o tratamento dado à pausa é de especial interesse no cálculo da TA (a qual se baseia em conteúdo sonoro mais restrito), uma vez que a TE se debruça sobre a oralidade plena.

A pausa é um fenômeno comumente observado na fala espontânea. Normalmente mostra-se associada à hesitação e ao planejamento da fala e é caracterizada por um tempo de silêncio (intitulada, por essa razão, pausa silenciosa) ou pela pronúncia de determinados sons (intitulada, nesse caso, pausa preenchida).

Na pausa preenchida é observada a utilização de sons preenchedores, entre eles “é”, “hum” e “ã”. Esse tipo de pausa é apontado como o fenômeno mais típico da fala espontânea (TROUVAIN, 2004).

Já a pausa silenciosa, seja ela fisiológica (para reposição do suporte aéreo, sendo eventualmente delatada por ruído no sinal acústico) ou linguística (usada no planejamento da fala ou como recurso entonacional), é caracterizada por Duez (1982) como qualquer intervalo no traçado oscilográfico onde a amplitude não se distingue do ruído de fundo, ou seja, onde não há vestígio de sinal vocal em intensidade significativa.

A duração da pausa silenciosa é fortemente discutida na literatura, sendo a duração da mesma bastante variável. Em Kowal et al. (1975) é assumida a extrema duração de 270ms, em Grosjean e Dechamps (1975), assim como na reanálise do estudo feita por Miller et al. (1984), a de 250ms. Laver (1994) e Goldman-Eisler (1968) têm como referência a duração de pelo menos 200ms; Van Donzel e Beinum (1996) apontam 150ms; Dancovičová (2001) advoga a adoção de 130ms, enquanto Butcher (1981) sugere a duração de 100ms. Os valores mais baixos encontrados na literatura são de 60ms, utilizado por Kendall (2009), e de 50ms, utilizado por Robb et al. (2004).

A larga variação nos valores apontados pode ser justificada pela incorporação tecnológica no cenário dos estudos. Como exemplo tem-se Hieke et al. (1983) que, ao adotarem os recursos da análise instrumental, constataram a existência de pausas abaixo do limite de percepção do observador, sendo essas por eles intituladas pausas breves.

Na delimitação do sinal de áudio com vistas ao cálculo da TA a concepção da duração da pausa silenciosa é de suma importância, pois intervalos de silêncio (ou de manifestação sonora restrita ao ruído de fundo) podem ser atribuídos inadvertidamente ao tempo não vocalizado necessário para reposicionamento dos articuladores, situação definida classicamente por Goldman-Eisler (1958) como a de uma pausa articulatória (sendo um exemplo desse tipo de pausa o tempo de fechamento das consoantes plosivas).

Trouvain (2004, p.28, tradução nossa) adverte:

A fase de fechamento das plosivas [...] não deve, é claro, ser considerada uma pausa, embora constitua um intervalo de silêncio. A fase de *clousure* tipicamente compreende de 30 a 70ms. Mas têm-se de fato problemas quando um intervalo interpausal começa com uma plosiva.

Por ser o principal elemento de diferenciação entre os dois tipos de taxas temporais de fala mais utilizados (TE e TA), os estudos sobre pausa são relativamente profusos na literatura, sendo possível perceber a preocupação com o rigor em seu manuseio.

Quanto ao limiar de diferenciação entre a pausa articulatória e a pausa silenciosa, foi admitido para fins deste estudo 130ms, em razão da especificidade (HIEKE et al., 1983) e do nível de aprofundamento (DANCOVIČOVÁ, 1997, 2001) de três obras que o indicam, além do fato de ser, tal valor de limiar, intermediário entre os referidos nas demais bibliografias consultadas.

Voltando à questão da pausa preenchida, é preponderante nos estudos sobre o tema, e por essa razão aqui adotada, a sua admissão como material de fala útil no cálculo da TE e o seu descarte, como não útil, no cálculo da TA (conforme APÊNDICE A). Exceção digna de nota reside no fato de Laver (1994), e autores de alguns trabalhos subsequentes (ao menos os sinalizados no referido apêndice), acolher a pausa preenchida no material de fala submetido ao cálculo da TA.

Também no escopo da discussão sobre que conteúdo de fala considerar como útil no cálculo das taxas estudadas está a problemática envolvendo a produção oral disfluente³³, havendo relativo consenso no que se refere a sua

³³ Segundo Duez (1982) são eventos de disfluência as repetições, os alongamentos e os falsos começos.

admissão no cálculo da TE. Com relação à TA, contudo, há posicionamentos contrários (adotá-la *versus* não adotá-la), podendo-se citar a favor de seu acolhimento Laver (1994) e um pequeno grupo de autores (ao menos os que têm apontado, no APÊNDICE A, a adoção de tal posicionamento).

Considerando que é prevalente a não admissão da fala disfluente como material de fala útil para fins de cálculo da TA e que a transcrição desse tipo de produção requerer mais tempo para execução (pois implica inúmeras revisões), além de áudio em boas condições de inteligibilidade e de qualidade acústica (normalmente não disponível nos áudios provindos de interceptação telefônica), sendo típica na fala naturalística a ocorrência de manifestações de disfluência, admitiu-se, neste estudo, que a fala utilizada no cômputo da TA (diferentemente do adotado no caso da TE) compreende exclusivamente a fala fluente.

Quanto à terceira questão, referente ao tamanho do intervalo de fala admitido para análise, elemento apontado na literatura como também capaz de interferir nos resultados das TEs e das TAs (JACEWICZ et al., 2010; JESSEN, 2007; MILLER et al., 1984; TROUVAIN, 2004), há na literatura tanto referência quanto à redução das mencionadas taxas quando considerados no exame intervalos de fala muito curtos (normalmente constituídos de partículas com segmentos alongados) como quanto ao aceleração da taxa tipicamente associado à produção de enunciados longos, levando ao conhecido fenômeno de encurtamento antecipatório³⁴.

Sobre a admissão de intervalos muito longos à análise, Quené (2008) afirma que frases longas, com muitas sílabas, tendem a ser faladas em taxas rápidas, com encurtamento médio da duração dos segmentos e, por consequência, das sílabas, efeito por vezes apontado como um universal fonético.

Já sobre a admissão de intervalos curtos à análise, tem-se referido na literatura (JACEWICZ et al., 2010) que a conduta é capaz de enviesar os resultados obtidos pois, mesmo havendo o controle da ocorrência de alongamentos, impede que ocorra a desejada neutralização da complexidade do tipo de sílaba, distribuição que se torna naturalmente balanceada quando tomados intervalos maiores de fala.

Malécot et al. (1972) consideraram como intervalo útil de fala as cadeias sonora compostas por duas ou mais sílabas. Jessen (2007), assim como Cao e Wang (2011), admite para cálculo da TA apenas intervalos com quatro ou mais

³⁴ Tendência que as sentenças com mais sílabas têm de serem produzidas com sílabas mais curtas (MALÉCOT et al., 1972).

sílabas, enquanto para Jacewicz et al. (2009) a duração mínima dos intervalos de fala é de cinco sílabas.

Em razão de ser um valor médio dentre os referidos e por ser o limiar adotado em um dos dois estudos de concepção metodológica mais próxima a desta pesquisa (JESSEN, 2007), assumiu-se o número de quatro sílabas como tamanho mínimo de intervalo de fala para fins de cálculo da TE e da TA.

Com base no exposto, percebe-se que a problemática na investigação das taxas temporais de fala concentra-se na montagem da configuração da pesquisa, pois os posicionamentos teórico-metodológicos assumidos tornam distintos os trabalhos, o que por sua vez, em muitos casos, inviabiliza a comparação entre os resultados obtidos nos estudos.

Um panorama da diversidade de opções acerca dos posicionamentos técnicos pré-requisitos à examinação das taxas temporais de fala é apresentado no já mencionado APÊNDICE A, onde são expostas as premissas assumidas por autores de nove pesquisas sobre TE e/ou TA (assim como as consideradas no estudo em tela) que tenham empregado cálculo em sílabas por segundo, a partir de fala espontânea não infantil, de indivíduo em condições normais de linguagem. Deduz-se, assim, que não constam em tal síntese estudos que tenham contemplado a verificação da(s) taxa(s) em grupos especiais (por exemplo, crianças ou indivíduos acometidos de patologia específica), a observação das taxas a partir de material de fala não espontâneo, o uso de unidade de contagem diversa da adotada neste estudo (a sílaba) ou a realização do cálculo da(s) taxa(s) via duração média da sílaba. Cabe comentar, ainda, que alguns dos trabalhos arrolados podem ter considerado outro tipo de material de fala além da fala espontânea, tendo sido, nesses casos, suprimidas as informações relativas ao tipo de material não almejado a fim de otimizar a caracterização pretendida.

Destaca-se que eventuais comparações interlínguas de valores de TE e de TA devem ser feitas com a ressalva da possível interferência dos tipos de moldes silábicos, que podem ser mais complexos e mais frequentes (um determinado tipo) em uma língua do que em outra. Falantes de uma língua com estruturas silábicas mais simples incorrem na produção de mais unidades silábicas por tempo de enunciado do que falantes de uma língua que admite moldes silábicos mais complexos (ou na qual esses sejam mais frequentes), sendo exemplo a língua espanhola em relação à inglesa (KOWAL et al., 1983).

Reitera-se que, dos estudos elencados, a pesquisa de Künzel (1997, sobre TE e TA) e a de Jessen (2007, exclusivamente sobre TA) são as que apresentam o delineamento mais próximo da realidade normalmente encontrada na Fonética Forense brasileira e, por essa razão, intencionalmente, no estudo em tela. Contudo, é importante pontuar que, diferentemente do aqui adotado, o primeiro contou com a participação de sujeitos estudantes universitários, escolaridade normalmente não encontrada nos locutores do confronto de vozes em âmbito criminal (ao menos no Brasil), e considerou a fala disfluente como material útil para cálculo da TA. Já o segundo, considerou um intervalo de fala singular (o trecho de memória) e ateu-se exclusivamente à fala masculina. Ainda, em ambos os estudos, diferentemente do aqui proposto, os locutores na ligação telefônica tinham ciência de que estavam sendo gravados.

2.2 VARIABILIDADE, POTENCIAL INDIVIDUALIZANTE E REFERÊNCIA AO USO FORENSE DA TE E DA TA

As variações temporais na fala podem ser relacionadas genericamente a diversos fatores, entre eles, individuais, fisiológicos, psicológicos, linguísticos, culturais e demográficos, havendo, ainda, conforme Yuan et al. (2006), a ocorrência de interação entre esses.

Há autores (MILLER et al., 1984) que associam as variações no tempo de fala também ao estado emocional do sujeito, tipo de situação comunicativa e à familiaridade do indivíduo com o material em discussão. De fato, Künzel (1997), a partir do trabalho com a intitulada taxa de sílaba e a TA, conclui que a situação comunicativa é (em detrimento das condições de gravação) relevante no que se refere à capacidade de interferir na variabilidade intrassujeito das taxas temporais.

Jacewicz et al. (2009) organizam as fontes de variação mencionadas em intra e intersujeitos, apontando que a variação no tempo de fala intrassujeito é sistematicamente afetada por fatores linguísticos como o comprimento do enunciado, a complexidade do discurso, o grau de formalidade e o estilo de comunicação (em ambientes ruidosos ou à longa distância), assim como por fatores extralinguísticos, entre eles o nível de afeto e o humor. Soma-se à complexa interação entre esses fatores intrínsecos outras fontes de variação, de ordem

intersujeitos, relacionadas às variáveis sociais sexo, educação, ocupação, *status* socioeconômico, local de residência e região geográfica de origem.

Interpretada como um componente da naturalidade da fala, a variação nos tempos de fala durante a emissão dos enunciados pode decorrer tanto de variações no tempo empregado na confecção dos gestos articulatórios dos fones, o que repercute nos resultados da TE e da TA, quanto de mudanças nas durações das pausas presentes no discurso, o que repercute exclusivamente nos resultados da TE (EFTING, 1988).

Miller et al. (1984) reforçam que deve ser observado, quando do cálculo da TE, o número e a duração das pausas presentes na elocução, pois já se tem conhecimento que tanto variações intrassujeito quanto intersujeitos devem-se verdadeiramente à taxa de pausa. Ainda, segundo os autores, a variação não diminui ou aumenta linearmente e sim muda seu curso inúmeras vezes ao longo do discurso. O dimensionamento da magnitude dessa variação, no entanto, ainda é pouco substancial na literatura, inexistindo referência, por exemplo, quanto ao que seria uma variação intrassujeito significativa.

Nos trabalhos de Goldman-Eisler (1968), Malécot et al. (1972) e Grosjean e Deschamps (1975) foi enfatizada a existência de baixa variação intrassujeito na TA e a alta variação decorrente das pausas, achados indicativos de maior estabilidade da TA em relação à TE.

Embora Verhoeven et al. (2004), ao pesquisar a fala espontânea de 160 indivíduos, tenham evidenciado inexistência de variação intrassujeito significativa na mensuração local de TE e de TA, Künzel (1997), em estudo dirigido à aplicação forense, também encontra menor variação intrassujeito a favor da TA. A avaliação da distribuição cumulativa das diferenças intrassujeito e intersujeitos feita pelo autor leva-o a concluir que a TA tem um poder de individualização muito mais forte do que o da TE, o que justificaria a preferência pela utilização da primeira no confronto forense de vozes. O autor identifica a TA como consideravelmente constante entre os indivíduos, considerando-a um promissor parâmetro individualizante na identificação de locutor.

Nesse sentido, Cao e Wang (2011, p.396, tradução nossa) sintetizam: “Estudos precedentes sobre o tempo de fala mostraram que a TA tem um maior poder discriminatório de falante que a TE no inglês [...] e no alemão [...].”

Cabe referir, no entanto, que os resultados favoráveis à invariabilidade intrassujeito da TE e da TA são menos expressivos na literatura. Os dados de Goldman-Eisler (1968), originalmente favoráveis à invariabilidade da TA, quando submetidos à reanálise de Miller et al. (1984) para revisão do desvio padrão das taxas obtidas a partir dos intervalos de fala pesquisados (turnos de fala), evidenciaram razoável variação, representada por extensão de coeficiente de variação indicativa de forte dispersão (de 11,5 a 25,1%).

A despeito dos posicionamentos contrários, Jacewicz et al. (2009, 2010) atribuem as diferenças individuais no tempo de produção da fala primeiramente à duração inerente dos movimentos articulatórios, razão pela qual consideram o tempo usado na fala uma característica própria do falante, assim como a qualidade vocal, o uso da entonação e da pausa.

A participação do componente neuromuscular é detalhadamente investigada por Tsao e Weismer (1997), que contrapõem tal proposição à hipótese de condicionamento sociolinguístico, sendo parte do *corpus* utilizado posteriormente reavaliado em Tsao et al. (2006). Os autores de ambas as obras concluem que seus resultados, obtidos a partir de leitura oral, suportam suficientemente (no primeiro estudo, contudo, não exclusivamente) a ideia de existência de uma base biológica responsável pelas diferenças intersujeitos encontradas.

Quanto à variação nos valores dos parâmetros utilizados no confronto de voz e fala, Rose (2002, p.10, tradução nossa) afirma: “dada a existência de variação intrassujeito, outro requerimento lógico à realização da identificação forense de locutor é que a variação intersujeitos seja maior que a intrassujeito”.

Nolan (1997), igualmente, afirma ser desejável que os parâmetros técnicos utilizados na comparação de falantes tenham baixa variação intraindividual, mas ampla variação interindividual, pois assim evidenciarão um significativo potencial discriminante de sujeito. Morrison (2009b, p.2387, tradução nossa) complementa: “se uma variável acústica tem uma variância intersujeitos maior que a intrassujeito, então ela contém informação que pode ser utilizada para auxiliar a distinguir falantes”, elucidando, dessa forma, que o poder discriminatório do parâmetro (neste caso da TE e da TA) é, por essência, decorrente da mencionada relação entre a variabilidade intrassujeito e intersujeitos.

Em suma, o indicativo mais importante da relevância forense de um parâmetro técnico-comparativo é a sua capacidade de discriminação entre amostras,

o que, considerada a indicação de uso de medida(s) de tempo de fala na Comparação de Locutor constante em Cao e Wang (2011), Eriksson (2012), Frech e Harrison (2007), French et al. (2010); Gold e French (2011), Hollien (2002), Jessen (2007), Künzel (1997) e Rose (2002), justifica a investigação do potencial individualizante da TE e da TA almejada neste estudo.

A condução adequada da admissão da TE e/ou da TA à prática forense, no entanto, implica o conhecimento dos valores de referência na língua (frequência e distribuição dos valores de taxa na população), ainda incipiente no que se refere ao PB, assim como informações quanto ao condicionamento gerado por variáveis intervenientes, visando o controle a minimização dos efeitos das mesmas. O presente estudo contribuirá, nesse sentido, fornecendo dados obtidos a partir de *corpus* forense do PB e com resultados referentes ao impacto causado por variáveis específicas (extralinguísticas e linguísticas) no comportamento das taxas estudadas.

2.3 POSSÍVEIS CONDICIONADORES DA TE E DA TA

2.3.1 Extralinguísticos

As variáveis extralinguísticas idade, sexo e dialeto recorrem em estudos que tem por tema os aspectos temporais da fala (BINNENPOORTE, 2005; BORSEL e MAESSCHALCK, 2008; BYRD, 1994; HEWLETT e RENDALL, 1998; JACEWICZ et al., 2009, 2010; KENDAL, 2009; MALÉCOT et al., 1972; QUENÉ, 2008; RAMIG, 1983; ROBB et al., 2004; SYRDAL, 1996; TAUROZA e ALLISON, 1990; TSAO e WEISMER, 1997; VERHOEVEN et al., 2004; YUAN et. al., 2006). Devido à impossibilidade de acesso ao histórico linguístico dos sujeitos participantes da presente pesquisa, a maior parte deles com exposição linguística diferenciada devido a recorrentes aprisionamentos, proceder-se-á ao relato de informações associadas, somente, às variáveis idade e sexo.

Quanto à idade, Malécot et al. (1972) apontam, em estudo sobre o francês espontâneo, que adultos mais velhos (com idades atingindo 69 anos) falam mais lentamente que adultos mais jovens (com idades a partir de 20 anos), sendo

observado decréscimo progressivo à taxa de meia sílaba por segundo (de 5,95 sil/s³⁵ a 5,52 sil/s), à medida que aumenta a idade do indivíduo.

Concordando com o achado, Ramig (1983), ao investigar 48 sujeitos do sexo masculino, com idades entre 25 e 75 anos, classificados em três grupos etários (25 a 35 anos, 45 a 55 anos e 65 a 75 anos), obteve como resultado valores de TE e de TA progressivamente menores com o avanço da idade dos sujeitos pesquisados. Na TE encontrou 3,62 sil/s na primeira faixa etária do estudo, 3,55 sil/s na segunda e 3,28 sil/s na terceira. Já na TA, encontrou 3,94 sil/s na primeira faixa etária do estudo, 3,84 sil/s na segunda e 3,60 sil/s na terceira.

Também em Verhoeven et al. (2004) há apontamento de existência de diferença significativa nas taxas temporais associada à idade, sendo os indivíduos pesquisados falantes do alemão. No estudo, os sujeitos mais jovens (de 21 a 40 anos) evidenciaram valor médio de TE e de TA globais maiores, respectivamente 4,23 sil/s e 4,78 sil/s, do que os encontrados em sujeitos mais velhos (de 45 a 59 anos), respectivamente com 4,01 sil/s e 4,52 sil/s.

Semelhante resultado, porém considerando-se outra forma de contagem (palavras por segundo), é encontrado em Yuan et al. (2006), ou seja, falantes mais velhos (do inglês e chinês) evidenciaram fala mais lenta se comparado ao tempo de fala empregado pelos falantes mais jovens.

Tal comportamento foi detectado também em dois dialetos do inglês americano, de Wisconsin e da Carolina do Norte em material provindo de leitura e apenas no dialeto do norte (de Wisconsin) em material referente à fala informal (JACEWICZ et al., 2009).

Quené (2008), com relação ao holandês, apresenta resultados com valores de TE e de TA mais altos para indivíduos mais jovens, no entanto o autor adverte que o achado pode decorrer do efeito indireto do comprimento da frase que, para ele, tende a ser relativamente encurtado em falantes mais velhos, o que favoreceria o emprego de taxas mais baixas.

Atentos ao alerta, Jacewicz et al. (2010) controlaram a variável comprimento da frase a fim de verificar se a mesma condiciona ou não os valores da TA. Nos dois dialetos do inglês americano pesquisados (o de Wisconsin e o da Carolina do Norte), a partir de uma amostra composta por sujeitos com idades entre 8 e 13 e entre 20 e

³⁵ A unidade “sil/s” (por ser usualmente empregada) será utilizada, neste estudo, na expressão dos valores de taxas (TEs e TAs).

91 anos, encontrou-se resultado similar, com menor TA na faixa-etária infantil (além de na faixa-etária composta por adultos mais velhos). Os autores observaram a ocorrência de aumento progressivo da TA até aproximadamente os 45 anos, esse seguido de contínuo (embora não constante) decaimento, o que denota um padrão de ascensão e queda da TA com o avançar da idade cronológica dos indivíduos.

Relativamente à variável sexo, no levantamento bibliográfico realizado foram encontrados estudos apontando a inexistência de diferenças entre os sexos no que se refere às taxas temporais empregadas na fala (BLOCK e KILLEN, 1996; JACEWICZ et al., 2009; MALÉCOT et al., 1972; TAUROZA e ALLISON, 1990), assim como estudos em que tal diferença é observada como sendo favorável ao sexo feminino (SYRDAL, 1996; YUAN et al., 2006) e, o evidenciado em maioria, ao sexo masculino (BYRD, 1994; HEWLETT e RENDALL, 1998; JACEWICZ et al., 2010; KENDAL, 2009; QUENÉ, 2008; VERHOEVEN et al., 2004).

Tauroza e Allison (1990), pesquisando igual taxa em 32 homens e 27 mulheres falantes do inglês britânico, não encontraram comportamento temporal específico para um ou outro sexo. Já Syrdal (1996), em estudo com 80 locutores do sexo masculino e 80 do sexo feminino, falantes de oito diferentes dialetos do inglês americano, encontrou diferença favorável, embora não estatisticamente significativa, ao segundo grupo (mulheres exibindo TEs maiores, média de 4,77 sil/s, do que as dos homens, média de 4,76 sil/s). Tal resultado atingiu índices significativos em Yuan et al. (2006), que consideraram material de conversação telefônica em inglês e chinês.

Apesar do relatado, imensa maioria dos estudos sobre TE e/ou TA que contemplam a variável sexo apontam a categoria masculino como associada aos valores mais altos desses tipos de taxa.

Hewlett e Rendall (1998), ao pesquisar a fala conversacional de 24 sujeitos com idades entre 15 e 18 anos, falantes de dois dialetos distintos do inglês, obtiveram valores de TE e de TA maiores para os indivíduos do sexo masculino.

Verhoeven et al. (2004) encontraram diferença significativa nas taxas temporais dos sujeitos pesquisados no que se refere à variável sexo. No estudo, os falantes do sexo masculino evidenciaram valor médio de TE e de TA globais maiores, respectivamente 4,23 sil/s e 4,79 sil/s, do que os encontrados na fala feminina, respectivamente 4,01 sil/s e 4,50 sil/s.

Em concordância, Kendal (2009), ao estudar a fala de 104 ingleses (sendo 47 homens e 57 mulheres), com observação de ampla faixa etária (10 a 90 anos), apresenta TA média de 4,99 sil/s para os falantes masculinos e 4,63 sil/s para os falantes femininos.

Tal tendência, porém com cálculo via duração média da sílaba, é observada também em Quené (2008). O autor destaca que as diferenças associadas à variável sexo podem ser significativas, no entanto, ressalta que estão abaixo do limite em que passam a ser percebidas pelo ouvinte.

Jacewicz et al. (2010), a partir do uso de um tratamento estatístico de dados mais refinado, aplicado aos dados apresentados em Jacewicz et al. (2009), apontam que os homens falam significativamente mais rápido do que as mulheres, sendo 5,09 sil/s a média obtida no primeiro grupo e 4,92 sil/s a obtida no segundo.

A variável tipo de gravação (como aqui admitida, denotando ciência ou não da gravação) assume visível destaque neste estudo em razão de ser uma das principais peculiaridades encontradas na rotina de Comparação Forense de Locutor. Tal variável figura no presente estudo, assim como a escolaridade e o *gap* temporal (interlúdio temporal existente entre as gravações), como variável independente extralinguística, carecendo-se, ao que se sabe, ao menos no Brasil (assegurada a procura também em textos internacionais), de estudo precedente que as tenham contemplado.

A especificidade da realidade forense (especialmente no que se refere ao confronto de parâmetros extraídos de fala gravada desavisadamente com os extraídos de fala cuja gravação é sabida) conduz a sério prejuízo em termos de possibilidade de comparação de resultados. Referência metodologicamente mais próxima ao desenvolvido neste estudo, reitera-se, são os trabalhos de Künzel (1997) e Jessen (2007).

Künzel (1997) preocupou-se em avaliar como se comportavam a TE e a TA globais em distintos tipos de fala (fala espontânea, recontagem e leitura) e condições de gravação, selecionando, no último caso, as formas de gravação atualmente mais recorrentemente encontradas no trabalho forense de Comparação de Locutor (diálogo ao telefone *versus* diálogo presencial). O autor avaliou a fala de 10 sujeitos (5 homens e 5 mulheres), não tendo encontrado diferença significativa,

em número de sílabas canônicas³⁶ por segundo, que pudesse ser atribuída ao fato de se estar dialogando com um entrevistador ao telefone ou presencialmente, o que sugere (grosso modo) a validade da comparação de perfis de voz e fala advindos desses dois sistemas de gravação.

Já Jessen (2007), ao que parece tentando reproduzir o trabalho recém-descrito, em cujo final há a indicação da necessidade de se trabalhar com um número maior de sujeitos, observou a TA global e localmente em 100 sujeitos (todos do sexo masculino) falantes nativos do alemão. Igualmente, o autor não encontrou diferenças significativas entre as médias das taxas obtidas a partir de fala espontânea ao telefone das obtidas a partir de fala espontânea presencial.

Ressalta-se, contudo, que em ambos os trabalhos a conversação telefônica gravada ocorre com conhecimento dos falantes, ou seja, a análise perpetrada pelos autores pode informar apenas quanto ao possível condicionamento do uso de um ou outro meio de comunicação (telefone *versus* presencial). O efeito da consciência de estar sendo gravado, que tanto se pretende compreender, segue, dado a inobservância de registro na literatura, sem ter sido até então abordado.

Quanto a estudos que tenham contemplado a obtenção de fala via conversação telefônica, podem ser mencionados, ainda, os que seguem, sendo expostas as respectivas particularidades e limitações:

- a) Binnenpoorte et al. (2005) – conta com 209 falantes do *Spoken Dutch Corpus*³⁷ (90 homens e 119 mulheres), observando a TE e a TA globais em fala espontânea presencial e ao telefone. Diferentemente do aqui considerado, no entanto, os autores eliminam do cômputo das taxas somente as pausas silenciosas, as taxas são calculadas em termos de palavras e fones por segundo e o tipo de gravação não é uma variável independente no estudo.
- b) Syrdal (1996) – avalia 160 falantes do *Switchboard Corpus*³⁸, de oito dialetos distintos do inglês americano, pareados quanto ao sexo e exclusivamente com, diferentemente do aqui considerado, material de

³⁶ Segundo Koreman (2005), a versão canônica de uma unidade refere-se a sua forma potencial pretendida. Para o autor é a forma regular, aceita e aprovada, não necessariamente, no entanto, um indicativo do que consta na estrutura subjacente.

³⁷ Banco de dados contemporâneo, com amostras de fala produzida na Holanda e em Flanders (OOSTDIJK, 2002), também utilizado em Quené (2008) e Verhoeven et al. (2004).

³⁸ Banco de dados com cerca de 2.400 conversações telefônicas (com registro de ambos os locutores do diálogo), compreendendo 543 falantes (302 do sexo masculino e 241 do sexo feminino) de diferentes áreas dos Estados Unidos (Cf. GODFREY et al., 1992).

conversação telefônica (5 a 7s), não contrapondo esse a qualquer outro tipo de material;

- c) Yuan et al. (2006) – abarca quatro *corpora* distintos, todos de conversação exclusivamente ao telefone, diferentemente do aqui considerado, já que em dois deles o diálogo é estabelecido entre pessoas conhecidas e o cálculo da TA dá-se em palavras por minuto.

Considerando o exposto, figura como objetivo específico no presente estudo a averiguação de existência de correlação entre os valores de TE e de TA encontrados e a idade e o sexo dos sujeitos pesquisados. De posse da informação e devido ao interesse tanto no reforço do caráter sociofonético desta pesquisa quanto no atendimento das demandas forenses, correlacionar-se-á também as taxas estudadas com as variáveis escolaridade, *gap* temporal e tipo de gravação.

2.3.2 Linguísticos

Com base no levantamento bibliográfico realizado, observa-se que, das variáveis linguísticas consideradas nos estudos em que a TE e/ou a TA consta(m) como variável dependente, a saber, o tipo de intervalo de fala (DANCOVIČOVÁ, 1997, 2001; HANSSON, 2002), a posição sequencial do intervalo de fala (QUENÉ, 2008), o tópico conversacional (CAO e WANG, 2011; YUAN, et al., 2006), a complexidade do discurso (JACEWICZ et al., 2009) e a intenção comunicativa (MALÉCOT et al., 1972), a única que se aplica ao presente trabalho, considerados os objetivos apresentados na Introdução, é o tamanho do intervalo de fala, sobre o qual se discorre a seguir.

Quanto à variável tamanho do intervalo de fala, Miller et al. (1984) sugerem que um razoável modelo de tempo de fala deve incluí-la como preditor da variação intrassujeito. Tal variável é apontada em Jacewicz et al. (2010) como capaz de afetar a variabilidade não só intrassujeito mas também a variabilidade intersujeitos.

Cabe documentar que neste estudo a variável tamanho do intervalo de fala refere-se à duração temporal dos turnos de fala e dos intervalos interpausais pesquisados. Apenas em duas situações, na verificação da correlação existente entre as variáveis tamanho do intervalo de fala e idade (para equivaler-se ao procedimento descrito por Jacewicz et. al., 2009, 2010 e por Quené, 2008) e na verificação da correlação existente entre as variáveis tamanho do intervalo de fala e

duração média das sílabas (visando observação do fenômeno de encurtamento antecipatório³⁹), o tamanho do intervalo de fala referir-se-á ao número de sílabas nele contido.

Jacewicz et al. (2010), sofisticando a análise estatística empregada em Jacewicz et al. (2009), com adaptações ao *corpus* original, detectaram que o decréscimo da variabilidade intersujeitos da TA é inversamente proporcional ao aumento do tamanho do enunciado, o que indica que os sujeitos individualmente tendem a convergir para um mesmo valor de taxa à medida que produzem frases mais longas. Miller et al., já em 1984, lançaram um alerta sobre a mencionada ocorrência de neutralização da variação na TA a partir da análise de intervalos de fala muito grandes.

Os autores apontam que o entendimento da TA como altamente idiossincrática, formulado por Goldman-Eisler (1968), pode ter decorrido da forte correlação existente entre o comprimento da sentença e a taxa obtida, uma vez que em tal pesquisa as taxas foram calculadas a partir de turnos de fala, um dos tipos de intervalo de fala admitidos no presente estudo. A constatação de um possível viés gerado pelo tamanho máximo do enunciado é relevante, já que uma das propostas do estudo em tela é a definição do potencial individualizante das taxas estudadas.

Dessa forma, projeta-se também como objetivo específico a verificação da correlação existente entre os valores de TE e de TA empregados pelos sujeitos pesquisados e as durações dos respectivos turnos de fala e intervalos interpausais.

Adiante, segue a descrição da metodologia empregada, com o apontamento do delineamento de pesquisa adotado, dos critérios considerados na composição do *corpus*, da forma de organização dos dados e das análises realizadas.

³⁹ A averiguação do encurtamento antecipatório restringir-se-á ao tipo de mensuração pausa-excludente (TA), a fim de evitar que sejam computados como tempo associado à sílaba os que provêm do silêncio (pausa silenciosa) ou de alongamentos (pausa preenchida).

3 METODOLOGIA

O delineamento do presente estudo é definido na Seção 3.1 deste capítulo (estabelecendo-se nesse o tipo de pesquisa realizado e as variáveis, objetivos e hipóteses assumidos). Após, há detalhamento da constituição da amostra admitida (Seção 3.2), com ênfase na seleção dos sujeitos participantes, e explanação sobre a sistemática de organização dos dados (Seção 3.3), ilustrando-se, entre outros, a rotina de segmentação e etiquetagem dos áudios empreendida.

Na Seção 3.4 tem-se a descrição do procedimento de verificação acústica e a apresentação, em linhas gerais, do tratamento estatístico efetuado (informação expandida na Seção 4.1 do Capítulo 4).

3.1 DELINEAMENTO DO ESTUDO

3.1.1 Tipo de pesquisa

A pesquisa em tela compreende um estudo observacional correlacional retrospectivo, atendo-se exclusivamente ao módulo de produção da fala. O método empregado prevê a análise quantitativa de parâmetros temporais de fala, especificamente a TE e a TA (apropriadamente definidas na Introdução e no Glossário deste trabalho), a partir de mensurações acústicas.

3.1.2 Variáveis

São variáveis dependentes neste estudo a TE e a TA e, independentes, as variáveis sócio-indexicais idade, sexo, escolaridade, além do *gap* temporal e do tipo de gravação, assim enquadrado por compreender espaço de variação estilística, opondo, segundo Labov (1972), a fala casual e a espontânea (ambas definidas na Subseção 1.3.1, Capítulo 1). Consta como variável independente linguística o tamanho do intervalo de fala.

Foram incluídas, ainda, as variáveis denominadas sujeito (enquadrada como independente não linguística), a fim de permitir a observação das variações das taxas entre os participantes da pesquisa, e duração média da sílaba (enquadrada

como independente linguística), para que fosse possível a verificação de existência do fenômeno de encurtamento antecipatório no PB.

Destaca-se que cada um dos sujeitos pesquisados apresenta uma idade associada à gravação desavisada e outra associada à gravação avisada, posto que existe, entre a feitura de ambas as gravações, o já mencionado *gap* temporal.

Quanto à variável sexo, são aqui contempladas as categorias masculino e feminino, destacando-se que o não pareamento do número dos sujeitos quanto ao sexo deve-se à existência de destacada prevalência do sexo masculino no que se refere à autoria de crimes, especialmente os que têm o tráfico de drogas e o homicídio como natureza⁴⁰.

O tamanho máximo do intervalo de fala a ser considerado no cálculo da TE e da TA, em razão de não ter sido encontrado valor de referência em estudos relativos ao tema, não foi limitado, tendo sido encontrada nos *corpora* certa extravagância (sobejo) apenas no que se refere à duração de alguns turnos de fala da gravação avisada, mais afeita à longas narrativas. Os excessos, no entanto, são naturalmente reduzidos no momento de qualificação do material de fala e definição dos turnos de fala e intervalos interpausais úteis.

Quanto ao tamanho mínimo, em concordância com o exposto na literatura, foram admitidos à examinação apenas os turnos de fala com quatro ou mais sílabas fonéticas, sendo esse o limiar adotado em uma das pesquisas de delineamento mais próximo ao deste estudo (JESSEN, 2007).

Reitera-se que no levantamento bibliográfico realizado não foi encontrado estudo sobre TE e TA que tenha estabelecido a escolaridade, o *gap* temporal e a ciência de estar sendo gravado, observada no presente trabalho sob o rótulo “tipo de gravação”, como variáveis independentes.

Os fatores da variável escolaridade, dadas às características do público pesquisado (suspeitos, indiciados ou réus em crimes de tráfico de entorpecentes e/ou homicídio), o qual tem comumente associado um restrito nível de instrução formal, limitou-se ao número de anos de escolarização do ensino fundamental (no formato de oito anos) e do ensino médio⁴¹.

⁴⁰ Atualmente 86% dos integrantes do VOXDC são do sexo masculino.

⁴¹ Conforme inquirido na entrevista, nenhum dos sujeitos participou de instrução escolar formal no período transcorrido entre as gravações desavisada e avisada.

A variável *gap* temporal, que representa o intervalo de tempo transcorrido entre a feitura da gravação desavisada e da gravação avisada, é uma realidade compulsória própria do confronto forense de vozes. Nesse tipo de perícia são cotejadas amostras de fala não contemporâneas com vistas à definição de existência (ou não) de unicidade de autoria, sendo comuns lapsos temporais de dois a quatro anos entre os registros questionado (do locutor que se deseja conhecer) e padrão (do locutor conhecido).

Quanto ao tipo de gravação, tem-se inicialmente que esse se refere à forma de captura do sinal de áudio, tendo sido a gravação desavisada obtida através de interceptação telefônica e a gravação avisada através da coleta ambiental. Considerando-se que na última os interlocutores (Peritos Criminais) objetivam, por princípio metodológico, levar o fornecedor da amostra à produção maximamente espontânea e que há registro na literatura de não interferência do meio de captura (KÜNZEL, 1997; JESSEN, 2007), entende-se como realmente contrastivo entre os tipos de áudios admitidos neste estudo o fato do locutor saber ou não que está sendo gravado, ou seja, a ciência da gravação. Em suma, os fatores da variável tipo de gravação são:

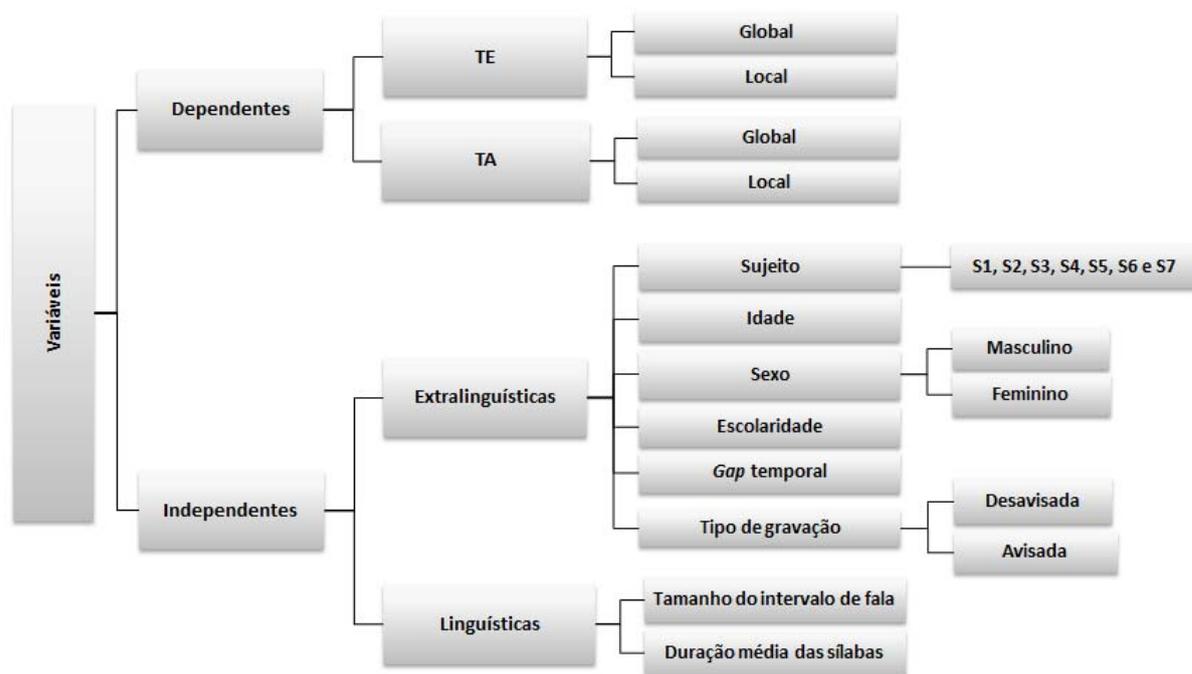
- a) desavisada – diálogo ao telefone celular entre locutores conhecidos ou que, ao menos, desenvolvem atividades ilícitas afins, gravado sem a ciência dos locutores que dele participam, sem rotina predefinida ou ingerência de terceiros, sendo obtido por meio de interceptação telefônica judicialmente autorizada;
- b) avisada – diálogo presencial entre um acusado e um ou mais peritos, em coleta técnica de padrão vocal para fins periciais, gravação ambiental nos moldes de uma entrevista semidirigida, na qual o locutor é solicitado a discorrer livremente sobre sucessivos temas de seu interesse ou propostos, sendo interpelado apenas ocasionalmente, estando todos os presentes cientes e cooperativos em relação à realização da gravação.

Cada tipo de material vocal tem propriedades particulares intrínsecas. No primeiro caso (interceptação telefônica), tem-se a inconcretude do interlocutor (não visualizado), turnos de fala via de regra mais curtos em razão do caráter funcional do sistema de comunicação (telefonia móvel) e um maior respeito à troca de turno típica dos diálogos. No segundo (gravação ambiental), há o incentivo por parte do

entrevistador para a exaustão das temáticas, o emprego auxiliar de formas não verbais de comunicação e a não incomum sobreposição de falas.

Na Figura 16 é apresentado um esquema com as variáveis (e respectivos fatores) admitidas neste estudo, destacando-se que as variáveis dependentes são verificadas global e localmente e que, conforme já relatado, às variáveis independentes-foco somam-se duas outras acessórias: sujeito e duração média das sílabas.

Figura 16 – Variáveis e fatores de variáveis admitidos no estudo



Fonte: A autora (2013).

Nota: As variáveis “Idade”, “Escolaridade”, “Gap temporal”, “Tamanho do intervalo de fala” e “Duração média das sílabas” não apresentam, no esquema acima, fatores associados em razão de serem escalares. Na definição das propriedades do *corpus* serão apresentados os valores mínimos e máximos de cada uma delas.

3.1.3 Objetivos

Assume-se como objetivo geral neste estudo a investigação do potencial individualizante (de indivíduo) da TE e da TA, com vistas a contribuir para a discussão acerca da efetividade da incorporação das medidas temporais de fala ao rol de parâmetros técnico-comparativos utilizados na realização da perícia de Comparação de Locutor.

Especificamente, objetiva-se:

- a) compor e implementar um método científico de mensuração (global e local) de TE e de TA, passível de ser integrado e dirigido à rotina em Fonética Forense, onde há, normalmente, limitações quanto à qualidade e quantidade de material de fala, além de um tempo de exposição linguística que se interpõe entre os áudios dos locutores a terem os perfis de voz e fala confrontados;
- b) investigar a variabilidade intrassujeito e intersujeitos de ambas as taxas temporais estudadas, a fim de explicitar o potencial individualizante da TE e da TA;
- c) observar, em cada um dos tipos de taxa (TE e TA), se há diferença significativa relacionada à forma de mensuração (global *versus* média das mensurações locais);
- d) verificar a correlação existente entre os valores de TE e de TA empregados pelos sujeitos pesquisados e as variáveis independentes idade, sexo, escolaridade, *gap* temporal e tipo de gravação;
- e) verificar a correlação existente entre os valores locais de TEs e de TAs empregados pelos sujeitos pesquisados e a variável independente tamanho do intervalo de fala;
- f) prever (mediante recursos estatísticos) o comportamento da TE e da TA em razão das variáveis independentes idade, sexo, escolaridade, *gap* temporal, tipo de gravação e tamanho do intervalo de fala;
- g) verificar a correlação existente entre as variáveis idade e tamanho do intervalo de fala (excepcionalmente referente ao número de sílabas presentes em um intervalo interpausal), a fim de descartar a possível interferência indireta deste na interpretação do comportamento da variável idade em relação à TA;
- h) verificar a correlação existente entre as variáveis duração média da sílaba e tamanho do intervalo de fala (excepcionalmente referente ao número de sílabas presentes em um intervalo interpausal), definindo se o aumento do número de sílabas do intervalo interpausal implica necessariamente na diminuição da respectiva duração média das sílabas, ou seja, se o fenômeno de encurtamento antecipatório mostra-se significativamente presente nos sujeitos pesquisados.

3.1.4 Hipóteses

O levantamento bibliográfico realizado permite que se estabeleçam as seguintes hipóteses de pesquisas:

- a) a TA, devido a sua maior estabilidade é, das taxas temporais de fala estudadas, a mais indicada para utilização no confronto de voz forense, sendo capaz de, em algum grau, distinguir falantes;
- b) não há diferença significativa entre os valores globais obtidos via divisão do número total de sílabas pela duração total do intervalo de fala dos obtidos via média das mensurações locais;
- c) quanto à idade, mantendo-se o controle da interferência da variável tamanho do intervalo de fala, conjectura-se que o sujeito de idade mais avançada produz fala com menores valores de TEs e TAs;
- d) quanto ao sexo, há a presunção de que os falantes com taxas mais altas são os do sexo masculino;
- e) quanto ao tamanho do intervalo de fala, prevê-se o emprego de valores de taxa (TE e TA) maiores em intervalos de fala com maiores durações, observando-se a ocorrência de progressiva redução na duração média das sílabas (encurtamento antecipatório).

As hipóteses relativas às variáveis escolaridade, *gap* temporal e tipo (ciência) de gravação, em razão de se desconhecer estudo prévio que as tenham abordado, foram definidas a partir das respectivas premissas empíricas:

- f) o indivíduo pouco escolarizado, menos influenciado pelo padrão culto da língua, exerceria menos controle sobre sua produção oral e, conseqüentemente, exibiria valores mais elevados de TEs e de TAs;
- g) o intervalo de tempo transcorrido entre as gravações do sujeito (desavisada e avisada) impactaria as TEs e as TAs, à medida que durante tal intercurso há avanço na idade do indivíduo e que é previsto o efeito da variável idade sobre as referidas taxas;
- h) variações estilísticas podem dever-se a diferenças nos elementos que compõem as situações comunicativas, entre eles a ciência ou não de estar sendo gravado, entendendo-se que essa é responsável, em algum grau, por um maior controle (por parte do locutor) sobre a fala.

3.2 AMOSTRA

3.2.1 Autorizações indispensáveis ao desenvolvimento do trabalho

Pretendendo-se a utilização dos dados de fala integrantes do banco de dados VOXDC, alocado na Seção de Perícias em Áudio e Imagens (SEPAI) do Departamento de Criminalística (DC) do Instituto-Geral de Perícias (IGP), órgão da Secretaria da Segurança Pública do Estado do Rio Grande do Sul (SSP – RS), solicitou-se autorização administrativa para tanto, o que foi concedido através do ofício cuja cópia é apresentada no já mencionado Anexo A.

Providenciou-se, ainda, a aprovação do projeto de pesquisa referente a este estudo junto ao Comitê de Ética em Pesquisa (CEP) da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), que deferiu a favor do proposto, segundo a cópia da carta de aprovação (OF.CEP-1107/10) apresentada no Anexo C.

3.2.2 Seleção dos sujeitos

Uma vez garantidas as autorizações indispensáveis ao desenvolvimento da pesquisa, procedeu-se à seleção dos sujeitos. No procedimento utilizou-se amostragem por conveniência, já que a estratificação encontrava-se limitada às características do conteúdo armazenado no banco de dados utilizado. Dessa forma, foram selecionados do banco de dados VOXDC, constituído então de amostras de fala de 61 indivíduos, todos os que atendiam aos critérios estabelecidos.

Formularam-se quatro critérios de inclusão dos sujeitos na pesquisa, devendo esses:

- a) integrar o banco de dados VOXDC, com Laudo Pericial de Comparação de Locutor positivo para o confronto realizado⁴²;
- b) ter informada a data de nascimento e as datas de gravação dos áudios a serem analisados, assim como a escolaridade (normalmente inquirida na ocasião da gravação avisada);

⁴² As informações sobre os locutores, as configurações técnicas empregadas no atendimento pericial e o resultado da perícia realizada encontram-se registradas originalmente no material de áudio colhido pelos peritos ou no próprio Laudo Pericial, de onde foram recolhidas para composição do relatório de conteúdo do referido banco de dados.

- c) ter disponível, na respectiva pasta digital do banco de dados, arquivos de áudio no formato *wave*, com resolução de 16 bits por amostra e taxa de amostragem associada à natureza da gravação (8000 Hz quando interceptação telefônica e 44100 Hz quando gravação ambiental), com generalizada inteligibilidade e qualidade acústica, nos quais conste fala espontânea em situação de diálogo advindo de sistemas de gravação e procedimento de coleta distintos (por regra: interceptações telefônicas judicialmente autorizadas e gravações ambientais tecnicamente produzidas);
- d) não apresentar sinais patológicos significativos relacionados à linguagem (compreensão e produção), bem como à audição periférica, sendo o comportamento do cedente avaliado informalmente pela acadêmica-autora⁴³ na ocasião do fornecimento do padrão sonoro.

As informações relativas a cada um dos sete sujeitos selecionados são apresentadas na Tabela 1, a seguir.

Tabela 1 – Dados dos sete sujeitos selecionados para participação no estudo

Sujeito	Sexo	Escolaridade (anos)	Idade na data mediana desavisada (anos)	Idade na data da avisada (anos)	Gap temporal entre os áudios (meses)
S1	Masculino	11	31	33	20
S2	Masculino	4	29	33	38
S3	Feminino	8	18	21	38
S4	Feminino	8	30	31	8
S5	Masculino	5	17	18	10
S6	Masculino	6	14	15	10
S7	Masculino	7	30	31	10

Fonte: A autora (2013).

Nota: Os valores mínimos e máximos de escolaridade, idade e *gap* temporal aparecem destacados por sombreamento.

3.3 ORGANIZAÇÃO DOS DADOS

3.3.1 Equipamentos e *softwares* utilizados

Seguem abaixo relacionados os *softwares* utilizados na execução do trabalho em tela, assim como, dos equipamentos utilizados, os que foram relevantes ao método empregado:

⁴³ Graduada em Fonoaudiologia, conforme currículo apresentado no ANEXO D.

- *softwares* de análise acústica Adobe Audition (versão 2.0) e Praat⁴⁴, (versão 5.3.23);
- *softwares* com recursos de análise estatística PASW Statistics (atualização do SPSS – *Statistical Package for the Social Sciences*, versão 18) e Microsoft Office Excel (versão 2003);
- fones de ouvido marca Roland (modelo RH-200) e Philips (modelo SBC HP 195);
- placa de captura de som marca Edirol (modelo UA-25);
- microfone cardioide marca Shure (modelo SM58);
- pedestal de mesa para microfone.

3.3.2 Montagem dos *corpora*

Inicialmente procedeu-se à organização do material de áudio de cada sujeito admitido na pesquisa, em pasta digital identificada conforme a procedência da gravação (desavisada ou avisada).

Quanto ao tamanho mínimo do *corpus* de cada sujeito, ressalta-se que a gravação consentida, por ser tecnicamente conduzida, normalmente contém o tamanho mínimo desejado para as análises pretendidas. Já os registros gravados desavisadamente tendem a ser, normalmente, quantitativamente restritos.

Entende-se como de fato preocupante apenas a escassez de material sonoro, uma vez que possíveis excessos são naturalmente dirimidos com a irremediável exclusão de trechos considerados metodologicamente inservíveis (as justificativas para exclusão de trechos de fala são arroladas adiante, na Subseção 3.3.4).

A fim de resolver os casos de limitação do material provindo de interceptação telefônica, situação encontrada em cinco dos sete sujeitos da pesquisa, optou-se pela justaposição de registros, procedimento que segue abaixo descrito.

3.3.3 Compilação das gravações desavisadas

Com o propósito de facilitar o manuseio dos áudios, os sujeitos que possuíam mais de uma gravação desavisada (interceptação telefônica) tiveram os sinais

⁴⁴ Disponível para *download* gratuito no sítio: <<http://www.praat.org>>. Acesso em: 23 set. 2012.

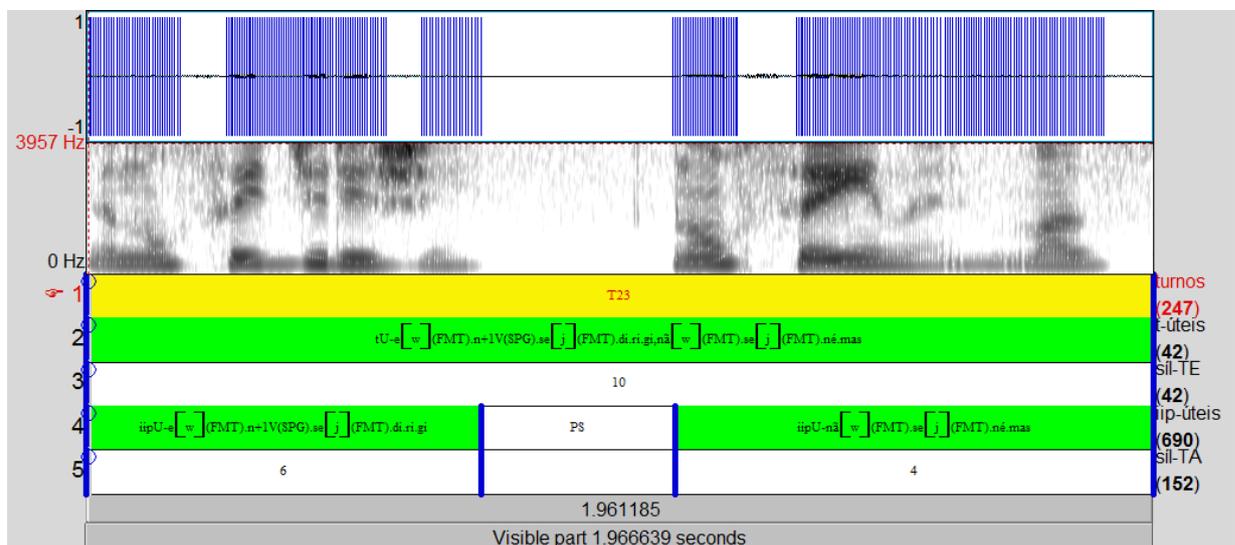
originais individuais justapostos, aproveitando-se a ocasião para exclusão dos tons de sinalização telefônica. Na compilação respeitou-se a ordem crescente de data de gravação, sendo o arquivo resultante, então único, salvo com a configuração proveniente do sistema de interceptação – áudio monoaural, no formato Windows PCM, com 8000 Hz de taxa de amostragem e 16 bits de resolução. Nesses casos, no cálculo da idade na gravação desavisada, como já mencionado, foi considerada a mediana das datas das ligações que compuseram o arquivo editado.

3.3.4 Segmentação e etiquetagem dos áudios

Os áudios que compreendem o *corpus* de cada um dos sujeitos da pesquisa, provindos da gravação desavisada e da gravação avisada, foram analisados utilizando-se na tarefa, além dos respectivos oscilogramas e espectrogramas, o recurso de vinculação de grade de texto disponibilizado pelo *software* Praat, através da opção “Annotate” e, em seguida, “To TextGrid” do menu dinâmico.

As informações depreendidas a partir do exame dos áudios (relativas aos locutores que participam no diálogo, ao conteúdo linguístico proferido pelo locutor-alvo e à qualificação do material para os fins pretendidos), assim como a quantificação do número de sílabas observadas no sinal útil como produzidas, foram gerenciadas em cinco camadas de etiquetagem verticalmente dispostas, como ilustrado na Figura 17.

Figura 17 – Exemplar da etiquetagem utilizada no procedimento do estudo



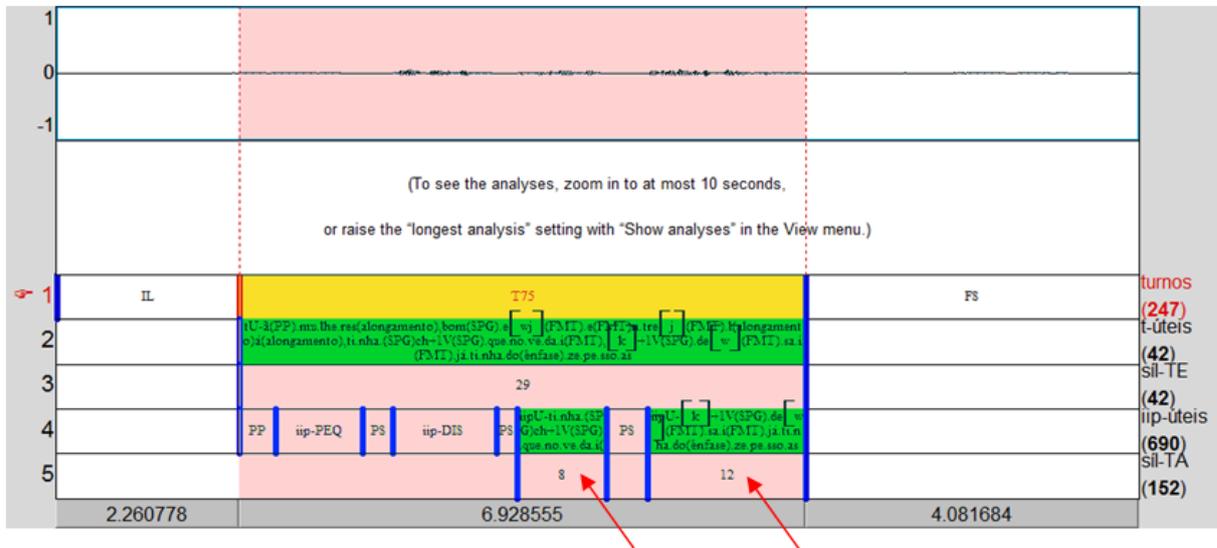
Fonte: A autora (2013).

Na Figura 17 observa-se, de cima para baixo, a imagem do sinal sonoro (oscilograma ou forma de onda), o respectivo espectrograma (em banda larga) e as cinco mencionadas camadas de etiquetagem, intituladas (da primeira à quinta) “turnos” (turnos de fala), “t-úteis” (turnos-úteis)⁴⁵, “síl-TE” (sílabas-TE), “iip-úteis” (intervalos interpausais-úteis) e “síl-TA” (sílabas-TA). Nessas, respectivamente, em linhas gerais: (a) se identifica os turnos úteis de fala do locutor-alvo, (b) procede-se à degravação e definição das sílabas constantes nos turnos de fala úteis (admitidos para cálculo da TE), (c) se explicita numericamente a quantidade de sílabas presentes em cada turno de fala útil, (d) se identifica os intervalos interpausais úteis (reproduzindo-se a relativa degravação e definição silábica exibida em “t-úteis”) e (e) se expõe numericamente a quantidade de sílabas presentes nos intervalos interpausais úteis (admitidos para cálculo da TA). A partir da imagem pode-se, ainda, contextualizar o trecho visualizado, relacionando-o com a integralidade do áudio, obtendo-se a indicação de que se trata, no caso, do 23º turno de fala (“T23”) de um dos dois registros de áudio de um dos sete sujeitos da pesquisa, em diálogo que compreende 247 turnos de fala (numeral exibido abaixo do rótulo “turnos”, na margem à direita da figura). Desses, considerando-se apenas os produzidos pelo locutor-alvo, 42 (numeral exibido abaixo do rótulo “t-úteis”) atendem às condições necessárias para admissão como turno passível de ser submetido ao cálculo de TE, estabelecidas na Subseção 3.2.2.

É importante salientar que a quantidade de intervalos interpausais úteis (dado apresentado abaixo do rótulo da camada “síl-TA”, no exemplo “152”) tende a ser consideravelmente maior do que a de turnos de fala úteis, uma vez que os turnos, internamente, podem apresentar diversas pausas. Na Figura 18 tem-se a ilustração dessa tendência, a partir da exibição de um turno de fala (“T75”) com dois intervalos interpausais úteis (indicados pelas setas sobrepostas).

⁴⁵ Na confecção das imagens exibidas neste documento, quando o conteúdo de fala apresentado na segunda camada de etiquetagem (“t-úteis”) mostrou-se ilegível, optou-se por assim mantê-lo em razão do comprometimento por parte da acadêmica-autora com a preservação da identidade das partes. Sendo necessária a explicitação do conteúdo proferido para a compreensão do descrito, não havendo violação de tal princípio, esse será reproduzido em nota, abaixo da respectiva figura.

Figura 18 – Exemplar de turno de fala com mais de um intervalo interpausal e da aplicação dos critérios de notação “IL”, “T” seguido de nº de ordem e “FS”



Fonte: A autora (2013).

Percebe-se, dessa forma, que cada uma das camadas acolheu informações de uma etapa do procedimento, sendo os respectivos critérios e condutas adotados a seguir detalhadamente descritos.

Como já mencionado em linhas gerais, em “turnos” realizou-se a identificação e delimitação dos turnos de fala dos locutores presentes no diálogo. Nessa camada, os turnos de fala do interlocutor foram rotulados como “IL” (interlocutor) e os nos quais se observa sobreposição de falas como “FS” (fala sobreposta). Já os turnos de fala do locutor-alvo foram sequencialmente identificados como “T1” (turno 1), “T2” (turno 2), Ilustrando parte do exposto, na Figura 18, recém-apresentada, visualiza-se forma de onda e respectivas camadas de etiquetagem de trecho de diálogo no qual ocorrem, da esquerda para direita, turno de fala interlocutora (“IL”), turno útil de fala exclusiva do locutor-alvo (“T75”) e trecho com sobreposição de falas (“FS”).

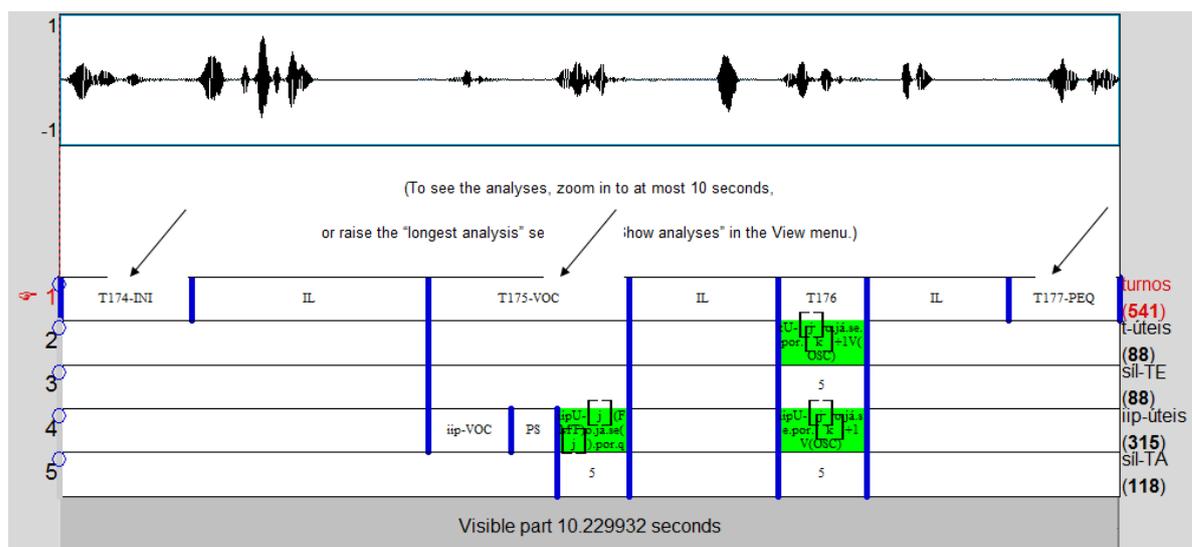
Os turnos de fala exclusiva do locutor-alvo foram, à medida que identificados, avaliados qualitativamente, notando-se adicionalmente após o número de série, sendo o caso, o código representativo da justificativa para exclusão do respectivo material das demais etapas de etiquetagem relativas ao cálculo da TE, primeiro tipo de taxa almejado no procedimento. Foram identificadas cinco situações específicas, além da ocorrência de fala sobreposta mencionada, que atribuem caráter inservível

ao material de fala a ser submetido ao cálculo da TE. Abaixo segue a caracterização de cada uma delas, bem como o rótulo utilizado no apontamento da exclusão:

- a) “INI” (ininteligível) – turnos de fala total ou parcialmente ininteligíveis ou com qualidade acústica que impede o estabelecimento da existência de algum segmento ou do(s) limite(s) do(s) segmento(s) fronteira(s);
- b) “VOC” (vocalização) – turnos de fala onde ocorre(m) vocalização(ões) não linguísticas como riso, tosse, pigarro ou clique labial prévio ao início da emissão de fala;
- c) “PEQ” (pequeno) – turnos de fala com número de sílabas inferior ao mínimo recomendável na literatura. Como já discutido em subseção anterior, a utilização de intervalos de fala curtos no cômputo das taxas temporais é desaconselhada, pois impede que ocorra a neutralização das diferenças entre os moldes silábicos das palavras examinadas. Assim, em concordância com Jessen (2007), foram descartados neste estudo os intervalos de fala com menos de quatro sílabas.

Na Figura 19 é exibido exemplar (forma de onda e respectivas camadas de etiquetagem) de trecho de diálogo no qual ocorrem, nas localizações indicadas pelas setas sobrepostas, turnos de fala enquadrados nos três critérios de descarte de material mais encontrados nos áudios trabalhados (“INI”, “VOC” e “PEQ”).

Figura 19 – Exemplos de turnos de fala notados como “INI”, “VOC” e “PEQ”

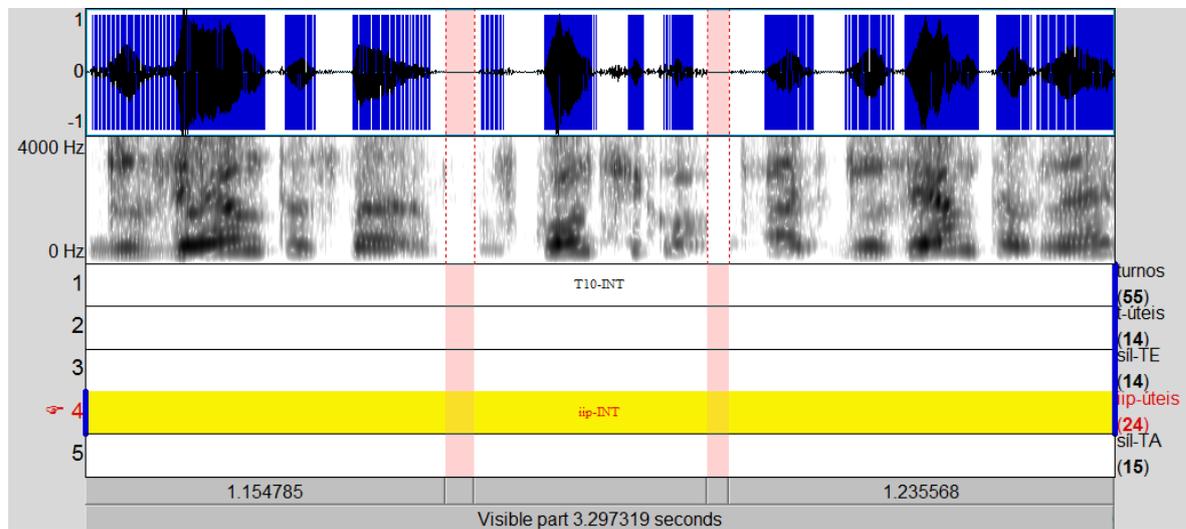


Fonte: A autora (2013).

- d) “INT” (interrompido) – turnos de fala incompletos devido à falha na transmissão telefônica, evento não incomum em se tratando de telefonia móvel, tecnologia utilizada na gravação desavisada.

Na Figura 20 é exibido exemplar (forma de onda, respectivo espectrograma em banda larga e camadas de etiquetagem) de parte de um dos turnos de fala de um dos locutores-alvo no qual são observadas (nas áreas selecionadas, limitadas pelas linhas tracejadas verticais) interrupções na transmissão do sinal, tendo-se como vestígio da intercorrência a ausência de manifestação sonora no trecho.

Figura 20 – Exemplar de turno de fala notado como “INT”



Fonte: A autora (2013).

- e) “ÉTI” (ético) – turnos de fala com impedimento ético à divulgação, ou seja, que apresentam conteúdo capaz de levar à identificação civil do(s) locutor(es) do diálogo ou de outro por eles referido. A autorização para uso do conteúdo armazenado no banco de dados VOXDC, constante no Ofício n.º 1815/2009-DG/IGP (Anexo A), implica a preservação da identidade dos locutores nos áudios questionados, bem como dos cedentes (fornecedores de material) e demais envolvidos no inquérito ou processo, sendo indevida por essa razão, inclusive, a exemplificação do que seria um intervalo assim qualificado.

O número de turnos de fala úteis identificados na amostra desavisada e avisada de cada um dos sete sujeitos pesquisados é apresentado na Tabela 2, onde constam também as respectivas durações. Na tabela é exibido, ainda, o número

total de turnos de fala úteis e o tempo total do material analisado, em cada um dos tipos de gravação.

Tabela 2 – Quantidade e duração dos turnos de fala úteis dos sujeitos do estudo

Sujeito	Nº de turnos de fala		Duração	
	Desavisada	Avisada	Desavisada	Avisada
S1	52	40	129.713	272.038
S2	88	12	180.880	101.260
S3	37	42	67.555	215.707
S4	14	13	27.265	56.444
S5	45	61	75.905	220.678
S6	31	64	48.235	176.988
S7	19	21	33.263	84.643
Total	286	253	562.816	1.127.758

Fonte: A autora (2013).

Nota: Na tabela, a expressão da duração total se dá em min.s.ms.

Os valores mínimos e máximos de turnos de fala e de duração aparecem destacados por sombreamento.

No que se refere a tamanho mínimo de *corpus* considerado para cálculo de taxa temporal de fala, encontrou-se referência na literatura a 17 intervalos interpausais (KENDAL, 2009) e a 20s de fala (CAO e WANG, 2011). Como visualizado na Tabela 2, dois dos sujeitos apresentam número total de turnos de fala inferiores ao indicado (S2 com o total de 12 turnos na gravação avisada e S4 com o total de 14 e 13 turnos, respectivamente, na gravação desavisada e avisada). Contudo, verificou-se que, nos três casos, a duração total de material de fala mostrava-se superior aos 20s apregoados.

Quanto ao estabelecimento das alocações das fronteiras dos intervalos, realizado de forma grosseira na camada “turnos”, mas definido com precisão nas camadas “síl-TE” (sílabas-TE) e “síl-TA” (sílabas-TA), considerou-se, alternada e oportunamente, a emissão auditivamente percebida, a forma de onda e a imagem espectrográfica do trecho de áudio sob exame.

Na identificação do que seria o local exato do início do primeiro segmento dos intervalos de fala, o mesmo se aplicando, com a devida adaptação, ao final do último segmento da cadeia de fala, foram encontradas situações técnicas conflitivas que levaram à postulação de critério único a ser considerado na marcação, a saber: o início do vestígio acústico da primeira fonte da(s) que compõe(m) o fone em questão, seja ela uma fonte de ruído ou de sonoridade. Tal critério não pode ser aplicado somente nas plosivas e africadas sonoras, que têm como primeiro vestígio

a barra de sonoridade, de difícil demarcação em áudios de baixa relação sinal/ruído⁴⁶ como os admitidos neste estudo, tendo sido considerado, nesse caso, o início da soltura do segmento (espícula vertical).

Nas fricativas considerou-se a zona de ruído visualizada (ressalvando-se a limitação em frequência do sinal interceptado), nas nasais o vestígio do vozeamento e nas líquidas o de soltura ou de fricção (dependendo da forma veiculada). Nas vogais, foi referência predominante a regularidade da forma de onda e, acessória ou substitutiva (na falta da primeira), os formantes do som (ou pistas remanescentes desses), visualizados no espectrograma.

As dificuldades encontradas na identificação do fim do último segmento dos intervalos de fala restringiram-se às vogais amortecidas e às fricativas surdas, nos casos de baixa dissociação entre o sinal de voz e o decorrente do ambiente.

Destaca-se que, percebendo-se prejudicada a determinação segura dos limites dos segmentos fronteiros, o intervalo de fala (turno de fala ou intervalo interpausal) foi descartado (como “INI”, posto que acusticamente ininteligível) do rol de úteis para fins de cálculo de taxa.

Ainda na descrição do procedimento metodológico, na camada intitulada “t-úteis” (turnos-úteis), é feita a exposição do conteúdo linguístico dos então denominados turnos úteis de fala do locutor-alvo. A transcrição da fala é antecedida pelo código “tU” (turno-útil), para que seja acionada automaticamente a coloração do intervalo (item “File” do menu da janela de edição da grade de texto, opção “Preferences” e, em seguida, “Paint intervals green whose label”). Cada um dos turnos de fala assim destacados foi submetido à transcrição alfabética da versão implementada da fala⁴⁷, recorrendo-se à transcrição fonética local e apontamentos específicos quando a informação fosse capaz de interferir na especificação quantitativa de sílabas existentes no turno de fala sob exame. A divisão silábica por meio de ponto foi adotada a fim de facilitar o trabalho de contagem, cujo resultado é apresentado na camada imediatamente seguinte, a “síl-TE” (sílabas-TE). Na Figura 21 tem-se uma amostra da transcrição flexibilizada empregada neste estudo⁴⁸,

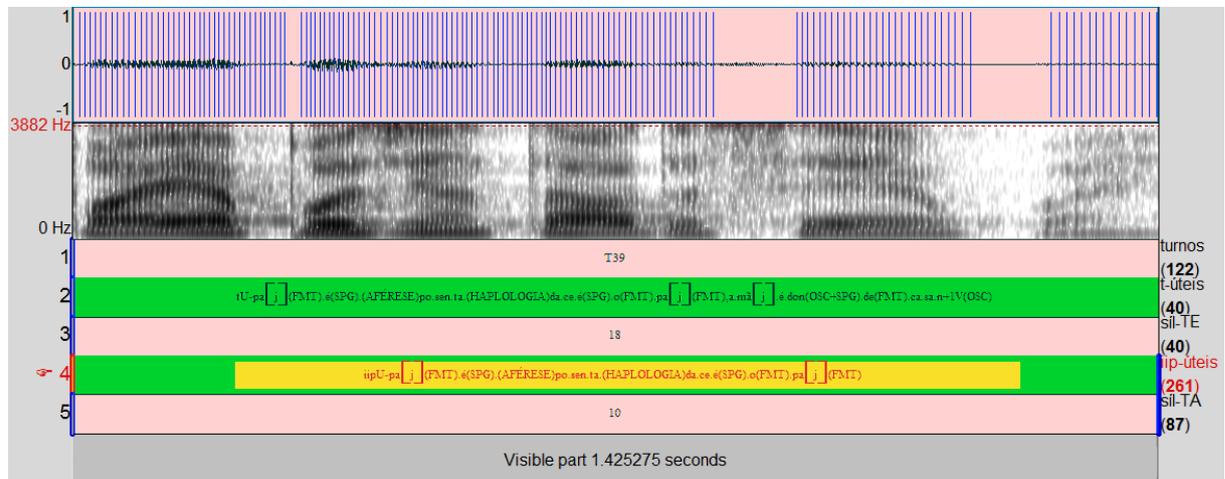
⁴⁶ A relação sinal/ ruído quantifica a qualidade do sinal. Em registros de fala, compara o valor da grandeza que desejamos medir (sinal vocal) com o da grandeza indesejável (ruído ambiental).

⁴⁷ A exemplo a sentença <lava ali mesmo> que, produzida como [lavalí'mez], foi notada como “la.va.li.mes”.

⁴⁸ A flexibilização do formato da transcrição foi necessária, dada a extensão do *corpus* (composto de 1.287 intervalos de fala).

sendo possível observar a notação da aplicação de processos (aférese e haplogogia), assim como da realização de semivogais no trecho.

Figura 21 – Exemplar de apontamentos adicionais e da transcrição fonética local empregada



Fonte: A autora (2013).

Nota: Consta na gravação “pa[j](FMT).é(SPG).(AFÉRESE)po.sen.ta.(HAPLOLOGIA)da.ce.é(SPG).o(FMT).pa[j](FMT)”, para a forma ortográfica <pai é aposentado da CEEE, o pai>.

Entre parênteses consta a fonte de informação que fundamentou a notação, sendo “FMT” relativo à disposição ou aos valores dos formantes e “SPG” à vestígios outros percebidos no espectrograma.

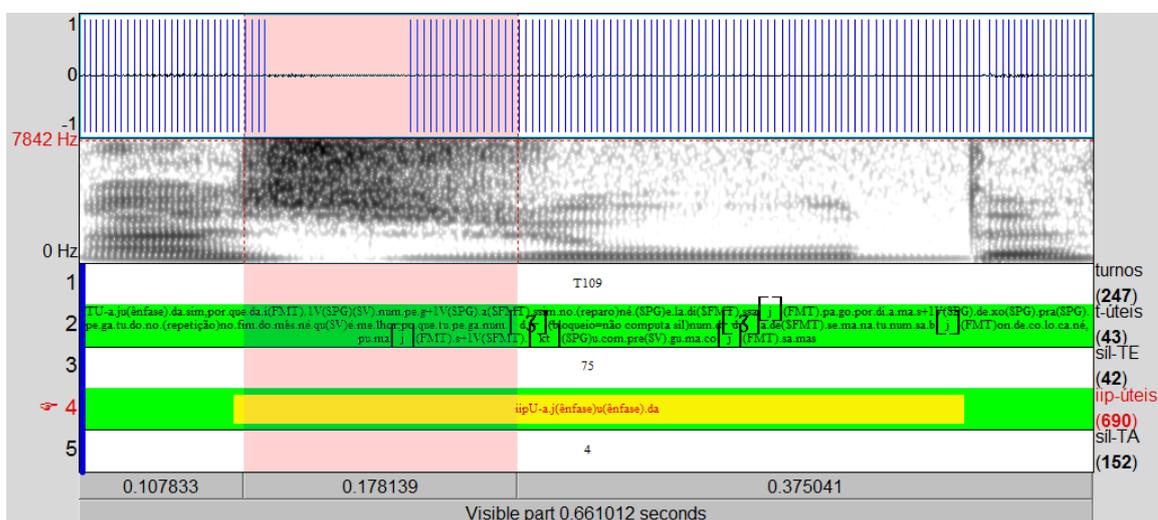
Ressalta-se que, quando se pretende um cálculo de taxa temporal a partir do nível fonético de fala, é imprescindível a identificação da aplicação de processos intra e interpalavras, uma vez que esses são capazes de interferir (por acréscimo ou supressão) no número total de sílabas do intervalo de fala. Com relação aos processos de redução, tem-se: próteses (acrécimo de segmento ou sílaba no início da palavra, como em “arrodeio”); epênteses (acrécimo de vogal, normalmente /i/, no meio da palavra, como em “p[i]neu”); paragoges (acrécimo de segmento ou sílaba no final da palavra, como em “narizi”); aféreses (supressão de segmento ou sílaba no início da palavra, como em “rancá”); síncofes (supressão de segmento ou sílaba no meio da palavra, como em “xícra”); apócopes (supressão de segmento ou sílaba no fim da palavra, como em “vamo”); elisões (apagamento de vogal postônica final, quando o vocábulo seguinte começa por vogal diferente, como em “desdás”); degeminações (fusão de duas vogais análogas, como em “casamarela”); ditongações (semivocalização de /i, u/ quando adjacente a outra vogal, como em “mini[nwaw]to”) e haplogias (queda total de uma sílaba quando duas sílabas semelhantes estão adjacentes, como em “perdicasa” para <perto de casa>).

É importante reiterar que as sílabas fonéticas, auditivamente percebidas e confirmadas à inspeção acústica do sinal, tiveram sua presença definida (manual e isoladamente)⁴⁹ em razão da existência no sinal de ao menos um dos vestígios vocálicos.

A diferenciação entre os vocoides silábicos [i, u] e os correspondentes não silábicos [j, w], os últimos não computáveis para fins de cálculo das taxas estudadas, foi estabelecida a partir da observação de trecho estável na manifestação do segundo formante do segmento (F2), critério indicado por Padgett (2008)⁵⁰.

Os alongamentos consonantais de fricativa (conforme ilustrado na Figura 22, na área selecionada) foram acostados ao núcleo silábico antecedente ou seguinte, de acordo com a emissão, não tendo sido computados como unidade na contagem. Já os alongamentos vocálicos foram interpretados como evento único (uma sílaba) quando imersos no texto, em visível ênfase (como ilustrado na Figura 23, na área selecionada), e como pausa preenchida quando utilizados como partícula preenchedora, associada à situação de hesitação (como ilustrado na Figura 24, na área selecionada).

Figura 22 – Exemplar de alongamento consonantal, na emissão de [ʒ]

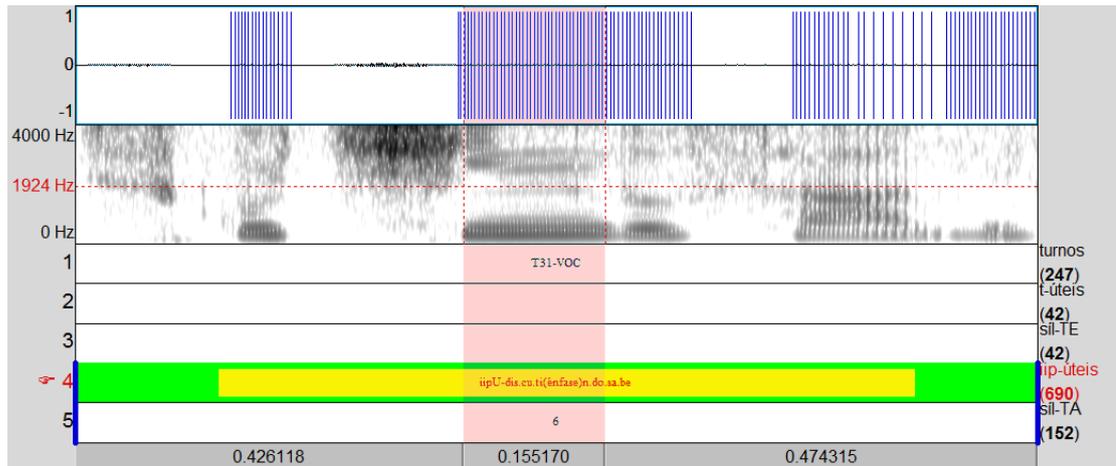


Fonte: A autora (2013).

⁴⁹ Por se tratar de fala verdadeiramente naturalística, previu-se que o grande número de fenômenos adversos à eficiente detecção silábica automatizada conduziria a um alto índice de erros de marcação (tanto falsos positivos quanto falsos negativos) e, por consequência, a um exaustivo trabalho de conferência, optando-se então pela detecção manual dos vocoides silábicos.

⁵⁰ Disponível em: <<http://people.ucsc.edu/~padgett/locker/glides07.pdf>>. Acesso em: 21 set. 2012.

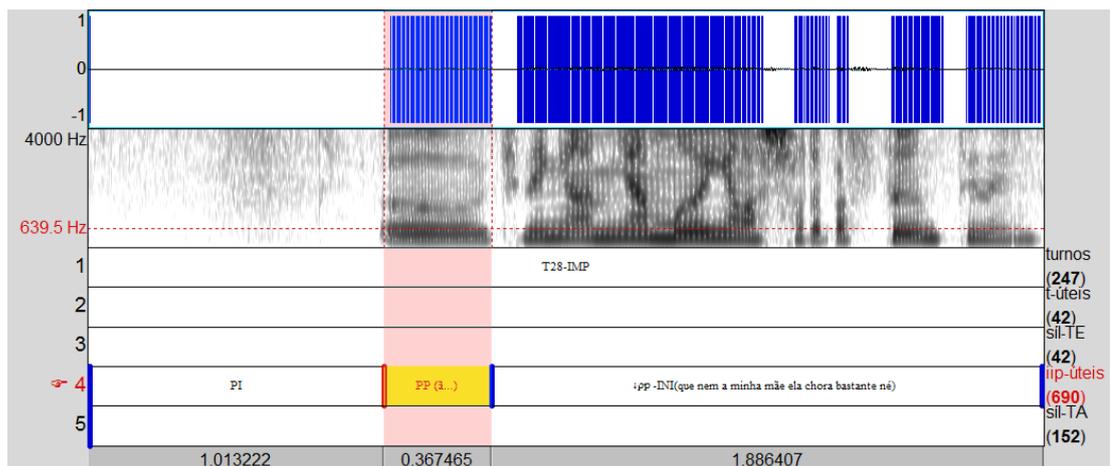
Figura 23 – Exemplo de alongamento vocálico em situação de ênfase



Fonte: A autora (2013).

Nota: Consta na gravação “dis.cu.ti(ênfase)n.do.sa.be”.

Figura 24 – Exemplo de alongamento vocálico em pausa preenchida (“ã...”)



Fonte: A autora (2013).

Adiante na apresentação do procedimento, na camada identificada como “síl-TE” (sílabas-TE) consta, no intervalo relativo a cada um dos turnos de fala úteis, o numeral equivalente ao número de sílabas produzidas, identificadas na gravação da camada precedente.

Em “iip-úteis” (intervalos interpausais-úteis), local de início do preparo do material a ser utilizado no cálculo da TA, assim como com a TE, fez-se necessário primeiramente o descarte dos intervalos interpausais não úteis. Na tarefa foram examinados amiúde todos os turnos de fala úteis visando localizar e delimitar as pausas silenciosas, notadas com “PS” (pausa silenciosa), cujo limiar para fins deste estudo é 130ms (conforme Subseção 2.1 do Capítulo 2); as pausas inspiratórias,

notadas com “PI” (pausa inspiratória)⁵¹, enquadrável como pausa silenciosa, mas subcategorizada em razão do ruído acústico típico que tem associado, e as pausas preenchidas, notadas com “PP” (pausa preenchida), atribuídas à hesitação.

Os intervalos interpausais intraturnos, evidenciados após o apontamento das pausas descritas, foram avaliados qualitativamente e devidamente rotulados, utilizando-se tanto o código “iipU” (intervalo interpausal útil) para significar intervalos interpausais úteis ao cálculo da TA quanto os códigos “iip-INI” (intervalo interpausal-ininteligível), “iip-VOC (intervalo interpausal-(com)vocalização)”, “iip-PEQ (intervalo interpausal-pequeno)”, “iip-INT” (intervalo interpausal-interrompido), “iip-ÉTI” (intervalo interpausal-(não)ético), para notar os intervalos enquadrados em um dos cinco critérios aplicados no descarte dos turnos de fala, e “iip-DIS” (intervalo interpausal-disfluente), para notar os intervalos nos quais são observados evento de disfluência. Os intervalos de fala identificados por “iip-U” (intervalo interpausal-útil), e somente esses, acolheram a transcrição correspondente, parcial da já expressa na camada “t-úteis” (turnos-úteis).

Na última camada, “síl-TA” (sílabas-TA), é feito o apontamento numérico do total de sílabas presentes em cada intervalo interpausal considerado útil, o que se baseou, como no caso da TE, na divisão silábica constante na camada imediatamente antecedente, a qual armazena o conteúdo degravado.

Como dito anteriormente, o cálculo da TA considerou apenas a fala fluente, sendo tal posicionamento concordante com o de Dancovičová (1997, 2001), Jessen (2007) e Jacewicz et al. (2009, 2010). A justificativa para o descarte de intervalos com disfluência⁵² reside no fato de não haver consenso quanto à interpretação dada aos eventos assim caracterizados.

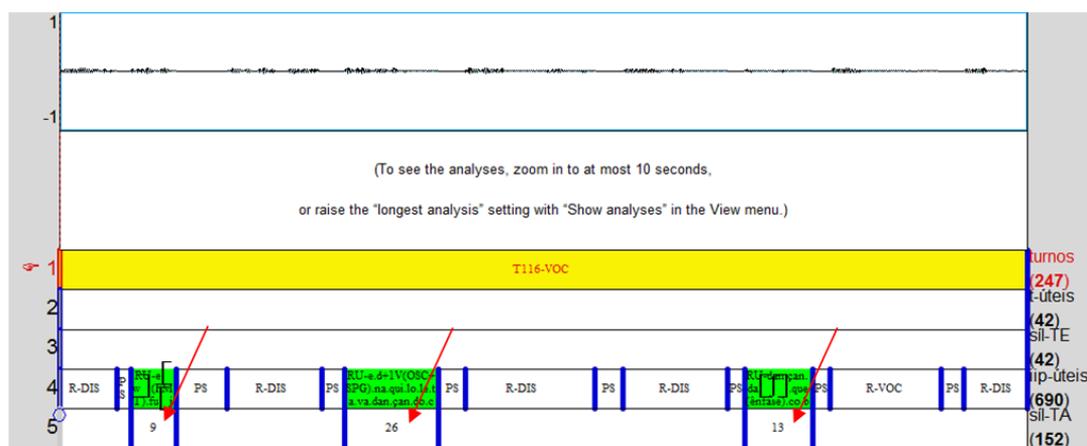
Na literatura, os eventos de disfluência são computados como unidade articulatória ordinária (sílabas), como pausa ou são simplesmente ignorados. Considerando que a TA propõe-se a avaliar o tempo empregado na integralização do gesto articulatório, seria um contrassenso admitir como produção de fala útil para fins de cálculo aquela prejudicada em termos de planejamento motor. Contudo, em se tratando de TE, por essa verificar o uso do tempo na fala de modo mais abrangente, contemplando questões cognitivas e linguísticas, essa foi admitida.

⁵¹ No material de fala útil dos sete sujeitos pesquisados, todas as pausas inspiratórias apresentaram duração superior à 130ms.

⁵² Robb et al. (2004) relatam que aproximadamente 6% do enunciado por falantes normais é disfluente.

Ressalta-se que um turno de fala definido como inservível, compreendendo mais de um intervalo interpausal, pode, desde que apenas parcialmente prejudicado, conter intervalo(s) interpausal(is) útil(eis), passível(is) de ser(em) utilizado(s) no cálculo da TA. Exemplo dessa situação é apresentado na Figura 25. Nela, tem-se a ilustração de um turno de fala com vocalização não linguística (“T116”), classificado (e descartado) como “VOC” (vocalização), no qual foram identificados três intervalos interpausais úteis (indicados pelas setas sobrepostas).

Figura 25 – Exemplo de descarte de turno de fala com aproveitamento de intervalos interpausais



Fonte: A autora (2013).

O número de intervalos interpausais úteis identificados na amostra desavisada e avisada de cada um dos sete sujeitos pesquisados é apresentado na Tabela 3, onde constam também as respectivas durações. Na tabela é exibido, ainda, o número total de intervalos interpausais úteis e o tempo total do material analisado, em cada um dos tipos de gravação.

Tabela 3 – Quantidade e duração total dos intervalos interpausais úteis dos sujeitos do estudo

Sujeito	Nº total de intervalos interpausais		Duração total	
	Desavisada	Avisada	Desavisada	Avisada
S1	60	87	87.534	122.704
S2	118	43	168.126	91.328
S3	23	152	24.517	263.350
S4	15	41	25.558	73.278
S5	34	43	46.616	69.764
S6	27	68	27.467	97.173
S7	12	25	15.728	30.047
Total	289	459	395.546	747.644

Fonte: A autora (2013).

Nota: Na tabela, a expressão da duração total se dá em s.ms.

Os valores mínimos e máximos de intervalos interpausais e de duração aparecem destacados, por sombreamento.

Observando os dados apresentados na Tabela 3 percebe-se que S7 possui apenas 12 turnos de fala na gravação desavisada e que a soma da duração dos mesmos não alcança o limite mínimo estabelecido como referência (20s), comprometendo o posicionamento assumido relativo ao tamanho e à duração total mínima para os turnos de fala. Apesar disso, dado o número já reduzido de sujeitos na pesquisa, optou-se por manter S7 como participante, tendo-se o cuidado de atentar para possíveis interferências provocadas pela restrição de material.

Sendo prevista a realização do cálculo das taxas através do *software* Excel, foi necessário transportar para o mesmo as informações relativas à duração dos intervalos de fala e respectivas quantidades de sílabas. A tarefa implicou a extração das camadas de etiquetagem “síl-TE” (sílabas-TE) e “síl-TA” (sílabas-TA) para a janela de objetos do Praat e a geração das respectivas listas de dados, as quais continham o tempo mínimo (correspondente ao início de cada intervalo), o tempo máximo (correspondente ao final de cada intervalo) e o texto digitado (equivalente ao número de sílabas contido em cada intervalo).

Foram então calculadas as durações individuais dos turnos de fala e intervalos interpausais, o que possibilitou o cálculo da TE e da TA (global, local nos 1.287 intervalos de fala e local média) das gravações desavisada e avisada de cada um dos sete sujeitos do estudo, referindo-se a mensuração global à divisão do resultado da soma do número de sílabas em todos os turnos de fala ou intervalos interpausais pelo resultado da soma da duração dos mesmos e a mensuração local média à média aritmética das taxas individualmente obtidas através da divisão do número de sílabas em cada intervalo de fala pela respectiva duração, como no esquema da Figura 26.

Figura 26 – Esquema da forma de cálculo global e local médio das taxas do estudo

$$\begin{aligned} \text{TE e TA} \\ \text{global} &= \frac{(\text{n}^\circ \text{ de sílabas no } t_1 \text{ ou iip}_1 + \text{n}^\circ \text{ de sílabas no } t_2 \text{ ou iip}_2 + \dots + \text{n}^\circ \text{ de sílabas no } t_n \text{ ou iip}_n)}{(\text{duração do } t_1 \text{ ou iip}_1 + \text{duração do } t_2 \text{ ou iip}_2 + \dots + \text{duração do } t_n \text{ ou iip}_n)} \\ \\ \text{TE e TA} \\ \text{local média} &= \frac{\frac{\text{n}^\circ \text{ de sílabas no } t_1 \text{ ou iip}_1}{\text{duração do } t_1 \text{ ou iip}_1} + \frac{\text{n}^\circ \text{ de sílabas no } t_2 \text{ ou iip}_2}{\text{duração do } t_2 \text{ ou iip}_2} + \dots + \frac{\text{n}^\circ \text{ de sílabas no } t_n \text{ ou iip}_n}{\text{duração do } t_n \text{ ou iip}_n}}{n} \end{aligned}$$

Fonte: A autora (2013).

Nota: Na figura, “t” refere turno de fala e “iip” intervalo interpausal, enquanto “n” representa o Último intervalo de fala da série.

3.4 INSTRUMENTOS DE ANÁLISE

3.4.1 Procedimento de verificação acústica

A análise acústica segmental e suprasegmental neste estudo contou com os recursos oferecidos pelos *softwares* Adobe Audition, especialmente utilizado na oitiva preliminar, seleção e compilação dos áudios, e Praat, utilizado nas demais etapas de manejo do *corpus*, como a segmentação e etiquetagem dos registros.

O sinal foi examinado no domínio do tempo, sob a perspectiva de oscilogramas e de espectrogramas em banda larga (*window length* de 5ms), sendo que sobre os primeiros estavam costumeiramente sobrepostos os pulsos glóticos (recurso da opção “Pulses” do *menu* da janela de edição do Praat) e, sobre os últimos, a curva de variação de f_0 (recurso da opção “Pitch” do *menu* da janela de edição do Praat), ambos desativados conforme a necessidade dos exames.

3.4.2 Tratamento estatístico dos dados

Com base nos valores globais e locais médios de TE e de TA, obtiveram-se as medidas descritivas relativas a ambas as espécies de taxas pesquisadas, tendo sido verificado o ajuste de tais variáveis ao modelo de normalidade.

Efetuuou-se o estudo de correlação entre as taxas globais e locais a fim de verificar se as formas de mensuração da TE e da TA entre si apresentam comportamento similar quanto ao sentido de variação, ou seja, se as variações têm relação direta ou inversa. Além disso, foi empregado um teste de comparação de médias pareadas, com o propósito de se identificar a existência de diferenças significativas entre os valores de taxa (TE e TA) gerados a partir da observação global ou local média.

Para verificar a significância da diferença entre as variabilidades relativas aos tipos de taxa e às formas de mensuração, utilizou-se o Teste F para comparação de variâncias entre duas amostras.

Quanto ao comportamento das taxas entre grupos por sexo (masculino e feminino) e por tipo de gravação (desavisada e avisada), procedeu-se à comparação de médias através de um Teste t para amostras independentes e de um Teste t para

amostras pareadas. Além disso, efetuou-se a Análise de Variância (ANOVA⁵³) quanto a essas duas variáveis (sexo e tipo de gravação), tendo sido seus resíduos (diferença entre o valor observado e o valor predito pelo modelo) avaliados quanto à normalidade (através do teste de Shapiro-Wilk e do gráfico Q-Q) e quanto à homogeneidade (através do teste de Levene e da análise do diagrama de dispersão dos valores preditos *versus* resíduos), importantes suposições para utilização de testes paramétricos e de ajustes de modelos lineares. Havendo violação dessas suposições, prevê-se a adoção dos seguintes procedimentos: para violação da normalidade, a transformação dos dados e, para violação da homogeneidade, a análise robusta de variância, através do Teste de Brown-Forsythe.

A correlação linear de Pearson foi utilizada para avaliar se o sentido de variação das taxas (TE e TA, global e local média) mostra-se direto ou inverso quando comparado ao sentido da variação das variáveis idade, escolaridade e *gap* temporal.

Uma exposição descritiva das taxas (com as quatro variáveis de resposta: TE global, TE local média, TA global e TA local média) foi efetuada, por tipo de gravação, para cada um dos sete sujeitos pesquisados. A existência de diferenças significativas intersujeitos (por tipo de taxa) foi verificada através de uma ANOVA. A análise dos resíduos dessa ANOVA considerou o empregado na análise similar precedente.

Já a variação intrassujeito foi definida com o emprego do coeficiente de correlação intraclasse (razão entre a variância intersujeitos e intrassujeito), ferramenta estatística utilizada para o estabelecimento da confiabilidade de medidas (LAUREANO, 2011), admitido como delator do potencial individualizante do parâmetro tipo de taxa e forma de mensuração.

Efetou-se o estudo baseado nas múltiplas tomadas locais de ambas as taxas (TE e TA), correlacionando a variável tamanho do intervalo de fala com os tempos de fala nos 1.287 intervalos, assim como idade e número de sílabas e número de sílabas e duração média das sílabas, restringindo-se a análise, nos dois últimos casos aos intervalos de fala pausa-excludente.

Foram ajustados modelos de regressão linear múltipla para as variáveis dependentes TE global, TE local média, TA global e TA local média, assumindo-se

⁵³ Simplificação para a expressão em inglês "*analysis of variance*".

como variáveis independentes idade, sexo, escolaridade, *gap* temporal e tipo de gravação. O critério para exclusão de variáveis independentes no modelo foi o nível descritivo do teste do parâmetro (p-valor⁵⁴). Para níveis descritivos do teste superiores a 0,25, a variável foi excluída do modelo, como recomenda Hosmer e Lemeshow (2000). O ajuste foi feito através do *software* PASW Statistics, com o método *backward*⁵⁵. A verificação das suposições do modelo de regressão foi realizada através dos mesmos procedimentos empregados na análise dos resíduos do modelo de análise de variância.

Para incluir na análise de regressão o tamanho do intervalo de fala como variável independente fez-se necessário a utilização das múltiplas tomadas locais de TE e de TA como variáveis de resposta. Os procedimentos de seleção de variáveis e de verificação de resíduos⁵⁶ foram os mesmos das análises antecedentes.

No próximo capítulo são apresentados os resultados obtidos a partir das análises realizadas, ao que segue a discussão acerca do significado dos achados.

⁵⁴ Neste estudo utiliza-se a expressão “p-valor” para referência à probabilidade de se obter uma diferença igual ou maior à existente entre a estimativa do valor observado na amostra e o valor suposto sob a hipótese nula. Já “nível de significância” será empregado para dizer da probabilidade de se cometer um erro, rejeitando a hipótese nula quando esta é verdadeira.

⁵⁵ Método em que todas as variáveis são incluídas na análise e posteriormente excluídas, uma a uma, a partir da análise do p-valor, até que se obtenha um modelo onde todas as variáveis inclusas possuam no máximo p-valor igual a 0,25.

⁵⁶ A inferência estatística no modelo de regressão linear baseia-se em determinados pressupostos, entre eles, como mencionado, o de distribuição normal, sendo a análise de resíduos uma das técnicas utilizadas na detecção de violações desses pressupostos.

4 RESULTADOS

Este capítulo divide-se em duas seções. A primeira delas (“4.1”) restringe-se exclusivamente à apresentação dos resultados obtidos na análise quantitativa de dados realizada⁵⁷, constando somente na segunda (“4.2”) a discussão linguística sobre o significado dos achados e as possíveis justificativas para os comportamentos encontrados, especialmente os não previstos.

Na mencionada exposição dos resultados (Seção 4.1), são feitas referências a termos comuns à área de estatística, mas pouco usuais fora dela. Esses termos, quando não explicados no próprio texto ou em apontamento acessório (como nota de rodapé ou legenda), foram incluídos e conceituados no glossário disponibilizado ao final do presente trabalho, para eventual consulta por parte do leitor.

4.1 APRESENTAÇÃO DOS RESULTADOS OBTIDOS

A observação local das taxas estudadas, dada a necessidade de valor único de saída para desenvolvimento da maior parte das análises pretendidas, é prevalentemente representada neste estudo através da média aritmética das tomadas individuais (conduta também adotada por Cao e Wang, 2011; Jessen, 2007⁵⁸; Künzel, 1997).

A inclusão da variável tamanho do intervalo de fala (tempo individual de cada um dos 1.287 intervalos de fala pesquisados), contudo, demanda a consideração não da mencionada média das tomadas locais, mas sim das observações múltiplas, individualmente. Por essa razão, são apresentadas, suplementarmente, algumas correlações e uma regressão linear baseadas nas tomadas múltiplas de ambos os tipos de taxa (TE e TA), compreendendo 539 observações no caso dos turnos de fala (intervalo de fala considerado no cálculo da TE) e 748 observações no caso dos intervalos interpausais (intervalo de fala considerado no cálculo da TA).

⁵⁷ Os conjuntos de comandos relativos às análises realizadas, com menção à tabela, figura ou subseção correspondente, são elencados no APÊNDICE B.

⁵⁸ Jessen (2007) enquadra a média das múltiplas tomadas locais como uma variação da mensuração global. Diferentemente, no presente estudo, assume-se que, tendo sido a taxa calculada individualmente a partir de cada um dos turnos de fala (no caso da TE) ou de cada um dos intervalos interpausais (no caso da TA), mesmo que posteriormente se realize a média dessas, tal versão simplificada de taxa representa um dado pertinente à mensuração local. Assim, tem-se neste estudo uma forma de mensuração global e duas formas de mensuração locais (a local média e as múltiplas tomadas locais).

Dessa forma, os resultados obtidos com o tratamento estatístico empreendido serão apresentados em dois blocos principais, a saber: “4.1.1”, onde a mensuração local é representada pela média das múltiplas tomadas locais (configurando a chamada TE ou TA local média), e “4.1.2”, onde a mensuração local é representada pelas múltiplas tomadas locais, isoladamente.

4.1.1 Análises baseadas nas TEs e TAs globais e TEs e TAs locais médias

4.1.1.1 Medidas relativas à tendência central e variabilidade dos dados

A fim de descrever o *corpus* no que se refere aos valores de taxas obtidos, efetuou-se a análise univariada, contando a mesma com os tradicionais descritores de tendência central e variabilidade, além dos relativos à normalidade dos dados, aqui referentes às quatro variáveis de resposta admitidas – TE global, TE local média, TA global e TA local média.

O resumo estatístico dos dados (exibido na Tabela 4) considerou o grupo completo de sujeitos pesquisados, constituído de sete falantes, cinco do sexo masculino e dois do sexo feminino, contando cada um deles com dois registros não contemporâneos de áudio, um obtido sem que os locutores tivessem ciência da gravação (gravação desavisada) e o outro de gravação sabida e consentida (gravação avisada).

Tabela 4 – Resumo estatístico relativo às quatro variáveis de resposta do estudo

Medida	TE global	TE local média	TA global	TA local média
Média	5,27 (0,196) ¹	5,47 (0,194) ¹	6,19 (0,142) ¹	6,20 (0,141) ¹
Variância	0,539	0,526	0,284	0,279
Desvio padrão ²	0,734	0,725	0,533	0,528
Mínimo	4,21	4,43	5,38	5,31
Máximo	6,51	6,61	6,93	6,99
Amplitude	2,30	2,18	1,55	1,68
Assimetria	0,090 (0,597) ¹	0,089 (0,597) ¹	-0,231 (0,597) ¹	-0,166 (0,597) ¹
Curtose	-1,217 (1,154) ¹	-1,310 (1,154) ¹	-1,090 (1,154) ¹	-1,004 (1,154) ¹
Teste de normalidade*	0,953 (0,607) ³	0,937 (0,383) ³	0,934 (0,347) ³	0,962 (0,754) ³

* Teste de Shapiro-Wilk, com 14 graus de liberdade.

¹ Desvio padrão da média.

² Desvio padrão amostral.

³ Estatística (p-valor).

Fonte: A autora (2013).

A avaliação dos dados apresentados no resumo estatístico da Tabela 4 permite a constatação de que as médias da TE são menores do que as da TA, condição prevista na literatura e atribuída logicamente à ampliação da duração do intervalo de fala provocada pelas pausas e pelos eventos de disfluência, presentes exclusivamente no material de fala considerado para cálculo da TE.

De acordo com os valores obtidos para as medidas “Variância”⁵⁹, “Desvio padrão” e “Amplitude” (intervalo), a TE apresenta variabilidade destacadamente maior do que a TA, embora a diferença entre os tipos de taxa nas formas de mensuração análogas, segundo o Teste F realizado, não possa ser considerada significativa (razão entre a TE global e a TA global: $F=0,284$; $GI\ 1= 13$; $GI\ 2= 13$ e $p= 0,539$ e entre a TE local média e a TA local média: $F=0,279$; $GI\ 1= 13$; $GI\ 2= 13$ e $p= 0,526$). A medida “Amplitude”, que tem maiores valores associados à TE, corrobora a maior variabilidade encontrada neste tipo de taxa.

Depreende-se, também, que tanto na TE quanto na TA são menos variáveis os resultados obtidos a partir da média das mensurações locais se comparados aos obtidos por meio da mensuração global. De acordo com um novo Teste F, no entanto, tal diferença (existente entre as formas de mensuração) não pode ser considerada estatisticamente significativa (razão entre a TE global e a TE local média: $F=1,025$; $GI1=13$; $GI2=13$ e $p=0,483$ e entre TA global e a TA local média: $F=1,018$; $GI1=13$; $GI2=13$ e $p=0,487$).

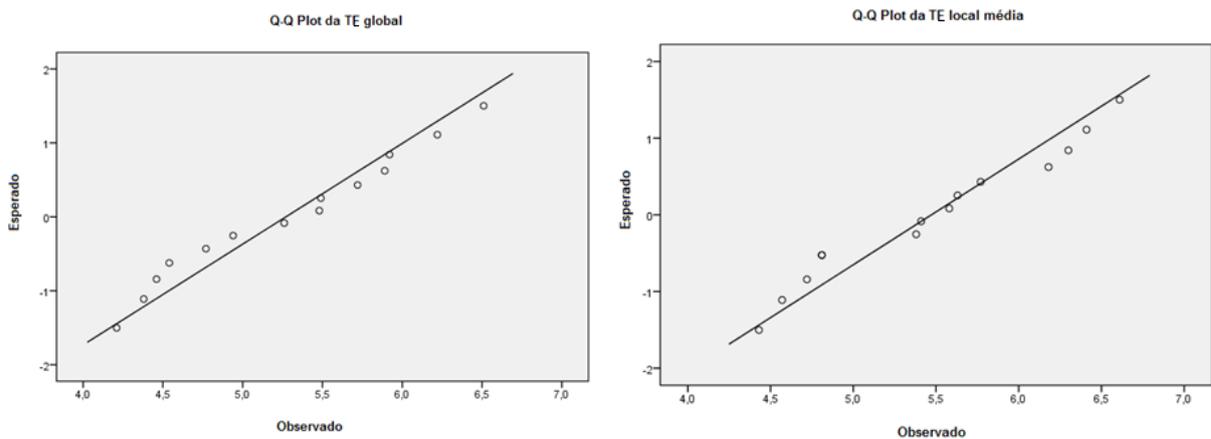
Os valores de assimetria alcançados (apresentados na Tabela 4) apontam a existência de desvios em relação à curva normal tendente para positivo nas TEs e para negativo nas TAs, já os valores negativos de curtose (também exibidos na referida tabela) delatam a ocorrência de achatamento nas duas formas de mensuração da TE e da TA, podendo-se informar que a discrepância, em ambas as medidas (assimetria e curtose), nas quatro variáveis de resposta (TE global, TE local média, TA global e TA local média), contudo, não compromete a normalidade dos dados.

A normalidade dos dados é visualmente ilustrada através dos gráficos Q-Q apresentados nas Figuras 27 e 28, tendo-se mostrado mais informativa através desses do que dos tradicionais histogramas devido à pequena quantidade de dados (exatamente 14, correspondentes, em cada tipo de taxa e forma de mensuração, às

⁵⁹ Destaca-se que, quanto maior é a variância de uma série de dados, maior é a dispersão dos valores que a compõem em relação à própria média da série de dados.

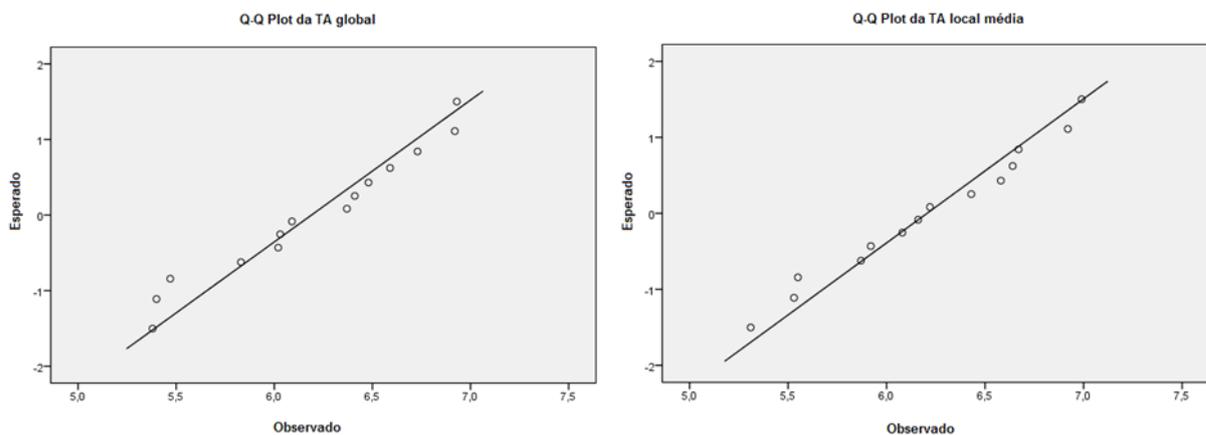
duas gravações de cada um dos sete sujeitos da pesquisa). Destaca-se que, nos quatro gráficos, os pontos concernentes aos valores de taxa mostram-se distribuídos ao longo da reta central que representa a normalidade, sendo o comportamento indicativo de que os conjuntos de dados apresentam configuração associada à distribuição normal, condição essa, reitera-se, importante quando se pretende a utilização de testes paramétricos e de ajustes de modelos lineares.

Figura 27 – Gráficos Q-Q dos dados relativos à TE global e local média



Fonte: A autora (2013).

Figura 28 – Gráficos Q-Q dos dados relativos à TA global e local média



Fonte: A autora (2013).

Cumprir informar, ainda, que nenhuma das quatro variáveis dependentes (TE global, TE local média, TA global e TA local média) rejeitou a hipótese de normalidade frente ao Teste de Shapiro-Wilk, conforme consta na medida “Teste de

normalidade” da Tabela 4, onde são visualizados p-valores não significativos⁶⁰ que corroboram a noção de aderência dos dados a um modelo de distribuição normal.

4.1.1.2 Análises relativas à existência de diferenças quanto à forma de mensuração

Quanto à existência de diferença entre os dois escopos de observação (global e local), têm-se, de antemão, a informação (provinda do Teste F para razão de variâncias mencionado na descrição do resumo estatístico, apresentado na subseção imediatamente anterior) de que a diferença existente entre os tipos de mensuração, tanto na TE quanto na TA, é não significativa.

Na Tabela 5 são exibidos os resultados da correlação intratixa efetuada, ou seja, da correlação entre as duas formas de mensuração de TE (relativa ao “Par 1” da coluna mais à esquerda) e de TA (relativa ao “Par 2” da coluna mais à esquerda), considerando-se conjuntamente as duas gravações dos sete sujeitos pesquisados. Observa-se, a partir das informações contidas na tabela, que em ambas as taxas estudadas a mensuração global e a média das mensurações locais apresentam forte correlação, sendo a condição caracterizada pela obtenção de coeficientes próximos a 1 (conforme visualiza-se na coluna “Correlação”, na TE: 0,968 e na TA: 0,987). O achado é indicativo de que na TE e na TA as duas formas de mensuração, global e local média, demonstram comportamento similar, de forma que, se uma aumenta ou diminui a outra, no mesmo sentido, igualmente o faz.

Tabela 5 – Correlação entre as formas de mensuração das taxas pesquisadas

Par	Correlação	P-valor (bicaudal)
Par 1- TE global e TE local média	0,968	0,000
Par 2 - TA global e TA local média	0,987	0,000

Legenda: P-valor= nível descritivo do teste.

Fonte: A autora (2013).

Para robustecer o resultado encontrado, o de não distinção entre as formas de mensuração em ambas as taxas estudadas, realizou-se um Teste t para

⁶⁰ Na verificação de normalidade dos dados deseja-se a não significância, visto que ela aponta a não rejeição (aceitação) da hipótese de aderência dos dados ao modelo de distribuição normal.

amostras pareadas por forma de mensuração (global e local média), nos dois tipos de taxa (TE e TA). Conforme os níveis descritivos do teste, visualizados na coluna “P-valor” da Tabela 6, há diferença significativa entre as duas formas de mensuração (ao nível de 5%) no caso da TE, não sendo o mesmo observado, contudo, no caso da TA.

Tabela 6 – Teste t para amostras pareadas quanto à forma de mensuração

Par	Média	Desvio padrão	Erro padrão médio	Estatística	GI	P-valor (bicaudal)
Par 1- TE global e TE local média	-0,201	0,184	0,049	-4,091	13	0,001
Par 2 - TA global e TA local média	-0,016	0,084	0,023	-0,692	13	0,501

Legenda: GI= graus de liberdade, P-valor= nível descritivo do teste.
Fonte: A autora (2013).

Sumarizando, de acordo com os dados das Tabelas 5 e 6, recém-apresentadas, a TE medida globalmente e a TE medida por meio da média das múltiplas tomadas locais, embora variem em um mesmo sentido, mostram-se significativamente distintas ($p= 0,001$). Já em relação à TA, tem-se como indiferente a forma de mensuração (se global ou se local média), sendo a variação das taxas obtidas através de uma ou outra forma de cálculo diretamente proporcional.

4.1.1.3 Análises relativas à variabilidade por grupo (sexo e tipo de gravação)

A análise da variabilidade intersujeitos considerou inicialmente grupos por sexo (masculino e feminino) e tipo de gravação (desavisada e avisada). Os resultados, respectivamente apontados nas Tabelas 7 e 8, evidenciam a ausência de diferença significativa (ao Teste t) entre os fatores da variável sexo (p -valores $> 0,05$) em ambas as taxas estudadas, independente da forma de mensuração, e a presença de diferenças significativas ($\alpha= 5\%$) entre os fatores da variável tipo de gravação na TE global e local média, diferentemente do que se observa na TA, em que, independente da forma de mensuração, o tipo de gravação tem associado níveis descritivos de teste (exibidos na coluna “P-valor” da Tabela 8) maiores que 0,05 (TA global com $p= 0,192$ e TA local média com $p= 0,176$).

Tabela 7 – Estatísticas descritivas e Teste t para amostras independentes por sexo das quatro variáveis de resposta

Taxa e mensuração	Sexo	N	Média	Desvio padrão	Estatística	GI	P-valor (bicaudal)
TE global	M	10	5,19	0,818	0,635	12	0,537
	F	4	5,47	0,502			
TE local média	M	10	5,44	0,806	0,245	12	0,811
	F	4	5,55	0,563			
TA global	M	10	6,15	0,493	0,357	12	0,727
	F	4	6,27	0,698			
TA local média	M	10	6,18	0,514	0,259	12	0,800
	F	4	6,26	0,638			

Legenda: M= masculino, F= feminino, N= número de observações, GI= graus de liberdade, P-valor= nível descritivo do teste.

Fonte: A autora (2013).

Tabela 8 – Estatísticas descritivas e Teste t para amostras pareadas por tipo de gravação das quatro variáveis de resposta

Par	Média	Desvio padrão	Erro padrão médio	Estatística	GI	P-valor (bicaudal)
Par 1- TE global desavisada e TE global avisada	-1,176	0,508	0,192	-6,128	6	0,001
Par 2 - TE local média desavisada e TE local média avisada	-1,121	0,605	0,228	-4,906	6	0,003
Par 3 - TA global desavisada e TA global avisada	-0,281	0,506	0,191	-1,470	6	0,192
Par 4 - TA local média desavisada e TA local média avisada	-0,301	0,520	0,197	-1,533	6	0,176

Legenda: GI= graus de liberdade, P-valor= nível descritivo do teste.

Fonte: A autora (2013).

Ainda tendo as variáveis sexo e tipo de gravação como fonte de variação, realizou-se a análise de variância (ANOVA de dois fatores) das TEs e das TAs globais e locais médias referentes aos dois registros dos sete sujeitos do estudo. Os resultados, respectivamente apontados nas Tabelas 9 e 10, evidenciam a existência de diferença significativa (ao nível de 1%) exclusivamente no que se refere à variável tipo de gravação na TE, independente de sua forma de mensuração, achado concordante com o apresentado nas Tabelas 7 e 8.

Tabela 9 – Análise de variância da TE global e local média quanto ao sexo e tipo de gravação

Fonte de variação	Mensuração	GI	Média quadrática	F	P-valor	Poder ^a
Tipo de gravação	Global	1	3,040	18,811	0,001	0,973
	Local média	1	2,761	13,102	0,005	0,902
Sexo	Global	1	0,228	1,411	0,262	0,190
	Local média	1	0,034	0,161	0,697	0,065
Tipo de gravação x sexo	Global	1	0,324	2,002	0,187	0,249
	Local média	1	0,298	1,415	0,262	0,190
Erro	Global	10	0,162	---	---	---
	Local média	10	0,211	---	---	---

Legenda: GI= graus de liberdade, F= razão de variâncias, P-valor= nível descritivo do teste.

^a Computado a partir de alpha =0,05.

Fonte: A autora (2013).

Tabela 10 – Análise de variância da TA global e local média quanto ao sexo e tipo de gravação

Fonte de variação	Mensuração	GI	Média quadrática	F	P-valor	Poder ^a
Tipo de gravação	Global	1	0,195	0,580	0,464	0,106
	Local média	1	0,150	0,468	0,509	0,095
Sexo	Global	1	0,039	0,115	0,741	0,061
	Local média	1	0,020	0,063	0,807	0,056
Tipo de gravação x sexo	Global	1	0,006	0,018	0,895	0,052
	Local média	1	0,082	0,255	0,625	0,074
Erro	Global	10	0,337	---	---	---
	Local média	10	0,320	---	---	---

Legenda: GI= graus de liberdade, F= razão de variâncias, P-valor= nível descritivo do teste.

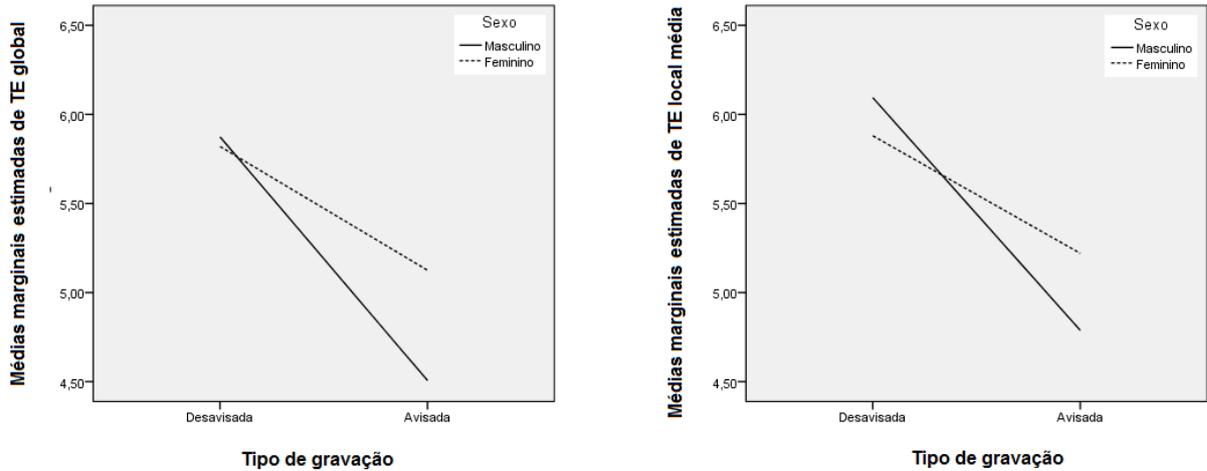
^a Computado a partir de alpha =0,05.

Fonte: A autora (2013).

Foram efetuadas as análises de resíduos para as análises de variância recém expostas, cabendo consignar que todas elas atenderam às suposições de normalidade e homogeneidade do modelo.

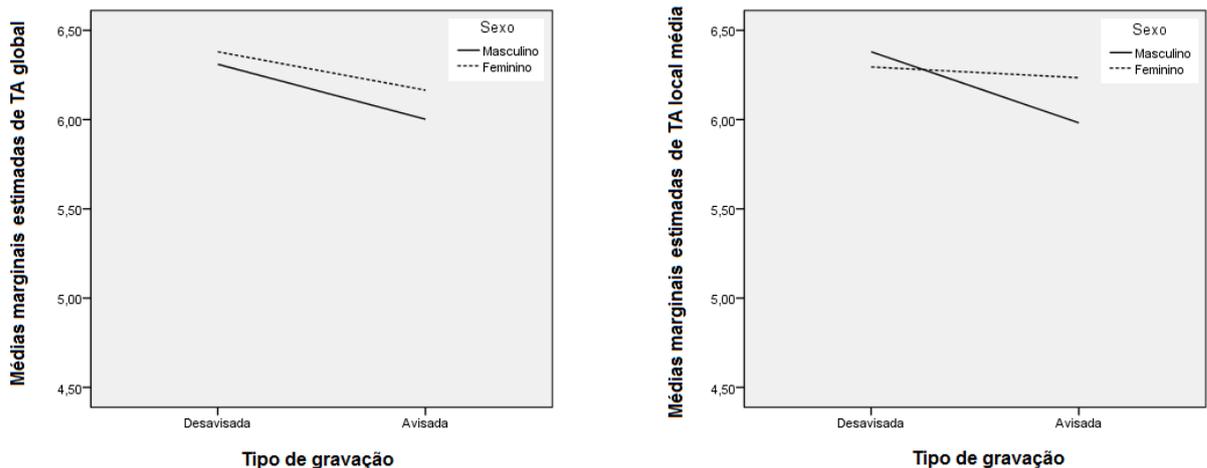
As médias das TEs globais e das TEs locais médias nos dois tipos de gravação (desavisada e avisada), assim como as médias das TAs globais e das TAs locais médias (da mesma forma, na gravação desavisada e avisada), por sexo, foram então visualmente avaliadas, tendo sido consideradas na plotagem dos gráficos de médias marginais exibidos, respectivamente, nas Figuras 29 e 30 a seguir apresentadas.

Figura 29 – Médias de TE global e local média por tipo de gravação e sexo



Fonte: A autora (2013).

Figura 30 – Médias de TA global e local média por tipo de gravação e sexo



Fonte: A autora (2013).

De maneira geral é possível deprender, a partir da avaliação dos gráficos exibidos nas Figuras 29 e 30⁶¹ que, independentemente do sexo, são empregadas menores taxa quando há ciência da gravação, o que é caracterizado pela inclinação descendente das oito retas apresentadas, cujos extremos referem, da esquerda para a direita, a média obtida pelo sexo na gravação desavisada e na gravação avisada.

Com referência à TE (Figura 29), percebe-se que o emprego das taxas globais e locais médias mais altas está associado prevalentemente ao sexo masculino na gravação desavisada e ao sexo feminino na gravação avisada.

⁶¹ A variável tipo de gravação, apresentada no eixo x, é medida em nível de escala nominal. A utilização de retas dá-se apenas para melhor ilustrar o comportamento dos sexos em cada nível da variável ("Desavisada" e "Avisada").

No que tange à extensão da diferença entre as médias alcançadas pelos sujeitos do sexo masculino e feminino em cada tipo de gravação, observa-se que há maior distanciamento entre as médias masculina e feminina na gravação desavisada quando em mensuração local e na gravação avisada quando em mensuração global.

Já relativamente à TA (Figura 30), destaca-se o comportamento discordante das médias marginais na observação global da TA se comparado ao das médias referentes à TA local média e à TE global e local média, sendo possível constatar (através do estabelecimento da relação entre as linhas do gráfico exibido no quadro à esquerda) que na TA global, diferentemente do que se observam nos três outros casos, a média de taxa mais alta na gravação desavisada é empregada pelo sexo feminino. A examinação da diferença entre as médias de TA global masculina e feminina (respectivamente 6,15 síl/s e 6,27 síl/s, conforme consta na Tabela 7) na gravação desavisada evidenciou (de acordo com o Teste F da análise de variância das TEs e das TAs exibido, respectivamente, nas Tabelas 9 e 10), que a diferença existente, contudo, não é estatisticamente significativa.

Quanto à extensão da diferença entre a média masculina e feminina em cada tipo de gravação, observa-se, na TA, proporção similar no distanciamento entre as médias masculina e feminina na gravação desavisada (porém no sentido inverso) e, no caso da gravação avisada, maior diferença entre as médias na observação local.

4.1.1.4 Análises relativas à variabilidade intrassujeito e intersujeitos

A TE e a TA globais e locais médias referentes a cada um dos sete sujeitos pesquisados foram verificadas intra e interindividualmente. Na Tabela 11 são apresentados os valores de TE e de TA global e local média, tendo sido primeiramente considerado cada um dos dois tipos de gravação em separado e, após, a junção dos mesmos na composição de um conjunto de dados indistinto quanto ao tipo de gravação. Na mencionada tabela é apresentada também a variabilidade intrassujeito, sendo a mesma indicada pelo desvio padrão da média das múltiplas tomadas locais de TE e de TA e, excepcionalmente, no caso da descrição indistinta (apresentada na categoria “Geral” da coluna “Gravação”), pelo desvio padrão da média estabelecida entre a taxa (TE ou TA, em mensuração global ou local) nos dois tipos de gravação (desavisada e avisada).

Tabela 11 – TE e TA global e local média dos sujeitos por tipo de gravação

Taxa	Gravação	Mensuração	S1	S2	S3	S4	S5	S6	S7
TE	Desavisada	Global	6.51	5.89	5.92	5.72	5.49	6.22	5.26
		Local média (dp)	6.61 (1,42)	6.30 (1,43)	6.18 (1,55)	5.58 (1,31)	5.77 (1,15)	6.41 (1,03)	5.38 (1,20)
	Avisada	Global	4.46	4.38	5.48	4.77	4.54	4.94	4.21
		Local média (dp)	4.43 (0,98)	4.57 (0,77)	5.63 (1,16)	4.81 (1,27)	4.81 (0,92)	5.41 (1,10)	4.72 (1,20)
	Geral	Global (dp)	5,48 (1,45)	5,13 (1,07)	5,70 (0,31)	5,24 (0,67)	5,01 (0,67)	5,58 (0,90)	4,73 (0,74)
		Local média (dp)	5,52 (1,54)	5,43 (1,22)	5,90 (0,39)	5,19 (0,54)	5,29 (0,67)	5,91 (0,71)	5,05 (0,47)
TA	Desavisada	Global	6.92	6.48	6.73	6.03	6.09	6.59	5.47
		Local média (dp)	6.99 (1,37)	6.58 (1,28)	6.67 (1,49)	5.92 (1,15)	6.16 (0,81)	6.64 (0,85)	5.53 (1,21)
	Avisada	Global	6.41	6.37	6.93	5.40	5.38	5.83	6.02
		Local média (dp)	6.43 (1,01)	6.22 (0,95)	6.92 (1,09)	5.55 (0,90)	5.31 (0,74)	5.87 (0,91)	6.08 (1,14)
	Geral	Global (dp)	6,66 (0,36)	6,42 (0,08)	6,83 (0,14)	5,71 (0,44)	5,73 (0,50)	6,21 (0,54)	5,74 (0,39)
		Local média (dp)	6,71 (0,40)	6,40 (0,25)	6,79 (0,18)	5,73 (0,26)	5,73 (0,60)	6,25 (0,54)	5,80 (0,39)

Legenda: dp= desvio padrão amostral, S= sujeito.

Fonte: A autora (2013).

Nota: A medida “Geral” considera os dois tipos de gravação.

Na verificação da suposição da normalidade e da homogeneidade de variâncias, constatou-se que a última foi violada, o que justificou que a comparação entre as médias dos sujeitos fosse efetuada através do teste robusto de igualdade de médias de Brown-Forsythe. Os resultados alcançados no mencionado teste são apresentados, por tipo de taxa e forma de mensuração, na Tabela 12.

Tabela 12 – Teste de Brown-Forsythe para igualdade de médias entre sujeitos

Taxa e mensuração	Estatística	GI 1	GI 2	P-valor
TE global	0,290	6	4,430	0,916
TE local média	0,285	6	3,555	0,915
TA global	2,946	6	4,839	0,132
TA local média	2,571	6	4,633	0,170

Legenda: P-valor= nível descritivo do teste, GI= grau de liberdade (“1” corresponde ao número de sujeitos menos 1 e “2” ao erro experimental).

Fonte: A autora (2013).

É possível constatar através da observação dos resultados exibidos na Tabela 12 que não existem diferenças significativas entre os sujeitos quanto às quatro variáveis de resposta. O achado pode ser explicado, no entanto, pelo elevado nível de heterogeneidade nas medições intrassujeito. Apesar de não ter sido encontrada diferença intersujeito significativa, observa-se que os níveis de significância do teste (exibidos na coluna “P-valor”) são consideravelmente menores para a TA (nas duas formas de mensuração), deduzindo-se com isso que, havendo um conjunto amostral maior, mais facilmente encontrar-se-ão diferenças significativas na TA do que na TE.

4.1.1.5 Análises relativas ao potencial individualizante das taxas

O estabelecimento da variabilidade intersujeitos e intrassujeito (através de uma análise de variância, tendo o sujeito como fonte de variação) possibilitou a verificação do poder discriminatório das taxas, nas formas de mensuração pesquisadas. Para estimar que fração da variabilidade total das medidas decorre das variações entre os indivíduos, ou seja, o quanto de variação é previsto para o grupo e o que pode ser atribuído à variação própria do indivíduo, utilizou-se o coeficiente de correlação intraclasse (CCI)⁶², sendo sua aplicação em estudos sociofonéticos referida por Thomas (2011)⁶³.

Ressalta-se que o CCI é determinado através da razão de duas somas de quadrados (soma de quadrados intrassujeito e intersujeitos). Na literatura sobre Reconhecimento de Locutor encontra-se referência ao uso da razão F (“F-ratio”) no estabelecimento do potencial individualizante dos parâmetros técnico-comparativos (NOLAN, 1983; ROSE, 2002), entendida como correspondente à distribuição de probabilidade F, sendo, assim como o CCI, uma razão entre duas variâncias independentes⁶⁴.

Na Tabela 13 são expostas, por tipo de taxa e forma de mensuração, as estimativas da variância intersujeitos e intrassujeito, bem como os CCIs obtidos.

⁶² O CCI é o coeficiente comumente utilizado na verificação da reprodutibilidade dos estudos, que estima, da fração da variabilidade total, quais são devidas às variações que aconteceram entre os indivíduos.

⁶³ O autor afirma que tal coeficiente é utilizado para determinar o grau de diferença entre conjuntos pareados de dados.

⁶⁴ O “F ratio of analysis of variance [...] é calculado através da razão entre a variância dos valores médios de um parâmetro nos falantes” (variação intersujeitos) “e a variância média intrassujeito”. (cf. NOLAN, 1983, p.12, tradução nossa).

Tabela 13 – Estimativas da variância e CCI para as quatro variáveis de resposta

Taxa e mensuração	Variância intersujeito	Variância intrassujeito	CCI
TE global	0,232	0,802	0,199
TE local média	0,224	0,786	0,196
TA global	0,441	0,150	0,716
TA local média	0,415	0,161	0,688

Legenda: CCI= coeficiente de correlação intraclasse.

Fonte: A autora (2013).

A observação dos dados da Tabela 13 indica que somente a TA (global e local média) apresenta a desejável superioridade da variância intersujeito em relação à variância intrassujeito. A interpretação dos valores do coeficiente de correlação intraclasse exibidos na coluna “CCI” corrobora a percepção de que só a TA (nas duas formas de mensuração contempladas neste estudo) apresenta poder discriminatório passível de ser enquadrado como satisfatório⁶⁵. Relativamente à TE, os CCIs obtidos (inferiores a 0,2) indicam pobre poder discriminatório de indivíduo, achado condizente com o fato da variância intrassujeito ser maior do que a variância intersujeitos nesse tipo de taxa (na medida global $0,802 > 0,232$ e na medida local média $0,786 > 0,224$).

4.1.1.6 Análises relativas ao condicionamento das taxas em razão das demais variáveis independentes

Na averiguação do impacto gerado pelas variáveis independentes nas taxas em estudo, empreendeu-se a correlação entre as variáveis independentes escalares (idade, escolaridade e *gap* temporal)⁶⁶ e as quatro variáveis de resposta generalizadamente admitidas (TE global, TE local média, TA global e TA local média). Os resultados alcançados são exibidos na Tabela 14.

⁶⁵ Conforme Lin (1989), CCIs menores do que 0,4 apresentam pobre poder discriminatório e maiores do que 0,4, poder discriminatório satisfatório. Classificação mais detalhada, no entanto, é encontrada em Pestana e Gageiro (2003), para quem: CCIs $< 0,40$ = pobre (reprodutibilidade), $CCI \geq 0,40$ e $< 0,75$ = satisfatório e $CCI \geq 0,75$ = excelente.

⁶⁶ Excluem-se temporariamente as variáveis escalares tamanho do intervalo de fala e duração média da sílaba que, por requerem a consideração das tomadas locais individualmente e não a média das mesmas, serão estudadas em separado. Ressalta-se, ainda, que a variação no comportamento da TE e da TA ocasionada pelas variáveis independentes qualitativas sexo e tipo de gravação foi verificada na análise por grupo apresentada neste capítulo, na Subseção 4.1.1.3.

Tabela 14 – Correlação entre as taxas pesquisadas e as variáveis idade, escolaridade e *gap* temporal

Taxa e mensuração	Medidas	Idade	Escolaridade	Gap temporal
TE global	Correlação de Pearson	-0,251	0,177	0,169
	P-valor (bicaudal)	0,386	0,546	0,564
TE local média	Correlação de Pearson	-0,360	0,037	0,199
	P-valor (bicaudal)	0,206	0,901	0,496
TA global	Correlação de Pearson	-0,022	0,293	0,675*
	P-valor (bicaudal)	0,940	0,310	0,008
TA local média	Correlação de Pearson	-0,018	0,321	0,631*
	P-valor (bicaudal)	0,951	0,263	0,016

* Correlação significativa ao nível de 0,05.

Fonte: A autora (2013).

Relativamente à TE, observa-se ausência de correlação significativa ($\alpha= 5\%$) entre as medidas global e local média e as três variáveis consideradas (idade, escolaridade e *gap* temporal). Apesar da não significância encontrada, nota-se que existe tendência à correlação positiva entre a TE, nas duas formas de mensuração, e as variáveis escolaridade e *gap* temporal, e tendência à correlação negativa entre a TE e a variável idade.

A aplicação de igual procedimento à TA permitiu a constatação da existência de correlação significativa (ao nível de 5%) entre a TA global e local média e, somente, a variável independente *gap* temporal. Com referência a indicadores de tendência, observam-se resultados compatíveis com os obtidos na TE (escolaridade com tendência à correlação positiva e idade com tendência à correlação negativa).

4.1.1.7 Análises relativas à predição do comportamento das taxas em razão da idade, do sexo, da escolaridade, do *gap* temporal e do tipo de gravação

A fim de prever o comportamento das taxas estudadas frente às variáveis independentes admitidas, aplicou-se o modelo de regressão linear automático (conforme método descrito na Subseção 3.4.2 do Capítulo 3), o qual considerou cinco variáveis no modelo de entrada, exatamente o(a): idade, sexo⁶⁷, escolaridade, *gap* temporal e tipo de gravação.

⁶⁷ Cumpre informar que o modelo de regressão linear generalizado empregado é robusto ao desbalanceamento existente neste estudo quanto ao número de sujeitos do sexo feminino e masculino.

Na Tabela 15 é apresentado o resumo estatístico dos modelos ajustados baseado em quatro variáveis de resposta e cinco variáveis preditoras (recém-referidas), sendo possível depreender, a partir das informações nela contidas, que os coeficientes de correlação múltipla (exibidos na coluna “R múltiplo”) são maiores nas mensurações de TE se comparados aos associados às mensurações de TA. No entanto, pode-se afirmar que em ambos os tipos de taxa a mensuração global apresenta correlação múltipla superior quando comparada à mensuração local, o que significa que nesse tipo de taxa (TE) e forma de mensuração (global) as variáveis preditoras apresentam maior poder explicativo, fato confirmado pelos coeficientes de determinação R^2 ajustados exibidos na coluna homônima.

Tabela 15 – Resumo estatístico da regressão linear com cinco variáveis independentes

Taxa e mensuração	R múltiplo	R^2 ajustado	Erro padrão
TE global	0,915 ¹	0,794	0,356
TE local média	0,901 ¹	0,729	0,378
TA global	0,845 ¹	0,587	0,342
TA local média	0,795 ²	0,523	0,365

¹ Preditoras: Constante + idade + tipo de gravação + *gap* temporal + escolaridade.

² Preditoras: Constante + tipo de gravação + *gap* temporal + escolaridade.

Fonte: A autora (2013).

Os coeficientes estimados para as variáveis preditoras dos modelos relativos, respectivamente, às variáveis de resposta TE e TA, nas duas formas de mensuração contempladas, são exibidos nas Tabelas 16 e 17.

Tabela 16 – Coeficientes da regressão linear relativos a TE global e local média

Fatores	Mensuração	Coefficientes	Erro padrão	Estatística t	P-valor
Constante ¹	Global	5,837	0,135	43,188	<0,001
	Local média	6,006	0,143	41,949	<0,001
Idade	Global	-0,226	0,108	-2,101	0,065
	Local média	-0,280	0,114	-2,461	0,036
Sexo	Global	---	---	---	---
	Local média	---	---	---	---
Escolaridade	Global	0,224	0,107	2,103	0,065
	Local média	0,142	0,113	1,255	0,241
<i>Gap</i> temporal	Global	0,175	0,101	1,735	0,117
	Local média	0,194	0,107	1,819	0,102
Tipo de gravação	Global	-1,133	0,192	-5,908	<0,001
	Local média	-1,068	0,203	-5,259	0,001

Legenda: P-valor= nível descritivo do teste.

¹ Fator caracterizado no Glossário.

² Variável excluída pelo modelo.

Fonte: A autora (2013).

Tabela 17 – Coeficientes da regressão linear relativos a TA global e local média

Fatores	Mensuração	Coeficientes	Erro padrão	Estatística t	P-valor
Constante	Global	6,317	0,130	48,665	0,000
	Local	6,356	0,138	46,109	0,000
Idade	Global	-0,134	0,103	-1,294	0,228
	Local	--- ¹	---	---	---
Sexo	Global	--- ²	---	---	---
	Local	---	---	---	---
Escolaridade	Global	0,241	0,102	2,350	0,043
	Local	0,204	0,102	2,003	0,073
Gap temporal	Global	0,401	0,097	4,132	0,003
	Local	0,352	0,102	3,468	0,006
Tipo de gravação	Global	-0,256	0,184	-1,390	0,198
	Local	-0,301	0,195	-1,546	0,153

Legenda: P-valor= nível descritivo do teste.

¹ Variável excluída pelo modelo nesse tipo de mensuração.

² Variável excluída pelo modelo.

Fonte: A autora (2013).

Na avaliação dos resultados relativos à TE (apresentados na Tabela 16) observa-se que o modelo selecionado apresentou como significativas ($\alpha=5\%$), conforme os níveis descritivos do teste exibidos na coluna “P-valor”, exclusivamente tipo de gravação na mensuração global e tipo de gravação e idade na mensuração local. Expandindo a interpretação dos dados da tabela, considerando os coeficientes do modelo de regressão cujo p-valor mostra-se significativo (exibidos na coluna “Coeficientes”), têm-se as seguintes formulações acerca do comportamento da TE:

- globalmente, a TE diminui 1,133 síl/s quando o sujeito sabe que está sendo gravado;
- localmente, a TE diminui 1,068 síl/s quando o sujeito sabe que está sendo gravado e 0,280 síl/s a cada ano a mais na idade do sujeito.

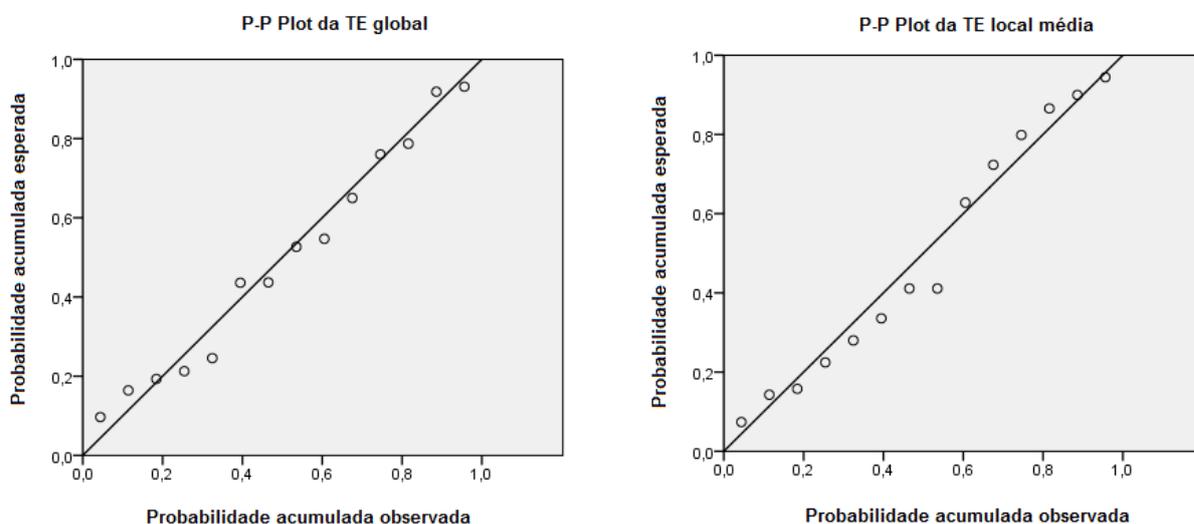
Relativamente à TA, o modelo selecionado apresentou como significativas ($\alpha=5\%$) as variáveis escolaridade e *gap* temporal na mensuração global e exclusivamente o *gap* temporal na mensuração local. A interpretação dos coeficientes com p-valores significativos (exibidos na coluna “Coeficientes” da Tabela 17) permite as seguintes formulações acerca do comportamento da TA:

- a) globalmente, a TA aumenta 0,241 síl/s a cada ano a mais de escolarização do sujeito e 0,401 síl/s a cada mês a mais existente entre as gravações avisada e desavisada;
- b) localmente, a TA aumenta 0,352 síl/s a cada mês a mais existente entre as gravações avisada e desavisada.

Cabe ressaltar que a partir da média das TAs locais a escolaridade apresentou nível descritivo do teste igual a 0,073, o que aponta a tendência de que, com o aumento do tamanho da amostra, se obtenha resultado significativo também nessa variável.

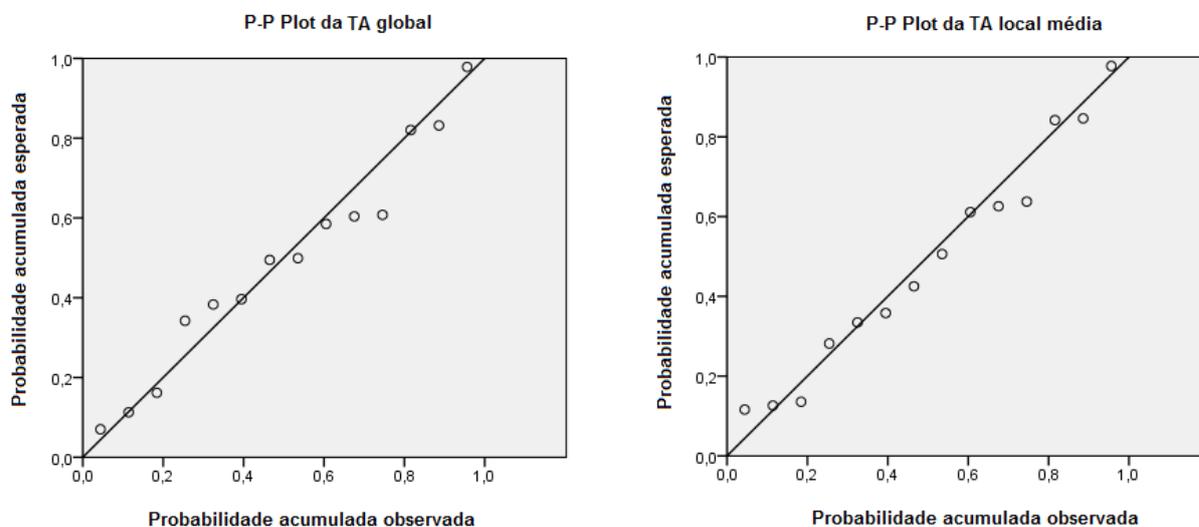
Finalizando a exposição dos resultados relativos à regressão linear envolvendo cinco das seis variáveis independentes admitidas (TE global, TE local média, TA global e TA local média), são apresentadas, respectivamente nas Figuras 31 e 32, as análises de normalidade dos resíduos relativos à TE.

Figura 31 – Gráficos P-P dos resíduos relativos a TE global e local média



Fonte: A autora (2013).

Figura 32 – Gráficos P-P dos resíduos relativos a TA global e local média



Fonte: A autora (2013).

Os resíduos da análise de regressão foram avaliados através de gráficos P-P (recém-apresentados) e do Teste de Shapiro-Wilk, tendo os resultados se mostrado, em ambos, compatíveis com os de uma distribuição normal. A normalidade é delatada pela disposição dos pontos relativos aos valores de taxa próximo à reta central, no caso dos gráficos P-P, e pelos p-valores (na TE: em tomada global $p=0,488$ e em tomada local $p=0,577$ e na TA: em tomada global $p=0,593$ e em tomada local $p=0,775$), no caso do Teste de Shapiro-Wilk.

4.1.2 Análises baseadas nas TEs e TAs globais e TEs e TAs locais múltiplas

Nesta subseção são apresentados os resultados obtidos a partir das múltiplas tomadas locais, tendo sido realizadas duas verificações extraordinárias: a da possível interferência indireta da variável tamanho do intervalo de fala no comportamento da variável idade e a da ocorrência (ou não) do fenômeno de encurtamento antecipatório. Após, segue exposição do resultado da observação do condicionamento gerado pela variável independente ainda não analisada (tamanho do intervalo de fala) e a reanálise da predição do comportamento das taxas (TE e TA) em razão das seis variáveis independentes.

4.1.2.1 Observação de efeitos específicos

A consideração das múltiplas tomadas locais oportunizou a verificação de dois efeitos que permeiam, segundo a literatura consultada, a rotina de investigação das taxas temporais estudadas, a saber, o de interferência indireta do comprimento do enunciado no comportamento da variável idade em relação à TA e o de encurtamento da duração silábica gerado pelo aumento do número de sílabas produzidas no intervalo de fala pausa-excludente (intervalo interpausal), sendo ambos adiante abordados.

4.1.2.1.1 Interferência indireta do tamanho do intervalo de fala no comportamento da variável idade

A partir do apontamento realizado, entre outros, por Quené (2008), de que o condicionamento provocado pela idade decorre, na verdade, do tamanho do

intervalo de fala analisado (esse, segundo o autor, maior em indivíduos mais jovens e menor em indivíduos mais velhos), procedeu-se ao cruzamento entre as variáveis idade e tamanho do intervalo de fala. Na tarefa, contudo, foi necessário conceber o tamanho do intervalo de fala como sendo o número de sílabas nele contidas (e não a sua duração) e, ainda, abandonar temporariamente o uso em curso de um valor médio de taxa local em favor da consideração das múltiplas tomadas (intervalos de fala) úteis de cada sujeito em cada uma de suas duas gravações.

A verificação de correlação pretendida (idade *versus* número de sílabas do intervalo) aplicou-se, assim como em Quené (2008), somente às TAs em razão de ser esse o tipo de fala pausa-excludente. Os resultados encontrados baseiam-se em 748 observações e apontam a inexistência de correlação significativa ao nível de 5% (correlação= 0,067 e $p= 0,065$), sendo possível concluir que o tamanho do intervalo de fala não interfere significativamente no comportamento da variável idade.

4.1.2.1.2 Encurtamento antecipatório

Com base na análise das observações locais, mantendo-se a restrição de análise ao tipo de taxa pausa-excludente (TA), verificou-se a correlação existente entre o número de sílabas identificado em cada um dos 748 intervalos interpausais e a duração média das sílabas nos mesmos. Nesse caso, obteve-se como resultado a correlação significativa de -0,312 ($\alpha= 5\%$, $p< 0,001$), segundo a qual o aumento no número de sílabas do intervalo interpausal condiciona a redução da duração média das sílabas nele contidas, o que equivale ao efeito de encurtamento antecipatório (conceituado na Subseção 2.1 do Capítulo 2).

4.1.2.2 Condicionamento das taxas em razão da variável independente tamanho do intervalo de fala

A partir da examinação das múltiplas observações locais, estabeleceu-se a correlação entre as TEs (em 539 turnos de fala úteis) e as TAs (em 748 intervalos interpausais úteis) e a variável tamanho do intervalo de fala (leia-se, a duração temporal das respectivas 1.287 observações). O resultado alcançado indica que há, no caso das TEs, correlação significativa ($\alpha= 5\%$) entre essas e os tamanhos de intervalo de fala (correlação de -0,340 e $p< 0,001$), diferentemente do que ocorre

com as TAs, que mostram-se com esses não correlacionadas (correlação= -0,007 e $p= 0,844$). Destaca-se que em ambos os casos (a despeito da não significância encontrada nas TAs), a correlação estimada é negativa, ou seja, o aumento no valor da TE e da TA implica a redução da duração temporal do intervalo de fala.

4.1.2.3 Reanálise da predição do comportamento das taxas em razão das variáveis independentes

A inclusão da variável tamanho do intervalo de fala na composição do modelo de regressão múltipla requereu a consideração das múltiplas tomadas locais da TE e da TA (e não mais a média dessas). Grosso modo, esta análise de regressão, em detrimento da anterior, é entendida como mais completa, uma vez que contempla o rol integral de variáveis independentes (excetuando-se por óbvio a variável sujeito, admitida com vistas à análise intersujeitos, e a variável duração média das sílabas, admitida para verificação exclusiva do fenômeno de encurtamento antecipatório).

Na Tabela 18 é apresentado o resumo estatístico dos modelos ajustados baseado em duas variáveis de resposta e seis variáveis preditoras (as consideradas no modelo de regressão precedente e a variável tamanho do intervalo de fala). A partir das informações nela contidas é possível depreender que os coeficientes de correlação múltipla (exibidos na coluna “R múltiplo”) são maiores na TE do que na TA, o que significa que nesse tipo de taxa (TE) as variáveis preditoras apresentam maior poder explicativo, fato confirmado pelos coeficientes de determinação R^2 ajustados exibidos na coluna homônima.

Tabela 18 – Resumo estatístico da regressão linear com seis variáveis independentes

Taxa e mensuração	R múltiplo	R² ajustado	Erro padrão
TE local	0,495 ¹	0,236	1,196
TA local	0,397 ²	0,158	1,104

¹ Preditoras: Constante + idade + sexo + escolaridade + *gap* temporal + tipo de gravação + tamanho do intervalo de fala.

² Preditoras: Constante + idade + sexo + escolaridade + *gap* temporal + tipo de gravação.
Fonte: A autora (2013).

Nas Tabelas 19 e 20 são apresentados os resultados obtidos na observação local exclusiva, reiterando-se que ela considera 539 turnos de fala úteis (na examinação da TE) e 748 intervalos interpausais úteis (na examinação da TA).

Tabela 19 – Coeficientes da regressão linear relativos às TEs locais

Fatores	Coeficientes	Erro padrão	Estatística t	P-valor
Constante	5,423	0,293	18,526	0,000
Idade	-0,027	0,009	-2,852	0,005
Sexo	0,353	0,163	2,168	0,031
Escolaridade	0,105	0,028	3,806	0,000
Gap temporal	0,026	0,005	4,893	0,000
Tipo de gravação	-0,754	0,118	-6,380	0,000
Tamanho do intervalo de fala	0,000	0,000	-6,143	0,000

Legenda: P-valor= nível descritivo do teste.

Fonte: A autora (2013).

Tabela 20 – Coeficientes da regressão linear relativos às TAs locais

Fatores	Coeficientes	Erro padrão	Estatística t	P-valor
Constante	5,127	0,231	22,200	0,000
Idade	-0,025	0,007	-3,490	0,001
Sexo	0,225	0,112	2,011	0,045
Escolaridade	0,145	0,020	7,191	0,000
Gap temporal	0,037	0,004	9,651	0,000
Tipo de gravação	-0,337	0,090	-3,762	0,000
Tamanho do intervalo de fala	--- ¹	---	---	---

Legenda: P-valor= nível descritivo do teste.

¹Variável excluída pelo modelo.

Fonte: A autora (2013).

Nessa nova perspectiva, os modelos selecionados excluíram apenas a variável tamanho do intervalo de fala na TA (embora selecionada na TE, tal variável não agrega nem destitui valor à taxa) e apontam como significativas ($\alpha= 5\%$, sendo os maiores p-valores associados à variável sexo) todas as demais variáveis admitidas nos modelos, em ambas as taxas estudadas.

Em termos de predição de comportamento pode-se formular a partir da observação dos coeficientes relativos às TEs locais (apresentados na coluna “Coeficientes” da Tabela 19) que:

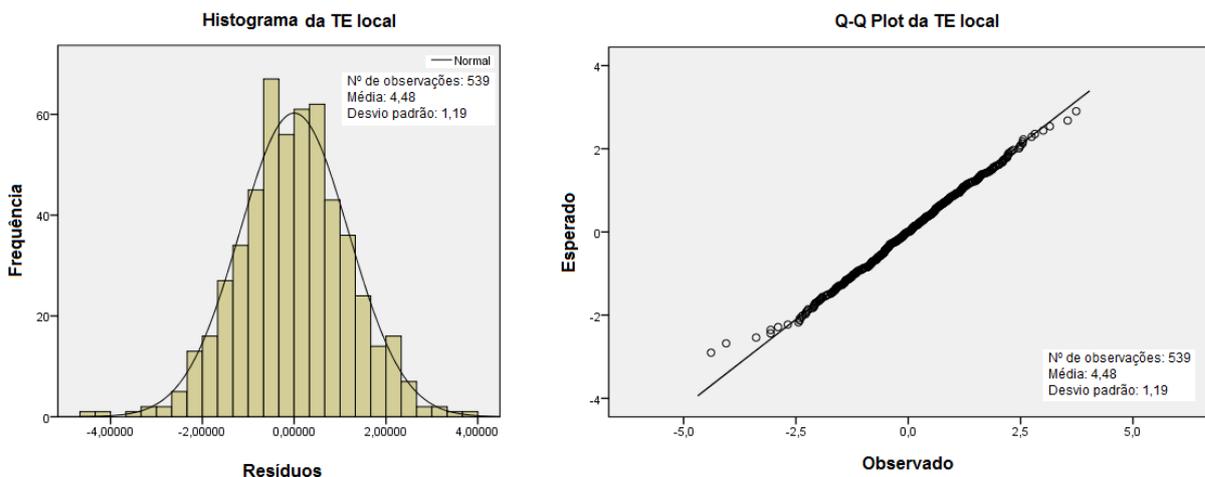
- a) a TE aumenta 0,353 síl/s quando o sujeito é do sexo masculino, 0,105 síl/s a cada ano a mais de escolaridade e 0,026 síl/s a cada mês a mais no *gap* temporal existente entre as gravações do sujeito;
- b) a TE diminui 0,027 síl/s a cada ano a mais de vida do sujeito e 0,754 síl/s se o sujeito sabe que está sendo gravado.

Com relação às TAs locais, a partir dos coeficientes exibidos na coluna homônima da Tabela 20, formula-se as seguintes predições de comportamento:

- a) a TA aumenta 0,225 síl/s quando o sujeito é do sexo masculino, 0,145 síl/s a cada ano a mais de escolaridade e 0,037 síl/s a cada mês a mais no *gap* temporal existente entre as gravações do sujeito;
- b) a TA diminui 0,025 síl/s a cada ano a mais de vida do sujeito e 0,337 síl/s se o sujeito sabe que está sendo gravado.

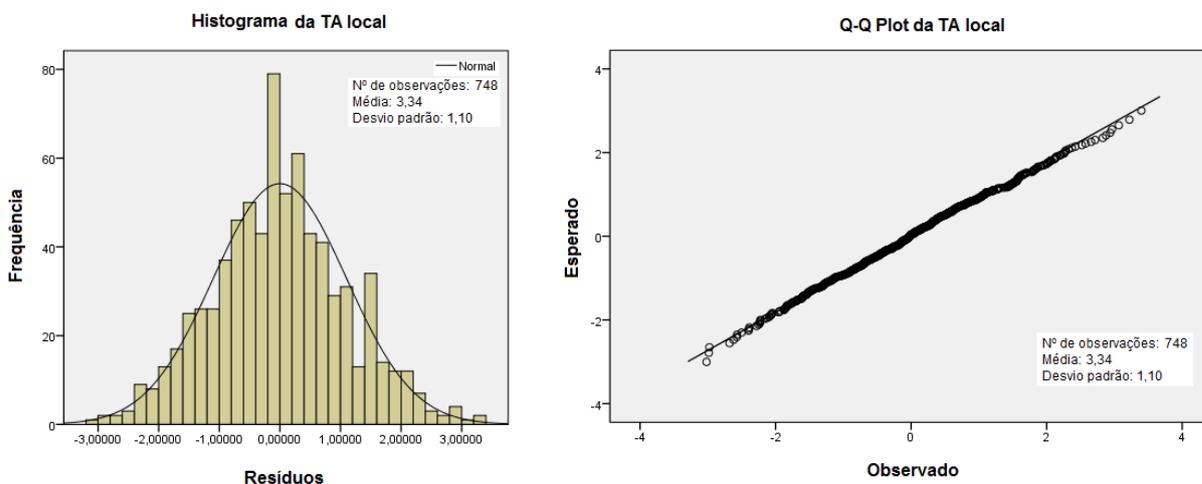
Finalizando a exposição dos resultados relativos à regressão linear múltipla baseada nas tomadas locais (em número de 1.287), apresenta-se nas Figuras 33 e 34 a análise de normalidade dos resíduos, avaliados através de histogramas e de gráficos Q-Q. O encaixamento das barras relativas aos dados dos sujeitos à curva normal (no caso dos histogramas) e a distribuição dos pontos concernentes aos valores de taxa ao longo da reta central que representa a normalidade (no caso dos gráficos Q-Q) evidenciam que tais resíduos mostram-se normalmente distribuídos.

Figura 33 – Histograma e gráficos Q-Q dos resíduos relativos às TE locais



Fonte: A autora (2013).

Figura 34 – Histograma e gráficos Q-Q dos resíduos relativos às TAs locais



Fonte: A autora (2013).

No APÊNDICE C é apresentado um quadro com a síntese dos resultados referentes ao condicionamento das variáveis dependentes TE e TA (global e local média) em razão das variáveis independentes foco do estudo (idade, sexo, escolaridade, *gap* temporal, tipo de gravação e tamanho do intervalo⁶⁸), destacando-se nele os resultados estatisticamente significantes.

4.2 DISCUSSÃO DOS RESULTADOS

Na discussão dos resultados a seguir exibida será debatida a pertinência dos posicionamentos teórico-metodológicos assumidos neste estudo, apontados tanto no texto da revisão de literatura (Capítulo 2) quanto no texto que discorre sobre a metodologia adotada (Capítulo 3). Após, na Subseção 4.2.2, tratar-se-á da resposta aos objetivos estabelecidos (Subseção 3.1.3, Capítulo 3), apresentando-se algumas das possíveis justificativas para os comportamentos (de variáveis) não previstos.

4.2.1 A pertinência das decisões teórico-metodológicas

Na Subseção 2.1 do Capítulo 2, discorre-se sobre os impasses teórico-metodológicos enfrentados durante o planejamento e desenvolvimento de uma pesquisa que se propõe a abordar o tempo de fala, sendo apresentada uma

⁶⁸ Reitera-se que a informação relativa à variável tamanho do intervalo de fala considerou as múltiplas mensurações locais e não a média dessas.

sequência de questionamentos atinentes à construção do método de investigação.

Na mencionada subseção discute-se sobre:

- a) que unidade linguística utilizar na contagem (se sentença, palavra, sílaba, unidade VV, mora, segmento ou outra) e o critério a ser adotado em sua identificação;
- b) qual o método de cálculo, se direto (o número de unidades no intervalo dividido pela sua duração) ou inverso (a duração do intervalo dividida pelo número de unidades produzidas);
- c) qual o *status* da unidade selecionada, se será observada a forma fonológica ou a fonética, havendo também a possibilidade de uso do registro ortográfico;
- d) qual o escopo de análise desejado ou possível, se global e/ ou local;
- e) que tipo e tamanho de intervalo de fala submeter à examinação, ou seja, se o cômputo se dará em turnos de fala, orações, intervalos interpausais, frases entonacionais ou outro e como controlar efeitos intervenientes observados em intervalos muito curtos e muito longos;
- f) que tipo de material de fala utilizar (se espontâneo ou não), sua forma de eliciação e condições de gravação;
- g) que conduta adotar com relação às pausas silenciosas (duração a ser considerada na caracterização), às pausas preenchidas (admiti-las ou não) e à fala disfluente (admiti-las ou não).

Em síntese, as decisões técnicas tomadas em resposta aos questionamentos elencados acima foram, respectivamente:

- a) a escolha da sílaba como unidade linguística a ser utilizada no cômputo das taxas e da presença de ao menos um dos vestígios acústicos associados às vogais plenas para cômputo de sílaba;
- b) o cálculo direto das TEs e das TAs, ressalvada a reversão indispensável à avaliação do fenômeno de encurtamento antecipatório;
- c) a versão fonética de sílaba, na fala como efetivamente implementada (a despeito da controvérsia conceitual, apresentada na Subseção 2.1 do Capítulo 2);
- d) a observação global e local das taxas (a última, em suas múltiplas tomadas e através da média aritmética dessas);

- e) a consideração de turnos de fala (no caso da TE) e intervalos interpausais (no caso da TA) com no mínimo quatro sílabas, não havendo restrição quanto à extensão máxima;
- f) a análise de diálogos espontâneos gravados sem a ciência dos locutores, no caso da interceptação telefônica de conversação ao telefone celular (gravação desavisada) e com a ciência dos locutores, no caso do registro ambiental de entrevista semidirigida (gravação avisada);
- g) o estabelecimento de 130ms como limiar na caracterização das pausas silenciosas e a admissão dos intervalos de fala com pausa preenchida e evento(s) de disfluência(s) como útil no cálculo da TE e como não útil no cálculo da TA.

O método estruturado a partir de tais proposições foi implementado, o que permitiu a avaliação da pertinência das decisões teórico-metodológicas tomadas neste estudo, sendo essa a seguir apresentada.

Quanto à adoção da sílaba fonética como unidade linguística de contagem, essa se mostrou apropriada no sentido da funcionalidade e, como previsto, favorável à comparabilidade entre os resultados obtidos e os já publicados. Exceções ao posto deveram-se à má qualidade acústica especialmente encontrada na gravação desavisada (provinda de interceptação telefônica), onde se tem banda de frequência restrita e típica baixa relação sinal/ ruído, o que conduziu à percepção de que a adoção de uma unidade menor, por exemplo, o segmento (BINNENPOORTE, 2005; FONAGY e MAGDICS, 1960; MIRGHAFORI et al., 1996; PFITZINGER, 1998; TROUVAIN et al. 2001; VERHASSELT e MARTENS, 1996), seria aparentemente inviável em registros sonoros com essas limitações qualitativas.

A imprecisão conceitual ainda vigente acerca da perspectiva fonética de sílaba, em contrapartida, dificultou sobremaneira a definição do critério a ser considerado na identificação da unidade no sinal acústico. A identificação automatizada dos elementos vocálicos, correspondentes a unidades silábicas, tem como referência o parâmetro intensidade, em decibéis (dB), pontuando como sílaba os picos de intensidade antecedidos e sucedidos por um decaimento pré-definido (JONG e WEMPE, 2009).

No estudo em tela, no entanto, em razão de se considerar a fala verdadeiramente naturalística, rica em excepcionalidades e eventos não regulares, acreditou-se que a adoção de semelhante recurso comprometeria a detecção

fidedigna das vogais silábicas presentes no texto oral, exigindo grande trabalho de retificação. Assim, optou-se pela identificação manual (não demeritória, tampouco mais subjetiva, uma vez que à marcação automática seguem conferência e retificações do investigador), que implicou o estabelecimento de um critério diferente da intensidade, uma vez que a mesma não foi controlada durante a captura dos sinais (gravação desavisada e avisada).

No empreendimento recorreu-se aos vestígios acústicos associados aos segmentos vocálicos (conforme Subseção 1.2.2.1, Capítulo 1) e ao que diferenciasse, segundo a literatura, vocóides silábicos de não silábicos, tendo sido de extrema relevância o postulado por Meneses (2012), que discute a caracterização das vogais desvozeadas no PB, também encontradas no presente estudo, e por Padgett (2008), que contrapõe os elementos acústicos definidores de vogais e semivogais.

Poder-se-ia simplesmente referir que foi adotada neste estudo a unidade VV⁶⁹ como elemento de contagem, já que, de fato, trabalhou-se na busca por vogais consecutivas. No entanto, conforme exposto na Subseção 2.1, Capítulo 2, tal unidade é, por conceito, pausa-includente, o que inviabiliza sua adoção em concepção original como unidade de contagem no estabelecimento da TA. Havendo o interesse na comparação do comportamento das duas taxas em estudo (TE e TA), tornou-se inevitável a admissão de um único tipo de unidade de contagem, no caso, a sílaba fonética, identificada a partir de vestígios acústicos associados às vogais plenas.

Quanto à direção da fórmula de cálculo, se direta ou indireta (via duração média das sílabas), foi compulsória a adoção da primeira, visto ser essa a amplamente utilizada nos estudos sobre o tema. Destaca-se que, no entanto, sendo de interesse, de posse do número de sílabas e da duração temporal relativa a cada intervalo de fala, facilmente o cálculo pode ser refeito, como, a exemplo, do que ocorreu por ocasião da verificação do fenômeno de encurtamento antecipatório, situação em que se correlacionou o número de sílabas produzido nos intervalos interpausais e a duração média da sílaba (obtida a partir do cálculo inverso da taxa).

O fato de a pesquisa ter abarcado ambos os escopos de observação (global e local, o último verificado tanto através das próprias múltiplas tomadas locais quanto

⁶⁹ Definida na Subseção 2.1 do Capítulo 2.

através da média aritmética dessas), deveu-se à importância de se definir o custo-benefício da aplicação de uma ou outra forma de mensuração na rotina de Comparação Forense de Locutor, em que se têm normalmente curtos prazos para a entrega de Laudos Periciais. Além disso, a obtenção do coeficiente indicativo do potencial discriminatório das taxas, foco principal deste estudo, requer informações acerca da variabilidade intrassujeito e intersujeitos, firmada (conforme Subseção 4.1.1.4 deste capítulo) global e localmente.

No que se refere à exigência de no mínimo quatro sílabas para a aceitação do intervalo de fala como útil, seja turno de fala ou intervalo interpausal, constatou-se que nos *corpora* dos sete sujeitos pesquisados a maior parte dos intervalos de fala com um número de sílabas menor do que o mínimo exigido abriga, como já apontado na literatura (JESSEN, 2007), segmento(s) alongado(s), o que repercute invariavelmente na diminuição da taxa (TE ou TA) obtida.

Já com respeito à opção de não restrição quanto ao tamanho máximo do intervalo de fala feita neste estudo, não foi observado impacto outro senão o do efeito de encurtamento antecipatório, fisiologicamente previsível e neste estudo foi confirmado como significativo, sendo a ingerência de tal efeito progressivamente maior à medida que aumenta o número de unidades silábicas no intervalo.

Neste estudo, os intervalos de fala excessivamente longos referem-se especificamente aos turnos de fala, já que a duração dos intervalos interpaisais é limitada, ao menos, pela capacidade expiratória do indivíduo. Na qualificação dos turnos de fala (como útil ou não útil) observou-se que os excessivamente longos acabaram sendo naturalmente descartados em razão de apresentarem fala sobreposta, ininteligibilidade à oitiva ou à inspeção acústica, vocalização não linguística, menos do que quatro sílabas, interrupção na transmissão do sinal e/ou impedimento ético à divulgação (critérios de descarte de material apresentados na Subseção 3.3.4 do Capítulo 3).

Nesse sentido, a adoção de turnos de fala mostrou-se um tanto problemática na gravação avisada, onde o entrevistador incentiva a narrativa longa, na intenção de provocar o envolvimento do falante no discurso e a diminuição da atenção dispensada à fala. Alguns dos extensos (e proveitosos) turnos de fala, nos quais se observava a ocorrência (mesmo que única) de qualquer um dos elementos justificadores de descarte de material, tiveram de ser descartados.

Já quanto à utilização de intervalos interpausais, destaca-se como conflitiva apenas a duração do limiar adotada (130ms), tendo o mesmo por vezes se mostrado excessivo (foram encontradas no *corpus* algumas pausas visivelmente linguísticas cujas durações são inferiores aos 130ms estabelecidos).

A utilização de fala espontânea (especialmente a definida por Labov, 1972, como casual, encontrada neste estudo na gravação desavisada), excepcionalmente bem-vinda uma vez que, na comparação forense, é sobre esse tipo de material que se debruçam os peritos, mostrou-se, como previsto, de análise complexa. No desenvolvimento do trabalho, diversas circunstâncias não convencionais, para as quais não se encontrou suficiente discussão e apontamento de conduta na literatura, tiveram de ser enfrentadas, como a situação das vogais com configurações acústicas não canônicas, dos alongamentos segmentais não associados à disfluência, mas sim a ênfases típicas da fala espontânea, e do acolhimento ou não da pausa preenchida no material a ser submetido ao cálculo da TA. A disfluência foi encontrada como elemento altamente recorrente nos *corpora* dos sujeitos pesquisados, merecendo cuidadosa abordagem, uma vez que a produção de segmentos consonantais desacompanhada de vogais agrega tempo à duração total do intervalo de fala, afetando diretamente o resultado da taxa (levando à diminuição da mesma).

Finalizando, cabe salientar que a investigação a partir de materiais de fala que se contrapõem quanto à ciência ou não da gravação (realidade inelutável na Comparação Forense de Locutor e que, ao que se sabe, tem investigação inédita, ao menos no Brasil), frente aos resultados obtidos e à possibilidade de contribuição efetiva à aplicação forense, revelou-se prestimosa.

4.2.2 Resposta aos objetivos estabelecidos

Adiante, responder-se-á de maneira direta aos objetivos formulados, apresentados na Introdução deste documento e na Subseção 3.1.3 do Capítulo 3.

O objetivo geral desta tese foi o de investigar o potencial individualizante (de indivíduo) da TE e da TA, com vistas à verificação da efetividade da incorporação dessas ao rol de parâmetros técnico-comparativos utilizados na realização da perícia de Comparação de Locutor. Entende-se que a proposição geral foi alcançada uma vez que, a despeito do reduzido número de participantes que atenderam aos pré-

requisitos de ingresso na pesquisa, obtiveram-se resultados passíveis de submissão a tratamento estatístico, o qual permitiu o estabelecimento do potencial discriminatório de falantes das taxas em investigação (exposto na Subseção 4.1.1.5 deste capítulo).

Quanto à pretensão basilar de idealizar e executar um método científico de mensuração de TE e de TA, de escopo global e local, não só passível de ser integrado, mas dirigido às necessidades da Fonética Forense, que convivesse satisfatoriamente com as limitações na qualidade e na quantidade de material de fala, assim como com o singular *gap* temporal existente entre os áudios dos locutores das amostras confrontadas, é apresentada, na Subseção 3.3 do Capítulo 3, uma detalhada descrição do método desenvolvido neste estudo, tendo sido o mesmo concebido com base no que é exclusivamente factível dentro da examinação comumente perpetrada na realização da perícia de Comparação de Locutor.

Quanto à variabilidade intersujeitos e intrassujeito de ambas as taxas temporais estudadas, o estudo conduziu à constatação de que, no que se refere ao tipo de taxa temporal, das duas pesquisadas, é a TA a taxa intraindividualmente mais estável (resultado compatível com a hipótese de trabalho estabelecida).

A TE média dos sete sujeitos amalgamados mostrou-se inferior à TA média assim também obtida ($5,37 \text{ síl/s} < 6,20 \text{ síl/s}$)⁷⁰, achado condizente com o apontado na literatura, tendo sido claramente observado durante a aplicação do procedimento de cálculo das TEs dos sujeitos o impacto causado pelos trechos de pausa (seja silenciosa, inspiratória ou preenchida). O achado é atribuído à pausa admitida no cálculo da TE, inerente ao turno de fala, que eleva o tempo total do intervalo, propiciando, conseqüentemente, valores de taxa inferiores aos da TA.

A rotina criada permitiu o cômputo global e local das taxas estudadas, devendo-se reiterar, contudo, que não foi objetivada, como feito entre outros por Dancovičová (1997, 2001) e Hansson (2002), a investigação da variação local com vista à detecção de padrão sistemático de variação ou à identificação de domínio linguístico de tal ocorrência.

Na verificação da existência de diferença entre as formas de mensuração das taxas constatou-se que não há, na TA, diferentemente do que ocorre na TE (conforme Teste t de amostra pareadas, exibido na Tabela 6), diferença significativa

⁷⁰ Valores resultantes da média aritmética entre as taxas globais e locais médias apresentadas na Tabela 4.

entre os escopos de observação global e local médio. Portanto, a hipótese de que não haveria diferença significativa entre os valores de TE e de TA globais em relação aos obtidos via média das mensurações locais mostrou-se apenas parcialmente procedente, uma vez que somente na TA é significativamente indiferente a forma de mensuração. Em ambos os tipos de taxa (TE e TA), a partir dos dados dos sujeitos pesquisados, é preferível o emprego da média das observações locais (menos variável) ao invés da observação global. Apesar da não significância encontrada no caso da TA, tanto na TE quanto na TA as formas de mensuração global e local média mostram-se fortemente correlacionadas, apontando um comportamento similar e em mesmo sentido.

As constatações relativas à variabilidade intersujeitos e intrassujeito e ao potencial individualizante da TE e da TA são plenamente compatíveis com a hipótese de trabalho formulada, a saber, a de que a TA, devido a sua maior estabilidade é, das taxas temporais de fala estudadas, a mais indicada para utilização no confronto de voz forense, sendo capaz de, satisfatoriamente (conforme CCI obtido, exibido na Tabela 13) distinguir falantes.

Especificamente quanto ao estabelecimento do potencial individualizante da TE e da TA, decorrente da relação entre a variabilidade intersujeitos e a variabilidade intrassujeito, considerando-se os *corpora* admitidos, teve-se revelado que a TA apresenta, quando comparada à TE, maior poder discriminatório de indivíduo, ou seja, um maior potencial individualizante.

Na análise de diferenças por grupo, especificamente quanto ao sexo e ao tipo de gravação, foram concordantes os resultados obtidos a partir da análise de variância com os obtidos a partir do Teste t de amostras pareadas. Em ambos os procedimentos foi encontrada não significância quanto ao sexo, mas significância quanto ao tipo de gravação, essa, contudo, restrita à TE. O achado é condizente também com os resultados obtidos na regressão linear realizada (exatamente a com cinco variáveis independentes, apresentada na Tabela 16), em que se têm os maiores valores de coeficiente associados às variáveis sexo e tipo de gravação, sendo esses destacadamente mais expressivos na TE do que na TA.

Contrastando-se os achados da correlação entre os valores de TE e de TA empregados a partir do *corpus* considerado e as variáveis independentes idade, sexo, escolaridade, *gap* temporal, tipo de gravação e tamanho do intervalo de fala

com as respectivas hipóteses de trabalho (apresentadas na Subseção 3.1.4, Capítulo 3), tem-se a síntese abaixo apresentada:

- a) a variável idade evidenciou correlação não significativa com a TE e com a TA, nas duas formas de mensuração, mas tendência concordante com o apregoado na literatura (JACEWICZ et al., 2009; MALÉCOT et al., 1972; RAMIG, 1983; VERHOEVEN et al., 2004; YUAN et al., 2006) e com o aqui admitido, de diminuição das taxas à medida que aumenta a idade do indivíduo. Tal comportamento, no entanto, diferentemente do previsto, não se confirmou quanto ao *gap* temporal (cujo aumento está associado ao avanço na idade do indivíduo), observando-se, nesse caso, significativamente na TA e em nível de tendência na TE, que o aumento do tempo existente entre a gravação desavisada e avisada dos sujeitos conduz ao aumento da TE e da TA⁷¹. Quanto à possível interferência indireta do tamanho do intervalo de fala no comportamento da variável idade em relação às taxas estudadas, verificou-se que, nos sujeitos pesquisados, diferentemente do apontado por Quené (2008), o número de sílabas do intervalo interpausal não está correlacionado com a idade do sujeito.
- b) a variável sexo apresentou variância não significativa tanto na TE quanto na TA, nas duas formas de mensuração (global e local média), divergindo, em primeira análise, da hipótese de trabalho (formulada a partir de Hewlett e Rendall, 1998; Jacewicz et al., 2009, 2010; Kendal, 2009; Verhoeven et al., 2004). Ao se verificar possíveis diferenças no comportamento do grupo por sexo em razão do tipo de gravação, no entanto, encontraram-se, em nível de tendência, na gravação desavisada, os homens como os falantes com maiores taxas. O achado sugere que na fala verdadeiramente naturalística, confirma-se o apontamento prevalente na literatura: a preponderância do sexo masculino na aplicação de maiores valores de taxas (TEs e TAs). Na gravação avisada (e, excepcionalmente, na mensuração global da TA),

⁷¹ A referida discrepância entre o comportamento das variáveis idade e *gap* temporal pode dever-se ao limitado número de dados, assim como a não estratificação da idade. De fato, a correlação de Pearson realizada entre as variáveis mostrou-se não significativa ($p=0,656$), contudo, a consideração de uma amostra maior e mais equilibrada permitiria verificar a hipótese de existência de um comportamento não linear (supõe-se, em parábola), envolvendo as variáveis idade e *gap* temporal em relação às taxas estudadas (TE e TA).

contudo, foram as mulheres que apresentaram as maiores taxas, superioridade encontrada também por Syrdal (1996) e Yuan et al. (2006) em dados provenientes de gravação de conversação telefônica consentida.

- c) a variável tipo de gravação, por sua vez, mostrou-se significativa à análise de variância, exposta nas Tabelas 9 e 10, nas duas formas de mensuração da TE, o que não ocorreu com a TA, embora em ambos os casos (tanto na TE quanto na TA) tenha se confirmando, em nível de tendência, a previsão de que a ciência de estar sendo gravado interfere na taxa empregada pelo indivíduo, diminuindo-a.
- d) a variável escolaridade apresentou correlação não significativa com as quatro variáveis de resposta (TE global, TE local média, TA global e TA local média), com tendência inversa à hipótese formulada, uma vez que a partir dos *corpora* dos sujeitos pesquisados o aumento da escolarização conduziu ao aumento de ambas as taxas, nas duas formas de mensuração.
- e) a variável tamanho do intervalo de fala mostrou correlação significativa com as TEs (globais e locais médias) mas não com as TAs (independente da forma de mensuração), evidenciando, no entanto, tendência condizente com o previsto em hipótese (o aumento da duração do intervalo de fala, seja ele um turno de fala ou um intervalo interpausal, conduz ao aumento do tempo de fala nele empregado).

Quanto ao fenômeno de encurtamento antecipatório, verificado através da correlação entre as variáveis duração média da sílaba e tamanho do intervalo de fala (equivalente, nesse caso, ao número de sílabas presentes em um intervalo interpausal), observou-se que o aumento do número de sílabas do intervalo interpausal implica necessariamente na diminuição da respectiva duração média das sílabas, ou seja, nos sujeitos pesquisados, o encurtamento antecipatório mostra-se significativamente presente. O achado é concordante com a hipótese de trabalho formulada, que previa a existência de uma relação inversamente proporcional entre o número de sílabas do intervalo interpausal e a duração média das sílabas nele contidas.

O comportamento das variáveis preditoras idade, sexo, escolaridade, *gap* temporal, tipo de gravação e tamanho do intervalo de fala nos dois modelos de

regressão linear realizados mostrou-se em essência similar. O primeiro deles contou com 14 observações (relativas aos dois tipos de gravação de cada um dos sete sujeitos admitidos) e cinco variáveis independentes enquanto o segundo, restrito às tomadas locais em razão da entrada da variável tamanho do intervalo de fala (determinada necessariamente por durações individuais), contou com 1.287 observações e seis variáveis independentes.

As variáveis *gap* temporal e tipo de gravação, delatadas como significantes no modelo de regressão baseado na média das tomadas locais (a primeira somente na TA e a segunda somente na TE), mostraram-se significativas preditoras também no segundo modelo de regressão, o baseado nas múltiplas tomadas locais. Neste também se encontrou como significativa preditora a variável tamanho do intervalo de fala, achado concordante com o da correlação efetuada a partir das múltiplas tomadas locais (apresentada na Subseção 4.1.2.3 do presente capítulo) e indicativo de que, na TE, procede a hipótese de que intervalos de fala maiores são produzidos a taxas mais altas.

Os comportamentos díspares encontrados, mencionados nessa subseção, naturalmente desencadeiam os seguintes questionamentos:

- a) por que somente na TE a forma de mensuração, o tipo de gravação e o tamanho do intervalo de fala são significativos?
- b) por que somente na TA o *gap* temporal é significativo?
- c) por que o aumento da escolaridade tende a fazer aumentar a TE e a TA?
- d) por que os homens⁷² tendem a controlar mais do que as mulheres as suas TEs e TAs quando sabem que estão sendo gravados?

4.2.3 Justificativas para os comportamentos não previstos

Nesta subseção são apresentadas e discutidas possíveis justificativas para os comportamentos obtidos que tenham se mostrado, de algum modo, divergentes em relação ao inicialmente previsto.

Na explanação, nos casos em que a questão envolver as variáveis inéditas tipo (ciência ou não) de gravação, *gap* temporal e escolaridade, a argumentação

⁷² Recupera-se que neste estudo os sujeitos, de ambos os sexos, são envolvidos (como suspeitos, indiciados ou réus) em crimes de tráfico de entorpecentes e/ou homicídio.

exploratória basear-se-á nos fundamentos considerados, buscando-se encontrar neles algum elemento elucidativo, que explique o observado. Quando a questão referir-se à restrição de comportamento a uma das taxas em especial, a proposição remeter-se-á, por lógica, às diferenças existentes entre as propriedades do material de fala admitido para cálculo da TE e da TA (a saber, na TE, acolhem-se as pausas de todos os tipos, assim como a fala disfluente, e na TA, somente a fala encadeada fluente, com alongamentos contextualizados, relacionados à ênfase) e/ou ao que reflete cada uma dessas medidas.

A análise das questões gênero-específicas (exatamente a relativa ao fato dos sexos lidarem diferentemente com a ciência da gravação) considerará preceitos sociolinguísticos (conforme Seção 1.3, Capítulo 1), e questões próprias à área de Criminologia.

4.2.3.1 Por que somente na TE a forma de mensuração, o tipo de gravação e o tamanho do intervalo de fala são significativos?

A existência de diferença significativa entre as formas de mensuração global e local média na TE e não na TA pode estar relacionada (a) à forma de cálculo de um e outro tipo de mensuração, (b) à diferença existente entre as propriedades da fala submetida ao cálculo da TE e da TA e/ou (c) à maior divergência existente entre as médias globais e locais na TE se comparada à TA, conforme exposto a seguir.

O primeiro caso refere-se à diferença que há entre uma TE ser obtida através do cálculo formado pela divisão do resultado da soma do número de sílabas produzidas em cada um dos turnos úteis de fala, separadamente na gravação desavisada e avisada, pelo resultado da soma das respectivas durações (esquema de cálculo formalizado na opção “TE e TA global” da Figura 26, Subseção 3.3.4, Capítulo 3) ou ser obtida por meio da média aritmética das TEs calculadas em cada um dos turnos úteis de fala, igualmente em um e outro tipo de gravação, com a divisão do número de sílabas neles contidas pelas suas respectivas durações (esquema de cálculo formalizado na opção “TE e TA local média” da mesma figura). Por óbvio uma dada forma de mensuração é igualmente aplicada na obtenção tanto de uma quanto de outra taxa, mas pode-se cogitar que diferenças relativas à formulação do cálculo revelem-se mais impactantes em um material de fala pausa-

includente (ressalvando-se que há ainda nesse tipo de material de fala os elementos de disfluência porventura produzidos) do que em um pausa-excludente.

O segundo caso diz respeito ao fato de que, na TE, diferentemente do que ocorre na TA, são acolhidos como úteis os intervalos de fala com disfluência, bem como as pausas silenciosas e preenchidas existentes no discurso.

A explicitação do terceiro caso, relativo à maior divergência entre as médias globais e locais na TE do que na TA, se beneficia da caracterização do *corpus* de um dos sujeitos da pesquisa. A exemplo tem-se o sujeito identificado neste estudo como “S1”, que teve admitido para fins de cálculo da TE (conforme consta na Tabela 21, a seguir) 52 turnos de fala úteis na gravação desavisada e 40 turnos de fala úteis na gravação avisada. Esses, quando submetidos ao cálculo global, evidenciam, respectivamente, as taxas de 6,51 síl/s e de 4,46 síl/s e, quando submetidos ao cálculo via média das taxas individuais, evidenciam, respectivamente, as taxas de 6,61 síl/s e de 4,83 síl/s. Já para o cálculo da TA, o mesmo sujeito contou com 60 tomadas desavisadas e 87 tomadas avisadas, que propiciaram a obtenção de, respectivamente, 6,92 síl/s e 6,41 síl/s na mensuração global e 6,99 síl/s e 6,43 síl/s na mensuração local média.

Tabela 21 – Propriedades da amostra de fala do sujeito identificado como “S1”

Propriedade	TE		TA	
	Desavisada	Avisada	Desavisada	Avisada
Nº total de intervalos de fala	52	40	60	87
Taxa global (síl/s)	6,51	4,46	6,92	6,41
Taxa local média (síl/s)	6,61	4,83	6,99	6,43

Legenda: síl= sílaba, s= segundos.

Fonte: A autora (2013).

A ilustração recém-apresentada permite a constatação de que tanto na TE quanto na TA, nos dois tipos de gravação explorados (desavisada e avisada), as taxas obtidas via média das tomadas locais alcançam valores maiores do que aquelas obtidas globalmente (o que acontece em 11 dos 14 sinais de fala submetidos ao cálculo da TE e em 7 dos 14 sinais de fala submetidos ao cálculo da TA). Ao analisar-se a diferença entre as médias associadas às duas formas de mensuração em cada uma das taxas (respectivamente 0,10; 0,37; 0,07 e 0,02), percebe-se que há uma diferenciação maior entre a taxa global e a taxa local média na TE do que na TA, tanto na gravação desavisada quanto na gravação avisada, o

que possivelmente justifica, a despeito da não diferenciação significativa entre as formas de mensuração evidenciada ao Teste F e à correlação, a significativa diferença encontrada entre as formas de mensuração ao Teste t.

Dessa forma, dos três argumentos oferecidos como possível resposta para a excepcionalidade em questão (a de existência de diferença significativa entre os tipos de mensuração global e local médio restrita à TE), é o último o mais afeito à percepção de que a explicação pretendida aparentemente fundamenta-se em efeitos matemáticos inerentes à análise estatística de dados, pois é previsível que maiores amplitudes entre médias ofereçam maior probabilidade de obtenção de diferenças significativas entre os fatores (de variável) aos quais essas estão associadas. Considerando-se, ainda, que ambas as formas de mensuração partem de um mesmo conjunto de dados, com igual número de turnos de fala (conservando-se em cada um desses o número de sílabas e a duração), entende-se como mais condizente o argumento quantitativo e, secundariamente, o relativo às diferenças existentes entre as propriedades da fala submetida a cálculo de um e outro tipo de taxa.

Adiante, a fim de elucidar justificativas para o fato da variável tipo de gravação, especificamente a ciência de estar sendo gravado, ter se mostrado significativa somente na TE, cabe resgatar, ressalvadas as discrepâncias decorrentes dos diferentes delineamentos de pesquisa, os resultados obtidos nos dois estudos de configuração teórico-metodológica mais próxima à do em tela, Jessen (2007) e Künzel (1997), tendo sido nesses também considerados os tipos de gravação presencial e ao telefone.

Künzel (1997), com relação à fala em língua alemã de 10 sujeitos, ao comparar os dois tipos de gravação pesquisados (presencial e ao telefone), diferentemente do obtido no presente estudo, não encontrou diferença significativa entre a TE (global e local média) gerada presencialmente ou ao telefone.

Jessen (2007), também investigando a fala alemã, porém em um grupo maior de participantes (composto de 100 sujeitos), limitada a observação, no entanto, ao sexo masculino e à TA, obteve resultado concordante com o de Künzel (1997) e com o do presente estudo: diferença não significativa no que se refere ao tipo de gravação na TA, com TAs locais maiores no registro presencial do que naquele ao telefone.

Ao compararem-se os resultados do estudo em tela com os divulgados pelos dois recém-mencionados, constata-se a limitada concordância existente entre os achados, uma vez que foi observada paridade somente no que se refere a não significância entre os tipos de gravação na TA. O presente estudo tem resultado discordante com o apontado por Künzel (1997), pois evidenciou a existência de diferença significativa quanto ao tipo de gravação na TE, e resultado discordante com o apontado pelos dois trabalhos em questão (KÜNZEL, 1997; JESSEN, 2007), pois encontrou categoricamente maiores valores de taxa no registro produzido ao telefone. É possível aventar que a falta de congruência entre os três estudos referidos (Künzel, 1997; Jessen, 2007 e o em tela) derive do fato de que, nos primeiros, o diálogo telefônico considerado foi produzido com ciência e anuência dos locutores.

Ressalta-se, entretanto, que não pode ser descartada a hipótese de que a diferença existente entre as taxas provindas dos distintos tipos de gravação decorra não só da mencionada ciência de gravação, mas também de particularidades que permeiam a gravação presencial (neste estudo intitulada avisada), entre elas, o tipo de situação comunicativa (entrevista semidirigida), o estilo de fala empregado (bem como possíveis variações desse ao longo do discurso), o conteúdo abordado (normalmente relato de vida pessoal, por vezes, com declarações emocionalmente fortes de privação familiar) e o grau de relação entre os locutores do diálogo (peritos, de um lado, e indivíduos associados à atividades ilícitas, de outro, sempre desconhecidos entre si).

A despeito da ciência ou não da gravação, pode-se pontuar que a variável tipo de gravação abarca também especificidades concernentes aos estilos de fala próprios a cada um dos materiais de fala analisados (conversação telefônica interceptada *versus* conversação presencial), enquadráveis a partir de Labov (1972) como, respectivamente, casual e espontâneo. Contudo, a partir do manejo dos dados e da aproximação com o referencial teórico especializado, formulou-se o convencimento de que o maior impacto do tipo de gravação excepcionalmente encontrado na TE (do que na TA) decorre, de alguma forma, das propriedades do material de fala analisado e relaciona-se, especialmente, com os mecanismos empregados pelos falantes na distribuição das pausas e/ou com a incorporação de um número maior de pausas quando da ciência da gravação.

Nessa perspectiva, o locutor, ao saber que está sendo gravado, controla as taxas temporais empregadas na fala (TE e TA), como demonstrado nas subseções pertinentes (“4.1.1.4”, “4.1.1.6” e “4.1.2.3” deste capítulo), cerceamento que se mostra mais expressivo (atingindo níveis significativos, segundo os p-valores relativos ao “Par 1” e ao “Par 2” exibidos na Tabela 8, apresentada na Subseção 4.1.1.3) quando o sujeito tem à disposição os recursos relativos à descontinuidade da cadeia de fala (pausas) e os atinentes à fluência (alongamentos segmentais, repetições de sílabas ou palavras, entre outros). Pode-se aventar, nesse caso, que são feitas adaptações estilísticas em decorrência do nível de atenção próprio ao contexto de manifestação da fala (LABOV, 1972), visando o ajuste da fala ao interlocutor (“*audience design*”, conforme Bell, 1984) e/ou em resposta a fatores intrínsecos do indivíduo (“*speaker design*”, conforme Schilling-Estes, 2002).

Quanto ao tamanho do intervalo de fala, apontado por alguns autores como significante na TA (JACEWICZ et al., 2009, 2010; QUENÉ, 2008), há, de acordo com a amostra em análise neste estudo, significância na correlação exclusivamente nas TEs (global e local média).

Volvendo às propriedades duracionais dos *corpora* dos sujeitos investigados, sendo as relativas aos turnos de fala (de onde se extraíram as TEs) apresentadas na Tabela 2 e as relativas aos intervalos interpausais (de onde se extraíram as TAs) apresentadas na Tabela 3 (ambas localizadas na Subseção 3.3.4 do Capítulo 3), verificou-se que a média da duração média dos múltiplos intervalos úteis de fala dos sete sujeitos da pesquisa, nas gravações desavisada e avisada, mostrou-se destacadamente maior na TE (120s e 755ms) do que na TA (81s e 656ms).

Ao considerar-se que a duração média dos turnos de fala é substancialmente maior do que a dos intervalos interpausais e que em intervalos de fala mais longos são empregadas maiores taxas, entende-se como coerente a obtenção de significância à correlação exclusivamente na TE. Em suma, conjectura-se que a significância extraordinária, existente entre o tamanho do intervalo de fala e a TE global e local, decorra verdadeiramente das propriedades duracionais inerentes aos turnos de fala, esses comprovadamente mais longos do que os intervalos interpausais.

Além disso, é possível que, adotando-se um maior rigor na exclusão dos intervalos interpausais nos quais constem alongamentos segmentais (tendo sido

esses, neste estudo, acolhidos como úteis quando denotando ênfase), constate-se significância também na correlação entre tamanho do intervalo de fala e TA.

Com o exposto, evidencia-se que um melhor entendimento acerca das excepcionalidades associadas à TE (mencionadas no questionamento que figura como título desta subseção) requer o estudo da relação existente entre as taxas temporais de fala (TE e TA) e as pausas observadas no discurso (caracterização, distribuição e frequência). Adicionalmente, faz-se necessária a investigação acerca das interferências causadas pelos elementos de disfluência (repetições, alongamentos e falsos começos), inevitavelmente presentes na fala espontânea e recorrentes na fala naturalística investigada neste estudo, visto que esses, como já mencionado, contribuem para o alargamento do tempo total do enunciado.

4.2.3.2 Por que somente na TA o *gap* temporal é significativo?

A existência de um tempo transcorrido (*gap* temporal) entre os áudios nos quais constam as falas a serem cotejadas é a realidade dominante no contexto de realização da perícia de Comparação de Locutor, desconhecendo-se, no entanto, como já informado anteriormente, estudo prévio que o tenha abordado.

A significância da mencionada variável restrita à TA (nas duas formas de mensuração), reforça a importância do critério de contemporaneidade entre os áudios durante a avaliação da adequação do material a ser utilizado nesse tipo de perícia. Em razão do *gap* temporal na prática em Fonética Forense normalmente não ultrapassar 4 a 5 anos e de os locutores periciados, via de regra, encontrarem-se em faixa etária de certa estabilidade laríngea (em que não se prevê muda vocal ou presbifonia⁷³), são valorizados na tarefa, normalmente, apenas os critérios de quantidade e qualidade do material, o que faz do resultado em questão um achado destacável na literatura da área.

Quanto à variável *gap* temporal ter se mostrado significante exclusivamente na TA, seria razoável estimar (considerando-se que a mesma reflete a habilidade do indivíduo em executar os gestos articulatórios associados aos sons da fala e de

⁷³ A muda vocal está relacionada às modificações orgânicas (no caso, laríngeas) típicas da puberdade, ocorrendo no homem por volta dos 13 a 15 anos e nas mulheres por volta dos 12 a 14 anos. Já a presbifonia é o envelhecimento vocal associado ao envelhecimento do indivíduo (especialmente ao de sua laringe), tendo início vinculado ao estado de saúde física e psíquica do indivíduo, considerando fatores constitucionais, raciais, hereditários, alimentares, sociais e ambientais (BEHLAU e PONTES, 1995).

concatená-los na formação da sequência segmental desejada e, ainda, que os participantes do presente estudo são relativamente jovens (14 a 33 anos) e, mesmo quando aprisionados, socialmente ativos), que o aumento da idade ocorrido no lapso temporal existente entre as gravações analisadas indiretamente interfira no comportamento da taxa em razão de um progressivo aprimoramento das habilidades neuromusculares do indivíduo.

Tal proposição inicial, no entanto, foi refutada devido à ausência de correlação ($\alpha = 5\%$) obtida entre as variáveis idade e *gap* temporal, conforme visualizado na Tabela 22.

Tabela 22 – Correlação entre as variáveis idade e *gap* temporal

Medidas	Idade x <i>Gap</i> temporal
Correlação de Pearson	0,131
P-valor (bicaudal)	0,656

Fonte: A autora (2013).

Ademais, a despeito da significância da variável *gap* temporal encontrada na TA, dada a contradição existente entre o comportamento das variáveis idade e *gap* temporal (expressa pela variação dos coeficientes em sentido inverso, conforme observado nas Tabelas 17 a 20, apresentadas nas subseções “4.1.1.7” a “4.1.2.3” deste capítulo), faz-se necessária a reavaliação, com mais rigor e em um conjunto de dados direcionado para tal fim, do real condicionamento gerado por essa variável.

A ressalva quanto ao resultado relativo à variável em questão justifica-se pela contradição que há na discordância entre o comportamento das variáveis idade e *gap* temporal em relação às taxas estudadas. Destaca-se que o aumento no tempo interveniente entre uma e outra gravação do sujeito (no âmbito forense, entre o áudio questionado e o áudio padrão) é irremediavelmente acompanhado do avanço da idade do sujeito, tendo sido previsto, em razão dessa sinergia inequívoca, que os coeficientes variariam em um mesmo sentido, o que efetivamente não ocorreu.

4.2.3.3 Por que o aumento da escolaridade tende a fazer aumentar a TE e a TA?

A hipótese de trabalho relativa à variável escolaridade inicialmente formulada, a saber, a de que o aumento da escolaridade, supostamente relacionado a uma maior aproximação do sujeito com a norma culta da língua, implicaria a redução das

TEs e das TAs empregadas pelos indivíduos não foi confirmada a partir dos dados dos sujeitos pesquisados.

Não havendo, ao que se sabe, estudo sobre taxas temporais de fala que tenha admitido tal variável (a escolaridade) como independente, encontra-se esta discussão limitada a suposições exploratórias, entre elas a que resgata o processamento cognitivo da linguagem e a que revê a hipótese de trabalho admitida.

Na perspectiva do processamento de linguagem, tem-se que o aumento da escolaridade leva o indivíduo a um maior contato com materiais escritos. A informação fonológica contida nesse tipo de material, para que se mantenha na memória operacional implica, segundo Gathercole e Baddeley (1993), o ensaio articulatório (implementado ou em nível subvocal).

Ao rever-se a hipótese de trabalho, baseada na previsão do impacto causado pela maior aproximação do falante em relação à norma culta da língua, pode ser considerado que o aumento na escolarização implica maior exposição comunicativa, o que promove a expansão do vocabulário (com conseqüente facilitação do acesso lexical), o incremento na capacidade de organização de ideias e os automatismos relacionados à linguagem, condições favoráveis à continuidade e à fluência da fala.

A possibilidade de interferência da escolarização nos modos de funcionamento cognitivo é defendida por Oliveira (1999), para quem a estrutura das organizações cognitivas decorre das demandas instituídas, propondo, a escola, atividades que exigem descontextualização⁷⁴, metacognição e reflexão por parte dos alunos.

Obviamente ambas as proposições explicativas apresentadas para justificar o comportamento da variável escolaridade (a relativa ao processo de ensaio articulatório e a que diz respeito à particularização do funcionamento cognitivo) carecem de investigação apropriada, delineada com vistas à verificação científica do referido.

⁷⁴ Diferença entre o pensamento descontextualizado e contextualizado é encontrada em Weber (2005, p.3). A autora caracteriza o primeiro como “abstraido de experiências pessoais e da realidade concreta” e o último, normalmente associado a sujeitos sem ou com pouca escolarização, como “atrelado à realidade vivenciada, concreta e imediata”.

4.2.3.4 Por que os homens tendem a controlar mais do que as mulheres as suas TEs e TAs quando sabem que estão sendo gravados?

As falas analisadas neste estudo permitem a constatação, em nível de tendência, de que, quando há ciência da gravação, os sujeitos do sexo masculino contêm mais suas TEs e TAs (globais e locais médias) do que os do sexo feminino, comportamento curiosamente diferente do que ocorre, no entanto, na fala verdadeiramente naturalística (gravação desavisada), situação na qual, em concordância com o predominantemente apontado na literatura (BYRD, 1994; HEWLETT e RENDALL, 1998; JACEWICZ et al., 2010; KENDAL, 2009; QUENÉ, 2008; VERHOEVEN et al., 2004), são os homens que, em prevalência, apresentam as maiores taxas.

Nesse sentido, Thomas (2011, p.189, tradução nossa) aponta que “alguns tipos de interação social induzem os locutores a planejar mais cuidadosamente as suas sentenças”. No entanto, tal achado vai, de certa forma, de encontro com a afirmação de Labov (1972) sobre mostrarem-se as mulheres mais sensíveis do que os homens aos valores sociolinguísticos explícitos. Sendo a gravação avisada uma situação de entrevista e havendo nela implícita hierarquia entre os participantes da conversação (de um lado suspeito, indiciado ou réu e, de outro, os peritos responsáveis pela coleta técnica de padrão vocal), seria razoável estimar que as mulheres apresentassem taxas menores do que as utilizadas pelos homens, dado o maior conservadorismo linguístico que lhes é associado (PAIVA, 2007) e o fato de que formas mais cuidadas implicam um tempo maior de implementação.

A partir do exposto, apesar de ser a amostra pouco representativa quanto ao gênero, depreende-se que a tendência relativa à diferença de comportamento entre os sexos no que se refere ao tempo de fala empregado quando da ciência da gravação provavelmente não esteja associada ao *status* social da variável em questão (tempo de fala, expresso através das medidas TE e TA) e sim às condições de uso da língua durante ações criminosas ou no planejamento dessas.

Cabe recuperar que neste estudo os sujeitos de ambos os sexos foram vinculados pela Autoridade Policial (Delegado), Autoridade Judiciária, Policial Militar encarregado de Inquérito Policial Militar ou Promotor de Justiça a alguma atividade ilícita, figurando em inquérito ou processo criminal como suspeitos, indiciados ou réus.

Não havendo, ao que se sabe, precedentes quanto à investigação dos efeitos da ciência ou não da gravação, ainda mais que tenham observado as especificidades associadas aos fatores da variável sexo, vislumbra-se que a justificativa para a disparidade encontrada envolva, no contexto da criminalística, aspectos socioculturais, entre eles, o grau de envolvimento do sujeito na ação delituosa (quão comprometido criminalmente ele está) e a importância que tal indivíduo confere ao fato de ser (ou não) um criminoso.

Inferências nesse sentido, se por um lado requerem um aporte teórico que ultrapassa a área da Sociolinguística, por outro, oportunizam reflexões exploratórias sobre questões relativas à Criminologia, especificamente à criminalidade relacionada ao gênero. Voegeli (2003) discute, a exemplo, que a criminalidade feminina manifesta-se não raro de forma dissimulada e encoberta, estando as mulheres em cooperação para realização do delito, mas afastadas dos atos de execução. Já Makki e Santos (2010), ao descreverem as características e as peculiaridades da mulher que cumpre pena, ressaltam que as mulheres cometem menos crimes que os homens (desvantagem também apontada por Voegeli, 2003) e que, ao serem presas, sofrem mais com a prisão e com a ausência dos filhos e familiares.

Gilligan (1982, p.183) aponta que a mulher tem como diferencial em relação ao gênero masculino “o cuidado e a conexão humana como valor ético prioritário”. Tal afirmação torna razoável a estimativa de que o sexo/ gênero feminino teria associado um comportamento mais autêntico, responsável, entre outros, por um menor controle das taxas temporais empregadas na fala gravada com a sua ciência.

Ao final da exposição e discussão dos resultados, flagram-se como de real contribuição, ao se atentar para a aplicação na realização da perícia de Comparação de Locutor, os achados que informam sobre o satisfatório poder discriminatório de locutor associado à TA e os que alertam quanto à relevância do *gap* temporal existente entre as gravações do sujeito. Quanto à contribuição à Sociofonética, destacam-se os resultados que indicam a tendência de comportamento gênero-específico relacionado à ciência da gravação e os indicativos de que o impacto decorrente da escolarização pode ser gerado por fator outro além da maior aproximação em relação à norma culta da língua.

Ressalvadas as limitações encontradas, segue, no próximo capítulo, a conclusão a que se chegou com o desenvolvimento do presente estudo.

CONCLUSÃO

O presente estudo destaca-se pelo ineditismo quando se propõe a analisar um *corpus* forense e a considerar a configuração habitualmente encontrada na perícia oficial brasileira (que contrapõe prevalentemente, na perícia de Comparação de Locutor, amostras de fala provindas de gravação desavisada e avisada). Visivelmente promove-se com ele a aproximação entre a pesquisa científica no campo das Ciências da Fala e a aplicação forense. Parte da amostra considerada abarca o tipo de fala mais indicado à observação sociolinguística, constante em um material autêntico (adequado à lei e à ética), gerado a partir de fala gravada sem a ciência do locutor.

O intrincado procedimento de manuseio dos dados, assim como os resultados apresentados e discutidos (respectivamente expostos nas subseções “3.3”, do Capítulo 3, e “4.1” a “4.2”, do Capítulo 4), evidenciou a complexidade conceitual e metodológica que permeia a abordagem científica do tempo de fala.

Na busca por esclarecimento acerca do potencial individualizante das taxas estudadas, foi possível constatar que, a despeito da diferença não significativa existente entre os tipos de taxa, a TA é, em detrimento da TE, a menos variável, sendo mais estável a taxa obtida por meio da média das TAs locais do que a obtida globalmente. O satisfatório poder discriminatório de indivíduo encontrado como associado à TA, diferentemente do enquadrável como fraco associado à TE, permite que se conclua pela pertinência da indicação da primeira (TA), em mensuração local média, como parâmetro temporal de fala a ser incorporado ao conjunto de elementos caracterizadores de voz e fala com vistas à definição de autoria de produções orais. Ainda, pode-se, com base nos resultados das correlações, das ANOVAs e dos Testes t realizados, desaconselhar o uso da TE na perícia de Comparação de Locutor, visto que esse tipo de taxa é significativamente afetado pela diferença existente entre os materiais confrontados (especificamente a que se refere ao fato de, na maioria das vezes, um provir de interceptação telefônica desavisada e o outro de gravação ambiental sabida e consentida) e pelo tamanho do

intervalo de fala, sendo preteridos, por essa razão, tempos de fala advindos de turnos de fala aos advindos de intervalos interpausais.

Entende-se como de suma relevância para a aplicação forense, dos resultados obtidos, especialmente os relativos à variável tipo de gravação, correspondente, neste estudo, à ciência ou não de estar sendo gravado. Dois comportamentos associados ao tipo de gravação merecem destaque: o primeiro deles é concernente ao fato de a ciência ou não da gravação ter se mostrado significativa exclusivamente na TE (o que se aventa ser justificado, conforme discussão empreendida, pelo tratamento dado à porção não fala do texto oral) e o segundo, ao fato de haver, a partir do *corpus* considerado, comportamento distinto entre os sexos (em nível de tendência) no que se refere à variável tipo de gravação.

Quanto ao tipo de gravação ser significativo exclusivamente na TE, infere-se que a situação decorra da presença e da recorrência das pausas (silenciosas e preenchidas) e/ou eventos de disfluência, elementos sabidamente diluidores dos tempos de fala que, conforme o método empregado neste estudo, foram acolhidos como fala útil na TE e como não útil na TA. A importância das pausas presentes na oralidade é confirmada pelo fato de se ter encontrado, exclusivamente na TE, correlação significativa entre a taxa e o tamanho do intervalo de fala, de onde se depreende que a taxa relativa ao tempo necessário para a execução dos gestos articulatórios dos segmentos encadeados é imune ao condicionamento gerado pelo tipo de gravação, recaindo tal condicionamento, em essência, sobre a porção não texto da fala.

Quanto aos sexos reagirem diferentemente à ciência da gravação, no grupo pesquisado, composto de sujeitos vinculados de algum modo à atividade ilícita, são os homens que, comparados às mulheres, mais controlam a TE e a TA (global e local média) quando sabem que estão sendo gravados. Esse resultado merece uma investigação sociolinguística melhor controlada, a partir de amostra representativa e fundamentos cognitivo-comportamentais e sociológicos (direcionados à questão do crime), capaz de elucidar com maior robustez se o tempo de fala é verdadeiramente um marcador sociocultural, que relaciona gênero e criminalidade (conforme proposto na Subseção 4.2.3.4 do Capítulo 4).

Ainda com relação ao sexo, é digno de nota que a primazia masculina referente ao emprego de maiores valores de taxas temporais de fala, apontada em alguns estudos sobre TE e TA baseados na fala espontânea (a exemplo Hewlett e

Rendall, 1998; Jacewicz et al., 2009, 2010; Kendal, 2009; Verhoeven et al., 2004), confirma-se como tendência na fala maximamente naturalística dos sujeitos pesquisados (gravação desavisada).

Cabe resgatar que a variável escolaridade, ao que se sabe não anteriormente investigada, evidenciou tendência contrária à estimada em hipótese, delatando que o avanço da escolarização (por razões que, infere-se, conforme discussão exposta na Subseção 4.2.3.3 do Capítulo 4, fogem à área de abrangência e ao delineamento da investigação proposta neste estudo), incita o aumento da TE e da TA, nas duas formas de mensuração (global e local média).

O *gap* temporal normalmente existente entre os áudios, variável que na prática de realização da Comparação de Locutor ainda é pouco valorada, recebe, a partir dos resultados disponibilizados neste estudo, maior destaque e uma perspectiva de uso mais criteriosa na avaliação da servibilidade dos materiais recebidos para perícia.

Considerando a aplicação pericial almejada, pode-se conduzir esta conclusão no sentido de responder, de forma objetiva, aos questionamentos que provavelmente formular-se-iam por ocasião da pretensão de uso de medidas temporais de fala na Comparação Forense de Locutor, sendo essas as que seguem.

a) Qual o tipo de taxa mais indicado para uso na Comparação Forense de Locutor?

Apesar de não ter sido encontrada diferença significativa entre os tipos de taxa estudados (TE e TA), nas formas de mensuração análogas, a TA mostrou-se o tipo de taxa mais indicado para uso na Comparação de Locutor, em razão de ser menos variável, de apresentar um melhor poder discriminante de indivíduo (CCI) quando comparado à TE e de se mostrar resistente ao condicionamento gerado (nas TEs) pelo tipo de gravação e pelo tamanho do intervalo de fala, devendo-se, no entanto, dada a correlação significativa existente entre a TA e a variável *gap* temporal, primar pela máxima contemporaneidade possível entre as amostras de fala utilizadas no cotejo.

b) No tipo de taxa indicado, qual a forma de mensuração mais efetiva para uso na Comparação Forense de Locutor?

Apesar de não ter sido encontrada diferença significativa entre as formas de mensuração global e local média na TA, em razão de ser

menos variável, é a média das tomadas locais a forma de mensuração mais indicada.

- c) Na realização da Comparação de Locutor, para o adequado emprego do tipo de taxa e da forma de mensuração indicada (TA local média), qual(is) variável(is) interveniente(s) deve(m) ser controlada(s)?

Reitera-se que o *gap* temporal existente entre as amostras do cotejo mostra-se significativamente correlacionado com a TA (global e local média). Cabe considerar, ainda, que na examinação das múltiplas tomadas locais foram identificados como significativos preditores o fator masculino da variável sexo (não problemático, uma vez que na Comparação de Locutor são contrapostas amostras produzidas por falantes supostamente do mesmo sexo) e o avanço na escolaridade (passível de ter a ocorrência verificada durante a coleta do padrão técnico), ambos responsáveis por aumento da TA, assim como a ciência da gravação, provocadora de diminuição da TA.

Recuperando-se a previsão de Künzel (1994) quanto à propensão de que cada vez mais na Comparação de Locutor devem ser admitidos parâmetros técnico-comparativos de cunho quantitativo, dada a necessidade de se ampliar o escopo objetivo dos exames empregados na realização desse tipo de perícia, conclui-se que não só foi adequada ao tema, mas oportuna, a abordagem quantitativa de análise dos dados admitida neste estudo.

A motivação que perpassou a presente investigação, do planejamento à implementação (considerando-se que eventualmente, quando se destaca como rápida ou lenta, já são utilizadas em laudos periciais no Brasil, referências informais quanto ao tempo de fala)⁷⁵, é condizente com a formulação apontada por Nolan (1983, p.1, tradução nossa) como sendo uma pergunta usual entre os cientistas da fala: “de que forma pode o trabalho de reconhecimento, em uma perspectiva objetiva, ser mais efetivo e mais confiavelmente conduzido”?

Com rigor metodológico foram analisadas a TE e a TA no *corpus* dos sete sujeitos pesquisados e definidos os potenciais individualizantes das taxas, tendo-se estabelecido, na tarefa, um procedimento adequado à realidade forense de

⁷⁵ Informação extraída do permanente debate em Fonética Forense e perícias em imagens mantido através do grupo de discussão para os Peritos Oficiais participantes da Capacitação Nacional para Peritos Criminais em Fonética Forense (CNPCFF), promovida pela Secretaria Nacional de Segurança Pública (SENASP)/ Ministério da Justiça (MJ).

Comparação de Locutor. A ausência de uma teoria própria, específica sobre tempo de fala (e não sobre o uso do tempo na fala, no emprego dos segmentos ou mesmo de traços prosódicos), que suportasse completamente a interpretação dos resultados obtidos, no entanto, não obscureceu o delineamento criterioso do estudo, tendo sido expressivo o cuidado com a emissão de conclusões.

Como sugestão para pesquisas futuras, com vistas ao aperfeiçoamento da proposta de utilização de medidas do tempo de fala na Comparação Forense de Locutor, pode-se mencionar a observação da TE a partir de intervalos menores do que o turno de fala, o estudo dirigido sobre qual seria o adequado limiar temporal distintivo entre pausas silenciosas e articulatórias no PB e, especialmente, o estudo do emprego linguístico das pausas. Pode-se propor, também, o aprofundamento quanto às modificações comportamentais relativas à ciência (ou não) de estar sendo gravado e quanto às variações estilísticas, especificamente o impacto dessas do emprego da TE e da TA. Ainda, indica-se a ampliação da verificação da existência de diferenças socioculturais motivadas pela variável sexo e sobre a correlação existente entre as taxas temporais de fala e o nível de instrução formal (escolaridade) do indivíduo.

REFERÊNCIAS BIBLIOGRÁFICAS

ABAURRE-GNERRE, M.B. Processos fonológicos segmentais como índices de padrões prosódicos diversos nos estilos formal e casual do português do Brasil. **Caderno de Estudos Lingüísticos**, v. 2, p. 23-44, 1981.

ANDRADE, C.R.F.; CERVONE, L.M.; SASSI, F.C. Relationship Between the Stuttering Severity Index and Speech Rate. **Revista Paulista de Medicina**, v. 121, n. 2, p. 81-84, 2003.

BARBOSA, P. A. **Incursões em torno do ritmo da fala**. Campinas: Pontes Editores, 2006.

BEHLAU, M.S, PONTES, P. O desenvolvimento ontogenético da voz: do nascimento à senescência. In: BEHLAU, M.S, PONTES, P. **Avaliação e tratamento das disfonias**. São Paulo: Lovise, 1995.

BELL, A. Language style as audience design. **Language in Society**, v. 13, n. 2, p. 145-204, 1984.

BELL, A. Back in style: reworking audience design. In: ECKERT, P.; RICKFORD, J.R. (Ed.). **Style and sociolinguistic variation**. Cambridge: Cambridge University Press, 2001.

BENTO, R.F.; MINITI, A.; MARONE, S.A.M. **Tratado de otologia**. São Paulo: Editora da USP, 1998.

BINNENPOORTE, D.; VAN BAEL, C.; DEN OS, E.; BOVES, L. Gender in Everyday Speech and Language: A Corpusbased Study. In: 9th Interspeech, 2005, Lisboa. **Proceedings**. Lisboa: 2005. 1-4.

BLOCK, S.; KILLEN, D. Speech Rates of Australian English-Speaking Children and Adults. **Australian Journal of Human Communication Disorders**, v. 24, n. 1, p. 39-44, 1996.

BONASTRE, J.F.; BIMBOT, F.; BOE, L.J.; CAMPBELL, J.P.; REYNOLDS, D.A.; MAGRIN-CHAGNOLLEAU, I. Person Authentication by Voice: A Need for Caution. In: 8th Eurospeech, 2003, Geneva. **Proceedings**. Geneva: 2003. 33-36.

BORSEL, J.V.; MAESSCHALCK, D. Speech Rate in Males, Females, and Male-to-Female Transsexuals. **Clinical Linguistics & Phonetics**, v. 22, n. 9, p. 679–685, 2008.

BRAID, A.C.M. **Fonética Forense**. 2.ed. Campinas: Millennium, 2003.

BRAUN, A. Fundamental Frequency: how speaker-specific is it? In: BRAUN, A.; KÖSTER, J.P. (Ed.) **Studies in Forensic Phonetics**. Trier: Wissenschaftlicher Verlag Trier, 1995.

BUTCHER, A. Aspects of Speech Pause: Phonetic Correlates and Communicative Functions. **Arbeitsberichte des Instituts für Phonetik der Universität Kiel**, v. 15, p. 1-233, 1981.

BYRD, D. Relations of Sex and Dialect to Reduction. **Speech Communication**, v. 15, p. 39-54, 1994.

BYRNE, C.; FOULKES, P. The 'Mobile Phone Effect' on Vowel Formants. **Speech, Language and the Law**, v. 11, n. 1, p. 83-102, 2004.

CAGLIARI, L.C. **Alfabetização & Lingüística**. São Paulo, Scipione, 1997.

CALDEIRA, T.P. **A cidade dos muros**. São Paulo: Edusp, 2000.

CAMACHO, R.G. Uma reflexão crítica sobre a Teoria Sociolinguística. **D.E.L.T.A.**, v. 26, n. 1, p. 141-162, 2010.

CAO, H.; WANG, Y. A forensic aspect of articulation rate variation in Chinese. In: XVII ICPHS, 2011, Hong Kong. **Proceedings**. Hong Kong: 2011. 396-399.

CLARK, J.; YALLOP, C.; FLETCHER, J. Prosody. In: CLARK, J.; YALLOP, C.; FLETCHER, J. **An Introduction to Phonetics and Phonology**. 3.ed. Oxford: Blackwell, 2007.

COSTA, C.G.B.C. **Influências prosódicas nos encontros vocálicos em fronteiras de palavras**. Dissertação (Mestrado em Linguística), UFMG, Belo Horizonte, 2008.

COUPLAND, N. **Language, situation, and the relational self: Theorising dialect-style in sociolinguistics**. Paper presented at Stanford Workshop on Stylistic Variation, Stanford, 1996.

CRYSTAL, D. **A Dictionary of Linguistics and Phonetics**. 2.ed. New York: Basil Blackwell, 1985.

CRYSTAL, T.H.; HOUSE, A.S. Articulation Rate and the Duration of Syllables and Stress Groups in Connected Speech. **Journal of the Acoustical Society of America**, v. 88, n. 1, p. 101-112, 1990.

CUTLER, A.; DAHAN, D.; VON DONSELAAR, W. Prosody in the Comprehension of Spoken Language: a Literature Review. **Language and Speech**, v. 40, n. 2, p. 141-202, 1997.

DA HORA, L.M.P. **Nem repressão nem educação: uma droga de cenário**. Tese (Doutorado em Educação), UFRJ, Rio de Janeiro, 2008.

DANKOVIČOVÁ, J. The Domain of Articulation Rate in Czech. **Journal of Phonetics**, v. 25, p. 287-312, 1997.

DANKOVIČOVÁ, J. The Linguistic Basis of Articulation Rate Variation in Czech. **Forum Phoneticum**, v. 71, 2001.

DI PAOLO, M.; YAEGER-DROR, M. (Ed.) **Sociophonetics: A Student's Guide**. London: Routledge, 2011.

DOREA, L.E.C.; STUMVOLL, V.P.; QUINTELA, V. **Criminalística**. 3.ed. Campinas: Millennium Editora, 2005.

DUEZ, D. Silent and Non-Silent Pauses in Three Speech Styles. **Language and Speech**, v. 25, Part 1, p. 11-28, 1982.

ECKERT, P. **Linguistic Variation as Social Practice**. Malden: Blackwell, 2000.

EEFTING, W. Temporal Variation in Natural Speech: Some Explorations. In: 7th FASE Symposium, 1988, Edinburgh. **Proceedings**. Edinburgh: 1988. 503-507.

ERIKSSON, A. Aural/ Acoustical vs. Automatic Methods in Forensic Phonetic case Work. In: NEUSTEIN, A.; PATIL, H.A. **Forensic Speaker Recognition: Law Enforcement and Counter-terrorism**. New York: Springer-Werlag, 2012.

ESCUADERO, P.; BOERSMA, P.; RAUBER, A.S.; BION, R.A.H. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. **Journal of Acoustical Society of America**, v. 126, n. 3, 2009.

FANT, G. **Acoustic Theory of Speech Production**. Mouton, The Hague, 1960.

FIRTH, J.R. **Papers in Linguistics**. Oxford: Oxford University Press, 1951.

FLETCHER, J. The Prosody of Speech: Timing and Rhythm. In: HARDCASTLE, W.; LAVER, J.; GIBBON, F.E. (Ed.). **The handbook of Phonetic Sciences**. 2.ed. Oxford: Wiley-Blackwell, 2010.

FONAGY, I.; MAGDICS, K. Speed of Utterance in Phrases of Different Lengths. **Language & Speech**, v. 3, p.179-192, 1960.

FOULKES, P.; DOCHERTY, G. The Social Life of Phonetics and Phonology. **Journal of Phonetics**, v. 34, p. 409-438, 2006.

FOULKES, P.; SCOBIE, J.M.; WATT, D. Sociophonetics. In: HARDCASTLE, W.; LAVER, J.; GIBBON, F.E. (Ed.). **The handbook of Phonetic Sciences**. 2.ed., Oxford: Wiley-Blackwell, 2010.

FRENCH, P.; HARRISON, P. Position Statement concerning use of impressionistic likelihood terms in forensic speaker comparison cases. **The International Journal of Speech, Language and the Law**, v. 14, n. 1, p. 137–144, 2007.

FRENCH, P.; NOLAN, F.; FOULKES, P.; HARRISON, P.; McDOUGALL, K. The UK position statement on forensic speaker comparison: a rejoinder to Rose and Morrison. **The International Journal of Speech, Language and the Law**, v. 17, n. 1, p. 143–152, 2010.

FUJIMURA, O.; ERICKSON, D. Acoustic Phonetics. In: HARDCASTLE, W.J.; LAVER, J. (Ed.). **The Handbook of Phonetic Sciences**. [S.l.]: Blackwell Publishers, 1997.

FUJISAKI, H. Prosody, models, and spontaneous speech. In: SAGISAKA, Y.; CAMPBELL, N.; Higuchi, N. (Ed.) **Computing Prosody**. New York: Springer, 1997.

GATHERCOLE, S.E.; BADDELEY, A.D. **Working Memory and Language**. Hove: Erlbaum, 1993.

GILLIGAN, C. **Uma voz diferente**. Rio de Janeiro: Rosa dos Tempos, 1982.

GODFREY, J.J.; HOLLIMAN, E.C.; MCDANIEL, J. Switchboard: Telephone Speech Corpus for Research and Development. In: **IEEE Transactions on Acoustics, Speech, and Signal Processing**, 1992, San Francisco. **Proceedings**. San Francisco: 1992. 517-520.

GOLD, E.; FRENCH, P. International practices in forensic speaker comparison. **International Journal of Speech, Language and the Law**, v. 18, p. 293-307, 2011.

GOLDMAN-EISLER, F. The Determinants of the Rate of Speech Output and Their Mutual Relations. **Journal of Psychosomatic Research**, v. 1, p. 137-143, 1956.

GOLDMAN-EISLER, F. Speech Production and the Predictability of Words in Context. **Quarterly Journal of Experimental Psychology**, v. 10, p. 96-106, 1958.

GOLDMAN-EISLER, F. **Psycholinguistics: Experiments in Spontaneous Speech**. New York: Academic, 1968.

GÖRSKI, E. M. **A variação estilística na ótica da sociolinguística laboviana: (re)dimensionando o papel do contexto**. Disponível em : <[http://www.cchla.ufrn.br/visiget/pgs/pt/anais/Artigos/Edair%20Maria%20G%C3%B6rski%20\(UFSC\).pdf](http://www.cchla.ufrn.br/visiget/pgs/pt/anais/Artigos/Edair%20Maria%20G%C3%B6rski%20(UFSC).pdf)>.

Acesso em: 12 nov. 2012.

GROSJEAN, F.; DESCHAMPS, A. Analyse Contrastive des Variables Temporelles de L'anglais et du Français: Vitesse de Parole et Variables Composantes, Phénomènes d'Hésitation. **Phonetica**, v. 31, p. 144-184, 1975.

GRUNWALD, T. **Reduktion und Kompensation als Funktion der Sprechgeschwindigkeit im Deutschen**. Hamburg: Buske, 1983. (Forum Phonicum, v. 28)

HANSSON, P. Articulation Rate Variation in South Swedish Phrases. In: Speech Prosody, 2002, Aix-en-Provence. **Proceedings**. Aix-en-Provence, 2002. 371-374.

HEWLETT, N.; RENDALL, M. Rural Versus Urban Accent as an Influence on the Rate of Speech. **Journal of International Phonetic Association**, v. 28, p. 63-71, 1998.

HIEKE, A.E.; KOWAL, S.; O'CONNEL, D.C. The Trouble with "Articulatory" Pauses. **Language and Speech**, v. 26, Part 3, 1983.

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

HOSMER, D.W.; LEMESHOW, S. **Applied Logistic Regression**. New York: John Wiley & Sons, 2000.

JACEWICZ, E.; FOX, R.A.; O'NEILL, C. Articulation Rate Across Dialect, Age, and Gender. **Language Variation and Change**, v. 21, p. 233-256, 2009.

JACEWICZ, E.; FOX, R.A.; WEI, L. Between-Speaker and Within-Speaker Variation in Speech Tempo of American English. **Journal of the Acoustical Society of America**, v. 128, n. 2, p. 839-850, 2010.

JESSEN, M. Forensic Reference Data on Articulation Rate in German. **Science & Justice**, v. 47, p. 50-67, 2007.

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

JONG, N.H.; WEMPE, T. Praat script to detect syllable nuclei and measure speech rate automatically. **Behavior Research Methods**, v. 41, n. 2, p. 385-390, 2009.

KAKITA, K. The Effect of Intrasentence Pause on the Articulation Rate of Prepausal and Postpausal Utterances. **Journal of the Acoustical Society of America**, v. 104, n. 3, p. 1818-1818, 1998.

KENDALL, T. **Speech Rate, Pause, and Linguistic Variation: An Examination Through the Sociolinguistic Archive and Analysis Project**. Doctoral Dissertation, Duke University, Durham, 2009.

KENT, R.D.; READ, C. **The acoustic Analysis of Speech**. San Diego: Singular Publishing Group, 1992.

KOISO, H.; SHIMOJIMA, A.; KATAGIRI, Y. Collaborative signaling of informational structures by dynamics speech rate. **Language and Speech**, v. 41, p. 323-350, 1998.

KOREMAN, J. Perceived Speech Rate: The Effects of Articulation Rate and Speaking Style in Spontaneous Speech. **Journal of the Acoustical Society of America**, v. 119, n. 1, p. 582-596, 2005.

KOWAL, S.; O'CONNELL, D.C.; SABIN, E.F. Development of Temporal Patterning and Vocal Hesitations in Spontaneous Narratives. **Journal of Psycholinguistic Research**, v. 4, p. 195-207, 1975.

KOWAL, S.; WIESE, R.; O'CONNELL, D. C. The Use of Time in Storytelling. **Language and Speech**, v. 26, p. 377-392, 1983.

KRAUSE, J.; BRAIDA, L. Investigating alternatives forms of clear speech: The effects of speaking rate and speaking mode of intelligibility. **Journal of the Acoustical Society of America**, v. 112, p. 2165-2172, 2002.

KÜNZEL, H.J. Current Approaches to Forensic Speaker Recognition. In: **ESCA**, 1994, Martigny. **Proceedings**. Martigny: 1994. 135-141.

KÜNZEL, H.J. Some General Phonetic and Forensic Aspects of Speaking Tempo. **Forensic Linguistics**, v. 4, n. 1, p. 48-83, 1997.

KUWABARA, H.; SAGISAKA, Y. Acoustic characteristics of speaker individuality: Control and conversion. **Speech Communication**, n. 16, p. 165-173, 1995.

LABOV, W. **The Social Stratification of English in New York City**. Washington DC: Center for Applied Linguistics, 1966.

LABOV, W. The Study of Language in Its Social Context. **Studium Generale**, v. 23, p. 30-87, 1970.

LABOV, W. **Sociolinguistic Patterns**. Philadelphia: University of Pennsylvania Press, 1972.

LABOV, W. Building on Empirical Foundations. In: LEHMANN, W.; MALKIEL, Y. (Ed.) **Perspectives on Historical Linguistics**. Amsterdam: John Benjamins, 1982.

LABOV, W. **Principles of Linguistic Change**. Oxford: Blackwell, 1994.

LABOV, W. The anatomy of style-shifting. In: ECKERT, P.; RICKFORD, J.R. (Ed.). **Style and sociolinguistic variation**. Cambridge: Cambridge University Press, 2001.

LABOV, W. **Padrões sociolinguísticos**. Tradução de: Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008.

LADEFOGED, P. **A Course in Phonetics**. 4.ed. Orlando: Harcourt Brace, 2001.

LADEFOGED, P. **Phonetic Data Analysis**. Oxford: Blackwell, 2003.

LADEFOGED, P.; MADDIESON, I. **The Sounds of the World's language**. Oxford: Blackwell, 1996.

LAUREANO, G.H.C. **Coeficiente de correlação intraclasse: comparação entre métodos de estimação clássico e bayesianos**. Monografia (Bacharelado em Estatística), UFRGS, 2011.

LAVAR, J. **The Phonetic Description of Voice Quality**. Cambridge: Cambridge University Press, 1980.

LAVAR, J. **Principles of Phonetics**. Cambridge: Cambridge University Press, 1994.

LIN L. A concordance correlation coefficient to evaluate reproducibility. **Biometrics**, v. 45, n. 1, p. 255-268, 1989.

LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge companion to sociolinguistics**. London: Routledge, 2007.

LLISTERRI, J. Speaking Styles in Speech Research. In: **ELSNET/ ESCA/ SALT**, Workshop on Integrating Speech and Natural Language, 1992, Dublin. **Proceedings**. Dublin: 1992. 1-28.

MAKKI, S.H.; SANTOS, M.L. Gênero e criminalidade: Um olhar sobre a mulher encarcerada no Brasil. In: **Âmbito Jurídico**, Rio Grande, v.78, julho, 2010. Disponível em: <http://www.ambito-juridico.com.br/site/index.php?n_link=revista_artigos_leitura&artigo_id=8080&revista_caderno=3>. Acesso em: 12 nov. 2012.

MALÉCOT, A.; JOHNSTON R.; KIZZIAR, P.A. Syllabic Rate and Utterance Length in French. **Phonetica**, v. 26, p. 235-251, 1972.

MARCUSCHI, L.A. **Análise da conversação**. 5.ed. São Paulo: Ática, 2001.

MARTINS, M.R.D. **Ouvir falar: Introdução à fonética do português**. Lisboa: Caminho, 1988.

McDOUGALL, K. **The Role of Formant Dynamics in Determining Speaker Identity**. Doctoral Dissertation, University of Cambridge, Cambridge, 2005.

MCMENAMIN, G.R. **Forensic Linguistics: Advances in Forensic Stylistics**. New York: CRC Press, 2002.

MACQUEEN, J.M.; CUTLER, A. Cognitive Processes in Speech Perception. In: HARDCASTLE, W.J.; LAVER, J. (Ed.). **The Handbook of Phonetic Sciences**. [S.I.]: Blackwell Pub, 1997.

MEIRELES, A.R. **Processos fonéticos-fonológicos decorrentes do aumento da velocidade de fala no português brasileiro**. Dissertação (Mestrado em Linguística), UFMG, Belo Horizonte, 2001.

MEIRELES, A.R. **Reestruturações rítmicas da fala no Português Brasileiro**. Tese (Doutorado em Linguística), UNICAMP, Campinas, 2007.

MEIRELES, A.; BARBOSA, P.A. Lexical reorganization in Brazilian Portuguese: an articulatory study". **Speech Communication**, v. 50, p. 916-924, 2008.

MEIRELES, A. R.; BARBOSA, P.A. O papel da taxa de elocução nos processos dinâmicos de mudança lingüística. **Revista (con) textos linguísticos (UFES)**, v. 3, p. 91-116, 2009.

MENESES, F.O. **As vogais desvozeadas no português brasileiro: investigação acústico-articulatória**. Dissertação (Mestrado em Linguística), UNICAMP, Campinas, 2012.

MILLER, J.L.; GROSJEAN, F.; LOMANTO, C. Articulation Rate and Its Variability in Spontaneous Speech: A Reanalysis and Some Implications. **Phonetica**, v. 41, p. 215-225, 1984.

MILROY, L. **Language and Social Networks**. 2.ed. Oxford: Blackwell, 1987.

MIRA-MATEUS, M.H. Estudando a melodia da fala: traços prosódicos e constituintes prosódicos. In: Encontro sobre o ensino das línguas e a linguística, 2004, Setúbal. **Anais**. Setúbal: 2004. 27-28.

MIRGHAFORI, N., FOSLER, E., MORGAN, N. Towards Robustness to Fast Speech. In: ICASSP, 1996, San Francisco. **Proceedings**. San Francisco: 1996, 335-338.

MIXDORFF, H.; PFITZINGER, H.R. Analysing Fundamental Frequency Contours and Local Speech Rate in Map Task Dialogs. **Speech Communication**, v. 46, p. 310-325, 2005.

MOLINA, G.P.; GOMES, L.F. **Criminologia**. São Paulo: Revista dos Tribunais, 1997.

MORISSON, A.L.C.; SAMPAIO, J.F.; RIBEIRO, J.F. Exames de registro de áudio e imagens: recomendações técnicas para a padronização de procedimentos e metodologias. In: TOCHETTO, D. e ESPINDULA, A. (Cord.). **Criminalística: Procedimentos e Metodologias**. 2.ed. Porto Alegre: [s.n.], 2009.

MORRISON, G.S. Forensic Voice Comparison and the Paradigm Shift. **Science and Justice**, v. 49, p. 298-308, 2009a.

MORRISON, G.S. Likelihood-ratio forensic voice comparison using parametric representations of the formant trajectories of diphthongs. **Journal of the Acoustical Society of America**, v. 125, n. 4, 2009b.

MULLANY, L. Speech Communities. In: LLAMAS, C.; MULLANY, L.; STOCKWELL, P. **The Routledge companion to sociolinguistics**. London: Routledge, 2007.

NÉSPOR, M.; VOGEL, I. **Prosodic Phonology**. Dordrecht: Foris Publications, 1986.

NOLAN, F. **The phonetic bases of speaker recognition**. Cambridge: Cambridge University Press, 1983.

NOLAN, F. Speaker Recognition and Forensic Phonetics. In: HARDCASTLE, W.J.; LAVER, J. (Ed.). **The Handbook of Phonetic Sciences**. [S.I.]: Blackwell Publishers, 1997.

NOLAN, F. Speaker Identification Evidence: Its Forms, Limitations, and Roles. In: The conference "Law and language: Prospect and retrospect", 2001, Levi. **Proceedings**. Levi: 2001. 12-15.

NOOTEBOOM, S. The Prosody of Speech: Melody and Rhythm. In: **The Handbook of Phonetic Sciences**. [S.I.]: Blackwell Publishers, 1997.

OLIVEIRA, M.K. Organização conceitual e escolarização. In. OLIVEIRA, M.B; OLIVEIRA, M.K. **Investigações Cognitivas**. Porto Alegre: ARTMED, 1999.

OOSTDIJK, N. The Design of the Spoken Dutch Corpus. In: PETERS, P.; COLLINS, P.; SMITH, A. (Ed.) **New Frontiers of Corpus Research**. Amsterdam: Rodopi, 2002.

PADGETT, J. Glides, vowels, and features. **Lingua**, v. 118, p. 1937-1955, 2008.

PAIVA, M.C. A variável gênero/sexo. In: MOLLICA, M.C. Fundamentação teórica: conceituação e delimitação. In: MOLLICA, M.C.; BRAGA, M.L. (Org.) **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2007.

PESTANA, M.H., GAGEIRO, J.N. **Análise de dados para ciências sociais: a complementaridade do SPSS**. 3.ed. Lisboa: Edições Sílabo, 2003.

PFITZINGER, H.R. Local Speech Rate as a Combination of Syllable and Phone Rate. In: 15th ICSLP, 1998, Sydney. **Proceeding**. Sydney: 1998, 3177-3180.

QUENÉ, H. Multilevel Modelling of Between-Speaker and Within-Speaker Variation in Spontaneous Speech Tempo. **International Journal of the Acoustical Society of America**, v. 123, n. 2, p. 1104–1113, 2008.

RAMIG, L. Effects of Physiological Aging on Speaking and Reading Rates. **Journal of Communication Disorders**, v. 16, p. 217–226, 1983.

RAMSARAN, S.M. **Phonetic and Phonological Correlates of Style in English: A preliminary Investigation**. Doctoral Dissertation, University of London, 1978.

RAMUS, F. Acoustic Correlates of Linguistic Rhythm: Perspectives. In: **Speech Prosody**, 2002, Aix-en-Provence. **Proceedings**. Aix-en-Provence: 2002. 11-13.

RISCHEL, J. Formal Linguistics and Real Speech. **Speech Communication**, v. 11, n. 4-5, p. 379-392, 1992.

ROBB, M.P., MACLAGAN, M.A., CHEN, Y. Speaking Rates of American and New Zealand Varieties of English. **Clinical Linguistics & Phonetics**, v. 18, n. 1, p. 1–15, 2004.

RODMAN, R.; McALLISTER, D.; BITZER, D.; CEPEDA, L.; ABBITT, P. Forensic speaker identification based on spectral moments. **Forensic Linguistics**, v. 9, n. 1, p. 22-43, 2002.

ROMAINE, S. What is a speech community? In: ROMAINE, S. **Sociolinguistic Variation in Speech Communities**. London: Edward Arnold, 1980.

ROMAINE, S. **Language in society: An introduction to sociolinguistics**. London: Blackwell, 1994.

ROMITO, L.; GALATÀ, V. Towards a protocol in speaker recognition analysis. **Forensic Science International**, v. 146S, p. S107-S111, 2004.

ROSE, P. **Forensic Speaker Identification**. London: Taylor & Francis, 2002.

ROSE, P. Technical Forensic Speaker Recognition: Evaluation, Types and Testing of Evidence. **Computer Speech & Language**, v. 20, n. 2-3, p. 159-191, 2006.

SAVILLE-TROIKE, M. **The Ethnography of Communication**. 3.ed. Oxford: Blackwell, 2003.

SCHILLING-ESTES, N. Investigating Stylistic Variation In: CHAMBERS, J.K.; TRUDGILL, P.; SCHILLING-ESTES, N. (Ed.) **Handbook of Language Variation and Change**. Oxford: Blackwell Publishing Ltd., 2002.

SCHILLING-ESTES, N. Constructing ethnicity in interaction. **Journal of Sociolinguistics**, v. 8, n. 2, p. 163-195, 2004.

SELKIRK, E.O. The syllable. In: HULST, F.V.; SMITH, N. (Ed.) **The structure of phonological representations**. Dordrecht: Foris, 1982.

SERAPIONI, M. Métodos qualitativos e quantitativos na pesquisa social em saúde: algumas estratégias para integração. **Ciência & Saúde Coletiva**, v. 5, n. 1, p. 187-192, 2000.

SILVA, J.A. **Estudo sociofonético de variações rítmicas no dialeto capixaba**. Dissertação (Mestrado em Linguística/ Estudos Linguísticos), UFES, 2010.

SIPTÁR, P. Fast Speech Processes in Hungarian. In: GOSY, M. (Ed.) **Temporal Factors in Speech: A Collection of Papers**. Budapest: Magyar Tudományos, 1991.

SWANN, J.; DEUMERT, A.; LILLIS, T.; MESTHRIE, R. **A dictionary of Sociolinguistics**. Edinburgh: Edinburgh University Press, 2004.

SYRDAL, A.K. Acoustic Variability in Spontaneous Conversational Speech of American English Talkers. In: ICSLP, 1996, Philadelphia. **Proceedings**. Philadelphia: 1996, 438-441.

TARALLO, F. **A pesquisa socio-lingüística**. 2.ed. São Paulo: Editora Ática, 1986.

TAUROZA, S.; ALLISON, D. Speech Rates in British English. **Applied Linguistics**, v. 11, n. 1, p. 90-105, 1990.

THOMAS, E. **Sociophonetics: an introduction**. New York: Palgrave Macmillian, 2011.

TILLMANN, H.G.; PFITZINGER, H.R. Local Speech Rate: Relationship Between Articulation and Speech Acoustics. In: 15th ICPHS, 2003, Barcelona. **Proceedings**. 2003, 3177-3180.

TROUVAIN, J. **Tempo Variation in Speech Production: Implications for Speech Synthesis**. Saarbrücken: Institute of Phonetics, Saarland University, 2004. (PHONUS 8)

TROUVAIN, J.; KOREMAN, J.; ERRIQUEZ, A.; BRAUN, B. Articulation Rate Measures and Their Relation to Phone Classification in Spontaneous and Read German Speech. In: ISCA, 2001, Sophia-Antipolis. **Proceedings**. Sophia-Antipolis: 2001, 155-158.

TRUDGILL, P. **A Glossary of Sociolinguistics**. Edinburgh: University Press, 2003.

TSAO, Y.C., WEISMER, G. Interspeaker Variation in Habitual Speaking Rate: Evidence for a Neuromuscular Component. **Journal of Speech, Language, and Hearing Research**, v. 40, p. 858–866, 1997.

TSAO, Y.C., WEISMER, G.; IQBAL, K. Interspeaker Variation in Habitual Speaking Rate: Additional Evidence. **Journal of Speech, Language, and Hearing Research**, v. 49, p. 1156–1164, 2006.

VAN DONZEL, M.; BEINUM, F. Pausing Strategies in Discourse in Dutch. In: 4th IEEE, 1996, Indiana. **Proceedings**. Indiana: 1996.

VERHASSELT, J.P.; MARTENS, J.P. A Fast and Reliable Rate of Speech Detector. In: Spoken Language, 1996, Philadelphia. **Proceedings**. Philadelphia: 1996. 2258-2261.

VERHOEVEN, J., DE PAUW, G., KLOOTS, H. Speech Rate in a Pluricentric Language: A Comparison Between Dutch in Belgium and the Netherlands. **Language and Speech**, v. 47, n. 3, p. 297–308, 2004.

VOEGELI, C.M.P.H. **Criminalidade & violência no mundo feminino**. Curitiba: Juruá, 2003.

WATT, D. The Identification of the Individual Through Speech. In: LLAMAS, C.; WATT, D. **Language and Identities**. Edinburgh: Edinburgh University Press, 2010.

WEBER, A. F. Difusão escrita de conhecimento técnico-científico para um público pouco escolarizado. **Linguagens & Cidadania**, v. 12, n. 5, 2005.

YUAN, J., LIBERMAN, M.; CIERI, C. Towards an Integrated Understanding of Speaking Rate in Conversation. In: International Conference on Spoken Language Processing, 2006, Pittsburgh. **Proceeding**. Pittsburgh: 2006.

ZELLNER, B. Fast and Slow Speech Rate: a Characterisation for French. In: 5th ICSLP, 1998, Sydney. **Proceeding**. Sydney: 1998, 3159-3163.

GLOSSÁRIO

Amostra questionada: Registro de áudio no qual consta a fala do locutor cuja identidade se deseja conhecer.

Amostra padrão: Registro de áudio no qual consta a fala do suspeito, indiciado ou réu (de identidade conhecida) que se presume ter produzido as falas questionadas.

Análise de variância (ANOVA): Teste que determina se a diferença existente entre duas médias amostrais deve-se a variações aleatórias ou se os dados vêm de populações com comportamentos de fato distintos.

Assimetria: Medida indicativa do formato da distribuição dos dados em relação à distribuição normal, refere o grau de desvio (afastamento da simetria) de uma distribuição.

Audiência: Interlocutor, aquele com quem se trava conversação.

Coefficiente de correlação intraclassa (CCI): Também referido como Coeficiente de Reprodutibilidade. É uma estimativa da função de variabilidade total das medidas devido às variações que acontecem entre os indivíduos.

Coefficientes de correlação múltipla (R múltiplo): Coeficiente que mede a força ou o grau de relacionamento existente entre uma variável dependente e um conjunto de outras variáveis.

Coefficientes de determinação R^2 ajustado: Coeficiente que mede a percentagem da explicação da variação da variável dependente pela variação das variáveis independentes.

Comparação de Locutor: Tipo de perícia no qual objetiva-se definir a autoria de determinada produção oral. Estabelece o confronto entre duas amostras de fala, visando definir se essas foram produzidas pelo mesmo indivíduo.

Constante: Termo independente do modelo, não sendo associado a variáveis.

Correlação: Procedimento estatístico utilizado para avaliar a força da associação entre duas variáveis.

Criminalística: Disciplina autônoma, integrada pelos diferentes ramos do conhecimento científico, auxiliar e informativa das atividades policiais e

judiciárias de investigação criminal, tendo por objeto o estudo dos vestígios materiais extrínsecos à pessoa física, no que tiver de útil à elucidação e a prova das infrações penais e, ainda, à identificação dos autores respectivos. (Conceito atribuído, conforme Dorea et al., 2005, p.2, ao Perito Criminalístico do RS Eraldo Rabello)

Criminologia: “É a ciência empírica e interdisciplinar, que tem por objeto o crime, o delinquente, a vítima e o controle social do comportamento delitivo; e que aporta uma informação válida, contrastada e confiável, sobre a gênese, a dinâmica e variáveis do crime – contemplado este como fenômeno individual e como problema social –; assim como a sua prevenção eficaz, as formas e estratégias de reação ao mesmo e as técnicas de intervenção positiva no infrator.” (MOLINA e GOMES, 1997)

Curtose: Medida indicativa do formato da distribuição dos dados em relação à distribuição normal, refere o grau de achatamento de uma distribuição.

Degravação: Transposição do texto oral para o meio gráfico comumente utilizada na rotina forense em registros de áudio.

Encurtamento antecipatório: Tendência que as sentenças com mais sílabas tem de serem produzidas com sílabas mais curtas.

Erro padrão: Desvio padrão da distribuição de probabilidade da estatística.

Estatística t: Tipo de estatística de teste de hipótese.

Fala disfluente: Fala na qual se observa ocorrência(s) de repetição(ões), alongamento(s) ou falso(s) começo(s).

FFT (*Fast Fourier Transform*): Algoritmo para determinação de espectro (com a intensidade no eixo das ordenadas e a frequência no eixo das abcissas) situacional da fonte glótica. Caracteriza-se por representar graficamente a frequência fundamental e seus harmônicos, o que difere do sinal de saída ou *output* vocal, que pressupõe o processo de filtragem.

Gap temporal: Intervalo de tempo transcorrido entre as gravações confrontadas.

Gráficos Q-Q: Gráfico quantil - quantil, utilizado para informar se dois conjuntos de dados pertencem à mesma distribuição de probabilidades.

Gráficos P-P: Gráfico probabilidade - probabilidade, utilizado para informar a respeito da normalidade de um conjunto de dados.

Graus de liberdade (GL): Número de determinações independentes menos o número de parâmetros estatísticos a serem avaliados na população.

Intervalo interpausal: Trecho de fala compreendido entre o silêncio do início de emissão e uma pausa silenciosa (com duração igual ou maior à 130ms) ou entre duas pausas silenciosas.

Locutor-alvo: Falante cuja produção oral está sendo analisada.

LPC (*Linear Predictive Coding*): Procedimento que permite a determinação dos parâmetros de um filtro digital cuja resposta em frequência se aproxima do espectro de um sinal sob análise.

LTAS (*Long-term Average Spectrum*): Espectro (com a intensidade no eixo das ordenadas e a frequência no eixo das abcissas) que reflete características tanto da fonte quanto do filtro. Considera um trecho suficientemente longo de sinal sonoro, para que ocorra a neutralização das propriedades segmentais.

Mensuração global da TE/ TA: Divisão do número de unidades linguísticas (sílabas, segmentos, etc) constantes em um intervalo de fala integral pela respectiva duração.

Mensuração local da TE/ TA: Divisão do número de unidades linguísticas (sílabas, segmentos, etc) constantes em um intervalo de fala menor que o texto oral integral pela sua respectiva duração.

Mensuração local média da TE/ TA: Média gerada a partir de um grupo de taxas obtidas localmente.

Nível de significância (α): Limite estabelecido como referência para se afirmar que um desvio decorre do acaso ou não.

Perito Criminal (Oficial): Servidor público de nível superior, admitido mediante concurso público, que atua em órgãos periciais estaduais, distrital ou na Polícia Federal, encarregado de fazer a prova técnica, através da análise científica de vestígios deixados durante a prática do(s) delito(s).

P-valor: Nível descritivo do teste. Probabilidade de se obter uma diferença entre a estimativa do valor do parâmetro e o valor suposto igual ou maior do que a observada.

Raiz Média quadrática: Raiz quadrada da média aritmética dos quadrados dos n termos.

Resíduos (Análise de): Diferença entre o valor observado e o valor predito pelo modelo. A análise dos resíduos verifica se o modelo utilizado é adequado aos dados.

Significante/ Significativo (Estatisticamente): Situação na qual o p-valor é menor que o nível de significância adotado.

Soma de quadrados: Numerador do cálculo da variância, sendo utilizado em modelos de análise de variância para obter a variabilidade de cada fator.

Taxa de articulação: Número de unidades linguísticas (sílabas, segmentos, etc) constantes em um intervalo de fala pausa-excludente dividido pela respectiva duração.

Taxa de elocução: Número de unidades linguísticas (sílabas, segmentos, etc) constantes em um intervalo de fala pausa-includente dividido pela respectiva duração.

Teste bicaudal: Teste de significância estatística no qual o desvio da hipótese nula é considerado possível em qualquer direção.

Teste de Levene: Teste usado para verificar se várias amostras têm ou não a mesma variância.

Teste de Shapiro-Wilk: Teste de normalidade baseado nos valores amostrais ordenados elevados ao quadrado.

Teste F para razão de variâncias: Teste para comparação de variâncias de dois grupos.

Teste robusto de igualdade de médias de Brown-Forsythe: Teste similar ao Teste F, sendo utilizado, porém, quando a suposição de homogeneidade de variâncias entre os grupos é violada.

Teste t de amostras independentes: Teste de hipótese utilizado para a comparação de médias entre grupos distintos e independentes.

Teste t de amostras pareadas: Teste de hipótese utilizado para verificação da diferença entre os valores observados em um mesmo sujeito em dois instantes distintos do tempo ou entre sujeitos pareados (em duas amostras distintas).

Turno de fala: “Tudo aquilo que um falante faz ou diz quando tem a palavra, incluindo o silêncio.” (MARCUSCHI, 2001)

Variabilidade intersujeito: Variabilidade determinada entre as médias dos sujeitos.

Variabilidade intrassujeito: Variabilidade não controlada e não explicada pelas fontes de variação, ou seja, a variabilidade que é inerente ao sujeito.

Variância: Soma dos desvios quadráticos de uma série em relação à média.

APÊNDICE A – Panorama de estudos sobre TE e TA com cálculo em sílabas por segundo, a partir de fala espontânea não infantil de indivíduo em condições normais de linguagem (continua)

Referência	Dados do estudo
DANCOVIČOVÁ (1997)	<p>Taxa(s): TA</p> <p>Tipo de material de fala: cerca de 2min de narrativa sobre figura de livro infantil</p> <p>Língua: tcheco</p> <p>Sujeitos: 3F, 20 a 25 anos, estudantes</p> <p>Tipo de intervalo de fala: frase entonacional (2 a 5 palavras), intervalo interpausal (2 sílabas ou mais) e oração</p> <p>Tipo de sílaba: fonética</p> <p>Forma de contagem: manual, por número de núcleos silábicos</p> <p>Duração da PS: 130ms</p> <p>Conduta com PP: descartou</p> <p>Conduta com disfluência: fluência na fala foi critério para admissão dos sujeitos</p> <p>Resultados: TA global média com diferença significativa somente entre 2 dos 3 sujeitos (7,07 : 5,15) e taxas locais apontam a frase entonacional como domínio da variação</p>
HEWLETT e RENDALL (1998)	<p>Taxa(s): TE ("<i>speaking rate</i>") e TA</p> <p>Tipo de material de fala: conversação (duração NIT) com o entrevistador sobre 2 ou 3 temas</p> <p>Língua: inglês urbano (de Edimburgo) e rural (de Orkney)</p> <p>Sujeitos: 24 (11M e 13F), 15 a 18 anos, ocupações NIT</p> <p>Tipo de intervalo de fala: turno de fala (11 sílabas ou mais) e intervalo interpausal</p> <p>Tipo de sílaba: fonética</p> <p>Forma de contagem: manual, critério NIT</p> <p>Duração da PS: 200ms</p> <p>Conduta com PP: descartou só na TA</p> <p>Conduta com disfluência: considerou em ambas as taxas (TA conforme Laver, 1994)</p> <p>Resultados: TE média (infere-se, global) sem diferença significativa entre os dialetos TA média (infere-se, global) é maior no dialeto rural do que no urbano (6,02 : 5,52)</p>
JACEWICZ et al. (2009)	<p>Taxa(s): TA ("<i>speaking rate</i>")</p> <p>Tipo de material de fala: 10 a 15min de narrativa sobre tema livre</p> <p>Língua: inglês de Wisconsin e da Carolina do Norte</p> <p>Sujeitos: 80 (40M e 40F), 20 a 65 anos, ocupações variadas</p> <p>Tipo de intervalo de fala: intervalo interpausal (5 sílabas ou mais)</p> <p>Tipo de sílaba: fonética</p> <p>Forma de contagem: manual, critério NIT</p> <p>Duração da PS: NIT</p> <p>Conduta com PP: descartou</p> <p>Conduta com disfluência: descartou</p> <p>Resultados: TA local média aponta diferenças significativas quanto ao dialeto (5,41 : 4,81) e a idade em um dos dialetos (5,58 : 5,25, respectivamente, adultos jovens e mais velhos)</p>
JACEWICZ et al. (2010)	<p>Taxa(s): TA ("<i>speaking rate</i>")</p> <p>Tipo de material de fala: ao menos 10 min sobre tema livre</p> <p>Língua: inglês de Wisconsin e da Carolina do Norte</p> <p>Sujeitos: 192 (95M e 97F), 8 a 91 anos, ocupações NIT</p> <p>Tipo de intervalo de fala: intervalo interpausal (5 sílabas ou mais)</p> <p>Tipo de sílaba: fonética</p> <p>Forma de contagem: manual, critério NIT</p> <p>Duração da PS: NIT</p> <p>Conduta com PP: descartou</p> <p>Conduta com disfluência: descartou</p> <p>Resultados: TA local média aponta diferenças significativas quanto ao dialeto (5,21 : 4,80), sexo (5,09M : 4,92F) e idade (comportamento de ascensão e queda da TA no progredir da idade com pico aos 45 anos)</p>

APÊNDICE A – Panorama de estudos sobre TE e TA com cálculo em sílabas por segundo, a partir de fala espontânea não infantil de indivíduo em condições normais de linguagem (continuação)

Referência	Dados do estudo
JESSEN (2007)	<p>Taxa(s): TA Tipo de material de fala: 2 a 6min de narrativa, presencial e ao telefone, sobre figura de jogo de linguagem Língua: alemão Sujeitos: 100M; 21 a 63 anos; cientistas forenses, funcionários administrativos, policiais e estudantes da academia de polícia Tipo de intervalo de fala: trecho de memória (4 sílabas ou mais) Tipo de sílaba: fonética Forma de contagem: manual, por número de sílabas auditivamente percebidas Duração da PS: NIT Conduta com PP: descartou Conduta com disfluência: descartou Resultados: diferença não significativa entre a TA local média presencial (5,19) e ao telefone (5,41)</p>
KENDAL (2009)	<p>Taxa(s): TA ("<i>speech rate</i>") Tipo de material de fala: entrevista sociolinguística, tendo sido obtido de 3 a 74min de duração Língua: inglês Sujeitos: 104 (47M e 57F), 10 a 90 anos, ocupações NIT Tipo de intervalo de fala: sentença Tipo de sílaba: fonética Forma de contagem: automática, baseada na notação ortográfica Duração da PS: 60ms Conduta com PP: descartou Conduta com disfluência: considerou (TA conforme Laver, 1994) Resultados: TA local média aponta diferenças significativas quanto ao dialeto, à etnicidade e à idade (decrésimo da taxa com o progredir da idade)</p>
KÜNZEL (1997)	<p>Taxa(s): TE ("<i>syllable rate</i>") e TA Tipo de material de fala: cerca de 90s de fala, presencial e ao telefone, em opinião pessoal sobre tema controverso Língua: alemão Sujeitos: 10 (5M e 5F), 20 a 26 anos, estudantes universitários de ciências sociais ou direito Tipo de intervalo de fala: intervalo interpausal (1 sílaba ou mais) Tipo de sílaba: canônica (conforme nota de rodapé em Jessen, 2007, p.59) Forma de contagem: manual, critério NIT Duração da PS: 100ms Conduta com PP: descartou só na TA Conduta com disfluência: considerou em ambas as taxas (TA conforme Laver, 1994) Resultados: diferença não significativa entre a TE global presencial (4,50) e ao telefone (4,55) e entre a TA local média presencial (5,96) e ao telefone (5,90)</p>
SYRDAL (1996)	<p>Taxa(s): TE ("<i>speaking rate</i>") Tipo de material de fala: 5 a 7min de conversação espontânea ao telefone Língua: 8 dialetos do inglês americano Sujeitos: 160 (80M e 80F), idades e ocupações NIT Tipo de intervalo de fala: trecho do integral Tipo de sílaba: NIT Forma de contagem: NIT Duração da PS: NIT Conduta com PP: NIT Conduta com disfluência: NIT Resultados: TE global aponta diferenças não significativas quanto ao dialeto e sexo (4,76M : 4,77F)</p>

APÊNDICE A – Panorama de estudos sobre TE e TA com cálculo em sílabas por segundo, a partir de fala espontânea não infantil de indivíduo em condições normais de linguagem (conclusão)

Referência	Dados do estudo
VERHOEVEN et al. (2004)	<p>Taxa(s): TE ("<i>speaking rate</i>") e TA</p> <p>Tipo de material de fala: ao menos 15min de narrativa sobre temas gerais</p> <p>Língua: 4 dialetos do holandês da Holanda e 4 dialetos do holandês falado na Bélgica</p> <p>Sujeitos: 160 (80M e 80F), 26 a 55 anos, professores de holandês</p> <p>Tipo de intervalo de fala: 10 trechos de fala</p> <p>Tipo de sílaba: supõe-se, ortográfica</p> <p>Forma de contagem: automática, baseada no número de vogais (supõe-se, da notação ortográfica)</p> <p>Duração da PS: NIT</p> <p>Conduta com PP: considerou em ambas as taxas (TA conforme Laver, 1994)</p> <p>Conduta com disfluência: NIT</p> <p>Resultados: TE e TA globais com diferença significativa quanto ao país de origem; ao sexo (TE 4,23M : 4,01F e TA 4,79M : 4,50F); e à idade, sendo os mais jovens os falantes mais rápidos (TE 4,23 e TA 4,78 nos jovens e TE 4,01 e TA 4,52 nos mais velhos)</p>
Este estudo	<p>Taxa(s): TE e TA</p> <p>Tipo de material de fala: média de 2min de fala, presencial (ciente) e ao telefone (não ciente), não contemporâneas, sobre temas gerais</p> <p>Língua: português brasileiro</p> <p>Sujeitos: 7 (5M e 2F); 14 a 33 anos; suspeitos, indiciados ou réus em inquérito ou processo criminal</p> <p>Tipo de intervalo de fala: turno de fala e intervalo interpausal</p> <p>Tipo de sílaba: fonética</p> <p>Forma de contagem: manual, baseada no número de vocoides silábicos</p> <p>Duração da PS: 130ms</p> <p>Conduta com PP: considerou na TE e descartou, com reservas*, na TA</p> <p>Conduta com disfluência: considerou na TE e descartou, com reservas*, na TA</p>

Legenda: NIT= Não Informado no Texto, min= minutos, s= segundos, ms= milissegundos, M= masculino, F= feminino, PS= pausa silenciosa, PP= pausa preenchida.

* Foram aceitos como úteis os intervalos de fala nos quais constasse(m) alongamento(s) segmental(is) denotando ênfase, como um recurso de expressividade.

Fonte: A autora (2013).

Nota: Entre aspas consta a denominação de taxa diferenciada empregada pelo autor no respectivo estudo.

APÊNDICE B – Conjunto de comandos referentes às análises realizadas

Tabela 4 e Figuras 26 e 27

```
EXAMINE VARIABLES=TE_global media_TEs_locais TA_global media_TAs_locais
BY tipo_de_gravacao
PLOT HISTOGRAM NPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

Tabela 5 e 6

```
T-TEST PAIRS=TE_global TA_global WITH media_TEs_locais media_TAs_locais
(PAIRED)
/CRITERIA=CI(.9500)
/MISSING=ANALYSIS.
```

Tabela 7

```
T-TEST GROUPS=Sexo('F' 'M')
/MISSING=ANALYSIS
/VARIABLES=TE_global media_TEs_locais TA_global media_TAs_locais
/CRITERIA=CI(.95).
```

Tabela 8

```
T-TESTPAIRS=TE_global.avisada media_TEs_locais.avisada TA_global.avisada
media_TAs_locais.avisada WITH TE_global.desavisada media_T
Fs_locais.desavisada TA_global.desavisada media_TAs_locais.desavisada
(PAIRED)
/CRITERIA=CI(.9500)
/MISSING=ANALYSIS.
```

Tabela 9 e Figura 28

```
UNIANOVA TE_global BY tipo_de_gravacao Sexo
/METHOD=SSTYPE(3)
/INTERCEPT=EXCLUDE
/PLOT=PROFILE(Sexo*tipo_de_gravacao)
/PRINT=HOMOGENEITY OPOWER
/PLOT=RESIDUALS
/CRITERIA=ALPHA(.05)
/DESIGN=tipo_de_gravacao Sexo tipo_de_gravacao*Sexo.
UNIANOVA media_TEs_locais BY tipo_de_gravacao Sexo
/METHOD=SSTYPE(3)
/INTERCEPT=EXCLUDE
/PLOT=PROFILE(Sexo*tipo_de_gravacao)
/PRINT=HOMOGENEITY OPOWER
/PLOT=RESIDUALS
/CRITERIA=ALPHA(.05)
/DESIGN=tipo_de_gravacao Sexo tipo_de_gravacao*Sexo.
```

Tabela 10 e Figura 29

```

UNIANOVA TA_global BY tipo_de_gravacao Sexo
/METHOD=SSTYPE(3)
/INTERCEPT=EXCLUDE
/PLOT=PROFILE(Sexo*tipo_de_gravacao)
/PRINT=HOMOGENEITY OPOWER
/PLOT=RESIDUALS
/CRITERIA=ALPHA(.05)
/DESIGN=tipo_de_gravacao Sexo tipo_de_gravacao*Sexo.
UNIANOVA media_TAs_locais BY tipo_de_gravacao Sexo
/METHOD=SSTYPE(3)
/INTERCEPT=EXCLUDE
/PLOT=PROFILE(Sexo*tipo_de_gravacao)
/PRINT=HOMOGENEITY OPOWER
/PLOT=RESIDUALS
/CRITERIA=ALPHA(.05)
/DESIGN=tipo_de_gravacao Sexo tipo_de_gravacao*Sexo.

```

Tabela 11

```

ONEWAY TE_global media_TEs_locais TA_global media_TAs_locais BY sujeito
/STATISTICS DESCRIPTIVES HOMOGENEITY BROWNFORSYTHE
/PLOT MEANS
/MISSING ANALYSIS
/POSTHOC=T3 ALPHA(0.05).
SORT CASES BY tipo_de_gravacao
SPLIT FILE SEPARATE BY tipo_de_gravacao.
ONEWAY TE_global media_TEs_locais TA_global media_TAs_locais BY sujeito
/STATISTICS DESCRIPTIVES HOMOGENEITY BROWNFORSYTHE
/PLOT MEANS
/MISSING ANALYSIS
/POSTHOC=T3 ALPHA(0.05).

```

Tabela 12

```

ONEWAY TE_global media_TEs_locais TA_global media_TAs_locais BY sujeito
/STATISTICS DESCRIPTIVES HOMOGENEITY BROWNFORSYTHE
/MISSING ANALYSIS
/POSTHOC=LSD T3 ALPHA(0.05).

```

Tabela 13

```

ONEWAY TE_global media_TEs_locais TA_global media_TAs_locais BY sujeito
/STATISTICS DESCRIPTIVES HOMOGENEITY BROWNFORSYTHE
/PLOT MEANS
/MISSING ANALYSIS
/POSTHOC=T3 ALPHA(0.05).

```

Tabela 14

```

CORRELATIONS
/VARIABLES=TE_global Escolaridade Gap_Temp Idade
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

```

CORRELATIONS
/VARIABLES=media_TEs_locais Escolaridade Gap_Temp Idade
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
CORRELATIONS
/VARIABLES=TA_global Escolaridade Gap_Temp Idade
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
CORRELATIONS
/VARIABLES=media_TAs_locais Escolaridade Gap_Temp Idade
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

Tabela 15, 17 e 18 e Figuras 30 e 31

```

REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.25)
/NOORIGIN
/DEPENDENT TE_global
/METHOD=BACKWARD tipo_de_gravacao Sexo escolaridade.p gap.p idade.p
/SAVE RESID.
EXAMINE VARIABLES=RES_1
/PLOT NPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.25)
/NOORIGIN
/DEPENDENT media_TEs_locais
/METHOD=BACKWARD tipo_de_gravacao Sexo escolaridade.p gap.p idade.p
/SAVE RESID.
EXAMINE VARIABLES=RES_2
/PLOT NPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.25)
/NOORIGIN

```

```

/DEPENDENT TA_global
/METHOD=BACKWARD tipo_de_gravacao Sexo escolaridade.p gap.p idade.p
/SAVE RESID.
EXAMINE VARIABLES=RES_3
/PLOT NPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.25)
/NOORIGIN
/DEPENDENT media_TAs_locais
/METHOD=BACKWARD tipo_de_gravacao Sexo escolaridade.p gap.p idade.p
/SAVE RESID.
EXAMINE VARIABLES=RES_4
/PLOT NPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.

```

Tabela extra 19 e 20, Figuras 32 e 33

```

FILTER BY TE.
EXECUTE.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.25)
/NOORIGIN
/DEPENDENT TXY
/METHOD=BACKWARD sexo escolaridade gaptemp tipo_de_grav idade duração
/SAVE RESID.
EXAMINE VARIABLES=RES_1
/PLOT HISTOGRAM NPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
FILTER BY TA.
EXECUTE.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA=PIN(.05) POUT(.25)
/NOORIGIN
/DEPENDENT TXY

```

```

/METHOD=BACKWARD sexo escolaridade gaptemp tipo_de_grav idade duração
/SAVE RESID.
EXAMINE VARIABLES=RES_2
/PLOT HISTOGRAM NPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.

```

Seção 4.1.2.1.1

```

SORT CASES BY tipo_TX.
SPLIT FILE SEPARATE BY tipo_TX.
CORRELATIONS
/VARIABLES=idade núm_sílabas
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

Seção 4.1.2.1.2

```

CORRELATIONS
/VARIABLES=TXY duração
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

Seção 4.1.2.2

```

COMPUTE duracaosilaba=1/TXY.
EXECUTE.
CORRELATIONS
/VARIABLES=núm_sílabas duracaosilaba
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

Tabela 22

```

CORRELATIONS
/VARIABLES=Gap_Temp Idade
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE

```

APÊNDICE C – Condicionamento da TE e da TA devido às variáveis independentes idade, sexo, escolaridade, *gap* temporal, tipo de gravação e tamanho do intervalo de fala

	TE		TA	
	Global	Local média	Global	Local média
Idade	NS, Tend – (correlação)	NS, Tend – (correlação)	NS, Tend – (correlação)	NS, Tend – (correlação)
Sexo	NS (ANOVA, Teste t) Tend M>F na Desav e F>M na Avis	NS (ANOVA, Teste t) Tend M>F na Desav e F>M na Avis	NS (ANOVA, Teste t) Tend F>M na Desav e na Avis	NS (ANOVA, Teste t) Tend M>F na Desav e F>M na Avis
Escolaridade	NS, Tend + (correlação)	NS, Tend + (correlação)	NS, Tend + (correlação)	NS, Tend + (correlação)
Gap temporal	NS, Tend + (correlação)	NS, Tend + (correlação)	Sig (correlação)	Sig (correlação)
Tipo de gravação	Sig (ANOVA, Teste t)	Sig (ANOVA, Teste t)	NS, Tend Desav>Avis (ANOVA, Teste t)	NS, Tend Desav>Avis (ANOVA, Teste t)
Tamanho do intervalo	Sig (correlação)	Sig (correlação)	NS, Tend – (correlação)	NS, Tend – (correlação)

Legenda: NS= não significativo, Sig= significativo, Tend= tendência, Tend + = tendência positiva (relação direta entre as variáveis), Tend – = tendência negativa (relação inversa entre as variáveis), M= masculino, F= feminino, Desav= gravação desavisada, Avis= gravação avisada, ANOVA= Análise de variância.

Fonte: A autora (2013).

Nota: Entre parênteses é indicado o teste a partir do qual se obteve a NS/ tendência ou Sig.
">" indica que maiores taxas são associadas ao fator à esquerda (do sinal).

Os resultados estatisticamente significativos aparecem destacados, em área sombreada.

ANEXO A – Cópia do Ofício n.º 1815/2009-DG/IGP



**GOVERNO DO ESTADO
RIO GRANDE DO SUL**

SECRETARIA DA SEGURANÇA PÚBLICA
INSTITUTO-GERAL DE PERÍCIAS
GABINETE DA DIREÇÃO-GERAL



Ofício n.º 1815/2009-DG/IGP

Porto Alegre, 07 de julho de 2009.

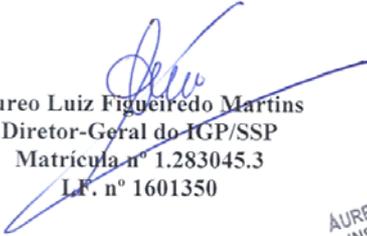
Prezada Senhora:

Ao cumprimentá-la cordialmente, vimos por meio deste, informá-la que esta Direção-Geral nada tem a opor quanto ao Projeto de Pesquisa Científica intitulado “TAXA DE ELOCUÇÃO E VARIAÇÃO LINGUÍSTICA”, aceito pelo PROGRAMA DE PÓS GRADUAÇÃO EM LETRAS / PUCRS e atualmente analisado, através do Processo n.º 2485-1205/09-4.

Desse modo, sabedores que o referido Projeto foi avaliado pela respectiva Banca de Seleção de Doutorado da PUC/RS e, estando de acordo com as atribuições constitucionais do IGP, o desenvolvimento de pesquisas em sua área de atuação, enviamos o expediente acima referido ao conhecimento da Diretora do Departamento de Criminalística/IGP, que dele ficou plenamente ciente.

Desse modo, renovamos votos de estima e consideração.

Atenciosamente,


Áureo Luiz Figueiredo Martins
Diretor-Geral do IGP/SSP
Matrícula n.º 1.283045.3
L.F. n.º 1601350

ÁUREO LUIZ FIGUEIREDO MARTINS
INSTITUTO-GERAL DE PERÍCIAS
Diretor - Geral
Mat. n.º 1283.0453 - IF n.º 1601350/01

Ilma. Sra.,
Cintia Schivinski Gonçalves,
M. D. Perita Criminal,
Nesta capital.

ANEXO B – Valores de referência relativos a vogais do PB adotados neste estudo

TABLE I. Geometric averages of vowel duration, F0, F1, F2, F3, and formant ceilings for female (F) and male (M) speakers of BP and EP. Between parentheses: the standard deviations, converted back to ratios of ms and Hz. Every cell represents ten speakers.

			/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
BP	Duration (ms)	F	99 (1.210)	122 (1.195)	141 (1.192)	144 (1.173)	139 (1.145)	123 (1.151)	100 (1.201)
		M	95 (1.216)	109 (1.200)	123 (1.232)	127 (1.186)	123 (1.209)	110 (1.189)	100 (1.205)
	F0 (Hz)	F	242 (1.096)	219 (1.098)	210 (1.092)	209 (1.088)	211 (1.093)	225 (1.098)	252 (1.087)
		M	137 (1.199)	131 (1.186)	124 (1.183)	122 (1.199)	122 (1.178)	132 (1.194)	140 (1.223)
	F1 (Hz)	F	307 (1.198)	425 (1.082)	646 (1.076)	910 (1.078)	681 (1.087)	442 (1.094)	337 (1.192)
		M	285 (1.077)	357 (1.077)	518 (1.089)	683 (1.095)	532 (1.160)	372 (1.100)	310 (1.070)
	F2 (Hz)	F	2676 (1.056)	2468 (1.061)	2271 (1.051)	1627 (1.062)	1054 (1.099)	893 (1.054)	812 (1.054)
		M	2198 (1.078)	2028 (1.076)	1831 (1.072)	1329 (1.088)	927 (1.108)	804 (1.092)	761 (1.100)
	F3 (Hz)	F	3296 (1.073)	3074 (1.048)	2897 (1.077)	2625 (1.119)	2653 (1.114)	2627 (1.158)	2691 (1.123)
		M	2952 (1.066)	2719 (1.077)	2572 (1.050)	2324 (1.084)	2335 (1.069)	2380 (1.060)	2309 (1.078)
	Ceiling (Hz)	F	6001 (1.086)	5933 (1.094)	5463 (1.166)	5577 (1.076)	5260 (1.137)	4938 (1.113)	5090 (1.095)
		M	5230 (1.155)	5063 (1.181)	5010 (1.137)	4463 (1.105)	4436 (1.077)	4522 (1.068)	4458 (1.064)

Fonte: Escudero et al. (2009, p.1383).

ANEXO C – Cópia do OF.CEP-1107/10



Pontifícia Universidade Católica do Rio Grande do Sul
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COMITÊ DE ÉTICA EM PESQUISA

OF.CEP-1107/10

Porto Alegre, 08 de outubro de 2010.

Senhora Pesquisadora,

O Comitê de Ética em Pesquisa da PUCRS apreciou e aprovou seu protocolo de pesquisa registro CEP 10/05206 intitulado **"Taxa de elocução e variação linguística"**.

Salientamos que seu estudo pode ser iniciado a partir desta data.

Os relatórios parciais e final deverão ser encaminhados a este CEP.

Atenciosamente,


Prof. Dr. Rodolfo Herberto Schneider
Coordenador do CEP-PUCRS
Prof^a Dr. Virginia Minghelli Schmitt
Coordenadora Substituta
Comitê de Ética em Pesquisa da PUCRS
CEP - PUCRS

Ilma. Sra.
Profa. Cláudia Regina Brescancini
FALE
Nesta Universidade

PUCRS

Campus Central
Av. Ipiranga, 6690 – 3ºandar – CEP: 90610-000
Sala 314 – Fone Fax: (51) 3320-3345
E-mail: cep@pucrs.br
www.pucrs.br/prppg/cep

ANEXO D – Currículo da acadêmica

Endereço para acesso à versão integral: <http://lattes.cnpq.br/0824959777665559>

Dados pessoais

Nome Cintia Schivinski Gonçalves
Filiação Albino Reinaldo Schivinski e Bernadete da Silva Schivinski
Nascimento 18/11/1973 - Porto Alegre/RS - Brasil
Carteira de Identidade 3049414489 SJS - RS - 09/09/1998
CPF 647.791.060-87

Endereço residencial Rua Carlos Reverbel, 160
 São José - Canoas
 92425-100, RS - Brasil
 Telefone: 51 81135892

Endereço profissional Secretaria da segurança Pública do RS, Instituto-Geral de Perícias
 Av. Princesa Isabel, 1056
 Santana - Porto Alegre
 90620-000, RS - Brasil
 Telefone: 51 32236677

Endereço eletrônico E-mail para contato : cintia-goncalves@igp.rs.gov.br
 E-mail alternativo : cintiasg@ig.com.br

Formação acadêmica/titulação

- 2000 - 2002** Mestrado em Letras/ AC Linguística Aplicada.
 PUCRS - Pontifícia Universidade Católica do Rio Grande do Sul, PUCRS, Brasil
 Título: Variáveis lingüísticas facilitadoras na reabilitação fonológica das líquidas não-laterais, Ano de obtenção: 2002
 Orientador: Profa Dr Regina Ritter Lamprecht
 Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
- 2003 - 2009** Especialização em Motricidade Orofacial.
 CFFa - Conselho Federal de Fonoaudiologia, CFFA (CEFAC), Brasil
 Título: Estudo da relação entre presença de frênulo lingual curto e/ou anteriorizado e a dorsalização de /r/ na articulação da fala
 Orientador: Profa Dr Adriana Velasquez Feijó
- 2002 - 2002** Especialização em Linguagem.
 CFFa - Conselho Federal de Fonoaudiologia, CFFA, Brasil
 Título: Cf. Resolução CFFa
- 1994 - 1998** Graduação em Fonoaudiologia.
 IPA/ IMEC - Instituto Metodista de Educação e Cultura, IPA/ IMEC, Brasil
 Título: Desvios fonéticos e fonológicos em paciente adulto
 Orientador: Profa Dr Carla Aparecida Cielo
- 1989 - 1991** Ensino Profissional de nível técnico em Publicidade e Propaganda.
 EEIP - Escola Estadual de 2o Grau Irmão Pedro, EEIP, Brasil
- 1989 - 1991** Ensino Médio (2º grau).
 EEIP - Escola Estadual de 2º Grau Irmão Pedro, EEIP, Brasil
- 1980 - 1988** Ensino Fundamental (1º grau).
 Escola Estadual de 1º Grau Araújo Porto Alegre, EEAPOA, Brasil

- 2007** Aperfeiçoamento em Capacitação Nacional de Peritos Criminais em Fonética Forense.
Ministério da Justiça-Secretaria Nacional de Segurança Pública, SENASP, Brasil
Título: Capacitação Nacional para Peritos Criminais em Fonética Forense (586h)
Bolsista do(a): Secretaria Nacional de Segurança Pública
- 2005 - 2005** Aperfeiçoamento em Formação Integrada para Futuros Servidores do IGP (718h).
FDRH - Fundação para o desenvolvimento de RH, FDRH, Brasil
Título: Prova teórica (para cargo de Perito Criminal Área Fonoaudiologia)
Bolsista do(a): Secretaria da Justiça e da Segurança do Rio Grande do Sul

Formação complementar (últimos 5 anos)

- 2012 - 2012** Curso de curta duração – Intensivo.
CEV Centro de Estudo da Voz S/S Limitada, CEV, São Paulo, Brasil
- 2012 - 2012** Curso de curta duração – Redação Técnica.
SENASP/ MJ, SENASP/ MJ, Brasil
- 2012 - 2012** Curso de curta duração – Treinamento para avaliação percepto-auditiva da voz.
Dialoguefono, DIALOGUEFONO, Brasil
- 2012 - 2012** Curso de curta duração – Análise Criminal.
SENASP/ MJ, SENASP/ MJ, Brasil
- 2012 - 2012** Curso de curta duração – Escola Itinerante de Direitos Humanos e Segurança Pública.
SENASP/ MJ, SENASP/ MJ, Brasil
- 2011 - 2011** Extensão universitária – SPSS aplicado à gestão.
PUCRS - Pontifícia Universidade Católica do Rio Grande do Sul, PUCRS, Brasil
- 2011 - 2011** Curso de curta duração – Reciclagem e Atualização Clínico-Científica em voz.
CEV Centro de Estudo da Voz S/S Limitada, CEV, São Paulo, Brasil
- 2010 - 2010** Extensão universitária – Fundamentos de Fonologia.
PUCRS - Pontifícia Universidade Católica do Rio Grande do Sul, PUCRS, Brasil
Bolsista do(a): Conselho Nacional de Desenvolvimento Científico e Tecnológico
- 2010 - 2010** Curso de curta duração – Fundamentos de estatística e análise com o SPSS.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil
- 2008 - 2008** Curso de curta duração – Treinamento Técnico do Uso da Arma de Fogo.
Academia de Polícia Militar, APM, Brasil
- 2007 - 2007** Curso de curta duração – Planejamento Estratégico.
Secretaria de Segurança Pública RS, SSP-RS, Brasil

Atuação profissional (últimos 5 anos)

1. Secretaria da Segurança Pública - SSP-RS

Vínculo institucional

- 2005 - Atual** Vínculo: Servidor público, Enquadramento funcional: Perita Criminal - Área Fonoaudiologia, Carga horária: 40, Regime: Dedicção exclusiva

2. Centro Universitário La Salle - Canoas - UNILASALLE

Vínculo institucional

2003 - 2010 Vínculo: Colaborador (curso noturno), Enquadramento funcional: Professor do ensino superior, Carga horária: 8, Regime: Parcial

Atividades

02/2003 - Atual Graduação, Psicopedagogia Clínica e Institucional

Disciplinas ministradas:

Aquisição e desenvolvimento da linguagem I - Linguagem Oral (4 cred.) , Aquisição e desenvolvimento da linguagem II - Linguagem Escrita (4 cred.)

3. Centro Universitário Feevale - FEEVALE

Vínculo institucional

2003 - 2005 Vínculo: Colaborador, Enquadramento funcional: Professor de terceiro grau (adjunto), Carga horária: 26, Regime: Parcial

Atividades

02/2003 - 12/2005 Graduação, Fonoaudiologia

Disciplinas ministradas:

Fundamentos de lingüística aplicada (4 cred.) , Princípios de motricidade oral e articulação (4 cred.) , Linguagem III/ desvios fonéticos e fonológicos (2 cred.) , Estágio de observação em desvios fonéticos e fonológicos (3 turmas de 4 cred.) , Estágio de observação em sistema estomatognático (4 cred.)

Áreas de atuação (últimos 5 anos)

1. Fonética Forense
2. Pesquisa e Ensino na Área de Desvios Fonéticos e Fonológicos
3. Avaliação, diagnóstico e tratamento em Motricidade Orofacial e Fala
4. Uso de Telerradiografia e Cefalometria Aplicado à Fonoaudiologia

Produções (últimos 5 anos)

Produções Bibliográficas

Artigos completos publicados em periódicos

1. GONÇALVES, C. S., PORTO, A. C.
Proposta de análise perceptivo-auditiva de voz e fala para uso em fonética forense. Revista IGP. , v.3, p.23 - 25, 2007.
2. GONÇALVES, C. S., FERREIRO, M. C.
Estudo da relação entre presença de frênulo lingual curto e/ou anteriorizado e a dorsalização de /r/ na articulação da fala. Revista Cefac Atualização Científica Em Fonoaudiologia. , v.8, p.56 - 60, 2006.
3. GONÇALVES, C. S., LAMPRECHT, R. R.
O potencial de facilitação lingüística na reabilitação fonológica de /r/ e /R/. Fonoaudiologia Brasil (Impresso). v.3, 2006.
4. GONÇALVES, C. S., OLIVEIRA, A. G., KLEY, F. H., KAUTZMANN, M. C.
Caracterização do perfil clínico de sujeitos com desvios fonéticos e fonológicos: dados preliminares. Jornal Brasileiro de Fonoaudiologia. , v.5, p.304 - 313, 2005.

Capítulos de livros publicados

1. GONÇALVES, C. S.

A fala da língua presa In: Anuário de Fonoaudiologia Feevale, 2ª ed. Novo Hamburgo : Editora Feevale, 2004, p. 61-66.

Trabalhos publicados em anais de eventos (resumo)

1. GOMES, D. H., LIMBERGER, C.M., CIMADON, É., GONÇALVES, C. S.

Crianças com desvios fonéticos e fonológicos versus crianças com desvios exclusivamente fonológicos: tipologia e especificidades na aplicação dos processos In: XII Congresso Brasileiro de Fonoaudiologia, 2004, Foz do Iguaçu.

Revista da Sociedade Brasileira de Fonoaudiologia-Suplemento especial. SBFa (Tema Livre), 2004.

2. OLIVEIRA, A.G., BACKES, B., KAUTZMANN, M.C., GONÇALVES, C. S.

Nível de consciência fonológica de sujeitos com desvios fonológicos: comportamento na testagem e relação com o estado do sistema fonológico In: XII Congresso Brasileiro de Fonoaudiologia, 2004, Foz do Iguaçu.

Revista da Sociedade Brasileira de Fonoaudiologia-Suplemento especial. SBFa (Tema Livre), 2004.

3. GONÇALVES, C. S., LAMPRECHT, R. R.

O potencial de facilitação lingüística das palavras-estímulo usadas na reabilitação fonológica das líquidas não-laterais In: XII Congresso Brasileiro de Fonoaudiologia, 2004, Foz do Iguaçu.

Revista da Sociedade Brasileira de Fonoaudiologia-Suplemento especial. SBFa (Pôster), 2004.

4. GOULART, B.N.G., GONÇALVES, C. S., DIAS, W.B.

Proposta de estágio interdisciplinar Fonoaudiologia-Odontologia In: XII Congresso Brasileiro de Fonoaudiologia, 2004, Foz do Iguaçu.

Revista da Sociedade Brasileira de Fonoaudiologia-Suplemento especial. SBFa (Tema Livre), 2004.

Artigos em jornal de notícias

1. GONÇALVES, C. S., PORTO, A. C., PINTO, M. O.

Jornal Nossa Voz (consultora). Jornal Nossa Voz. RS, p.08 - 09, 2006.

2. GONÇALVES, C. S.

Escute seu filho (consultora). Zero Hora. Porto Alegre, v.678, Caderno Vida , 2004.

Apresentação de trabalho e palestra

1. GONÇALVES, C. S., PINTO, M. O.

Fonética acústica - instrumentos de análise, 2012. (Conferência ou palestra, Apresentação de Trabalho)

2. GONÇALVES, C. S.

Identificação, caracterização e interpretação da fala não-normal no confronto forense de voz e fala, 2012. (Seminário, Apresentação de Trabalho)

3. GONÇALVES, C. S., PORTO, A. C., CARNEIRO, D. O., PETRY, T.

Fonoaudiólogos - Peritos Criminais Oficiais: Atuação, Histórico e Expansão no Brasil, 2011. (Congresso, Apresentação de Trabalho)

4. GONÇALVES, C. S., PORTO, A. C.

A análise perceptivo-auditiva na verificação de falantes em fonética forense, 2006. (Seminário, Apresentação de Trabalho)

5. GONÇALVES, C. S., PORTO, A. C.

A análise perceptivo-auditiva na verificação de falantes em fonética forense, 2006. (Seminário, Apresentação de Trabalho)

6. GONÇALVES, C. S.

Problemas de Linguagem Oral no Processo Ensino-Aprendizagem, 2006. (Conferência ou palestra, Apresentação de Trabalho)

7. GONÇALVES, C. S.

Avaliando sistemas fonológicos: formalização de uma proposta fundamentada na Teoria Autossegmental e da Sílabas, 2004. (Congresso, Apresentação de Trabalho)

Demais produções técnicas (Curso de curta duração ministrado)

1. GONÇALVES, C. S.

Capacitação para a realização da perícia de Comparação de Locutor, 2012. (Aperfeiçoamento)

2. GONÇALVES, C. S.

Comportamento humano social e ocupacional - Módulo Fonoaudiologia, 2012. (Aperfeiçoamento)

3. GONÇALVES, C. S.

Formação para a realização de Exames de Comparação de Locutor, 2012. (Aperfeiçoamento)

4. GONÇALVES, C. S.

Perícia em Fonoaudiologia, 2009. (Extensão)

5. GONÇALVES, C. S., PORTO, A. C.

Perícias em registro de áudio, 2009. (Aperfeiçoamento)

6. GONÇALVES, C. S., DIAS, W.B.

Alterações Dento-faciais e Fonoaudiologia, 2004. (Extensão)

7. GONÇALVES, C. S., SANTOS, R.S.C.

Materiais e Práticas em Motricidade Oral - 1a ed., 2004. (Extensão)

8. GONÇALVES, C. S., SANTOS, R.S.C.

Materiais e Práticas em Motricidade Oral - 2a ed., 2004. (Extensão)

Orientações e Supervisões (últimos 5 anos)

Orientações e supervisões concluídas

Monografias de conclusão de curso de aperfeiçoamento/especialização

1. GHYSIO, M.R.S.. Práticas metodológicas que enfatizam a consciência fonológica e sua relação com os processos iniciais de escrita. 2008. Monografia (Alfabetização e letramento) - Centro Universitário La Salle - Canoas

Trabalhos de conclusão de curso de graduação

1. BISOGNIN, F.. Letramento em crianças com TDAH: uma abordagem psicopedagógica. 2010. Curso (Psicopedagogia Clínica e Institucional) - Centro Universitário La Salle - Canoas

2. BÖHM, G.C.. O trabalho psicopedagógico institucional na aprendizagem da escrita. 2007. Curso (Psicopedagogia Clínica e Institucional) - Centro Universitário La Salle - Canoas

Eventos (últimos 5 anos)

Participação em eventos

1. I Encontro Sul Letras, 2012. (Encontro)

O uso da taxa de fala e de articulação como parâmetro técnico-comparativo na perícia de Comparação de Locutor.

2. IX Seminário Nacional de Fonética Forense, I Seminário Nacional em Análise Forense de Imagens, 2012. (Seminário)

3. 19º Congresso Brasileiro de Fonoaudiologia e 8º Congresso Internacional de Fonoaudiologia, 2011. (Congresso)

4. III Colóquio de Linguística e Literatura - PPG em Letras PUCRS, 2010. (Outra)
Minicurso Análise acústica da fala.

5. VIII Seminário Nacional de Fonética Forense, 2010. (Seminário)

6. 9th International Conference on Computational Processing of the Portuguese Language, 2010. (Outra)

7. VII Seminário Nacional de Fonética Forense e IV Seminário Nacional de Perícias em Crimes de Informática, 2008. (Seminário)

8. X Seminário Nacional de Documentoscopia, 2008. (Seminário)

9. Atividades Integradas de Extensão UNILASALLE - IV Edição, 2007. (Outra)
Minicurso Linguística aplicada a fins forenses.

10. 4o Fórum de Fonoaudiologia na saúde pública, 2007. (Outra)

11. Atividades Integradas de Extensão UNILASALLE - III Edição, 2006. (Outra)
Minicurso Desvios fonológicos na aprendizagem.

12. I Seminário de Estudos e Pesquisas do Instituto Geral de Perícias, 2006. (Seminário)

13. I Seminário de Integração Pericial SETEC/ Polícia Federal e IGP, 2005. (Seminário)

14. XII Congresso Brasileiro de Fonoaudiologia, 2004. (Congresso)
Crianças com desvios fonéticos e fonológicos versus crianças com desvios exclusivamente fonológicos (+ 3 outros trabalhos).

15. Apresentação (Outras Formas) no(a) Atividades Integradas de Extensão UNILASALLE - I Edição, 2004. (Outra)
Minicurso Estrutura da memória e sua interferência no desenvolvimento da linguagem.

16. II Encontro Brasileiro da Síndrome do Respirador Bucal, 2004. (Encontro)

Bancas (últimos 5 anos)

Participação em banca de qualificação de trabalhos de conclusão

Graduação

1. GONÇALVES, C. S.

Participação em banca de MASCARELLO, F. Análise da produção da líquida não-lateral dorsal por falantes bilingües do português brasileiro e do dialeto vênето, 2007 (Fonoaudiologia) Universidade Feevale.