UNIVERSIDADE FEDERAL DE MINAS GERAIS

FACULDADE DE LETRAS

Programa de Pós-Graduação em Estudos Linguísticos

OLIVER RENATO DE ARAÚJO GOBBO

MARCADORES DISCURSIVOS EM UMA PERSPECTIVA INFORMACIONAL: ANÁLISE PROSÓDICA E ESTATÍSTICA

Belo Horizonte

2019

Oliver Renato de Araújo Gobbo

MARCADORES DISCURSIVOS EM UMA PERSPECTIVA INFORMACIONAL: ANÁLISE PROSÓDICA E ESTATÍSTICA

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Mestre em Linguística.

Área de concentração: Linguística Teórica e Descritiva

Linha de pesquisa: Estudos Linguísticos baseados em *corpora*

Orientador: Prof. Dr. Tommaso Raso

Belo Horizonte

2019

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

G574m

Gobbo, Oliver Renato de Araújo.
Marcadores discursivos em uma perspectiva informacional [manuscrito] : análise prosódica e estatística / Oliver Renato de Araújo Gobbo. – 2019.
163 f., enc..

Orientador: Tommaso Raso.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Estudos Linguísticos Baseados em Corpora.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 131-135.

Apêndices: f. 136-163.

Inclui CD-ROM com apêndices.

1. Linguística de corpus – Teses. 2. Análise prosódica (Linguística) – Teses. 3. Marcadores discursivos – Teses. I. Raso, Tommaso. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD : 410

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

FOLHA DE APROVAÇÃO

MARCADORES DISCURSIVOS EM UMA PERSPECTIVA INFORMACIONAL: ANÁLISE PROSÓDICA E ESTATÍSTICA

OLIVER RENATO DE ARAUJO GOBBO

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Estudos Linguísticos Baseados em Corpora.

Aprovada em 28 de fevereiro de 2019, pela banca constituída pelos membros:

Prof(a). Tommaso Raso - Orientador
UFMG

Prof(a). Plínio Almeida Barbosa
UNICAMP

Prof(a). Heliana Ribeiro de Mello
UFMG

Belo Horizonte, 28 de fevereiro de 2019.

Agradecimentos

Meus pais, Siléia e Renato, por tudo, tudo mesmo.

Nathalia, pela força e pela paciência com minhas longas ausências;

Funcionários do Poslin e da Fale, em especial à Elizete e ao Edson, pela alegria redentora nos momentos solitários;

Bárbara, Camila, Luis e Luis II, Fred, Giulia, Lucia, Saulo, Thais, e todos os colegas de LEEL, pela companhia, pela alegria, por toda ajuda, pela troca de experiências, de conhecimentos e tantas discussões valiosas;

Marcelo Vieira, Marcos Vinícius, Pablo Arantes, Plinio Barbosa e Renato Assunção, meus professores de prosódia e estatística, conhecimentos que levarei por toda minha vida;

Heliana Mello, pelo saber, pela ternura;

Tommaso Raso, pela oportunidade, por toda sabedoria partilhada. Pela paciência infinita, por iluminar meu caminho nas sombras da angústia e da ignorância.

“Data! Data! Data!” he cried impatiently. “I can’t make bricks without clay.”

“It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.”

"Having gathered these facts, Watson, I smoked several pipes over them, trying to separate those which were crucial from others which were merely incidental."

"…when you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth.”

(Sr. Arthur Conan Doyle)

RESUMO

Os marcadores discursivos (MDs) são uma categoria linguística muito problemática, apesar de muitos estudos dedicados a eles ao longo das últimas quatro décadas. Não há na literatura consenso sobre a terminologia empregada, sobre a definição da categoria e sobre como identificar os MDs e suas funções.

Uma proposta para estudo dos MDs é apresentada pela Teoria da Língua em Ato (L-AcT) (CRESTI, 2000; MONEGLIA, 2005; RASO, 2012; MONEGLIA e RASO, 2014), elaborada através da observação de *corpora* de fala espontânea. A L-AcT adota a visão de que a veiculação de funções comunicativas é feita através da prosódia, e não do item lexical, como é comumente defendido em outras propostas sobre os MDs. Sob esse ponto vista, os MDs seriam unidades informacionais de tipo dialógico (UDs) e não textual, pois se dirigirem ao interlocutor com diversas funções interacionais. São sempre isolados em unidades entonacionais e não compõem o conteúdo proposicional do enunciado.

Sob essa perspectiva, o presente trabalho analisa os MDs em um *subcorpus* da seção informal do *corpus* de fala espontânea C-ORAL-BRASIL (Mini BR Informal). Em uma primeira etapa, foi realizada uma revisão do Mini BR Informal em busca de todas as UDs que apresentam a forma prosódica compatível com os três tipos de MDs já satisfatoriamente descritos na literatura (RASO, 2014; RASO e VIEIRA, 2016): conativo, alocutivo e incipitário. Essas UDs foram submetidas a análise de discriminante linear (LDA) com objetivo de investigar estatisticamente os parâmetros prosódicos que mais diferenciam os tipos entre si. O modelamento estatístico obteve uma acurácia de 84,6% na classificação e indicou os parâmetros com maior efeito de distinção entre os tipos: a média de intensidade; a média de f0; a taxa de variação inicial da curva de f0 e a posição do máximo de f0.

As demais unidades, que não se enquadram nos três tipos já descritos, foram submetidas a análise de *clusters* para identificação de padrões estabelecidos estatisticamente entre seus parâmetros prosódicos. O resultado da análise identificou três grandes grupos com similaridade sob os parâmetros avaliados, levantando a hipótese que estes grupos correspondam a três tipos de UDs.

Palavras-chave: marcadores de discurso, prosódia, corpus de fala, estatística.

ABSTRACT

Discourse markers (DMs) are still a very problematic linguistic category, despite many dedicated studies over the past four decades. There is no consensus in the literature on terminology, on the category definition and on how to identify the DMs and their functions.

A proposal for the study of DMs is presented by the Language into Act Theory (L-AcT) (CRESTI, 2000; MONEGLIA, 2005; RASO, 2012; MONEGLIA and RASO, 2014), elaborated through spontaneous speech corpora observation. L-AcT supports that communicative functions are carried out through prosody, not through the lexical item, as it is traditionally adopted in DMs studies. From this perspective, DMs would be informational units of dialogic type (DUs) and not textual type, as they address the interlocutor performing several interactional functions. DMs are also always isolated in intonational units and they do not compose the utterance propositional content.

The present work investigates DMs in the informal section of C-ORAL-BRASIL spontaneous speech *subcorpus* (Mini BR Informal). In a first step, the Mini BR Informal was reviewed in a search for DUs presenting prosodic form matching three types already satisfactorily described in the literature (RASO, 2014; RASO and VIEIRA, 2016): conative, allocutive and incipt. These DUs were submitted to linear discriminant analysis (LDA) intended to investigate statistically which prosodic features better differentiate the types. The statistical model achieved 84.6% classification accuracy and pointed to four mostly discriminating features: average f0; average intensity; initial f0 slope and maximum f0 position.

Other units, which do not match the three types of DMs already described, were submitted to cluster analysis with the aim of statistically determining patterns based on their prosodic features. Results identified three major groups that merged by the similarity of the investigated features, proposing the hypothesis these groups relate to three other DUs types.

Keywords: discourse markers, prosody, spoken corpus, statistics.

LISTA DE FIGURAS

Figura 1: ALL com função de marcar coesão social (verde: intensidade; azul: f0) 29

Figura 2: ALL em posição interna. Em azul, o perfil tonal ajustado às unidades vizinhas (verde: intensidade; azul: f0) 29

Figura 3: Exemplo de CNT com perfil modulado-descendente vizinhas (verde: intensidade; azul: f0) 30

Figura 4: Exemplo de enunciado com dois CNTs consecutivos de perfil descendente vizinhas (verde: intensidade; azul: f0) 31

Figura 5: Exemplo de dúvida entre CNT e ALL vizinhas (verde: intensidade; azul: f0) 31

Figura 6: INP com perfil ascendente: inicia com consoante vozeada (iii) vizinhas (verde: intensidade; azul: f0) 33

Figura 7: INP com perfil descendente: inicia com obstruinte desvozeada e finaliza em ditongo (i) vizinhas (verde: intensidade; azul: f0) 34

Figura 8: Enunciado com duas UDs PHA (né) (verde: intensidade; azul: f0) ... 35

Figura 9: EXP com perfil nivelado 36

Figura 10: EXP com perfil ascendente 36

Figura 11: Exemplo de DCT no início de enunciado 37

Figura 12: ALL reclassificado como AUX_n por sobreposição 46

Figura 13: ALL reclassificado como AUX_n por ausência de informações acústicas 47

Figura 14: ALL reclassificado como CNT 47

Figura 15: Extrato da tabela gerada pelo Winpitch 52

Figura 16: Extrato da tabela de unidades tonais 53

Figura 17: Exemplo de TextGrid gerado para a segmentação silábica e para alinhamento das unidades tonais (bfammn05-90) 54

Figura 18: Exemplo de TextGrid após o alinhamento e a segmentação (bfammn05-90)..... 55

Figura 19: Posição da UD em relação à unidade de referência, em UT 58

Figura 20: Ilustração do escore Z de 100ms para vogais com duração intrínseca diferentes..... 61

Figura 21: Objeto pitch originalmente gerado (limites entre 140 e 450Hz)..... 69

Figura 22: Objeto pitch manualmente ajustado. 70

Figura 23: TextGrid com espectrograma de parte do enunciado bfammn02[009]. 70

Figura 24: Esquema da definição dos parâmetros de variação de f0 em relação à tônica. 73

Figura 25: Distribuição dos parâmetros de duração por tipo de UD. 80

Figura 26: Distribuição do número de sílabas por tipo de UD. 80

Figura 27: Correlações entre os parâmetros de duração. 82

Figura 28: Distribuição dos parâmetros de intensidade por tipo de UD. 84

Figura 29: Correlações entre os parâmetros de intensidade. 85

Figura 30: Distribuição dos parâmetros de f0 por tipo de UD. 87

Figura 31: Correlações entre dos parâmetros de f0. 88

Figura 32: Distribuição dos parâmetros de variação de f0 por tipo de UD. 90

Figura 33: Correlações entre os parâmetros de variação de f0. 91

Figura 34: Distribuição dos parâmetros de alinhamento relativos à intensidade por tipo de UD. 93

Figura 35: Correlações entre parâmetros de alinhamento relativos à intensidade. 94

Figura 36: Distribuição dos parâmetros de alinhamento relativos à f0 por tipo de UD. 96

Figura 37: Correlações entre os parâmetros de alinhamento relativos à f0. 97

Figura 38: Ilustração de problemas de subjauste (linha rosa) e sobreajuste (linha verde) em modelos de ajuste de curvas (esquerda) e separação de grupos (direita). 102

Figura 39: Exemplo de árvore de combinações para 5 parâmetros usadas no método Leaps and Bounds..... 104

Figura 40: Estimativa do erro quadrático médio (C_p) para as 60 melhores combinações para cada número de parâmetros. 106

Figura 41: Percentual de vezes em que cada parâmetro foi escolhido entre as melhores combinações. 107

Figura 42: Esquema de tabela de confusão. 109

Figura 43: Acurácia global e por tipo máxima obtido em função da quantidade de parâmetros. 111

Figura 44: Separação dos dados pelos discriminantes. 113

Figura 45: Acurácia global e por tipo máxima obtido em função da quantidade de parâmetros, sem N_sil e Dur. 115

Figura 46: Exemplo de dendograma 116

Figura 47: Esquema da escolha do número de classes 117

Figura 48: Distribuição dos parâmetros de duração por agrupamento..... 122

Figura 49: Distribuição dos parâmetros de intensidade por agrupamento. 122

Figura 50: Distribuição dos parâmetros de f0 por agrupamento..... 123

Figura 51: Distribuição dos parâmetros de variação de f0 por agrupamento. 123

Figura 52: Distribuição dos parâmetros de alinhamento relativos à intensidade por agrupamento. 124

Figura 53: Distribuição dos parâmetros de alinhamento relativos à f0 por agrupamento. 124

Figura 54: Dendograma do agrupamento hierárquico aglomerativo com todos os parâmetrso para todos AUXs 119

Figura 55: Exemplos de fronteiras de classificação linear (esquerda) e quadráticas (direita)..... 126

LISTA DE TABELAS

Tabela 1: Número de palavras e distribuição por contexto social e tipologia no C-ORAL-BRASIL 40

Tabela 2: Número de palavras e distribuição por contexto social e tipologia no Mini BR Informal 42

Tabela 3: Característica dos textos do Mini BR Informal 43

Tabela 4: Quantidade de MDs por etiqueta, antes e após a revisão 45

Tabela 5: Quantidade e frequência de UDs por tipologia interacional 48

Tabela 6: Quantidade e frequência de UDs por contexto social 48

Tabela 7: Quantidade e frequência de UDs por contexto e tipologia 48

Tabela 8: Preenchimento lexical das UDs 51

Tabela 9: Média e desvios padrões dos segmentos utilizados para o cálculo da duração padronizada 63

Tabela 10: Quantidade de UDs analisadas em comparação com a quantidade de UDs no Mini BR Informal 78

Tabela 11: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de duração. 81

Tabela 12: Sumário dos parâmetros de duração por tipo de UD 81

Tabela 13: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de intensidade. 84

Tabela 14: Sumário dos parâmetros de intensidade por tipo de UD 85

Tabela 15: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de f0 87

Tabela 16: Sumário dos parâmetros de f0 por tipo de UD 88

Tabela 17: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de variação de f0. 90

Tabela 18: Sumário dos parâmetros de variação de f0 por tipo de UD 91

Tabela 19: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de alinhamento relativos à intensidade. 93

Tabela 20: Sumário dos parâmetros de alinhamento relativos à intensidade por tipo de UD 94

Tabela 21: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de alinhamento relativos à f0..... 96

Tabela 22: Sumário dos parâmetros de alinhamento relativos à f0 por tipo de UD 97

Tabela 23: Coeficientes de correlação entre todos os parâmetros. 99

Tabela 24: Descrição dos parâmetros estimados. 100

Tabela 25: Tabela de confusão para a melhor combinação de 9 parâmetros 112

Tabela 26: Coeficientes das funções discriminantes..... 114

Tabela 27: Parâmetros e acurácias máximas, sem N_sil e Dur 115

Tabela 28: Médias dos parâmetros para as UDs conhecidas e para os grupos propostos..... 120

LISTA DE ABREVIATURAS E SIGLAS

ALL Alocutivo
APC Apêndice de Comentário
APT Apêndice de Tópico
AUX Auxílio Dialógico
CMM Comentários Múltiplos
COB Comentários Ligados
COM Comentário
CNT Conativo
DCT Conector Discursivo
EXP Expressivo
f0 Freqüência Fundamental
INT Introdutor Locutivo
INP Incipitário
INT Introdutor Locutivo
MD Marcador Discursivo
PAR Parentético
PB Português do Brasil
PHA Fático
SCA Unidade de Escansão
TMT Tomada de Tempo
TOP Tópico
UI Unidade Informacional
UD Unidade Dialógica
UFMG Universidade Federal de Minas Gerais
UT Unidade Tonal

LISTA DE SÍMBOLOS

(*) início de turno

(ABC) identificação do falante

(_r) discurso reportado

(//) quebra entonacional terminal; fim de enunciado

(/) quebra entonacional não-terminal; fim de unidade tonal interna ao enunciado

(+) enunciado interrompido

(< >) sobreposição de fala

([/nº]) *retracting* ou falha na execução do enunciado

(&) início de palavra interrompida

(&he) hesitação ou silêncio preenchido

(“ ”) citação

(hhh) comportamento paralingüístico

SUMÁRIO

1 INTRODUÇÃO 18

2 OS MARCADORES DISCURSIVOS 19

3 A TEORIA DA LÍNGUA EM ATO 21

3.1 As unidades informacionais 22

4 AS UNIDADES DIALÓGICAS 26

4.1 Alocutivo (ALL) 28

4.2 Conativo (CNT) 30

4.3 Incipitário (INP) 32

4.4 Fático (PHA) 34

4.5 Expressivo (EXP) 36

4.6 Conector Discursivo (DCT) 37

5 METODOLOGIA DE ESTIMATIVA DOS PARÂMETROS PROSÓDICOS 39

5.1 A base dos dados 39

5.1.1 O *Corpus* C-ORAL-BRASIL 39

5.1.2 O Mini BR Informal 42

5.2 Revisão do Mini BR Informal 43

5.3 Tratamento dos dados 52

5.3.1 Preparação dos dados 52

5.3.2 Alinhamento e segmentação 54

5.3.3 Referência para as medidas da fala 56

5.4 Estimativa dos parâmetros prosódicos 59

5.4.1 Parâmetros de duração 59

5.4.2 Parâmetros de intensidade 64

5.4.3 Parâmetros de f0 66

5.4.4 Parâmetros de variação de f0 71

5.4.5 Parâmetros de alinhamento 73
6 ESTATÍSTICA DESCRITIVA E INFERENCIAL DOS PARÂMETROS ESTIMADOS 76
6.1 Parâmetros de duração 79
6.2 Parâmetros de intensidade 82
6.3 Parâmetros de f0 86
6.4 Parâmetros de variação de f0 89
6.5 Parâmetros de alinhamento 92
7 ANÁLISE DOS DADOS 101
7.1 Triagem de parâmetros 101
7.2 Análise das UDs classificadas 107
7.3 Análise das UDs não classificadas 116
8 CONSIDERAÇÕES FINAIS 125
REFERÊNCIAS 131
APÊNDICES 136

1 INTRODUÇÃO

Os marcadores discursivos (MDs) são uma categoria linguística muito problemática, apesar de muitos estudos dedicados a eles ao longo das últimas quatro décadas. Não há na literatura consenso sobre a terminologia empregada, sobre a definição da categoria e sobre como identificar os MDs e suas funções.

Uma proposta para estudo dos MDs é apresentada pela Teoria da Língua em Ato (L-AcT) (CRESTI, 2000; MONEGLIA, 2005; RASO, 2012; MONEGLIA e RASO, 2014), elaborada através da observação de corpora de fala espontânea. A L-AcT adota a visão de que a veiculação de funções comunicativas é feita através da prosódia, e não do item lexical, como é comumente defendido em outras propostas sobre os MDs. Sob esta abordagem teórica, a prosódia permite não somente prever quando um item lexical se comporta como MD como também caracterizar a função específica por ele desempenhada.

Este trabalho busca contribuir para a compreensão dos MDs, sob a abordagem proposta pela L-AcT, através de uma análise prosódica e estatística sobre MDs na fala espontânea do português brasileiro, com base no *corpus* C-ORAL-BRASIL (Raso e Mello 2012). Para tanto, buscaremos atingir os seguintes objetivos:

I. Validar através da modelagem estatística a forma dos MDs já descritos em seus parâmetros prosódicos;

II. Identificar grupos com padrões prosódicos semelhantes entres os marcadores que ainda não foram descritos prosodicamente.

2 OS MARCADORES DISCURSIVOS

Marcadores discursivos podem ser descritos como itens ou sequências lexicais que não contribuem para o conteúdo proposicional do enunciado e não alteram seu valor de verdade. (AIJMER, 2009; FOOLEN, 1996). Essa definição negativa é um dos poucos consensos sobre o tema na literatura. Para Follen (1996), uma definição clara e positiva não é fácil de encontrar, tendo em vista a grande variedade de funções que podem ser atribuídas aos MDs.

O termo marcador discursivo, inaugurado por Schiffrin (1987), parece ser o mais empregado, porém longe da unanimidade. É bastante comum que análises teóricas sobre o tema se iniciem por reflexões acerca da terminologia. Fraser (2006) menciona 15 termos utilizados, elegendo como mais frequentes *discourse markers*, *discourse operators*, *discourse connectors* e *cue phrases*. Para Follen (1996), que identifica 26 termos diferentes, as variações de terminologia refletem variações de abordagem e de ponto de vista que dificultam a comparação e a avaliação dos diferentes resultados de pesquisas.

Por outro lado, também não há consenso sobre as várias funções que os MDs podem desempenhar. De modo geral, é possível distinguir entres as funções dedicadas à estruturação textual do discurso e as funções interacionais. Quando as funções interacionais são destacadas, o termo marcador pragmático é preferido por alguns autores, embora seja mais comum utilizar este termo como mais abrangente, abarcando formas com várias funções textuais e interacionais (AIJMER 2009; FRASER 2006; FOOLEN, 1996; HANSEN, 2006).

Brinton (1996) estabelece uma lista de características dos MDs baseada em critérios lexicais, fonológicos, sintáticos, semânticos, funcionais, sociolinguísticos e estilísticos:

1. São curtos e foneticamente reduzidos;
2. Formam um grupo tonal distinto;
3. São formas marginais e por isso difíceis de classificar em classes de palavras tradicionais;
4. São restritos à posição inicial da sentença;
5. Ocorrem fora da estrutura sintática ou são apenas vagamente ligados a ela;
6. São opcionais;
7. Possuem pouco ou nenhum significado proposicional;

8. São multifuncionais, operando em vários níveis do discurso simultaneamente;
9. São mais característicos do discurso oral do que do escrito e são associados à informalidade;
10. Ocorrem com alta frequência;
11. São estilisticamente estigmatizados;
12. São mais frequentes na fala feminina.

Algumas das características listadas são consensuais para a maioria dos autores, particularmente (5), (6), (7), (9) e (10). Apesar do consenso, observamos que algumas dessas características contém redundâncias e pouco contribuem para uma compreensão da categoria. Há discussão sobre a posição (4); exatamente por ser desconectado da estrutura sintática, há exemplos de marcadores discursivo em qualquer posição. A multifuncionalidade dos MDs (8) ocorre se considerarmos que um item lexical pode desempenhar diferentes funções, porém o ponto deve ser completamente reavaliado na abordagem não lexical que será adotada neste estudo.

Sobre os critérios sociolinguísticos (11 e 12), a nosso ver, os MD são típicos da fala em geral, sendo a escolha lexical dependente de fatores de gênero, idade ou outra variável sociolinguística. Ainda que alguns itens lexicais sejam estigmatizados, isso não caracteriza os MDs.

A realização dos MDs dentro de uma unidade tonal distinta (2) é uma característica pouco explorada. Segundo Aijmer (2009), as características prosódicas, importantes para distinguir os diferentes usos dos MDs, têm sido amplamente negligenciadas e estudadas sistematicamente apenas recentemente.

Uma abordagem teórica em que a prosódia tem papel fundamental na identificação dos MDs é a oferecida pela L-AcT, e será utilizada como base para a realização deste trabalho.

3 A TEORIA DA LÍNGUA EM ATO

A análise de marcadores discursivos proposta neste trabalho será desenvolvida com base na Teoria da Língua em Ato (L-AcT) (CRESTI, 2000; MONEGLIA, 2005; RASO, 2012; MONEGLIA e RASO, 2014). Elaborada através da profunda observação de dados de corpora da fala espontânea, a L-AcT destaca o papel determinante da prosódia na organização da fala.

O fluxo da fala é naturalmente dividido em unidades prosódicas discretas denominadas unidades tonais (UTs). Essas unidades são delimitadas através de certas alterações prosódicas perceptualmente relevantes para o ouvinte. As rupturas são denominadas quebras ou fronteiras prosódicas, e podem ser percebidas como conclusivas ou como não conclusivas (MONEGLIA, 2005).

As quebras prosódicas conclusivas, denominadas terminais, têm alta saliência perceptual e segmentam a fala em enunciados (KARCEVSKY, 1931; CRYSTAL, 1975). Conforme a tradição pragmática fundada por Austin (1962), o enunciado é a unidade de referência da fala e, segundo a definição aprimorada por Cresti (2000), é a mínima entidade linguística que pode ser interpretada pragmática e prosodicamente em isolamento, sendo autônoma comunicativamente. O enunciado corresponde ao ato de fala, ou seja, através de sua realização uma ação linguística é cumprida.

Os enunciados podem ser simples, quando contêm uma única unidade tonal, ou complexos, quando há duas ou mais unidades tonais, separadas por quebras prosódicas não terminais (percebidas como não conclusivas). O exemplo abaixo mostra quatro enunciados, separados por quebras terminais indicadas por dupla barra (//):

Exemplo 1: bfamdl04[243-246]:

```
*SIL: [243] porque enquanto descansa / carrega pedra // [244] cê sabia
disso //
*KAT: [245] carrega o quê //
*SIL: [246] carrega pedra //
```





No exemplo, o enunciado 243 é um enunciado complexo, realizado por duas unidades tonais distintas, separadas por uma quebra não terminal (/). Ao ouvirmos isoladamente, a primeira unidade não pode ser interpretada como completa, pois há um sinal de continuidade. Já a segunda unidade é interpretável pragmática e prosodicamente em isolamento. O mesmo falante repete o conteúdo no enunciado 246, dessa vez como um enunciado simples. Em ambos os casos essa unidade prosódica, por sua função informacional, é denominada Comentário. O Comentário (COM) carrega a força ilocucionária e, por essa razão, confere autonomia pragmática e prosódica ao enunciado.

O comentário é necessário e suficiente para veicular a ilocução, portanto está presente em todos os enunciados, sozinho ou acompanhado de outras unidades. Sua forma prosódica é variável dependendo do ato de fala que ele veicula[1] e de traços cognitivos e pragmáticos, e não de fatores léxico-sintáticos, como proposto por Searle (1969). Não há requerimentos sintáticos ou lexicais para o conteúdo do comentário, ele pode ser realizado por sentenças completas, sintagmas ou até uma única palavra de qualquer classe (uma interjeição, um nome próprio, uma negação, etc). Além disso, o mesmo conteúdo locutivo pode ter valor comunicativo diferente dependendo da ilocução expressa pela prosódia (RASO, 2012, p. 121).

3.1 As unidades informacionais

Além do comentário, necessariamente presente no enunciado, pode haver uma ou mais unidades prosódicas que, mesmo sendo opcionais, possuem valor informacional. De modo geral, cada unidade prosódica corresponde a uma unidade informacional, caracterizada pelo perfil melódico, pela sua distribuição no enunciado e pela função que executa dentro do enunciado. A sequência de unidades informacionais constitui o padrão informacional do enunciado, encerrado pela quebra terminal.

Cresti (2000) descreveu as unidades informacionais agrupando-as em duas categorias: as unidades textuais, que compõem o texto do enunciado; e as unidades

[1] Uma ordem, uma saudação, uma pergunta, etc

dialógicas, que não compõem o texto e se dirigem ao interlocutor, ou à interação dialógica. Estas últimas serão discutidas na próxima seção.

Além do comentário (COM), são unidades informacionais textuais:

- Tópico (TOP): define o âmbito de aplicação da força ilocucionária. Está sempre posicionado à esquerda do COM. É a única unidade, além do COM, que possui um foco funcional ou um núcleo prosódico, localizado sempre à direita da unidade.

- Apêndices de comentário (APC) e de tópico (APT): integram textualmente as unidades de COM e TOP, respectivamente, sempre à direita delas. O perfil prosódico do APC é nivelado ou descendente, o APT pode ter movimento ascendente, porém sem foco funcional;

- Introdutor locutivo (INT): introduz uma meta-ilocução: discurso reportado, instruções, exemplificações emblemáticas e pensamentos narrados. Possui um perfil descendente finalizado com f0 muito mais baixa em claro contraste com a f0 da unidade seguinte. A velocidade de elocução é muito mais rápida que o resto do enunciado.

- Parentético (PAR): tem função metalinguística, fornecendo instruções sobre como o enunciado deve ser interpretado. Pode ocorrer em qualquer posição, exceto no início absoluto do enunciado. O perfil prosódico é nivelado, a taxa de elocução é maior e o valor de f0 é diferente do enunciado, quase sempre mais baixo.

Além do enunciado complexo do Exemplo 1, exemplos de tópico podem ser observados nos enunciados a seguir. O Exemplo 2 contém um introdutor locutivo e o Exemplo 3 ilustra um parentético e um apêndice de comentário.

Exemplo 2: bfamdl04[153]

```
*SIL: [153] eu acho assim /INT se a pessoa nũ tem condições de fazer /TOP
ele paga pra fazer //COM
```

[153]

Exemplo 3: bfammn02[088]:

```
*DFL: [091] <e isso> /TOP assim /PAR era muito importante /COM pra ela
//APC
```

[091]

Nem sempre a correspondência entre perfil prosódico e padrão informacional é biunívoca (isomorfia). Por vezes uma unidade informacional textual pode estar dividida em duas ou mais unidades tonais, geralmente por veicular um conteúdo muito grande ou por imperícia do falante, ênfase ou hesitação. Nesse caso, a última unidade tonal carrega a marca prosódica da unidade informacional, enquanto a(s) anterior(es) possuem perfil neutro. Essas unidades tonais neutras são denominadas unidades de escansão (SCA).

Há também casos de quebra de isomorfia entre enunciado e ilocução, ou seja, uma mesma unidade terminada pode veicular mais de uma ilocução, o que ocorre em duas situações:

- Comentários Múltiplos (CMM): comentários separados por quebra não terminal, mas organizados em um padrão prosódico fortemente convencionalizado. Ainda que cada unidade possua força ilocucionária, o CMM é interpretado como um todo. Exemplos de CMM são listas, pedido de confirmação e relações de natureza lógica.

- Comentários Ligados (COB): comentários justapostos ligados por um perfil prosódico de continuidade, formando unidades terminadas complexas denominadas estrofes (*stanzas*). São típicos da tipologia monológica (embora possam ocorrer em qualquer texto) visto que a interatividade é reduzida e a expressão do pensamento em curso se manifesta durante a fala. Dentro dessas unidades complexas é possível identificar subpadrões estruturais a partir de cada uma das unidades ilocucionárias.

O Exemplo 4 mostra dois enunciados adjacentes, um simples e outro composto por 3 unidades de comentários formando um comentário múltiplo. No Exemplo 5, podemos observar uma lista com sete comentários ligados e três ocorrências de unidades de escanção (SCA).

Exemplo 4: bfammn04[018-019]

```
*REG: [18] porque tinha um neném que já tinha nascido //COM [19] eles
lavaram o João /CMM puseram roupa /CMM e no bercinho //CMM
```



Exemplo 5: bfammn05[140]

*CAR: [140]porque ele ama /COB ele chama ela de /SCA fuminho /COB né /COB meu pretinho /COB papai nũ güenta carregar mais /COB mas /DCT nũ güenta pegar mais /COB porque tá muito grande /COB mas é [/1]SCA é essa é a história /SCA e é a vida /SCA que nós temos aqui em casa //COM$



[140]

4 AS UNIDADES DIALÓGICAS

As unidades dialógicas (UDs) desempenham funções pragmáticas necessárias para regular a interação dialógica, ou seja, são dirigidas ao interlocutor. Distinguem-se das unidades textuais por não serem composicionais sintaticamente com o resto do enunciado nem contribuírem para a interpretação de seu significado. Por essas características, as unidades informacionais dialógicas correspondem aos MDs de acordo com parte das teorias apesentadas anteriormente.

Entretanto, diferentemente de outras abordagens, pela L-AcT é a prosódia que desempenha o papel determinante na identificação da função dialógica, e não o item lexical. Pela teoria, não é possível prever a função pragmática de um item lexical isoladamente, sem conhecer o contexto prosódico. Consequentemente não há categorização prévia de qualquer item lexical como MD (RASO, 2014). Um item ou expressão lexical pode aparecer em três possíveis contextos:

> Pode realizar uma função ilocucionária, sendo assim pragmaticamente e prosodicamente autônomo, permitindo interpretabilidade em isolamento; pode aparecer dentro de uma unidade tonal/informacional, sendo deste modo semanticamente e sintaticamente composicional com o restante do conteúdo da unidade, porém não interpretável isoladamente; pode preencher uma unidade tonal, embora sem realizar uma função ilocucionária e portanto não ser autônomo e interpretável em isolamento (RASO e VIEIRA, 2016).[2]

Apenas no terceiro contexto o item ou expressão seria uma UD. Os três exemplos a seguir mostram enunciados de um diálogo no qual o item *não* aparece nos três contextos:

Exemplo 6: bpubdl01[014-015]

*PAU: **não** // tá dando a altura / daquele que a <Isa> marcou <lá>/ né //



[014-015]

[2] Tradução nossa para: *it can fulfill an illocutionary function, therefore being pragmatically and prosodically autonomous, allowing its interpretability in isolation; it can appear inside a tone/information unit, being therefore semantically and syntactically compositional with the rest of unit's content, but not interpetable in isolation; it can fulfill a tone/information unit, though not performing an illocutionary function and therefore not being autonomous and interpretable in isolation*.

Exemplo 7: bpubdl01[119]

*PAU: ah / **não** acaba **não** / acaba //



[119]

Exemplo 8: bpubdl01[197]

*PAU: ah / **não** / ea disse que é pa ficar / por algum tempo //



[197]

No Exemplo 6, *não* realiza uma ilocução e pode ser interpretado isoladamente. O material que o segue não é o escopo da negação, o falante não nega que *tá dando altura*. No Exemplo 7, *não* pertence à mesma unidade tonal do verbo e é composicional com ele, negando-o. No Exemplo 8, *não* está em uma unidade tonal isolada sem autonomia prosódica, não pode ser interpretado isoladamente e não nega nada (não é composicional nem constitui uma ilocução), sendo, portanto, uma UD.

Estudos no *corpus* C-ORAL-ROM italiano, inicialmente por Cresti (2000) e em seguida por Frosali (2006) levaram à proposta de seis unidades informacionais dialógicas: incipitário, alocutivo, conativo, fático, expressivo e conector discursivo. Essas unidades foram propostas a partir de critérios entonacionais, funcionais e distribucionais, além de características lexicais.

Observou-se também que as UDs ocorrem com frequência na fala espontânea. Frosali (2006, p. 418) verificou, no *corpus* C-ORAL-ROM italiano, a presença dessas unidades em 27% de todos os enunciados e em 50% dos enunciados complexos (único tipo de enunciado que permite UD). No português, a anotação informacional de uma amostra do C-ORAL-BRASIL (Mini BR Informal[3]) (MELLO e RASO, 2012, p. 194) identificou similarmente 27% dos enunciados totais com UD, porém considerando-se apenas os enunciados complexos, a proporção chega a 65%.

Raso (2014) aplica a Teoria da Língua em Ato para identificar e classificar as UDs na fala espontânea do português brasileiro a partir de dados do Mini BR Informal. Esse estudo identifica as unidades lexicais que preenchem as unidades informacionais dialógicas e realiza uma análise comparada ao *corpus* italiano equivalente.

[3] Para mais detalhes, ver seção 5.1.2.

Posteriormente, Raso e Vieira (2016) refinam a descrição prosódica de quatro UDs e apresentam uma proposta de investigação estatística dos parâmetros fonéticos. Nesse estudo, três unidades foram satisfatoriamente descritas: o incipitário, o alocutivo e o conativo.

A seguir serão apresentados a descrição mais detalhada das seis UDs conforme a tipologia proposta por Cresti (2000) e Frosali (2006), o resultado do refinamento da descrição de incipitário, alocutivo e conativo e as questões e dificuldades encontradas nas unidades fático, expressivo e conector discursivo (RASO e VIEIRA 2016).

Serão utilizados dados e exemplos do Mini BR Informal obtidos de Raso (2014) e Raso e Vieira (2016) como também dados de distribuição e preenchimento lexical para cada unidade obtidos da anotação original do Mini BR Informal[4] (anterior à revisão que será apresentada neste estudo).

4.1 Alocutivo (ALL)

O alocutivo realiza a função de identificar o interlocutor e marcar coesão social. É uma UD associada a uma tipologia lexical: sempre é realizada por nomes próprios, títulos e epítetos. O perfil melódico é nivelado ou descendente em posição final no enunciado, e a intensidade é baixa.

Apesar de a UD ser sempre preenchida lexicalmente por nomes próprios, títulos e epítetos, não é verdade que estes sempre desempenhem a função de alocutivos. Os mesmos itens podem ser usados para veicular outro tipo de UD (ver Exemplo 17) ou mesmo uma ilocução, como em um chamamento. Raso (2014) mostra como a categoria vocativo não encontra evidência empírica de existência na fala. Nomes próprios tradicionalmente enquadrados nessa categoria se apresentam como unidades informacionalmente muito distintas, ora como ilocuções de chamamento, ora como UDs, tanto do tipo ALL como CNT.

As figuras abaixo mostram exemplo de alocutivo:

[4] As descrições dos perfis prosódicos das unidades dialógicas ainda estão em aprimoramento, como mostra o estudo de RASO e VIEIRA (2016). Assim todos os dados quantitativos são sujeitos a revisão.

Exemplo 9: bfamdl04[072]

*KAT: que que a sio' falou /=COM= **mãe //=ALL=**

[072]

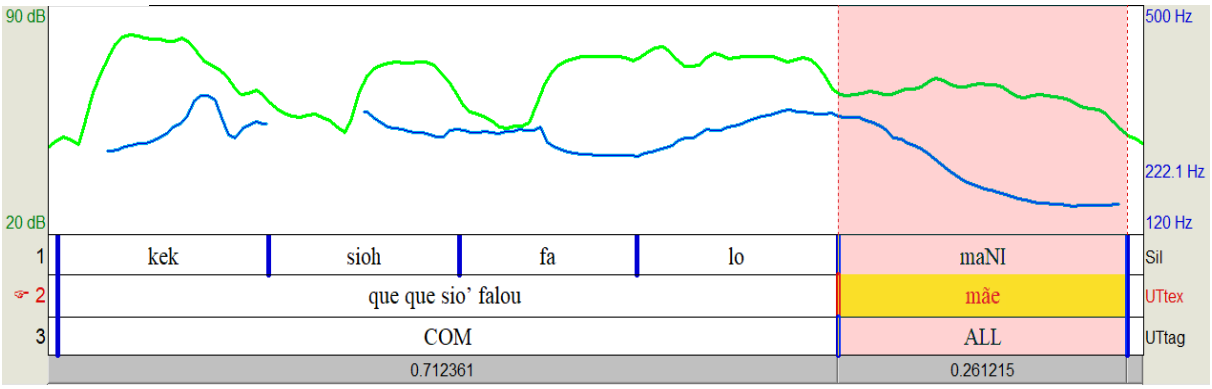


Figura 1: ALL com função de marcar coesão social (verde: intensidade; azul: f0)

Apesar de a maior parte dos alocutivos apresentar um perfil descendente no final do enunciado, alguns casos levaram Raso e Vieira (2016) a propor a hipótese de que o ALL na verdade tenha perfil nivelado e baixo, que se ajusta[5] às unidades tonais vizinhas nas regiões próximas às fronteiras, como no Exemplo 14. Nos casos em que a unidade é a última na estrutura informacional, este perfil se torna descendente pela ausência de uma unidade tonal seguinte (como no Exemplo 13).

Exemplo 10: bfamcv03[247]

*TON: é ocê /=COM= **Onofre /=ALL=** que tá cagando assim //=APC

[247]

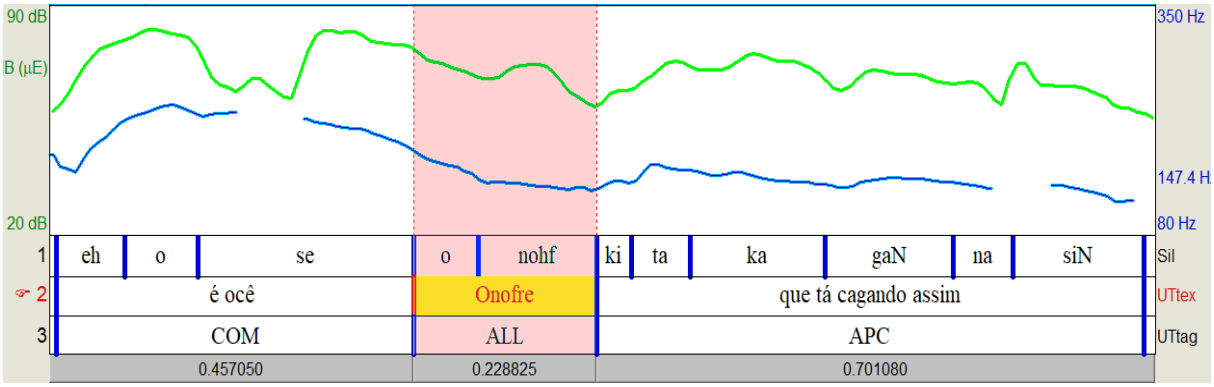


Figura 2: ALL em posição interna. Em azul, o perfil tonal ajustado às unidades vizinhas (verde: intensidade; azul: f0)

[5] Continua o perfil melódico anterior e prepara o posterior;

O alocutivo supostamente não possui posição determinada no enunciado, porém os dados mostram preferência muito forte pela posição final no português brasileiro. Além disso, nenhuma ocorrência de ALL foi encontrada no início do enunciado, sugerindo a hipótese de que esta posição não seja possível. Os nomes próprios predominam no preenchimento lexical do alocutivo. Outros itens encontrados no Mini BR Informal são: sô; meu filho/minha filha; cara; velho; mãe; filho/filha/filhote; bicho; dona; gente; eh meu pai; boba; homem; meu; moça[6].

4.2 Conativo (CNT)

Unidade dialógica com função de encorajar o ouvinte a se engajar na interação dialógica, ou tenta interromper um comportamento não colaborativo. (CRESTI, 2018). O perfil é descendente a partir da tônica. A duração é curta e a intensidade, mais alta que a do ALL, além de não decair imediatamente após o início da UD. Três exemplos são mostrados a seguir:

Exemplo 11: bfamdl02[038]

*BAL: **tá vendo /=CNT=** a setinha tem que tar no cento-e-dez //=COM=

[038]

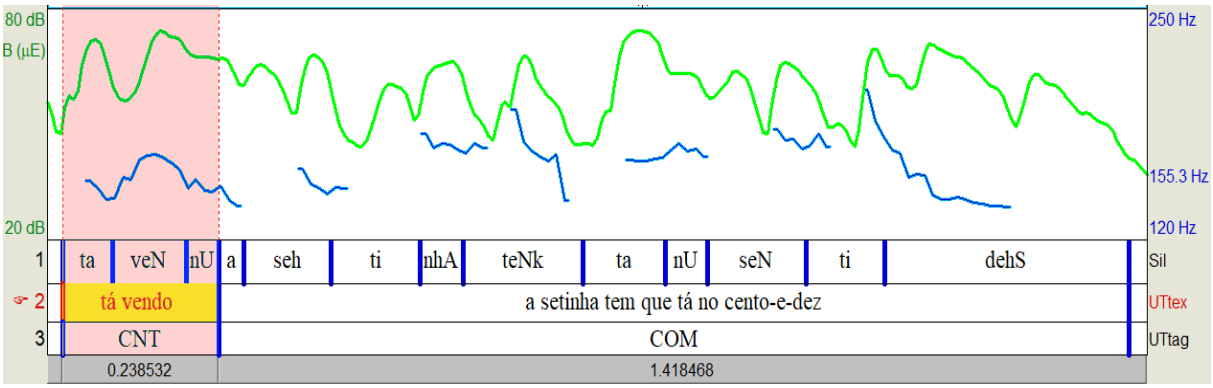


Figura 3: Exemplo de CNT com perfil modulado-descendente vizinhas (verde: intensidade; azul: f0)

[6] A lista atualizada com itens lexicais após a revisão encontra-se na Tabela 8.

Exemplo 12: bfamcv04[241]

*HEL: **não /=CNT= pera aí /=CNT=** eu nũ entendi <o> “um” //=COM=

[241]

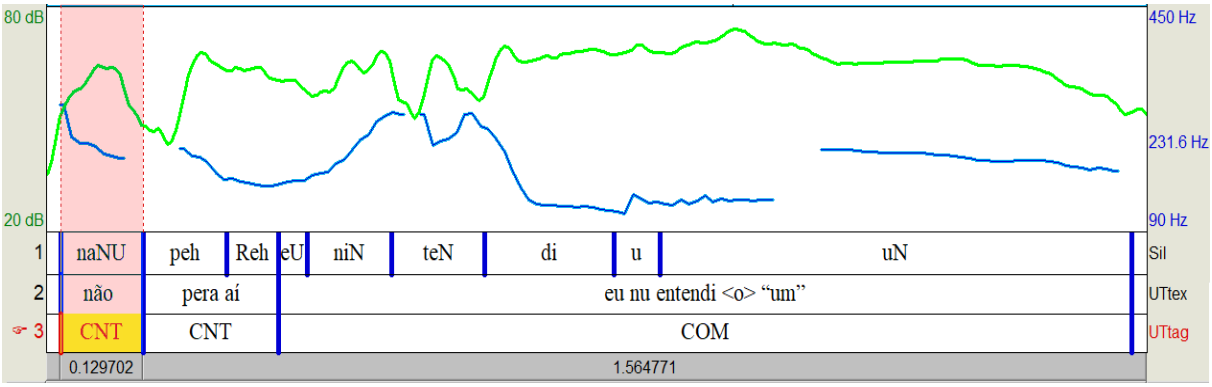


Figura 4: Exemplo de enunciado com dois CNTs consecutivos de perfil descendente vizinhas (verde: intensidade; azul: f0)

Exemplo 13: bfamcv03[063]

*HEL: não /=CMM= sô /=AUX= o cinco /=CMM= **Renato //=CNT=**

[063]

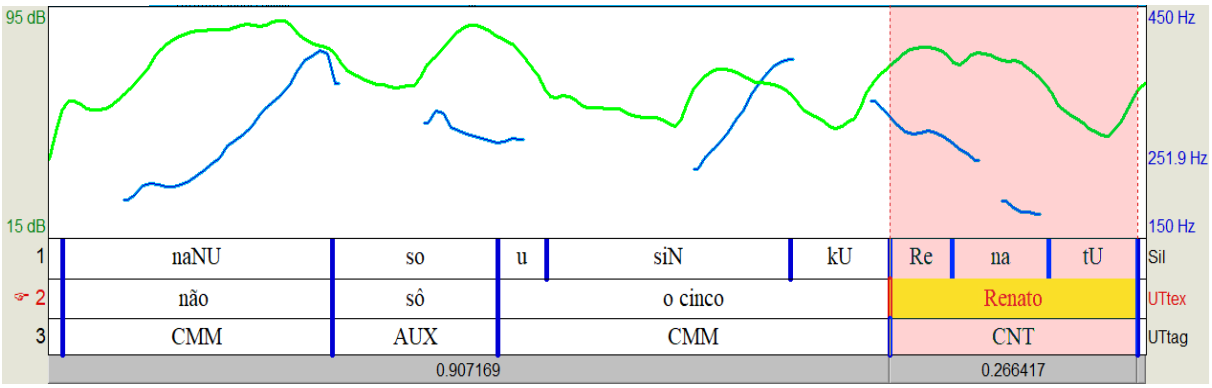


Figura 5: Exemplo de dúvida entre CNT e ALL vizinhas (verde: intensidade; azul: f0)

Os dois primeiros exemplos se encaixam bem na descrição. Já o exemplo da figura 9 é um caso de conativo que pode ser confundido com alocutivo, por ter perfil descendente e ser preenchido por um nome próprio.

Em Raso e Vieira (2016) as diferenças entre CNT e ALL são investigadas melhor. A elevação da f0 em relação ao fim da unidade anterior não é característica de alocutivo. Outras diferenças são encontradas: o CNT tem o mesmo perfil em qualquer posição, diferentemente do ALL, como já descrito acima: mesmo em posição final, onde a semelhança é maior, o movimento descendente no ALL começa no início da palavra e termina nivelado, enquanto que no CNT ele ocorre a partir da sílaba tônica e

decresce até o final da unidade, com inclinação maior. Por fim, o CNT é comum em posição inicial, onde nenhum ALL foi observado.

De acordo com a anotação do Mini BR Informal, o conativo tem distribuição livre, com preferência para o início ou final dos enunciados. Na maior parte dos casos o conativo é realizado por um verbo, sendo *olha* (ou as variações *o'* e *olha p'cê vê*) o mais frequente. Outras ocorrências encontradas são: *pera aí/pera pera aí; ô; não; tá vendo; por favor; quer ver; aí; aqui; bem; bom; só um segundo; cê vê; então; hhh tá; quer dizer; cê sabe; uai; viu; xá eu ver*[7].

As próximas três unidades que serão comentadas têm como referência apenas os trabalhos de Cresti (2000), Frosali (2008) e Raso 2014. Entendemos que ainda são necessários mais estudos para confirmar a existência dessas categorias e descrever suas formas.

4.3 Incipitário (INP)

Sua função é sinalizar tomada de turno ou início de enunciado e de subpadrão de *stanza*, marcando contraste. É caracterizado pela alta intensidade, curta duração e elevado valor de f0 em relação à amplitude de variação de f0 no restante do enunciado. O perfil pode ser ascendente, descendente ou ascendente-descendente, sempre alcançando um pico de alta frequência.

De acordo com Raso (2014), o item lexical que é mais utilizado no Mini BR Informal para a função de incipitário é o advérbio *não*. Outros itens que preenchem a unidade são: *é*, *ah*, *então*, *bom*, *aí*, *agora*, *eu*, *mas*, *porque*, *que eu*, *se bem*, *também*, *tá*[8]. A distribuição do INP não é livre; pela própria função, essa unidade aparece no início dos enunciados. Foram encontradas 100 ocorrências (96%) em posição inicial e 4 casos em posição interna, em início de subpadrão de estrofe. A revisão realizada neste estudo, que será apresentada na seção 5.2, resultou em uma redução no número de ocorrências de INP no *Mini BR Informal* de 102 para 40[9].

---

[7] A lista atualizada com itens lexicais após a revisão encontra-se na Tabela 8.

[8] A lista atualizada com itens lexicais após a revisão encontra-se na Tabela 8.

[9] Ver capítulo 5.2

As variações no perfil melódico do incipitário foram estudadas em detalhe em Raso e Vieira (2016). De acordo com o estudo, o que é realmente necessário para seu perfil melódico é atingir um pico de frequência em relação à ilocução, sendo as variações da curva consequência de efeitos micromelódicos das consoantes sobre a vogal tônica:

i) se o INP termina em ditongo ou possui mais sílabas após a tônica, a frequência decresce após a tônica (perfil descendente);

ii) se a vogal tônica é precedida por consoante obstruinte não vozeada e não há sílabas antes da tônica, a frequência se mantém alta desde o início (perfil plano);

iii) se a sílaba tônica começa com consoante vozeada seguida de vogal, a frequência começa mais baixa e ascende para o pico na vogal tônica (perfil ascendente).

iv) a combinação dos contextos (i) e (ii) produz um perfil nivelado-descendente;

v) a combinação de (i) e (iii) gera o perfil ascendente-descendente.

Os exemplos abaixo ilustram algumas das possibilidades de perfil tonal para o incipitário:

Exemplo 14: bfammn06[049]

*JOR: **bom /=INP=** e eu fiquei nesse [/1]=SCA= nessa empresa durante algum tempo /=COB=

[049]

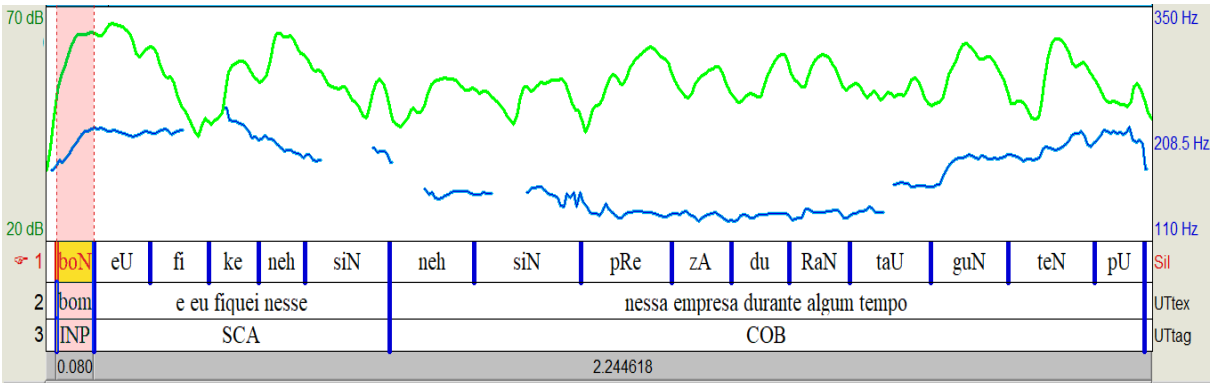


Figura 6: INP com perfil ascendente: inicia com consoante vozeada (iii) vizinhas (verde: intensidade; azul: f0)

Exemplo 15: ifamcv06[032][10]

*ILA: **poi /=INP=** &he /=TMT= alla FLOG /=TOP cioè /=PHA= non si sente mai nulla //=COM

[032]

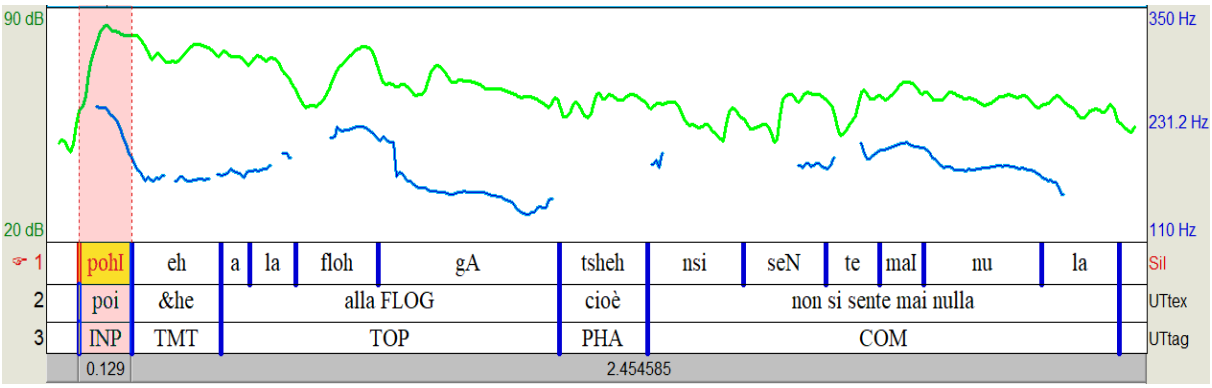


Figura 7: INP com perfil descendente: inicia com obstruinte desvozeada e finaliza em ditongo (i) vizinhas (verde: intensidade; azul: f0)

Os efeitos microprosódicos sobre o nível suprassegmental parecem explicar satisfatoriamente as variações de perfil do INP nos exemplos escolhidos. Além disso, o exemplo em italiano sugere que a descrição proposta seria aplicável não apenas em português.

4.4 Fático (PHA)

Essa unidade tem a função de sinalizar ao interlocutor que o canal de comunicação está aberto. É caracterizado por um perfil prosódico nivelado ou descendente, com baixa intensidade e duração muito curta. A realização fonética é normalmente reduzida.

O Exemplo 16 mostra um enunciado com duas UDs realizadas com o item lexical *né*:

[10] Exemplo retirado do minicorpus italiano do projeto C-ORAL-ROM. Essa configuração não foi encontrada no PB.

Exemplo 16: bfammn04[029]

*REG: e aí /=INP= &n [/1]=EMP= nem anestesia /=CMM= nem nada /=CMM= **né** **/=PHA=** porque a gente também nũ /=SCA= teve [/2]=EMP= &n [/1]=EMP= nũ teve que dar ponto /=CMM= nem <nada /=CMM= **né>** **//=PHA=**

[029]

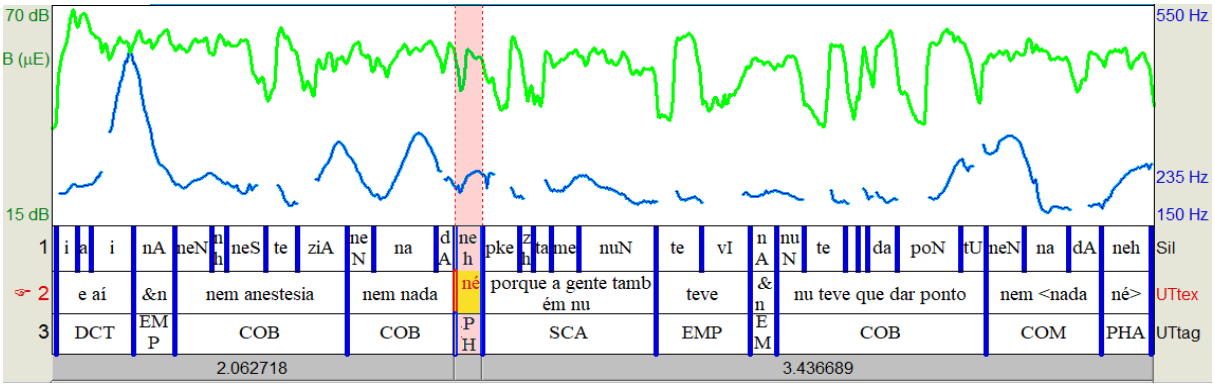


Figura 8: Enunciado com duas UDs PHA (né) (verde: intensidade; azul: f0)

De acordo com os dados da anotação do Mini BR Informal em sua publicação, o fático seria a mais frequente das UDs do português brasileiro, com 456 ocorrências (42% do total de 1089 UDs). O estudo de preenchimento lexical mostraria um elevado grau de concentração de realizações de PHA com a contração *né*. Outros itens encontrados foram: *uai; não/mas não; é/é não; ué; hein; assim; viu; então; agora; aí/aí não; sabe; ah; porque; uhn; eh; bom; tá; entendeu; o'; o; ahn; aqui; hhh; No'; que; sô; quer dizer; mas; também.*

No Exemplo 16, apesar de conterem o mesmo item lexical, as duas unidades PHA mostram perfil melódicos diferentes, o que não é interessante para a teoria. Raso (2014, p. 440) indica que o perfil ascendente do fático ocorre em posição final, uma possibilidade que deve ser verificada e validada.

No Exemplo 16, o segundo PHA pode ser percebido como uma ilocução de pedido de confirmação, que também é realizada frequentemente por 'né'. Ouvindo isoladamente, parece que há margem para dúvida entre UD e ilocução (existência de autonomia prosódica e pragmática)

A quantidade significativa de dúvidas na anotação de fáticos é reconhecida em Raso e Vieira (2016), sendo este o motivo para exclusão da unidade do estudo de detalhamento da descrição fonética.

4.5 Expressivo (EXP)

Segundo Frosali (2006) e Raso (2014), a UD expressivo (EXP) funciona como suporte emocional à ilocução e como meio de demonstrar coesão social. Frosali (2006, p. 420) indica que o expressivo é caracterizado por um movimento ascendente de f0, duração abaixo da média do enunciado, intensidade média e realização fonética reduzida. Entretanto, no Mini BR Informal, não é clara uma identidade prosódica para esta unidade. A seguir dois exemplos com perfis distintos, como pode ser visto nas figuras abaixo:

Exemplo 17: bfamdl01[167]

*REN: **ah /=EXP=** eu gosto desses eucalipto normal /=CMM= mas //=CMM=

[167]

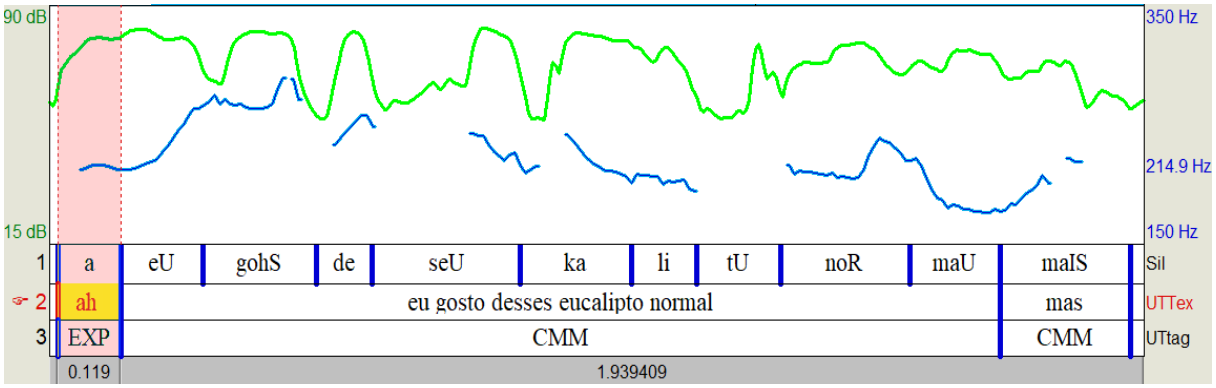


Figura 9: EXP com perfil nivelado

Exemplo 18: bpubmn01[077]

SHE: né /=COB= pa falar assim /=INT= Nossa /=EXP= mas isso aqui /=SCA= que cê tá fazendo é uma besteira muito grande /=COB= né //=PHA=

[077]

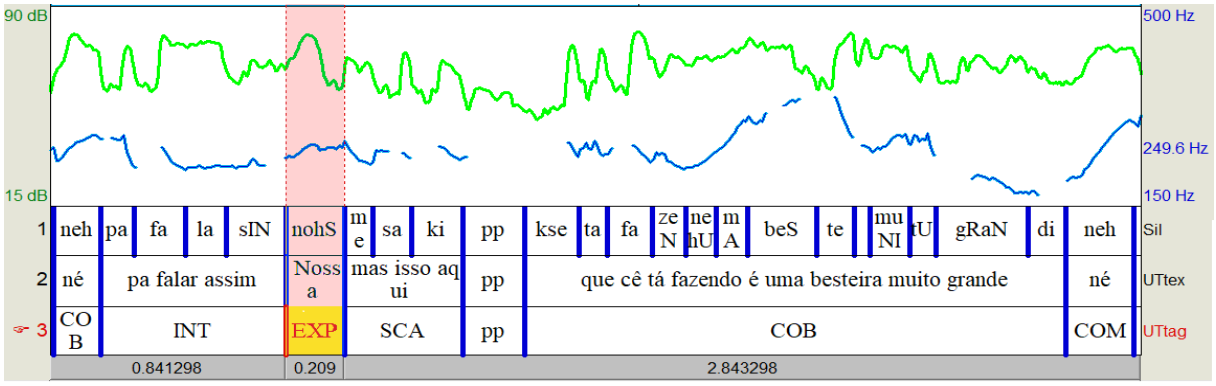


Figura 10: EXP com perfil ascendente

Assim como para o fático, a ocorrência de unidades tonais com características tão diferentes veiculando uma mesma unidade informacional entra em contradição com a base teórica proposta. Para Raso (2014, p 440), a complexidade do expressivo pode ser devida à interferência de propriedades prosódicas de diferentes tipos de ilocução aos quais o expressivo confere suporte emocional. É ainda possível a existência de subcategorias ou até mesmo a existência de um novo tipo de UD. Também é sugerido que as unidades anotadas como EXP possuam no português brasileiro uma terceira função, a de abertura de turno, sem contraste.

Novamente, a análise do Mini BR Informal indica um item lexical predominante: a interjeição *ah*. Outros casos de expressivo identificados foram: *Nossa/No'/Nossa Senhora; gente/ô gente; uai; hein; ô; não/não não; sô; oh; ai; pois é; né; é; eh; ah não; cara; então o'; menino; mas; o'; pô; pelo amor de Deus; porra; sabe; uhn; viu*;

4.6 Conector Discursivo (DCT)

Possui a função de conectar e marcar continuidade entre dois enunciados ou dois subpadrões de estrofe. O preenchimento lexical é limitado a algumas classes de palavras: conjunções, advérbios e talvez algumas preposições. Seu perfil prosódico é variado, a intensidade é alta e a duração, longa:

Exemplo 19: bfammn01[014]

*MAI: **e** **/=DCT=** essa cobra ea [/1]=SCA= ea [/1]=EMP= ea anda é assim o' //=COM=

[014]

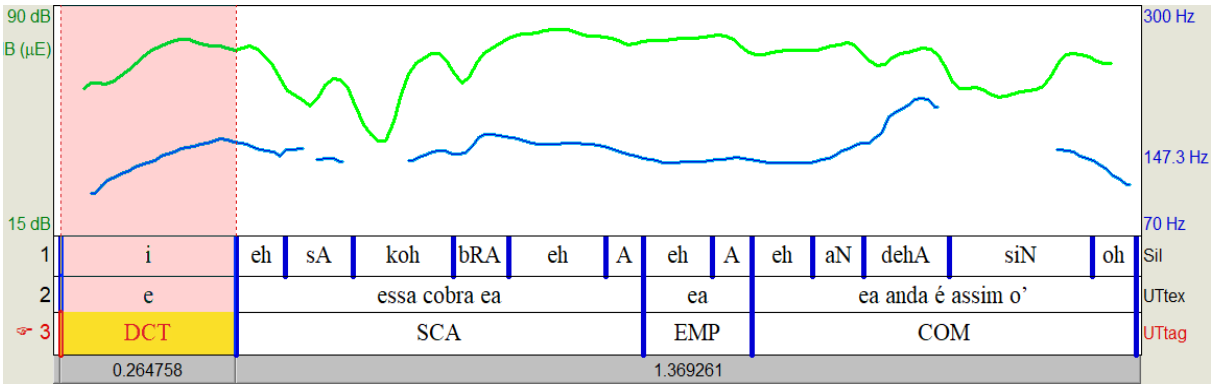


Figura 11: Exemplo de DCT no início de enunciado

Diferenciando-se das outras UDs, essa unidade não possui um item amplamente predominante. Cinco termos totalizam 85% dos casos no Mini BR Informal: *e*, *porque*,

mas, *então* e *aí*. As demais ocorrências são: *que; e então; e aí; ou; que eu; só que; aí eu; depois; mas o; ou então; isso; portanto; pra; daí; se.*

Como consequência de sua função, o DCT não se distribui livremente no enunciado, aparecendo sempre no início ou interno ao enunciado, mas no início de subpadrão de estrofe. Diferentemente do INP, que também se restringe às posições iniciais, há mais equilíbrio entre a distribuição nas duas posições (58% no início e 42% no meio do enunciado), o que seria coerente com a função.

Raso e Vieira (2016) discutem questões importantes sobre o status do DCT como UD: i) seu conteúdo semântico é normalmente significativo para o entendimento do enunciado e parece se dirigir muito mais ao texto do que ao interlocutor, o que contraria a definição de UD; ii) apesar de sempre apresentar longa duração, não é possível identificar um perfil melódico claro e iii) é difícil distinguir entre o DCT e uma unidade textual de escansão. Essas questões levantam a hipótese de que as unidades tonais anotadas como conector discursivo sejam, na verdade, unidades de escanção preenchidas por palavras ou expressões adverbiais ou conjuntivas.

5 METODOLOGIA DE ESTIMATIVA DOS PARÂMETROS PROSÓDICOS

Para alcançar os objetivos propostos, foram utilizados dados de UDs do *corpus* C-ORAL-BRASIL. Esses dados foram submetidos a uma revisão de etiquetagem e tratados para permitir a coleta de informações prosódicas por análise acústica. As medidas acústicas foram usadas na estimativa de parâmetros prosódicos que formaram a base de dados para a análise estatística. A metodologia utilizada neste trabalho está apresentada em quatro seções:

i) Descrição do *corpus* utilizado;
ii) Revisão da categorização das UDs
iii) Tratamento dos dados para análise acústica;
iv) Estimativa de parâmetros prosódicos;

5.1 A base dos dados

Todos os dados analisados neste estudo foram obtidos de uma amostra do C-ORAL-BRASIL. A seguir, encontra-se uma breve descrição do *corpus* e do *subcorpus* utilizados.

5.1.1 O *Corpus* C-ORAL-BRASIL

A base de dados utilizada neste estudo é uma amostra do C-ORAL-BRASIL (RASO e MELLO, 2012): *corpus* de fala espontânea do português brasileiro, principalmente da variante mineira, compilado pelos pesquisadores do LEEL (Laboratório de Estudos Empíricos e Experimentais da Linguagem) da Faculdade de Letras da Universidade Federal de Minas Gerais.

O C-ORAL-BRASIL foi estruturado com arquitetura equivalente à adotada no projeto C-ORAL-ROM (CRESTI e MONEGLIA, 2005), que reúne *corpora* de fala espontânea de línguas românicas da Europa: espanhol, italiano, francês e português europeu. Seus textos foram selecionados com o objetivo principal de retratar as variações situacionais da fala.

A compilação do *corpus* foi dividida em duas seções: uma informal, em contexto natural, já concluída e publicada; e outra formal em contexto natural, mídia e

conversas telefônicas, em fase final de compilação. No presente trabalho, serão utilizados apenas as gravações da seção informal.

A seção informal do *corpus* é constituída por 139 textos, contendo 208130 palavras[11], ou 1500 palavras por texto em média. A maioria dos falantes é mineira com grande concentração em Belo Horizonte e região metropolitana.

Os textos são organizados em dois contextos sociais: familiar/privado, quando os falantes interagem entre amigos e familiares, e público, quando os falantes interagem representando um papel social (no trabalho, por exemplo). A maioria dos textos (105, ~75%) são em contexto familiar/privado. Cada contexto é subdividido equilibradamente em três tipologias interacionais.

As três tipologias de interação da fala referem-se a: i) monólogo: um falante elabora o texto, sem ou quase sem interação com outros falantes; ii) diálogo: dois interlocutores interagem entre si; iii) conversação: três ou mais interlocutores interagem entre si. É importante notar que na fala espontânea in formal não há tipologias perfeitas, um monólogo pode ser acompanhado de alguma interação, por exemplo. O que determina a tipologia é a clara prevalência da construção no texto. A tabela abaixo quantifica a distribuição das palavras do *corpus* por contexto e por tipologia:

Tabela 1*: Número de palavras e distribuição por contexto social e tipologia no C-ORAL-BRASIL*

	cv[12]	**dl**	**mn**	**todas**
fam[13]	51.887	55.361	52.116	159.364
pub	14.547	17.997	16.222	48.766
todos	66.434	73.358	68.338	208.130

	cv	**dl**	**mn**	**todas**
fam	25%	27%	25%	77%
pub	7%	9%	8%	23%
todos	32%	35%	33%	100%

Os textos foram transcritos com base em normas ortográficas, porém incluindo a representação não-ortográfica dos principais fenômenos de lexicalização e gramaticalização em curso, como cliticização (enfraquecimento) dos pronomes sujeito e perda da morfologia verbal, entre outros (MELLO et al., 2012, p 131).

[11] São consideradas palavras sequências de caracteres alfabéticos da transcrição ortográfica separadas por espaço, excluindo-se palavras interrompidas, tomadas de tempo e ruídos paralinguístico. Para os critérios de transcrição usados no C-ORAL-BRASIL, ver MELLO et al. (2012).

[12] cv = conversação; dl = diálogo; mn = monólogo.

[13] fam = familiar/privado; pub = público

Para a segmentação dos enunciados foi utilizado como critério a percepção das rupturas prosódicas perceptualmente relevantes para marcar os limites das unidades tonais no fluxo da fala. As fronteiras terminais e não terminais são representadas respectivamente por barra dupla (//) e simples (/); e os enunciados interrompidos, pelo sinal de adição (+). A barra simples entre colchetes seguida de número ([/n]) indica uma interrupção dentro do enunciado com *retracting*, cancelamento das últimas n palavras pronunciadas. Sinais ortográficos de pontuação não são utilizados.

O extrato a seguir mostra três turnos totalizando treze enunciados, sendo três interrompidos e sete terminados. Destes, cinco são simples, com apenas uma unidade tonal, e dois são compostos, formados por duas ou mais unidades tonais. Em um dos enunciados, notam-se duas ocorrências de *retracting*. É possível também observar a indicação de sobreposição na fala (<>) e exemplos de transcrição não-ortográfica: a cliticização do pronome sujeito *ela* para *ea* e a redução do verbo *olha* para *o':*

Exemplo 20: bfamcv02[095-107]:

```
*TER: aqui // a mãe + não / ea tá falando do [/1] dos [/1] dos presente //
escuta só // a Dani // a Dani // a Dani falou assim / ô mãe // o' // a mãe
da Fafica / vai dar / o fogão // o tio do Anderson / <vai> dar +
*JAE: <geladeira> //
*TER: não / a geladeira já tá comprada // que a Fafica &de + já comprou //
já tá lá na casa dela //
```

[095-107]

O C-ORAL-BRASIL conta também com anotação morfossintática, realizada através do etiquetador automático (*parser*) PALAVRAS (BICK, 2000) seguido de revisão manual.

5.1.2 O Mini BR Informal

O Mini BR Informal[14] é um *subcorpus* do C-ORAL-BRASIL, composto por 20 textos[15]. A Tabela 2 e a Tabela 3[16] mostram as características dos textos selecionados e a distribuição do número de palavras por contexto social e por tipologia conversacional.

O Mini BR Informal é apropriado para estudos da estrutura informacional, como o presente trabalho. Suas unidades tonais receberam uma etiqueta que identifica as unidades informacionais sob uma perspectiva de estudos pragmáticos em acordo com a Teoria da L-AcT, descritas no capitulo 3. O extrato abaixo ilustra o mesmo trecho do Exemplo 20 com a etiquetagem informacional:

Exemplo 21: bfamcv02[095-107]:

```
*TER: [94] aqui //=COM=$[95]a mãe +=TOP=$[96] não /=INP= ea tá falando do
[/1]=SCA= dos [/1]=EMP= dos presente //=COM=$[97] escuta só //=COM=$[98] a
Dani //=COM=$[99] a Dani //=COM=$[100] a Dani falou assim /=INT= ô mãe
//=COM_r=$[101] o' //=COM_r=$[102] a mãe da Fafica /=TOP_r= vai dar /=SCA= o
fogão //=COM_r=$[103] o tio do Anderson /=TOP_r= <vai> dar +=EMP=$
*JAE: [104] <geladeira> //=COM=$
*TER: [105] não /=PHA= a geladeira já tá comprada //=COM=$[106] que a Fafica
&de [/1]=EMP= já /=SCA= comprou //=COM=$[107] já tá lá na casa dela //=COM=$
```

[095-107]

A sigla entre dois símbolos de igualdade (=) após a marca de fronteira tonal etiqueta a unidade anterior à fronteira. Neste exemplo, foram anotadas as unidades de comentário (COM), tópico (TOP), incipitário (INP), escansão (SCA), introdutor locutivo (INT), fático (PHA) e unidades interrompidas (EMP). A extensão '_r' identifica discurso reportado; o número entre colchetes indica o número sequencial do enunciado no texto e '$' marca o fim do enunciado.

Tabela 2: Número de palavras e distribuição por contexto social e tipologia no Mini BR Informal

	cv	**dl**	**mn**	**todas**
fam	6.348	8.325	8.599	23.272
pub	3.426	3.006	1.614	8.046
todos	9.774	11.331	10.213	31.318

	cv	**dl**	**mn**	**todas**
fam	20%	27%	27%	74%
pub	11%	10%	5%	26%
todos	31%	36%	33%	100%

[14] Em RASO e MELLO (2012) a amostra do C-ORAL-BRASIL é denominada *minicorpus*.

[15] Os textos foram selecionados de modo a "representar uma ampla variação diafásica e diastrática, garantindo a qualidade acústica e a diversificação dos falantes" (RASO e MELLO, 2012, p. 73).

[16] tabela adaptada de RASO e MITTIMANN, 2012, p. 220-221

Tabela 3*: Característica dos textos do* Mini BR Informal

Texto	Situação	Falantes		Duração	Nº de
		M	F	hh:mm:ss	palavras
Total		28	27	03:58:36	31.318
Conversações		15	9	01:07:28	9.774
bfamcv01	amigos avaliam um campeonato de futebol organizado por eles e planejam o próximo	4	0	00:07:00	1.467
bfamcv02	senhoras conversam sobre os preparativos do casamento de uma parente	0	3	00:07:51	1.725
bfamcv03	amigos jogam sinuca	5	0	00:06:50	1.390
bfamcv04	amigos jogam “Imagem e Ação”, após explicar as regras do jogo para uma das participantes	2	2	00:07:30	1.766
bpubcv01	funcionários de banco de sangue explicam como o sangue coletado é armazenado*	1	3	00:08:30	1.798
bpubcv02	reunião ordinária em uma sede regional de partido político	3	1	00:29:47	1.628
Diálogos		6	8	01:45:28	11.331
bfamdl01	colegas de apartamento fazem as compras do mês*	0	2	00:14:39	2.131
bfamdl02	colegas de faculdade batem papo enquanto organizam o material de gravação	1	1	00:07:26	1.572
bfamdl03	casal faz uma viagem de carro*	1	1	00:10:30	1.637
bfamdl04	domésticas, mãe e filha, fazem a limpeza da cozinha após o almoço*	0	2	00:19:32	1.249
bfamdl05	corretor de imóveis leva a irmã para visitar apartamento*	1	1	00:11:28	1.736
bpubdl01	engenheiro e pedreiro trabalham em uma obra	2	0	00:26:08	1.568
bpubdl02	cliente e vendedor interagem durante a compra de calçados*	1	1	00:15:45	1.438
Monólogos		7	10	01:05:40	10.213
bfammn01	senhor narra história fantástica sobre uma cobra	2	0	00:05:02	1.086
bfammn02	sobrinha de Carlos Drummond de Andrade conta histórias da família ao neto	1	1	00:07:23	1.677
bfammn03	narrativa de “causos” divertidos para a família*	3	3	00:07:08	1.206
bfammn04	senhora conta sua experiência no hospital após ter dado à luz no carro*	0	1	00:06:57	1.450
bfammn05	senhora fala sobre a adoção da filha após a morte de sua filha biológica*	0	2	00:09:52	1.580
bfammn06	pai conta seu percurso profissional à sua filha	1	1	00:10:02	1.600
bpubmn01	entrevista de avaliação sobre aulas de inglês na rede pública de ensino	0	2	00:19:16	1.614

5.2 Revisão do Mini BR Informal

A primeira etapa deste trabalho foi a revisão de todas as etiquetas das UDs no Mini BR Informal. A revisão foi realizada manualmente pelo autor deste trabalho e seu orientador em todas as UDs, com o auxílio dos *softwares* de análise acústica Winpitch (MARTIN, 2014); e Praat (BOERSMA e WEENINK, 2018). A revisão buscou: i) confirmar se as unidades são de fato UDs, ou seja, confirmar seu isolamento em uma

UT e seu valor não-ilocucionário; ii) verificar se as etiquetas das UDs atendem à função e à forma prosódica das unidades conativo (CNT), alocutivo (ALL) e incipitário (INP), conforme a descrição realizada em Raso (2014) e Raso e Vieira (2016); (iii) etiquetar as UDs restantes como AUX (unidade de auxílio dialógico.

Além disso, também foram marcadas UDs que não podem ser classificadas com certeza por não apresentarem informações acústicas suficientes, seja por sobreposições intensas, presença de ruído, intensidade excessivamente baixa ou ausência de curva de f0 confiável. Essas unidades foram anotadas com a etiqueta AUX_n e desconsideradas no presente estudo.

As UDs etiquetadas como conector discursivo (DCT) foram excluídas dessa revisão pelo entendimento de que essas unidades requerem um estudo dedicado a ser realizado no futuro. Excetuando-se alguns casos de evidente necessidade, as etiquetas originais dessas unidades foram mantidas.

Paralelamente à revisão das UDs, foi realizada, por equipe distinta de três etiquetadores, a revisão de transcrição, segmentação e anotação informacional de todo o Mini BR Informal, com exceção da categorização das UDs. Ao final das duas revisões, foi realizada a consolidação ambas para publicação (BOSSAGLIA, CAVALCANTE e RASO, no prelo).

Um panorama da revisão pode ser observado na Tabela 4. O número total de UDs foi reduzido de 1084 para 1025 (redução de 5,4%). As alterações na quantidade de UDs é efeito também da reavaliação da presença ou da qualidade da fronteira prosódica.

Os itens lexicais frequentes no preenchimento das UDs, como *né* e interjeições em geral, representam um desafio para a segmentação. Por não serem composicionais sintaticamente por natureza, podem induzir uma percepção condicionada de presença de fronteira, mesmo em ausência de pistas prosódicas adequadas.

Tabela 4*: Quantidade de MDs por etiqueta, antes e após a revisão*

Etiquetagem Original								Etiquetagem Revisada						
Texto	DCT	INP	ALL	CNT	PHA	EXP	Todas	DCT	INP	ALL	CNT	AUX	AUX_n	Todas
bfamcv01	5	6	3	8	8	10	40	5	2	2	9	15	6	39
bfamcv02	8	6	19	2	28	11	74	7	4	8	1	26	27	73
bfamcv03	1	3	39	6	31	15	95	1	0	15	6	51	28	101
bfamcv04	14	7	6	8	19	6	60	14	2	2	8	21	11	58
bfamdl01	3	7	10	8	27	14	69	3	2	4	11	37	11	68
bfamdl02	11	7	7	4	15	14	58	11	7	6	9	16	7	56
bfamdl03	6	16	6	5	34	8	75	6	10	1	5	32	21	75
bfamdl04	0	5	7	2	8	0	22	0	1	4	1	10	6	22
bfamdl05	2	5	4	8	33	6	58	2	0	1	4	24	24	55
bfammn01	5	0	1	0	12	2	20	5	0	0	2	11	2	20
bfammn02	19	6	2	0	26	8	61	19	3	2	1	19	4	48
bfammn03	18	3	4	0	18	3	46	16	1	2	4	8	7	38
bfammn04	6	3	8	1	25	7	50	6	1	5	3	34	2	51
bfammn05	26	5	4	0	8	1	44	26	0	2	1	10	0	39
bfammn06	10	3	0	1	6	0	20	10	1	0	2	6	0	19
bpubcv01	8	3	3	3	18	7	42	8	0	1	3	18	8	38
bpubcv02	6	2	12	0	23	4	47	6	1	6	3	19	12	47
bpubdl01	4	5	1	0	40	9	59	4	3	1	1	48	1	58
bpubdl02	7	6	2	11	21	6	53	7	2	0	8	24	10	51
bpubmn01	19	4	1	2	58	7	91	17	0	1	2	48	1	69
Todos	178	102	139	69	458	138	1084	173	40	63	84	477	188	1025
							Variação (%):	-2,8	-61	-53	+22			-5,4

Uma outra razão para a redução no número de UDs é a reclassificação para outra unidade informacional não dialógica. O caso mais comum de alteração é o reconhecimento de força ilocucionária em diversos itens. Um caso extremo é o *né,* que pode veicular um pedido de confirmação através de forma prosódica convencionalizada de comentário múltiplo (CMM).

As unidades etiquetadas como INP e ALL foram as mais afetadas na revisão, com redução de 61 e 53% de suas ocorrências, respectivamente. Em contrapartida, houve aumento do número de marcadores com a função prosódica de conativo (15 unidades ou 22% a mais).

Os INPs foram em larga escala reclassificados por não atenderem pelo menos um dos critérios da forma: intensidade, f0 e taxa de elocução[17] maiores do que da unidade de referência (COM). É importante mencionar que nos casos marginais, como por exemplo se apenas um dos critérios não fosse atingido por diferença pequena, ainda

[17] em dB, Hz e sílabas/s, respectivamente.

assim a classificação pode ter sido mantida se a função de INP fosse evidente. Entretanto, a maioria dos casos duvidosos foram etiquetados como AUX.

A redução dos ALLs também foi muito significativa. Em razão de suas características de hipoarticulação e baixa intensidade, ocorreram casos em que a curva de f0 não era detectável na análise acústica, impedindo a confirmação. Essas unidades foram reclassificadas com a etiqueta AUX_n. Os exemplos a seguir mostram dois casos em que a classificação não é possível.

No Exemplo 22, a sobreposição torna o conteúdo praticamente inaudível e causa uma anomalia na curva de f0 (em azul na Figura 12):

Exemplo 22: bfamcv02[045]

```
*RUT: uai /=AUX= e cê acha isso velho /=COM= <Terezinha> //=AUX_n=
```

[045] [045]a

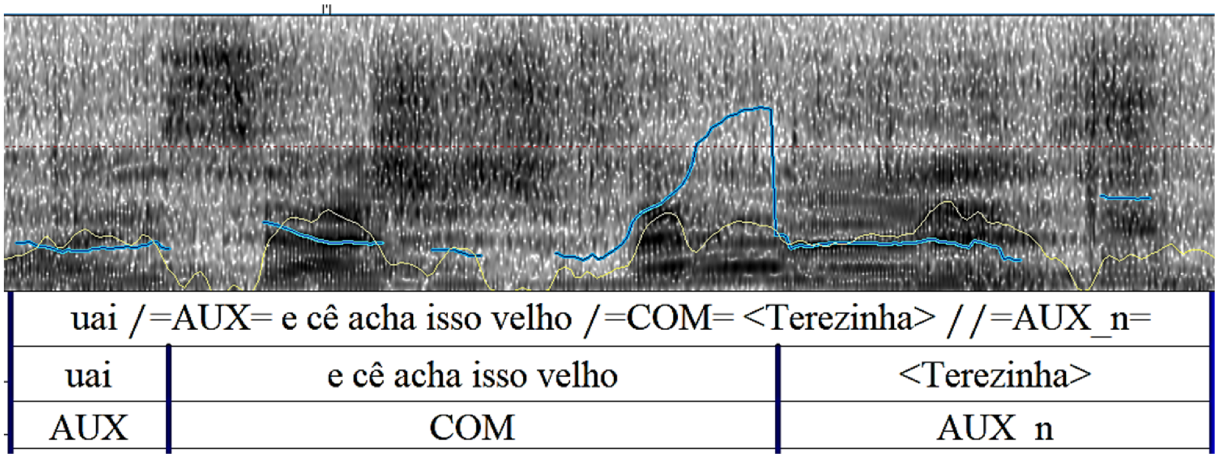


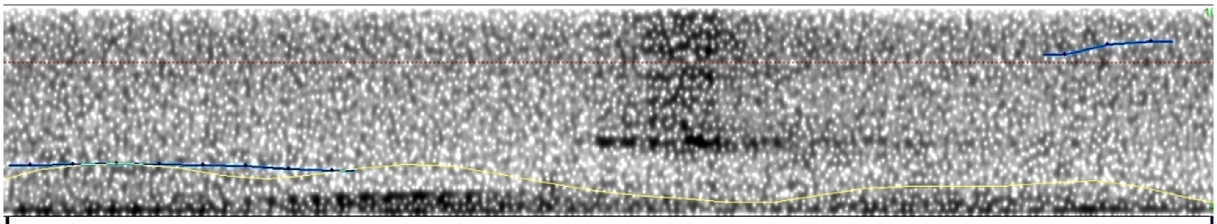
Figura 12: ALL reclassificado como AUX_n por sobreposição

No Exemplo 23, não há sobreposição, mas a qualidade do áudio compromete a compreensão do enunciado e a confiabilidade da curva de f0. A pequena porção com curva de f0 (em azul na Figura 13) ocorre no final da unidade, possivelmente resultante de ruído de fundo:

Exemplo 23: bpubcv02[194]

```
*CAR: uhn /=COM= Osvaldo //=AUX_n=
```

[194] [194]a



uhn /=COM= Osvaldo //=AUX_n=	
uhn	Osvaldo
COM	AUX_n

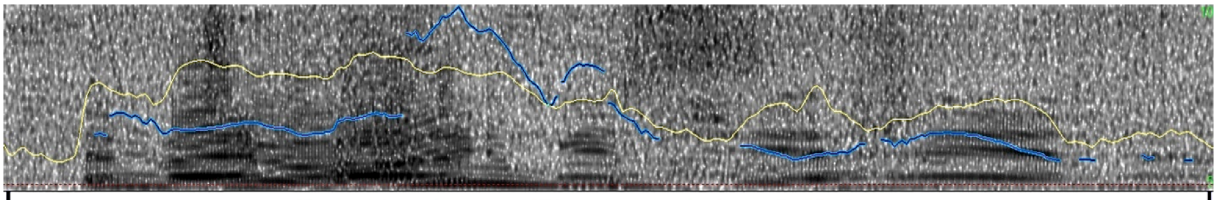
Figura 13: ALL reclassificado como AUX_n por ausência de informações acústicas

Alguns ALLs foram reclassificados como CNT, apesar de serem lexicalmente preenchidos por nomes próprios ou epítetos. O Exemplo 24 ilustra um caso bastante claro. A curva de f0 (em azul) tem movimento inicialmente ascendente e desce após a sílaba tônica. A intensidade (em amarelo) também não desce imediatamente. Pela oitiva do enunciado percebe-se a intenção de captar a atenção do interlocutor:

Exemplo 24: bfamdl01[028]

```
*FLA: <cê vai embora que> dia /=COM= Rena //=CNT=
```

[028] [028]a



<cê vai embora que> dia /=COM= Rena //=ALL=	
<cê vai embora que> dia	Rena
COM	CNT

Figura 14: ALL reclassificado como CNT

Pelos critérios utilizados na etiquetagem original do Mini BR Informal, quase todo nome próprio ou epíteto em unidade tonal própria não ilocucionária foi classificado como ALL. Na revisão, alguns ALL foram reclassificados como CNT. Os casos duvidosos foram etiquetados como AUX.

Nas tabelas abaixo, pode-se observar a distribuição da ocorrência de UDs por contexto social e por tipologia de interação. Como os tipos não são igualmente distribuídos no Mini BR Informal, foi calculada a frequência, em número de UDs a cada 1000 palavras, para permitir a comparação.

Tabela 5: Quantidade e frequência de UDs por tipologia interacional

	Número de ocorrências							Frequência a cada 1000 palavras						
	DCT	INP	ALL	CNT	AUX	AUX_n	Todas	DCT	INP	ALL	CNT	AUX	AUX_n	Todas
cv	41	9	34	30	150	92	356	4,2	0,9	3,5	3,1	15,3	9,4	36,3
dl	33	25	17	39	191	80	385	2,9	2,2	1,5	3,4	16,9	7,1	34,0
mn	99	6	12	15	136	16	284	9,7	0,6	1,2	1,5	13,3	1,6	27,8
Todos	173	40	63	84	477	188	1025	5,5	1,3	2,0	2,7	15,2	6,0	32,7

Tabela 6: Quantidade e frequência de UDs por contexto social

	Número de ocorrências							Frequência a cada 1000 palavras						
	DCT	INP	ALL	CNT	AUX	AUX_n	Todas	DCT	INP	ALL	CNT	AUX	AUX_n	Todas
fam	131	34	54	67	320	156	762	5,6	1,5	2,3	2,9	13,8	6,7	32,7
pub	42	6	9	17	157	32	263	5,2	0,7	1,1	2,1	19,5	4,0	32,7
Todos	173	40	63	84	477	188	1025	5,5	1,3	2,0	2,7	15,2	6,0	32,7

Tabela 7: Quantidade e frequência de UDs por contexto e tipologia

	Número de ocorrências							Frequência a cada 1000 palavras						
	DCT	INP	ALL	CNT	AUX	AUX_n	Todas	DCT	INP	ALL	CNT	AUX	AUX_n	Todas
famcv	27	8	27	24	113	72	271	4,3	1,3	4,3	3,8	17,8	11,3	42,5
famdl	22	20	16	30	119	69	276	2,6	2,4	1,9	3,6	14,3	8,3	33,2
fammn	82	6	11	13	88	15	215	9,5	0,7	1,3	1,5	10,2	1,7	25,0
pubcv	14	1	7	6	37	20	85	4,1	0,3	2,0	1,8	10,8	5,8	24,8
pubdl	11	5	1	9	72	11	109	3,7	1,7	0,3	3,0	24,0	3,7	36,3
pubmn	17	0	1	2	48	1	69	10,5	0,0	0,6	1,2	29,7	0,6	42,8
Todos	173	40	63	84	477	188	1025	5,5	1,3	2,0	2,7	15,2	6,0	32,7

Percebe-se um equilíbrio na ocorrência total de UDs entre os contextos sociais (32,7/1000 palavras), embora essa distribuição equilibrada não se reflita em todas as categorias. ALLs, CNTs e INPs foram mais frequentes em contextos familiares do que em públicos. Por outro lado, as unidades etiquetadas como AUX parecem ser mais frequentes em contexto público, embora o oposto aconteça nas conversas.

Por tipologia, percebe-se uma maior frequência de DCTs nos monólogos, de INPs nos diálogos e de ALL nas conversações. O número de CNTs é menor nos monólogos e equilibrado nas outras tipologias. Essas diferenças são coerentes com as funções das

UDs: O DCT é frequente nos monólogos pela maior presença das *stanzas* contendo número grande de UI; por identificar o falante o ALL é natural que ocorra mais nas interações entre mais de dois participantes. O CNT, que é um convite à interação, tem menor aplicação em discursos monológicos.

Para as unidades sem classificação, não há tendência clara. Em contexto familiar, há mais AUXs nas conversas; em contexto público, ocorreram mais nos monólogos. As unidades com baixa qualidade acústica (AUX_n) foram mais frequentes nas conversações, o que é esperado pela maior sobreposição e pela participação de falantes sem microfone dedicado durante as gravações.

Apesar das diferenças destacadas, é importante observar que a quantidade de textos por tipologia e por contexto é pequena e insuficiente para suportar generalizações.

A seguir, algumas características quantitativas do Mini BR Informal a respeito do número de enunciados e de unidades tonais após a revisão:

- 5467 enunciados:
  - 3138 enunciados simples (57,4%);
  - 2329 enunciados complexos (42,6%)
- 10707 unidades tonais:
  - 1,96 unidades por enunciado, em média;
  - 2,36 por enunciado, nos enunciados complexos.
  - Máximo de UT em 1 enunciado: 35
- 1024 UDs:
  - 9,6% das unidades tonais;
  - 24% das unidades tonais não ilocucionárias[18];
  - 0,095 UD por enunciado, em média (~1 UD a cada 10 enunciados)
  - 0,46 UD por enunciado complexo (~1 UD a cada 2 enunciados complexos)
- 873 enunciados contendo UDs:
  - 16,0% dos enunciados;
  - 37,5% dos enunciados complexos;
  - 741 enunciados com 1 UD (85% dos enunciados com UDs);

[18] Todas unidades informacionais exceto COM, CMM e COB.

- 114 enunciados com 2 UDs (13% dos enunciados com UDs);
- Máximo de UDs em 1 enunciado: 5

O exemplo a seguir mostra uma *stanza* com cinco UDs, que também é o enunciado com maior número de unidades tonais do Mini BR Informal (35).

Exemplo 25: bpubmn01[048]

*SHE: **agora /=AUX=** se eu for olhar tempo mesmo /=TOP= **né /=AUX=** de [/1]=EMP= de [/1]=EMP= de sentar e escrever /=APT= aquilo /=APT(1)= pra dar aquela aula ideal /=APT= **né /=AUX=** quela [/1]=EMP= quela aula igual se dá no EDUCONLE que cê acha que é &o [/1]=SCA= é tudo de bom /=PAR= **né /=AUX=** nu tem /=COB= porque também eu acho que /=SCA= &he /=TMT= &v [/1]=EMP= depende muito dos alunos de tar preparado /=COB= por exemplo /=PAR= cê prepara aquela aula /=COB= maravilhosa /=COB= e acha que vai dar tudo certo /=COB= de repente ocê encontra lá o aluno /=COB= **né /=AUX=** que /=SCA= &he /=TMT= &he /=TMT= nu [/1]=EMP= nu [/1]=EMP= foi pra escola /=COB= nu almoçou /=COB= tá [/1]=EMP= tá com fome /=COB= ou brigou com o pai /=COB= ou aquelas questões todas familiares que vão influenciar //=COM=



[048]

O preenchimento lexical das UDs após a revisão é mostrado na Tabela 8. De forma geral, não houve mudanças significativas; a de maior destaque é a reclassificação das UDs preenchidas por nomes próprios, exclusivas do ALL antes da revisão do Mini BR Informal, em CNT (8) ou etiquetadas como AUX (8) ou AUX_n (27). O item *né* é destacadamente o mais frequente; responde a 20% de todas as unidades e a 43% das UDs AUX que serão investigadas neste trabalho.

Tabela 8: Preenchimento lexical das UDs

DCT		INP		ALL		CNT		AUX		AUX_n	
itens: 21	173	itens: 11	40	itens: 14	63	itens: 31	84	itens: 45	477	itens: 41	188
porque	34	não	16	(nome)	35	o'	15	né	206	né	47
e	33	ah	9	minha filha	7	ô	9	ah	39	(nome)	27
mas	30	é	3	cara	5	(nome)	8	uai	34	não	14
então	27	então	3	(pt)+(nome)	3	não	6	não	30	sô	12
aí	20	ah não	2	gente	2	sô	5	é	17	uai	12
que	8	bom	2	meu filho	2	aqui o'	4	hein	17	ah	9
e então	4	aí	1	velho	2	pera aí	4	ué	14	o'	9
e aí	3	é não	1	bicho	1	cara	3	agora	11	é	7
ou	2	mas	1	boba	1	olha	3	Nossa	9	Nossa	4
aí eu	1	se bem	1	filha	1	tá	3	viu	9	então	3
daí	1	também	1	homem	1	bom	2	(nome)	8	gente	3
depois	1			mãe	1	quer ver	2	sabe	7	velho	3
e aí então	1			meu	1	tá vendo	2	assim	6	viu	3
isso	1			meu pai	1	ah	1	então	6	aí	2
mas o	1					aí	1	gente	6	assim	2
ou então	1					aqui	1	No'	6	filho	2
portanto	1					bem	1	porque	6	hein	2
pra	1					bicho	1	sô	6	mãe	2
que eu	1					cê sabe	1	bom	4	minha filha	2
se	1					então o'	1	ai	3	uhn	2
só que	1					mãe	1	mas	3	agora	1
						mas não não	1	olha	3	aí não	1
						o' (nome)	1	aí	2	aqui	1
						ocê o'	1	eh	2	aqui o'	1
						olha p' cê ver	1	ô	2	bom	1
						pera pera aí	1	uhn	2	cara	1
						pô	1	ah não	1	eh	1
						por favor	1	ahn	1	hein (nome)	1
						quer dizer	1	cê vê	1	meu	1
						uai	1	é não	1	meu filho	1
						viu	1	entendeu	1	moça	1
								filhote	1	né gente	1
								mas assim	1	No'	1
								menino	1	ô	1
								Nossa Senhora	1	ô gente	1
								Nossa Sio'	1	pera aí	1
								o'	1	por favor	1
								oh	1	porra	1
								olha p' cê ver	1	só um segundo	1
								pelo amor de Deus	1	tá	1
								pois é	1	xá eu ver	1
								porque né	1		
								que	1		
								tá	1		
								também	1		

(nome): nome próprio; (pt)+(nome): pronome de tratamento seguido de nome próprio

5.3 Tratamento dos dados

Com o Mini BR Informal revisado, uma série de procedimentos foi aplicada para permitir a coleta de informações acústicas das UDs. O tratamento foi realizado em três etapas principais que serão descritas e comentadas a seguir:

vi) Preparação dos dados para análise no Praat;
vii) Segmentação silábica e alinhamento das fronteiras prosódicas;
viii) Indicação da unidade de referência de cada UD.

5.3.1 Preparação dos dados

A partir de recurso do software Winpitch, *software* utilizado para o alinhamento som-texto do Mini BR Informal, foram geradas planilhas eletrônicas com a lista numerada de enunciados para cada um dos 20 textos, com a identificação do falante, o tempo de início e de fim do enunciado e a transcrição com a etiqueta informacional. A Figura 15 mostra um extrato dos dados de alinhamento dos enunciados:

File	Rank	Speaker	Text	E_Start	E_End	E_Dur
bfamcv01	1	LEO	o Juninho <foi> //=COM=	0,000	1,056	1,056
bfamcv01	2	GIL	<ô /=CNT= mas> /=DCT= voltando à questão /=TOP= falando em [/2]=EMP= e	0,512	9,139	8,626
bfamcv01	3	LUI	<não> //=COM=	8,507	9,139	0,631
bfamcv01	4	LEO	<não> //=COM=	8,965	9,307	0,342
bfamcv01	5	LUI	<eu acho não> //=COM=	9,139	9,992	0,853
bfamcv01	6	LEO	<com certeza> //=COM=	9,478	10,388	0,910
bfamcv01	7	LUI	<com certeza es nu vão participar /=COM= uai> //=AUX_n=	9,992	11,555	1,562
bfamcv01	8	LEO	<eles são piores do que o> Durepox //=COM=	10,388	12,523	2,135
bfamcv01	9	EVN	é /=CMB= pois <é> //=COM=	12,523	13,804	1,281
bfamcv01	10	LUI	<agora> manda uma barrinha <minha> //=COM=	13,519	14,886	1,366
bfamcv01	11	EVN	<porque o Durepox> /=TOP= pelo menos jogava bola //=COM=	14,105	16,186	2,080
bfamcv01	12	GIL	não /=AUX= e o Durepox /=TOP= eu vou +=EMP=	16,186	17,598	1,411
bfamcv01	13	GIL	tinha um cara //=COM=	17,598	18,452	0,854
bfamcv01	14	GIL	era &aque [/2]=EMP= era &aque [/2]=EMP= era aquele cara <lá /=TOP= que e	18,452	22,749	4,296
bfamcv01	15	EVN	<era aquele cara /=CMB= é> //=COM=	19,846	21,155	1,309
bfamcv01	16	LUI	<escroto> /=TOP= <e como> ele era amigo dos caras /=TOP= a galera meio c	21,155	25,538	4,382
bfamcv01	17	LUI	mas nu [/2]=EMP= mas nu [/2]=EMP= eles nu eram todos <escrotos> /=SCA=	25,538	30,548	5,009
bfamcv01	18	GIL	<é> //=COM=	28,071	28,470	0,398
bfamcv01	19	EVN	No' /=AUX= o Galáticos é mesmo /=INT= todo mundo é <babaca> //=COM=	30,719	32,512	1,792
bfamcv01	20	GIL	<não> /=COB= e assim /=COB= esse que é o ponto //=COM=	32,192	33,849	1,656
bfamcv01	21	GIL	eles <reclamavam> +=EMP=	33,955	35,321	1,365
bfamcv01	22	LEO	<não /=AUX_n= se bem que tinha> [/1]=SCA= tinha uns dois que falaram assir	34,235	38,337	4,102

Figura 15: Extrato da tabela gerada pelo Winpitch[19]

[19] *File*: texto; *Rank*: número do enunciado; *Text*: transcrição com segmentação e etiqueta; *E_Start* e *E_End*: tempo de início e de fim do enunciado; *E_Dur*: duração do enunciado.

No	File	Rank	Speaker	Tot_UT	No_UT	Text	Fron	Tag	W_C
1	bfamcv01	1	LEO	1	1	o Juninho <foi>	//	COM	3
2	bfamcv01	2	GIL	8	1	<ô	/	CNT	1
2	bfamcv01	2	GIL	8	2	mas>	/	DCT	1
2	bfamcv01	2	GIL	8	3	voltando à questão	/	TOP	3
2	bfamcv01	2	GIL	8	4	falando em	[/2]	EMP	2
2	bfamcv01	2	GIL	8	5	e também falando em povo mascarado	/	TOP	6
2	bfamcv01	2	GIL	8	6	esse povo do Galáticos é muito palha	/	COB	7
2	bfamcv01	2	GIL	8	7	eu acho que es nu deviam mais participar	/	COM	8
2	bfamcv01	2	GIL	8	8	e <tal>	//	UNC	2
3	bfamcv01	3	LUI	1	1	<não>	//	COM	1
4	bfamcv01	4	LEO	1	1	<não>	//	COM	1
5	bfamcv01	5	LUI	1	1	<eu acho não>	//	COM	3
6	bfamcv01	6	LEO	1	1	<com certeza>	//	COM	2
7	bfamcv01	7	LUI	2	1	<com certeza es nu vão participar	/	COM	6
7	bfamcv01	7	LUI	2	2	uai>	//	AUX_n	1
8	bfamcv01	8	LEO	1	1	<eles são piores do que o> Durepox	//	COM	7
9	bfamcv01	9	EVN	2	1	é	/	CMB	1
9	bfamcv01	9	EVN	2	2	pois <é>	//	COM	2
10	bfamcv01	10	LUI	1	1	<agora> manda uma barrinha <minha>	//	COM	5
11	bfamcv01	11	EVN	2	1	<porque o Durepox>	/	TOP	3
11	bfamcv01	11	EVN	2	2	pelo menos jogava bola	//	COM	4
12	bfamcv01	12	GIL	3	1	não	/	AUX	1

Figura 16: Extrato da tabela de unidades tonais[20]

A partir da tabela de enunciados, foi produzida uma lista das unidades tonais, através de *script* em linguagem R (R CORE TEAM, 2016), com o auxílio da interface RStudio© (RSTUDIO, 2017). A Figura 16 ilustra algumas linhas do banco de dados[21]. O *script* foi utilizado antes e após a revisão do Mini BR Informal para produzir as comparações tabeladas na seção 5.2.

Com a informação de início e fim, o trecho correspondente a cada enunciado foi recortado do arquivo de áudio original, através de *script* programado para o *software* de análise acústica Praat. Foram adicionados 300ms antes e depois dos tempos indicados para uma visualização do contexto adjacente imediato. Nos textos gravados em estéreo, ambos os canais foram recortados separadamente, sendo o canal correspondente ao falante do enunciado[22] selecionado para a análise e o outro, arquivado para consulta, se necessário.

Como preparação para a etapa de segmentação silábica e alinhamento das unidades tonais, o mesmo *script* utilizou a tabela de unidades tonais para criação de um arquivo *TextGrid* (objeto do Praat utilizado para segmentação e anotação) para cada

[20] *No*: número sequencial; *File*: texto; *Rank*: número do enunciado; *Speaker*: falante; *Text*: transcrição; *Tot_UT*: total de Uts no enunciado; *No_UT*: número da UT no enunciado; *Fron*: tipo de fronteira ao final da UT; *Tag*: etiqueta informacional; *W_C*: número de palavras (*word count*).

[21] As tabelas completas de enunciados e de unidades tonais estão disponíveis no Apêndice B.

[22] Os canais foram escolhidos pela intensidade média por falante no texto. Esses dados encontram-se no Apêndice E.

enunciado. Os *TextGrids* foram estruturados com 3 camadas (*tiers*): i) a transcrição etiquetada, ii) a transcrição ortográfica de cada unidade tonal e iii) a etiqueta informacional da unidade, distribuídas em períodos de tempos iguais (ainda desalinhados com o áudio). A figura a seguir mostra o espectrograma e a forma de onda de um enunciado (bfammn05-90) com seu correspondente *TextGrid;* camada 1 selecionada:

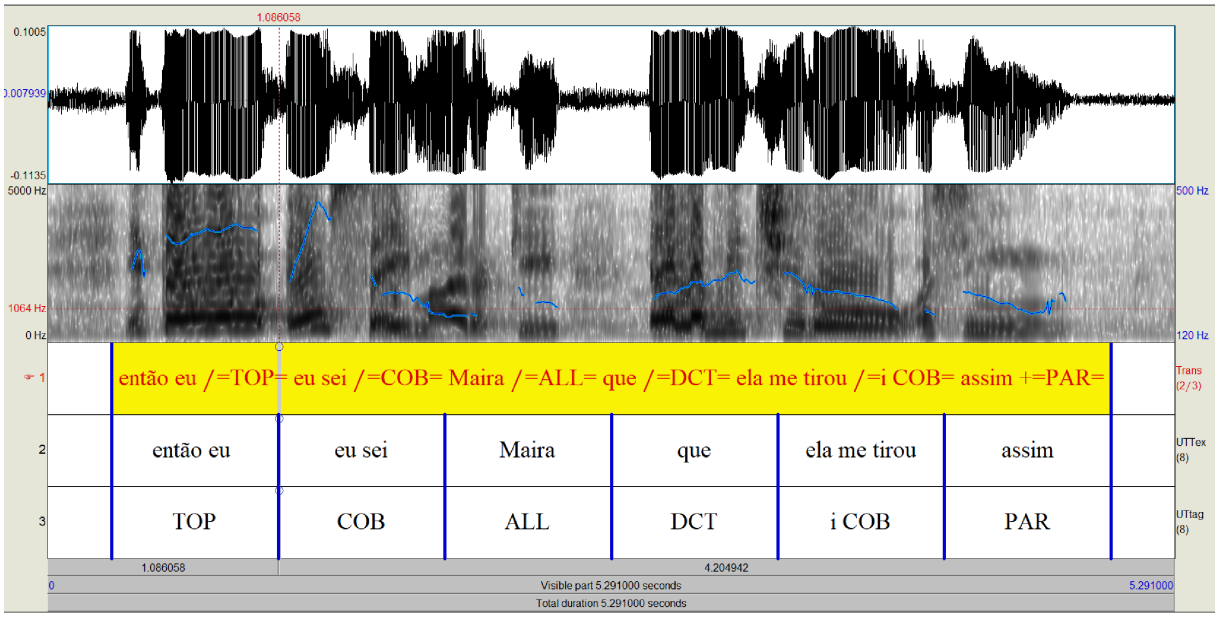


Figura 17: Exemplo de TextGrid gerado para a segmentação silábica e para alinhamento das unidades tonais (bfammn05-90)

5.3.2 Alinhamento e segmentação

O alinhamento das unidades tonais e a segmentação silábica foram executados em cada um dos enunciados que continham UDs, com exceção daquelas etiquetadas como AUX_n e DCT. Essa etapa contemplou as seguintes tarefas:

i. Alinhar com o áudio as fronteiras das UTs dentro do enunciado;
ii. Segmentar o enunciado em sílabas;
iii. Transcrever foneticamente;
iv. Delimitar as pausas nas fronteiras tonais;
v. Delimitar a vogal tônica na UD;

A segmentação silábica e o alinhamento tonal foram realizados em 592 enunciados contendo 664 UDs. Nem todas unidades tonais desses enunciados foram alinhadas e

segmentadas. Como apenas as unidades etiquetadas como CNT, ALL, INP e AUX foram consideradas no escopo do trabalho, o alinhamento e a segmentação silábica se restringiu a estas UDs e às unidades ilocucionárias associadas às UDs.

A identificação das fronteiras dos segmentos foi realizada pelo autor deste trabalho segundo orientações descritos por Machač e Skarnitzl (2014). Foram utilizadas pistas visuais do espectrograma, da forma da onda sonora e das curvas de intensidade, além da oitiva.

A figura abaixo mostra o mesmo enunciado da Figura 17 após a segmentação. Foram inseridas duas camadas no *TextGrid*, uma para a transcrição silábica e outra para a marcação da vogal tônica da UD[23]. As pausas foram marcadas com a etiqueta 'pp'. Nos poucos casos em que havia mais de uma sílaba tônica na UD, foi marcada como referência a de maior saliência.

Exemplo 26: bfammn05[090]

```
*CAR: então eu /=TOP= eu sei /=COB= Maira /=ALL= que /=DCT= ela me tirou
/=i COB= assim +=PAR=
```

[090]

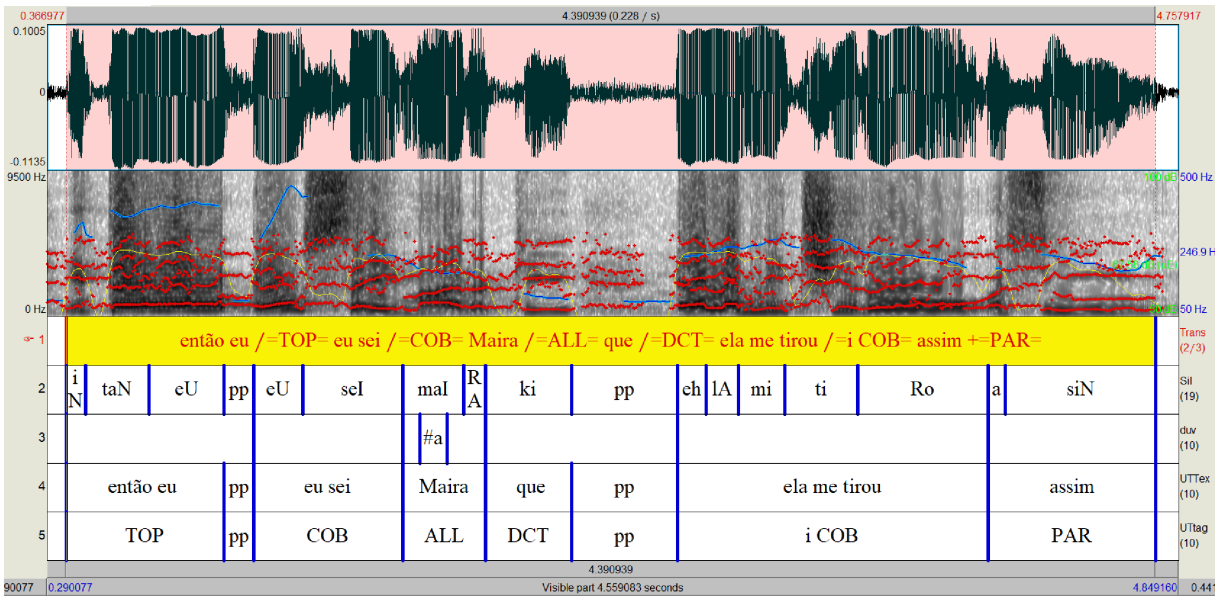


Figura 18: Exemplo de TextGrid após o alinhamento e a segmentação (bfammn05-90)

[23] O simbolo # foi utilizado como convenção particular para identificar a vogal tônica mais longa, informação relevante nos casos em que a UD contém mais de uma sílaba tônica.

A transcrição fonética utilizou-se de caracteres ASCII por facilidade de anotação e para ser compatível com o procedimento de estimativa da duração normalizada (BARBOSA, 2013), que será descrito na seção 5.4.1. Utilizou-se uma transcrição fonética larga, sem representação distinta para alofones[24]. A notação utilizada, bem como sua equivalência em símbolos IPA está indicada na Tabela 9.

Apesar da ótima qualidade acústica das gravações do Mini BR Informal (RASO e MELLO, 2012, p. 74), a natureza espontânea e a diversidade de situações e de ambientes em que foram tomadas produzem imagens espectroscópicas muito mais poluídas do que aquelas obtidas em falas lidas em cabines com proteção e abafamento acústico apropriado. Em consequência, a análise oitiva torna-se decisiva em maior extensão que o desejado, por vieses induzidos pela percepção subjetiva. As características inerentes de muitas UDs de baixa intensidade, alta taxa de elocução e de redução do material fonético amplificam dificuldades que podem levar a imprecisões.

As mesmas razões dificultam muito a aplicação de segmentação automática nas gravações analisadas, o que poderia ser uma alternativa para aumentar a velocidade e reduzir a subjetividade. Na impossibilidade de automatização, a realização da etapa por mais um ou dois segmentadores experientes em paralelo seria recomendado, assim como medidas de controle de qualidade para validação da segmentação.

5.3.3 Referência para as medidas da fala

De modo geral, os valores absolutos das medidas acústicas da fala não podem ser comparados diretamente. As medidas de intensidade, por exemplo, são muito dependentes de fatores que não têm ligação direta com aspectos linguísticos, como o tipo de microfone utilizado e a distância do microfone até a fonte, fazendo com que gravações diferentes sejam completamente incomparáveis. De forma semelhante, a idade, o gênero e as dimensões do trato vocal interferem na tessitura vocal do falante, e idiossincrasias rítmicas podem fazer que medidas de taxa de elocução de um falante particular sejam diferentes sem que transmitam funções linguísticas. Atitude, postura

[24] Mais informações sobre a transcrição podem ser encontradas em Barbosa at al (1999).

interpessoal, estilo de elocução e estado emocional também são veiculados pela prosódia e, consequentemente, geram variabilidade nas medidas acústicas sem (ou com pouca) relevância para a estrutura informacional.

Os efeitos das particularidades de cada enunciação, são contornados tomando-se medidas relativas a cada contexto específico. Desta forma, a informação relevante não é o valor absoluto da medida acústica, mas o quanto este valor se distancia do valor contextual da mesma medida, indicado pela referência. Exceto para algumas medidas que serão explicitadas oportunamente, para a maioria dos parâmetros analisados, foi adotada a diferença proporcional da medida da UD em relação à referência, conforme a expressão:

$$M_{UD}^{d} = \frac{M_{UD}^{a} - M_{Re}^{a}}{M_{Re}^{A}} \quad (5.1)$$

Onde:

M_{UD}^{d}= medida da diferença proporcional da UD em relação à referência;

M_{UD}^{a}= medida absoluta da UD;

M_{Re}^{a}= medida absoluta da referência

Neste trabalho foi adotada como referência a unidade informacional ilocucionária (comentário), que constitui o núcleo do enunciado, seguindo o procedimento usado em Raso (2014) e Raso e Vieira (2016). Supõe-se que a unidade de comentário, por ser a única necessária e suficiente para transmitir a ilocução, deva conter outras informações prosódicas que caracterizam as particularidades do enunciado. Por essa razão, todas as unidades dependem, direta ou indiretamente, da ilocução. Adicionalmente, é a única unidade presente em todos os enunciados, com exceção de alguns enunciados interrompidos[25]. Por essas razões, consideramos que a unidade de comentário é a melhor referência para as medidas das UDs.

Muitos enunciados possuem mais de uma unidade ilocucionária (comentários múltiplos ou comentários ligados); as referências das UDs nestes casos foram escolhidas caso a caso. De um modo geral, foi considerado o encaixamento da UD no

[25] Cerca de 20 UDs foram desconsideradas por pertencerem a enunciados que foram interrompidos antes da enunciação do comentário.

subpadrão interno ao enunciado, principalmente nos casos de discurso reportado, e a proximidade com a UD. A Figura 19 mostra a posição das UDs estudadas, em número de UTs, em relação à unidade de comentário de referência. Mais de 80% das UDs são adjacentes ao comentário (-1 ou 1[26]) e apenas 20% delas se separam por 4 ou mais UTs do comentário, antes ou depois.

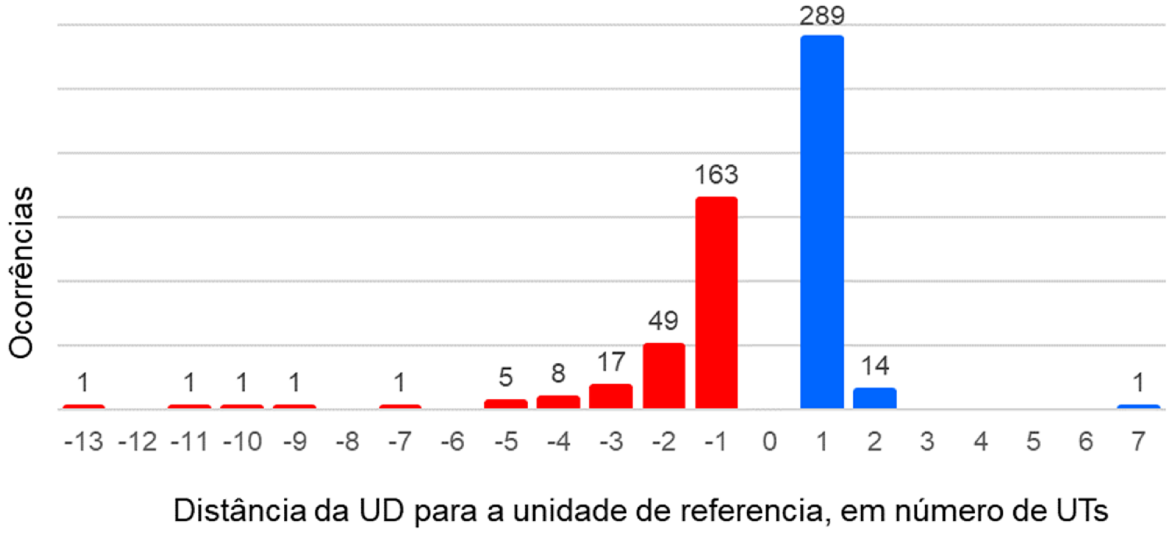


Figura 19: Posição da UD em relação à unidade de referência, em UT

Apesar de considerarmos a melhor referência, a utilização do comentário pode levar a algumas questões a se discutir: (i) Algumas ilocuções são marcadas por amplos movimentos melódicos ou ênfases duracionais que podem afetar as medidas de diferença proporcional sem que necessariamente reflitam aspectos formais da UD em questão; (ii) apesar de a maioria das UDs estarem próximas do comentário e percentualmente se apoiarem diretamente nele, em alguns casos a UD está separada por muitas unidades informacionais (tipicamente unidades de escansão) (iii) algumas UDs, o INP particularmente, podem contrastar f0 e intensidade com a unidade adjacente, que não necessariamente é a ilocucionária; (iv) nos casos com mais de uma unidade ilocucionária, é possível que mais de uma ou todas sejam a melhor referência, e não a mais próxima como foi adotado. Este é o caso dos CMMs, em que duas ou mais unidade formam um padrão ilocucionário único.

[26] Números negativos significam que a UD antecede o comentário.

Um estudo direcionado para a escolha da unidade de referência pode ser interessante, pois a questão é relevante para muitas análises que se utilizam de medidas acústicas para caracterizar a estrutura informacional do enunciado. Alternativamente, outras opções de referência poderiam ser o enunciado todo (ou excluindo a UD), o subpadrão ilocucionário, as unidades tonais adjacentes ou dados de referência para o falante. Todas essas opções apresentam desvantagens. Os falantes frequentemente mudam de atitude ou modalidade ao longo do discurso ou mesmo dentro do enunciado (tipicamente em meta-ilocuções) e, consequentemente, alteram seus patamares de f0, taxa de elocução e intensidade.

5.4 Estimativa dos parâmetros prosódicos

Nesta seção serão apresentados as definições e os procedimentos utilizados para a obtenção dos valores dos parâmetros prosódicos que foram analisados estatisticamente. Ao todo foram estimados 25 parâmetros[27], que serão organizados em cinco subgrupos com intuito principal de facilitar sua apresentação:

i) parâmetros de duração;
ii) parâmetros de intensidade;
iii) parâmetros de f0
iv) parâmetros de variação de f0
v) parâmetros de alinhamento

5.4.1 Parâmetros de duração

Para a extração das medidas de duração foi criado um *script* no Praat para a coleta dos tempos iniciais e finais das unidades tonais de interesse e das sílabas a partir dos *TextGrids* segmentados. A partir dessas medidas foram calculados os parâmetros seguintes:

i) *Dur*: a duração total da unidade;
ii) *N_sil*: o número de sílabas fonéticas da unidade;

[27] O termo parâmetro é utilizado neste trabalho com sentido de atributos (*feartures*) dos dados (as UDs) ou varíaveis dos modelos estatísticos.

iii) *Dm_sil*: a duração silábica média;

iv) *Zm_sil*: a duração normalizada silábica média;

v) *Z_ut*: a duração normalizada das UDs.

Para poder comparar a duração de unidades preenchidas por lexemas de comprimento diferente o que é considerado é a duração silábica média da unidade, calculada como a razão entre a duração da UT e o número de sílabas:

$$Dm_{sil}^{UT} = \frac{T_f^{UT} - T_i^{UT}}{N_sil} = \frac{Dur^{UT}}{N_sil} \quad (5.2)$$

Onde:

Dm_{sil}^{UT} = Duração silábica média, em ms/sílaba;

Dur^{UT} = Duração da UT;

N_sil= Número de sílabas;

T_f^{UT} *e* T_i^{UT} = Tempo final e inicial da UT.

Apesar de as medidas poderem ser diretamente tomadas a partir da diferença entre o tempo de fim e de início, não é simples determinar se um segmento ou uma sequência sonora foram enunciadas com maior ou menor duração. Não apenas sons diferentes possuem duração intrinsicamente diferentes como um mesmo som pode apresentar uma grande variabilidade duracional em contextos fonéticos e taxa de elocução diferentes (FOX, 2000, p.12-13). Para minimizar os efeitos das durações intrínsecas dos segmentos, foi realizada a normalização da duração através da transformação linear em desvio padronizado, ou escore Z:

$$Z^D = \frac{X^D - \mu^D}{\sigma^D} \quad (5.3)$$

Onde:

Z^D= *Z score* da duração;

X^D= valor medido de duração do segmento;

μ^D= valor médio esperado de duração do segmento;

σ^D= desvio-padrão da duração do segmento.

O desvio padronizado indica em quantas unidades de desvio-padrão a medida se afasta da média. Um valor de Z = 0 indica que a medida é igual ao valor médio; Z = 1 significa que a medida é a média mais 1 desvio-padrão; Z = 2 é a média mais dois desvios-padrão, e assim por diante. Valores negativos de Z correspondem a medidas menores que o valor médio esperado. A Figura 20 ilustra como uma duração de 100ms pode ter interpretação diferente dependendo de sua variabilidade de cada vogal[28]. Essa duração é mais curta que a média (Z = -0,7) para a vogal [a], e é mais longa que a média (Z = +0,7) para a vogal [i], cujas médias são, respectivamente 132 e 87ms. O Z escore também mostra que para a vogal [a], dado seu desvio-padrão de 45ms, 100ms está numa região de maior probabilidade de ocorrência. Já para a vogal nasal [uN], que dura, em média, quase o mesmo tempo que [a], 100ms é muito mais curto, em função de sua menor variância, resultando em um Z = -2,8.

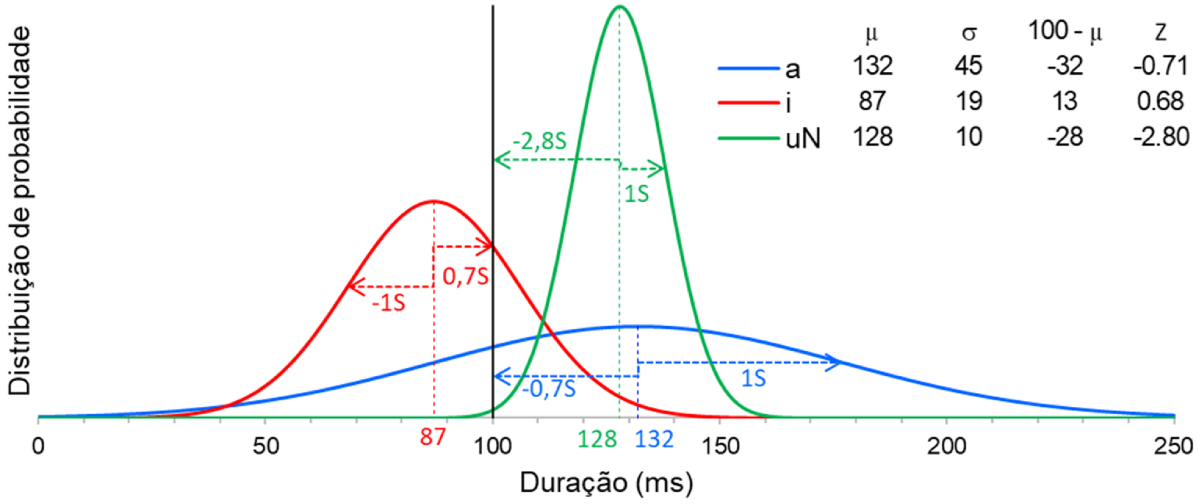


Figura 20: Ilustração do escore Z de 100ms para vogais com duração intrínseca diferentes.

A medida de duração normalizada da UD foi realizada com o auxílio de uma adaptação do *script SGDetector* (BARBOSA, 2013). O procedimento calcula a duração normalizada de sílabas a partir dos valores tabelados de média e desvio-padrão para o português falado (Tabela 9). A média esperada da sílaba é estimada pela soma das médias de seus segmentos e o desvio-padrão calculado considerando a variância da soma das durações igual a soma das variâncias:

[28] Assumindo que as variações de duração da vogal se distribuem segundo uma função gaussiana.

$$\mu_{sil}^{D} = \sum \mu_{i}^{D} \tag{5.4}$$

$$v_{sil}^{D} = \sum v_{i}^{D}$$

$$\left(\sigma_{sil}^{D}\right)^{2} = \sum \left(\sigma_{i}^{D}\right)^{2}$$

$$\sigma_{sil}^{D} = \sqrt{\sum \left(\sigma_{i}^{D}\right)^{2}} \tag{5.5}$$

Onde:

μ_{sil}^{D} *e* μ_{i}^{D}= média de duração da sílaba e do segmento *i*;

v_{sil}^{D} *e* v_{i}^{D}= variância da duração da sílaba e do segmento *i*;

σ_{sil}^{D} *e* σ_{i}^{D}= desvio-padrão da duração da sílaba e do segmento *i*.

A partir da média e desvio-padrão para a sílaba, o desvio padronizado (Z) é calculado pela equação (5.3).

A duração normalizada da UT foi estimada de duas formas: i) como a média aritmética dos escores Z de suas sílabas e ii) a partir da soma das médias e das variâncias de suas sílabas:

$$Z_{m_sil}^{UT} = \frac{\sum Z_{sil}^{D}}{N_sil} \tag{5.6}$$

$$Z^{UT} = \frac{\sum Dur_{sil_i} - \sum \mu_{sil_i}^{D}}{\sqrt{\sum \left(\sigma_{sil_i}^{D}\right)^{2}}} \tag{5.7}$$

Onde:

$Z_{m_sil}^{UT}$= duração silábica normalizada média;

Z^{UT} = duração normalizada da unidade tonal.

A equação (5.7) é combinações das equações (5.3), (5.4) e (5.5), de modo que o cálculo da duração normalizada feito pela da soma de suas sílabas é igual ao cálculo realizado pela soma de seus segmentos individuais.

Os três parâmetros *Dm_sil*, *Z_ut* e *Zm_sil* representam tentativas de estimar a mesma propriedade: a taxa de elocução. Eles foram considerados, apesar da redundância, para que o estimador mais conveniente seja revelado através das análises estatísticas. O mesmo acontece entre alguns dos parâmetros que serão descritos a seguir.

Tabela 9*: Média e desvios padrões dos segmentos utilizados para o cálculo da duração padronizada*[29]

ASCII	IPA	média	desvio	ASCII	IPA	média	desvio	ASCII	IPA	média	desvio	ASCII	IPA	média	desvio
a	a	132	45	A	ɐ	59	15	aN	'aN	134	15	p	p	80	16
e	e	110	19	E	ɐ̣	51	13	eN	'eN	134	15	t	t	81	20
eh	ɜ	113	17	I	ɪ	55	14	iN	'iN	128	16	k	k	80	22
i	i	87	19	O	o̞	51	14	oN	'oN	137	18	b	b	59	18
o	o	111	20	U	ʊ	47	15	uN	'uN	128	10	d	d	58	20
oh	ɔ	119	18	EU	ɐ̣ʊ̯	106	14	AN	ɐN	93	34	g	g	52	20
u	u	103	14	IA	ɪɐ	80	17	EN	eN	82	26	f	f	89	28
aI	aɪ̯	135	12	IE	ɪe	49	14	IN	iN	89	25	s	s	96	24
aU	aʊ̯	145	15	II	ɪɪ̯	49	14	ON	oN	88	19	sh	ʃ	104	19
ehI	ɜ	134	13	IO	ɪo̞	61	16	UN	uN	98	16	v	v	57	19
ehU	ɜʊ̯	143	15	Io	ɪo	131	14	aNI	aNɪ̯	143	18	z	z	64	25
eI	eɪ̯	126	13	IU	iʊ̯	61	16	aNU	'aNʊ̯	140	13	zh	ʒ	65	17
eU	eʊ̯	134	17	UA	ʊɐ	70	14	oNI	oNɪ̯	120	26	m	m	61	13
iU	iʊ̯	145	23	UE	ʊe	106	14	ANU	ɐNʊ̯	96	10	n	n	51	17
oI	oɪ̯	131	14	UI	ʊɪ̯	106	18	uNI	uNɪ̯	120	10	nh	ɲ	84	27
oU	oʊ̯	133	14	UO	ʊo	63	18	eNI[30]	eNɪ̯	134	15	l		47	13
ohI	ɔɪ̯	133	12	UU	ʊʊ̯	63	16					lh	ɬ	72	24
ohU	ɔʊ̯	138	19									r	r	78	24
uI	uɪ̯	122	13									R	ɾ, ɹ, ɽ	30	10
uU	uʊ̯	145	23									S	s (coda)	74	31

A padronização minimiza os efeitos da duração intrínseca e da variabilidade normal dos segmentos sobre a duração da UD, mas ainda se faz necessário tornar os valores comparáveis levando-se em conta as particularidades de cada contexto enunciativo. Para tanto, as durações foram calculadas em relação à unidade de referência, ou seja, o comentário. A duração silábica média da UD em relação a referência foi calculada pela diferença proporcional conforme a equação (5.1); para as durações

[29] Barbosa (2006, p.489)

[30] Mesmos valores da vogal nasal [eN].

normalizadas, entretanto, o parâmetro foi calculado pela diferença absoluta com a unidade de referência:

$$Z_{m_sil} = Z^{UD}_{m_sil} - Z^{Re}_{m_sil} \quad (5.8)$$

$$Z_{ut} = Z^{UD} - Z^{Re} \quad (5.9)$$

A diferença absoluta foi usada no lugar da diferença proporcional porque o valor da duração normalizada no comentário pode ser 0 ou próximo a 0, o que levaria a valores de diferenças proporcionais muito elevados, ou até ao erro (divisão por zero). Por exemplo, se a UD e o comentário apresentarem valores de Z_{UT} iguais a 0,1 e 0,001, respectivamente (taxas de elocução próximas à média de referência), o valor da diferença proporcional pela equação (5.3) será 99 (duração da UD 9900% maior que o comentário).

5.4.2 Parâmetros de intensidade

As medidas de intensidade foram coletadas através de procedimento automatizado no Praat utilizando os dados de alinhamento das fronteiras tonais dos *TextGrids* segmentados. Quatro parâmetros de intensidade foram estimados pelas medidas de intensidade:

vi) *In*: a intensidade média;
vii) *Insd*: o desvio-padrão da intensidade;
viii) *Inmax*: a intensidade máxima;
ix) *Inmin*: a intensidade mínima.

O *software* calcula a intensidade a partir da análise das amplitudes das ondas sonoras. De acordo com Boersma (2013), para cada ponto da onda sonora, o valor de pressão é dividido pelo limiar de pressão detectável pelo ouvido humano (2×10^{-5}Pa) elevado ao quadrado, obtendo um valor adimensional proporcional à potência e à intensidade sonora. Em seguida a curva é suavizada para eliminar as flutuações periódicas da onda, ponderando os valores no entorno com pesos proporcionais a uma função gaussiana. O nível de intensidade acústica, em dB, é então obtido tomando-se o logaritmo decimal e multiplicando-o por 10 (constante convencional da unidade).

Os valores absolutos obtidos no Praat não correspondem aos valores reais captados pelo microfone, a não ser que este tenha seu ganho calibrado. Mesmo calibrado, a intensidade no microfone é tanto menor quanto maior for a distância da fonte, pois a energia sonora se dissipa no ar. Por essa razão é sempre necessário analisar a intensidade em relação a outros valores de intensidade na mesma gravação (BOERSMA 2013, p. 380). Ainda assim é preciso garantir que não houve movimento do falante em relação ao microfone e que o controle de ganho não foi alterado entre esses dois pontos comparados[31].

Em função dos critérios sob os quais o C-ORAL-BRASIL foi arquitetado, com gravações em ambientes os mais naturais possíveis, a satisfação dessas condições nem sempre são garantidas. Entretanto, como foi utilizado o comentário do mesmo enunciado como referência, sempre a alguns poucos segundos de diferença no tempo, é razoável admitir que as medidas são comparáveis.

As medidas foram obtidas criando-se um objeto *intensity* no Praat, que gera os pontos de intensidade a partir do arquivo de som. O objeto requer a definição das seguintes especificações (BOERSMA e WEENINK, 2018):

- *Mínimo de f0* (Hz): define a mínima frequência do sinal acústico. É usada para determinar o tamanho da janela de análise, igual a 3,2 dividido pela f0 mínimo (ex.: para 80Hz, a janela será 0,25s). Se o valor for alto, o intervalo de tempo será pequeno, poucos pontos vizinhos serão considerados na inferência e, como consequência, a suavização será menor. Se o valor for muito baixo, muitos pontos vizinhos serão considerados; a suavização pode ser excessiva e ocultar variações relevantes para a análise. Foi utilizado um valor para cada enunciado, igual à f0 mínimo usado na análise de periodicidade que será descrita na próxima seção;
- Intervalo de tempo (s): o intervalo de tempo entre os pontos da curva. Foi utilizado parâmetro 0, que leva ao cálculo do intervalo de 0,8 dividido pela f0 mínimo (ex. para 80Hz, a janela será 0,01s);

[31] A medida de ênfase espectral (TRAUNMÜLLER e ERIKSSON, 2000) é uma medida de esforço vocálico, alternativa à intensidade, independente do tipo e da posição do microfone.

- Subtrair média (lógico): Remove o *DC offset* (Polarização de corrente contínua), ruído de fundo de baixa frequência resultante do processo de captação e processamento do som. Essa opção foi selecionada.

Com o objeto *intensity* criado, as medidas são tomadas, para cada unidade, entre o tempo inicial e final estabelecido no alinhamento das fronteiras da UD. Foram utilizados os seguintes métodos disponíveis no *software*: a intensidade média foi calculada utilizando como método a integral da curva de energia; os pontos máximos e mínimos foram calculados pelo ajuste de uma parábola aos pontos vizinhos aos máximos e mínimos locais.

5.4.3 Parâmetros de f0

As medidas de *f0* das unidades foram obtidas com o software Praat e utilizadas para o cálculo dos seguintes parâmetros:

x) *f0*: f0 média;
xi) *f0sd*: desvio-padrão de f0;
xii) *f0max*: f0 máxima;
xiii) *f0min*: f0 mínima.

Foram elaboradas instruções na linguagem do aplicativo para a coleta em cada um dos trechos sonoros correspondentes aos enunciados estudados. No entanto, o procedimento não foi totalmente automatizado, incluindo etapas de intervenção manual para a análise visual da curva de f0, ajuste dos critérios de análise de periodicidade e correções pontuais da curva.

O método de análise de periodicidade utilizado foi a autocorrelação, que se baseia na avaliação do coeficiente de correlação[32] de um valor instantâneo do sinal acústico com valores do mesmo sinal defasados por diferenciais de tempo. Se o sinal é periódico, os valores serão correlacionados em intervalos de tempo equivalente ao período da vibração. O procedimento foi realizado no Praat pela função *To Pitch*[33] *(ac)*, que cria

[32] Ver seção 6 para mais informações sobre o coeficiente de correlação.

[33] A rigor, a análise de periodicidade realizada pelo Praat estima a frequência fundamental da onda sonora (f0), uma propriedade física que é correlacionada, no caso da fala, com a frequência de vibração

um objeto *pitch*, que contém todos os valores de f0 estimados para cada ponto de tempo, bem como todos os demais candidatos. A função requer as seguintes especificações[34] (valores padrões entre parênteses) (BOERSMA e WEENINK, 2018):

- Mínimo de f0[35] (75Hz): além determinar o limite mínimo, é utilizado na determinação do tamanho da janela de processamento do sinal, calculada para conter 3 períodos. Para 75Hz a janela é de 3÷74Hz = 0,04s.
- Intervalo de tempo (0,75÷mínimo de f0): o intervalo de tempo entre as medidas, ou entre os pontos da curva resultante (0,01s para 75Hz);
- Máximo número de candidatos (15): máximo número de candidatos considerados para cada ponto no tempo;
- Muito acurado (não): Se selecionado, altera o formato da janela de analise para uma curva gaussiana. Caso contrário, é usada a forma da função de Hann.
- Limiar de silêncio (0,03): pontos que não possuírem amplitude acima desse limite, em relação à amplitude global máxima, são considerados silenciosos;
- Limiar de vozeamento (0,45): a força do candidato desvozeado, relativa à máxima correlação possível. Quanto maior este valor, mais candidatos serão desclassificados por serem considerados desvozeados;
- Custo de oitava (0,01): estabelece um fator de favorecimento para os candidatos com frequências mais altas;
- Custo do salto de oitava (0,35): grau de desfavorecimento nas mudanças de f0, em relação à máxima correlação;
- Custo voazeado/não-vozeado (0,14): grau de desfavorecimento das transições de vozeamento;
- Máximo de f0 (600Hz): candidatos com frequência acima desse valor são descartados;

---

das cordas vocais. Já o *pitch* é o fenômeno psicológico de percepção de altura de um som pelo ouvinte, fortemente relacionado à frequência fundamental do som. A f0 é a forma mais usual de estimar o *pitch.*

[34] O termo *parâmetro*, mais apropriado, foi evitado pelo uso frequente com significado diferente neste trabalho.

[35] Nomes originais dos parâmetros: *pitch ceiling; pitch floor; time step; max. number of candidates; very accurate; silence threshold; voicing threshold; octave cost; octave-jump cost* e *voiced/unvoiced cost*.

As primeiras quatro especificações são usadas para estimar os valores candidatos de f0; os demais entram no algoritmo de definição do caminho da curva, ou seja, na escolha do melhor candidato ou de nenhum deles (trechos desvozeados). Os detalhes acerca do procedimento utilizado podem ser encontrados em Boersma (1993).

Os limites de f0 são as especificações mais importantes para a curva de f0, pois devem ser ajustadas para se adequar à amplitude de frequência de cada falante. Para a determinação de uma estimativa inicial, foi utilizado o procedimento proposto por Hirst (2007), pelo qual primeiro cria-se uma curva de f0 com um amplo *range* e, a partir dela, estimam-se as especificações de mínimo e máximo de f0 pelos 1º e 3º quartis das frequências, respectivamente:

$$Min_{f0} = 0{,}75Q_1 \tag{5.10}$$

$$Max_{f0} = 1{,}5Q_3 \tag{5.11}$$

Onde:

Min_{f0} e Max_{f0}= máximo e mínimo de f0;

Q_1 e Q_3= primeiro e terceiro quartis.

Foram utilizados os valores de 50Hz e 750Hz, intervalo de tempo 0,01 e os valores padrões sugeridos pelo Praat para as demais especificações para a estimativa de *Min*$_{f0}$ e *Max*$_{f0}$. Em seguida, foi produzido um novo objeto de *pitch* com os limites calculados.

As curvas foram inspecionadas visualmente, uma a uma, e suas especificações redefinidas com objetivo de obter uma curva de f0 aceitável pelo menos dentro dos limites das UDs e de suas unidades de referência. As alterações mais frequentes foram realizadas nos limites de f0, principalmente quanto ao mínimo de f0. De modo geral o procedimento de Hirst superestima esses valores. O limiar de vozeamento padrão foi modificado em 17% dos enunciados; as demais especificações foram alteradas pontualmente em poucos casos particulares. Todas os valores utilizados na criação do objeto de *pitch* estão registrados no Apêndice B.

Além do ajuste das especificações, também foram realizadas as alterações manuais dos melhores candidatos para a curva de f0 e a alteração do *status* de vozeamento (remoção do valor de f0 estimado) de segmentos não vozeados que apresentavam

valores de f0 claramente anômalos. Um exemplo de alteração manual pode ser visualizado no exemplo a seguir. A Figura 21 mostra os pontos selecionados automaticamente, com o trecho correspondente à UD na região rosa, na qual há um claro salto de oitava. A Figura 22 mostra o resultado da correção feita manualmente. Antes dessa intervenção, duas alternativas de ajuste de especificações foram testadas, insatisfatoriamente: i) a elevação do mínimo de f0, com resultado equivalente na escolha dos pontos mais altos, porém com a perda do movimento descendente final da unidade de comentário; ii) a elevação do custo de salto de oitava, que resultou no abaixamento dos pontos iniciais da curva, mantendo o salto 150Hz após a fronteira da UD neste exemplo.

Exemplo 27: bfammn02[009]

DFL: e então /=DCT= tinha muito texto do tio Carlos /=COB= então ele falava /=INT= **ah /=AUX_r= ele é tio da minha tia //=COM_r=**

[009]

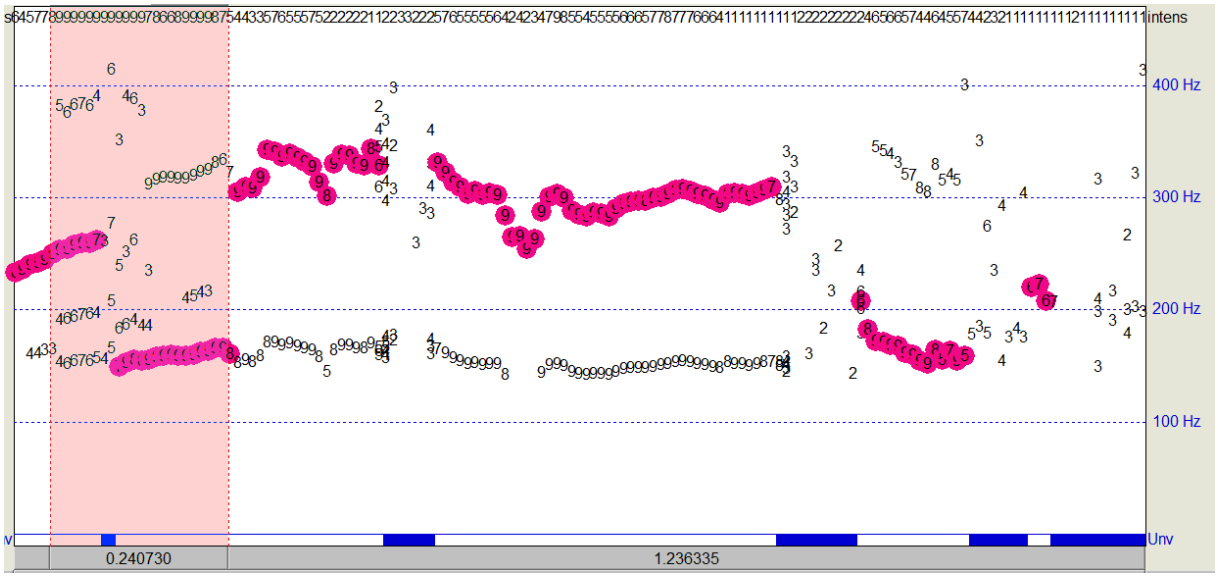


Figura 21: Objeto pitch originalmente gerado (limites entre 140 e 450Hz).

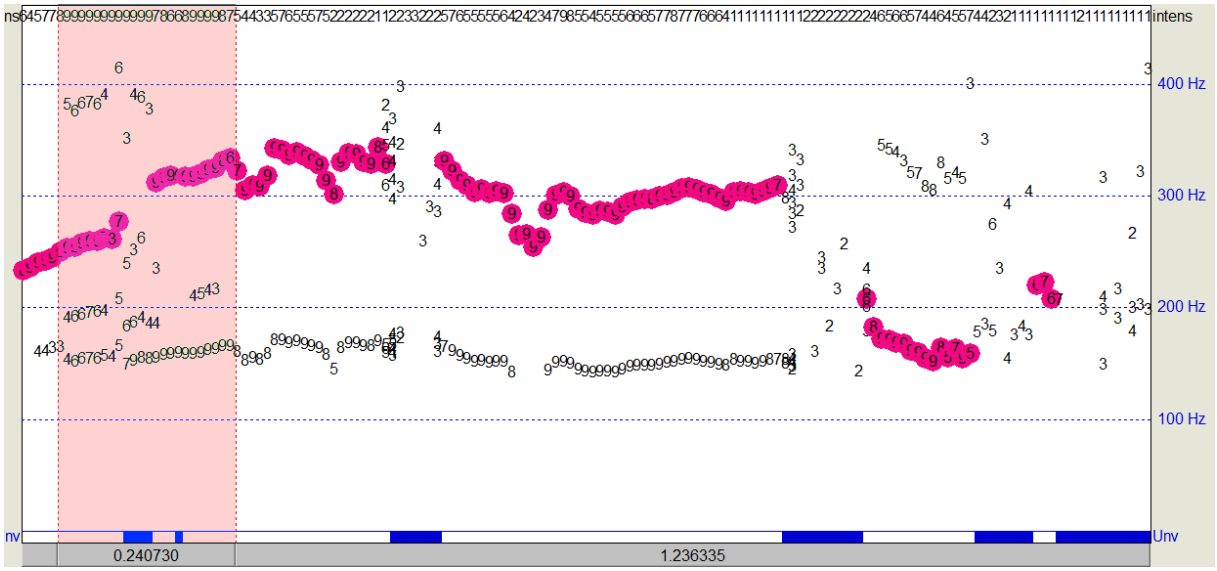


Figura 22: Objeto pitch manualmente ajustado.

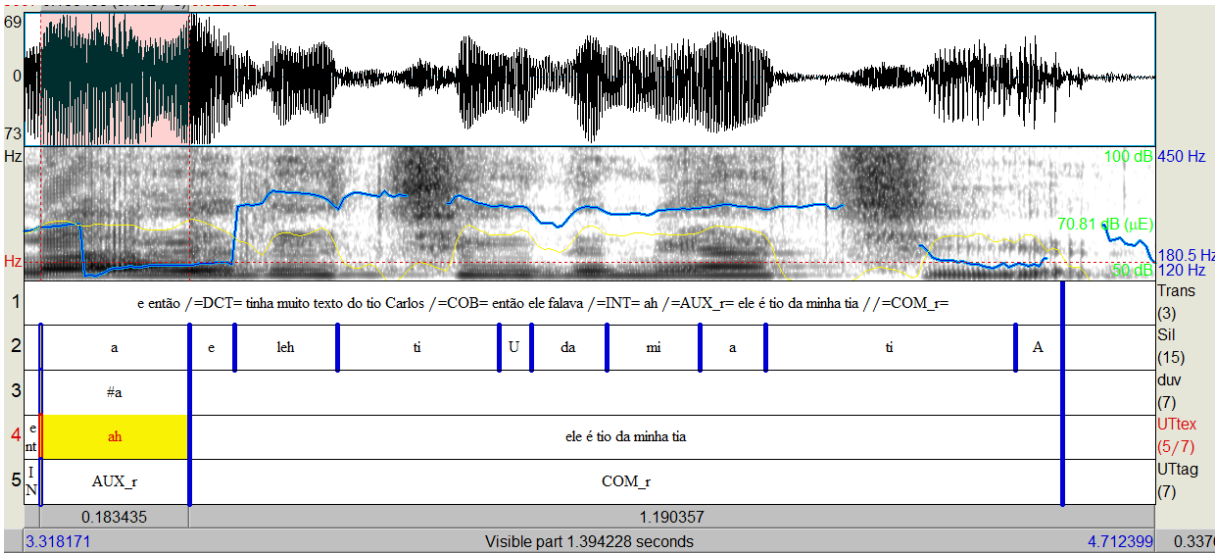


Figura 23: TextGrid com espectrograma de parte do enunciado bfammn02[009].

O ajuste das especificações e dos melhores candidatos manualmente é uma etapa indispensável, porém custosa pelo tempo requerido e indesejada por adicionar um nível de subjetividade às medidas. A decisão dos valores especificados e da curva de f0 são, na maioria das vezes, escolhas difíceis; frequentemente mais de uma alternativa parece razoável. Em muitos casos, nenhum ajuste foi capaz de gerar uma curva visualmente aceitável. Esse motivo levou ao descarte de cerca de 40 enunciados.

Com a curva de f0 definida, as medidas de f0 foram extraídas através de *script* em linguagem do Praat: a f0 médio, o desvio-padrão, o máximo e o mínimo de f0 de cada

unidade tonal. Os valores foram tomados em semitons[36] (st) com referência 1Hz, para amenizar diferenças entre os falantes e para que todos os dados fossem positivos[37], evitando divisão por zero (ou valor próximo a zero) no cálculo da diferença proporcional, conforme a equação (5.1), utilizada nos quatros parâmetros f0.

5.4.4 Parâmetros de variação de f0

As medidas descritas a seguir são calculadas a partir dos dados de f0 obtidos pelos procedimentos descritos na seção anterior. Foram determinados 4 parâmetros para estimar a variação de *f0* da UD:

xiv) *f0exc*: excursão de f0;
xv) *f0slo*: taxa de variação de f0, em st/s;
xvi) *if0sit*: variação de f0 do início da UD até a vogal tônica, em st;
xvii) *if0stf*: variação de f0 da vogal tônica ao fim da UD, em st;

A excursão de f0 indica a amplitude do movimento de f0 na UT. É dada por:

$$f0_{exc}^{UT} = f0_{max}^{UT} - f0_{min}^{UT} \quad (5.12)$$

Para este parâmetro foi aplicada a diferença proporcional em relação à unidade de comentário, segundo a equação (5.1), ao contrário dos demais parâmetros de variação, para os quais foram analisados os valores absolutos.

A taxa de variação de f0 (*f0 slope)* é a razão entre a excursão de f0 e o intervalo de tempo em que ela ocorre:

$$f0_{slo} = \frac{f0_{exc}^{UT}}{f0T_{max}^{UT} - f0T_{min}^{UT}} = \frac{f0_{max}^{UT} - f0_{min}^{UT}}{f0T_{max}^{UT} - f0T_{min}^{UT}} \quad (5.13)$$

Onde:

$f0T_{max}^{UT}$= Tempo do máximo de f0 (s)

$f0T_{min}^{UT}$= Tempo do mínimo de f0 (s)

---

[36] A frequência em semitons é igual a 12 vezes o logaritmo de base 2 da frequência em Hz. Ex.: 64, 128, 256 e 512Hz (2^6; 2^7, 2^8 e 2^9) equivalem a 72, 84, 96 e 108st (12×6, 12×7 e 12×8), respectivamente.

[37] Na faixa de f0 da fala humana. Os valores em semitons (1Hz) são negativos para f0 menor que 1Hz.

O valor de *f0 slope* estima a taxa de variação da amplitude da curva melódica dentro da UD. Quanto maior, mais brusca é a variação, mesmo que não necessariamente implique em um movimento amplo. Valores positivos e negativos indicam movimento ascendente e descendente de f0, respectivamente.

Os movimentos de f0 dentro das UDs são geralmente simples. É é raro que apresentem mais de um pico ou vale. O mesmo não ocorre com a unidade de comentário, que pode ser longa o suficiente para conter dois ou mais movimentos. Por essa razão, os pontos de mínimo e de máximo globais do comentário podem estar muito distantes no tempo e pertencer a movimentos diferentes, resultar em valores de pouco significado calculado pela equação (7.1). Esta é a razão para o uso do valor absoluto da UD ao invés do valor relativo. Por outro lado, o uso do valor absoluto implica menor comparabilidade pois, mesmo a medida tendo sido tomada em semitons, as amplitudes das curvas melódicas tendem a ser maiores, em valor absoluto, em falantes com tom de voz mais alto, sem que essa diferença seja comunicativamente significativa.

Apesar do curto tempo de duração da UD, Raso e Vieira (2016) observaram que o movimento de descida ou subida de f0 é relevante para a distinção formal das UDs. Para tentar traduzir os movimentos melódicos distintos que possam ocorrer até a sílaba tônica e depois dela, os seguintes parâmetros foram definidos:

$$if0s_{it} = \frac{if0_t^{UT} - if0_i^{UT}}{t_c^T - t_i^{UT}} = \frac{if0_t^{UT} - if0_i^{UT}}{\left(T_c^T - T_i^{UD}\right)/Dur^{UT}} \qquad (5.14)$$

$$if0s_{tf} = \frac{if0_f^{UT} - if0_t^{UT}}{t_f^{UT} - t_c^T} = \frac{if0_f^{UT} - if0_t^{UT}}{1 - \left(T_c^T - T_i^{UD}\right)/Dur^{UT}} \qquad (5.15)$$

Onde:

$if0s_{it}$ = Taxa de variação de f0 do início da UT até o centro da vogal tônica;

$if0s_{tf}$= Taxa de variação de f0 do centro da vogal tônica ao fim da UT;

$if0_i$ = f0 no início da UT;

$if0_t$ = f0 no centro da vogal tônica;

$if0_f$ = f0 no final da UT;

T_c^T = Tempo do centro da tônica (s);

T_i^{UD} = Tempo de início da UT (s);

$t_c^T = (T_c^T - T_i)/Dur^{UT}$ = Tempo até o centro da tônica relativo à duração da UT;

t_i^{UT} = Tempo inicial relativo à duração da UT, igual a 0;

t_f^{UT} = Tempo final relativo à duração da UT, igual a 1;

A figura a seguir ilustra a informação representada pelo parâmetro:

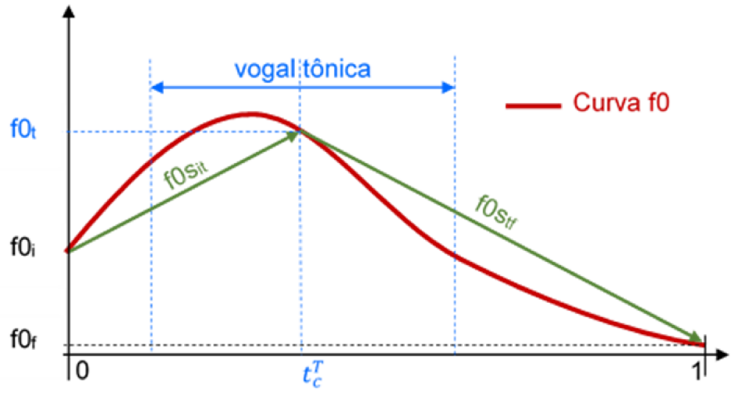


Figura 24: Esquema da definição dos parâmetros de variação de f0 em relação à tônica.

O índice '*i*' antes de f0 indica que, para esses parâmetros (apenas), as medidas de f0 utilizadas foram obtidas a partir da interpolação dos pontos do objeto *pitch*, de modo a garantir a existência de valores em todos os pontos da curva ao longo da duração da UD. A curva também foi suavizada, usando um filtro com largura de banda de 4Hz.

5.4.5 Parâmetros de alinhamento

As medidas de alinhamento tentam relacionar a posição dos extremos de intensidade e de f0 em relação à duração da UD e em relação à posição de sua sílaba tônica. Ao todo foram determinados 8 parâmetros para cada UD:

xviii) *Intmax*: Tempo do máximo de intensidade em relação à duração da UT;
xix) *Intmin*: Tempo do mínimo de intensidade em relação à duração da UT
xx) *InmaxTa*: Tempo do máximo de intensidade em relação à vogal tônica;
xxi) *InminTa*: Tempo do mínimo de intensidade em relação à vogal tônica;
xxii) *f0tmax*: Tempo do máximo de f0 em relação à duração da UT;

xxiii) *f0tmin*: Tempo do mínimo de f0 em relação à duração da UT

xxiv) *f0maxTa*: Tempo do máximo de f0 em relação à vogal tônica;

xxv) *f0minTa*: Tempo do mínimo de f0 em relação à vogal tônica;

As medidas necessárias para o cálculo dos parâmetros foram coletadas no mesmo procedimento de tomada dos valores de máximo e mínimo de intensidade e f0 descritos acima. Os alinhamentos em relação à duração da tônica são dados pela expressão:

$$Xtpcr = \frac{XT_{pcr}^{UT} - T_i^{UT}}{Dur^{UT}} \quad (5.16)$$

Onde:

$Xtpcr$= Tempo do ponto crítico da medida X relativo à duração da UT.

X = Medida (f0 ou intensidade);

pcr = Ponto crítico (máximo ou mínimo);

XT_{pcr}^{UT}= Tempo do ponto crítico da medida X na UT;

T_i^{UT} = Tempo do início da UT.

Os parâmetros de tempo relativo ou de alinhamento ao longo da duração da UT podem assumir valores apenas entre 0 e 1, sendo 0 quando o ponto crítico coincide com o início da unidade e 1 quando se alinha ao final da unidade.

Os parâmetros de tempo em relação à vogal tônica são calculados por:

$$XpcrTa = \frac{XT_{pcr}^{UT} - T_i^{T}}{Dur^{T}} = \frac{XT_{pcr}^{UT} - T_i^{T}}{T_f^{T} - T_i^{T}} \quad (5.17)$$

$XpcrTa$= tempo do ponto crítico da medida X relação à vogal tônica.

T_i^T = tempo de início da vogal tônica;

T_f^T = tempo de fim da vogal tônica;

Dur^T = duração da vogal tônica.

Diferentemente dos parâmetros de tempo em relação à duração da UD, o tempo em relação à vogal tônica também pode assumir valores negativos ou maiores que 1. Valores entre 0 e 1 significam que o ponto crítico ocorreu dentro da tônica; se igual a

0 coincide com o início desta; se igual a 1, com o fim. Quando o extremo de f0 ou intensidade ocorrem antes ou após a vogal tônica, os valores dos parâmetros serão negativos ou maiores que 1, respectivamente.

6 ESTATÍSTICA DESCRITIVA E INFERENCIAL DOS PARÂMETROS ESTIMADOS

Nessa seção serão apresentados os valores dos parâmetros pela análise acústica dos enunciados que serão utilizadas como variáveis nas análises. Foram utilizados ao todo 25 parâmetros; eles serão apresentados em cinco seções, organizados por classe de medida, de forma a facilitar a descrição e a visualização:

Será descrito como as medidas foram coletadas e tratadas e será apresentado um sumário dos dados, para cada etiqueta de UD, através de estatísticas descritivas:

i) média aritmética;
ii) desvio-padrão[38];
iii) mediana;
iv) média truncada: média aritmética excluindo dados extremos (10%[39]);
v) mínimo;
vi) máximo;
vii) amplitude: diferença entre os valores máximo e mínimo;
viii) assimetria: indica se a sua distribuição possui concentração de dados à esquerda (valores negativos) ou à direita da moda (valores positivos);
ix) curtose: indica o achatamento da curva de distribuição. Curtose igual a 0 significa que a distribuição dos dados tem achatamento igual ao de uma distribuição normal; valores positivos significam que os dados são mais concentrados (ou afunilados) em torno da moda e valores negativos indicam dados mais espalhados (distribuição achatada).

A distribuição dos valores de cada parâmetro por tipo de etiqueta também poderá ser visualizada através de gráficos do tipo *boxplot*. Para cada parâmetro foi aplicado o teste não-paramétrico de Kruskal-Wallis [40] (KRUSKAL e WALLIS, 1952) para comparação da significância das diferenças entre os quatro tipos de etiquetas das UDs. O teste compara os dados ordenados transformados em uma estrutura de níveis

[38] Desvio-padrão amostral: raiz quadrada da média dos desvios entre os valores e a média dos valores dos dados elevados ao quadrado;

[39] Os 5% menores e os 5% maiores;

[40] Calculado pela função *kruskalmc* do pacote *pgirmess* versão 1.6.9 para linguagem R (GIRAUDOUX at al, 2018).

e avalia a probabilidade de os grupos pertencerem à mesma população, para um determinado critério de significância. A estatística foi escolhida por não exigir o pressuposto de normalidade na distribuição dos dados e ser menos afetada pela presença de valores extremos, casos encontrados entre os parâmetros que serão apresentados. No entanto, a transformação aplicada para permitir a comparação pode causar perda de informação sobre a estrutura interna dos dados, por essa razão alguns resultados devem ser interpretados com cautela. Os resultados entre os pares de etiquetas de UD serão apresentados na forma de tabelas que indicam se a diferença foi considerada significativa para o nível de significância de 5 %.

As tendências e diferenças significativas entre os pares de grupos de dados serão descritas e brevemente comentadas. É importante observar que apenas os grupos ALL, CNT e INP contém dados que foram classificados por conterem características comuns, enquanto os dados etiquetados como AUX não constituem uma categoria; são apenas dados reunidos por não se enquadraram nas outras 3 categorias.

A interação entre os parâmetros foi analisada pelo coeficiente de correlação de Pearson (1895). O coeficiente (*r*) mede o quanto dois conjuntos de dados estão linearmente relacionados entre si, e mede a orientação dessa relação, numa escala de -1 a 1. Valores próximos a 1 significam forte relação positiva (quando uma variável aumenta a outra também aumenta); valores próximos a -1 indicam forte relação negativa; variáveis pouco relacionadas entre si tem coeficientes de correlação próximo a 0). Dados de variáveis com alto coeficiente de correlação (em valor absoluto) se concentram formando uma linha reta quando projetados em um gráfico de dispersão.

As correlações poderão ser visualizadas por pares, dentro de cada grupo de parâmetros, através de diagramas da matriz de correlação gerados com o auxílio da função *chart.Correlation* do pacote *PerformanceAnalytics* para linguagem R (PETERSON et al., 2018). Além dos coeficientes, também é possível visualizar no diagrama o gráfico de dispersão entre pares; o histograma de cada de parâmetro e o nível de significância da correlação (uma correlação pode ser forte, mas não significativa, ou vice-versa).

Todas as análises estatísticas descritivas, bem como as ilustrações apresentadas, foram realizadas em linguagem de programação R utilizando a interface RStudio©. Os

dados completos para todos os parâmetros de todas as UDs estão reunidos no Apêndice B.

Ainda que a maior parte dos procedimentos utilizados tenha sido automatizada, o as informações mostradas a seguir são resultados da iteração de muitas etapas de concepção, programação e depuração dos *scripts* e de conferência manual dos cálculos, por amostragem ou por valores anômalos.

Neste processo, muitos dados tiveram que ser descartados por impossibilitarem a aplicação dos procedimentos ou por apresentarem características que impedem a confiabilidade das informações. Das 664 UDs originais selecionadas para o estudo, 560 dados foram usados para análise. 104 UDs foram descartadas pelas razões: (i) por conterem sobreposições ou ruídos muito interferentes; (ii) por impossibilidade de segmentação; (iii) por impossibilidade de ajuste aceitável da curva de *f0*; (iv) por apresentarem os mesmos problemas na unidade de referência ou (vi) por não possuírem uma unidade de referência compatível com o critério adotado. O descarte de dados inadequados não foi exaustivo nem rigoroso, propositalmente, com objetivo de preservar a quantidade nesta primeira etapa de análise dos dados. A Tabela 10 resume a quantidade de UDs que foram analisadas neste estudo, em comparação com a quantidade de UDs quantificadas após a revisão do Mini BR Informal:

Tabela 10*: Quantidade de UDs analisadas em comparação com a quantidade de UDs no* Mini BR Informal

UDs	DCT	ALL	CNT	INP	AUX	AUX_n	Total
No Mini BR Informal	173	63	84	40	477	188	1025
Não consideradas	173	0	0	0	0	188	361
Descartadas	0	12	15	4	73	0	104
Utilizadas na análise	0	51	69	36	404	0	560

Ao final deste capítulo, na Tabela 24, encontra-se a lista com a descrição resumida de todos os parâmetros estimados, incluído a indicação da unidade e qual o tipo de valor adotado para cada parâmetro: se foi usado o valor calculado para a própria UD, a diferença absoluta com a unidade de comentário ou a diferença proporcional com a unidade de comentário.

6.1 Parâmetros de duração

A Tabela 12 sumariza os dados obtidos para os cinco parâmetros de duração. A distribuição dos dados pode ser visualizada na Figura 25. A Tabela 11 mostra os resultados do teste de Kruskal-Wallis par a par com os parâmetros de duração e, a Figura 27, o diagrama de correlações.

A grande maioria das UDs possui apenas uma sílaba, cerca de 80% do total (446/560) e, muito poucas UDs superam 3 sílabas. As UDs etiquetadas como AUX, INP e CNT reproduzem essa tendência com médias de 1,11; 1,22 e 1,56 sílabas por unidade, respectivamente. O ALL não acompanha o padrão e apresenta moda de 2 sílabas. O teste de Krustal-Wallis mostrou diferença significativa no número de sílabas para o ALL em relação às demais categorias e para o CNT em relação às UDs não classificadas.

A duração da UD em relação à unidade de referência (*Dur*) segue o padrão do número de sílabas para o ALL, CNT e INP. Em média os ALLs duram 53% menos que a referência; os CNT 71% e os INPs duram 82% menos em média. Apenas o ALL se diferencia de forma significativa das demais categorias quanto a duração da UD.

Os ALLs também apresentam duração silábica média (*Dm_sil*) significativamente maior que os INPs e CNTs, indicando que a duração não seria somente consequência do número maior de sílabas. Por se tratar de uma unidade que ocorre muito frequentemente na posição final, sua duração média é afetada por efeitos de alongamento final (BECKMAN e EDWARDS, 1990).

Quando analisados os dois parâmetros normalizados, o CNT passa a apresentar maior duração em média. Quando comparadas as duas formas de padronização, percebe-se uma maior amplitude de valores na duração da unidade (*Z_ut*) em relação a duração normalizada média (*Zm_sil*). Esses parâmetros não diferenciaram de forma significativa as categorias já classificadas, como seria esperado, especialmente para INP, cuja taxa de elocução elevada é uma de suas características.

Como era esperado, foram encontradas fortes correlações entre essas variáveis, principalmente entre as durações médias. Os dois parâmetros normalizados mostram alto grau de colinearidade (r = 0,89).

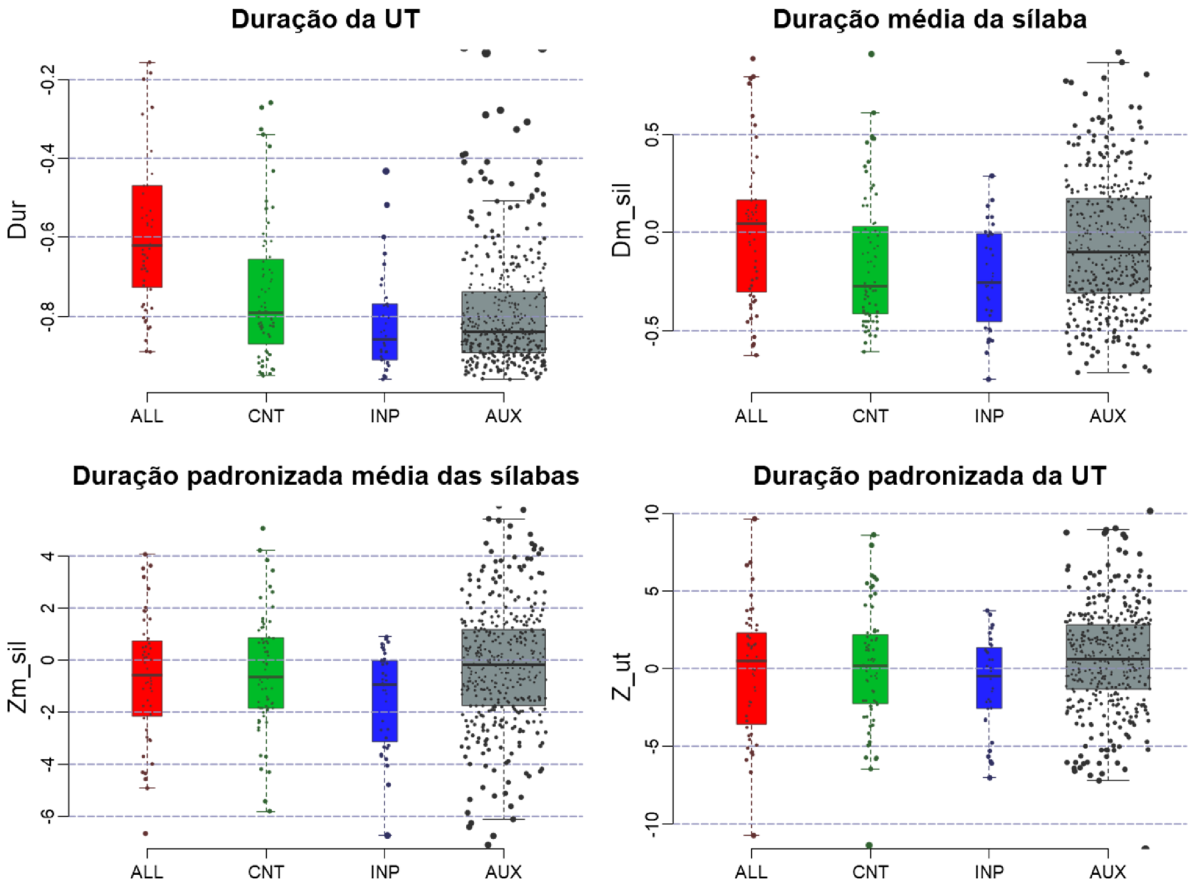


Figura 25: Distribuição dos parâmetros de duração por tipo de UD.

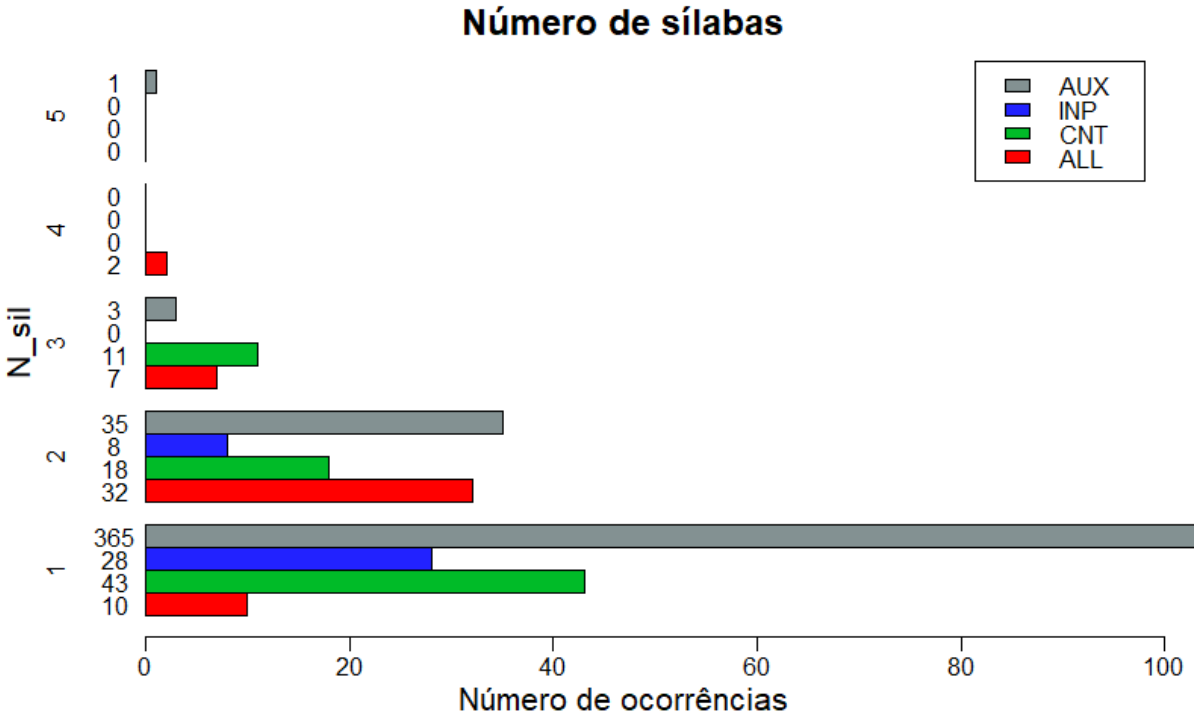


Figura 26: Distribuição do número de sílabas por tipo de UD.

Tabela 11: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de duração.

Par	N_Sil	Dur	Dm_sil	Zm_sil	Z_ut
ALL-CNT	✓	✓	✗	✗	✗
ALL-INP	✓	✓	✓	✗	✗
ALL-AUX	✓	✓	✗	✗	✗
CNT-INP	✗	✓	✗	✗	✗
CNT-AUX	✓	✓	✗	✗	✗
INP-AUX	✗	✗	✓	✓	✗

✓ diferença significativa
✗ diferença não significativa

Tabela 12: Sumário dos parâmetros de duração por tipo de UD

Parâmetro	Etiqueta	Número	Média	Desvio padrão	Mediana	Média truncada	Mínimo	Máximo	Amplitude	Assimetria	Curtose
N_sil	ALL	51	2,020	0,707	2	1,976	1	4	3	0,640	0,774
	CNT	69	1,580	0,755	1	1,491	1	3	2	0,842	-0,783
	INP	36	1,222	0,422	1	1,167	1	2	1	1,281	-0,367
	AUX	404	1,111	0,379	1	1,000	1	5	4	4,674	31,515
Dur	ALL	51	-0,530	0,344	-0,621	-0,583	-0,890	1,180	2,069	2,657	9,837
	CNT	69	-0,708	0,253	-0,787	-0,748	-0,950	0,695	1,645	2,850	11,775
	INP	36	-0,818	0,125	-0,858	-0,835	-0,959	-0,433	0,526	1,251	1,134
	AUX	404	-0,789	0,158	-0,838	-0,817	-0,959	0,124	1,083	2,183	6,282
Dm_sil	ALL	51	0,035	0,430	0,045	-0,002	-0,625	1,147	1,772	0,710	-0,147
	CNT	69	-0,118	0,406	-0,259	-0,170	-0,608	1,837	2,445	2,000	6,132
	INP	36	-0,233	0,256	-0,257	-0,236	-0,748	0,287	1,035	0,021	-1,026
	AUX	404	-0,046	0,369	-0,099	-0,075	-0,713	1,504	2,217	0,949	1,414
Zm_sil	ALL	51	-0,710	3,161	-0,560	-0,609	-12,973	7,759	20,732	-0,819	3,293
	CNT	69	-0,515	3,103	-0,358	-0,417	-11,331	7,107	18,438	-0,761	2,938
	INP	36	-1,475	1,873	-0,922	-1,332	-6,737	0,905	7,643	-0,737	-0,255
	AUX	404	-0,202	2,432	-0,177	-0,214	-13,752	6,969	20,721	-0,296	2,355
Z_ut	ALL	51	-0,473	4,755	0,490	-0,255	-18,591	9,660	28,251	-1,010	2,558
	CNT	69	0,040	3,952	0,475	0,187	-12,588	8,627	21,215	-0,541	0,970
	INP	36	-0,886	2,895	-0,459	-0,762	-7,015	3,744	10,759	-0,440	-0,805
	AUX	404	0,585	3,294	0,630	0,641	-14,634	10,163	24,797	-0,305	1,179

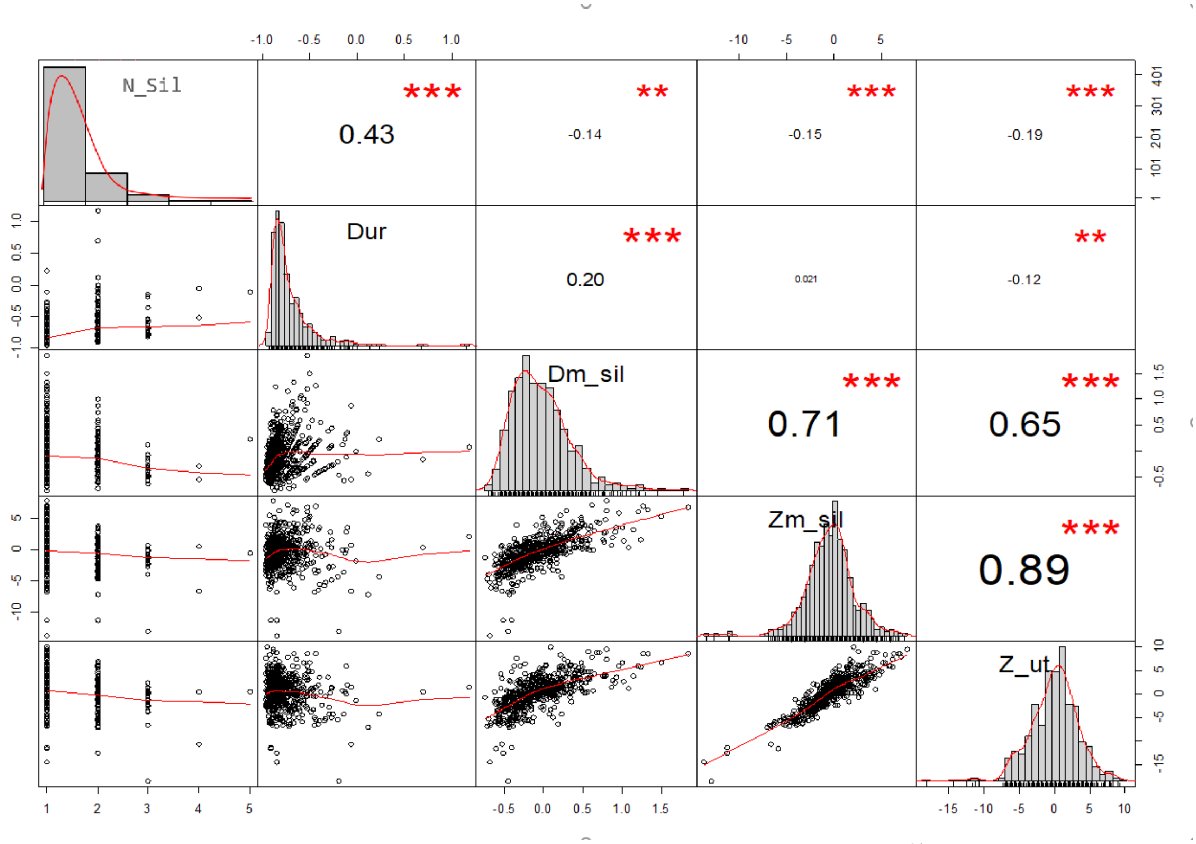


Figura 27: Correlações entre os parâmetros de duração[41].

6.2 Parâmetros de intensidade

O sumário das distribuições e sua visualização gráfica, por categoria de UD, são exibidos na Figura 28 e na Tabela 14. A Tabela 13 contém o resultado do teste de significância da diferença entre os grupos e a Figura 29 exibe as correlações entre os parâmetros.

Foram observadas diferenças significativas de intensidade média entre as quatro categorias cujas médias seguem a ordem ALL, AUX, CNT e INP. Esta última possui intensidade maior que o comentário (média positiva); o CNT apresenta média próxima a 0 e as intensidades são reduzidas para ALL e AUX.

O desvio-padrão de intensidade da UD em comparação com a unidade de comentário (*Insd*) é menor e diferente significativamente apenas entre a o grupo AUX e os demais.

[41] Os asteriscos indicam o nível de significância da correlação. ***) p-valor < 0,001; **) entre 0,001 e 0,01; *) entre 0,01 e 0,05; •) entre 0,05 e 0,1.

O desvio-padrão da intensidade é, via de regra, menor na UD do que na unidade ilocucionária.

A intensidade máxima diferencia significativamente dois supergrupos[42] de etiquetas, CNT-INP e ALL-AUX, com o valor médio destes menores que os do primeiro. Em geral, a maioria das unidades tem pico de intensidade menor que a unidade de referência. A intensidade mínima destaca significativamente o ALL, com valores inferiores às demais categorias. Os mínimos de intensidade são a maioria positivos, o oposto do máximo, corroborando com a observação de menor variância na UD.

Observa-se muito forte correlação entre o a intensidade máxima e o máximo de intensidade (r = 0,90) e correlação negativa entre o desvio-padrão e o mínimo de intensidade (r = -0,60). Também há alguma correlação significativa positiva entre as demais variáveis, exceto entre *In* e *Insd*.

[42] O termo foi usado para dizer, de forma simplificada, que há diferença significativa entre as duas categorias de um supergrupo com as categorias do outro, e que não há diferença significativa entre as categorias do mesmo supergrupo. Não implica que existam relações entre as categorias do mesmo supergrupo.

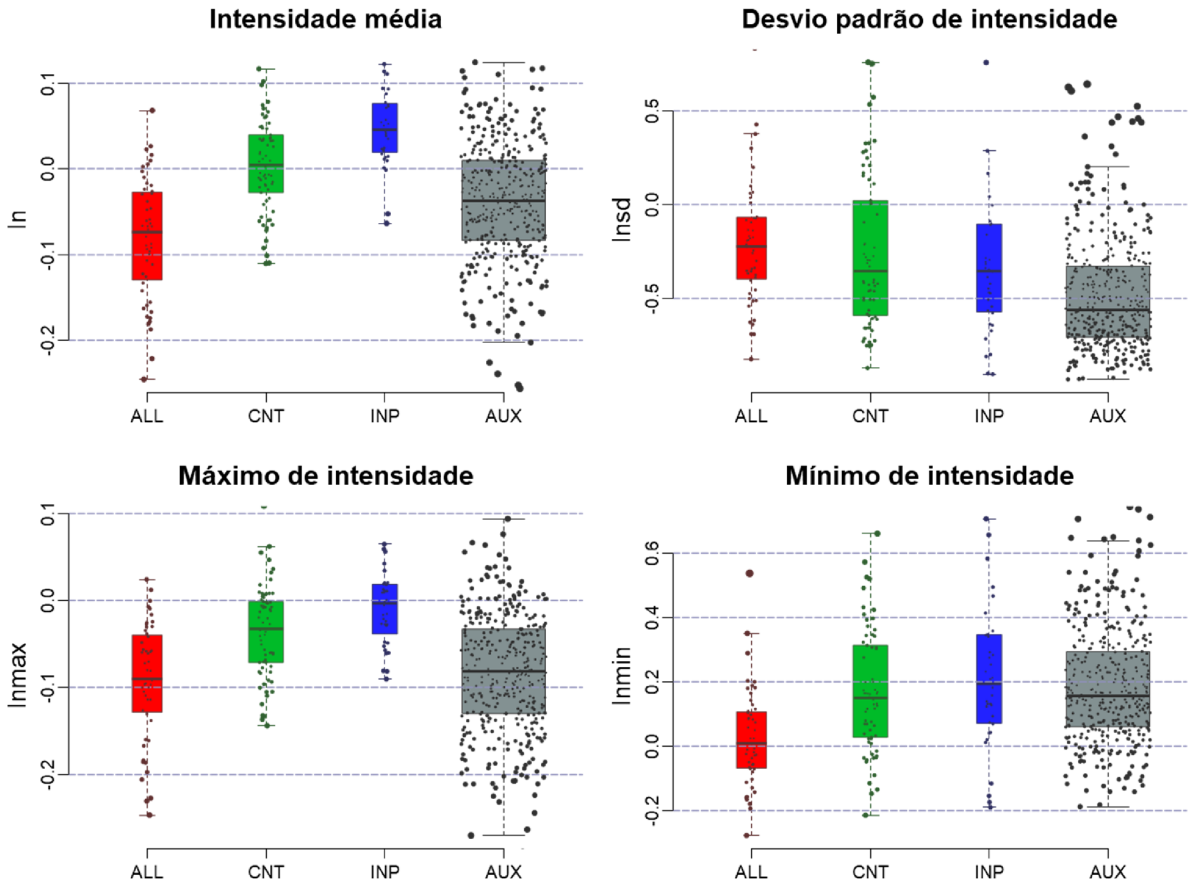


Figura 28: Distribuição dos parâmetros de intensidade por tipo de UD.

Tabela 13*: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de intensidade.*

Par	In	Insd	Inmax	Inmin
ALL-CNT	✓	✗	✓	✓
ALL-INP	✓	✗	✓	✓
ALL-AUX	✓	✓	✗	✓
CNT-INP	✓	✗	✗	✗
CNT-AUX	✓	✓	✓	✗
INP-AUX	✓	✓	✓	✗

✓ diferença significativa
✗ diferença não significativa

Tabela 14: Sumário dos parâmetros de intensidade por tipo de UD

Parâmetro	Etiqueta	Número	Média	Desvio padrão	Mediana	Média truncada	Mínimo	Máximo	Amplitude	Assimetria	Curtose
In	ALL	51	-0,082	0,070	-0,074	-0,080	-0,245	0,068	0,314	-0,184	-0,685
	CNT	69	0,005	0,057	0,005	0,005	-0,110	0,190	0,300	0,230	0,302
	INP	36	0,051	0,050	0,045	0,050	-0,063	0,172	0,236	0,302	0,490
	AUX	404	-0,038	0,069	-0,037	-0,037	-0,256	0,161	0,416	-0,188	0,055
Insd	ALL	51	-0,162	0,421	-0,219	-0,218	-0,820	1,319	2,138	1,558	3,035
	CNT	69	-0,254	0,430	-0,353	-0,298	-0,867	1,073	1,940	0,953	0,246
	INP	36	-0,280	0,476	-0,354	-0,345	-0,900	1,281	2,181	1,442	2,091
	AUX	404	-0,472	0,323	-0,557	-0,515	-0,927	1,099	2,026	1,527	3,217
Inmax	ALL	51	-0,092	0,067	-0,090	-0,088	-0,247	0,024	0,271	-0,503	-0,581
	CNT	69	-0,036	0,052	-0,033	-0,036	-0,144	0,109	0,253	0,000	-0,323
	INP	36	-0,009	0,041	-0,004	-0,009	-0,090	0,065	0,155	-0,161	-0,849
	AUX	404	-0,084	0,068	-0,082	-0,082	-0,290	0,129	0,419	-0,159	-0,086
Inmin	ALL	51	0,025	0,149	0,008	0,016	-0,278	0,537	0,815	0,794	1,330
	CNT	69	0,169	0,190	0,150	0,162	-0,215	0,661	0,876	0,393	-0,477
	INP	36	0,220	0,236	0,195	0,211	-0,190	0,765	0,956	0,419	-0,270
	AUX	404	0,189	0,188	0,155	0,174	-0,188	1,017	1,205	0,899	1,203

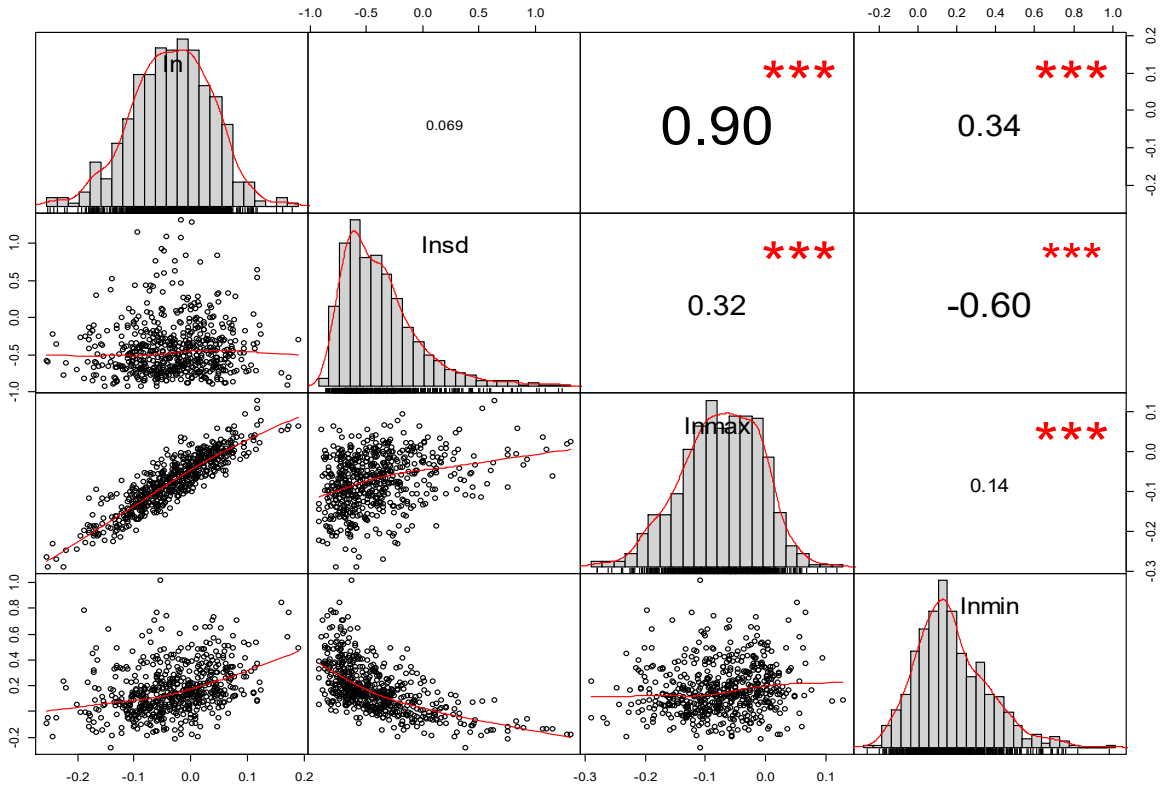


Figura 29: Correlações entre os parâmetros de intensidade.

6.3 Parâmetros de f0

Os resultados obtidos estão sumarizados na Tabela 18 e ilustrados na Figura 30. Tabela 15 indica o resultado do teste de significância da diferença entre os grupos e A Figura 33 mostra as correlações entre os parâmetros.

Podemos observar que o parâmetro f0 média (*f0*) distingue de forma significativa as categorias ALL, CNT e INP, que apresentam médias de f0 crescentes nessa ordem, em relação ao comentário. O ALL é mais baixo que o comentário (5,3%), o CNT também tem frequência menor, porém próxima à do comentário ($f0 \approx 0$) e o INP é mais alto (2,8%). O parâmetro não diferencia AUX de CNT.

Os valores de f0 mínima apresentam exatamente as mesmas tendências, com e exceção de serem ligeiramente acima de 0 para CNT e AUX. O padrão f0 máxima também acompanha a média, porém os picos de f0 são em geral menores na UD do que no comentário. As diferenças não são significativas somente entre ALL e AUX e entre CNT e INP. Este último resultado não era esperado; pode ser consequência da variância dos valores de CNT. Apesar da média mais baixa (-2,9% contra 0,8% do INP),

O desvio-padrão de f0, assim como o de intensidade, é menor que os desvios na unidade de referência, porém com uma grande quantidade de valores no extremo da tendência oposta nas unidades não classificadas (AUX). O parâmetro distingue os supergrupos[42] ALL-CNT de AUX-INP.

Na Figura 33 é possível observar as correlações entre as variáveis, com tendências semelhantes aos parâmetros de intensidade. Reforçando as observações dos *boxplots*, a f0 média correlaciona-se fortemente com a f0 mínima e máxima (r = 0,69 e 0,78, respectivamente.

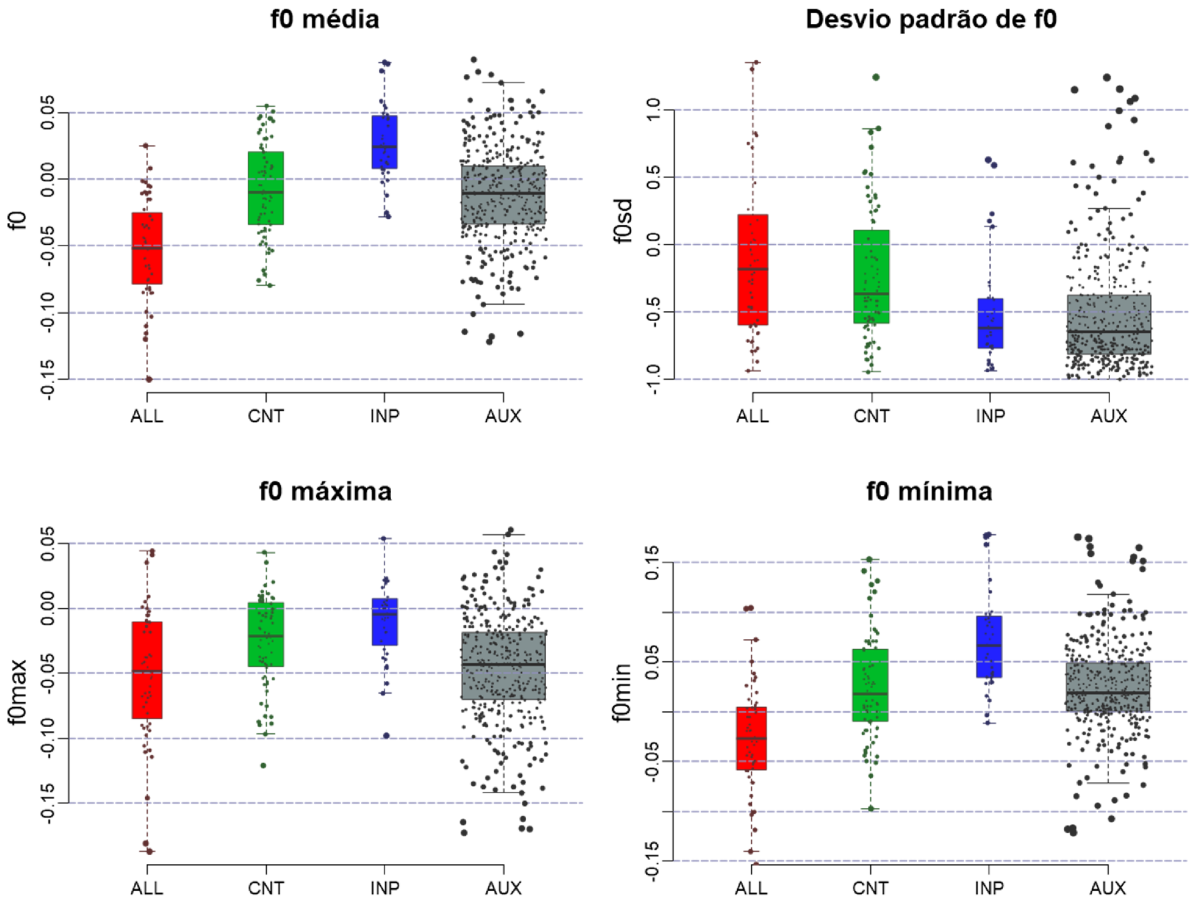


Figura 30: Distribuição dos parâmetros de f0 por tipo de UD.

Tabela 15*: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de f0*

Par	f0	f0sd	f0max	f0min
ALL-CNT	✓	✗	✓	✓
ALL-INP	✓	✓	✓	✓
ALL-AUX	✓	✓	✗	✓
CNT-INP	✓	✓	✗	✓
CNT-AUX	✗	✓	✓	✗
INP-AUX	✓	✗	✓	✓

✓ diferença significativa
✗ diferença não significativa

Tabela 16: Sumário dos parâmetros de f0 por tipo de UD

Parâmetro	Etiqueta	Número	Média	Desvio padrão	Mediana	Média truncada	Mínimo	Máximo	Amplitude	Assimetria	Curtose
f0	ALL	51	-0,053	0,038	-0,052	-0,051	-0,150	0,025	0,175	-0,264	-0,427
	CNT	69	-0,012	0,040	-0,011	-0,010	-0,203	0,055	0,258	-1,417	5,302
	INP	36	0,028	0,028	0,024	0,027	-0,028	0,088	0,116	0,259	-0,407
	AUX	404	-0,012	0,039	-0,011	-0,012	-0,227	0,144	0,371	-0,596	4,094
f0sd	ALL	51	0,038	0,829	-0,184	-0,086	-0,935	2,364	3,299	1,237	0,696
	CNT	69	-0,226	0,491	-0,331	-0,266	-0,944	1,243	2,186	0,777	-0,089
	INP	36	-0,518	0,402	-0,620	-0,578	-0,931	0,629	1,560	1,459	1,381
	AUX	404	-0,500	0,495	-0,645	-0,591	-0,998	2,367	3,365	2,517	8,426
f0max	ALL	51	-0,051	0,050	-0,048	-0,049	-0,187	0,044	0,231	-0,469	0,181
	CNT	69	-0,029	0,041	-0,022	-0,025	-0,221	0,043	0,264	-1,601	4,997
	INP	36	-0,008	0,034	-0,005	-0,008	-0,098	0,093	0,190	0,096	1,388
	AUX	404	-0,047	0,045	-0,043	-0,045	-0,321	0,091	0,411	-0,948	3,826
f0min	ALL	51	-0,027	0,055	-0,027	-0,027	-0,154	0,104	0,258	0,083	0,109
	CNT	69	0,022	0,051	0,018	0,019	-0,097	0,153	0,251	0,506	0,093
	INP	36	0,072	0,050	0,067	0,068	-0,011	0,178	0,190	0,649	-0,294
	AUX	404	0,024	0,044	0,019	0,023	-0,122	0,176	0,297	0,259	1,678

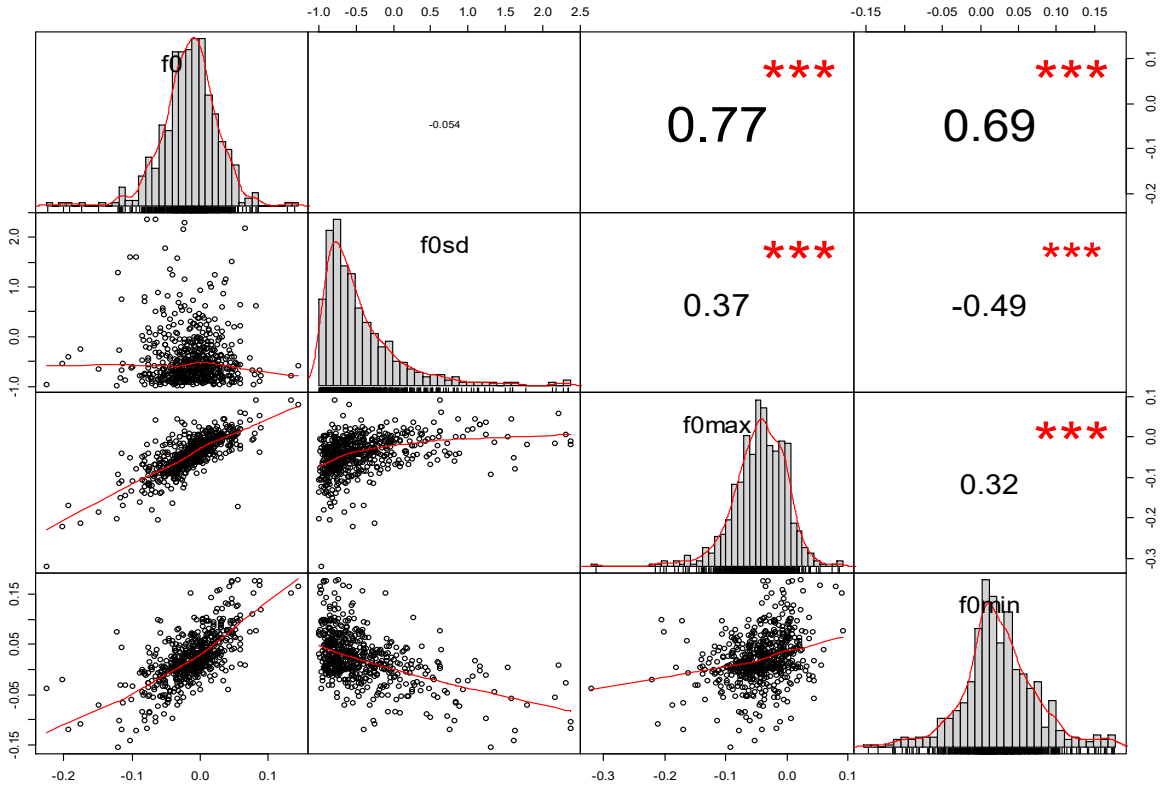


Figura 31: Correlações entre dos parâmetros de f0.

6.4 Parâmetros de variação de f0

A Tabela 18 sumariza os dados obtidos para os cinco parâmetros de variação de f0. A distribuição dos dados pode ser visualizada na Figura 32. A Tabela 17 mostra os resultados do teste de Kruskal-Wallis para a par e a Figura 33 o diagrama de correlações.

A amplitude de f0, assim como o desvio-padrão, mostrou-se menor que a do comentário em média (*f0exc* < 0), apesar de alguns valores serem muito positivos, principalmente na categoria AUX. A excursão de f0 é maior no ALL, com grande dispersão (*f0exc* = -0,10 ± 0,67). É importante observar que o ALL não apresenta em princípio um movimento de variação de f0; o movimento descendente encontrado em muitas unidades deste tipo é devido a fenômenos de terminalidade do enunciado.

As excursões das categorias INP e AUX apresentam os menores valores (≈-6%), sem se distinguirem significativamente entre elas de acordo com o teste de Kruskal-Wallis. O teste também não distingue entre CNT e ALL, apesar dos valores médios diferentes.

A taxa de variação de f0 (*f0slo*) distingue significativamente dois supergrupos: CNT e ALL, com movimento descendente de f0, e INP e AUX, com movimento ascendente, com grande dispersão de valores em todos os grupos. Os valores médios descendentes são maiores em valor absoluto do que os ascendentes (-36 e -48st/s em ALL e CNT contra +15 e +10st/s para INP e AUX).

Uma tendência semelhante é observada na variação de f0 do início à tônica (*if0sit*), porém mostrando valores destacadamente maiores para INP. Da tônica ao fim da UD, as diferenças são de menor magnitude, com perda de significância da diferença entre ALL e INP. ALL e CNT continuam apresentando tendências de movimento descendente após a tônica; para as categorias INP e AUX, os valores de *if0stf* oscilam entre movimentos ascendente e descendente.

A Figura 33 mostra correlações baixas, negativas, mas significativas entre a excursão de f0 e os demais parâmetros. Entre os 3 parâmetros de inclinação, a correlação positiva é forte (entre 0,46 e 0,63).

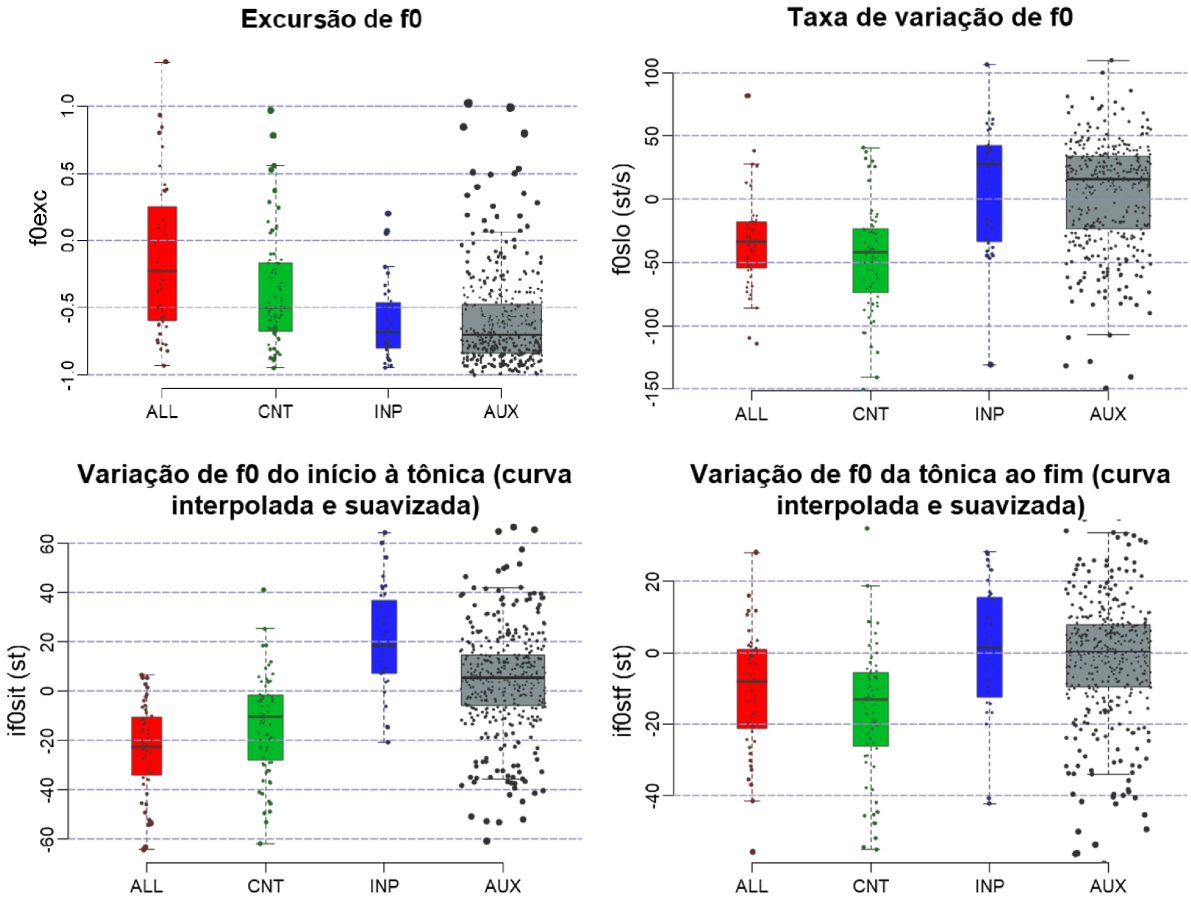


Figura 32: Distribuição dos parâmetros de variação de f0 por tipo de UD.

Tabela 17*: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de variação de f0.*

Par	f0exc	f0slo	if0sit	if0stf
ALL-CNT	✗	✗	✗	✗
ALL-INP	✓	✓	✓	✗
ALL-AUX	✓	✓	✓	✓
CNT-INP	✓	✓	✓	✓
CNT-AUX	✓	✓	✓	✓
INP-AUX	✗	✗	✓	✗

✓ diferença significativa
✗ diferença não significativa

Tabela 18: Sumário dos parâmetros de variação de f0 por tipo de UD

Parâmetro	Etiqueta	Número	Média	Desvio padrão	Mediana	Média truncada	Mínimo	Máximo	Amplitude	Assimetria	Curtose
f0exc	ALL	51	-0,102	0,670	-0,224	-0,189	-0,930	2,744	3,673	1,700	4,375
	CNT	69	-0,361	0,415	-0,462	-0,406	-0,946	0,966	1,912	1,052	0,784
	INP	36	-0,601	0,295	-0,682	-0,639	-0,943	0,201	1,144	1,078	0,419
	AUX	404	-0,589	0,383	-0,698	-0,655	-0,999	1,751	2,750	2,574	9,508
f0slo	ALL	51	-35,70	69,44	-33,90	-34,17	-404,2	143,2	547,4	-2,416	13,87
	CNT	69	-47,96	78,72	-41,79	-45,72	-523,3	175,1	698,4	-2,665	17,87
	INP	36	15,30	58,51	28,20	12,84	-130,8	212,8	343,6	0,586	2,157
	AUX	404	9,761	66,83	15,71	7,746	-311,3	663,2	974,6	2,890	26,82
if0sit	ALL	51	-23,44	18,31	-22,64	-22,64	-64,26	6,531	70,79	-0,411	-0,606
	CNT	69	-11,43	27,71	-11,05	-13,11	-61,84	146,7	208,5	2,535	13,16
	INP	36	18,11	25,65	18,64	19,31	-71,13	64,53	135,7	-0,913	2,155
	AUX	404	4,373	20,76	5,486	4,890	-85,35	79,12	164,5	-0,244	1,612
if0stf	ALL	51	-10,53	18,81	-8,088	-10,12	-59,76	39,66	99,42	-0,194	0,579
	CNT	69	-14,15	30,48	-14,08	-16,72	-64,42	168,9	233,4	3,148	16,83
	INP	36	-2,793	24,76	1,378	0,337	-79,75	28,20	108,0	-1,339	1,824
	AUX	404	-1,104	19,18	0,061	-0,675	-75,40	84,61	160,0	0,016	3,249

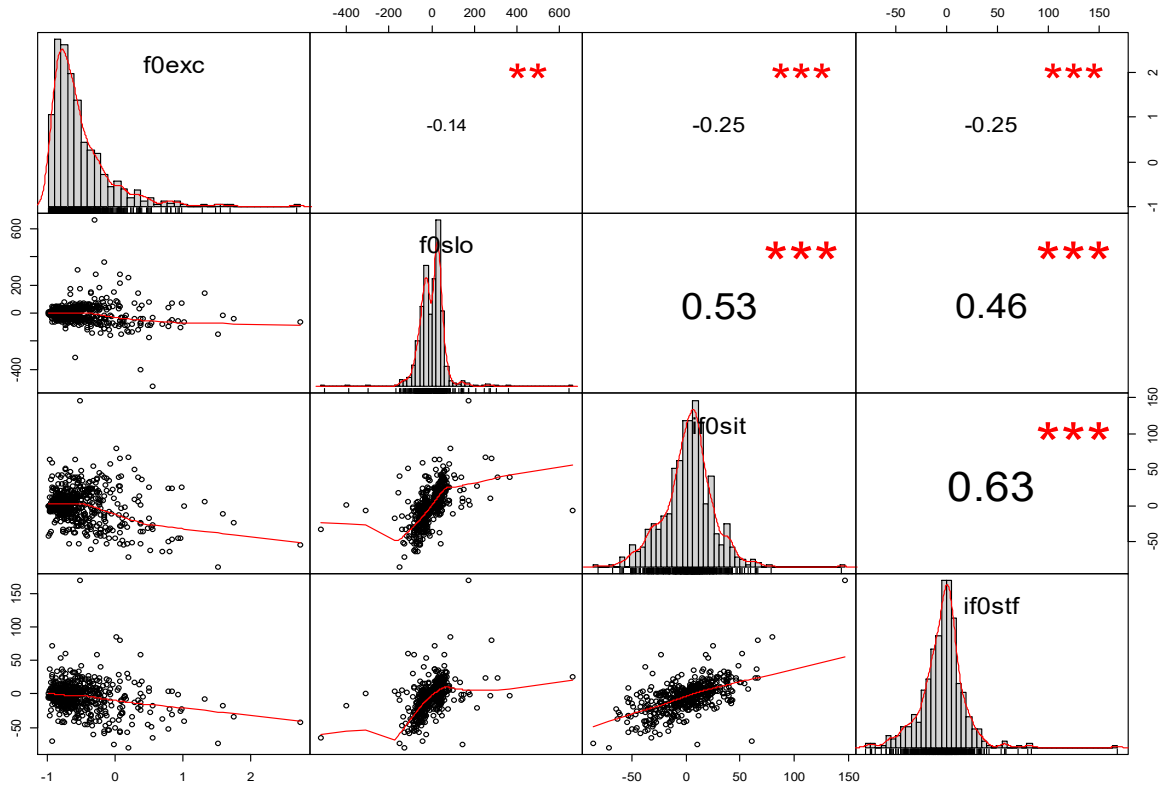


Figura 33: Correlações entre os parâmetros de variação de f0.

6.5 Parâmetros de alinhamento

Os parâmetros de alinhamento, diferentemente dos anteriores, apresentam distribuições multimodais, conforme pode ser observado nos histogramas da Figura 35 e da Figura 37. Esta observação indica que fatores relacionados à distribuição da estrutura informacional, à estrutura silábica e às marcas de fronteira estariam sendo captados pelas distribuições dos dados, por consequência da forma como os parâmetros foram definidos (seção 5.4.5).

Os valores máximos de intensidade se concentram no meio ou no início da unidade; os valores mínimos são mais prováveis em proximidade dos limites da unidade tonal e, conforme o esperado, raros em proximidade do centro da vogal tônica. O sumário dos parâmetros de alinhamento relativos à intensidade é mostrado na Tabela 20, e podem ser visualizados graficamente na Figura 34. A Tabela 19 mostra os resultados do teste de significância das diferenças. Os diagramas de correlação são apresentados na Figura 35.

Todos os parâmetros de alinhamento relativos à intensidade distinguem o ALL das demais UDs, já que o ALL apresenta, em média, o máximo de intensidade mais próximo ao início da unidade e o ponto de mínimo mais próximo ao final (com grande dispersão) quando comparado com as demais categorias. O INP exibe tendência oposta, com picos de intensidade entre o meio e o fim da unidade.

Nenhum dos parâmetros de alinhamento de pontos críticos de intensidade mostrou diferenças significativas entre CNT e AUX, e apenas o alinhamento do máximo de intensidade com a vogal tônica distingue INP de CNT. Os picos de intensidade dos INPs ocorrem, em média, após o centro da vagal tônica (*InmaxTa* = 0,5), e dos CNTs, antes do centro da tônica.

As correlações entre os alinhamentos em relação à posição na UT e em relação à tônica é muito forte e positiva para ambos os pontos críticos. Entre máximos e mínimos, as correlações são negativas, sendo moderada para a posição na unidade (r = -0,57) e fraca para a posição em relação à tônica (r = -0,15), ou seja, se o máximo de f0 encontra-se no começo da unidade, é mais provável que o mínimo encontre-se no mais próximo ao fim, e vice versa.

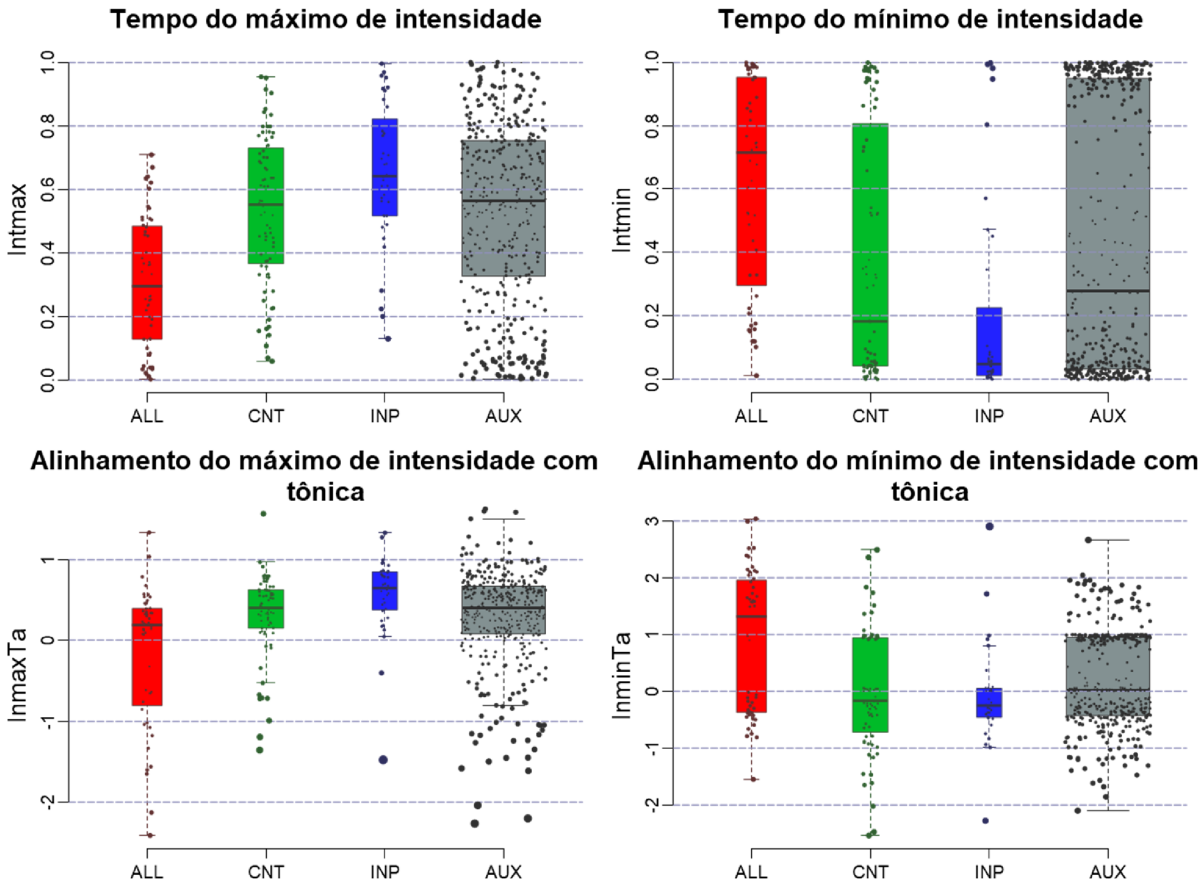


Figura 34: Distribuição dos parâmetros de alinhamento relativos à intensidade por tipo de UD.

Tabela 19*: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de alinhamento relativos à intensidade.*

Par	Intmax	Intmin	InmaxTa	InminTa
ALL-CNT	✓	✓	✓	✓
ALL-INP	✓	✓	✓	✓
ALL-AUX	✓	✓	✓	✓
CNT-INP	✗	✗	✓	✗
CNT-AUX	✗	✗	✗	✗
INP-AUX	✗	✓	✓	✗

✓ diferença significativa
✗ diferença não significativa

Tabela 20: Sumário dos parâmetros de alinhamento relativos à intensidade por tipo de UD

Parâmetro	Etiqueta	Número	Média	Desvio padrão	Mediana	Média truncada	Mínimo	Máximo	Amplitude	Assimetria	Curtose
Intmax	ALL	51	0,310	0,206	0,295	0,304	0,001	0,708	0,707	0,112	-1,191
	CNT	69	0,524	0,233	0,527	0,527	0,058	0,954	0,896	-0,137	-0,917
	INP	36	0,646	0,226	0,641	0,660	0,129	0,996	0,867	-0,395	-0,505
	AUX	404	0,526	0,279	0,562	0,536	0,003	1,000	0,997	-0,335	-0,956
Intmin	ALL	51	0,618	0,331	0,714	0,636	0,012	0,999	0,988	-0,355	-1,428
	CNT	69	0,391	0,386	0,288	0,369	0,000	0,999	0,999	0,483	-1,483
	INP	36	0,208	0,332	0,048	0,150	0,000	0,998	0,998	1,508	0,676
	AUX	404	0,445	0,420	0,277	0,432	0,000	1,000	1,000	0,275	-1,745
InmaxTa	ALL	51	-0,266	1,130	0,195	-0,099	-5,068	1,338	6,407	-1,856	4,541
	CNT	69	0,305	0,546	0,360	0,349	-1,358	1,808	3,166	-0,749	1,739
	INP	36	0,561	0,493	0,651	0,614	-1,479	1,334	2,814	-1,973	5,973
	AUX	404	0,291	0,677	0,407	0,368	-4,239	2,311	6,550	-1,679	6,793
InminTa	ALL	51	0,850	1,260	1,313	0,840	-1,551	3,040	4,591	-0,026	-1,522
	CNT	69	-0,058	1,062	-0,206	-0,062	-2,538	2,493	5,031	0,098	-0,199
	INP	36	-0,100	0,832	-0,246	-0,167	-2,279	2,907	5,186	1,138	3,952
	AUX	404	0,205	0,890	0,016	0,174	-2,102	4,009	6,111	0,493	0,501

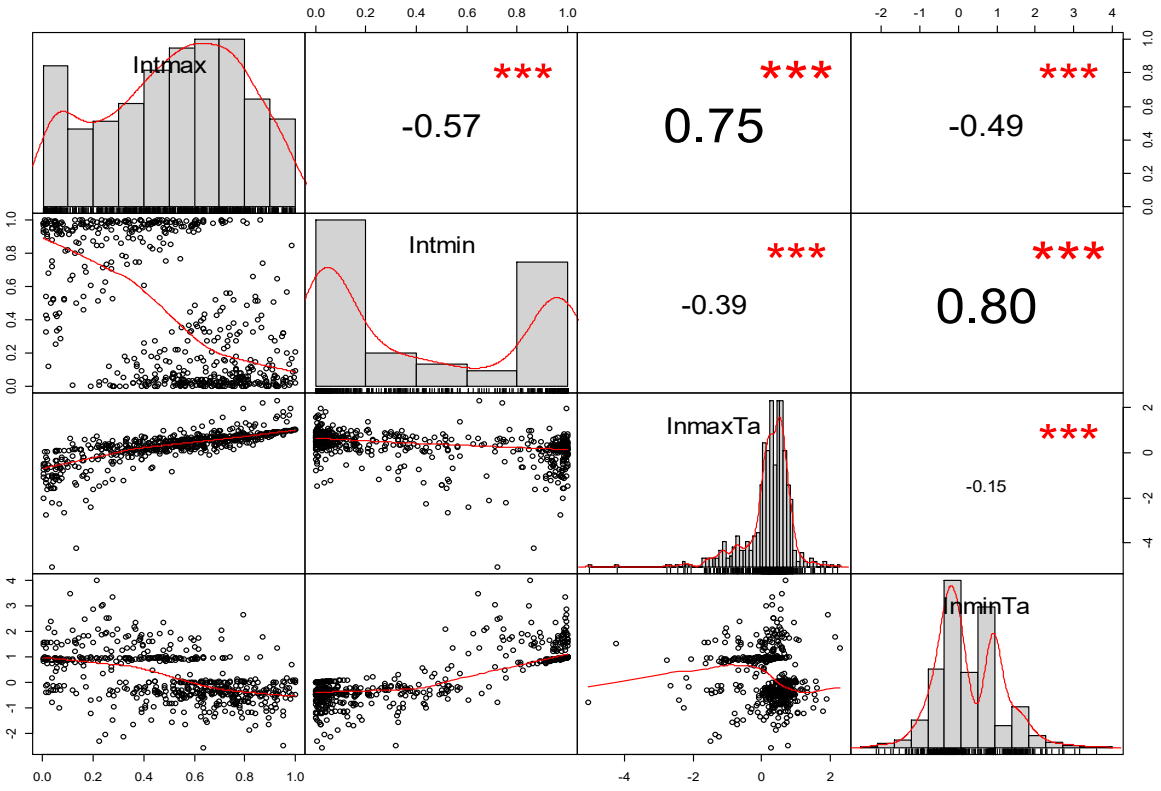


Figura 35: Correlações entre parâmetros de alinhamento relativos à intensidade.

O sumário das distribuições e sua visualização gráfica dos parâmetros de alinhamento de f0, por categoria de UD, são exibidos na Figura 36 e na Tabela 22. A Tabela 21 contém o resultado do teste de significância da diferença entre os grupos e a Figura 37 exibe as correlações entre os parâmetros.

Os resultados mostram distinção significativa entre dois supergrupos[42] de categorias: ALL-CNT e INP-AUX (com exceção da diferença ALL-INP para *f0tmin*).

O máximo de f0 é, na média, mais próximo ao início da UT no CNT e no ALL, principalmente para este último. A diferença entre as duas UDs é maior em relação à mediana: metade dos ALLs tem seu pico de f0 posicionados dentro dos 3% iniciais da UD; no CNT a mesma quantidade ocorre em até 27% do tempo. De forma semelhante, a maioria dos picos de f0 nos ALLs se posicionam antes da vogal tônica, enquanto nos CNTs a mediana é próxima a 0 (ao início da tônica). Em relação ao mínimo de f0 nessas duas UDs, há maior probabilidade que ocorra no fim da UD e no fim da vogal tônica. A grande dispersão dos dados e a bimodalidade distribucional podem ser responsáveis por essas diferenças não terem sido apontadas como significativas no teste de Kruskal-Wallis.

Nos grupos INP e AUX a tendência é diferente, os pontos críticos de f0 ocorrem na maior parte dos dados do meio ao fim para o máximo e do início ao meio para o mínimo. Entretanto, a dispersão é grande, com concentração de dados nos limites das fronteiras, tanto da inicial quanto da final. Mais de 50% de seus valores tem o ponto máximo dentro da sílaba tônica ($0 < f0maxTa < 1$). Quanto ao alinhamento do mínimo de f0, aproximadamente em metade dos dados dessas unidades eles ocorrem antes do início da tônica ($f0minTa < 0$) e em 25% após a sílaba tônica ($f0minTa > 1$).

Os alinhamentos dos pontos críticos de f0 seguem o mesmo padrão de correlação dos parâmetros de intensidade, com valores um pouco superiores. Os alinhamentos dos mesmos pontos críticos são muito correlacionados entre as duas referências (início da unidade ou vogal tônica) e de moderado a forte, negativamente, entre máximos e mínimos.

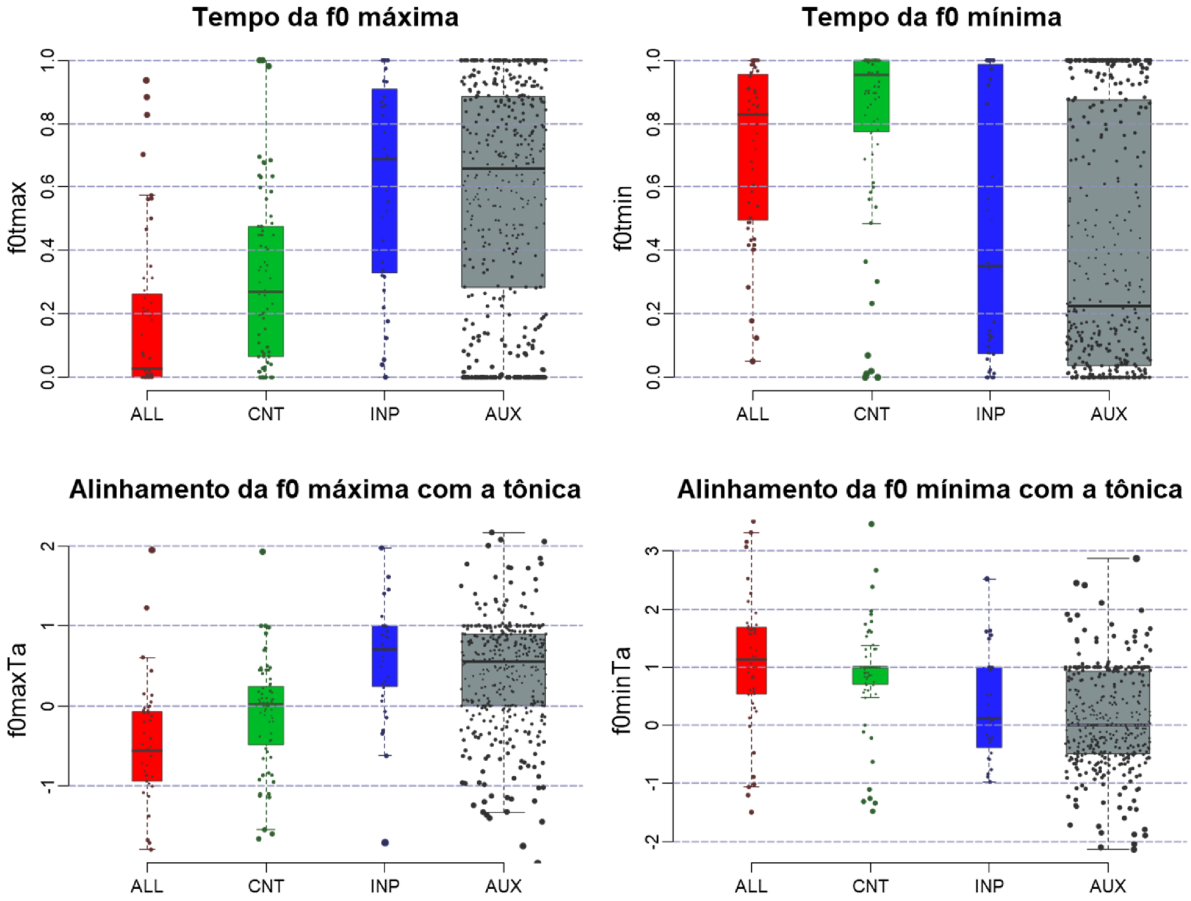


Figura 36: Distribuição dos parâmetros de alinhamento relativos à f0 por tipo de UD.

Tabela 21*: Diferenças significativas pelo teste de Kruskal-Wallis para os parâmetros de alinhamento relativos à f0.*

Par	f0tmax	f0tmin	f0maxTa	f0minTa
ALL-CNT	✗	✗	✗	✗
ALL-INP	✓	✗	✓	✓
ALL-AUX	✓	✓	✓	✓
CNT-INP	✓	✓	✓	✓
CNT-AUX	✓	✓	✓	✓
INP-AUX	✗	✗	✗	✗

✓ diferença significativa
✗ diferença não significativa

Tabela 22: Sumário dos parâmetros de alinhamento relativos à f0 por tipo de UD

Parâmetro	Etiqueta	Número	Média	Desvio padrão	Mediana	Média truncada	Mínimo	Máximo	Amplitude	Assimetria	Curtose
f0tmax	ALL	51	0,182	0,258	0,029	0,131	0,000	0,937	0,937	1,445	1,078
	CNT	69	0,298	0,265	0,261	0,272	0,000	1,000	1,000	0,781	-0,079
	INP	36	0,613	0,324	0,689	0,633	0,001	1,000	0,999	-0,367	-1,253
	AUX	404	0,577	0,345	0,657	0,597	0,000	1,000	1,000	-0,442	-1,182
f0tmin	ALL	51	0,719	0,263	0,828	0,747	0,051	1,000	0,949	-0,719	-0,554
	CNT	69	0,811	0,287	0,932	0,870	0,000	1,000	1,000	-1,707	1,839
	INP	36	0,467	0,421	0,351	0,460	0,000	1,000	1,000	0,232	-1,774
	AUX	404	0,409	0,398	0,225	0,386	0,000	1,000	1,000	0,466	-1,485
f0maxTa	ALL	51	-0,604	0,873	-0,565	-0,566	-2,930	1,949	4,879	-0,303	1,677
	CNT	69	-0,153	0,745	0,000	-0,121	-3,095	1,928	5,023	-0,817	2,701
	INP	36	0,578	0,687	0,706	0,614	-1,709	1,976	3,685	-0,889	1,703
	AUX	404	0,406	0,746	0,557	0,456	-2,505	3,068	5,572	-0,619	1,672
f0minTa	ALL	51	1,125	1,400	1,127	1,126	-2,777	4,995	7,772	-0,030	0,804
	CNT	69	0,843	1,111	1,000	0,914	-4,010	4,218	8,228	-1,028	5,263
	INP	36	0,198	1,095	0,112	0,256	-2,736	2,521	5,257	-0,573	0,677
	AUX	404	0,122	0,867	0,004	0,117	-2,137	5,438	7,574	0,601	2,890

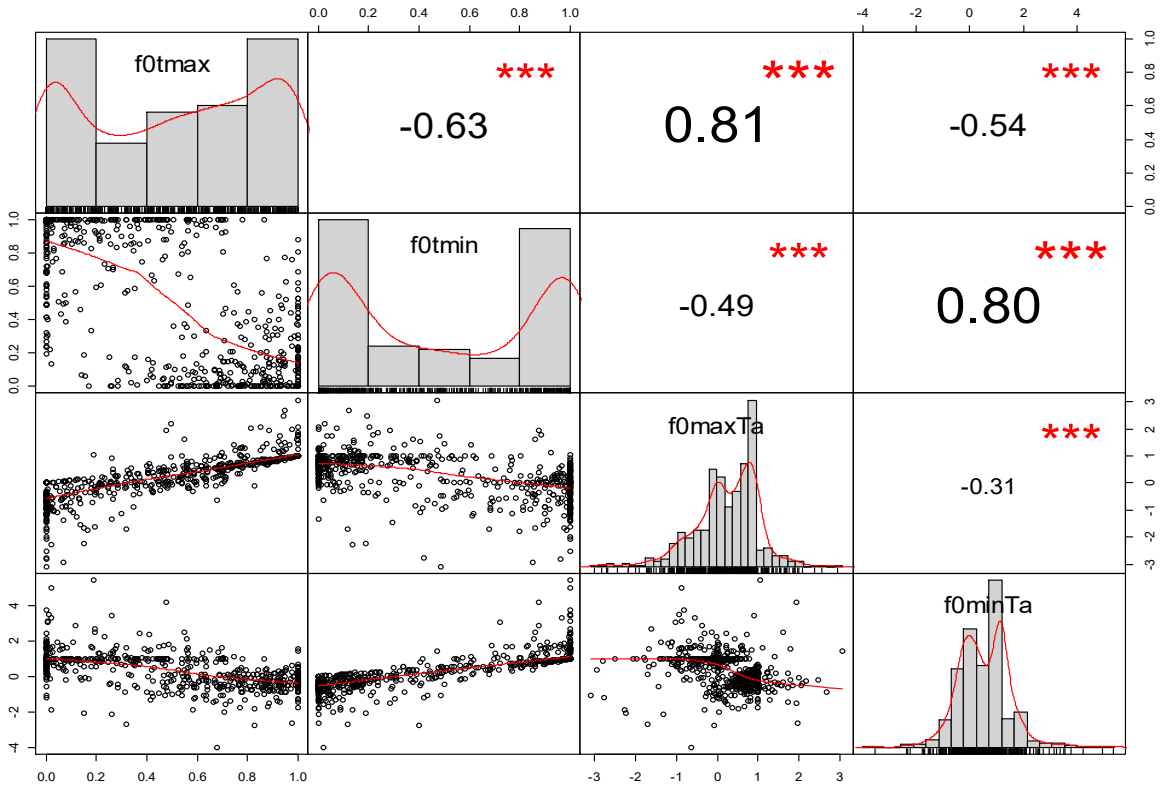
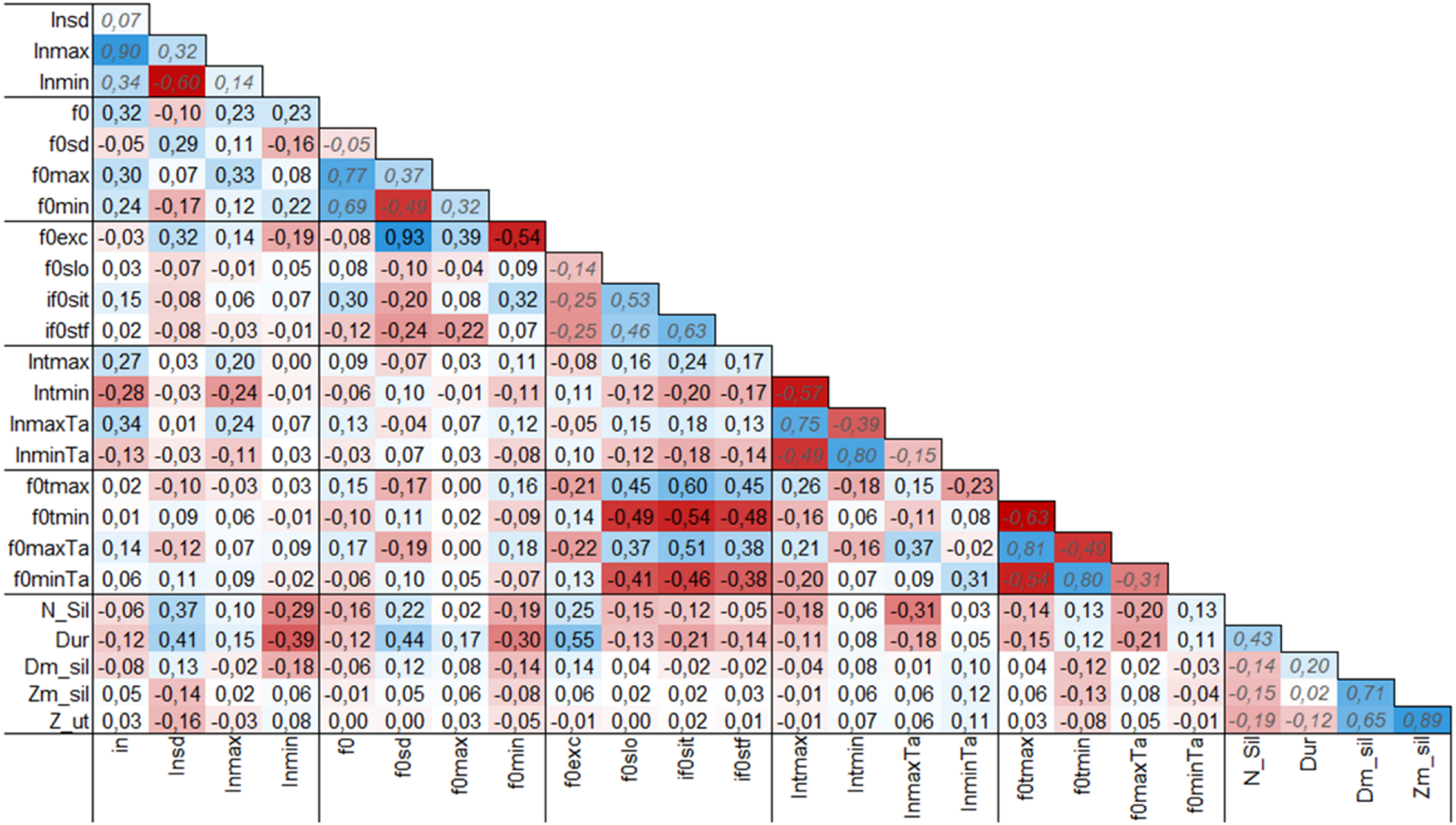


Figura 37: Correlações entre os parâmetros de alinhamento relativos à f0.

Além das correlações apontadas entre os grupos de parâmetros neste capítulo, a Tabela 23 reúne os coeficientes de correlações entre todos os 25 parâmetros. Foi encontrada correlação muito forte entre o desvio padrão e a excursão de f0 (r = 0,93). Além disso, há correlações moderadas entre os parâmetros de alinhamento relativos à f0 e os parâmetros de variação de f0; entre o desvio-padrão de intensidade e a duração e o número de sílabas da UD.

Tabela 23: Coeficientes de correlação entre todos os parâmetros.



	lnsd	lnmax	lnmin	f0	f0sd	f0max	f0min	f0exc	f0slo	if0sit	if0stf	lntmax	lntmin	lnmaxTa	lnminTa	f0tmax	f0tmin	f0maxTa	f0minTa	N_Sil	Dur	Dm_sil	Zm_sil
lnsd	0,07																						
lnmax	0,90	0,32																					
lnmin	0,34	-0,60	0,14																				
f0	0,32	-0,10	0,23	0,23																			
f0sd	-0,05	0,29	0,11	-0,16	-0,05																		
f0max	0,30	0,07	0,33	0,08	0,77	0,37																	
f0min	0,24	-0,17	0,12	0,22	0,69	-0,49	0,32																
f0exc	-0,03	0,32	0,14	-0,19	-0,08	0,93	0,39	-0,54															
f0slo	0,03	-0,07	-0,01	0,05	0,08	-0,10	-0,04	0,09	-0,14														
if0sit	0,15	-0,08	0,06	0,07	0,30	-0,20	0,08	0,32	-0,25	0,53													
if0stf	0,02	-0,08	-0,03	-0,01	-0,12	-0,24	-0,22	0,07	-0,25	0,46	0,63												
lntmax	0,27	0,03	0,20	0,00	0,09	-0,07	0,03	0,11	-0,08	0,16	0,24	0,17											
lntmin	-0,28	-0,03	-0,24	-0,01	-0,06	0,10	-0,01	-0,11	0,11	-0,12	-0,20	-0,17	-0,57										
lnmaxTa	0,34	0,01	0,24	0,07	0,13	-0,04	0,07	0,12	-0,05	0,15	0,18	0,13	0,75	-0,39									
lnminTa	-0,13	-0,03	-0,11	0,03	-0,03	0,07	0,03	-0,08	0,10	-0,12	-0,18	-0,14	-0,49	0,80	-0,15								
f0tmax	0,02	-0,10	-0,03	0,03	0,15	-0,17	0,00	0,16	-0,21	0,45	0,60	0,45	0,26	-0,18	0,15	-0,23							
f0tmin	0,01	0,09	0,06	-0,01	-0,10	0,11	0,02	-0,09	0,14	-0,49	-0,54	-0,48	-0,16	0,06	-0,11	0,08	-0,63						
f0maxTa	0,14	-0,12	0,07	0,09	0,17	-0,19	0,00	0,18	-0,22	0,37	0,51	0,38	0,21	-0,16	0,37	-0,02	0,81	-0,49					
f0minTa	0,06	0,11	0,09	-0,02	-0,06	0,10	0,05	-0,07	0,13	-0,41	-0,46	-0,38	-0,20	0,07	0,09	0,31	-0,54	0,80	-0,31				
N_Sil	-0,06	0,37	0,10	-0,29	-0,16	0,22	0,02	-0,19	0,25	-0,15	-0,12	-0,05	-0,18	0,06	-0,31	0,03	-0,14	0,13	-0,20	0,13			
Dur	-0,12	0,41	0,15	-0,39	-0,12	0,44	0,17	-0,30	0,55	-0,13	-0,21	-0,14	-0,11	0,08	-0,18	0,05	-0,15	0,12	-0,21	0,11	0,43		
Dm_sil	-0,08	0,13	-0,02	-0,18	-0,06	0,12	0,08	-0,14	0,14	0,04	-0,02	-0,02	-0,04	0,08	0,01	0,10	0,04	-0,12	0,02	-0,03	-0,14	0,20	
Zm_sil	0,05	-0,14	0,02	0,06	-0,01	0,05	0,06	-0,08	0,06	0,02	0,02	0,03	-0,01	0,06	0,06	0,12	0,06	-0,13	0,08	-0,04	-0,15	0,02	0,71
Z_ut	0,03	-0,16	-0,03	0,08	0,00	0,00	0,03	-0,05	-0,01	0,00	0,02	0,01	-0,01	0,07	0,06	0,11	0,03	-0,08	0,05	-0,01	-0,19	-0,12	0,65	0,89

Tabela 24: Descrição dos parâmetros estimados.

	Parâmetro	Descrição	Unidade	Valor da UD	Diferença para COM	Diferença proporcional
i	*In*	Intensidade média				X
ii	*Insd*	Desvio-padrão de intensidade				X
iii	*Inmax*	Máximo de intensidade				X
iv	*Inmin*	Mínimo de intensidade				X
v	*f0*	f0 média				X
vi	*f0sd*	Desvio-padrão de f0				X
vii	*f0max*	f0 máxima				X
viii	*f0min*	f0 mínima				X
ixi	*f0exc*	Excursão de f0				X
x	*f0slo*	Taxa de variação de f0	st/s	X		
xi	*if0sit*	Taxa de variação de f0 do início à tônica	st	X		
xii	*if0stf*	Taxa de variação de f0 da tônica ao fim	st	X		
xiii	*Intmax*	Tempo do máximo de intensidade em relação à duração		X		
xiv	*Intmin*	Tempo do mínimo de intensidade em relação à duração		X		
xv	*InmaxTa*	Tempo do máximo de intensidade em relação à tônica		X		
xvi	*InminTa*	Tempo do mínimo de intensidade em relação à tônica		X		
xvii	*f0tmax*	Tempo do máximo de f0 em relação à duração		X		
xviii	*f0tmin*	Tempo do mínimo de f0 em relação à duração		X		
xix	*f0maxTa*	Tempo do máximo de f0 em relação à tônica		X		
xx	*f0minTa*	Tempo do mínimo de f0 em relação à tônica		X		
xxi	*N_Sil*	Número de sílabas		X		
xxii	*Dur*	Duração da UT				X
xxiii	*Dm_sil*	Duração média da sílaba				X
xxiv	*Zm_sil*	Duração padronizada média das sílabas			X	
xxv	*Z_ut*	Duração padronizada da UT			X	

7 ANÁLISE DOS DADOS

Neste capítulo serão apresentadas as análises estatísticas que foram aplicadas aos parâmetros calculados para atingir os objetivos propostos. Foram escolhidas as técnicas de análise do discriminante linear (LDA), para modelar a categoria das unidades em função dos parâmetros nas unidades ALL, CNT e INP; e a análise hierárquica de agrupamento (análise de *cluster*), para identificar grupos com padrões de parâmetros semelhantes entre as unidades etiquetadas como AUX.

7.1 Triagem de parâmetros

Além de construir um modelo capaz de separar as unidades em tipos através da análise de seus parâmetros prosódicos, é preciso resolver algumas questões acerca dos parâmetros que foram levantados: o número de parâmetros é excessivo; alguns parâmetros são redundantes ou fortemente correlacionados entre si.

O excessivo número de parâmetros gera modelos insatisfatórios. Quando há sobreajuste (*overfitting*), o modelo se ajusta muito bem aos dados, porém perde capacidade de extrapolação e de previsibilidade para novos dados. No extremo, se o número de parâmetros do modelo é igual ao número de instâncias, o modelo irá se ajustar perfeitamente aos dados, com erro zero, e ainda assim ser incapaz de captar a natureza das relações entre as variáveis. Um modelo com sobreajuste tem complexibilidade excessiva e incorpora variâncias residuais (ruídos) presentes nos dados, de forma que os valores preditos pelo modelo também contêm alta variância.

Por outro lado, um modelo com número muito reduzido de parâmetros pode ser genérico demais para captar a estrutura subjacente aos dados (subajuste ou *underfitting*). Neste caso, os dados previstos terão pequena variância, porém terão maior tendência a desvios sistemáticos (viés) e erros maiores, em média, em relação aos dados medidos.

A Figura 38 ilustra o efeito de subajuste e o sobreajuste de modelos em problemas de regressão (figura à esquerda) e de classificação (à direita). A linha verde é resultante de sobreajuste; ela passa sobre todos os pontos na regressão e separa perfeitamente os grupos na classificação, porém contém oscilações irreais que dificultam a compreensão e a previsibilidade do modelo. Já as linhas rosas mostram um subajuste

ao tentar adequar um modelo linear à uma estrutura de dados que não é linear. Neste caso, o modelo é simples, os dados previstos contêm baixa variância, porém os erros são grandes e a estrutura das relações entre as variáveis não é captada. O modelo representado pela linha preta exibe uma solução intermediária, com parâmetros suficientes, que não se ajusta ou classifica 100% dos pontos, mas traduz de forma mais crível e compreensível a estrutura dos dados.

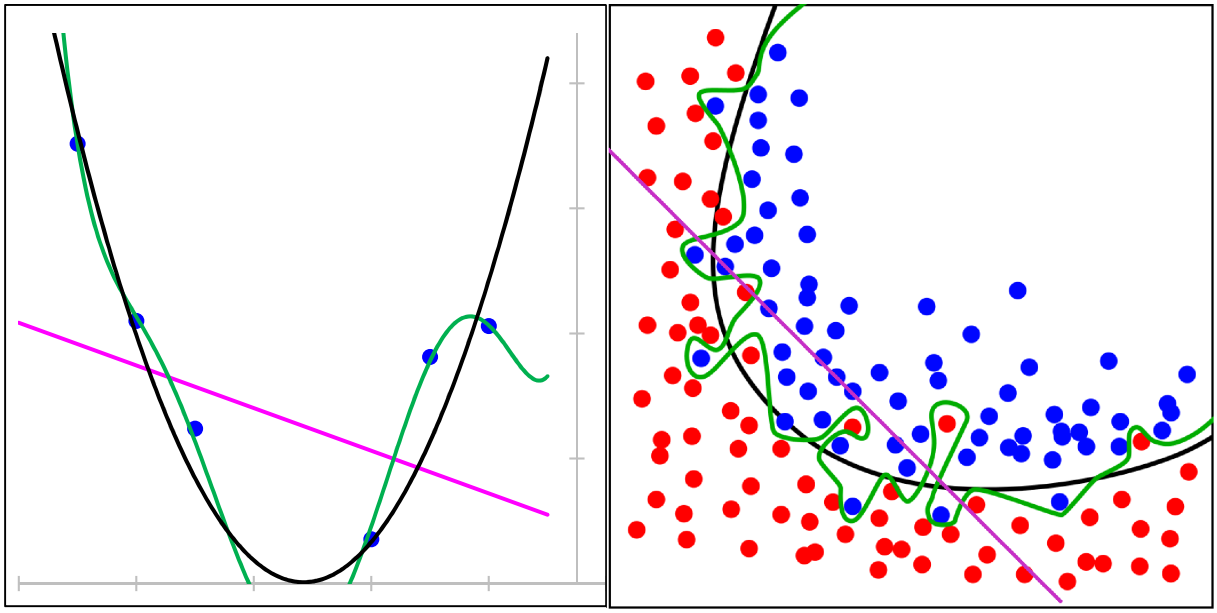


Figura 38: Ilustração de problemas de subjauste (linha rosa) e sobreajuste (linha verde) em modelos de ajuste de curvas (esquerda) e separação de grupos (direita[43]).

Outra questão relacionada ao excesso de parâmetros é a denominada maldição da dimensionalidade (BELLMAN, 1961). Por exemplo, considerando como critério para elaborar um experimento 5 sujeitos por grupo, se tivermos como único parâmetro o gênero, precisaremos de 10 sujeitos, 5 homens e 5 mulheres. Se adicionarmos o parâmetro idade com 4 faixas etárias, serão necessários 40 sujeitos. Com mais três níveis de escolaridade, 120 sujeitos; com mais cinco classes sociais 600. O número de dados necessário cresce exponencialmente com o número de parâmetros, sob pena de não obtermos a significância necessária para as conclusões.

Assim como no planejamento de experimentos e em amostragem para análises sociolinguísticas, em problemas de classificação o número de elementos de treinamento requeridos para que um classificador tenha um bom desempenho é uma função exponencial da dimensão do espaço de seus parâmetros. Na prática, quanto

[43] LIANG (2018, p.12)

maior a razão entre a quantidade de dados e o número de parâmetros, melhor será a estimativa do erro de classificação (JAIN; DUIN e MAO, 2000).

Além disso, parâmetros redundantes ou irrelevantes podem diminuir a acurácia de modelos, aumentar o tempo de processamento e dificultar a compreensão dos resultados obtidos.

> The objective of variable selection is three-fold: improving the prediction performance of the predictors, providing faster and more cost-effective predictors, and providing a better understanding of the underlying process that generated the data (GUYON e ELISSEEF, 2003).

Uma das estratégias possíveis para escolha do subconjunto de parâmetros é simplesmente testar todas as possibilidades. É uma estratégia viável até certa quantidade de variáveis, pois o número de combinações possíveis é igual 2^p. Em nosso caso, para os 25 parâmetros, 2^{25} = 33.554.432 modelos possíveis, o que já é uma quantidade bastante custosa computacionalmente para equipamentos com mediana capacidade de processamento.

Para reduzir a quantidade de combinações a serem testadas, foi realizado uma triagem das possibilidades utilizando o algoritmo conhecido como *Leaps and bounds* (FURNIVAL e WILSON, 1974), que é uma forma de busca exaustiva de parâmetros para ajuste de modelos lineares.

A grande vantagem do método é organizar todas as possibilidades em pares de combinações estruturadas hierarquicamente. A Figura 39 ilustra como o algoritmo estrutura a árvore com todas as 32 possíveis combinações de 5 parâmetros organizadas em pares:

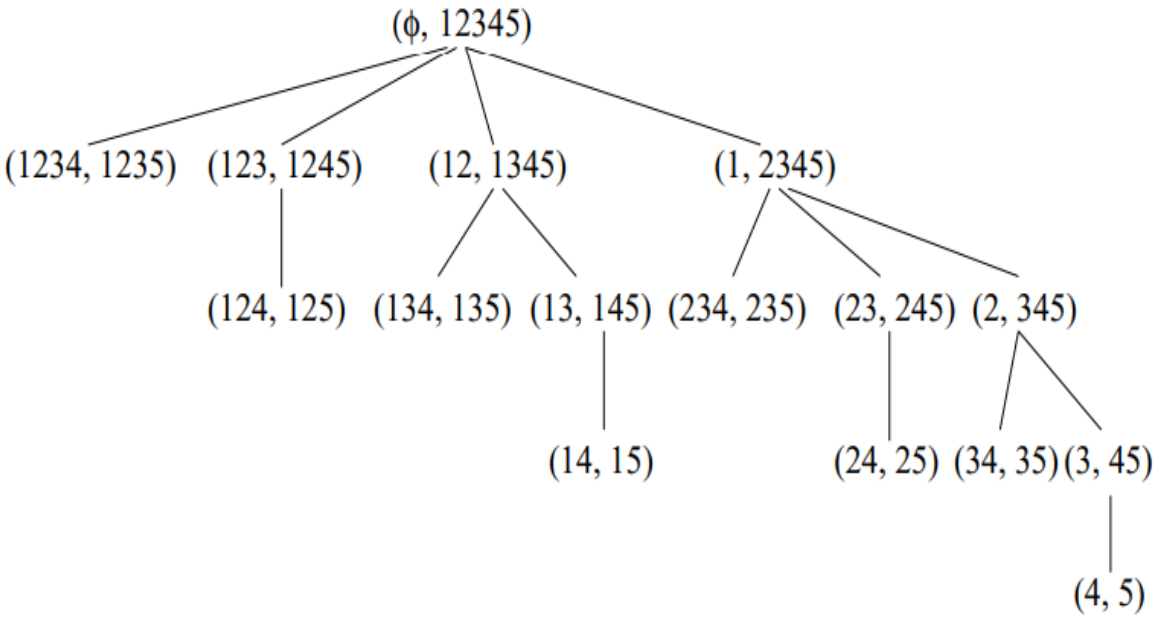


Figura 39: Exemplo de árvore de combinações para 5 parâmetros usadas no método Leaps and Bounds[44].

Cada conjunto de números representa os parâmetros de um modelo particular entre todos possíveis. O par (23, 245) contém dois modelos, um com os parâmetros 2 e 3 e outro com os parâmetros 2, 4 e 5). Do modo como é construída a árvores, cada combinação de um ramo derivado é um subconjunto da segunda combinação do par. Por exemplo, o par (124, 125) é subconjunto de 1245.

Primeiramente, o algoritmo avalia o modelo com a combinação de todos os parâmetros (12345) e os 4 pares derivados no primeiro nível. A partir daí os ramos derivados de cada par só serão avaliados se o modelo do lado direito do par apresentar um erro menor que o do lado esquerdo. Por exemplo, se, no par (13, 145), o modelo com os parâmetros 13 for melhor que com os parâmetros 145, o par derivado (14, 15) será desconsiderado, pois estes 2 modelos também serão piores que o modelo 13.

Dessa forma, o algoritmo consegue chegar no melhor subconjunto de parâmetros testando uma quantidade muito menor que todas as possibilidades. Ainda assim, é um algoritmo limitado computacionalmente para um número de parâmetros maior que 40 (BRUSCO e STAHL, 2005).

[44] Adaptada de Xuelei e Xiaoming (2006). Este artigo contém uma descrição mais acessível do método proposto por Furnival e Wilson (1974).

Como critério para a escolha entre modelos de tamanho diferentes, foi utilizada a estimativa do erro quadrático médio de predição dado pela estatística C_p de Mallows (1973). Diferentemente da somados erros quadráticos, o erro C_p não necessariamente é menor à medida que mais variáveis são adicionadas ao modelo, sendo penalizado tanto pela adição de parâmetros quanto pela colinearidade (redundância) entre eles. É definido como:

$$C_p = \frac{SQR_p}{\hat{\sigma}^2} - n + 2p = \frac{\sum_{i=1}^{N}\left(Y_i - Y_{pi}\right)^2}{\hat{\sigma}^2} - n + 2p \qquad (7.1)$$

Onde:

SQR_p= Soma dos quadrados dos resíduos;

n= número de observações;

p= número de parâmetros preditores;

Y_{pi}= valor predito para a i-ésima observação Y_i com no modelo com p preditores;

$\hat{\sigma}^2$ = estimativa da covariância dos resíduos.

O melhor modelo, entre dois possíveis, é aquele que apresenta o menor C_p, condicionado a que $C_p \leq p$ para que o modelo não seja enviesado.

A implementação do algoritmo foi realizada utilizando a função *regsubset* do pacote *leaps* (LUMLEY e MILLER, 2017) para linguagem R, que determina os melhores modelos para cada quantidade de parâmetros de parâmetros. A função foi aplicada com os 25 parâmetros determinados para as 156 UDs já classificadas (51 ALLs; 69 CNTs e 36 INPs). Foi solicitado a seleção dos 60 (no máximo [45]) melhores combinações escolhidas entre todos os subconjuntos possíveis de *p* parâmetros, resultando em 1371 combinações. O número 60 foi arbitrariamente escolhido como grande suficiente sem tempo excessivo de processamento. A Figura 40 mostra o valor de *Cp* em função da quantidade de parâmetros para todas as combinações escolhidas por *Leaps and Bounds*:

[45] Há no máximo 25 combinações com 1 ou com 25 parâmetros e apenas 1 combinação com os 25. O total de combinações é 22x60 +2x25 + 1 = 1371)

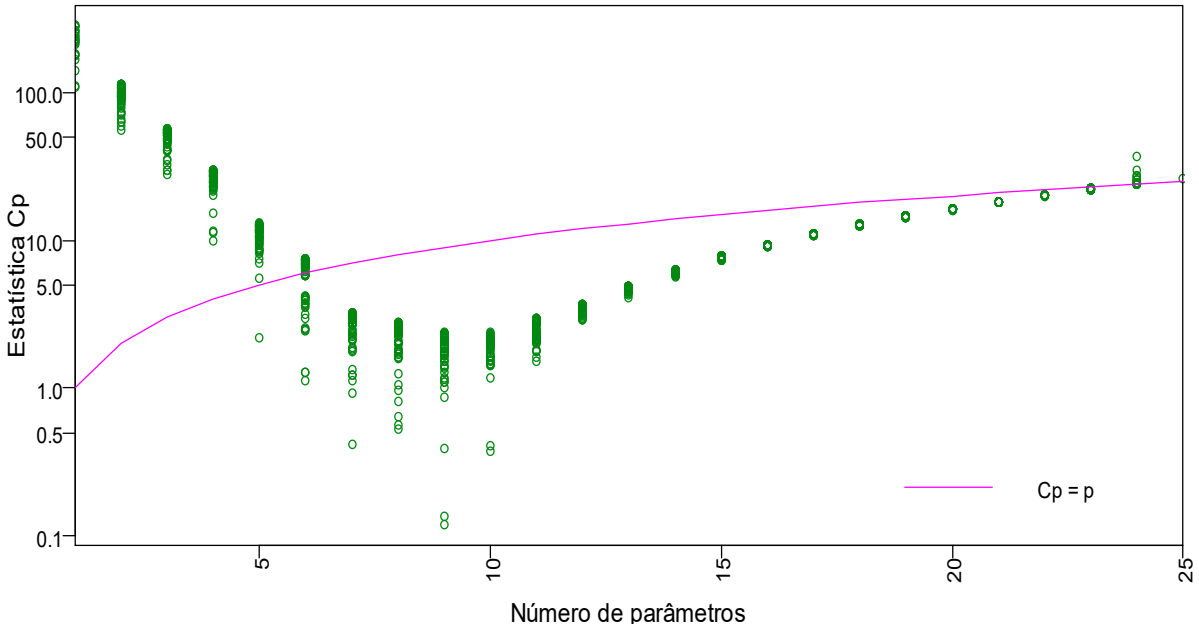


Figura 40: Estimativa do erro quadrático médio (C_p) para as 60 melhores combinações para cada número de parâmetros.

Os resultados sugerem uma quantidade de parâmetros ótima em torno de 9 parâmetros. Adicionalmente foi calculado o percentual de vezes em que cada parâmetro foi selecionado nos 1371 modelos. Conforme mostra a Figura 41, a intensidade média (*In*) foi o parâmetro mais frequente, seguido por f0 média (*f0*), número de sílabas (*N_sil*) e inclinação inicial da curva de f0 (*if0sit*). No outro extremo, o desvio-padrão de intensidade foi o parâmetro com menor percentual de seleção.

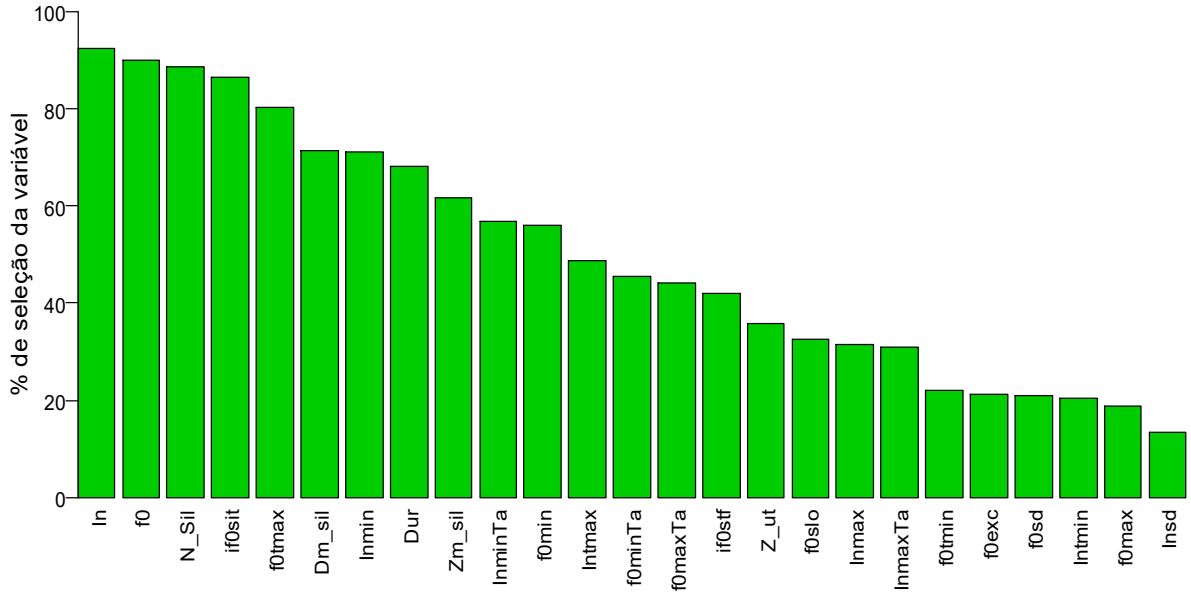


Figura 41: Percentual de vezes em que cada parâmetro foi escolhido entre as melhores combinações.

Cada uma das 1371 combinações de parâmetros selecionadas pelo método *Leaps and Bounds* foi submetida a uma análise de discriminante linear (LDA).

7.2 Análise das UDs classificadas

A análise de discriminante linear (LDA) é uma técnica de classificação que consiste em encontrar os coeficientes de uma transformação linear sobre as variáveis que minimize a probabilidade de classificação errada.

Para uma variável com *n* categorias, a dimensão do conjunto de dados é reduzida para *n* - 1 variáveis. Em nosso caso, são 3 categorias (INP, ALL e CNT), então teremos duas funções discriminantes lineares. Assim, para cada instância *i (ou para cada UD)*, serão aplicadas duas funções lineares:

$$LD1^i = C1_1 \cdot X_1^i + C1_2 \cdot X_2^i + \cdots + C1_p \cdot X_p^i \quad (7.2)$$

$$LD2^i = C2_1 \cdot X_1^i + C2_2 \cdot X_2^i + \cdots + C2_p \cdot X_p^i \quad (7.3)$$

Onde:

$LD1^i$ = Discriminante 1 para a UD i;

$LD2^i$ = Discriminante 2 para a UD i

$X_1;\ X_2;\ \ldots;\ X_p\ =$ Cada um dos *p* parâmetros utilizados; em nosso caso *In*; *f0*; etc;

$C1_1; C1_2; \ldots; C1_p =$ Coeficientes da função discriminante 1 para cada parâmetro *p;*

$C2_1; C2_2; \ldots; C2_p =$ Coeficientes da função discriminante 2 para cada parâmetro.

Para encontrar todos os pares de coeficientes ($C1_p$; $C2_p$) que projetam a matriz de dados na direção de maior separação entre grupos foi utilizada a função *lda* do pacote *Mass* (VENABLES e RIPLEY, 2002) em linguagem R.

Após o cálculo dos coeficientes, a classe de um dado *i* pode ser prevista estimando-se[46] a probabilidade relativa deste dado pertencer a cada uma das classes, baseado na distância de seus valores $LD1^i$ e $LD2^i$ para os valores médios de cada classe. A classificação é dada para a classe de maior probabilidade.

É importante observar que o método garante que a solução oferece a máxima separação se as populações apresentarem distribuições normais multivariadas com a mesma matriz de covariância entre os grupos. Essa condição não deve ser satisfeita para nossos dados, tendo em vista algumas que algumas distribuições são claramente não normais, como apresentadas na seção 6. Mesmo sem essa garantia, é comum que o método estime bons discriminantes, em comparação com técnicas com menos pressupostos:

> In practice these assumptions are never correct, and often some of the components of X are qualitative variables. It is generally felt that logistic regression is a safer, more robust bet than the LDA model, relying on fewer assumptions. It is our experience that the models give very similar results, even when LDA is used inappropriately, such as with qualitative predictors (FRIEDMAN, HASTIE, e TIBSHIRANI, 2009, p. 129).

Para avaliar a eficiência da separação, utilizamos a medida de acurácia, dada pela soma das diagonais da matriz de confusão. A figura abaixo esquematiza uma tabela de confusão gerada por um teste de classificação com três categorias (A, B, C):

[46] a classificação foi feita via função *predict* do pacote R *MASS* (VENABLES e RIPLEY, 2002).

		Dados reais			
		A	B	C	
Dados Previstos	A	**q(A|A)**	q(A|B)	q(A|C)	QPA
	B	q(B|A)	**q(B|B)**	q(B|C)	QPB
	C	q(C|A)	q(C|B)	**q(C|C)**	QPC
		QRA	QRB	QRC	QT

Figura 42: Esquema de tabela de confusão.

$\boldsymbol{q(A|A)}$ = quantidade de dados classificados como A tal que sejam realmente A;

$q(A|B)$ = quantidade de dados classificados como A tal que sejam na realidade B;

$q(B|A)$ = quantidade de dados classificados como B tal que sejam na realidade A;

QRA = Quantidade real de dados A;

QPA = Quantidade de dados previstos como A;

QT = Quantidade total de dados.

As demais quantidades podem ser definidas por indução. Idealmente, para um bom teste de classificação espera-se que a maioria dos dados estejam na diagonal em negrito. Assim, a acurácia global da classificação será definida como:

$$AC_G = \frac{\boldsymbol{q(A|A)} + \boldsymbol{q(B|B)}}{QT} \tag{7.4}$$

A acurácia global representa o percentual de acerto da classificação. Além da acurácia, outros dois critérios serão utilizados para avaliar o sucesso da classifica para cada grupo, tomando como exemplo o grupo A:

$$SE_A = \frac{\boldsymbol{q(A|A)}}{QRA} \tag{7.5}$$

$$ES_A = \frac{\boldsymbol{q(\sim A|\sim A)}}{QR\sim A} = \frac{\boldsymbol{q(B|B)} + \boldsymbol{q(C|C)} + q(B|C) + q(C|B)}{QRB + QRC} \tag{7.6}$$

$$AC_A = \frac{SE_A + ES_A}{2} \quad (7.7)$$

Onde:

SE_A= Sensibilidade da classificação de A;

ES_A= Especificidade da classificação de A;

AC_A= Acurácia da classificação de A.

$\boldsymbol{q}(\sim \boldsymbol{A} | \sim \boldsymbol{A})$= quantidade de elementos que não são do grupo A tal que não foram classificados como sendo do grupo A;

$QR\sim A$= quantidade de elementos que não pertencem ao grupo A.

A sensitividade indica qual o percentual de elementos do grupo A foram classificados como sendo do grupo A; também é denominada de taxa de verdadeiros positivos ou revocação. A especificidade indica qual o percentual de elementos que não eram do grupo A e que não foram classificados como sendo do grupo A; também é conhecida como taxa de verdadeiros negativos. A acurácia balanceada é a média entre os dois valores, os quais deseja-se que sejam os mais altos possíveis (muitos falsos positivos, poucos falsos negativos).

As tabelas de confusão foram calculadas com o auxílio da função *confusionMatrix* do pacote *caret* (KUHN, 2018) para linguagem R.

A LDA, a classificação e a matriz de confusão foram determinadas para cada uma das 1371 combinações de parâmetros pré-selecionadas pelo algoritmo *Leaps and Bounds,* sobre as 156 UDs já classificadas: 51 ALLs; 69 CNTs e 36 INPs. Com a matriz de confusão, foi calculada a acurácia global e a acurácia por tipo para cada modelo. A Figura 43 mostra os resultados dos melhores modelos para cada quantidade *p* de parâmetros:

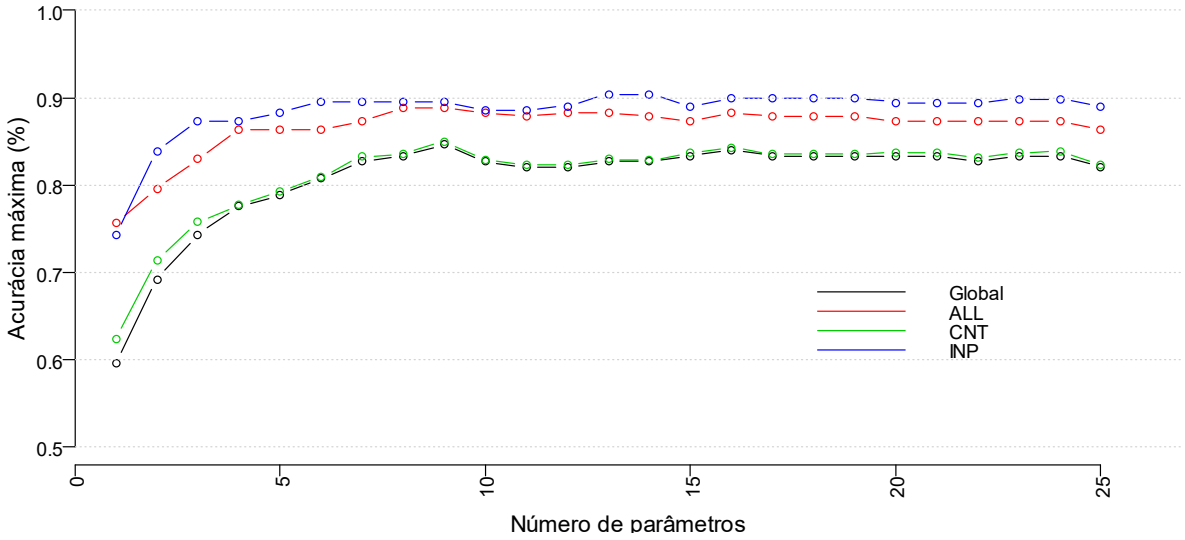


Figura 43: Acurácia global e por tipo máxima obtido em função da quantidade de parâmetros.

Observa-se que 3 parâmetros fazem grande diferença para a classificação de INP e que com 6 parâmetros a acurácia deste tipo chega próximo a 90%. Com o ALL a evolução da acurácia é semelhante com 4 e 8 parâmetros.

O CNT é a unidade com maior dificuldade de classificação, seus valores de acurácia são inferiores aos das demais classes e determinam a acurácia global da classificação, por esta ser a classe mais frequente. Os parâmetros adicionados aumentam gradualmente a sua acurácia até um máximo em torno de 85% com 9 parâmetros.

O aumento da quantidade de parâmetros acima de 9 não resulta em maior acurácia global. Apenas para o INP, há um resultado levemente superior com 13 parâmetros, porém em detrimento do resultado para ALL e CNT.

A máxima acurácia global com o mínimo de parâmetros foi obtida com 9 parâmetros: *In*; *Inmin*; *InminTa*; *f0*; *f0maxTa*; *f0minTa*; *if0sit*; *N_sil* e *Dur*. A tabela de confusão a seguir mostra como os dados foram classificados:

Tabela 25: Tabela de confusão para a melhor combinação de 9 parâmetros

		Dados reais			
		ALL	CNT	INP	*QP*
Dados Preditos	ALL	**42**	6	0	48
	CNT	9	**61**	7	77
	INP	0	2	**29**	31
	QR	51	69	36	156
	Sensitividade	82,4%	88,4%	80,6%	
	Especificidade	94,3%	81,6%	98,3%	AC_G
	Acurácia	88,3%	85,0%	89,4%	84,6%

Observa-se que os melhores resultados para ALL e INP se devem às suas elevadas especificidades: 94,4 e 89,5%, respectivamente. Das 31 unidade classificadas como INP, apenas 2 não são INP; entre as classificadas como ALL, 6 UDs não correspondem à categoria. Não houve confusão entre INP e ALL, a função distinguiu completamente as duas unidades. A sensitividade, entretanto, foi menor, em ambas as unidades. Dos 51 ALL, 42 foram classificados corretamente (82,4%); para o INP foram 29 acertos em 36 (80,6%).

O CNT apresentou a menor acurácia devido sua menor especificidade, 64 acertos em 72 unidades (88,9%). Sua acurácia menor é devida à especificidade mais baixa, muitas unidades foram incorretamente classificadas como CNT, 9 ALLs e 7 INPs.

A separação entre os grupos promovida pela LDA pode ser visualizada na Figura 44. No gráfico, os dados estão posicionados de acordo com os valores de discriminantes LD1 e LD2. O nome do rótulo indica a sua classe de fato, a cor indica a classificação do modelo: vermelho para ALL, verde para CNT e azul para INP. Os valores com cores intermediárias são aqueles mal classificados:

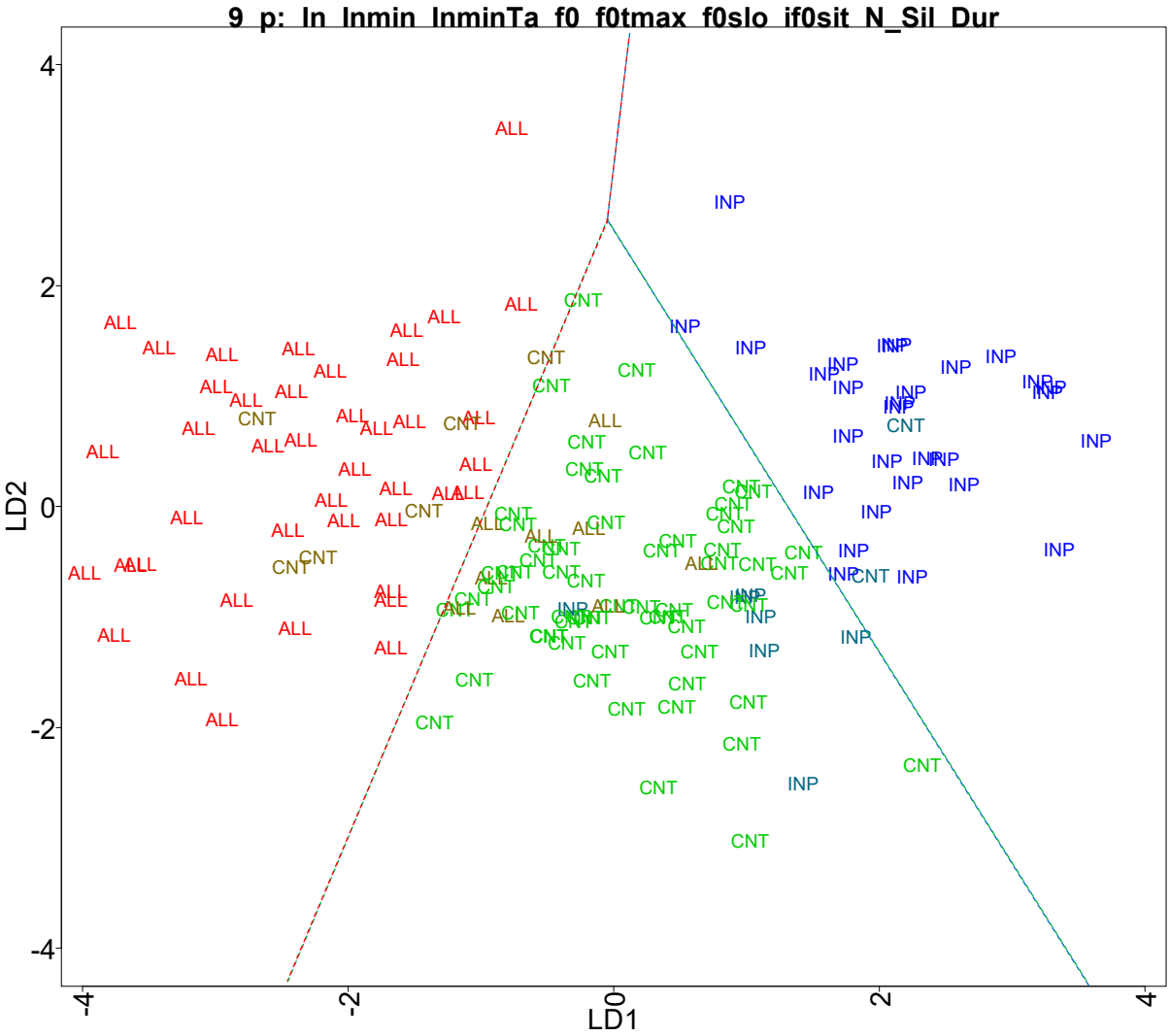


Figura 44: Separação dos dados pelos discriminantes.

A visualização mostra uma clara separação entre os tipos ALL e INP, com os CNTs se posicionando entre eles. A fronteira parece mais clara entre CNT e INP, grupos com menor confusão (9 elementos). A maioria destes encontram-se próximos à fronteira; porém um CNT e um INP estão posicionados nas regiões centrais dos respectivos grupos equivocados. Entre ALL e CNT, a fronteira é menos clara, muitos dos 15 dados mal classificados dispersam-se entre os corretamente classificados de ambos os lados. Esses resultados são coerentes com as dificuldades de diferenciação ALL e CNT durante a revisão da etiquetagem informacional do Mini BR Informal.

O primeiro discriminante (*LD1*) é o que promove a maior separação. Os ALLs concentram-se na região com *LD1* < 1, e os INPs na região com *LD1* > 1,5. O segundo discriminante parece ajudar a separar os CNTs, que são mais frequentes em *LD2* < 0, ao contrário dos outros dois tipos. Pelos resultados gerados pela função *lda*, o *LD1*

é responsável por 91,2% da variância entre os grupos, enquanto o *LD2* responde por 8,8%[47]. A tabela a seguir mostra os coeficientes das funções discriminantes para cada parâmetro:

Tabela 26: Coeficientes das funções discriminantes

	Padronizados		Não padronizados	
Parâmetro	LD1	LD2	LD1	LD2
In	**0,712**	**-0,454**	8,995	-5,739
f0	**0,596**	0,286	12,58	6,035
if0sit	**0,361**	0,154	0,012	0,005
f0tmax	**0,339**	**0,504**	1,060	1,575
f0slo	0,049	**0,549**	0,001	0,007
InminTa	-0,187	**0,744**	-0,161	0,640
Dur	-0,224	0,138	-0,781	0,482
Inmin	-0,232	-0,128	-1,138	-0,628
N_sil	**-0,520**	-0,023	-0,707	-0,032

Os coeficientes estão mostrados de duas formas, com os dados padronizados[48] e com os dados não padronizados. Os dados não padronizados foram os utilizados pelas funções; os dados padronizados permitem a comparação do peso relativo de cada parâmetro na função discriminante. Os valores mostram que os parâmetros com maiores efeitos para o primeiro discriminante, são a intensidade média (*In*), a f0 média (*f0*) e o número de sílabas (*N_sil*).

É interessante que o procedimento utilizado selecionou os parâmetros *N_sil* e *Dur* sobre os parâmetros de duração estimados para traduzir a taxa de elocução (*Dm_sil*, *Zm_sil* e *Zut*). Apesar de serem úteis à classificação e de estarem relacionados conjuntamente à taxa de elocução, esses parâmetros carregam informações lexicais e distribucionais que dificultam a compreensão da forma prosódica. Por essa razão o procedimento foi reaplicado aos dados excluindo-se os parâmetros de duração da UD

[47] A proporção entre as variâncias é chamada de *proportion of trace.*

[48] Foi utilizada a padronização pela transformação *Z-score*, conforme equação (5.3).

e de número de sílabas. A Figura 45 mostra a mudança na curva de acurácia em função da quantidade de parâmetros:

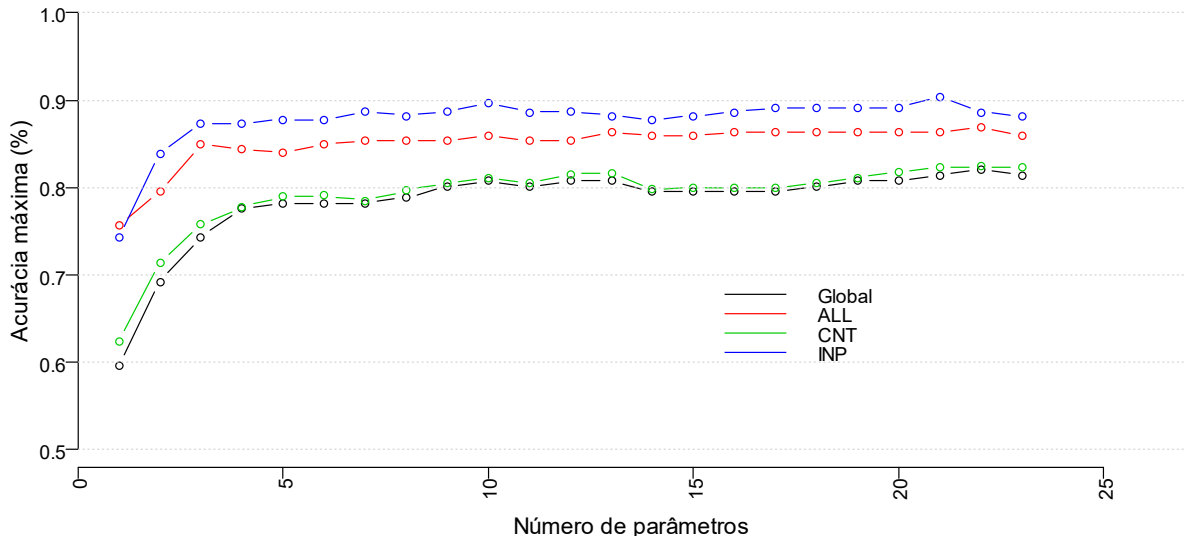


Figura 45: Acurácia global e por tipo máxima obtido em função da quantidade de parâmetros, sem N_sil e Dur.

Observa-se que houve queda na acurácia global máxima obtida para próximo de 80%, também em torno de 10 parâmetros. As acurácias para os tipos também se reduzem, principalmente para ALL. Os parâmetros que maximizaram a acurácia para 3, 4, 5 e 10 parâmetros podem ser visualizados na Tabela 27:

Tabela 27: Parâmetros e acurácias máximas, sem N_sil e Dur

		Acurácia			
p	Parâmetros	Global	ALL	CNT	INP
3	*In*; *f0*; *if0sit*	74,4%	81,0%	74,6%	84,6%
4	*In*; *f0*; *f0minTa*; *if0sit*	76,6%	82,0%	76,9%	86,4%
5	*In*; *f0*; *f0min*; *f0tmax*; *if0sit*	78,2%	83,4%	78,9%	87,4%
10	*In*; *Inmax*; *Inmin*; *Intmin*; *InmaxTa* *f0*; *f0min*; *f0maxTa*; *f0minTa*; *if0sit*	80,8%	85,9%	81,1%	87,8%

A tabela indica que os parâmetros com maior contribuição para a discriminação são a intensidade média (*In*), a f0 média (*f0*) e a inclinação inicial da curva de f0 (*if0sit*). No modelo com 10 parâmetros, os parâmetros adicionais selecionados foram aqueles aqueles relacionados com o alinhamento dos pontos críticos em relação à vogal tônica e à duração da UD.

7.3 Análise das UDs não classificadas

Para as UDs que foram etiquetadas como AUX por não se enquadrar nas categorias já descritas, foi realizada uma busca por padrões na matriz de parâmetros prosódicos estimados. Para tanto, aplicou-se o método de agrupamento hierárquico aglomerativo (HASTIE; TIBSHIRANI e FRIEDMAN, 2009).

O método consiste em considerar-se inicialmente cada elemento um grupo e determinar a distância entre cada par de elementos. Os pares com a menor distância são então reunidos em um único grupo. Repete-se o processo até que todos os elementos sejam reunidos em um único grupo, formando uma hierarquia de aglomerações sucessivas. A hierarquia de agrupamento pode ser graficamente representada por um dendograma, como exemplificado na figura abaixo:

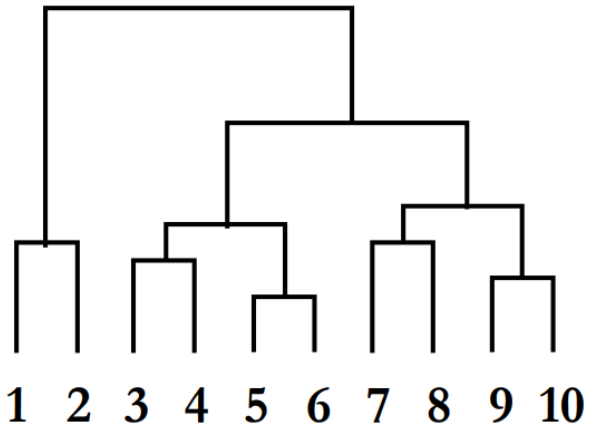


Figura 46: Exemplo de dendograma

Usando a Figura 46 para ilustrar o método, observamos no primeiro nível (ou mais baixo) 10 grupos, cada um contendo um único elemento. No nível seguinte, os elementos 5 e 6 se aglomeram no grupo (5-6). No próximo forma-se o grupo (9-10); no seguinte o (3-4), e assim sucessivamente, até que todos elementos estejam reunidos no nível mais alto, o 9º, neste exemplo com 10 observações. Usualmente, a distância entre os nós (altura) do dendograma é proporcional ao grau de dissimilaridade entre os grupos reunidos, que sempre aumenta a cada nível de agrupamento.

O método tem a vantagem de não necessitar do conhecimento do número de classes *a priori*. Em contrapartida, a quantidade de classes resultante será uma decisão a ser

tomada pelo analisador, que é feita pela escolha da altura de corte no dendograma. A Figura 47 ilustra duas possibilidades de escolha para o exemplo acima. À esquerda, a linha verde corta o dendograma na altura entre os níveis 8 e 9, resultando em dois grupos: (1-2) e (3-4-5-6-7-8-9-10). À direita, a linha vermelha corta um nível abaixo, resultando em 3 grupos: (1-2) e (3-4-5-6) e (7-8-9-10). A escolha do número de grupos frequentemente é uma decisão difícil. Para Rokach (2010), a avaliação se um agrupamento é bom ou não é uma questão problemática e controversa; não há um critério universalmente aceito e é comum que a decisão seja feita pela análise visual do dendograma.

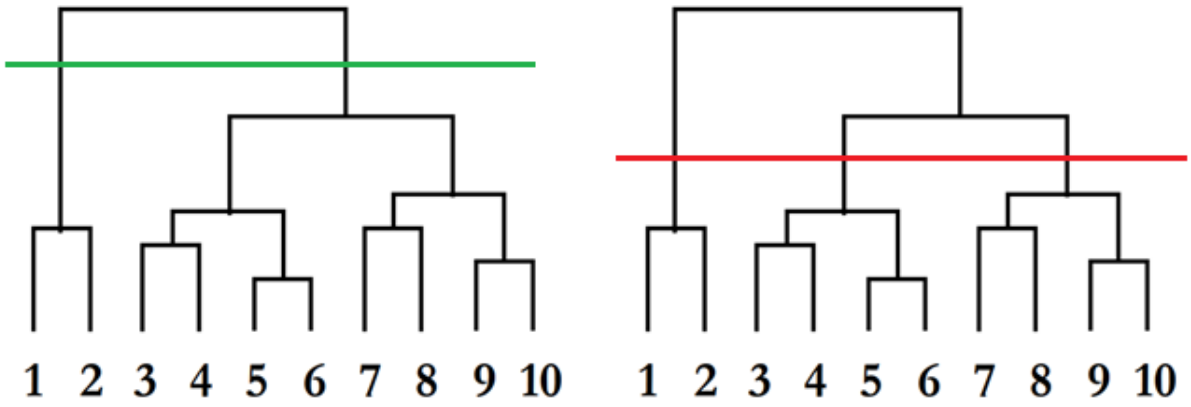


Figura 47: Esquema da escolha do número de classes

A análise do dendograma deve sempre ser cautelosa. Sua representação não é um sumário dos dados em si, e sim uma representação gráfica do algoritmo utilizado, que impõe uma estrutura hierárquica quer ela exista ou não. Além disso, a estrutura pode ser sensível aos dois requisitos para implementação da aglomeração: i) a definição de medida de dissimilaridade entre os pares de elementos ou grupos e ii) um critério de ligação de elementos ou grupos aos grupos similares.

A matriz de dissimilaridade para um conjunto de *n* dados é uma matriz de dimensões *n* x *n* que contém a distância entre cada par de dados. Neste trabalho foi utilizada a métrica mais frequente. Entre dois elementos *i* e *j* de um conjunto de dados com número *p* de parâmetros, a distância é dada por:

$$d^{ij} = \sqrt{\sum_{k=1}^{p} \left(X_i^k - X_j^k\right)^2} \tag{7.8}$$

Onde:

d^{ij}=distância entre o elemento i e o elemento j

X_i^k=valor do parâmetro X^k para o elemento i;

X_j^k=valor do parâmetro X^k para o elemento j.

Para que a medida de distância não seja enviesada com relação aos parâmetros com maior magnitude, é muito importante que os dados sejam previamente padronizados. Utilizamos, para tanto, a padronização pelo escore Z, conforme a equação (5.3).

A reunião entre os elementos ou entre grupos a cada etapa do método é feita entre os elementos com menor dissimilaridade, ou com menor distância euclidiana. A partir do momento que se decide quais elementos serão reunidos, é preciso decidir qual critério para definir o valor da reunião na matriz de dissimilaridade, ou um critério de ligação. Os critérios mais comuns são: a ligação simples (escolha do elemento com a menor distância); ligação completa (escolha do elemento com a maior distância); ligação média (média aritmética entre as distâncias de todos os elementos). Neste trabalho foi usado o método de ligação de Ward (1972), que liga ou grupos de modo a minimizar a soma das diferenças quadráticas entre os elementos do grupo formado. Este método apresenta a vantagem de formar grupos mais homogêneos entre si (OIKONOMAKOU e VAZIRGIANNIS, 2010).

Para a implementação do método, utilizamos a função *agnes* do pacote *Cluster* em linguagem R (MAECHLER et al., 2018). Foram utilizados todos os 25 parâmetros para encontrar um padrão de agrupamento em 404 UDs sem classificação. O dendograma gerado é mostrado na Figura 48.

Baseado na observação visual da estrutura do dendograma, escolhemos avaliar um possível agrupamento com 3 classes. O corte na altura 34 gera 3 grupos com 169, 133 e 102 elementos. A Tabela 28 sumariza a média desses grupos para todos os parâmetros, bem como as dos grupos já classificados, para comparação. Está também indicado o resultado do teste Kruskal-Wallis para diferença entre os pares de grupos.

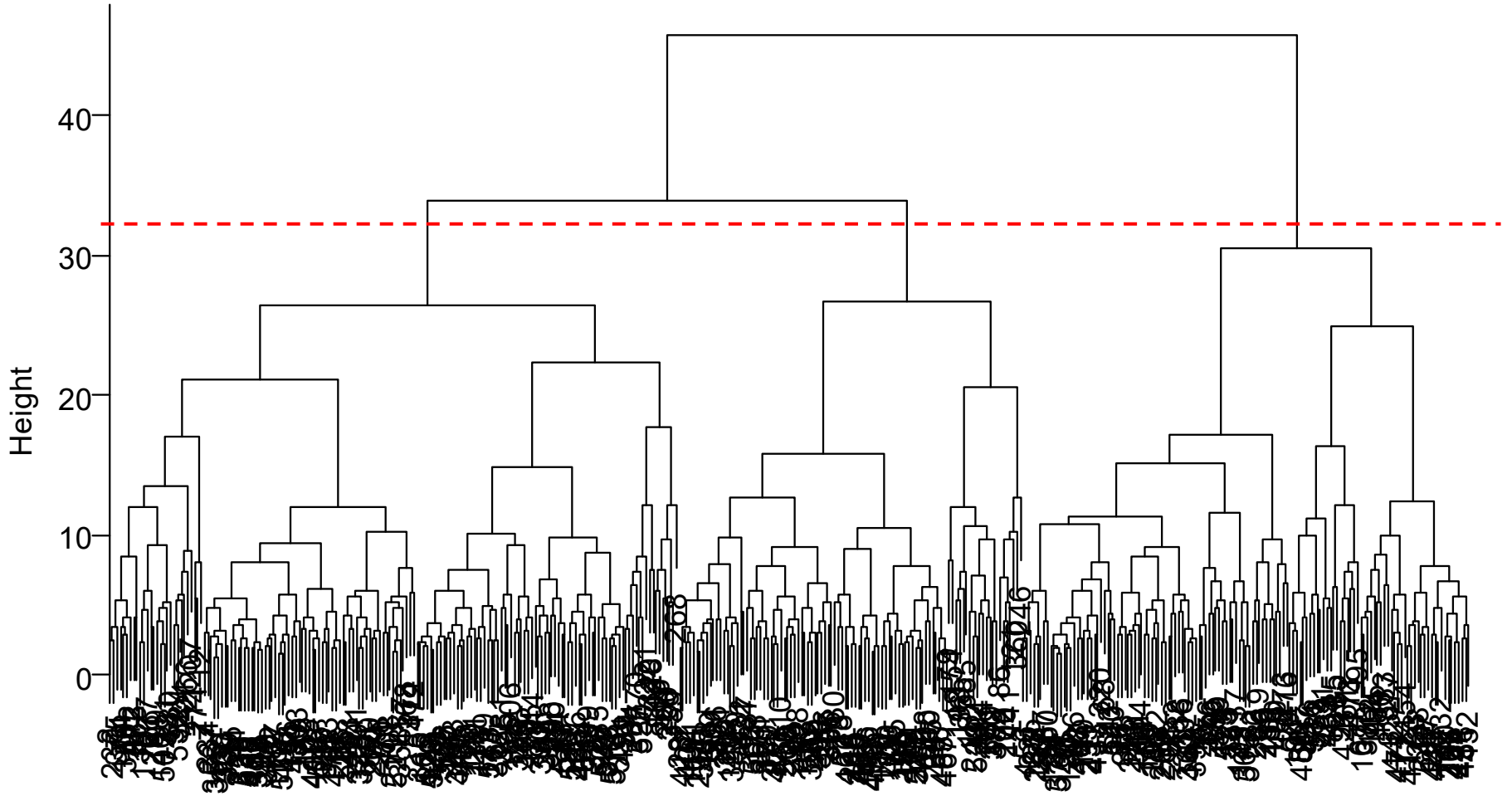


Figura 48: Dendograma do agrupamento hierárquico aglomerativo com todos os parâmetros para todos AUXs

Tabela 28: Médias dos parâmetros para as UDs conhecidas e para os grupos propostos

N	Grupo	N_Sil	Dur	Dm_sil	Zm_sil	Z_ut	In	Insd	Inmax	Inmin	f0	f0sd	f0max	f0min	f0exc	f0slo	if0sit	if0stf	Intmax	Intmin	InmaxTa	InminTa	f0tmax	f0tmin	f0maxTa	f0minTa
51	ALL	2.02	-0.53	0.04	-0.71	-0.47	-0.082	-0.16	-0.09	0.02	-0.053	0.04	-0.051	-0.027	-0.10	-35.7	-23.4	-10.5	0.31	0.62	-0.27	0.85	0.18	0.72	-0.60	1.12
69	CNT	1.58	-0.71	-0.12	-0.52	0.04	0.005	-0.25	-0.04	0.17	-0.012	-0.23	-0.029	0.022	-0.36	-48.0	-11.4	-14.2	0.52	0.39	0.31	-0.06	0.30	0.81	-0.15	0.84
36	INP	1.22	-0.82	-0.23	-1.48	-0.89	0.051	-0.28	-0.01	0.22	0.028	-0.52	-0.008	0.072	-0.60	15.3	18.1	-2.8	0.65	0.21	0.56	-0.10	0.61	0.47	0.58	0.20
169	G1	1.07	-0.78	0.03	0.27	0.87	-0.029	-0.44	-0.07	0.17	-0.021	-0.42	-0.055	0.014	-0.54	43.0	13.0	7.4	0.68	0.18	0.57	-0.31	0.78	0.19	0.75	-0.31
133	G2	1.05	-0.81	-0.16	-0.78	0.14	-0.032	-0.50	-0.08	0.22	-0.008	-0.50	-0.042	0.029	-0.57	-39.2	-10.0	-13.1	0.48	0.47	0.19	0.26	0.24	0.84	-0.21	0.86
102	G3	1.25	-0.77	-0.02	-0.22	0.69	-0.059	-0.48	-0.10	0.18	-0.004	-0.63	-0.041	0.036	-0.68	18.5	8.8	0.5	0.33	0.85	-0.04	0.99	0.68	0.22	0.65	-0.13
G1-G2		✗	✓	✓	✓	✗	✗	✗	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓	✗	✓	✓
G1-G3		✗	✗	✗	✗	✗	✗	✓	✗	✓	✓	✗	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓	✗	✓	✗	✗
G2-G3		✗	✓	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓	✗	✓	✓

Os resultados mostram que os grupos propostos se distinguem significativamente entre si em 6 parâmetros: *Inmin*; *Intmax*; *Intmin*; *InmaxTa*; *f0min* e *if0sit*. Por outro, lado os parâmetros *In*; *Inmax*; *f0sd*; *if0stf* e *Z_ut* não distinguem significativamente nenhum dos grupos propostos. Os demais parâmetros são capazes de perceber diferença significativas em um ou dois pares de grupos.

Apesar do resultado do teste, algumas médias são bem diferentes entre si (ex.: *Z_ut*; *if0stf*), sugerindo a necessidade de uma caracterização mais detalhada das diferenças entre os grupos.

É interessante notar que os grupos propostos não se distinguem significativamente pela intensidade média (*In*), parâmetro que se mostrou o mais importante para discriminar as unidades já classificadas.

Os *boxplots* das figuras a seguir ilustram as distribuições nos 3 agrupamentos gerados. De modo geral, nota-se o destaque do grupo G2 em relação aos demais em um maior número de parâmetros, principalmente nos parâmetros de variação de f0 (Figura 52) e nos parâmetros de alinhamento relativos à f0 (Figura 54). O G2 apresenta variações negativas, enquanto em G1 e G3 as variações são positivas. Os mínimos de f0 em G2 ocorrem no fim da UD, após a sílaba tônica, enquanto em G1 e G3 os mínimos são mais frequentes no início da UD, um pouco antes da vogal tônica. O alinhamento do máximo de f0 exibe tendência oposta.

Os grupos G1 e G3 apresentam diferenças maiores principalmente quanto aos alinhamentos em relação à intensidade (Figura 53). O mínimo de intensidade ocorre mais frequentemente no início da UD em G1, antes da tônica, e no fim em G3; alinhado ao fim da vogal tônica. Já o máximo de intensidade tende ao centro da tônica e ao final da unidade em G1 e mais no início da vogal tônica em G3.

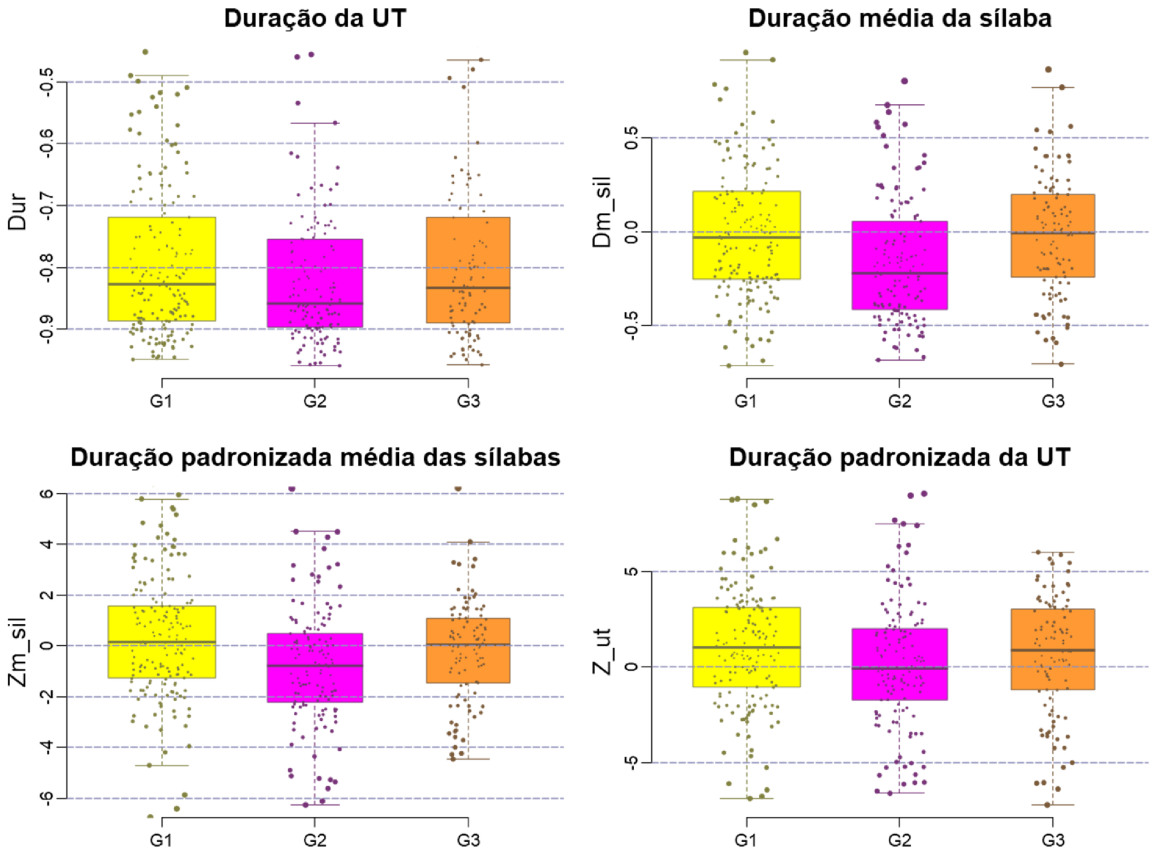


Figura 49: Distribuição dos parâmetros de duração por agrupamento.

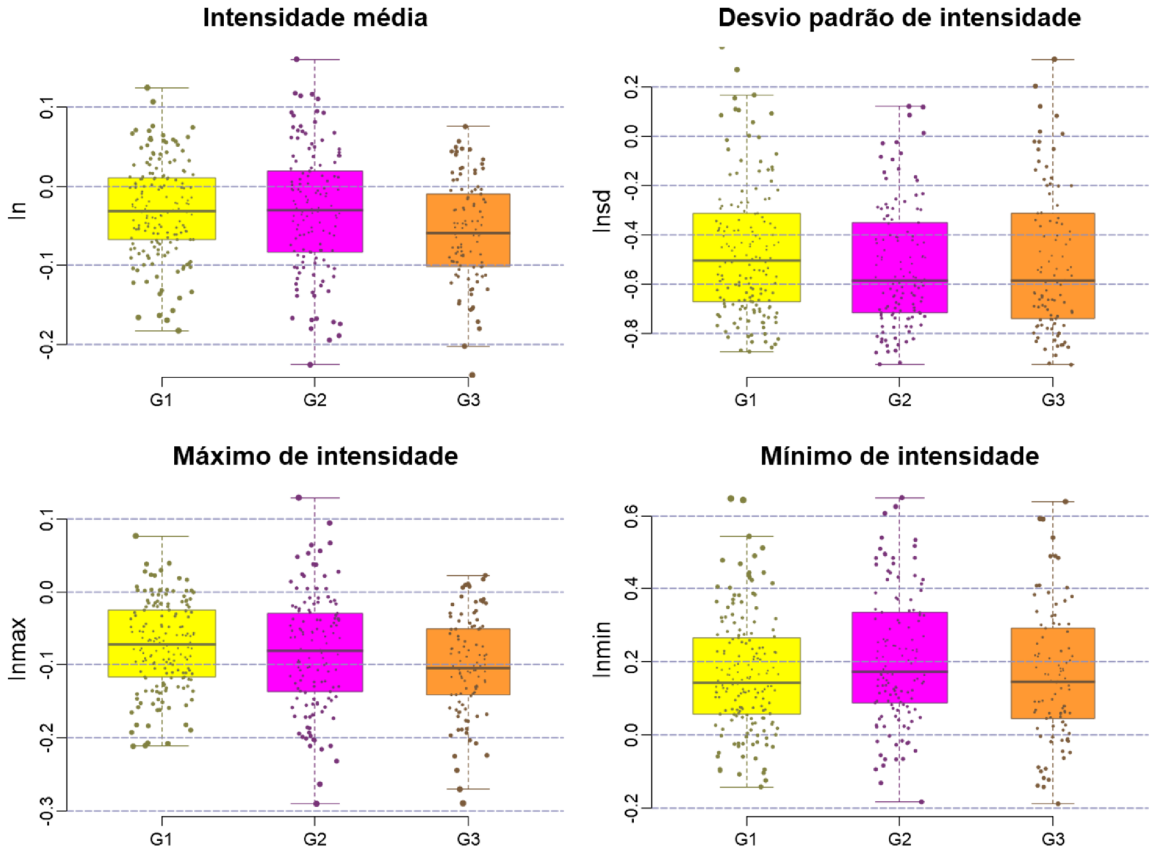


Figura 50: Distribuição dos parâmetros de intensidade por agrupamento.

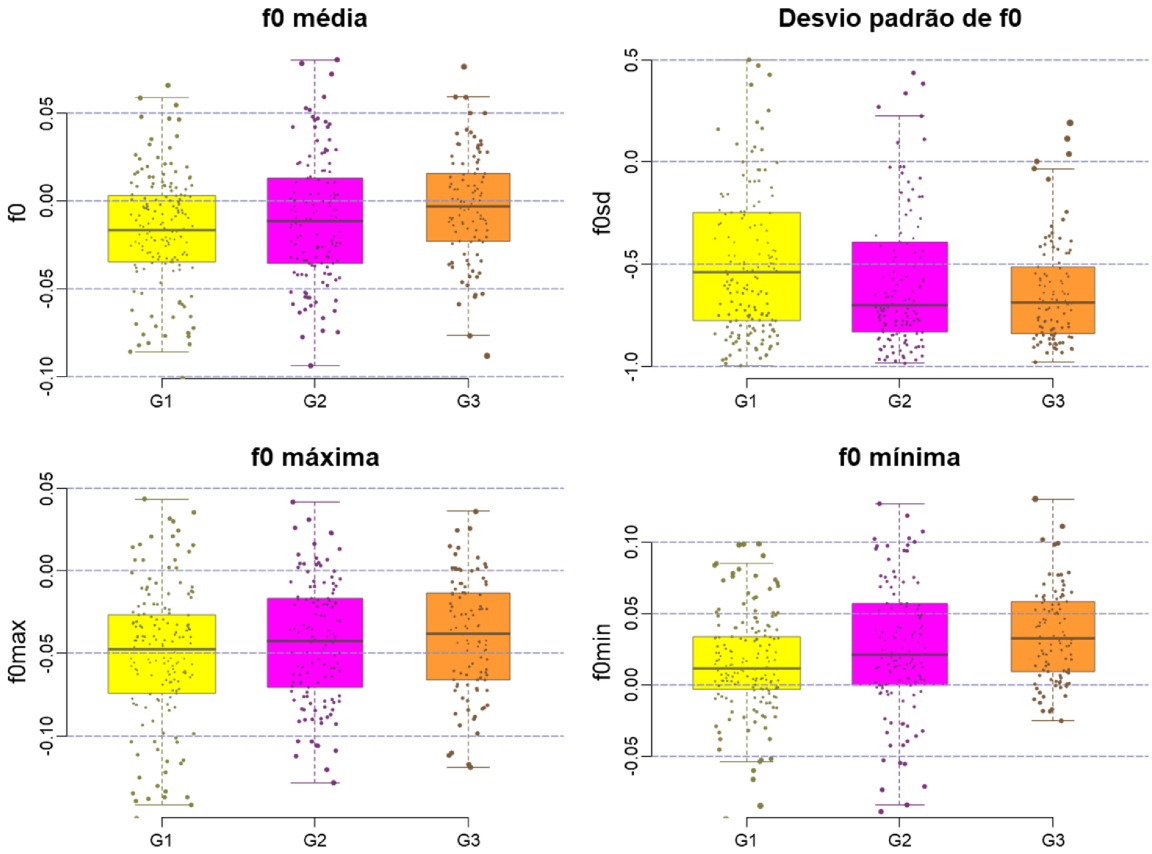


Figura 51: Distribuição dos parâmetros de f0 por agrupamento.

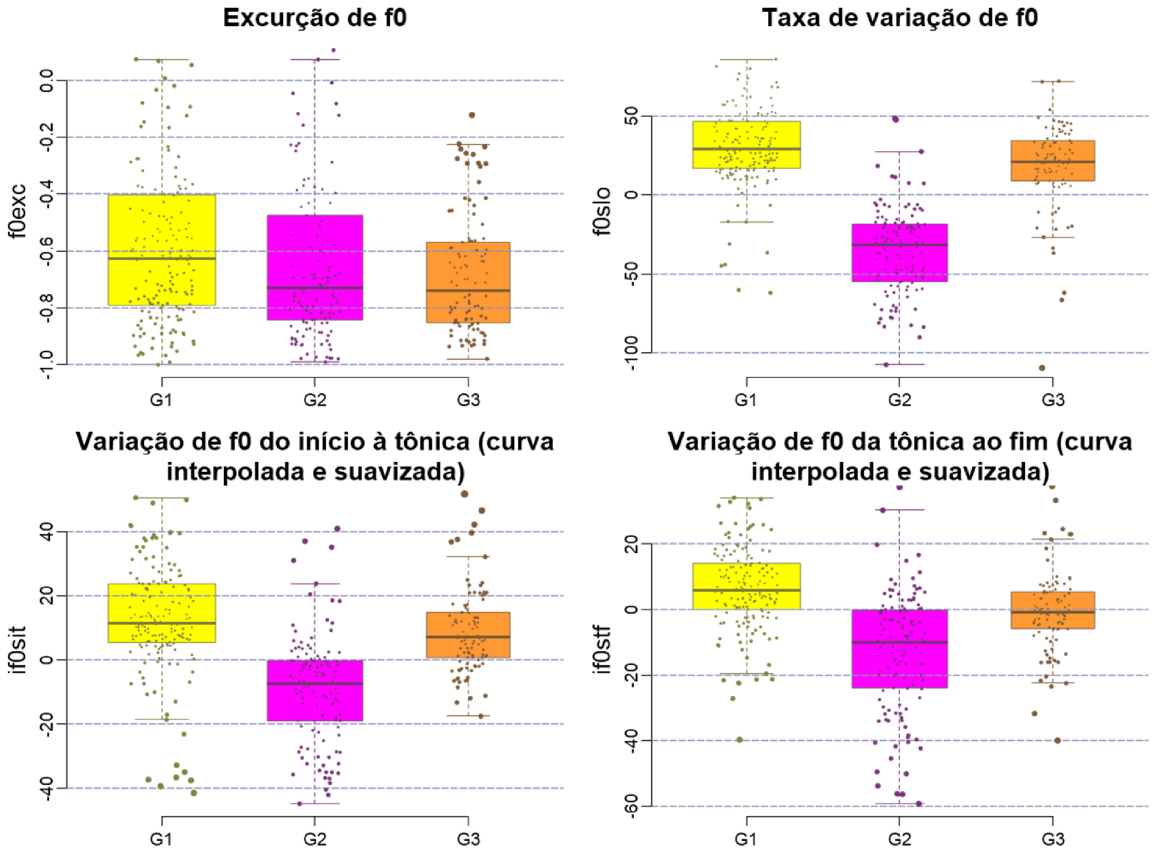


Figura 52: Distribuição dos parâmetros de variação de f0 por agrupamento.

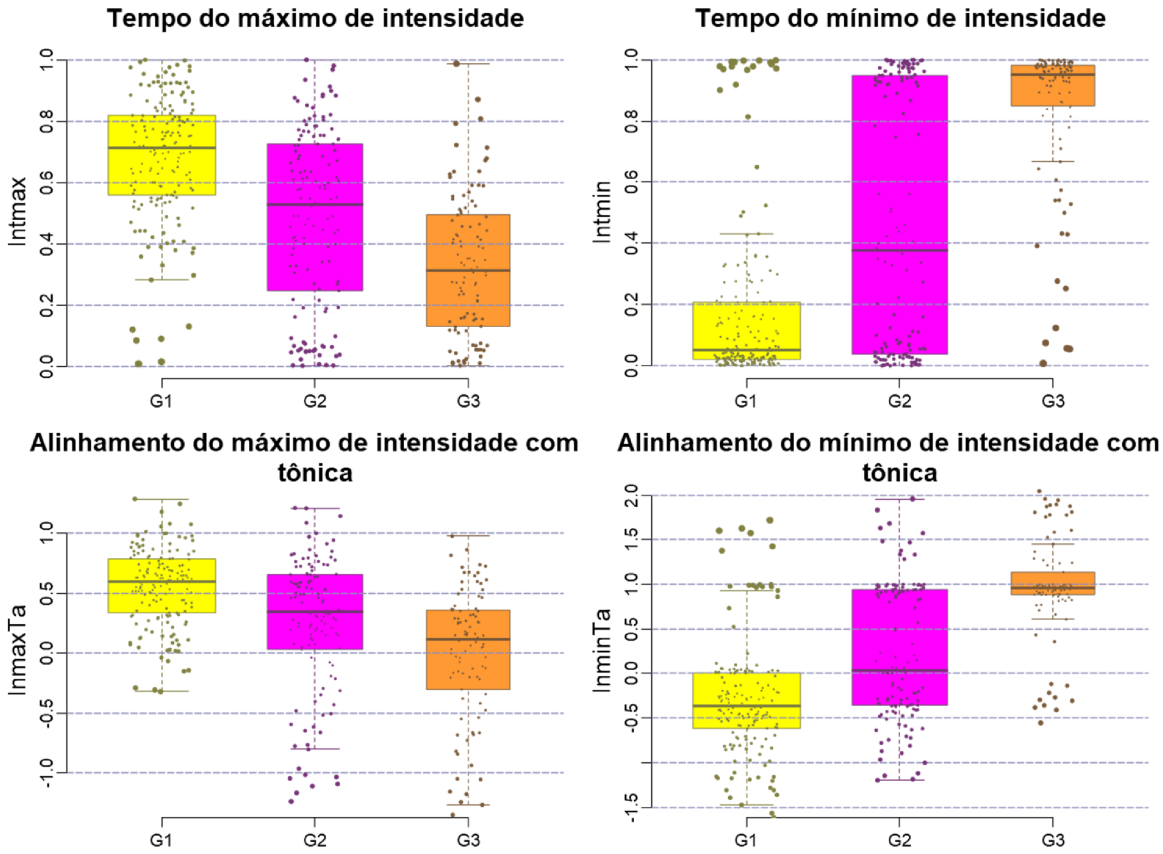


Figura 53: Distribuição dos parâmetros de alinhamento relativos à intensidade por agrupamento.

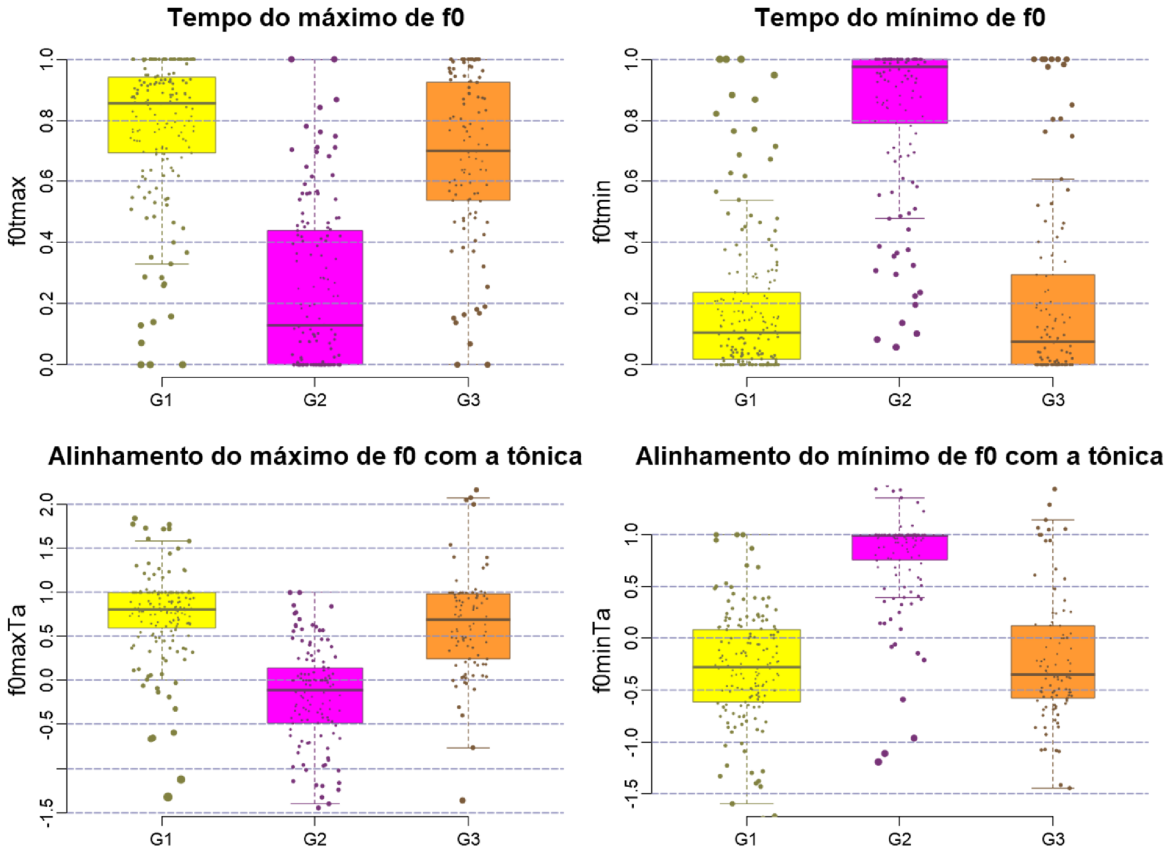


Figura 54: Distribuição dos parâmetros de alinhamento relativos à f0 por agrupamento.

8 CONSIDERAÇÕES FINAIS

Para concluir este trabalho, retomaremos os objetivos propostos inicialmente:

I. Validar através da modelagem estatística a forma das UDs já descritas em seus parâmetros prosódicos;
II. Identificar grupos com padrões de parâmetros semelhantes entre as unidades etiquetadas como AUX.

Sobre o objetivo (I), a aplicação da análise de discriminante linear (LDA) mostrou que as categorias de UDs do Mini BR Informal podem ser, em grande extensão, diferenciadas pelas características de seus parâmetros prosódicos em relação à unidade de comentário. A classificação obteve uma acurácia de 84,6% e indicou os parâmetros com maior efeito de distinção entre os tipos: a média de intensidade (*In*); a média de f0 (*f0*); a taxa de variação inicial da curva de f0 (*if0sit*) e a posição do máximo de f0 com relação à duração da unidade (*f0tmax*).

Sobre o objetivo (II), a análise de *cluster* realizada como os parâmetros levantados mostrou-se potencialmente útil, pois foi capaz de identificar grupos com diferenças significativas entre si, levantando a hipóteses de que eles representem diferentes categorias de UDs dentro do grupo etiquetado com AUX.

A classificação aplicada às UDs já descritas também se mostra útil como ferramenta de auxílio para a revisão da etiquetagem manual das UDs. As mal classificadas ou as distantes do centro de seus grupos são boas candidatas a terem sua etiqueta verificada. Adicionalmente, as UDs etiquetadas como AUX também podem ser avaliadas quanto à probabilidade de pertencerem a esses grupos através dos valores das funções discriminantes encontradas.

De modo geral, os resultados corroboram as conclusões obtidas por Raso e Vieira (2016) a respeito da importância da intensidade e da f0 média e das hipóteses sobre a importância do movimento de f0 em relação à vogal tônica para identificar as UDs dos tipos ALL, INP e CNT.

Não obstante, acreditamos que a investigação dos parâmetros mais relevantes para a separação entre os tipos, ou para sua identidade, pode vir a ser refinada através da aplicação de outras técnicas ou abordagens estatísticas.

Uma possibilidade interessante seria investigar a separação entre os grupos par a par (classificação binária), pois os parâmetros que melhor distinguem ALL de CNT podem ser diferentes dos que melhor distinguem CNT de INP; a avaliação pareada permitiria o foco em cada uma das fronteiras dos agrupamentos separadamente. Outra possibilidade é avaliar um tipo contra todas as unidades, como por exemplo ALL e ~ALL (não alocutivo). Esta abordagem seria útil para buscarmos os parâmetros que mais identificam cada grupo. Nesta opção, o conjunto de dados poderia ser ampliado pela inclusão das unidades etiquetadas como AUX.

Outras técnicas de classificação podem ser usadas para investigar os parâmetros. A análise do discriminante quadrático (QDA) é uma opção que não pressupõe igualdade das covariâncias entre os grupos e é capaz de separar os grupos por fronteiras curvas, ao contrário da LDA, que separa por retas (ver Figura 55, abaixo). Como consequência desvantajosa, a possibilidade de sobreajuste é maior na QDA. Entre outras técnicas que também podem apresentar melhor aderência, podemos citar o modelamento pela função logística, principalmente para classificação binária, e a *random forest* (Ho, 1995). Esta última apresenta como vantagens as capacidades de pesar a importância relativa de cada parâmetro e de reduzir o sobreajuste.

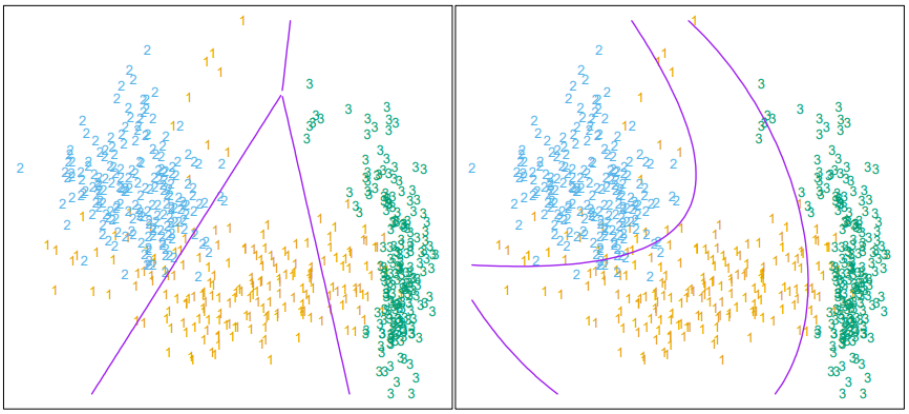


Figura 55: Exemplos de fronteiras de classificação linear (esquerda) e quadráticas (direita).[49]

Para buscar uma melhor classificação também é importante melhorar a qualidade das estimativas dos parâmetros e avaliar a necessidade de redefinição, inclusão ou exclusão de parâmetros.

[49] FRIEDMAN, HASTIE, e TIBSHIRANI (2008).

Um exemplo de parâmetro que pode ser substituído é a intensidade pela ênfase espectral (TRAUNMÜLLER e ERIKSSON, 2000), contribuição relativa das frequências altas à intensidade total. É uma medida de esforço vocal que não apresenta os problemas da medida de intensidade quanto à dependência da distância do microfone à fonte. Por outro lado, a ênfase espectral não dispensaria o uso de uma referência para comparação entre enunciados.

Uma das maiores dificuldades em compreender as formas prosódicas das UDs é a variabilidade do contorno melódico em função da sequência segmental. O INP é um tipo muito afetado pelo conteúdo segmental, podendo ter perfil de f0 ascendente, nivelado, descendente; ou ainda: ascendente-descendente; ascendente-nivelado; nivelado-descendente e ascendente-nivelado-descendente. Os parâmetros de f0 ou de variação de f0 escolhidos devem ser capazes de representar o seu formato e ao mesmo tempo ser independentes dessa variabilidade. Uma sugestão é a suavização da curva de f0. Neste trabalho, a suavização da curva de f0 do objeto de *pitch* foi utilizada apenas nos parâmetros *if0sit* e *if0stf*. Apesar de ser normalmente utilizada para reduzir os efeitos micromelódicos sobre a curva de f0, a aplicação dessas correções para as UDs, que são UTs curtas e isoladas, é algo que deve ser avaliado com cautela. Embora afetem pouco a percepção do contorno melódico do enunciado, podem alterar muito os valores de f0 localmente sem necessariamente aumentar a qualidade das estimativas.

Uma opção interessante pode ser utilizar as medidas de f0 ou de variação de f0 apenas dentro da vogal tônica. Além de complementar ou substituir os parâmetros de f0 utilizados neste estudo, o enfoque na vogal tônica poderia reduzir a variabilidade nas curvas de f0.

As medidas de duração silábica, normalizadas ou não, apresentaram pouca interferência na discriminação dos grupos, ao contrário do que era esperado. Também não foi possível concluir quais das três medidas é a mais relevante. Podemos supor que as análises tiveram dificuldade de lidar com a alta variabilidade dos parâmetros; essa hipótese pode ser avaliada aplicando-se a suavização das medidas de duração da UD a partir da média móvel ponderada com os valores de duração das sílabas vizinhas (BARBOSA, 2013). O inconveniente da suavização da duração é trazer informações de fora da UD para a sua estimativa, pelo fato de que a maioria das UD conterem apenas uma sílaba (um ponto de duração). Outra avaliação que pode ser

testada para melhorar a estimativa da duração é a segmentação em unidades VVs[50] (BARBOSA, 2006).

Os parâmetros de alinhamento dos pontos críticos de intensidade e f0 se mostram relevantes tanto para a discriminação quanto para o agrupamento. No entanto, parece haver margem para melhorar essa informação. Os altos coeficientes de correlação entre o posicionamento em relação à duração da UD e em relação à vogal tônica indicam redundância de informação. De fato, muitas UDs contêm apenas a vogal tônica (Ex.: *ah; é; ô; o')*, fazendo com que ambos os parâmetros tenham o mesmo valor. Parece ser o caso de que apenas uma das referências deva ser utilizada, não necessariamente a mesma para cada ponto crítico. Além disso, a distribuição de alguns desses parâmetros sugere que é possível a categorização dessas variáveis.

Além da aplicação de outras técnicas de classificação ou regressão e de melhorar a estimativa de parâmetros, uma estratégia recomendada é ampliar o conjunto de dados. Uma maior quantidade de dados aumenta a precisão dos modelos, reduz o intervalo de confiança, diminui a chance de sobreajuste e aumenta o poder dos testes de hipótese. Alternativamente, partindo-se de uma oferta maior de dados, pode-se aumentar a exigência de qualidade acústica e ainda assim obter quantidade suficiente de dados para as análises quantitativas. O *minicorpus* da seção formal do C-ORAL-BRASIL, em fase final de compilação, é a melhor recomendação para complementar o conjunto de dados das UDs.

Os próprios textos do corpus C-ORAL-BRASIL, além dos selecionados para compor o Mini BR Informal, podem ser usados para ampliação da base de dados, apesar de estes textos não terem sido etiquetados informacionalmente. Uma busca pelos itens lexicais mais frequentes das UD (Tabela 8) pode oferecer uma base de dados útil, mesmo que enviesada. Por exemplo, se buscarmos pelos nomes próprios em unidades tonais isoladas não ilocucionárias, é possível que se obtenha uma base com léxico controlado para investigação das diferenças entre ALL e CNT. Analogamente, o lexema *não* pode ser utilizado para comparar INP e CNT sem os efeitos da variabilidade gerada pela cadeia segmental.

[50] Separar as sílabas do início da vogal até o início da vogal seguinte.

Apesar de não ter sido o objetivo da aplicação da técnica neste trabalho, os resultados indicam potencial para uma possível classificação automática. A inclusão de variáveis distribucionais e lexicais podem aumentar muito o grau de previsibilidade de uma categoria de UD. Mas, antes disso, é preciso continuar a investigação para classificação das unidades ainda não categorizadas.

A análise de *cluster* apresentada nesse trabalho, mesmo que promissora, é apenas o início da investigação. Diferentemente das técnicas de classificação supervisionada, na qual se sabe de antemão quantas e quais são as categorias, a análise de agrupamento não tem critérios consensuais de avaliação. Rokach (2010) elenca uma dezena de possibilidades e separa os critérios de avaliação em dois tipos: i) critérios de qualidade interna e ii) critérios de qualidade externa. No primeiro tipo, são realizadas medidas de compactação dos grupos através de medidas quantitativas de similaridade entre os elementos do grupo, de separabilidade entre os grupos ou uma combinação de ambas. Nenhuma informação que não seja de dentro do conjunto de dados é utilizada na avaliação. Um exemplo de avaliação interna por separabilidade foi apresentado neste trabalho através do teste de Kruscal-Wallis (Tabela 28), que testa de há diferença significativa entre os grupos para cada parâmetro. No segundo tipo, os grupos são avaliados quanto à possibilidade de pertencerem a alguma categoria previamente estabelecida, como por exemplo, testar se os grupos formados têm identidade com as categorias ALL, CNT ou INP, ou ainda, avaliar a plausibilidade do agrupamento através de conhecimentos que estão não estão contidos nos valores de seus parâmetros. Para resolver a questão sobre quantas e quais são as categorias de UDs serão necessários ambos os tipos de validação.

As avaliações internas podem ser usadas como critério para seleção da técnica de *cluster*, para escolha do número de *clusters*, e para seleção e redução de parâmetros. Adicionalmente, as observações realizadas acima sobre as possibilidades de aperfeiçoamento das estimativas dos parâmetros prosódicos também se aplicam às técnicas de agrupamento (ex.: substituir intensidade por ênfase espectral). As possibilidades são muitas e a validação interna é um guia importante, ainda que a própria escolha do critério de validação seja sujeita a possibilidades amplas. Entretanto, por mais que validemos os grupos internamente, é fundamental conectar a compreensão do processamento matemático dos dados com a realidade linguística.

A tarefa fundamental e inevitável para a avaliar a qualidade do agrupamento será verificar a identidade pragmática dos grupos propostos, ou seja, submeter a teste a hipótese de que os elementos dos grupos reunidos pela similaridade de seus padrões prosódicos veiculam valores informacionais equivalentes.

> The validation of clustering structures is the most difficult and frustrating part of cluster analysis. Without a strong effort in this direction, cluster analysis will remain a black art accessible only to those true believers who have experience and great courage (JAIN e DUBES, 1988, p. 222)

REFERÊNCIAS

AIJMER, K; SIMON-VANDENBERGEN, A. M. Pragmatic markers. In: VERSCHUEREN, J. (ed.) *Handbook of Pragmatics*, John Benjamins, v. 13, 2009.

AUSTIN, L. J. *How to do things with words.* Oxford: Oxford University Press, 1962.

BARBOSA, P. Semi-automatic and automatic tools for generating prosodic descriptors for prosody research. In: BRIGITTE, B; Hirst, D. (eds.), *Proceedings of the Tools and Resources for the Analysis of Speech Prosody*, vol. 13, p. 86–89. Aix-en-Provence: Laboratoire Parole et Language, 2013.

BARBOSA, P. *Incursões em torno do ritmo da fala*. Campinas: Pontes, 2006.

BARBOSA, P. et al. Aiuruetê: a high-quality concatenative text-to-speech system for brazilian portuguese with demisyllabic analysis-based units and a hierarchical model of rhythm production. *Proceedings of the Eurospeech'99*, Budapest, Hungary, v. 5, 1999. p. 2059-2062.

BECKMAN, M. E.; EDWARDS, J. Lengthenings and shortenings and the nature of prosodic constituency. In KINGSTON, J.; BECKMAN, M.E. (eds.), *Papers in Laboratory Phonology I*: Between the Grammar and the Physics of Speech, p. 152-178. Cambridge University Press, 1990.

BELLMAN, R. E. *Adaptive Control Processes*. Princeton University Press: Princeton, 1961.

BICK, E. *The parsing system PALAVRAS:* automatic grammatical analysis of Portuguese in a constraint grammar framework. Aarhus: Aarhus University Press, 2000. Disponível em: <http://beta.visl.sdu.dk/~eckhard/pdf/PLP20-amilo.ps.pdf>.

BOERSMA, P.; WEENINK, D. *Praat*: Doing phonetics by computer. Programa de computador. Versão 6.0.38, 2018. Disponível em :<http://www.praat.org/>.

BOERSMA, P. Acustic Analysis. In: PODESVA R; SHARMA D. *Research Methods in Linguistics*. Cambridge University Press: Cambridge, 2013. p. 375-396.

BOERSMA. P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings 17.* Institute of Phonetic Sciences, University of Amsterdam, 1993. p. 97-110.

BRINTON, L. *Pragmatic Markers in English*: Grammaticalization and Discourse Functions. Berlin-Nova York: Monton de Gruyter, 1996.

CRESTI, E. The illocution-prosody relation and the information pattern in the spontaneous speech according to the Language into Act Theory (L-AcT). In: MORONI M; HEINZ, M (eds.), *Prosody*: Grammar, information structure, interaction. Special issue Linguistik, v 88, 2018. Disponível em: <https://bop.unibe.ch/index.php/linguistik-online/index>.

CRESTI, E. *Corpus di italiano parlato.* Firenze: Accademia della Crusca, 2000.

CRESTI, E.; MONEGLIA, M. *C-ORAL-ROM*: Integrated Reference Corpora for Spoken Romance Languages. Amsterdam-Philadelphia: John Benjamins, 2005.

CRYSTAL, D. *The English tone of voice*. London: Edward Arnold, 1975.

FISHER, K. Towards an understanding of the spectrum of approaches to discourse particles. In: FISHER, K. (ed.). *Approaches to discourse Marks*, 2006.

FOOLEN, A. Pragmatic markers. In: VERSCHUEREN, J. (ed.) *Handbook of Pragmatics*, John Benjamins, v. 2, 1996.

FOX, A. *Prosodic features and prosodic structure*: The phonology of suprasegmentals. Oxford University Press Inc: New York, 2000.

FRASER, B. A theory of discourse marks. In: FISHER, K. (Ed.). *Approaches to discourse Marks*, 2006.

FRIEDMAN, J; HASTIE, T; TIBSHIRANI, R;. Linear Methods for Classification. In: HASTIE, T; TIBSHIRANI, R; FRIEDMAN, J. *The Elements of Statistical Learning*: Data mining, inference and prediction. Springer-Verlag: New York, 2 ed, 2009. p. 101-137.

FROSALI, F. L'unità di informazione di ausilio dialogico: Valori percentuali, caratteri intonativi, lessicali e morfo-sintattici in un corpus di italiano parlato (C-ORAL-ROM). In: CRESTI, E. (ed.), *Prospettive nello studio del lessico italiano*. Firenze: University Press, 2006. p. 417-424.

FURNIVAL, G. M. WILSON, R. W. Regressions by Leaps and Bounds. *Technometrics*, v. 16, n. 4, 1974. p. 499-511.

GIRAUDOUX et al, *Spatial Analysis and Data Mining for Field Ecologists*: Package pgirmess. Versão 1.6.9, 2018. Programa de computador. Disponível em: <https://cran.r-project.org/web/packages/pgirmess/pgirmess.pdf>.

GUYON, I; ELISSEEFF, A. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, v. 3, 2003. Disponível em: < http://www.jmlr.org/ papers/volume3/guyon03a/guyon03a.pdf>.

HASTIE, T; TIBSHIRANI, R; FRIEDMAN, J. Unsupervised Learning. In: HASTIE, T; TIBSHIRANI, R; FRIEDMAN, J. *The Elements of Statistical Learning*: Data mining, inference and prediction. Springer-Verlag: New York, 2 ed, 2009. p. 485-585.

HIRST, D. A Praat plugin for MOMEL and INTSINT with improved algorithms for modelling and coding intonation. *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, 2007, p.1233-1236.

Ho, T. K. Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, 1995. p. 278-282. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=598994>.

JAIN, A. K, DUIN, R. MAO, J. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, 2000. p 4-37.

KARCEVSKY, S. Sur la phonologie de la phrase. *Travaux du Cercle linguistique de Prague IV*, 1931, p. 188-228.

KRUSKAL, W; WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*. n. 47 (260), 1952. p. 583-62.

KUHN, M. *Classification and Regression Training*: Package caret. Versão 6.0-81, 2018. Programa de computador. Disponível em: < https://cran.r-project.org/web/packages/caret/caret.pdf>.

LENK, U. Introduction: Pragmatic particles and discourse markers. In: VERSCHUEREN, J. (ed.) *Handbook of Pragmatics*, John Benjamins. v. 3, 1997.

LIANG, P. Lecture 4: Machine learning III. *Artificial Intelligence*: Principles and Techniques, Stanford University, 2018. Disponível em: <https://web.stanford.edu/class/cs221/lectures/learning3.pdf>.

LUMLEY, T; MILLER, A. Regression subset selection, including exhaustive search: Package leaps. Versão 3.0, 2017. Disponível em : <https://cran.r-project.org/web/packages/leaps/index.html >

MARTIN, P. *WinpitchW8*. Versão 2.00, 2014. Programa de computador. Disponível em: <http://www.winpitch.com/>.

MELLO, H; RASO T; MITTIMAN, M.; VALE, H.; CÔRTES, P. Transcrição e segmentação prosódica do *corpus* C-ORAL-BRASIL: critérios de implementação e validação. In: RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal.* Belo Horizonte: Editora UFMG, 2012.

MONEGLIA, M. The C-ORAL-ROM resource. In: CRESTI, E.; MONEGLIA, M. (eds.). CORAL-OM: *Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2005. p. 1-70.

MONEGLIA, M.; RASO, T. Notes on Language into Act Theory. In: RASO, T.; MELLO, H. (eds.). *Spoken Corpora and Linguistic Studies*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2014. p. 469–494.

OIKONOMAKOU N; VAZIRGIANNIS, M. A Review of Web Document Clustering Approaches. In: MAIMON O; ROKACH L. (eds) *Data Mining and Knowledge Discovery Handbook*. Springer-Verlag: New York, 2 ed, 2010. p. 931-948.

PEARSON. K. Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, n. 58, 1895. p. 240-242.

PETERSON, B. et al. *Econometric Tools for Performance and Risk Analysis*: Package PerformanceAnalytics. Versão 1.5.2, 2018. Programa de computador. Disponível em: <https://cran.r-project.org/web/packages/PerformanceAnalytics/PerformanceAnalytics.pdf>.

R CORE TEAM. *R*: A language and environment for statistical computing. R Foundation for Statistical Computing, Viena, Austria, 2016. Disponível em: <http://www.R-project.org/>.

RASO, T.; VIEIRA, M. A description of Dialogic Units/Discourse Markers in spontaneous speech corpora based on phonetic parameters. *Chimera*: Romance Corpora and Linguistic Studies, v. 3, p. 221-249, 2016. Disponível em: <https://revistas.uam.es/index.php/chimera/issue/archive>.

RASO, T. Prosodic constraints for discourse markers Tommaso. In: RASO, T.; MELLO, H. (eds.). *Spoken Corpora and Linguistic Studies*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2014.

RASO, T. O C-ORAL-BRASIL e a teoria da língua em ato. In: RASO, T.; MELLO, H. (eds.). *C-ORAL-BRASIL I:* Corpus de referência do português brasileiro falado informal*.* Belo Horizonte: *Editora* UFMG, 2012.

RASO, T.; MELLO, H. (Eds.). *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal.* 1. ed. Belo Horizonte: Editora UFMG, 2012.

ROKACH L. A survey of Clustering Algorithms. In: MAIMON O; ROKACH L. (eds) *Data Mining and Knowledge Discovery Handbook*. Springer-Verlag: New York, 2 ed, 2010. p. 269-298.

ROUSSEEUW, P et al. *Finding Groups in Data*: Cluster Analysis Extended: Package cluster. Versão 2.0.7-1, 2018. Programa de computador. Disponível em: <https://cran.r-project.org/web/packages/cluster/cluster.pdf>.

RSTUDIO. *RStudio*: Integrated development environment for R. Versão 1.0.153, 2017. Programa de computador. Disponível em: <http://www.rstudio.org/>.

SCHIFFRIN, *Discourse Marks*. Cambridge University Press, 1987.

SEARLE, J. *Speech acts*: an essay in the Philosophy of Language. Cambridge: Cambridge University Press, 1969.

SIEGEL, S.; CASTELLAN, N. *Non parametric statistics for the behavioural sciences*. New York: MacGraw Hill Int, 1988. p. 213-214.

TRAUNMÜLLER, H; ERIKSSON A. Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America*, n. 107, 2000. p. 3438-3451. Disponível em: <https://www.researchgate.net/publication/12444344_Acoustic_effects_of_variation_in_vocal_effort_by_men_women_and_children>.

VENABLES, W. N.; RIPLEY, B. D. *Modern Applied Statistics with S*. 4 ed. Springer: New York, 2002.

WARD, B. *What's Wrong with Economics*. Basic Books: New York, 1972.

APÊNDICES

APÊNDICE A: Áudio dos exemplos no texto

APÊNDICE B: Matriz de parâmetro s

APÊNDICE C: Dados usados no Praat

APÊNDICE D: *Scripts*

APÊNDICE E: Valores de f0 e intensidade por falante do Mini BR Informal

APÊNDICE A: Áudio dos exemplos no texto
(lista apenas, audios disponíveis apenas em meio eletrônico)

Exemplo ..... Página
Exemplo 1: bfamdl04[243-246]: 21
Exemplo 2: bfamdl04[153] 23
Exemplo 3: bfammn02[088]: 23
Exemplo 4: bfammn04[018-019] 24
Exemplo 5: bfammn05[140] 25
Exemplo 6: bpubdl01[014-015] 26
Exemplo 7: bpubdl01[119] 27
Exemplo 8: bpubdl01[197] 27
Exemplo 9: bfamdl04[072] 29
Exemplo 10: bfamcv03[247] 29
Exemplo 11: bfamdl02[038] 30
Exemplo 12: bfamcv04[241] 31
Exemplo 13: bfamcv03[063] 31
Exemplo 14: bfammn06[049] 33
Exemplo 15: ifamcv06[032] 34
Exemplo 16: bfammn04[029] 35
Exemplo 17: bfamdl01[167] 36
Exemplo 18: bpubmn01[077] 36
Exemplo 19: bfammn01[014] 37
Exemplo 20: bfamcv02[095-107]: 41
Exemplo 21: bfamcv02[095-107]: 42
Exemplo 22: bfamcv02[045] 46
Exemplo 23: bpubcv02[194] 47
Exemplo 24: bfamdl01[028] 47
Exemplo 25: bpubmn01[048] 50
Exemplo 26: bfammn05[090] 55
Exemplo 27: bfammn02[009] 69

APÊNDICE B: Matriz de parâmetros

(Disponível apenas em meio eletrônico)

APÊNDICE B1: ENUs UTs.xlsx: Lista dos enunciados e das unidades tonais do Mini BR Informal *revisado;*

APÊNDICE B2: UD Re Par.xlsx: Dados das UDs, das unidades de referência e dos parâmetros analisados;

APÊNDICE B3: Pitch.esp.xls: Especificações usadas para criação do objeto *pitch* do Praat.

APÊNDICE C: Dados usados no Praat
(Disponível apenas em meio eletrônico)

APÊNDICE C1: Áudio: Arquivos de aúdio dos enuciados recortados*;*
APÊNDICE C2: TextGridUD: Arquivos de TextGrid (segmentação);
APÊNDICE C3: DurSil: Arquivos de TextGrid (durações silábicas).
APÊNDICE C4: Pitch: Arquivos de Pitch.

APÊNDICE D: *Scripts*

APÊNDICE D1: *Scripts* R

APÊNDICE D1.1: Cria lista de enunciados e de unidades tonais

APÊNDICE D1.2: Calcula os parâmetros de acordo com a o tipo de referência e realiza a estatística descritiva

APÊNDICE D1.3: *Leaps* and *bounds* e LDA

APÊNDICE D1.4: Análise de *cluster* via agnes

APÊNDICE D2: *Scripts* PRAAT

APÊNDICE D2.1: Corta os audios do Mini BR Informal por enunciado e cria textgrid contendo a transcrição, as unidades tonais desalinhadas

APÊNDICE D2.2: Calcula a duração normalizada por sílaba e por unidade tonal

APÊNDICE D2.3: Tabela por UT a segmentação silábica, a duração e a duração normalizada

APÊNDICE D2.4: Calcula as medidas acústicas das unidades tonais

```
# APÊNDICE D1.1 Cria lista de enunciados e de unidades tonais

library(readxl)
setwd("C:/Users/olive/Documents/tmpLEEL/Work/MC_R_UD")
sheets <- excel_sheets("MiniCorpus_R_UD.xlsx")

mc <- lapply(sheets,read_excel, path = "MiniCorpus_R_UD.xlsx")
mc_data <- data.frame(0)
mc_ut <- data.frame(0)

for(t in 1:length(sheets)) {
  mc[[t]] <- cbind(rep(sheets[t], nrow(mc[[t]])), mc[[t]])
  if (t == 1) {mc_data <- mc[[1]]}
  else {mc_data <- rbind(mc_data, mc[[t]])}
}

names(mc_data) <- c("File","Rank","Speaker", "Text", "E_Start", "E_End", "E_Dur",
"Channel", "Tag")

mc_data[[4]]<-gsub("\\+", "+@", mc_data[[4]])
mc_data[[4]]<-gsub("\\//", "@@", mc_data[[4]])
mc_data[[4]]<-gsub("\\[/([0-9])\\]", "@\\1", mc_data[[4]])
mc_data[[4]]<-gsub("/", "!@", mc_data[[4]])

igual <- gregexpr("=",mc_data[[4]])
u <- 1

mc_ut<-data.frame(1, mc_data[1,c(1,2,3)],1,1,"1","1","1", stringsAsFactors=FALSE)
names(mc_ut) <- c("Enu","File","Rank","Speaker", "Tot_UT", "No_UT", "Text",
"Fron", "Tag")

for (e in 1:nrow(mc_data)){
  nt = length(igual[[e]])/2
  igual[[e]]<-c(0,as.integer(igual[[e]]))

  if (nt < 1) {
    mc_ut[u,] <- c(e, mc_data[e,c(1,2,3)], nt, 1, mc_data[[e,4]], "", "")
    u <- u + 1
    print(e)
      }
  else {
    for (t in 1:nt){
      mc_ut[u,] <- c(e, mc_data[e,c(1,2,3)], nt, t,
                     substr(mc_data[[e,4]],igual[[e]][2*t-1]+1,igual[[e]][2*t]-3),
                     substr(mc_data[[e,4]],igual[[e]][2*t]-2,igual[[e]][2*t]-1),
                     substr(mc_data[[e,4]],igual[[e]][2*t]+1,igual[[e]][2*t+1]-1))
      u <- u + 1
     }
  }
}

mc_ut$Fron<-gsub("\\+@", "+", mc_ut$Fron)
mc_ut$Fron<-gsub("@@", "//", mc_ut$Fron)
mc_ut$Fron<-gsub("@([0-9])", "\\[/\\1\\]", mc_ut$Fron)
mc_ut$Fron<-gsub("!@", "/", mc_ut$Fron)
mc_ut$Text<-gsub("\\+@", "+", mc_ut$Text)
mc_ut$Text<-gsub("@@", "//", mc_ut$Text)
mc_ut$Text<-gsub("@([0-9])", "\\[/\\1\\]", mc_ut$Text)
mc_ut$Text<-gsub("!@", "/", mc_ut$Text)
```

```
mc_data$Text<-gsub("\\+@", "+", mc_data$Text)
mc_data$Text<-gsub("@@", "//", mc_data$Text)
mc_data$Text<-gsub("@([0-9])", "\\[/\\1\\]", mc_data$Text)
mc_data$Text<-gsub("!@", "/", mc_data$Text)

mc_ut<-cbind(mc_ut,t(as.data.frame.list(lapply(gregexpr(" ",mc_ut$Text),
length)))-(substr(mc_ut$Text,nchar(mc_ut$Text),nchar(mc_ut$Text))==" ") -
(substr(mc_ut$Text,1,1)==" ") + 1)

names(mc_ut) <- c(names(mc_ut[1:9]),"W_C")
mc_ut$Tag <- as.factor(mc_ut$Tag)

write.table(mc_ut, "MC_R_UD_UT.txt", sep="\t", row.names = FALSE, qmethod =
"double", quote= FALSE)
write.table(mc_data, "MC_R_UD_ENU.txt", sep="\t", row.names = FALSE, qmethod =
"double", quote= FALSE)

mc_enu<-cbind("Enu"=1:nrow(mc_data),mc_data)
mc_enu$n_UD <- 0
na_mc_enu<-names(mc_enu)
na_mc_enu[10]<-"Tot_UT"
names(mc_enu)<-na_mc_enu
n_enu<-nrow(mc_enu)

tag_ud <- c("CNT", "INP", "ALL", 'AUX')
tag_ud <- c(tag_ud,paste(tag_ud,"_r", sep = ""))
n_tag_ud <- sapply(1:length(tag_ud),function(x)
which(levels(mc_ut$Tag)==tag_ud[x]))

enu_ut <- data.frame()
enu_fron  <- data.frame()
enu_tag <- data.frame()

for (e in 1:n_enu){

  e_ut = which(mc_ut$Enu==mc_enu$Enu[e])
  for(u in mc_ut$No_UT[e_ut]) {
    enu_ut[e, u] <- mc_ut$Text[e_ut][u]
    enu_fron[e, u] <- mc_ut$Fron[e_ut][u]
    enu_tag[e, u] <-  mc_ut$Tag[e_ut][u]

    mc_enu$n_UD[e] <- mc_enu$n_UD[e] + sum(as.integer(enu_tag[e,u]) == n_tag_ud)

  }
  mc_enu$Tot_UT[e] <- sum(!is.na(t(enu_ut[e,])))

}

names(enu_ut)<-1:ncol(enu_ut)
names(enu_fron)<-1:ncol(enu_fron)
names(enu_tag)<-1:ncol(enu_tag)

write.table(mc_enu,"MC_ENU.txt",sep="\t",row.names = FALSE, qmethod =
"double",quote= FALSE)
write.table(enu_ut,"ENU_UT.txt",sep="\t",row.names = FALSE, qmethod =
"double",quote= FALSE)
write.table(enu_fron,"ENU_FRON.txt",sep="\t",row.names = FALSE, qmethod =
"double",quote= FALSE)
```

```
write.table(enu_tag,"ENU_TAG.txt",sep="\t",row.names = FALSE, qmethod =
"double",quote= FALSE)

enu_sob_sta <- rep(1,n_enu)%*%t(matrix(mc_enu$E_Start))
enu_sob_end <- rep(1,n_enu)%*%t(matrix(mc_enu$E_End))
```

```
# APÊNDICE D1.2 Calcula os parâmetros de acordo com a o tipo de referência e
realiza a estatística descritiva

library("readxl")
library("MASS")
library("caret")
library("psych")
library("Cairo")
library("PerformanceAnalytics")
library("pgirmess")

dir <- "C:/Users/olive/OneDrive/Documentos/Work/MC_R_UD/"
file <- "UT_SIL_V6.xlsx"

var   <- read_excel(path = paste(dir, file, sep = ""), sheet = "Var", na = "",
range = "A1:l60")
UDdat <- read_excel(path = paste(dir, file, sep = ""), sheet = "UD_Dat", na =
"NA")
Redat <- read_excel(path = paste(dir, file, sep = ""), sheet = "Re_Dat", na =
"NA")

hUDdat <- UDdat[, vinf]
hUDdat$Tag <- substr(hUDdat$Tag, 1 , 3)

vinf <- which(var$inf == 1)
vval <- which(var$val == 1)
vdif <- which(var$dif == 1)
vpdf <- which(var$pdf == 1)

aci <- which(UDdat$AUX == 0)
a_c <- which(hUDdat$Tag == "CNT" | hUDdat$Tag == "ALL")
a_i <- which(hUDdat$Tag == "ALL" | hUDdat$Tag == "INP")
c_i <- which(hUDdat$Tag == "CNT" | hUDdat$Tag == "INP")
aux <- which(UDdat$AUX == 1)
neh <- which(UDdat$né == 1)
ann <- which(UDdat$AUX == 1 & UDdat$né == 0)
all <- 1:nrow(UDdat)

sel = all

dUDdat <- cbind("Tag" = factor(hUDdat$Tag[all], levels = c("ALL", "CNT", "INP",
"AUX")),
                UDdat[all, vval],
               (UDdat[all, vpdf] - Redat[all, vpdf])/abs(Redat[all, vpdf]),
                UDdat[all, vdif] - Redat[all, vdif])
o <- order(sapply(1:ncol(dUDdat), FUN = function(x) which(var$Var ==
names(dUDdat)[x])))
dUDdat <- dUDdat[, o]

sum.dUDdat <- describeBy(dUDdat[, -1 ], group = dUDdat$Tag, mat = TRUE)
names(sum.dUDdat) <- c("No", "Etiqueta", "Variável", "Número", "Média", "Desvio
padrão", "Mediana", "Média truncada", "Desvio da mediana", "Mínimo", "Máximo",
"Amplitude", "Assimetria", "Curtose", "Erro padrão")
sum.dUDdat <- cbind( "Var" = substr(row.names(sum.dUDdat), 1,
nchar(row.names(sum.dUDdat)) - 1), sum.dUDdat[, -c(1, 3, 9, 15)])
row.names(sum.dUDdat) <- 1:nrow(sum.dUDdat)
write.table(sum.dUDdat, paste(dir, "VarSum.txt", sep = ""), sep = "\t", row.names
= FALSE, quote = FALSE)
```

```
shapT <- sapply(2:ncol(dUDdat), function(x) shapiro.test(dUDdat[, x])[[2]])
names(shapT) <- names(dUDdat[,-1])

krTok <- matrix(0, 6, ncol(dUDdat) - 1)

for (v in 2:ncol(dUDdat)) {
  krt <- kruskalmc(dUDdat[, v], dUDdat$Tag, p = 0.05)
  krTok[, v - 1] <- krt$dif.com[, 3]
}
krTok <- data.frame(row.names(krt$dif.com), krTok)
names(krTok) <- names(dUDdat)
write.table(krTok, "clipboard", sep = "\t")


col = c("#FF0000", "#00BB28", "#2222FF", "#839192")
lcol = c("#603030", "#306030", "#303060", "#303030")

for (k in 2:ncol(dUDdat)) {

  data = dUDdat[, c(1,k)]
  j <- which(names(dUDdat)[k] == var$Var)
  ylab <- var$Var[j]
  if (!is.na(var$uni[j])) ylab <- paste(ylab, " (", var$uni[j], ")", sep = "")
  tit <- paste(strwrap(var$Descrição[j], 45), collapse="\n")


  Cairo(file  = paste(dir, "Plots/",  k, "BP_", var$Var[j], ".png", sep = ""),
        type  = "png", pointsize = 18)
  par(mgp = c(2, 0.6, 0))
  par(mar = c(2.5, 3.5, 3, 1))
  boxplot(dUDdat[, k] ~ data$Tag , col = col, horizontal = FALSE, varwidth = TRUE,
          ylab = ylab, main = tit, cex.main = 1.5, outline = FALSE, boxwex = .7,
          cex.lab = 1.4, frame.plot = FALSE, border = lcol)
  grid(nx = NA, ny = NULL, col = "#9090BB", lty = 5,
       lwd = 1.5*par("lwd"), equilogs = TRUE)
  tlev<-levels(data$Tag)
  lprops<-summary(data$Tag)/nrow(data)


  for(i in 1:length(tlev)){

    lev <- tlev[i]
    val <- data[data$Tag == lev, 2]
    q <- quantile(val)

    jit <- jitter(rep(i, length(val)), amount = lprops[i]/2)
    points(jit, val, pch = 20, cex = abs(scale(val))/3, col = lcol[i])

  }

  dev.off()

  par(mgp = c(1.7, 0.5,0))
  par(mar = c(3, 3.5, 2, 1))
  bp <- barplot(table(dUDdat[, c(1, 43)]), beside = T, col = col, legend.text =
tlev, xlim =  c(-6, 100),
                ylab = "N_sil", horiz = TRUE, main = "Número de sílabas", xlab =
"Número de ocorrências",
                cex.sub = 1.5, cex.axis = 1.0, cex.main = 1.5, cex.lab = 1.3)
```

```
  text(-3, bp,  labels = table(dUDdat[, c(1, 43)]))

}
dev.off()

Cairo(file  = paste(dir, "Plots/",  "BarP_N_sil.png", sep = ""),
      type  = "png", pointsize = 12)
dev.off()


gg <- c("d", "in", "f0", "vf", "ai", "af")
for (g in gg){
  chart.Correlation(dUDdat[, var$Var[which(var$t4 == g)]], histogram=TRUE, pch=18)
  Cairo(file  = paste(dir, "Plots/CO",  g, ".png", sep = "", width = 1000),
        type  = "png", pointsize = 12)
}
dev.off()

par(mgp = c(1.7, 0.5,0))
par(mar = c(3, 3.5, 2, 1))
bp <- barplot(table(dUDdat[, c(1, 43)]), beside = T, col = col, legend.text =
tlev, xlim =  c(-6, 100),
              ylab = "N_sil", horiz = TRUE, main = "Número de sílabas", xlab =
"Número de ocorrências",
              cex.sub = 1.5, cex.axis = 1.0, cex.main = 1.5, cex.lab = 1.3)
text(-3, bp,  labels = table(dUDdat[, c(1, 43)]))


gg <- c("d", "in", "f0", "vf", "ai", "af")
for (g in gg){

  chart.Correlation(dUDdat[, var$Var[which(var$t4 == g)]], histogram=TRUE, pch=12,
cex.lab = 3)
}
```

```
# APÊNDICE D1.3 Leaps and bounds e LDA

library("readxl")
library("MASS")
library("caret")
library("leaps")
library("cluster")
library("pgirmess")

var_d <-var[var$Var%in%names(dUDdat),]
var_l <- var_d[which(var_d$t2!="s" & var_d$t2!="i" & var_d$t2 != "x" & var_d$t1 !=
"i"), ]

sel <- aci

lUDdat <- dUDdat[sel, c("Tag", var_l$Var)]
lUDdat$Tag <- factor(lUDdat$Tag)

REGS <- regsubsets(lUDdat[, -1], lUDdat$Tag, int = TRUE, names = names(lUDdat[-
1]),
                   nbest = 60, nvmax = 30, really.big = TRUE)
RegsW <- summary(REGS)$which
RegsS <- as.numeric(row.names(RegsW))

plot(RegsS, summary(REGS)$cp, col = "#008810", log = "y",
     xlab = "Número de parâmetros", ylab = "Estatística Cp", cex.lab = 1.2)
points(1:nrow(var_l), 1:nrow(var_l), type = "l", col = 6, cex = 0.7)
legend(x = nrow(var_l) - 7, y = 0.3, legend = c("Cp = p"), col = 6, lty = 1, bty
="n")

par(las = 2)
par(mar = c(4, 3.5, 2, 1))
barplot(sort(colSums(RegsW)/nrow(RegsW)*100, decreasing = TRUE)[-1], col = 3, ylim
= c(0, 100),
        ylab = "% de seleção da variável", cex.lab = 1.2)

nmod <-nrow(RegsW)
TACC <- rep(list(matrix(0, nmod, length(levels(lUDdat$Tag)))), 11)
OACC <- rep(0, nmod)

for (n in 1:nmod){

  LDA <- lda(Tag ~., data = lUDdat[, c("Tag", names(which(RegsW[n,]))[-1])])
  mTag <- predict(LDA, lUDdat)$class
  conf <- confusionMatrix(mTag, lUDdat$Tag)
  OACC[n] <- conf$overall[1]
  for (m in 1:11) TACC[[m]][n,] <- conf$byClass[, m]

}
par(xaxs="i")

lsVAR <- names(which(RegsW[which.max(OACC),]))[-1]
LDA <- lda(Tag ~., data = lUDdat[, c("Tag", lsVAR)])
mTag <- predict(LDA, lUDdat)$class
conf <- confusionMatrix(mTag, lUDdat$Tag)
lsVAR; conf

par(mar = c(4, 5.3, 2, 1))
par(mgp = c(3, 0.5,0))
```

```
cores = rgb(t((col2rgb(as.numeric(lUDdat$Tag) + 1) + col2rgb(as.numeric(mTag) +
1))/2/255))
plot(LDA, col = cores, main =  paste(c(length(lsVAR), "p:", lsVAR), collapse = "
"),
     ylim = c(-3, 3), xlim = c(-4, 4), cex.lab = 3, cex = 2, cex.main = 3,
cex.axis = 3)
pot <- LDA$svd^2/sum(LDA$svd^2)

sdata <- data.frame("Tag" = lUDdat$Tag, scale(lUDdat[, lsVAR]));
sLDA <- lda(Tag ~., data = sdata)
cLDA <- cbind(coef(sLDA), coef(LDA))
cLDA <- cLDA[order(cLDA[, 1], decreasing = TRUE), ]
write.table(cLDA, "clipboard", sep = "\t")

nv <- REGS$np - 1
maxoa <- sapply(1:nv, function(x) max(OACC[which(RegsS == x)]))
maxaa <- sapply(1:nv, function(x) max(TACC[[11]][which(RegsS == x), 1]))
maxca <- sapply(1:nv, function(x) max(TACC[[11]][which(RegsS == x), 2]))
maxia <- sapply(1:nv, function(x) max(TACC[[11]][which(RegsS == x), 3]))

par(mgp = c(2, 0.5,0))
par(mar = c(3, 3.5, 2, 1))
 plot(1:nv, maxoa, col = 1, type = "b", ylim = c(0.5, 1), xlab = "Número de
parâmetros",
      ylab = "Acurácia máxima (%)", cex.lab = 1.2, frame.plot = FALSE, xlim = c(0,
27))
lines(1:nv, maxaa, col = 2, type = "b")
lines(1:nv, maxca, col = 3, type = "b")
lines(1:nv, maxia, col = 4, type = "b")
legend(x = nv - 8, y = 0.7, legend = c("Global", "ALL", "CNT", "INP"), col = 1:4,
lty = 1, bty = "n")
grid(nx = NA, ny = NULL, col = "lightgray", lty = "dotted",
     lwd = par("lwd"), equilogs = TRUE)

k = 10
kRegsW <- RegsW[which(RegsS ==  k),]
kOACC <- OACC[which(RegsS ==  k)]
kVAR <- names(which(kRegsW[which.max(kOACC),]))[-1]
kLDA <- lda(Tag ~., data = lUDdat[, c("Tag", kVAR)])
kVAR
kPRED <- predict(kLDA, lUDdat)
kconf <- confusionMatrix(kPRED$class, lUDdat$Tag)
kconf

cores <- rgb(t((col2rgb(as.numeric(lUDdat$Tag) + 1) +
col2rgb(as.numeric(kPRED$class) + 1))/2/255))
plot(kLDA, col = cores, main =  paste(c(k, "p", ":", kVAR), collapse = "  "))

if (sum(max(kOACC)==kOACC) > 1) {
t(sapply(1: nrow(kRegsW[which(kOACC == max(kOACC)),]), function(x)
names(which(kRegsW[which(kOACC == max(kOACC)),][x,]))[-1]))
}
```

```
#APÊNDICE D1-4 Analise de cluster via agnes

library("readxl")
library("MASS")
library("caret")
library("leaps")
library("cluster")
library("pgirmess")

var_l <- var_d[which(var_d$t2!="s" & var_d$t2!="i" & var_d$t2 != "x" & var_d$t1 !=
"i"), ]
#var_l <- var_l[c(-21, -22),]

op <- 4
myVAR <-c( "In", "Inmin", "Intmax", "InminTa", "f0", "f0max", "f0min", "f0tmax",
"f0exc", "f0slo",  "f0minTa", "if0sit", "if0stf", "N_Sil", "Dur", "Dm_sil")
rVAR <- names(sort(colSums(RegsW)/nrow(RegsW)*100, decreasing = TRUE))
cVAR <- list(myVAR, lsVAR, rVAR[1:15], var_l$Var)[[op]]

sel <- aux
cUDdat <- dUDdat[sel, cVAR]

g = 3
tClag <- agnes(scale(cUDdat), method = "ward")
par(mar = c(1, 4, 2, 1))
plot(tClag, which.plots = 2, cex.main = 1.5,  main = "Agrupamento Aglomerativo")
#Todos parâmetros; todas UDs AUX") #,  main =  paste(c(ncol(cUDdat), "p's:",
names(cUDdat)), collapse = "  "))
tGrag <- data.frame("Grupo" = cutree(tClag, k = g))
tgru <- table(tGrag$Grupo)
#row.names(gru) <-sapply(1:nrow(gru), function(x) names(which.max(gru[x,])))
#gru3 <- rowsum(gru, rownames(gru))
sVAR; tgru
gUDdat <- cbind(tGrag, cUDdat)
Caux.m <- cbind("N" = as.vector(tgru), aggregate(na.exclude(gUDdat)[-1],
list(na.exclude(gUDdat)$Grupo), mean))
aci.m <- cbind("N" = summary(lUDdat$Tag), aggregate(na.exclude(lUDdat)[-1],
list(na.exclude(lUDdat)$Tag), mean))
aci.m
Caux.m
write.table(Caux.m, "clipboard", sep = "\t")
ACI.m <- aggregate(. ~ Tag,  mean, data = dUDdat[aci, c("Tag", cVAR)])
write.table(ACI.m, "clipboard", sep = "\t")

ckrt <- 0
ckrTok <- matrix(0, ncol(combn(g, 2)), ncol(cUDdat))

for (v in 1:ncol(cUDdat)) {
  ckrt <- kruskalmc(cUDdat[, v], tGrag, p = 0.05)
  ckrTok[, v - 1] <- ckrt$dif.com[, 3]
}
ckrTok <- data.frame(row.names(ckrt$dif.com), ckrTok)
names(ckrTok) <- c("par", names(cUDdat))
ckrTok

write.table(ckrTok, "clipboard", sep = "\t")
```

```
# APÊNDICE D2.1
# corta os audios do Mini BR Informal por enunciado e cria um textgrid contendo
# a transcrição, as unidades tonais (desalinhadas), as etiquetas informacionais

clearinfo
form Escolha
  sentence Local C:\Users\LEEL\Documents\Oliver\Work\MC_R_UD\
  choice Fam: 1
    button fam
    button pub
  choice Diatype: 1
    button cv
    button dl
    button mn
  choice Number: 1
    button 01
    button 02
    button 03
    button 04
    button 05
    button 06
endform

audio$ = "b" + fam$ + diatype$ + number$

#local$ = "C:\Users\OG\Documents\Mestrado\Work\"

enu_ut   = Read from file: local$ + "ENU_UT.txt"
enu_tag  = Read from file: local$ + "ENU_TAG.txt"
spk_ch  = Read from file: local$ + "SPK_Ch.txt"
n_spk = Get number of rows
mc_enu   = Read from file: local$ + "MC_ENU.txt"
n_enu = Get number of rows

lista = Create Strings as file list: "list_arqs", local$ + "MC_R_UD\" + "*.wav"

n_arq = Get number of strings

for a to n_arq
  selectObject: lista
  file$ = Get string: a
  Open long sound file: local$ + "MC_R_UD\" + file$
  Rename: file$ - "-final.wav"
endfor

for enu to n_enu
  selectObject: mc_enu
  file$ = Get value: enu, "File"
  enu_sta = Get value: enu, "E_Start"
  enu_end = Get value: enu, "E_End"
  enu_tex$ = Get value: enu, "Text"
  enu_spk$ = Get value: enu, "Speaker"
  enu_rank = Get value: enu, "Rank"
  n_ut = Get value: enu, "Tot_UT"
  enu_ran$ = mid$(string$(enu_rank + 1000), 2, 3)
  enu_dur = enu_end - enu_sta
  d = 0.3
  enu_tem_ud = index_regex(enu_tex$, "(AUX|CNT|ALL|INP)(_r|)[^_n]") > 0
```

```
  if enu_tem_ud and audio$ == file$

    selectObject: spk_ch
      for s to n_spk
        sfile$ = Get value: s, "audio"
        spk$ = Get value: s, "spk"
        ch$ =  Get value: s, "Ch"
        if spk$ = enu_spk$ and sfile$ = file$
          enu_ch$ = ch$
          print 'spk$' 'enu_spk$' 'enu_ch$''newline$'
        endif
    endfor

    sub_fol$ = local$ + "MC_ENU\" + file$ + "\"

    createDirectory: sub_fol$
    createDirectory: sub_fol$ + "och\"

    print
'file$''tab$''enu_ran$''tab$''enu_sta''tab$''enu_end''tab$''n_ut''tab$''newline$'

    selectObject: "LongSound " + file$
    s_end = Get end time
    Extract part: max(0, enu_sta - d), min(enu_end +d, s_end), 1
    ch = Get number of channels

    if ch = 2

      if enu_ch$ = "I"
       Save as WAV file: sub_fol$ + file$ + "__" + enu_ran$ + "__" + enu_spk$ +
"__c12.wav"
      endif

      Extract one channel: 1

      if enu_ch$ = "1"
        Save as WAV file: sub_fol$ + file$ + "__" + enu_ran$ + "__" + enu_spk$ +
"__ch1.wav"
      else
        Save as WAV file: sub_fol$ + "och\" + file$ + "__" + enu_ran$ + "__" +
enu_spk$ + "__ch1.wav"
      endif

      selectObject: "Sound " + file$
      Extract one channel: 2

      if enu_ch$ = "2"
        Save as WAV file: sub_fol$ + file$ + "__" + enu_ran$ + "__" + enu_spk$ +
"__ch2.wav"
      else
        Save as WAV file: sub_fol$ + "och\" + file$ + "__" + enu_ran$ + "__" +
enu_spk$ + "__ch2.wav"
      endif

    else
      Save as WAV file: sub_fol$ + file$ + " [" + enu_ran$ + "] " + enu_spk$ +
".wav"
    endif
```

```
    selectObject: "Sound " + file$
    To TextGrid: "Trans UTTex UTtag", ""
    Insert boundary: 1, enu_sta
    Insert boundary: 1, enu_end
    Insert boundary: 2, enu_sta
    Insert boundary: 3, enu_sta
    Set interval text: 1, 2, enu_tex$

    for ut to n_ut
      selectObject: enu_ut
      ut_tex$ = Get value: enu, string$(ut)
      selectObject: enu_tag
      ut_tag$ = Get value: enu, string$(ut)

      selectObject: "TextGrid " + file$
      Set interval text: 2, ut + 1, ut_tex$
      Set interval text: 3, ut + 1, ut_tag$
      Insert boundary: 2, enu_sta + ut*(enu_dur)/n_ut
      Insert boundary: 3, enu_sta + ut*(enu_dur)/n_ut

    endfor

    Shift times to: "start time", 0
    Save as text file: sub_fol$ + file$ + " [" + enu_ran$ + "] " + enu_spk$ +
".TextGrid"
    selectObject: "Sound " + file$
    plusObject: "TextGrid " + file$
    Remove

  endif
endfor

#select all
#Remove
```

```
# APÊNDICE D2.2: calcula a duração normalizada por sílaba e por unidade tonal
# resultados como TextGrid

select all
if numberOfSelected() <> 0
  Remove
endif

clearinfo

#local$ = "C:\Users\OG\Documents\Mestrado\Work\"
local$ = "C:\Users\LEEL\Documents\Oliver\Work\"

tab = Read from file: local$ + "BP_v1.TableOfReal"
r_tab = Get number of rows

lista = Create Strings as file list: "list_arqs", local$ +
"Data\TextGrid\*.TextGrid"
n_arq = Get number of strings

for a to n_arq

  selectObject: lista
  file$ = Get string: a
  txg = Read from file: local$ + "Data\TextGrid\" + file$
  Duplicate tier: 2, 6, "Zsil"
  Duplicate tier: 2, 7, "Msil"
  Duplicate tier: 2, 8, "Vsil"

  n_sil = Get number of intervals: 2

  for s from 1 to n_sil
    selectObject: txg

    sil$ = Get label of interval: 2, s
    ti_sil = Get start time of interval: 2, s
    tf_sil = Get end time of interval: 2, s
    dur_sil = (tf_sil - ti_sil)*1000

    if sil$ <> "pp" and sil$ <> "" and sil$ <> "xxx"

      tc_sil = length(sil$)
      sum_mean = 0
      sum_var = 0

      nc_sil = 1

      while nc_sil <= tc_sil
        tc_seg = 1
        terminate = 0
        seg$ = mid$(sil$, nc_sil,1)

        if nc_sil < tc_sil

          if mid$(sil$, nc_sil + 1, 1) = "h"  or mid$(sil$, nc_sil + 1, 1) = "N"
            tc_seg = tc_seg + 1
            seg$ = seg$ + mid$(sil$, nc_sil + 1, 1)
          endif
```

```
          if (nc_sil + tc_seg <= tc_sil)
            ic$ = mid$(sil$, nc_sil, 1)

            if (mid$(sil$, nc_sil + tc_seg, 1) = "I"  or  mid$(sil$, nc_sil +
tc_seg, 1) = "U")  and
            ...(ic$ = "i" or ic$ = "e"  or ic$ = "a"  or ic$ = "o"  or ic$ = "u"
or ic$ = "I" or
            ... ic$ = "E" or ic$ = "A"  or ic$ = "O"  or ic$ = "U")
              seg$ = seg$ + mid$(sil$, nc_sil + tc_seg, 1)
              tc_seg = tc_seg + 1
            endif

          endif

        endif

        selectObject: tab

        match = 0
        for j to r_tab
          label$ =  Get row label: j

          if seg$ = label$
            mean = Get value: j, 1
            sd = Get value: j, 2
            sum_mean = mean + sum_mean
            sum_var = sd*sd + sum_var
            match = 1
          endif

        endfor

        if match = 0
          print erro em 'file$' sílaba 's' 'sil$': 'tab$'segmento 'seg$' 'tab$'não
encontrado'newline$'
        endif

        nc_sil = nc_sil + tc_seg

      endwhile

      z = (dur_sil - sum_mean)/sqrt(sum_var)

      selectObject: txg

      Set interval text: 6, s, fixed$(z, 3)
      Set interval text: 7, s, fixed$(sum_mean, 0)
      Set interval text: 8, s, fixed$(sum_var, 0)

#      print 'file$' sílaba 's' 'sil$': 'tab$''dur_sil'ms,'tab$'z = 'z''newline$'

    endif

  endfor

  Save as text file: local$ + "Data\DurSil\" + file$

endfor
```

```
# APÊNDICE D2.3
# tabela por UT a segmentação silábica(1),  a duração (2) e a duração normalizada
(3)
# também regista as somas das médias (4) e das variâncias padrões de cada sílaba
(5) posição da sílaba tônica


select all
if numberOfSelected() <> 0
 Remove
endif

clearinfo

r = 1

for i to 5
  tab'i' = Create Table with column names: "UT_SIL" + string$(i), r, "No ID Audio
Enu Spk N_UT T_UT Tag Text N_Sil"
endfor

tab6 = Create Table with column names: "UT_SIL6", r, "No ID Audio Enu Spk N_UT
T_UT Tag Text N_Sil Bad Tsil Sti Stf Sdur Tvow Vti Vtf Vdur"

c = Get number of columns

local$ = "C:\Users\olive\OneDrive\Documentos\Work\"
#local$ = "C:\Users\LEEL\Documents\Oliver\Work\"

lista = Create Strings as file list: "list_arqs", local$ +
"Data\DurSil\*.TextGrid"
n_arq = Get number of strings

for a to n_arq
  selectObject: lista
  file$ = Get string: a
  txg = Read from file: local$ + "Data\DurSil\" + file$
  n_ut = Get number of intervals: 5
  n_pp = Count intervals where: 5, "is equal to", "pp"
  pp = 0

  for u from 1 to n_ut
    selectObject: txg
    ti = Get start time of interval: 5, u
    tf = Get end time of interval: 5, u
    tag$ = Get label of interval: 5, u
    txt$ = Get label of interval: 4, u
    if tag$ <> "" and tag$ <> "pp"
      u_t = Extract part: ti, tf, "no"
      n_sil = Get number of intervals: 2
      n_to = Get number of intervals: 3
      print 'file$' 'tag$'

      for s to n_sil
        selectObject: u_t
        val$[1] = Get label of interval: 2, s
        val$[3] = Get label of interval: 6, s
        val$[4] = Get label of interval: 7, s
        val$[5] = Get label of interval: 8, s
```

```
  ti_sil = Get start time of interval: 2, s
  tf_sil = Get end time of interval: 2, s
  if val$[1] <> ""
    val$[2] = string$((tf_sil - ti_sil)*1000)
  else
    val$[2] = "?"
  endif

 for i to 5
    selectObject: tab'i'
    cols = Get number of columns
    if cols < s + c
      Append column: "S" + string$(s)
    endif
    Set string value: r, "S" + string$(s), val$[i]
  endfor

endfor

selectObject: tab2
no_sil$ = Get value: r, "S1"

for t to n_to
  selectObject: u_t
  lab$ = Get label of interval: 3, t

  if index(lab$, "#")
    tti = Get start time of interval: 3, t
    ttf = Get end time of interval: 3, t
    nsil2 = Get interval at time: 2, tti/2 + ttf/2
    tsil$ = Get label of interval: 2, nsil2
    sti = Get start time of interval: 2, nsil2
    stf = Get end time of interval: 2, nsil2
    selectObject: tab6
    Set string value: r, "Tsil", tsil$
    Set numeric value: r, "Sti", sti/(tf - ti)
    Set numeric value: r, "Stf", stf/(tf - ti)
    Set numeric value: r, "Sdur", (stf - sti)/(tf - ti)
    Set string value: r, "Tvow", mid$(lab$, 2, length(lab$))
    Set numeric value: r, "Vti", tti/(tf - ti)
    Set numeric value: r, "Vtf", ttf/(tf - ti)
    Set numeric value: r, "Vdur", (ttf - tti)/(tf - ti)
  endif

  if index(lab$, "X")
    selectObject: tab6
    Set numeric value: r, "Bad", 1
  endif

endfor


selectObject: u_t
Remove
print 'newline$'

audio$ = mid$(file$, 1, 8)
enu$ = mid$(file$, 11, 3)
```

```
      spk$ = mid$(file$, 16, 3)
      n_ut$ = mid$(string$(100 + u - pp), 2, 2)
      t_ut$ = mid$(string$(100 + n_ut - n_pp - 2), 2, 2)

      for i to 6
        selectObject: tab'i'
        Set numeric value: r, "No", r
        Set string value: r, "ID", mid$(audio$, 5, 1) + mid$(audio$, 2, 1) +
mid$(audio$, 8, 1) + enu$ + n_ut$ + t_ut$
        Set string value: r, "Audio", audio$
        Set string value: r, "Enu", enu$
        Set string value: r, "Spk", spk$
        Set numeric value: r, "N_UT", u - pp
        Set numeric value: r, "T_UT", n_ut - n_pp - 2
        Set string value: r, "Tag", tag$
        Set string value: r, "Text", txt$

        if no_sil$ = "?"
          Set numeric value: r, "N_Sil", 0
        else
          Set numeric value: r, "N_Sil", n_sil
        endif

        Append row
      endfor

      r = r + 1

    else
      pp = pp + 1
    endif

  endfor

  selectObject: txg
  Remove

endfor
```

```
# APÊNDICED2.4
# medidas acústicas das unidades tonais
select all
if numberOfSelected() <> 0
 Remove
endif

clearinfo

n_row = 0

v$[1] = ""
v$[2] = "s"
v$[3] = "i"
va1$ = "No ID Audio Enu Spk N_UT T_UT Tag Text Bad dur In Insd Inmax Inmin Intmax
Intmin InmaxTa InminTa "
va2$ = "f0 f0sd f0max f0min f0tmax f0tmin f0exc f0slo f0maxTa f0minTa "
va3$ = "if0sit if0stf if0sif"
var$ = va1$ + va2$ + replace$(va2$, "f0", v$[2] + "f0", 0) + replace$(va2$, "f0",
v$[3] + "f0", 0) + va3$

tab = Create Table with column names: "UT_ACU", n_row, var$

 local$ = "C:\Users\olive\OneDrive\Documentos\Work\MC_R_UD\"
#local$ = "C:\Users\OG\Documents\Mestrado\Work\"
#local$ = "D:\LEL4017M04\Documents\Mestrado\Work\"
#local$ = "C:\Users\LEEL\Documents\Oliver\Work\"

lista = Create Strings as file list: "list_arqs", local$ + "Data\Audio\*.wav"
n_arq = Get number of strings
par = Read from file: local$ + "Pit_PAR.txt"
ton = Read from file: local$ + "UT_SIL6.txt"

for a to n_arq
  selectObject: lista
  file$ = Get string: a
  audio$ = mid$(file$, 1, 8)
  enu$ = mid$(file$, 11, 3)
  ide$ = mid$(audio$, 5, 1) + mid$(audio$, 2, 1) + mid$(audio$, 8, 1) + enu$
  spk$ = mid$(file$, 16, 3)

  selectObject: par
  p = Search column: "FileEnu", audio$ + enu$
  p_flo = Get value: p, "p_flo"
  p_cel = Get value: p, "p_cel"
  bad = Get value: p, "Bad"

  if p = 0
    print "Warning! Enunciado " 'file$' " não encontrado na tabela de parâmetros
de pitch"'newline$'
  endif

  som = Read from file: local$ + "Data\Audio\" + file$
  int = To Intensity: p_flo, 0, "yes"

  if fileReadable(local$ + "Data\Pitch\" + file$ - ".wav" + ".Pitch") = 0
    print "Warning! Não encontrado arquivo de pitch para o enunciado " 'file$' bad
= 'bad'"newline$'
    selectObject: som
```

```
    pit = To Pitch: 0, p_flo, p_cel
  else
    pit = Read from file: local$ + "Data\Pitch\" + file$ - ".wav" + ".Pitch"
  endif

  spit = Smooth: 4   ; pitch suaviado com bandwidth 4Hz
  selectObject: pit
  Interpolate
  ipit = Smooth: 4   ; pitch interpolado e suaviado

  txg = Read from file: local$ + "Data\TextGrid\" + mid$(file$, 1, 18) +
".TextGrid"
  n_ut = Get number of intervals: 5
  n_pp = Count intervals where: 5, "is equal to", "pp"
  pp = 0

  for u to n_ut
    selectObject: txg
    tag$ = Get label of interval: 5, u

    if tag$ <> "" and tag$ <> "pp"
      ti = Get start time of interval: 5, u
      tf = Get end time of interval: 5, u
      txt$ = Get label of interval: 4, u
      dur = (tf - ti)

      n_ut$ = mid$(string$(100 + u - pp), 2, 2)
      t_ut$ = mid$(string$(100 + n_ut - n_pp - 2), 2, 2)
      id$ = ide$ + n_ut$ + t_ut$
      selectObject: ton
      t = Search column: "ID", id$
      vti = Get value: t, "Vti"
      vtf = Get value: t, "Vtf"
      ttm = ti + dur*(vti + vtf)/2
      if0tt = 0
      if0ti = 0
      if0tf = 0
      tiv = ti
      tfv = tf
      selectObject: ipit

      if ttm <> undefined
        if0tt = Get value at time: ttm, "semitones re 1 Hz", "Linear"

        while (if0ti = undefined or if0ti = 0) and tiv < tf
          if0ti = Get value at time: tiv, "semitones re 1 Hz", "Linear"
          tiv = tiv + 0.005
        endwhile

        while (if0tf = undefined or if0tf = 0) and tfv > ti
          if0tf = Get value at time: tfv, "semitones re 1 Hz", "Linear"
          tfv = tfv - 0.005
        endwhile

      endif


      selectObject: int
      in = Get mean: ti, tf, "energy"
```

```
insd = Get standard deviation: ti, tf
inmax = Get maximum: ti, tf, "Parabolic"
inmin = Get minimum: ti, tf, "Parabolic"
intmax = Get time of maximum: ti, tf, "Parabolic"
intmin = Get time of minimum: ti, tf, "Parabolic"
intmax = (intmax - ti)/dur
intmin = (intmin - ti)/dur

selectObject: tab
Append row
n_row = n_row + 1
Set numeric value: n_row, "No", n_row
Set string value: n_row, "ID", id$
Set string value: n_row, "Audio", audio$
Set string value: n_row, "Enu", enu$
Set string value: n_row, "Spk", spk$
Set numeric value: n_row, "N_UT", u - pp
Set numeric value: n_row, "T_UT", n_ut - n_pp - 2
Set string value: n_row, "Tag", tag$
Set string value: n_row, "Text", txt$
Set numeric value: n_row, "Bad", bad
Set numeric value: n_row, "Dur", dur*1000
Set numeric value: n_row, "In", in
Set numeric value: n_row, "Insd", insd
Set numeric value: n_row, "Inmax", inmax
Set numeric value: n_row, "Inmin", inmin
Set numeric value: n_row, "Intmax", intmax
Set numeric value: n_row, "Intmin", intmin
Set numeric value: n_row, "InmaxTa", (intmax - vti)/(vtf - vti)
Set numeric value: n_row, "InminTa", (intmin - vti)/(vtf - vti)

Set numeric value: n_row, "if0sit", (if0tt - if0ti)/(ttm - ti)
Set numeric value: n_row, "if0stf", (if0tf - if0tt)/(tf - ttm)
Set numeric value: n_row, "if0sif", (if0tf - if0tt)/(tf - ti)

for i to 3
  ob$ = v$[i] + "pit"
v$ = v$[i]
  x = 0.25 ; %fração de tempo
  selectObject: 'ob$'
  'v$'f0 = Get mean: ti, tf, "semitones re 1 Hz"
  'v$'f0sd = Get standard deviation: ti, tf, "semitones"
  'v$'f0max = Get maximum: ti, tf, "semitones re 1 Hz", "Parabolic"
  'v$'f0min = Get minimum: ti, tf, "semitones re 1 Hz", "Parabolic"
  'v$'f0tmax = Get time of maximum: ti, tf, "semitones re 1 Hz", "Parabolic"
  'v$'f0tmin = Get time of minimum: ti, tf, "semitones re 1 Hz", "Parabolic"
  'v$'f0exc = 'v$'f0max - 'v$'f0min
  'v$'f0slo = 'v$'f0exc/('v$'f0tmax - 'v$'f0tmin)
  'v$'f0tmax = ('v$'f0tmax - ti)/dur
  'v$'f0tmin = ('v$'f0tmin - ti)/dur
  'v$'f0maxTa = ('v$'f0tmax - vti)/(vtf - vti)
  'v$'f0minTa = ('v$'f0tmin - vti)/(vtf - vti)

  selectObject: tab
Set numeric value: n_row, v$ + "f0", 'v$'f0
Set numeric value: n_row, v$ + "f0sd", 'v$'f0sd
Set numeric value: n_row, v$ + "f0max", 'v$'f0max
Set numeric value: n_row, v$ + "f0min", 'v$'f0min
Set numeric value: n_row, v$ + "f0tmax", 'v$'f0tmax
```

```
       Set numeric value: n_row, v$ + "f0tmin", 'v$'f0tmin
       Set numeric value: n_row, v$ + "f0exc", 'v$'f0exc
       Set numeric value: n_row, v$ + "f0slo", 'v$'f0slo
       Set numeric value: n_row, v$ + "f0maxTa", 'v$'f0maxTa
       Set numeric value: n_row, v$ + "f0minTa", 'v$'f0minTa
      endfor

      appendInfoLine: a, tab$, audio$, tab$, enu$, tab$, tag$, tab$, fixed$(ti,
3), tab$, fixed$(tiv, 3), tab$, fixed$(ttm, 3), tab$, fixed$(tfv, 3), tab$,
fixed$(tf, 3),
                  ... tab$, fixed$(if0ti, 1), tab$, fixed$(if0tt, 1), tab$,
fixed$(if0tf, 1), tab$, (tiv > ttm) + (tfv < ttm)
    else
      pp = pp + 1
    endif

  endfor

  select all
  minusObject: lista, tab, par, ton, ipit
  Remove

endfor

selectObject: tab
n_col = Get number of columns

for c from 10 to n_col
  c_lab$ = Get column label: c

  for r to n_row
    val = Get value: r, c_lab$

    if val = undefined
      Set string value: r, c_lab$, "NA"
    endif

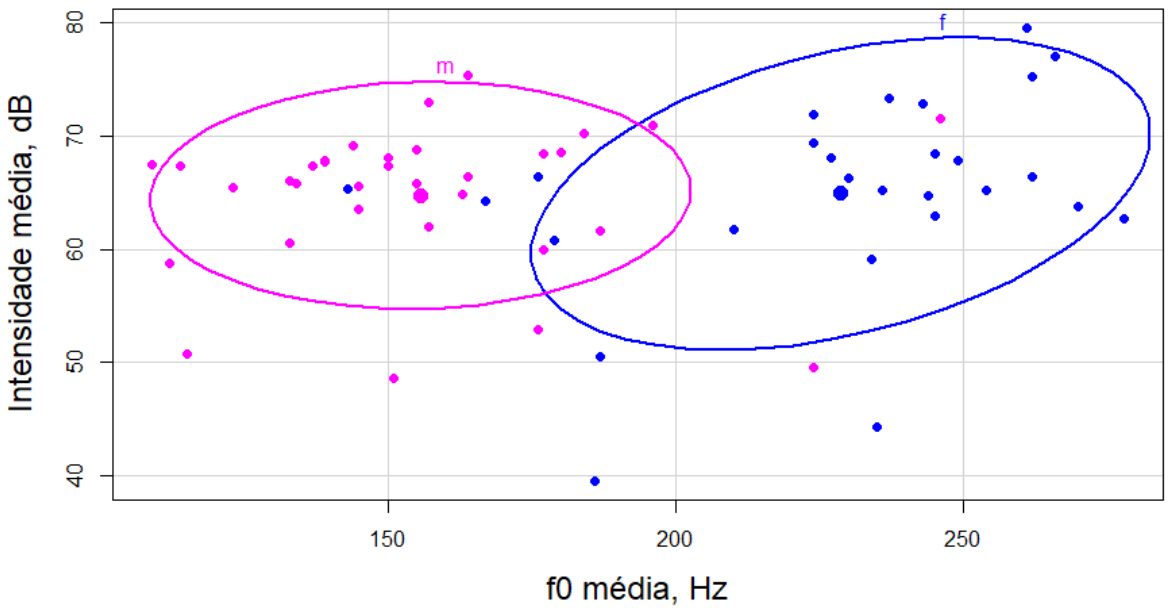
  endfor
endfor

Save as tab-separated file: local$ + "UT_ACU.txt"
```

APÊNDICE E: Valores de f0 e intensidade por falante do Mini BR Informal

Audio	spk	Gen	média Hz	sd Hz	1o_qt Hz	med Hz	3o_qt Hz	dur s	voiced	int dB	int1 dB	int2 dB	Ch
bfamcv01	EVN	m	150	31	126	145	169	63,0	61%	67,4	70,1	58,8	1
bfamcv01	GIL	m	184	51	141	175	218	141,1	58%	70,2	63,0	72,7	2
bfamcv01	LEO	m	164	34	136	158	186	66,0	48%	66,4	68,8	60,6	1
bfamcv01	LUI	m	177	40	145	176	205	49,6	60%	68,5	62,0	71,0	2
bfamcv02	JAE	f	245	41	215	249	271	25,6	41%	68,4			M
bfamcv02	RUT	f	244	46	211	239	271	131,1	51%	64,7			M
bfamcv02	TER	f	230	45	197	225	258	190,3	57%	66,3			M
bfamcv03	CAR	m	155	40	123	147	179	88,6	40%	65,8	68,5	57,4	1
bfamcv03	CEL	m	246	53	205	239	280	66,1	56%	71,6	65,3	74,1	2
bfamcv03	ONO	m	205	34	177	204	232	3,2	37%	49,0	50,1	47,4	I
bfamcv03	REN	m	157	42	126	148	179	35,6	38%	62,0	64,5	55,8	I
bfamcv03	TON	m	180	45	143	169	211	71,3	51%	68,6	65,9	70,2	2
bfamcv04	BRU	f	278	53	242	268	305	166,5	33%	62,7			M
bfamcv04	CEL	m	163	39	131	156	190	17,1	56%	64,9			M
bfamcv04	HEL	f	262	59	217	250	293	72,2	50%	66,4			M
bfamcv04	LUC	m	150	42	117	141	176	26,5	48%	68,1			M
bfamdl01	CLA	f	86	39	41	116	117	0,2	63%	56,6			M
bfamdl01	CLB	f	296	53	255	293	326	3,0	65%	72,7			M
bfamdl01	FLA	f	266	64	220	252	299	288,6	58%	77,0			M
bfamdl01	MDS	f	224	120	118	159	320	20,0	69%	71,9			M
bfamdl01	REN	f	261	56	223	250	288	322,4	56%	79,6			M
bfamdl02	BAL	m	196	43	164	189	222	195,7	68%	71,0	57,7	73,9	2
bfamdl02	BEL	f	234	43	203	229	257	152,0	61%	59,1	59,7	58,3	I
bfamdl03	LAU	m	123	38	89	124	149	142,0	56%	65,4			M
bfamdl03	LUZ	f	176	89	90	174	246	308,4	52%	66,4			M
bfamdl03	POR	m	150	69	64	171	185	6,5	23%	51,4			I
bfamdl04	ERN	f	216	36	192	210	229	18,5	15%	46,5	48,0	44,3	I
bfamdl04	GUI	m	224	41	200	222	244	15,2	27%	49,5	51,5	45,6	I
bfamdl04	HRM	f	110	0	110	110	110	6,1	0,3%	54,2	55,7	51,7	I
bfamdl04	KAT	f	249	60	207	232	271	91,5	56%	67,8	52,1	70,8	2
bfamdl04	SIL	f	245	58	203	230	277	341,2	45%	62,9	65,3	57,2	I
bfamdl04	TOM	m	151	32	125	148	175	12,8	29%	48,6	50,4	45,5	I
bfamdl05	ANE	f	143	91	67	84	223	166,4	49%	65,3	65,5	65,1	I
bfamdl05	CES	m	114	45	71	110	147	285,0	52%	67,4	62,4	69,7	2
bfamdl05	ENC	m	115	43	84	95	155	12,4	44%	50,7	45,2	53,0	I
bfamdl05	PED	m	61	5	58	61	64	8,5	92%	51,7	51,0	52,3	I
bfamdl05	RAQ	f	95	16	84	96	104	1,8	50%	52,9	46,8	55,3	I
bfammn01	DUD	m	144	24	125	139	157	10,3	68%	69,1			M
bfammn01	MAI	m	139	31	116	137	158	281,9	45%	67,7			M
bfammn02	DFL	f	179	62	123	179	217	375,4	63%	60,8	52,4	63,4	I
bfammn02	LUC	m	133	30	111	123	147	28,5	63%	60,6	61,0	60,2	I
bfammn03	ALO	m	134	31	112	129	150	350,4	62%	65,8			M
bfammn03	ANA	f	186	28	169	180	197	10,0	65%	39,5			M
bfammn03	JOS	m	112	27	99	106	122	2,2	55%	58,8			M
bfammn03	JUL	m	187	53	156	170	219	3,6	74%	61,6			M
bfammn03	LUA	f	235	38	202	233	248	9,1	22%	44,3			M
bfammn03	MAR	f	356	51	314	330	411	0,6	45%	34,8			M
bfammn04	BAL	m	176	23	163	173	181	2,4	86%	52,9			M
bfammn04	REG	f	270	61	226	256	300	400,5	79%	63,8			M
bfammn05	CAR	f	236	45	204	226	259	557,7	68%	65,2			M
bfammn05	JUN	m	177	38	149	168	201	9,2	81%	60,0			M
bfammn05	MAR	f	187	54	125	211	228	5,2	32%	50,5			M
bfammn06	JOR	m	145	28	125	140	163	569,6	63%	63,6	55,5	66,3	2
bfammn06	TAT	f	227	37	199	219	249	13,1	56%	68,1	71,1	46,5	1

bpubcv01	BRU	m	139	24	123	132	153	34,4	61%	67,9	61,2	70,4	2
bpubcv01	EMM	m	137	25	118	132	152	75,6	56%	67,4	69,0	64,9	1
bpubcv01	FLA	f	254	44	222	249	278	164,8	46%	65,2	61,0	67,3	2
bpubcv01	FUN	f	374	34	353	379	407	0,3	92%	66,6	63,2	68,5	2
bpubcv01	MAR	f	167	60	111	170	215	96,5	43%	64,3	62,2	65,7	I
bpubcv02	CAR	m	133	35	106	122	157	85,2	61%	66,1	61,4	68,3	2
bpubcv02	JEA	f	210	75	119	227	264	9,0	48%	61,7	63,0	59,8	I
bpubcv02	OSV	m	157	33	130	153	180	123,2	70%	73,0	63,4	75,8	2
bpubcv02	TIQ	m	155	23	138	148	170	13,8	65%	68,8	58,1	71,6	2
bpubdl01	PAU	m	109	19	95	106	120	384,4	48%	67,5			M
bpubdl01	ROG	m	145	35	118	139	167	243,2	39%	65,6			M
bpubdl02	EUG	m	164	42	132	157	191	242,6	59%	75,4			M
bpubdl02	JAN	f	262	56	220	253	300	86,2	51%	75,3			M
bpubdl02	WOA	f	237	67	187	233	281	38,1	33%	73,4			M
bpubmn01	LUA	f	224	31	203	218	238	72,7	73%	69,4			M
bpubmn01	SHE	f	243	49	206	234	274	471,1	73%	72,9			M
média			188	45	153	181	217	139		65			
media f		f	229	57	183	220	264	170		65			
média m		m	190	44	157	182	218	139		79			



Obs.: Os dados em cinza não foram utilizados no cálculo das médias, por terem duração muito pequena ou percentual de vozeamento muito abaixo da média. Também não foram incluídos na figura.

As elipses delimitam região com 68,27% (1 desvio-padrão) de probabilidade normal em torno da média.