

unesp  UNIVERSIDADE ESTADUAL PAULISTA

“JÚLIO DE MESQUITA FILHO”

**Faculdade de Ciências e Letras
Campus de Araraquara - SP**

MARIANE CARVALHO VISCHI

A DISFLUÊNCIA COMUM E GAGA



ARARAQUARA – S.P.

2017

MARIANE CARVALHO VISCHI

A DISFLUÊNCIA COMUM E GAGA

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – Unesp/Araraquara, como requisito para obtenção do título de Doutor em Linguística e Língua Portuguesa.

Exemplar apresentado para exame de Qualificação.

Linha de pesquisa: Análise Fonológica, Morfosintática, Semântica e Pragmática.

Orientador: Luiz Carlos Cagliari

Bolsa: CNPq

ARARAQUARA – S.P.

2017

VISCHI, MARIANE
A DISFLUÊNCIA COMUM E GAGA / MARIANE VISCHI - 2017
329 f.

Tese (Doutorado em Linguística e Língua
Portuguesa) - Universidade Estadual Paulista "Júlio
de Mesquita Filho", Faculdade de Ciências e Letras
(Campus Araraquara)

Orientador: LUIZ CARLOS CAGLIARI

1. DISFLUÊNCIA COMUM. 2. DISFLUÊNCIA GAGA. 3.
ARTICULAÇÃO. 4. EMPATIA NEGATIVA. I. Título.

Ficha catalográfica elaborada pelo sistema automatizado
com os dados fornecidos pelo(a) autor(a).

MARIANE CARVALHO VISCHI

A DISFLUÊNCIA COMUM E GAGA

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras – UNESP/Araraquara, como requisito para obtenção do título de Doutor em Linguística e Língua Portuguesa.

Linha de pesquisa: Análise Fonológica, Morfossintática, Semântica e Pragmática.

Orientador: Luiz Carlos Cagliari

Bolsa: CNPq

Data de defesa: ___/___/___

MEMBROS COMPONENTES DA BANCA EXAMINADORA:

Presidente e Orientador: Professor Dr. Luiz Carlos Cagliari

Universidade Estadual Paulista “Júlio de Mesquita Filho”

Membro Titular: **Nome e título**
Universidade.

Membro Titular: **Nome e título**
Universidade.

Membro Titular: **Nome e título**
Universidade.

Membro Titular: **Nome e título**
Vínculo
Universidade.

Local: Universidade Estadual Paulista
Faculdade de Ciências e Letras
UNESP – Campus de Araraquara

Dedico esta tese à Maria Laura, ao Bruno e aos meus pais.

Agradecimentos

“A minha alma engrandece o Senhor, e meu espírito se alegra em Deus, meu salvador”. Lc, 1.47.

A Deus, por guiar os meus caminhos e me dar forças para vencer mais esta etapa em minha vida.

Aos meus pais, toda gratidão e amor pelo incentivo, dedicação e apoio em todos os momentos.

Ao Bruno, pela companhia e cumplicidade de 12 anos, pelo amor e reconhecimento de todo o meu esforço.

Aos meus sogros, pela disponibilidade em me ajudar em todos os momentos.

Ao meu orientador, pela confiança e dedicação durante esses anos, por iluminar meu caminho durante a elaboração desta tese, e por compartilhar comigo ideias grandiosas. Ao senhor, deixo registrada toda a minha admiração.

A todos meus amigos e familiares, pelo incentivo e compreensão em minhas ausências.

À Cintia, minha cunhada mais que querida, por me ouvir nos momentos de desespero e alegria, e, sobretudo, por nunca me julgar.

À minha irmã, pelo espírito alegre, afetuoso e cheio de calma.

Às minhas queridas amigas Raisa e Paloma, pela paciência, carinho e atenção.

Aos meus queridos informantes, que prontamente se disponibilizaram para participar deste estudo.

Ao CNPq, pelo apoio financeiro.

A todos, que fazem parte da minha vida, e não foram mencionados, mas que estão em meu coração.

“Aquilo que eu escuto eu esqueço
Aquilo que eu vejo eu lembro,
Aquilo que eu faço eu aprendo”
Confúcio

RESUMO

O presente trabalho investiga e compara a disfluência na fala de pessoas gagas e não gagas. O primeiro tipo foi denominado de disfluência gaga e o segundo, de disfluência comum. De uma maneira geral, as disfluências são caracterizadas pelo não fluir normal dos segmentos sonoros da fala, devido à presença de hesitações e de interrupções durante o processo comunicativo. Inicialmente, foi feita uma revisão bibliográfica sobre a relação da fluência com a disfluência a partir de uma compilação dos principais estudos sobre o assunto. Na sequência, apresentamos os trabalhos da literatura especializada e da linguística para a disfluência gaga, bem como dos aspectos que tratam da articulação da fala e dos elementos suprasegmentais. A metodologia consistiu na gravação semiespontânea de seis informantes gagos e seis não gagos. Como parte da metodologia para a formação do *corpus*, foi pedido a eles que narrassem algum fato que tivesse marcado suas vidas. Não foi definida de uma temática específica. Os dados foram analisados auditivamente e acusticamente. Para a análise acústica utilizou-se o *software* PRAAT. Para análise e interpretação dos dados foram usados modelos teóricos referentes à fonologia autosegmental (Pierrehumbert, 1980; Ladd, 1996, entre outros) e, como apoio teórico, o da fonologia funcional (Halliday, 1970; Cagliari, 2007). Também foram utilizados os trabalhos relacionados ao estudo da fluência e disfluência da fala de uma maneira geral (Scarpa, 1995; Cruz, 2009; Koch, 2009; Scarpa; Fernandes-Svartman, 2012; Merlo, 2015, entre outros), bem como da disfluência gaga (Van Riper, 1971; Friedman 1986; Azevedo, 2000, entre outros). A investigação acústica estuda os segmentos fonéticos, as pausas, as repetições, as hesitações, os alongamentos e os bloqueios de fala. A interpretação fonológica investiga o comportamento dos fonemas nos dois tipos de disfluências. Os resultados mostraram, em relação à análise da F0 e padrão da intensidade, diferenças pouco significativas para a comparação entre os dois tipos de disfluências. A duração foi o evento que mais distanciou os dois tipos de fala, sendo maior para os falantes com disfluência gaga. Em relação aos tipos de disfluências (repetições, inserções, hesitações, entre outros), tanto os informantes gagos quanto os não gagos realizaram o mesmo tipo de disfluência. A diferença entre eles deve-se à quantidade e à complexidade das disfluências. Nesse caso, os informantes gagos tendem a apresentar um maior número de disfluências por grupo tonal, bem como uma quantidade maior de disfluências complexas. Outra diferença entre os dois grupos está na duração e quantidade de pausas, que tende a ser maior para os informantes com disfluência gaga. A análise fonológica, por sua vez, mostrou que gagos e não gagos tendem a manter o padrão entoacional para as repetições, inserções e hesitações. Esses resultados confirmam a hipótese de que a gagueira estaria mais relacionada a uma questão articulatória que leva, por meio de uma empatia fonética negativa, a quebra na expectativa do interlocutor. Já a disfluência comum, por estar mais relacionada a fenômenos estilísticos, não é sentida da mesma maneira. Mesmo assim, o caminho mais adequado não é o de distanciar negativamente esses dois tipos de fala, uma vez que os eventos são os mesmos, o que muda é o modo como eles são refletidos e sentidos.

Palavras-chave: Fluência. Disfluência comum. Disfluência gaga. Articulação. Empatia negativa.

ABSTRACT

The present work investigates and compares the dysfluency speech of stuttering and non-stuttering people. The first type was called stutterer dysfluency and the second common disfluency. In general, dysfluencies are characterized by no expected flow of the phonetic segments of speech due to the presence of hesitations and interruptions during the communication process. A general review of the literature on the relationship between fluency and dysfluency is presented with comments. A special review presents works of linguists, discussing stutterer dysfluency in particular in relation to segmental and suprasegmental articulatory problems. The methodology consists of recordings of semi spontaneous speech by six stutterers and non-stutterers informants. As part of the methodology for the formation of the corpus, the informants were asked to report any fact that marked their lives. Data were analyzed auditorily and acoustically. For acoustic analysis, the PRAAT software was used. Theoretical models regarding autosegmental phonology (Pierrehumbert, 1980; Ladd, 1996 among others) and functional phonology (Halliday, 1970, Cagliari, 2007) were used for theoretical support. The study of fluency and dysfluency discusses in particular the works from Scarpa (1995); Cruz (2009); Koch (2009) and Merlo (2015). On the other hand, stutterer dysfluency reminds the works of Van Riper (1971); Friedman (1986); Azevedo (2000), among others. The acoustic analyses take into consideration the phonetic segments, pauses, repetitions, hesitations, speech stretching and blocking. The phonological interpretation investigates the phoneme behavior in both types of disfluency. The results revealed, in relation to the analysis of F0 and intensity standard, minor differences in comparing the two types of disfluency. The duration was the event that more distanced the two types of speech behavior, being higher for speakers with stutterer dysfluency. The difference between them is due to the amount of occurrences and the types of complex disfluencies. Stuttering informants tend to have a larger number of dysfluency by tonal group as well as a greater amount of complex dysfluency. The phonological analyses, in turn, showed that stuttering and non-stuttering tend to maintain the intonation pattern for repeats, insertions and hesitations. These results confirm the hypothesis that stuttering is more related to an articulation process which leads, through a negative phonetic empathy, to break the expectation of the interlocutor. On the other hand, common dysfluency which is more related to stylistic phenomena is also involved with the same kind of speech problems but in a different dimension. Indeed, the most appropriate conclusion is to set a scale for the events with the two types of disfluency in opposite positions. This scale is interpreted as such by the persons involved in the problem, being framed by different expectations.

Keywords: Fluency. Common dysfluency. Stutterer dysfluency. Articulation. Negative empathy.

LISTA DE FIGURAS

Figura 1	Imagem do comportamento da laringe durante a produção de uma palavra gaguejada (A) e outra (B), fluente.	76
Figura 2	Exemplificação do comportamento da laringe durante a produção da vogal alongada [i] na palavra <i>tea</i> em (A) e da sua produção fluente (B).	76
Figura 3	Levelt e o modelo de Auto-monitoramento da fala.	86
Figura 4	Representação simplificada do modelo autosegmental de Pierrehumbert.	96
Figura 5	Distribuição do <i>pitch accent</i> , acento frasal e tom de fronteira no enunciado “ <i>Anna</i> ”	97
Figura 6	Representação do bloqueio -v- da palavra <i>vasupo</i> no enunciado “ <i>vasupo a -va vasupo a pi -v- vasupo ela vasupo aí</i> ”, produzido pelo informante gago AP.	111
Figura 7	Exemplo de segmento com redução de forma de onda em BLO sem PS para o informante gago PA.	112
Figura 8	Exemplo de segmento com interrupção de forma de onda em BLO sem PS para o informante não gago CR.	113
Figura 9	Representação da possibilidade das três ocorrências de bloqueio observadas para as disfluências comuns.	114
Figura 10	Representação da PS de bloqueio silenciosa para os informantes não gagos.	114
Figura 11	As quatro possibilidades de pausas de BLO para as disfluências gagas.	115
Figura 12	Primeira etapa: seleção do som desejado para análise na tela do PRAAT <i>objects</i> .	119
Figura 13	Segunda etapa: processo para a seleção do silêncio análise na tela do PRAAT <i>objects</i> .	120
Figura 14	Terceira etapa: janela do <i>software</i> para a notação dos valores referentes ao silêncio.	121
Figura 15	Tela do PRAAT referente aos valores utilizados para anotação do silêncio em <i>minimum silent interval duration</i> (s) e <i>minimum sounding interval duration</i> .	122
Figura 16	Tela com a identificação dos <i>tiers</i> referentes à segmentação das palavras, dos silêncios (<i>silent</i>), dos períodos de sonorização (<i>sounding</i>) e enunciado completo, na frase “Então, assim pra mim nor fico nor normal, mais a gentis abi qui num é, né”, produzido pelo informante gago CAM.	123
Figura 17	Procedimento para análise da intensidade máxima no enunciado “meu pai escutava im caza num, naquelas toca é toca discu antigu” produzido pelo informante não gago BR.	125
Figura 18	Procedimento para análise da intensidade no momento de transição, início ou fim de segmento no enunciado “Aí, assim tem coiza qui eu não não consigu fala” produzido pelo informante gago CAM.	126
Figura 19	Identificação dos pulsos por meio das linhas azuis verticais no enunciado “i de de depressão juntu”, produzido pelo informante gago CAM.	128

Figura 20	Exemplificação do padrão entoacional da repetição para o informante CR, no enunciado resumido EN6 “eli crio dentro da cadeia”.	134
Figura 21	Padrão entoacional da repetição no enunciado EN19 “graças a Deus foi/ foi foi feliz” produzido por CR.	135
Figura 22	Representação detalhada do espectrograma das repetições no enunciado “graças a Deus foi/ foi foi feliz”, produzido por CR.	136
Figura 23	Exemplo de HES no enunciado EN4 “ah como bom comu se falo assisti várius filmi”.	137
Figura 24	Exemplo de alongamento em EN24 “/mais u Rodrigu tá senu um:/ nu meu modu di pensa, di analisa”, produzido por CR.	138
Figura 25	Exemplo de AL _p em fim e início de palavra em EN11 “vivu: :minha vida vivi tenu situação engraçada”, produzida por CR	139
Figura 26	Representação da possibilidade dos três tipos de bloqueio observados para as disfluências comuns.	141
Figura 27	Representação das hesitações nos enunciados EN4, EN20 e EN28 realizado por CR.	147
Figura 28	Caracterização dos pulsos, intensidade e frequência fundamental na disfluência do tipo bloqueio, presentes em EN1, EN4, EN26 e EN31, respectivamente.	150
Figura 29	Representação da interrupção da forma de onda para os enunciados EN1 (60a) e EN31 (60b).	151
Figura 30	Identificação da frequência fundamental, intensidade, forma de onda e espectrograma no enunciado “graças a Deus foi/foi foi feliz”.	157
Figura 31	Exemplificação dos parâmetros acústicos e do padrão entoacional para MA no enunciado “trinta i dois anu, trinta i patru anu”.	164
Figura 32	Representação da duração, intensidade, F0 e pulsos para a RP _{pi} no enunciado “# queu comecei desdi d udus dozi, né”, realizado pelo informante não gago MA.	169
Figura 33	Exemplificação de uma forma de onda crescente para uma repetição do tipo 4RP _p realizada pelo informante não gago LE no enunciado “ah é quando se nu nu nu nu um nu compreendi u otu né#”.	176
Figura 34	Exemplo de repetição de sílaba dupla no enunciado / aí consiguiu lu lu luga pa dona da farmácia/ realizado pelo informante não gago LE.	177
Figura 35	Exemplo de diminuição da duração para uma RP tripla no enunciado “iscutanu# pra pra pra hora qui tocassi Rolling Stonis eu grava” produzido pelo informante não gago BR.	185
Figura 36	Representação da curva entoacional para a INS _p /sabi/ em final de frase no enunciado “/qué a primera conversa nossa daí eu to meu timidu sabi”.	206
Figura 37	Representação da curva entoacional para a INSp /sabi/ em final de frase, para o enunciado EN5 “/ah queu ficu # é conversanu muito # é na sala sabi#”	206
Figura 38	Possibilidades de pausas de BLO para o informante gago PA.	209

Figura 39	Comparação da forma de onda e do espectrograma em um segmento bloqueado (imagem esquerda) /-eu/ e em um sem bloqueio em momento fluente do informante gago PA.	210
Figura 40	Análise do bloqueio em repetição de estrutura para PA no enunciado “daí a genti ia cantanu di lá até -nu até nu campu”.	212
Figura 41	Observação das durações para as 3RP _p s da palavra /si/ no enunciado EN24 “si si si não num é festa du vinhu”.	215
Figura 42	Comparação da F0 ao longo das cinco repetições produzidas por PA no enunciado “#da da da da #da eu vortei essi anu”	218
Figura 43	Representação de um trecho do enunciado para a visualização da repetição quántupla de palavra (5RP _p) realizada pelo informante gago JO.	233
Figura 44	Representação de um erro de gravação no enunciado EN35 “é qui todú mu sabi queu so, todú mundu fica falanu cumigu, vo isconde porque?”, realizado por JO.	239
Figura 45	Comparação entre os momentos de em meio de palavra para o informante gago AP.	248
Figura 46	Representação do segmento [t] sem BLO no enunciado EN26 “in-tão eu fiz cursu militar”.	249
Figura 47	Espectro da forma de onda da consoante bloqueada [t] da palavra com disfluência /in-tão/.	250
Figura 48	Espectro da forma de onda da consoante sem bloqueio [t] da palavra sem disfluência /militar/.	250
Figura 49	Representação de uma PS de bloqueio com presença de articulação, produzida pelo informante AP no enunciado EN32 “mais eu achu qui eu achu qui -eu achu eu achu qui cada coisa du #i vasupo eu separu cada coisa nu seu lugar”	251
Figura 50	Representação do AL _p inicial realizado pelo informante AP no enunciado “/apartei :muitas briga, né”.	258
Figura 51	Repetição do tipo 4RP _v realizada por MAU no enunciado EN26 “#aí você cumeça a a a a pensa”.	265

LISTA DE GRÁFICOS

Gráfico 1	Relação das disfluências simples para o informante não gago CR.	131
Gráfico 2	Os subtipos das disfluências identificados para o informante não gago CR.	132
Gráfico 3	Especificação das pausas para o informante não gago CR.	145
Gráfico 4	Análise das intensidades inicial, final, mínima e máxima para a hesitação para o informante não gago CR.	148
Gráfico 5	Os tipos de disfluências realizadas pelo informante não gago MA.	161
Gráfico 6	Distribuição dos eventos disfluentes realizados pelo informante não gago LE.	173
Gráfico 7	Distribuição dos eventos disfluentes realizados pelo informante não gago BR.	182
Gráfico 8	Distribuição dos eventos disfluentes realizados pelo informante não gago NE.	190
Gráfico 9	Distribuição dos eventos disfluentes realizados pelo informante não gago CL.	196
Gráfico 10	Os tipos de disfluência realizados pelo informante gago PA.	204
Gráfico 11	Distribuição dos eventos disfluentes realizados pelo informante gago VA.	221
Gráfico 12	Distribuição dos eventos disfluentes realizados pelo informante gago JO	231
Gráfico 13	As variações de disfluência gaga na gravação de AP.	241
Gráfico 14	Distribuição dos eventos disfluentes realizados pelo informante gago MAU.	262
Gráfico 15	Distribuição dos eventos disfluentes realizados pelo informante gago CAM.	273

LISTA DE QUADROS

Quadro 1	Os tipos de sons preenchedores mais comuns em algumas línguas	42
Quadro 2	Os estágios de desenvolvimento da disfluência gaga, adaptado de Valerie Laporte (1997).	62
Quadro 3	Esquema para a análise das categorias temporais.	127
Quadro 4	Os tipos de INs para os informantes com disfluência gaga e comum.	289
Quadro 5	Descrição das INs para os informantes com disfluência gaga.	291
Quadro 6	Descrição das INs para os informantes com disfluência comum.	292
Quadro 7	Descrição geral dos fenômenos observados para a disfluência gaga e comum.	312

LISTA DE TABELAS

Tabela 1	Distribuição do número de Grupos tonais por informante.	107
Tabela 2	Comparação entre o número de GTs dos informantes gagos e não gagos.	283
Tabela 3	Número de ocorrência das disfluências para os informantes gagos e não gagos.	284
Tabela 4	Distribuição do número de disfluências e pausas por enunciado para os informantes gagos e não gagos.	284
Tabela 5	Comparação entre os tipos de disfluências complexas para os informantes gago JO e não gago MA.	286
Tabela 6	Distribuição das pausas de acordo com o seu contexto no GT.	287
Tabela 7	Duração das PSs de acordo com a sua posição no enunciado.	288
Tabela 8	Ocorrência das PSs silenciosas e respiratórias.	288
Tabela 9	Duração das PSs respiratórias e silenciosas	289
Tabela 10	Duração total das INSs para os informantes gagos e não gagos.	293
Tabela 11	Descrição da quantidade e contexto das CORs de acordo com o grupo de informantes	297
Tabela 12	Tendência da duração das CORs simples e complexas conforme o grupo de informantes.	298
Tabela 13	Descrição da porcentagem de ocorrência dos BLOs em consoante para os informantes com disfluência comum e gaga.	300
Tabela 14	Descrição da porcentagem de ocorrência dos BLOs em vogais para os informantes com disfluência comum e gaga	301
Tabela 15	Porcentagem de ocorrência dos tipos de RPs.	303
Tabela 16	Duração total das repetições que aumentaram para os informantes com disfluência gaga e comum.	303
Tabela 17	Duração total das repetições que diminuíram para os informantes com disfluência gaga e comum.	304
Tabela 18	Descrição das F0 mínima e máxima e média aritmética para os informantes com disfluência comum e gaga.	306
Tabela 19	Média dos aspectos prosódicos temporais dos informantes com disfluência gaga e comum.	307

LISTA DE ABREVIATURAS

AM -	Teoria métrica autosegmental da entoação.
AP -	Informante gago.
BR -	Informante não gago.
CA -	Informante não gago.
CAM -	Informante gago.
CL -	Informante não gago.
CMN -	Contorno melódico notável.
EVA -	Método multiparamétrico de avaliação vocal objetiva assistida.
F0 -	Frequência fundamental.
GT -	Grupo Tonal.
JO -	Informante gago.
LE -	Informante não gago.
MA -	Informante não gago.
MAU -	Informante gago.
NE -	Informante não gago.
NURC -	Norma Urbana Culta.
PA -	Informante gago.
PB -	Português do Brasil.
PCO -	Princípio do contorno obrigatório.
RP -	Received Pronunciation.
STS -	Sílaba tônica saliente.
TCLE -	Termo de Consentimento Livre e Esclarecido.
ToBI -	<i>Tones and Break Indices.</i>
VA -	Informante gago.
VOT -	<i>Voice onset time</i>

SUMÁRIO

INTRODUÇÃO.....	18
1 FUNDAMENTAÇÃO TEÓRICA.....	26
1.1 Sobre A (Dis)Fluência	26
1.3 As hesitações (ou disfluências)	36
1.4 Os tipos de disfluências	38
1.5 GAGUEIRA: uma visão geral	48
1.5.1 Os primeiros relatos.....	48
1.5.2 O ponto de vista da literatura patológica e fonoaudiológica.....	50
1.5.3 A gagueira sob a ótica dos estudos linguísticos	54
1.5.4 Os estudos sobre a gagueira a partir da perspectiva linguística no Brasil	58
1.6 Caracterização do fenômeno.....	61
1.7 Sobre a articulação	68
1.8 O modelo de levelt (1989)	85
2 QUADRO TEÓRICO	90
2.1 O modelo autossegmental métrico	90
2.1.1 Características do modelo teórico	92
2.2 O modelo de Halliday (1970) e Cagliari (2007)	98
3 METODOLOGIA	103
3.1 Informantes	103
3.1.1 Os informantes gogos.....	104
3.1.2 Os informantes não gogos	105
3.2 MATERIAIS E PROCEDIMENTOS	106
3.2.1 A coleta de dados.....	106
3.2.2 Dados da análise acústica.....	107
3.3 Metodologia para análise dos fenômenos prosódicos e acústicos.....	118
4 RESULTADOS E DISCUSSÃO.....	131
4.1 Informantes com disfluência comum	131
4.2 Informantes com disfluência gaga.....	203
4.3 Comparação geral dos resultados.....	282
4.3.1 GTs, PSs, tipos de disfluências, tessitura e aspectos temporais.....	282
CONCLUSÃO.....	309
REFERÊNCIAS	317

INTRODUÇÃO

Hesitations thus form an integral part of speech production in the positive sense, a view quite in opposition to the attitude that there is fluency on one hand and hesitancy on the other.

Heike (1981, p. 150)

As pessoas são naturalmente diferentes, seja na aparência física, na personalidade, nos desejos, nos gostos, nas crenças, entre outras características. É isso que nos torna singulares dentro de um todo. Com a fala, essa questão também não poderia ser de outra forma. O tema da Variação Linguística já deixou de ser novidade nos livros didáticos, bem como nas disciplinas dos cursos de Letras. Mas, para chegarmos a esse ponto, foram (e ainda são) necessárias muitas discussões por parte da comunidade científica. Infelizmente, junto com o diferente surge o preconceito, que pode ser social, racial, étnico, linguístico, enfim, há vários tipos. Mas, o que o preconceito tem em comum com esta tese? Basicamente, é o fato de as pessoas que são gagas serem marcadas socialmente, como veremos a seguir.

Esta tese trata da disfluência na fala. Esse tema, até pouco tempo, não fazia parte do rol dos estudos linguísticos, principalmente, no que diz respeito aos trabalhos desenvolvidos no Brasil, que ficaram mais reduzidos à literatura especializada, como os estudos de fonoaudiólogos.

De uma maneira geral, conforme Jakubovicz (1992), classificar a disfluência é muito mais complicado do que parece, pois tudo o que é considerado como problemático na fala é visto como disfluyente. Esta autora considera quatro tipos de disfluência: a não fluência normal, a gagueira, a disprosódia orgânica e a taquilalia. A não fluência normal ocorre, devido à repetição das palavras e sintagmas e à prolongação esporádica dos sons e pausas. No segundo tipo de disfluência, há alteração na frequência, e os alongamentos apresentam uma duração maior, além de haver um número maior de repetições. A disprosódia orgânica, por sua vez, diz respeito à desorganização da fala e da linguagem e está ligada a fatores neurológicos. Por fim, a última é um problema relacionado com uma disfunção no sistema nervoso central, sendo identificada por um ritmo de fala mais acelerado do que o normal, pela incompreensão da linguagem e pela alteração na sequência de sons.

Entre os tipos de disfluência sugeridos, tomamos como base a não fluência, que aqui é subdividida em disfluência comum e disfluência gaga. Concomitantemente a isso, será preciso olhar para o que entendemos como fluência na fala.

Tradicionalmente caracteriza-se como fala fluente ou normal aquela em que não há oscilações, inserções, esforço muscular, entre outros, durante o seu fluxo contínuo (JABUCKOVICZ, 1992). Por outro lado, uma fala com disfluência comum, seria aquela em que aparecem alterações típicas em uma fala considerada normal, como inserções, repetições, prolongamentos, entre outros. Já a fala com disfluência gaga, de uma maneira geral, é caracterizada na literatura especializada por repetições de sílabas e sons, bloqueios, prolongamentos, pausas, entre outros. Divisão essa que, de antemão, já nos parece contraditória, assim como já notara Cavalcante e Brandão (2012).

Essa nomenclatura é contraditória por dois motivos principais. Primeiro porque, e assim como veremos no decorrer deste estudo, a fala fluente, a não ser por uma fala decorada, não existe. Conforme Cavalcante e Brandão (2012, p. 61), o sujeito fluente é “uma abstração ideal que inexistente, mas que é necessária para que possamos investigar a linguagem”. Sobre isso, Scarpa acrescenta

[...] a fluência é uma abstração metodológica, baseada na leitura ensaiada ou profissional de um texto escrito ou em textos orais decorados e ensaiados. O sujeito fluente é abstrato e integra-se em algum estilo de fala ou de comportamento social. No entanto, é com esta abstração ou esta ilusão – necessária, em termos de recorte epistemológico – que se [têm constituído os *corpora* analisados por modelos linguísticos] (SCARPA, 2006, p. 174).

As disfluências são, portanto, parte integrante da dinâmica da fala. Nas palavras de Scarpa e Fernandes-Svartman (2012, p. 30) são “atividades da língua”.

O segundo motivo concentra-se na presunção inadequada em estabelecer uma separação entre a disfluência gaga e a disfluência comum, pois quando levamos em consideração os eventos observados nos dois tipos de fala, vemos que mesmo sendo diferentes, eles se assemelham em muitos pontos. A diferença principal entre esses dois tipos de fala está, de um lado, relacionado a questões acústicas e, de outro, ao modo como o interlocutor, nesse caso, o não gago, sente a fala do outro, ou seja, a do gago. É nesse ponto que entra a questão do indivíduo marcado socialmente.

Em ensaio autobiográfico, intitulado “As pequenas memórias” de 2006, José Saramago diz:

O meu problema foi, e continua a ser, o tartamudeio, a gagueira. Aqueles que gozam da sorte de uma palavra solta, de uma frase fluida, não podem imaginar o sofrimento dos outros, esses que no mesmo instante em que abrem a boca para falar já sabem que irão ser objeto da estranheza ou, pior ainda, do riso do interlocutor.

Esse relato representa o que a maioria das pessoas gagas sentem, ou seja, o medo, a rejeição, a vergonha, o preconceito, entre outros. Fato que também se acentua por meio de piadas e músicas veiculadas pela mídia, que são, muitas vezes, de tom pejorativo.

Observe, no trecho da canção abaixo, que já em 1930, Noel Rosa, em o “Gago apaixonado”, retratava de maneira bem humorada o desespero de um homem que de tanto sofrer por amor estava ficando gago.

Mu-mu-mulher, em mim fi-fizeste um estrago
 Eu de nervoso estou-tou fi-ficando gago
 Não po-posso com a cru-crueldade da saudade
 Que que mal-maldade, vi-vivo sem afago
 [...]
 Tem tem pe-pena deste mo-moribundo
 Que que já virou va-va-va-va-ga-gabundo
 Só só só só por ter so-so-sofri-frido
 Tu tu tu tu tu tu tu tu
 Tu tens um co-coração fi-fi-fingido

É esse humor que gera o preconceito e, conseqüentemente, contribui para o que pode vir a ser a maior diferença entre gagos e não gagos.

É diante dessa perspectiva que tratamos a disfluência gaga por meio da sua singularidade e heterogeneidade (cf. Carneiro e Scarpa, 2012). A análise das disfluências comuns também ajuda a constatar o fato de que não existe uma fala completamente fluente, isto é, livre de hesitações, pausas, bloqueios, repetições, entre outros, ou de qualquer outro tipo de desvio. Conforme Carneiro e Scarpa (2012, p. 156), os fenômenos linguísticos caracterizadores da fluência ou da disfluência “indicam diferentes relações do sujeito com a língua, pois ambos são partes do funcionamento da fala, já que estão presentes em falas gagas e não gagas”.

O estudo também faz análises acústicas, com o intuito de caracterizar, compreender e descrever melhor os dois tipos de fala. Para dar conta disso, os parâmetros acústicos observados nos dois tipos de disfluência foram: a análise da frequência fundamental (F0), taxa de elocução, intensidade¹ e duração das pausas

silenciosas e respiratórias. Além disso, também foi feita a análise entoacional dos fenômenos a partir do modelo teórico de Pierrehumbert (1980) e Beckham e Pierrehumbert (1986). A divisão dos enunciados foi feita em grupos entoacionais de acordo com o modelo de Halliday (1970) e Cagliari (2007). Em seguida, foram desenvolvidas nomenclaturas que ajudassem a reconhecer os eventos disfluente, bem como BLO_c, para a presença de bloqueio em consoante, AL_v, para caracterizar o alongamento em vogais, entre outros. Com isso, será possível observar a frequência e distribuição desses eventos nos enunciados dos informantes.

A análise dos parâmetros acústicos, bem como a análise dos padrões entoacionais ajudarão a propor as similaridades e heterogeneidade entre os dois tipos de fala. Também nos ajudarão a observar as singularidades dentro de cada grupo, a partir da verificação de diferenças significativas ou não entre os elementos repetidos. Em decorrência disso, a análise acústica descritiva também nos ajudará a explicar o porquê de sentirmos tanta diferença nos dois tipos de fala.

Como sabemos, os fenômenos acústicos estão intimamente relacionados à articulação. Lindblom (1985) diz que a precisão articulatória está inversamente relacionada à fluência: em situações que exigem cuidadosa articulação, a fluência diminui. Por outro lado, em ocasiões que exigem cuidadosa fluência, a articulação fica comprometida. Segundo Scarpa (1995), isso significa que é difícil falar rápido e de modo acurado ao mesmo tempo, o que está relacionado a “uma propriedade emergente da dinâmica auto-organizadora do processamento fonético” (SCARPA, 1995, p. 7). Ainda sobre essa questão, Mata (1999, p. 348) afirma que

O visível controle dos níveis de F0 [frequência fundamental] relativamente ao contexto prosódico adjacente é, sem dúvida, um importante indicador da fluência e “boa formação” do discurso. As pausas preenchidas, por exemplo, podem não ser consideradas como disfluente. Nos casos em que o contorno global de F0 não é perturbado, elas funcionam como importantes elementos de transição.

O ritmo também é essencial na fala, e está intimamente relacionado aos fenômenos articulatórios. Ele caracteriza-se por uma simetria, “uma harmonia resultante de certas combinações e proporções regulares” (CAGLIARI, 2007, p. 132). Segundo o autor, não há um parâmetro que gere o ritmo na fala, sendo manifestado pelos elementos que apresentam momentos de redução e hesitação na dinâmica da fala. Sendo assim, também consideramos que os desvios na fala representam uma quebra na estrutura rítmica no momento da articulação dos sons, gerando os dois tipos de disfluência. O que

também pode ocorrer devido a um erro durante a programação da duração das sílabas antes da articulação dos sons (CAGLIARI, 2007).

Esta tese está centrada, portanto, em três hipóteses principais. A primeira é a de que uma das funções das disfluências seria a de atrasar o fluxo da fala, a fim de ajudar o informante a recuperar a próxima palavra. A segunda pressupõe que os momentos disfluentes sejam mais frequentes, em um mesmo enunciado, na fala gaga, o que contribuiria para a maior diferenciação entre a fala gaga e não gaga. A terceira hipótese é uma consequência da segunda e considera que a quantidade de disfluências gagas causa um desconforto no ouvinte, o que gera o preconceito e as piadas mal intencionadas. Esse fato, por sua vez, ocasiona um processo de empatia² fonética negativa, que leva a uma quebra na expectativa do interlocutor durante a fala gaga. A empatia diz respeito à capacidade do ouvinte em sentir biologicamente, de modo positivo ou negativo, a fala do outro. A citação abaixo ajuda a entender um pouco esse evento, uma vez que o que temos são sujeitos que estão sob o efeito de fenômenos linguísticos que, incontrolavelmente,

[...] põem sua fala à deriva, enveredando no inesperado; põem em destaque o movimento, articulatório ou corporal daquele falante; põem, falante e ouvinte, à escuta dos movimentos da língua; deslocam do sentido para o corpo. Geram tensão no corpo do falante. Geram tensão no ouvinte (CARNEIRO e SCARPA 2012, p. 157).

No decorrer deste trabalho também mostraremos alguns estudos desenvolvidos dentro e fora do Brasil relacionados ao fenômeno da disfluência comum e gaga, a partir de diferentes perspectivas.

No Brasil, os trabalhos do ponto de vista da linguística ainda são escassos. A precursora dos estudos nessa área foi Ester Mirian Scarpa (1995, 2006, 2012). A autora observou as disfluências tendo em vista a fala infantil. Os seus estudos estão mais voltados para questões relacionadas à fluência e disfluência comum. Depois, há os trabalhos que tratam de fenômenos mais pontuais, como por exemplo, a dissertação de Delfino (2009) sobre as disfluências de reparo e a tese de Merlo (2012), que observou a não aleatoriedade das pausas demarcativas e das hesitações, dentre outros estudos.

² O termo empatia já havia sido utilizado nos estudos fonéticos por Abercrombie (1969) e Cagliari (1982, 2007) para os estudos do ritmo da fala, e mais recentemente por Quezada et al (2012), que propuseram a união entre fonética e empatia. No entanto, o termo “empatia negativa” aparece pela primeira vez, neste trabalho, para as disfluências comum e gaga.

Em Portugal aparecem os estudos de Moniz (2006), que procurou caracterizar as disfluências em Português Europeu, os estudos recentes de Cruz (2009) e Cruz e Frota (2010), que trataram a gagueira de acordo com o ponto de vista da prosódia, e o trabalho de Condélipis (2010), que observou a gagueira do ponto de vista linguístico e psicolinguístico.

Há, ainda, trabalhos em língua inglesa, como o de Goldman-Eisler (1968); o de Levelt (1989), o de Shiriberg (1999); o de Guitar (2006); o de Bailey e Ferreira (2004), entre outros trabalhos que serão apresentados.

Vale lembrar que os estudos feitos a partir de considerações patológicas, sociais e comportamentais não foram deixados de lado, mas sim tomados como ponto de referência para um melhor conhecimento do tema.

Ao chamar a atenção para a questão da disfluência gaga e da disfluência comum, o presente trabalho pretende contribuir para o aumento do número de estudos sobre o tema na área da Linguística, no Brasil, bem como discutir os limites entre a fluência e a disfluência (Scarpa; Novaes-Pinto, 2012, p. 5), sem atribuir um caráter desviante a elas. Além disso, o *corpus* que será formado aqui poderá ajudar fonoaudiólogos a caracterizar tipos e ocorrências das disfluências e, ainda, auxiliar os professores alfabetizadores ou não a reconhecerem os dois tipos de fala.

Intencionamos igualmente, aqui, a partir dos resultados obtidos, se não extinguir, ao menos diminuir o preconceito por parte da sociedade de um modo geral e ajudar os próprios falantes em relação à sua fala. Essa intenção caracteriza a inovação deste trabalho, pois ao comparar, por meio de uma análise linguística, as disfluências comum e gaga também junta dados quantitativos e qualitativos. Os trabalhos recentes usam a disfluência comum apenas como apoio, não há uma análise comparativa mais detalhada sobre o assunto. Um olhar sobre muitos parâmetros, mesmo que superficialmente, ajuda a entender melhor o fenômeno e a planejar futuras pesquisas, investigando características e fatores específicos.

A partir disso, o objetivo geral desta tese consiste em comparar, a partir de uma perspectiva linguística do aspecto acústico, as disfluências comum e gaga, a fim de mostrar que, mesmo sendo fenômenos diferentes, é possível encontrar semelhanças dentro da singularidade que cada uma delas possui.

Diante desses objetivos esta tese se divide em quatro capítulos.

O primeiro capítulo trata da fundamentação teórica. Primeiro, discute-se a disfluência comum a partir de sua relação com a fluência. O capítulo inicia-se

questionando se a fluência da fala realmente existe. A partir disso, desenvolve-se a discussão, assim como a definição proposta para esse evento, a partir de diferentes pontos de vista. Na sequência, caracterizamos os diferentes tipos de disfluência, como repetições, pausas, alongamentos, bloqueios, correções, falso começo, entre outros. Por fim, trazemos algumas suposições sobre a não intencionalidade das disfluências.

Também é apresentado nesse capítulo uma visão geral sobre a gagueira, bem como o que teria sido o primeiro relato do fenômeno e as possíveis teorias que envolvem esse tipo de disfluência. Em seguida, abordamos o ponto de vista da literatura especializada e, na sequência, o ponto de vista dos estudos linguísticos sobre a gagueira no Brasil. Posteriormente, é feita uma caracterização do fenômeno, que segue principalmente os conceitos fornecidos pela literatura especializada. Nesse capítulo discutimos a hipótese de que a disfluência gaga estaria relacionada ao que denominamos de empatia fonética negativa, isto é, o sentimento do ouvinte em relação à fala do outro.

Na sequência fazemos uma revisão sobre os aspectos que envolvem a articulação da fala, a fim de trazermos uma introdução geral às questões fonéticas relativas à produção da fala. Também é feita uma apresentação dos fenômenos suprasegmentais, como a tessitura, variação melódica, duração, ritmo, já que são elementos que estão intimamente relacionados à articulação e, além disso, são responsáveis por garantir o fluir da fala. Aqui também é abordado o ponto de vista cognitivo do processo de produção de fala proposto por Levelt (1989).

O segundo capítulo contempla o quadro teórico usado para embasar este trabalho. O modelo teórico apresentado é o desenvolvido por Pierrehumbert (1980) e a sua modificação, desenvolvida por Beckham e Pierrehumbert (1986). O modelo autosegmental favorece a organização hierárquica dos constituintes prosódicos. É um sistema que também conjuga os fatos fonológicos e fonéticos. Em seguida, por meio dos trabalhos de Ladd (1996) e Gussenhoven (2002), fazemos alguns apontamentos sobre os possíveis problemas com a teoria. Em seguida, apresentamos uma revisão sucinta do modelo teórico de Halliday (1970) e Cagliari (2007)³, a fim de mostrarmos como é feita a distribuição dos enunciados em grupos tonais.

³ O modelo funcional de Halliday (1970) e Cagliari (1982/2007) foi utilizado neste trabalho para a segmentação dos enunciados em grupos tonais (GTs), e não para a análise dos padrões entoacionais. Para essa etapa optamos por utilizar o modelo de Pierrehumbert (1980) e Beckham (1986), por serem um dos modelos entoacionais mais utilizados na pesquisa linguística. Sendo assim, não trataremos das questões específicas do modelo funcional.

No terceiro capítulo seguimos com a apresentação da metodologia utilizada durante a pesquisa, expondo os dados acústicos de leitura, bem como os dados do *corpus* e dos informantes. Em relação à análise acústica, destacamos a metodologia desenvolvida e proposta nesta tese para a análise dos dois tipos de disfluências. Também apresentamos nesse capítulo os procedimentos utilizados para a análise dos elementos suprasegmentais e as adaptações necessárias ao modelo entoacional de Pierrehumbert (1980) para a análise dos padrões entoacionais.

No quarto capítulo apresenta-se a análise e discussão dos resultados obtidos por meio das análises acústicas para cada um dos informantes gogos e não gogos, a fim de apresentarmos melhor os resultados obtidos para cada um deles. Nesse capítulo também é feita uma comparação geral dos resultados obtidos para cada um dos grupos.

Por fim, de acordo com os resultados e discussões obtidos apresenta-se as conclusões obtidas com este trabalho, bem como as suas contribuições futuras.

1 FUNDAMENTAÇÃO TEÓRICA

1.1 Sobre A (Dis)Fluência

In all spheres of life, hasty engagement and unreflective enactment are evidence of brutalization. Hesitation alone is human
(Werner, 2006, p. 92).

Disfluência é o termo mais novo que aparece relacionado às perturbações da fluência na fala. De uma maneira geral, são as interrupções que ocorrem durante o fluxo da fala. O termo surgiu na área patológica e fonoaudiológica com o intuito de nomear as interrupções que não são tratadas como normais, entre elas, a gagueira, termo de domínio popular e tema deste trabalho. Outra forma de se fazer referência ao fenômeno era com o termo “hesitação”.

De acordo com Bortfeld et al (2001), desde 1950 os estudos sobre a disfluência comum têm sido tratados como um caminho para a produção linguística intermediária e do processo cognitivo de planejamento da fala.

Atualmente, os estudos linguísticos sobre a disfluência comum são muito importantes, pois ajudam a caracterizar os diferentes tipos de fala, auxiliam os fonoaudiólogos nos diagnósticos clínicos e ainda permitem que o professor alfabetizador saiba distinguir a disfluência comum da disfluência gaga.

Para Delfino (2009), os estudos sobre a disfluência nos apresentam novas pistas sobre a maneira como ocorre o processamento da linguagem, uma vez que é a fala mais próxima da real e, grande parte dos estudos considera apenas a fala filtrada, livre de correções.

Consideramos aqui que as disfluências podem ser de dois tipos: comum ou gaga. Os termos disfluência comum e disfluência gaga não são novos. Já podemos encontrar na literatura diferentes trabalhos que optaram por essa nomenclatura.

Nesta seção, faremos uma discussão sobre a disfluência comum, relacionando-a com a fluência.

É impossível falarmos em disfluência comum, sem antes tratar da fluência, uma vez que a primeira depende da última. Os estudos sobre a fluência só ganharam força em função da análise da sua contraparte, a gagueira. Portanto, a sua definição passou a ser tomada a partir do seu lado negativo, a "fluência é melhor definida como uma unidade de resposta *destituída de disfluências, prolongamentos e pausas*" (HEDGE, 1978 apud SCARPA, 1995, p. 2, grifos da autora).

Outra caracterização de fluência pelo seu lado desviante, segundo Scarpa, é dada por Plunkett (1993), que diz que ela “se refere ao nível de integração de uma sequência de fonemas que diferencia um desempenho que flui macio de um (desempenho) aos trancos” (PLUNKETT apud SCARPA, 1995, p. 167)⁴. De acordo com Scarpa, a caracterização de Plunkett, mesmo sendo um pouco imprecisa, logo é diminuída devido à sua referência aos aspectos constitutivos da dinâmica da fala.

De acordo com Plunkett (1993), Lindblom (1985) postulou que há uma relação inversa entre precisão articulatória e fluência, ou seja, não é possível haver, ao mesmo tempo, precisão articulatória e fluência. “Este tipo de relação de troca entre precisão articulatória e fluência é explicada como uma propriedade emergente da dinâmica auto-organizadora do processamento fonético” (SCARPA, 1995, p. 167).

Scarpa (1995) explica que a questão fonética diz respeito às possibilidades de adaptações da fala, ou seja, a possibilidade de passar de um estilo de fala rápido para um vagaroso, mudanças na altura, de alto para baixo, adaptar a fala em situações formais e informais. Em outros termos, seria a capacidade de adaptar/modular a fala de acordo com os contextos em que estamos inseridos.

Para Scarpa (1995, p. 164),

A definição pela negativa explica parcialmente o fato de que a grande concentração de trabalhos ou reflexão sobre a fluência se encontra, na verdade, no campo de estudo dos distúrbios e da terapia da fala. Os fonoaudiólogos e foniatras se interessam pela gagueira – ou pelos ‘distúrbios de fluência’ – isto é, pelos casos em que o centro de interesse do investigador é exatamente explorar as causas e características da face desviante [...] o oposto radical do termo neutro e ideal da fluência.

Além disso, segundo a autora, caracterizar a fluência pelo seu lado desviante traz problemas quanto à sua definição, pois, não está claro se ela seria interpretada pelo ouvinte como fluente ou se diz respeito ao que seria uma fala ‘normalmente fluente’. É preciso especificidades na definição, uma vez que o que o ouvinte interpreta como uma fala fluente distingue-se do que seria uma fala normalmente fluente.

O termo fluência pode ser encontrado em várias áreas. No entanto, não há um consenso sobre a sua definição. De acordo com Merlo (2006), tanto a Linguística quanto a Fonoaudiologia a colocam em uma posição marginal de importância. Porém, no ensino de línguas estrangeiras a fluência recebe um tratamento diferente, ocupando um lugar essencial. Isso ocorre devido à sua relação com o julgamento de proficiência dos alunos e adaptação às

⁴ Tradução de SCARPA, Ester. Marian.

diferentes situações comunicativas. Já no campo da Neurologia e da Psiquiatria, a fluência tem como função principal caracterizar eventos patológicos.

De um modo geral, a fluência é entendida como aquela em que a linguagem flui normalmente, sendo facilmente compreendida pelo ouvinte. Em outras palavras, é o fenômeno em que os elementos segmentais e suprasegmentais realizam-se em conformidade com os padrões prosódicos (ABERCROMBIE, 1967). Quando ocorre uma ruptura nesses padrões, temos o que se denomina disfluência, que pode ser comum ou gaga.

Um dos primeiros estudos a investigar a fluência foi o de Fillmore (1979). Segundo Scarpa (1995), o autor reconhece a variedade de habilidades linguísticas que o termo fluência admite, de modo que uma pessoa fluente seria classificada como bem articulada, eloquente e clara.

Tomando como base essas características, Scarpa (1995) apresenta quatro tipos de fluência propostos por Fillmore. Para ele, a fluência está relacionada ao uso efetivo da língua, fazendo, portanto, parte da gramática da língua. De acordo com Moniz (2006), o autor trata essa definição como um conceito satélite em quatro dimensões. A partir de Scarpa (1995, p. 6) seguem os quatro tipos de fluência dos quais o autor fala.

1º) O primeiro tipo demonstra capacidade e habilidade de falar extensamente, sem muitas pausas, com habilidade para preencher os vazios com a fala. O falante desse tipo dificilmente planeja ou faz pequenas pausas para pensar sobre o que vai falar. Os falantes característicos desse tipo de fala são os locutores esportivos;

2º) O segundo tipo está relacionado à capacidade de domínio dos recursos semânticos e sintáticos da língua. Os falantes têm facilidade de produzir sentenças coerentes, “semanticamente densas”;

3º) O terceiro tipo de fluência caracteriza pessoas que conseguem dominar qualquer tipo de assunto, ou seja, têm habilidade em dizer palavras apropriadas em diferentes contextos conversacionais. Por outro lado, a pessoa não fluente nesse contexto, seria aquela que, em contextos familiares, não consegue falar muito bem, diferentemente do que aconteceria se estivesse em situações de confronto ou na presença de estranhos;

4ª) O quarto tipo mostra a capacidade e a habilidade “de demonstrar uso criativo e imaginativo da linguagem, de expressar as ideias de modo diferente, no uso de trocadilhos, na criação de metáforas etc” (SCARPA, 1995, p. 6).

Kormos e Denés (2004) trazem uma extensa definição sobre a fluência, a partir da visão de alguns estudiosos. Dentre eles, apresentaremos as definições mais apropriadas para o

desenvolvimento desta tese. Segundo as autoras, Rehbein (1987) acredita que a fluência depende do contexto. De acordo com o autor, seria a capacidade simultânea de planejar e enunciar a fala.

Lennon (1990), por sua vez, trata a fluência como “*an impression on the listener’s part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently*”⁵ (LENNON, 1990, p. 149). Já Schmidt (1992) considera que a fluência é, na produção da fala, uma habilidade processual automática, que não requer tanta atenção ou esforço por parte do falante. Essa definição também foi adotada por Kormos e Denés (2004).

Para Andrade (2001) a fluência diz respeito a uma fala contínua e suave. Já para Oliveira e Gargantini (2003) a fala fluente é aquela que requer pouco esforço motor, emocional e cognitivo em sua produção. Jakubovicz (1997) diz também, que além de ser produzida sem esforço, a fala fluente deve ser percebida pelo ouvinte como normal.

Como vemos, há na literatura muitas tentativas para caracterizar a fluência na fala. Quando falamos em tentativas, não estamos dizendo que a fluência na fala não exista. Sim, ela existe. No entanto, ela não é frequente ao longo do discurso. A não ser por uma fala decorada, não existe fala completamente fluente. Os falantes são, a todo o momento, interrompidos por alguma desorganização no discurso, seja ela intencional ou não. Os falantes são frequentemente disfluente (ARNOLD *et al*, 2013).

Por esse motivo, entendemos que a fala comum, espontânea – aquela do convívio familiar, das conversas entre amigos, e até mesmo as formais, como as da sala de aula, de um discurso, de uma entrevista de emprego, dos sermões dos padres – está repleta de disfluências. Porém, são “lapsos normais”, comuns na língua. Por isso, são tratados aqui como disfluência comum.

Quando falamos em “lapsos normais”, não queremos dizer que a fala gaga não é normal. No entanto, a maior frequência desses eventos causa uma empatia⁶ negativa no ouvinte. Em outras palavras, causa um estranhamento na fala do outro, fazendo com que a fala gaga seja sentida como “errada”.

A fluência ocorre com mais frequência em uma fala controlada, como as falas ensaiadas dos atores de teatro e novela, dos locutores esportivos, dos radialistas.

⁵ “uma impressão por parte do ouvinte de que os processos psicolinguísticos de planejamento e produção de fala estão funcionando de forma fácil e eficiente” [Tradução nossa].

⁶ Esse fenômeno será tratado no capítulo referente às disfluências gagas.

Os narradores esportivos, por exemplo, são treinados para produzirem uma fala muito concatenada. Por isso, nos dão a impressão de estarem falando muito rápido. Em geral, eles são treinados para acompanhar a velocidade da competição esportiva. Os atores de televisão e teatro, por sua vez, possuem um texto que deve ser decorado. E, no caso de um erro, é possível, para os atores de televisão, gravar a cena novamente. Por isso, poderíamos dizer, com 99,9% de certeza, que esse grupo seria a representação do que vem a ser uma fala fluente.

No entanto, mesmo a fala mais controlada não está totalmente livre de lapsos. Perceba, por exemplo, que os comentários feitos pelos locutores, durante uma partida de futebol são muito mais disfluente do que durante a narração de uma partida. Isso pode ocorrer porque nesse momento as ideias estão em fase de desenvolvimento, diferente da narração em que temos uma fala mais treinada. Um ouvinte atento já deve ter percebido períodos de disfluências nesses tipos de fala. Contudo, se formos levar em consideração o que dizem O’Connel e Kowal (2010), a fluência sequer existiria em uma fala mais controlada, “*in the sense of absolute continuity, perfect fluency of sequential ordering is impossible by reason of the need to breath*” (O’CONNEL; KOWAL; 2010, p. 27). Segundo os autores, as pausas para respirar não são opcionais, o que faz a fala disfluente. Eles dizem que nós respiramos, em média, a cada cinco segundos, ou doze vezes por minuto.

De acordo com Koch (2006), a fala não planejada apresenta-se “em se fazendo”, em sua origem. É uma constante construção. Segundo a autora, o planejamento e a verbalização ocorrem simultaneamente na fala, pois no momento da interação é que ela nasce, é o seu próprio rascunho.

Koch compara a fala a um filme, devido ao seu caráter dinâmico e, por isso, descontínuo. O texto escrito, por sua vez, por ser estático é comparado a um quadro. Para a autora, a descontinuidade frequente no fluxo discursivo deve-se a fatores de ordem cognitivo-interacional.

[...] o texto falado não é absolutamente caótico, desestruturado, rudimentar. Ao contrário, ele tem uma estruturação que lhe é própria, ditadas pelas circunstâncias sociocognitivas de sua produção e é à luz desta que deve ser descrito e avaliado. (KOCH, 2006, p.46).

A fim de buscar uma melhor conceituação para a fluência, Merlo (2006) procurou, em alguns dicionários de Linguística, uma definição para o fenômeno. Mas, ela relata que dos

⁷ “no sentido da continuidade absoluta, a fluência perfeita da ordem sequencial é impossível em razão da necessidade da respiração”[Tradução nossa].

nove dicionários consultados, apenas três apresentaram uma definição para o verbete. Uma das definições propõe que a fluência se individualiza por meio das características [...] *which give speech the qualities of being natural and normal, including native-like use of pausing, rhythm, intonation, stress, rate of speaking, and use of interjections and interruptions* [...] (RICHARDS, PLATT; WEBER 1985, p. 107-8 apud MERLO 2006, p. 13)⁸.

É natural, quando as pessoas falam, produzirem disfluências. De acordo com Ferreira e Bailey (2004), a disfluência comum é uma perturbação não prejudicial, pois *Conversation would be impossible if efficient mechanisms for dealing with uhus, ums, repeats and repairs were not a basic part of architecture of the human comprehension system*⁹ (FERREIRA; BAILEY, 2004, p. 231).

Os autores apresentam três motivos para o estudo da disfluência comum. Um deles está associado, justamente, ao fato da normalidade da disfluência na fala. O que é confirmado pelo dado dos autores de que em cada 100 palavras, podem ocorrer, em média, de seis a dez disfluências. Na maior parte das vezes, as disfluências comuns são imperceptíveis aos ouvintes.

A outra razão deve-se ao fato das disfluências introduzirem um novo conteúdo lexical, através das repetições, correções, entre outros. De acordo com Ferreira e Bailey (2004), esse novo conteúdo formaria um enunciado agramatical, que permite diferentes possibilidades de estudo. Por exemplo, eles dizem que o fragmento “*what museum is it is it in*”, apresenta um conteúdo não gramatical, uma vez que a repetição de *is it* não está sintaticamente licenciada. Mesmo assim, esse lapso não ocasiona uma falha na compreensão humana. Fato que sugere outros meios para distinção dessas rupturas do resto do *input*.

Por fim, o terceiro motivo propõe que os mecanismos sintáticos (*parser*) para a manipulação das disfluências ajudariam a compreender a estrutura em que esses mecanismos são incorporados. Por exemplo, em uma correção do tipo “*you will put – you should drop the ball*”, no nível descritivo, a sintaxe integra a sequência “*you will put*” como parte do enunciado, de modo que o sistema linguístico não inutiliza as expressões que não estão sintaticamente corretas na sentença.

Ao citar Satir (1988), Starkweather e Givens-Ackerman (1997) afirmam que a comunicação envolve aspectos verbais e não verbais. Aos aspectos não verbais estão relacionadas questões como a respiração correta, a postura corporal, os gestos, a expressão

⁸ [...] que conferem à fala as qualidades de ser normal e natural, incluindo o uso nativo, como o da pausa, ritmo, entoação, tonicidade, velocidade de fala, e o uso de interjeições e interrupções. [Tradução nossa].

⁹ “A conversação seria impossível se mecanismos eficientes para lidar com *uhus, ums*, repetições e reparações não fossem uma parte básica da arquitetura da compreensão humana” [Tradução Nossa].

facial, o ritmo e as pausas. Elementos que segundo as autoras estão relacionados à fluência normal da fala. No entanto, é impossível mantermos a normalidade constante desses fenômenos, o que desencadeia a disfluência comum e a disfluência gaga.

Merlo (2006) faz a sua distinção a partir dos termos: fala fluente e não fluente. Para a autora elas são distintas pelo fato da primeira apresentar baixa frequência de oscilações, hesitações, reformulações, maior habilidade gramatical e menor complexidade semântica, além de uma facilidade de emissão e taxa de elocução confortável. Merlo tem razão quanto à separação utilizada. No entanto, não concordamos com os termos utilizados por ela, uma vez que acreditamos, como dissemos, que a fala fluente normal não existe, além de generalizar a sua ocorrência. Talvez a nomenclatura mais adequada fosse entre *fala controlada* e *fala com disfluência comum*. Mesmo assim, essa nominalização não se encaixaria, uma vez que mesmo sendo pequena, é possível que ocorra disfluência na fala controlada.

De acordo com Starkweather e Givens-Ackerman (1997), as pessoas apresentam graus de fluência variados. Segundo os autores esses graus de fluência podem variar conforme a linguagem, quando há facilidade em encontrar palavras, construir sentenças e fazer sequenciamentos fonológicos sem muito esforço ou, em relação à fala, devido à facilidade do falante em iniciar uma conversa e manter a sua suavidade durante o fluxo de fala. Para Starkweather e Givens-Ackerman (1997), a fala é construída por meio do conjunto entre fluência e disfluência, variando de acordo com o momento de fala e as condições do falante.

Kent e Curlee (1991 apud Oliveira e Gargantini 2003) acreditam que a fala é formada por componentes linguísticos e paralinguísticos. Para eles, a fluência ocorre quando esses componentes estão em sintonia. Caso ocorra uma falha no processo, teríamos o que chamamos de disfluência comum. “Para se garantir a recuperação da fluência e a qualidade rítmica da emissão, é fundamental o reequilíbrio automático dessa defasagem temporal entre as distintas operações” (Oliveira e Gargantini 2003, p. 56). Segundo Oliveira e Gargantini (2003) o processo de recuperação da fluência envolve questões ambientais e pessoais. Nos processos pessoais inserem-se questões emocionais e, nos ambientais, o contexto da interação durante a comunicação.

De acordo com Souza e Silva; Crescitelli (2006) as disfluências comuns são melhor percebidas por meio da transcrição ortográfica ou fonética, do que quando analisadas somente por análises auditivas. Segundo elas, isso ocorre porque “as variações de ritmo, velocidade e entonação, além do preenchimento de pausas, sem dúvida, ajudam a preencher os vazios que,

na verdade, são o pano de fundo do texto falado, e não imperfeições” (SOUZA e SILVA; CRESCITELLI, 2006, p.86).

Ao procurar distinguir passagens fluentes de passagens não fluentes, Scarpa (1985, p. 171) diz que os trechos fluentes são os “já ajeitados, conhecidos, analisados ou, na grande maioria dos casos, congelados, vêm em bloco”. Já os trechos disfluentes “são aqueles em construção, instáveis, com tentativas infrutíferas de segmentação em blocos prosódicos; supõem passos mais complexos tanto paradigmática quanto sintagmaticamente na elaboração do enunciado”.

Conforme Ferreira e Bailey (2004), os falantes tendem a ser disfluentes em pontos sintáticos específicos. Por exemplo, as pausas preenchidas e as palavras repetidas são mais comuns em início de sentença. Conforme os autores, quando há coocorrência de informações em uma frase, essa situação fica melhor resolvida em um evento disfluente, pois o ouvinte saberá distinguir melhor a estrutura sintática dos elementos envolvidos. Por exemplo, na frase “*Bush spoke to Blair and uh the media reported de story*”, Ferreira e Bailey explicam que é mais fácil para o ouvinte entender o termo “*the media*” como o sujeito na nova frase, do que para quem ouve a mesma sentença sem o evento disfluente. Isso acontece, segundo eles, porque o evento disfluente, *uh*, antecede a frase substantiva, que tem um papel temporariamente ambíguo. Ou seja, ela é ao mesmo tempo, o final do objeto de *spoke to* e o início de uma nova sentença. Assim, como as disfluências tendem a aparecer no início do enunciado, ela permite que sua análise ocorra mais tardiamente, de modo que o mecanismo de processamento das frases tem uma sugestão do que seria a estrutura correta. Então, o processamento linguístico usa a disfluência como uma informação.

Ao relacionar as disfluências ao processamento linguístico, Delfino (2009) acredita que as disfluências indicam tanto a necessidade do falante em planejar sua fala, quanto o modo como o sujeito organiza as informações que recebe do mundo. Nesse caso, funciona também como pistas para o julgamento metalinguístico do ouvinte.

Para Shireberg (2001) a fala mais planejada leva a uma maior produção de pausas e repetições, enquanto que a menos planejada é mais tendenciosa à produção de correções e truncamentos. O que é perfeitamente plausível, já que diferentemente do texto escrito, podemos fazer correções e revisões.

De acordo com Scarpa (1995), os gestos fonéticos organizam-se e adaptam-se em função dos estilos de fala como, por exemplo, uma mudança de fala rápida para lenta, baixa

para alta, informal para bem articulada, íntima para pública. Até involuntariamente modulamos nossa fala em função de fatores emocionais e fisiológicos.

[...] a fala é adaptativa [...] não lança mão de recursos mais potentes nem maiores do que o necessário. Não é mais elaborada do que precisa ser. A fala normal é, segundo Lindblom, um '*pianissimo* fisiológico', o que aponta para o fato de que aparentemente severas restrições governam a seleção de valores linguístico-fonéticos. (SCARPA, 1995, p. 6).

Quanto a essas restrições, Scarpa diz, seguindo as afirmações de Lindblom (1885), que elas não se centram na invariância do sinal. Portanto, não são nem necessárias, nem suficientes ao acesso lexical. Isso porque, ela diz, além da estrutura das línguas apresentarem redundâncias, a percepção da fala é voltada para dois tipos de informações: uma que depende do sinal e outra que independe dele. Na independência do sinal, a informação que é interessante, é de caráter linguístico.

Deixando a percepção de lado, e centrando-se mais na língua, Lindblom sugere, segundo Scarpa (1995), que essas restrições estão relacionadas “ao princípio de contraste perceptual suficiente”, que está em equilíbrio com o “princípio da simplificação articulatória”. Conforme Scarpa, o último princípio diz que “os gestos fonéticos não são feitos mais distintivos do que precisam ser” (SCARPA, 1995, p. 7). Ou seja, a fala poupa-se de gestos articulatórios muito extremos. Como ela exemplifica, é possível pronunciarmos um [i] com a mandíbula abaixada, mas por ser um esforço desnecessário, pode ser descartado. Assim, a autora diz que essa tensão (ou equilíbrio) entre simplicidade e complexidade está na origem dos processos essenciais à dinâmica da fala. Esse fato mostra, segundo Scarpa “que, quando demandas por contrastes perceptuais diminuem, o gesto se simplifica, cedendo lugar à disfluência” (SCARPA, 1995, p. 7). Portanto, ela conclui que a fluência e a disfluência estão inseridas no mesmo processo de dinâmica do processamento da fala, de modo que o que gera a fluência também gera a disfluência. Daí, a dificuldade, que mencionamos, em se falar somente em disfluência.

Para Scarpa (1995) essa falta de mensuração adequada para classificar os termos fluência/disfluência, também se deve ao seu caráter material e comportamental, uma vez que as interpretações ligadas a este termo vão além do seu significado. Isso nos leva a acreditar, baseados em Perkins 1971, Finn e Ingham 1991, que a fluência seria “o barômetro para o sistema de fala como um todo, cujos limites são estabelecidos pela adequação do desempenho das dimensões semânticas, sintática e morfológica da fala”. Mesmo sendo a fala o ambiente

favorável para acontecer o fenômeno, é difícil medir corretamente os aspectos da fluência que os diferenciam da disfluência.

Para Koch (2007), as disfluências devem-se a situações de interação entre o falante e o interlocutor. Segundo ela, no texto falado, há uma “co-produção discursiva” entre os interlocutores, que cooperam, “co-negociam” e “coargumentam” durante a produção do texto falado. Nessa interação, ela explica que ocorrem pressões pragmáticas, que se sobrepõem às normas da sintaxe. Então, é nesse momento que acontece o que chamamos de disfluência comum. Segundo a autora, esses fenômenos são de extrema relevância e apresentam funções cognitivo-interacionais.

Koch refere-se às pressões pragmáticas como “estratégias” cognitivo-conversacionais usadas pelo falante. São essas estratégias que mostram que o texto falado apresenta uma estruturação própria. A autora (KOCH, 2009, p. 81) descreve essas habilidades de estruturação em quatro tipos:

- 1) Sempre que perceber que o parceiro já compreendeu o que você pretendia comunicar-lhe, a continuação de sua fala se tornará, na maioria das vezes, desnecessária;
- 2) Logo que perceber que seu interlocutor não o está entendendo devidamente, suspenda o fluxo da informação e repita, parafraseie, mude o planejamento ou insira explicações e/ou exemplos;
- 3) Ao perceber que formulou algo de maneira inadequada, interrompa-se imediatamente e proceda a uma correção;
- 4) Ao se dar conta de que disse algo que é ou poderia ser ofensivo à face do seu interlocutor ou que foi excessivamente categórico naquilo que disse, proceda imediatamente a um reparo, acrescentando ou inserindo expressões atenuadoras ou modalizadoras.

São essas estratégias que conferem ao texto falado o seu caráter não caótico, desempenhando funções cognitivas e interacionais, afirma Koch. Para a autora, as principais estratégias para o processamento do texto falado são: a inserção, a reformulação, que pode ser retórica ou saneadora, e a hesitação. Dentro dessa perspectiva, Scarpa (1995) diz que os fenômenos disfluente são atividades epilinguísticas. Ou seja, são atividades “que o sujeito faz com a linguagem, atuando consigo mesmo ou através da intermediação do outro, de discursos anteriores, da própria forma linguística etc” (SCARPA, 1995, p. 178).

A seguir, apresentaremos cada um dos recursos considerados por Koch, acrescentando mais alguns que também estão presentes na fala. Como veremos, Koch (2009) e Marchuschi (2006) dividem a hesitação em subtipos, como repetições, pausas, alongamentos, entre outros.

Neste trabalho consideramos a hesitação, seja prosódica ou não, como todo fenômeno que se caracteriza, na fala, como uma disfluência, ou melhor, como uma estratégia discursiva. Vale ressaltar que as disfluências apresentadas aqui são características (e foram evidenciadas) tanto da fala com disfluência comum quanto da fala gaga.

1.3 As hesitações (ou disfluências)

As hesitações são caracterizadas por Koch (2009, p. 91) como “constitutiva do próprio processo de construção do texto falado”. A autora explica que, no texto falado, o planejamento e a verbalização são simultâneos. “Não existe, assim, trecho de fala sem hesitações, ao passo que podem existir trechos, mais ou menos longos, sem inserções ou reformulações, quer retóricas, quer saneadoras” (KOCH, 2009, p. 91).

Por ser um fenômeno indissociável da fala, a autora diz que as hesitações são diferentes da inserção e da reformulação (como veremos) pois, na maior parte das vezes, não podem ser cognitivamente controladas, mas apenas parcialmente, principalmente quando assinala a dificuldade do falante durante a verbalização dos enunciados.

Segundo ela, as hesitações manifestam-se através de pausas (preenchidas ou não), alongamento de vogais, consoantes ou sílabas (iniciais ou finais), repetição de palavras, truncamentos, reformulações, entre outros. Esses eventos podem ser comparados ao que Chomsky (1965) chamou de “erros de performance”, contribuindo para a descontinuidade da fala.

Esses fenômenos revelam, então, a dificuldade de construção do texto falado, devido às inúmeras atividades que o falante realiza a fim de produzir sentidos.

Na interação verbal face a face, ele (o falante) tem de estar constantemente alerta, controlando ‘pari passu’ tal processo, sempre pronto a facilitar a compreensão dos parceiros, a atender suas solicitações, a (re-)negociar com eles o sentido do que está sendo dito, através, em grande parte, de estratégias de identificação referencial. Tudo isso exige o domínio não só de habilidades linguísticas, como também de estratégias de ordem cognitiva, social e cultural. (KOCH, 2009, p. 92).

Marchuschi (2006) divide as disfluências em:

1. **fenômenos prosódicos**, como pausas (prolongadas) e alongamentos vocálicos;
2. **expressões hesitativas**, como éh, ah, ahn, ente outros;
3. **itens funcionais**, como artigos, preposições, conjunções, pronomes, verbos de ligação;
4. **itens lexicais**, como substantivos, advérbios, adjetivos, verbos;

5. **marcadores discursivos acumulados**, como sei lá, quer dizer, sabe, então, né, entre outros;
6. **fragmentos lexicais**, como palavras iniciadas e não concluídas.

Segundo o autor, disfluências causadas por fenômenos prosódicos, fragmentos lexicais ou expressões hesitativas, normalmente ocorrem “como um sintoma de busca de um elemento funcional ou um item lexical, mas por vezes realiza-se com esses elementos” (2006, p.50). As disfluências podem acumular-se em um enunciado.

Marchuschi (2006) também concorda que as disfluências são fenômenos intrínsecos e típicos da oralidade. Segundo ele, elas atuam como uma estratégia do falante para ganhar mais tempo durante o discurso, e são influenciados por pressões situacionais. São “os procedimentos adotados pelos falantes para resolverem os problemas que surgem devido ao processamento *on-line* de formas e conteúdos” (MARCHUSCHI, 2006, p. 49).

O autor diz ainda que elas não são aleatórias no discurso, constituindo rupturas na fala em lugares sintáticos e prosodicamente desmotivados. Marchuschi também explica que as hesitações podem apresentar motivações discursivas, preservando a fluência, uma vez que mesmo com hesitações a fala pode continuar fluente. Sendo assim, o autor propõe duas funções para as hesitações ou disfluências: 1) indicação da orientação/reorientação de seleções sintagmáticas e 2) atividade de busca e confirmação das seleções lexicais.

Sobre a primeira função, Marcuschi diz que ela é a mais comum, pois está relacionada ao planejamento linguístico. Seria “um indicador da busca de um item lexical com a antecipação de um elemento que lhe convém formalmente” (MARCUSCHI, 2006, p.69). De acordo com o autor, a nossa competência linguística nos permite saber quando usar uma determinada regra. Por exemplo, sabemos que certos nomes femininos têm artigos no feminino. No entanto, ele diz que essa capacidade não garante o uso, já que há outros tipos de seleções, “um item lexical determina por antecipação seleções específicas de outros itens” (MARCUSCHI, 2006, p.69). Segundo autor, a questão seria saber se o que temos, na hesitação, seria um indício de correção ou um caso em que não encontramos o item que buscávamos. Nesse caso, teríamos uma tentativa de preenchimento. Porém, ele diz que “na fala quase sempre executamos construções curtas, sequenciadas e com alto índice de pré-fabricação. Por isso não hesitamos tanto quanto seria de se esperar” (MARCUSCHI, 2006, p.70).

Em relação à segunda função, Marcuschi diz que ela está relacionada à frequência do uso e às “restrições derivadas do contexto e da estrutura da língua” (MARCUSCHI, 2006, p.70).

Sendo assim, o autor acha difícil pensar nas hesitações apenas como descontinuadoras da fala, pois o texto falado pode funcionar como uma pista para observação das estratégias de formulação, bem como das atividades verbais em seu aspecto formal, cognitivo e interacional.

Merlo e Barbosa (2012) também consideram as hesitações como marcas linguísticas que têm como função textual desacelerar o texto falado e, como função psicolinguística, dar tempo para que o falante decida sobre como a informação será verbalizada. Para eles, a desaceleração não significa uma perda na fluência, e sim um recurso que ajuda a manter a fluência, uma vez que auxiliam o falante a produzir enunciados mais precisos. Segundo os autores esse fenômeno ocorre com mais frequência em situações comunicativas menos familiares e automatizadas do informante.

A seguir apresentamos uma descrição dos tipos de disfluências tratados pela literatura.

1.4 Os tipos de disfluências

Há vários tipos de disfluências. Dentre elas, temos a inserção, que se caracteriza pela introdução de um determinado material linguístico no discurso. Ela ocorre a partir de uma pausa temporária no discurso pelo falante. O intuito da inserção, segundo Koch (2009), é: a) introdução de explicações ou justificativas; b) fazer menção a um conhecimento prévio que, normalmente, é pré-requisito para o entendimento do assunto; c) fazer exemplificações; d) introduzir comentários metaformativos; e) formular questões retóricas; f) introduzir comentários de caráter jocoso; g) servir como apoio para a argumentação; e h) fazer ressalvas.

Marcuschi (2006) caracteriza essa disfluência como “marcadores discursivos acumulados”. Segundo ele, e diferentemente de Koch (2009), são os mais difíceis de serem localizados no discurso, uma vez que podem se confundir com outras manifestações de hesitação. Para o autor, as inserções são manifestações que se acumulam em certos momentos do enunciado, como *né, sei lá, assim, sabe*, entre outros.

Merlo (2012), referindo-se a Chafe (1980), diz que as disfluências acontecem porque o texto falado não está pronto, é um ato de criação que se apresenta pouco familiar ao sujeito. Elas caracterizam-se por serem fones de preenchimento, variando de acordo com a língua (MERLO, 2012). Conforme a autora, em português brasileiro, a inserção é representada, comumente, por vogais semiabertas [ɛ] e pela consoante nasal [m].

Em nossos dados, consideramos fenômenos de inserção, além da vogal semiaberta, palavras como *né, então, aí, tão, sabe, assim*, entre outras. Tomamos essa atitude porque, em nossos dados, consideramos essas palavras assim como Scliar-Cabral, Martim e Chiari (1981) como recursos fáticos, utilizadas como “retroalimentação do interlocutor”, a fim de “manter e tomar o turno”, o que evidencia as diferentes relações do sujeito com o enunciado.

Outro tipo de disfluência é a reformulação que também pode ser chamada de correção. Koch (2009) divide a reformulação em dois tipos: a reformulação retórica e a reformulação saneadora. A primeira ocorre através de repetições e parafraseamentos. A sua função principal é a de reforçar a argumentação. Estas também são consideradas como uma estratégia de construção do texto falado. Um exemplo de reformulação retórica por meio de repetição seria: “eu lembrou qui eu era piquininhu *tipu tipu* eu tinha uns novi anus” (Exemplo nosso). Nesse exemplo, temos a repetição da palavra “tipu”.

Diferentemente da inserção, de macrofunção cognitiva, a reformulação retórica tem característica interacional. Esse recurso tem a função de facilitar “a compreensão através da desaceleração do ritmo da fala, dando ao(s) parceiro(s) tempo maior para o processamento do que vai ser dito” (KOCH, 2009, p. 89). A reformulação retórica também pode apresentar a função cognitiva.

A reformulação saneadora, por sua vez, pode acontecer em forma de correções, de repetições ou paráfrases saneadoras. No caso das correções, o falante necessita solucionar problemas encontrados em um segmento, após a sua materialização, seja o enunciado produzido por ele ou pelo interlocutor. As repetições ou paráfrases, por sua vez, são provocadas pelo interlocutor.

Fávero *et al* (2006) falam em dois tipos de correção: as de infirmação e as de retificação. Na primeira, o segmento anterior é anulado, e na segunda, ele é parcialmente corrigido. Além dessas, os autores também falam em 1) correções fonético-fonológicas, referentes à correção da pronúncia ou articulação, como por exemplo, pranta > planta; 2) correção lexical, que ocorre quando a palavra pronunciada não é a pretendida e, 3) correção morfossintática, quando há uma má formação da frase por meio da regência e concordância.

Para os autores, as correções também são um produto do planejamento local. Não são ocasionais nem aleatórias, uma vez que o falante procura a palavra mais adequada, que o leve a uma intercompreensão.

Ferreira e Bailey (2004) chamam as correções de *reparo* e os trechos com informações inadequadas de *reparandum*. Em suas análises, os autores observaram que as sentenças

corrigidas apresentam características acústicas diferentes das que contém a informação errada. Delfino (2010) segue a mesma nomenclatura dos autores. De acordo com ele, as disfluências de reparo são os trechos de frases em que o falante, ao perceber que passou uma informação errada, se autocorrige. Os resultados de Delfino mostraram que as disfluências de reparo estão relacionadas tanto com a estrutura sintática e semântica quanto com a prosódica.

Ao citar Heike (1981), Moniz (2006) diz que o autor considera que o fato do falante reconhecer o erro e corrigi-lo caracteriza eventos fluentes na língua. Segundo Moniz, a autocorreção permite que a mensagem recebida não seja prejudicada ao ser recebida pelo interlocutor.

Segundo Levelt (1989)¹⁰, a correção existe porque o falante monitora o que está dizendo e como está pronunciando. É ele quem inicia o próprio reparo. Em casos como esse, não há necessidade do *feedback* do interlocutor, uma vez que as pessoas descobrem os seus erros sozinhas. Conforme o autor, quando a fala é interrompida, os falantes usam recursos para sinalizar ao ouvinte que houve um problema, como por exemplo, uma pausa repentina ou o uso de recursos de edição, como *er, rather, no, and I mean*, no caso do Inglês.

Os reparos são construídos de uma forma altamente regular e estão sujeitos a regras linguísticas e conversacionais estritas. Não é necessário que esteja presente no reparo partes do enunciado original. Levelt (1989) também diz que o contexto do discurso é um determinante para o tipo de falha que o falante tentará prevenir, interceptar ou corrigir. Prosodicamente, ao citar Cutler (1983), o autor explica que as correções podem ou não ser marcadas. Dessa maneira, o *reparo* será prosodicamente não marcado quando o falante iniciar a correção com a mesma variação melódica do item com problema. Além disso, a amplitude e a duração relativa imitam o enunciado problemático. Por outro lado, o reparo será prosodicamente marcado quando a prosódia do item que deverá ser corrigido divergir substancialmente da sua correção. Cutler (1983) observou que os reparos dos erros fonéticos são, como regra, não marcados.

A probabilidade do reparo ocorrer depende de quão perturbador será o erro para o ouvinte. De acordo com Levelt (1989) a correção é mais frequente em final de frase.

A repetição, por sua vez, refere-se a qualquer repetição não intencional de um segmento. A sucessividade desse segmento pode ocorrer em palavras isoladas ou em sequência, e em parte de palavras, como fone ou sílaba. A repetição também se caracteriza

¹⁰ É importante salientar que Levelt (1989) denomina todos os tipos de disfluências como autocorreção. O autor parte do princípio de que esses fenômenos seguem regras linguísticas específicas que regulam a estruturação dos constituintes, a fim de que a mensagem chegue completa ao ouvinte.

por ser uma estratégia de estruturação do texto falado, e está presente constantemente na conversação do cotidiano.

Para Koch (2009), as repetições são essenciais, seja em situações formais ou rituais, seja em situações cotidianas, como, por exemplo, nas formas de agradecimento. Elas caracterizam-se por ser um recurso utilizado com o propósito de ganhar mais tempo durante o planejamento da fala, retomar a enunciação, garantindo assim a sua continuidade.

Entre os tipos de repetições citadas por Koch (2009), tem-se a que ela denomina de “autorrepetição”. Esse tipo de repetição é produzida pelo mesmo falante, por exigências de ordem cognitivo-interacional. Pode ser direcionada ao próprio falante, ao interlocutor ou a ambos.

Quando a repetição está orientada ao falante, ela apresenta três funções: 1) ganhar tempo para o planejamento de fala, 2) garantir a posse de fala e 3) simplificar a produção discursiva. Koch (2009, p. 134-135) diz que elas ocorrem com mais frequência: a) após um falso começo ou como preenchedoras de pausas; b) fazendo ponte em uma interrupção, provocada, por exemplo, por um pedido de esclarecimento do interlocutor e c) com a finalidade de substituição ou correção da formulação inicial.

As concepções sobre as estratégias discursivas utilizadas no texto falado são importantes para o estudo da disfluência comum, pois confirmam a sua constituição na fala de todas as pessoas.

Marcuschi (2006) diz que as repetições podem ser de itens funcionais ou lexicais. Os elementos funcionais são todos os elementos linguísticos sem significação referencial (artigo, preposição, conjunção, pronomes). Como por exemplo, a repetição de *pra, do*, entre outras. De acordo com o autor, as palavras reduplicadas são, em sua maioria, monossílabos. Elas situam-se em posições sintáticas preferenciais, “no momento de construir um sintagma, como é o caso das preposições e dos artigos” (2006, p. 54). Os itens lexicais, por sua vez, não ocorrem com tanta frequência quanto os funcionais. São comuns em verbos de uma ou mais sílabas, mas também podem aparecer com substantivos, advérbios e adjetivos.

Marcuschi (2006) diz que a repetição contribui para:

- a) a organização discursiva;
- b) a monitoração da coerência textual;
- c) o favorecimento da coesão;
- d) a geração de sequências mais compreensíveis;
- e) a continuidade da organização tópica;

f) o auxílio nas atividades de interação.

Segundo o autor, esse conjunto de funções proporciona uma textualidade menos densa e com envolvimento interpessoal maior. Isso faz da repetição um processo essencial na textualização da língua falada. A repetição pode resultar de fatores de ordem cognitiva, textual, sintática ou interacional. Conforme Marcuschi (2006), em média, de cada cinco palavras, uma é repetida.

Já a pausa é um dos fenômenos prosódicos mais utilizados para a demarcação dos grupos entoacionais. De acordo com Cagliari (1992), a função das pausas é destacar os grupos tonais e sinalizar como os interlocutores devem interpretar o que o outro diz.

Segundo Cruttenden (1986), as pausas podem ser preenchidas ou não (nesse caso o silêncio). Moniz (2006) explica que as pausas preenchidas caracterizam-se por meio da vocalização de segmentos. Na maior parte das vezes, ela diz que as pausas preenchidas podem ser de vogais do inventário silábico de uma língua.

Veja os exemplos das possibilidades de sons de preenchimento para algumas línguas, a partir da adaptação da tabela de Clark e Fox Tree (2002, p. 92).

Quadro 1- Os tipos de sons preenchedores mais comuns em algumas línguas

Tipos de línguas	Sons de preenchimento	Referências
Alemão	äh, ähm	Fischer (1999)
Holandês	uh, um	Swerts (1998)
Sueco	eh, aːh, aːːh, m, mm, hmm, oːh, a, oːh	Allwood <i>et al.</i> (1990) and Eklund (1999)
Norueguês	e, e=, e==, eh, eh=, m, m=, m==, hm, mm, em=, øhø, aj	Svennevig (1999)
Espanhol	eh, em, este, pues	Brody (1987) and van der Vlugt (1987)
Francês	eu, euh, em, eh, oe, n, hein	Duez (1982, 1991, 1993)
Hebreu	eh, e-h, em, e-m, ah, a-m	Maschler (1997)
Japonês	eeto, etto, ano, anoo, uun, uunto, konoo, sonoo, jaa	Cook (1993), Emmett (1996, 1998) and Hinds (1975)

Fonte: Adaptado de Clark e Fox Tree (2002).

Acrescentamos a essas, as possibilidades de pausas preenchidas verificadas no Português Europeu por Mata (1999): *aa, aam, mm*. E, para o Português Brasileiro, *éh, ahn, ah* e *mm*. (MARCUSCHI, 2006).

De acordo com Clark e Fox Tree (2002), os sons preenchedores ajudam o falante a encontrar a palavra que deverá ser usada, a decidir o que dizer, bem como indicam quando os

falantes querem manter ou ceder o turno da fala. Além disso, também mostram que os falantes monitoram o seu plano de fala. Os autores explicam que se o falante percebe um atraso na fala, eles formulam onde e quando suspendê-la, produzindo uma pausa preenchida. O argumento dos autores é de que os preenchimentos são palavras convencionais que os falantes planejam, formulam e produzem, assim como outra palavra qualquer.

Cruttenden (1986) diz que as pausas podem ocorrer em três lugares do enunciado. Sendo eles:

- 1) Nas principais fronteiras dos constituintes maiores (principalmente entre orações e entre sujeito e predicado);
- 2) Antes de palavras de alto conteúdo lexical, dentro do sintagma nominal, verbal ou adverbial;
- 3) Depois da primeira palavra do grupo entoacional. Essa é a posição típica para outros ‘*errors of performance*’, por exemplo, correções de falso começo e repetições.

Conforme Cruttenden, as pausas do tipo 1 são tomadas como indicação de uma fronteira de grupo entoacional. As pausas do tipo 2 e 3 são, geralmente, assumidas como exemplos do fenômeno de hesitação, sendo que as do tipo 2 ocorrem como mecanismo para a reorganização do enunciado.

Nem todos os silêncios podem ser considerados pausas e nem todas as pausas, hesitações. Marcuschi (2006) divide o silêncio em interturno e intraturno. Segundo o autor, os silêncios interturno são uma manifestação discursiva e não pausas. Esse tipo de silêncio ocorre quando o falante espera a vez do outro para falar. O silêncio intraturno, por sua vez, é considerado como uma pausa, já que apresenta uma determinada duração e um padrão entoacional característico. Conforme Marcuschi (2006) esse tipo de silêncio será uma pausa se ocorrer principalmente em contextos sintáticos ou junções fonêmicas em que a pausa não é prevista. O autor diz ainda que os falantes menos fluentes são aqueles que têm dificuldade em manter o turno, com muitas pausas em sua fala. Consequentemente, os falantes com menos pausas apresentam um discurso mais concatenado.

Marcuschi (2006) também observou que as pausas ocorrem com mais frequência no início do enunciado. Para ele, esse fato deve-se à busca pelo foco, uma vez que a pausa tem a função de marcar o processamento momentâneo e problemático do texto falado. O autor chama esses sons preenchedores de *expressões hesitativas*.

Merlo e Barbosa (2012) também corroboram com Marcuschi em relação a não aleatoriedade das pausas, bem como a sua tendência à posição inicial devido ao foco

informativo do texto. No entanto, os autores consideram que não há uma tendência de maior esforço cognitivo para a localização das pausas e hesitações durante o enunciado, podendo ocorrer de maneira ininterrupta do começo ao fim do texto falado.

Conforme Merlo e Barbosa (2012, p. 12), as pausas silenciosas fluentes estão “localizadas preferencialmente em constituintes maiores”, “apresentando ampla variação de *pitch* antes e depois da pausa”.

Textualmente, a função das pausas, eles dizem, é a segmentação da produção falada em unidades semânticas, sintáticas e (ou) prosódicas. Do lado psicolinguístico os autores afirmam que esse fenômeno está vinculado ao processamento da informação que será transformado em fala. Isto é, “funções executivas, como por exemplo, recuperações de informações da memória de longo prazo, escolha de estratégias para organizá-las e monitoramento do desempenho” (MERLO; BARBOSA, 2012, p.12).

Merlo (2012) também constatou que as pausas demarcativas e hesitações não são supérfluas na língua falada, uma vez que se distribuem de forma cíclica, bem como não são aleatórias ou erradas, pois ajudam o falante a manter a fluência e a expressar sua intenção comunicativa. Assim, ela considera que a pausa demarcativa está relacionada ao planejamento conceitual e as hesitações, à formulação linguística.

Para O’Connell e Kowal (2010) os termos pausa preenchida e pausa silenciosa são inapropriados. De acordo com eles, não há nada entre as pausas que possa contrastá-las. Isso porque, eles acreditam que as chamadas pausas de preenchimento não preenchem um vazio. Assim, adotam o termo preenchedores em substituição à pausa preenchida. Pois, para eles, a função da pausa silenciosa é a de demarcação, enquanto que essa não é a função dos preenchedores.

O alongamento, também chamado de prolongamento, é um tipo de disfluência que se caracteriza, segundo Merlo (2012), por um aumento, seja ele em início ou final de palavra com relação à duração de um fone. Segundo a autora, esse evento não tem função de indicar proeminência ou fronteira prosódica. Assim como as pausas, nem todos os alongamentos são um tipo de hesitação, uma vez que alguns caracterizam uma coesão rítmica, ou seja, são um tipo de estratégia prosódica de coordenação rítmica, como por exemplo, a formação de listas.

Segundo Marcuschi (2006), os alongamentos também podem vir acompanhados de uma elevação do tom, funcionando como ênfase. Caso ocorram no interior de uma palavra, diz o autor, eles recaem nas sílabas tônicas, podendo ser coesivos ou sintáticos. A vogal

alongada será hesitativa quando aparecer, principalmente, em final de palavra, sobretudo em monossílabos ou em sílabas finais átonas, como *um.*, *com.*, entre outros.

Por sua vez, o fragmento lexical, de acordo com Marcuschi (2006) evidencia-se pelo aparecimento de um item lexical duvidoso ou de difícil acesso ocorre durante a formulação textual. Seria a produção incompleta de uma palavra, como por exemplo, no enunciado “essa ul/ a última eu não lembro” (MARCUSCHI, 2006, p. 55).

Nakatani e Hirschiberg (1994) consideram os fragmentos como palavras monossilábicas, com produção incompleta. De acordo com elas, são fenômenos associados a efeitos de glotalização. As autoras apresentam quatro tipos de fragmentos: 1) fragmentos de palavras completadas com mesmo material sintático; 2) fragmentos de palavras que se completam com outras do mesmo constituinte sintático; 3) fragmentos de palavras que são abandonados ou introduzidos com um novo material sintático e, 4) fragmentos de palavras repetidas.

Há também o que se denomina de falso começo. Duez (1997) define o falso começo¹¹ como um enunciado que é interrompido e substituído por outro que o completará. Diferentemente dos fragmentos lexicais, o falso começo caracteriza-se pelo abandono do enunciado por meio informante¹².

Souza e Silva e Crescitelli (2006), por sua vez, denominam o falso começo como interrupção. Para elas, o falso começo representa diferentes estratégias de recuperação do texto falado, que pode ocorrer por meio da repetição, correção ou paráfrase. De acordo com as autoras, os enunciados interrompidos são resultado de como a língua falada insere-se no tempo, ou seja, indicando a simultaneidade do planejamento e da materialização verbal. Além disso, elas acreditam que as interrupções são indicativos do caráter reflexivo da linguagem, mostram “a possibilidade que esta [a língua] tem de poder olhar para si mesma, de se voltar sobre aquilo que acabou de ser dito” (SOUZA E SILVA; CRESCITELLI, 2006, p. 86).

Outro tipo de fenômeno que interfere na produção de fala, e por isso caracteriza-se como um tipo de disfluência é o aumento da velocidade (aceleração). De acordo com Levelt (1989), a alta velocidade na produção leva à redução e à assimilação. Conforme o autor, o falante pode aumentar sua velocidade de fala através da produção de mensagens curtas, ou por meio da redução dos alomorfes. Tudo isso é planejado acima do nível fonológico. O falante

¹¹ Neste trabalho consideramos um falso começo quando o falante inicia a produção de uma palavra, mas não consegue terminá-la, deixando-a incompleta no discurso.

¹² Tendo em vista a caracterização trazida por Nakatani e Hirschiberg (1994), o que denominamos de falso início entraria no item três da divisão feita pelas autoras. No entanto, para fins de análise adotamos como nomenclatura o falso começo.

também pode ganhar velocidade reduzindo pequenas palavras átonas, como pronomes e preposições.

O bloqueio, segundo Merlo (2006), é um tipo de disfluência que diz respeito a uma posição articulatória fixa, que leva à obstrução temporária de um enunciado. Neste trabalho, caracterizamos os bloqueios como uma obstrução de um fone, frequentemente uma oclusiva que ocorre, normalmente, em início de palavra. Na maior parte das vezes os bloqueios vêm antecedidos de uma pausa.

Tendo em vista esses eventos, Marcuschi (2006) resume a hesitação em: pausas preenchidas e não preenchidas, repetições hesitativas e falsos inícios. As pausas preenchidas caracterizam então as expressões hesitativas e os alongamentos que não recaem em sílabas tônicas. Normalmente são seguidos de pausas breves. As pausas não preenchidas, por sua vez, são os silêncios prolongados. Eles ocorrem como rupturas em lugares não previstos pela sintaxe. As repetições hesitativas dizem respeito à recorrência de itens lexicais, funcionais e de marcadores discursivos acumulados, semanticamente não significativos. Por fim, os falsos inícios são todos os inícios de unidades sintáticas que apresentam algum problema de realização e, por isso, são refeitos ou retomados.

A fim de verificar esses eventos na oralidade, Marcuschi analisou onze textos do projeto NURC (Norma Urbana Culta). Por meio da análise, o autor encontrou que a expressão hesitativa mais utilizada no Português do Brasil (PB) é *eh* e a classe de palavras com menor incidência de hesitação é a do adjetivo. O autor também observou que 41,2% de hesitações são de itens funcionais. Esse dado justifica o fato das hesitações serem momentos de planejamento *on-line* que interferem no processamento da fala.

Outra questão levantada na análise do autor é a de que as hesitações incidem, na maioria das vezes, em artigos, preposições, conjunções e pronomes. Para Marcuschi isso mostra que esses elementos representam momentos críticos na construção sintagmática. Entre esses fenômenos, o autor também observou que as hesitações em artigos são mais frequentes do que em preposições. De acordo com ele, “isso sugere que é mais fácil encontrar a preposição certa (em caso de dúvida) do que o artigo, já que encontrar o artigo significa ter decidido que nome foi escolhido” (MARCUSCHI, 2006, p. 60).

Resumindo, o autor conclui que:

- i. os itens funcionais são as formas linguísticas mais comuns para constituir as hesitações;
- ii. a expressão hesitativa de excelência do PB é *éh*;
- iii. as repetições são mais frequentes do que as hesitações;

iv.o predomínio dos itens funcionais na construção de hesitações mostra que eles são “um indicador de planejamento sintático e cognitivo e não uma estratégia de formulação textual” (MARCUSCHI, 2006, p. 61).

Marcuschi (2006) também chegou à conclusão de que a posição mais comum para que as hesitações ocorram seria no início da construção dos sintagmas. De acordo com ele, essa frequência indica a dificuldade de construção dos constituintes, o que atribui a relação da hesitação ao processamento linguístico.

A partir dessas evidências o autor conclui que as hesitações interferem na enunciação discursiva ao refletirem condicionamentos pragmáticos e também nas atividades cognitivas, quando interferem no processo da compreensão. Mesmo assim, elas não prejudicam a gramaticalidade do enunciado, pois quando retiradas do texto, observa-se que elas não apresentam um papel sintático.

Para Marcuschi (2006) as hesitações não apresentam funções sistemáticas no plano da formulação textual, diferentemente das estratégias de construção do texto falado (repetição, paráfrase, correção, parentização, tematização e rematização). Para ele, isso não significa que as hesitações sejam apenas uma disfunção da fala. Mas sim, ele diz “que seu papel é muito mais o de sugerir os sintomas de um processamento em curso do que o de propor alternativas de formulação textual-discursiva” (MARCUSCHI, 2006, p. 67). Portanto, ele acredita que a hesitação não é uma atividade formulativa, mas um tipo de índice problemático da formulação.

Ainda conforme o autor, é difícil dizer se as hesitações são uma propriedade da língua ou do falante, uma vez que elas são irregulares entre os falantes. Uns falam mais pausadamente e outros, mais rapidamente. “Há os que gaguejam muito e os que são controlados e não gaguejam quase nada” (MARCUSCHI, 2006, p. 70). Por isso, a necessidade em distinguir o que cabe ao falante e o que cabe ao processo de enunciação.

Para o autor,

[...] dadas as funções que a hesitação desempenha no intercâmbio comunicativo, não é pacífica a ideia de que a hesitação seja uma simples descontinuidade da fala. Pois o texto falado, em sua materialidade, pode ser visto como uma pista relevante para a observação tanto das estratégias de formulação como das atividades verbais nos aspectos formais, interacionais e cognitivos (MARCUSCHI, 2006, p.70).

As disfluências podem ser intencionais ou não. Quando falamos na sua intencionalidade, levamos em consideração o que Jakobson (1995) denomina de caráter fático

da linguagem, que tem como função principal testar o canal, prolongar ou interromper a conversação. Portanto, são significativas no discurso. Já as estratégias discursivas (no nosso caso disfluência comum) não intencionais (marginais) refletem as dificuldades de planejamento. Quando elas ocorrem, não há nenhuma alteração significativa nos elementos disfluentes.

Koch (2009) afirma que reformulação pode ser cognitivamente controlada. A correção também não é uma estratégia discursiva que o falante consegue controlar. Ela simplesmente acontece durante a enunciação. Na maior parte das vezes, o falante percebe que cometeu um erro. Nesses casos, caso ele ache necessário para compreensão da enunciação, corrige o enunciado. No entanto, ele não consegue controlar o erro antes da sua produção. Esse fato pode estar relacionado a um erro de programação antes da articulação da produção das sílabas (CAGLIARI, 2007, p. 141).

Há momentos em que o falante necessita criar situações que o ajudem a dar continuidade ao discurso e, por isso, utiliza-se de alguns recursos que podem ser intencionais, como a variação da entoação, por exemplo, ou não intencionais, como uma inserção alongada, uma correção, entre outros. É uma característica inerente a ele, que é rápida e, em grande parte das vezes, imperceptível aos ouvidos do interlocutor. Em outras palavras, são atitudes inconscientes, próprias da organização da fala.

Tendo em vista estas questões, é certo que a fala gaga não é um caso estritamente patológico, bem como a fala com disfluência comum não é uma questão apenas cognitiva. Indivíduos gagos e não gagos criam estratégias cognitivas durante a construção da fala, bem como questões situacionais influenciam determinado tipo de ocorrência. O item 1.5, abaixo, mostra essas questões a partir de diferentes pontos de vista.

1.5 GAGUEIRA: uma visão geral

1.5.1 Os primeiros relatos

“O homem corajoso não é quem não sente medo, mas quem vence esse medo”.
Nelson Mandela

A disfemia, mais comumente conhecida como gagueira, é um distúrbio da fala que pode acometer qualquer indivíduo. De acordo com a ABRA (Associação Brasileira de Gagueira), estima-se que 70 milhões de pessoas no mundo apresentem essa disfunção, sendo no Brasil, aproximadamente dois milhões de gagos.

Os relatos sobre esse fenômeno coincidem com o surgimento da fala. No livro da Bíblia, Êxodo, escrito por volta de 1400 a.C, conta a história do povo de Israel e da vida de Moisés. Nele, há passagens traduzidas que indicam que o profeta tinha dificuldades com a fala. Nos trechos, Moisés diz ao senhor, ‘Eu não tenho o dom da palavra; nunca o tive [...] tenho a boca e a língua pesadas’ (BÍBLIA, Êxodo, 4, 10); ‘Os israelitas não me ouviram; como me ouvirá o faraó, a mim que não tenho o *dom da palavra?*’ (BÍBLIA, Êxodo, 6, 12 grifos nossos). Na Bíblia, o excerto grifado é traduzido como “sou incircunciso dos lábios”. Conjectura-se que as indicações “boca e língua pesadas” e “incircunciso dos lábios” refiram-se à dificuldade de Moisés com a fala, que poderia estar relacionada com a gagueira. Posteriormente, por volta de 300 a.C, também há relatos de que Aristóteles apresentava uma desordem da fala. No entanto, agora já não temos suposições, mas sim fatos, “Era um pouco gago e tinha uma pronúncia defeituosa das palavras com a letra S, o que era motivo de zombaria por parte de seus conhecidos” (MARTINS, p. 7, s/d). Para Aristóteles, a gagueira era resultado de uma língua fraca. Relatos parecidos também eram dispensados aos filósofos gregos Hipócrates (460-370 a.C), Demóstenes (384-322 a.C) e Galeno (129-199 a.C). Hipócrates e Galeno nessa época já estudavam a produção da fala. O primeiro ainda apresentava noções muito vagas. No entanto, definia o fenômeno como um distúrbio da fala. O segundo, por sua vez, já possuía conhecimentos precisos sobre a sua fisiologia. O filósofo acreditava que a gagueira era resultado de línguas curtas. De acordo com Gomes-Kelly (2002), a fim de evitá-la ou suavizá-la, Demóstenes colocava pedrinhas na boca antes de discursar.

Isaac Newton (1643-1727), Charles Darwin (1809-1882), Machado de Assis (1839-1908), Winston Churchill (1874-1985), Nelson Gonçalves (1919-1998), Samuel L. Jackson (1948), Bruce Willis (1955) e Murilo Benício (1971), são outros exemplos, entre tantos gagos que se tornaram famosos.

De acordo com Büchel e Sommer (2004), as teorias sobre a origem da gagueira sofreram mudanças com o passar dos anos. Segundo os autores, na Grécia Antiga, a gagueira era atribuída à aspereza ou à secura da língua e ao seu comprimento. A crença na etiologia da língua como causadora da gagueira seguiu, segundo Gomes-Kelly (2002), até o século XVII. No século XVIII, a gagueira passa a ser vista como um problema relacionado à aprendizagem, às emoções e aos processos mentais.

No século XIX, o distúrbio também tem relação com a configuração do aparelho fonador. Nesse caso, o tratamento consistia em uma cirurgia de remodelação da boca e da

garganta. Método parecido com as incisões feitas na língua, no século XVII, para a extração de uma parte dela (Gomes-Kelly, 2002). Também há relatos de tratamentos com pesos presos à língua e técnicas de relaxamento.

Ainda nessa época, José Leite de Vasconcelos (1858-1941), médico e linguista português, escreve o livro *A evolução da linguagem: ensaio anthropologico* (1886). Na obra, o autor refere-se aos distúrbios da fala da seguinte forma:

Pessoas, que normalmente tem dificuldade em se exprimir, fallão desembaraçadas e expeditas quando experimentão uma emoção qualquer (entusiasmo, ira, etc.): nota-se esse phenomeno a cada passo [...] Chama-se em physiologia lucta vocal ao esforço dos agentes inspiradores para conservarem o ar no pulmão em opposição ao dos agentes expiradores para o expulsarem [...] (VASCONCELOS, 1886, p.11).

A citação nos leva a crer que Vasconcelos sugeria que os distúrbios da fala, de uma maneira geral, estavam relacionados a problemas no trato vocal e a fatores psicológicos, sendo o último fator evidenciado um século mais tarde.

No século XX, é atribuída ao fenômeno uma desordem de natureza psicogênica, sendo tratado por meio de terapias comportamentais. No entanto, o tratamento falha, uma vez que não foram encontrados padrões psicológicos que pudessem estar relacionados à gagueira. Büchel e Sommer (2004) dizem que também há relatos que consideram a gagueira como um fenômeno comportamental, resultado de estímulos externos prejudiciais, como por exemplo, a reação negativa dos pais diante da disfluência na criança. Os autores acreditam que mesmo que essa seja uma teoria falha, ela foi importante, pois contribuiu para o conhecimento dos sintomas secundários do fenômeno, como o nervosismo e a ansiedade.

1.5.2 O ponto de vista da literatura patológica e fonoaudiológica

No âmbito da literatura fonoaudiológica, Van Riper (1971) foi o pioneiro nos estudos sobre a gagueira. Segundo Azevedo (2000), o autor refere-se a ela como um quebra-cabeça, com peças faltando, complicado e de difícil resolução. Van Riper insere a gagueira no nível articulatorio. Para ele, ela ocorre quando há uma interrupção na fala devido a uma falha motora de um som, sílaba ou palavra (AZEVEDO, 2000). Portanto, um comportamento verbal que incide sobre o ritmo da fala.

Ao fazer uma revisão de literatura sobre os estudos desse fenômeno, Azevedo (2000) cita os trabalhos mais expoentes no campo da fonoaudiologia. São eles, o de Friedman (1986), Cunha e Gomes (1996) e Andrade (1999).

Friedman (1986) trata a gagueira sob o ponto de vista da psicologia social. Nesse caso, a gagueira deve-se à relação comunicativa da criança com os pais, ou seja, no processo comunicativo, a criança passa por um período de gagueira natural, que é repreendido inadequadamente pelos pais. Cunha e Gomes (1996) lançam mão de um olhar psicanalítico sobre o fenômeno, relacionando-o a um tipo de neurose, e não a um fenômeno da linguagem. Andrade (1999), por sua vez, trata a gagueira sob uma perspectiva biológica, levando em consideração questões psicológicas, linguísticas e sociais.

Azevedo (2000) critica a posição dos três autores em relação à atribuição de um lugar para a gagueira. De acordo com ela, os autores localizam a disfunção no sujeito, devido às tensões e hesitações, além de a considerarem ou como um problema de produção de fala, articulação, respiratória, ou como uma questão rítmica. Para ela, a gagueira é um fenômeno que se relaciona às condições de produção e ao espaço discursivo.

A fonoaudióloga desenvolve seu estudo tendo em vista o arcabouço teórico da Análise do discurso de linha francesa, desenvolvido por Eni Orlandi (1987). Dentro dessa perspectiva ela trata três aspectos do fenômeno: a sua origem, o seu lugar e a oposição entre fala e língua.

Em relação à origem, ela acredita que a gagueira insere-se no momento em que há um “deslocamento do sujeito falante em relação à sua fala e à do outro” (AZEVEDO, 2000, p. 120). De acordo com a autora, quando o adulto interpreta a fala da criança como gaga, ele a repreende com um discurso, na maioria das vezes, autoritário. Para ela, isso impediria a criança de identificar o erro no espaço discursivo. Isso faz com que a criança não seja capaz de identificar o erro e de movimentá-lo em seu discurso. Conforme Azevedo, o estranhamento familiar deveria ser significado e circunscrito discursivamente, a fim de que a criança consiga reconhecer o erro e, conseqüentemente, assemelhar-se à fala do outro.

Sobre o lugar da gagueira, a autora diz que os sujeitos gagos que participaram de sua análise identificaram as disfluências gagas como estando presentes neles próprios, na língua e no outro. Para ela, a gagueira não está em nenhum desses lugares, mas no espaço do discurso, em uma relação direta com a produção e a exterioridade. Em outras palavras durante o turno conversacional. Conforme a autora,

Neste espaço a gagueira acontece pelas formações imaginárias, onde se apresentam a relação de forças (o lugar a partir do qual o sujeito fala constitui o seu dizer), a relação de sentido (intertextualidade) e a antecipação (capacidade de colocar-se no lugar do seu ouvinte). (AZEVEDO, 2002, p. 120-121).

Por fim, sobre o terceiro aspecto – oposição fala e língua – ela diz que há uma tensão natural no discurso entre a fala e a língua, devido ao caráter abstrato da última. Segundo a

autora, esse fato leva à possibilidade do erro, da falha e da gagueira. Seguindo a ótica discursiva, Azevedo diz que o discurso só faz sentido no dizer do outro, pois é ele quem o reorganiza. Fato que gera tensão na fala de quem se pronuncia. Nesse caso, o gago, criando, então, a oposição entre língua e fala. Sendo assim, o sujeito, “não compreendendo esta dinâmica fundante [...] se aliena na língua e, nas tentativas de falar bem o ‘p’, o ‘f’, o ‘m’, a fala se perde, fica à deriva” (AZEVEDO, 2000, p. 121). O sujeito gago é então silenciado por meio das condições de produção, do autoritarismo do seu locutor, da tensão entre língua e fala. A gagueira é para Azevedo um lugar de subjetivação discursiva, em que o sujeito se vê alienado e silenciado em decorrência dos seus deslizes.

Azevedo (2002) constrói sua dissertação com o intuito de compreender a gagueira a partir de um novo ponto de vista: o do funcionamento da linguagem. A intenção da autora é contribuir com o tratamento clínico dos gagos, a partir de sua inclusão como sujeito da linguagem. Essa dissertação representa um ganho para os estudos fonoaudiólogos e para os linguistas, uma vez que é difícil encontrar trabalhos da literatura especializada que unam as duas vertentes ou que incluam de alguma forma a teoria linguística.

Atualmente, as propostas terapêuticas baseiam-se ou na Psicologia experimental, social ou psicanalista, na Filosofia fenomenológica ou, na Biologia (AZEVEDO, 2000). Não estamos dizendo com isso que a clínica não contribua para o tratamento do distúrbio. No entanto, ela deixa de lado o sujeito. Como diz Azevedo (2000, p. 8) “deixam escapar a linguagem e, com ela, excluem o sujeito, uma vez que ambos se encontram indissolivelmente atrelados, pois sujeito e linguagem se constituem mutuamente”.

A partir do mesmo princípio de Azevedo, temos o trabalho de Freire e Passos (2012). As autoras também levam em consideração o aspecto discursivo da gagueira, propondo o fenômeno como um efeito discursivo da interpretação do outro. Nesse sentido, Freire e Passos entendem, assim como Azevedo, que durante o processo de aquisição da linguagem, o discurso dos pais, na maioria das vezes, autoritário, pode impedir que a fala da criança se assemelhe à fala do outro, afetando assim o seu discurso. Em outras palavras, o autoritarismo dos pais, que ao corrigir a fala do filho, rotula-a como gaga “pode cristalizar o manifesto, aprisionando-o como patológico”¹³ (Freire e Passos, 2012, p.159). As evidências das fonoaudiólogas partem de entrevistas realizadas com as mães de pacientes gagos. As autoras reconhecem a insuficiência em se considerar apenas a interpretação dos pais nesse processo.

¹³ Entenda por manifesto, as pausas, repetições, reformulações e autocorreções que podem, normalmente, ser encontradas na fala da criança durante o processo de aquisição da linguagem. Fenômenos que são, na maioria das vezes, interpretados como eventos gagos por parte dos pais, nesse caso.

Para elas, também seria necessário levar em consideração questões relacionadas à sobredeterminação simbólica da linguagem e de seu funcionamento. No entanto, as autoras não trataram desse assunto no artigo. Elas apenas inferem que o discurso é sobredeterminado simbolicamente, o que descaracteriza questões referentes à origem da gagueira.

Os outros trabalhos, por sua vez, tratam mais da questão patológica do fenômeno, deixando de lado questões linguísticas. Entre eles, há diferentes perspectivas. Cruz e Frota (2011) os dividem do seguinte modo, os que estão relacionados 1) a fatores genéticos, como os trabalhos de Andrews *et al.* 1983; Mulligan, Anderson, Jones, Williams e Donaldson 2003; 2) aos que consideram a organização cerebral anômala para o discurso e linguagem, destacando-se Moore 1990, Braun *et al.* 1997, Fox *et al.* 2000, Inghan 2001, De Nil 2007; 3) à perturbação temporal dos articuladores durante a produção da fala, sendo eles, Zimmermann 1980, Kleinow e Smith 2000 e Max 2007; e, por fim, 4) aos que levam em consideração o tremor fisiológico dos músculos relacionados à produção de fala, que têm como principal expoente Zimmermann (1980).

Em 2006, Guitar fala em dois tipos de gagueira: a de desenvolvimento e a adquirida. A gagueira de desenvolvimento ocorre na infância, no período de aquisição da linguagem. A adquirida acontece mais tardiamente e subdivide-se em dois tipos: psicogênica e neurológica. A primeira está relacionada a fatores psicológicos, resultado de um conflito emocional, como por exemplo, a perda de uma pessoa querida, maus tratos, entre outros. A segunda surge em decorrência de alguma lesão cerebral. Esse tipo de gagueira é mais comum em adultos (ROCHA, 1999).

Podemos dizer que dentro da literatura especializada, os estudos se concentram em identificar os traços que ocasionam o aparecimento do fenômeno. Eles dividem-se entre aqueles que utilizam técnicas de neuroimagem, comparando a ativação cerebral em gogos e não gogos e, aqueles que correlacionam o aparecimento da gagueira a marcas genéticas. (MEIRA, s/d, ABRA). Conforme Meira, esses estudos ainda são inconclusivos, mas, ainda assim, podem levar a algumas evidências. Por exemplo, o primeiro grupo de pesquisadores (que olha para as técnicas de neuroimagem) acredita que a maior diferença entre a fala de gogos e não gogos está no planejamento e execução de áreas motoras de fala, que são ativadas com intensidade diferente, diz Meira. Quanto ao segundo grupo (que olham para as marcas genéticas), as pesquisas sugerem que o cromossomo 12 poderia ser responsável pela causa da gagueira. No entanto, os resultados ainda são inconclusivos. De acordo com Celeste (2010), em 2006, um ano após a descoberta da possibilidade genômica para a gagueira, foi publicada

uma pesquisa mostrando diferentes resultados dessa questão para homens e mulheres e a prevalência da gagueira em indivíduos do sexo masculino.

Também há trabalhos como o de Bloodstein (1993) que relacionam a gagueira ao ambiente. Nesse caso, o meio pode influenciar ou não o aparecimento dos eventos gagos. De acordo com Celeste (2010), a hipótese do autor partiu das informações dos próprios indivíduos gagos, que relatavam não apresentar gagueira quando estavam sozinhos. A partir disso, Bloodstein considera que a gagueira é uma interrupção do fluxo da fala, relacionada a situações de medo ou fuga.

1.5.3 A gagueira sob a ótica dos estudos linguísticos

Na primeira metade do século XX, a gagueira começa a ser analisada, segundo Cruz e Frota (2011), sobre diferentes perspectivas linguísticas, como por exemplo, a partir do ponto de vista da sintaxe, da fonética, da fonologia e da pragmática. De acordo com Arbisi-Kelm (2006) o grande inovador nos estudos sobre a gagueira foi S. F. Brown em 1938, analisando fatores linguísticos como, qualidade do segmento e comprimento da palavra. Desde então, segundo o autor, a maioria dos estudos preocupados com os fatores linguísticos da gagueira dividiram-se em duas áreas.

A primeira área está relacionada ao estudo do planejamento da sentença e tem como maiores representantes, Kleinow e Smith 2000; Kadi-Hanifi e Howell 1992; Yaruss 1999; Maner, *et al.*, 2000; Tornick e Bloodstein 1976; Melnick e Conture 2000. A segunda diz respeito ao estudo da disfluência gaga no nível da palavra, destacando-se os trabalhos de Brown 1945; Quarrington, *et al.*, 1962, Soderberg, 1962; Dworzynski, *et al.*, 2003 (ARBISI-KELM, 2006, p. 3).

Arbisi-Kelm (2006, p. 3-4) apresenta ainda estudos que se basearam em outros pontos de vista, como os trabalhos que levaram em consideração as evidências subsegmentais e a análise do acento lexical. O objetivo do primeiro grupo foi fazer uma correlação entre a taxa de disfluência e os VOTs (*voice onset time*). Os resultados mostraram que os gagos apresentaram VOTs mais longos do que os não gagos. Entre os representantes dessa área estão: Agnello 1975; Adams & Hayden 1976; Hillman & Gilbert 1977 e Healey & Ramig 1986. Por fim, para a análise do acento lexical destacam-se os trabalhos de Natke, *et al.*, 2002; Prins, *et al.*, 1991 e Brown 1938. Para Arbisi-Kelm esses estudos foram os mais bem sucedidos, pois descobriram correlações significativas entre as sílabas tônicas e a produção de sílabas disfluentes.

Mais especificamente, o maior interesse dos estudos linguísticos sobre a gagueira teve início na língua inglesa, na segunda metade do século XX, a partir dos estudos do funcionamento cerebral e do processamento da linguagem por indivíduos gogos.

As abordagens prosódicas da gagueira ganharam maior força nos anos 70 do século XX. O primeiro trabalho que tratou da fluência na língua inglesa, mostrando que o discurso normal está repleto de hesitações foi o de Goldman-Eisler (1968). Os trabalhos que o sucederam começaram seus estudos a partir da gagueira e não de uma disfluência.

No âmbito da fonologia prosódica, em 1986, segundo Frota e Cruz (2009), Bergmann, primeiramente, discute a capacidade dos gogos em produzir aspectos prosódicos. Depois, ao assumir que eles (gogos) são capazes de produzir corretamente os padrões acentuais, Bergmann questiona se essa característica elimina a hipótese do tratamento da gagueira como um distúrbio prosódico. A partir disso, segundo Frota e Cruz (2009), o autor, em suas análises, chega a três conclusões, a primeira de que os gogos não apresentam um padrão tonal reduzido, como era pensado, a segunda a de que eles têm dificuldade de executar o traço prosódico, mas o acentuam corretamente e, a terceira conclusão que o autor chega refere-se ao fato de a disfluência gaga estar mais presente nas sílabas tônicas, sendo os intervalos mais variáveis.

Segundo Bergmann (1986), a última conclusão pode ser explicada por meio de três razões. A primeira alude ao fato de que as sílabas tônicas são responsáveis por determinar os padrões temporais das frases. A segunda é a consideração de que sílabas tônicas e átonas são diferentes em relação aos movimentos articulatórios, uma vez que as primeiras exigem uma maior precisão. Por fim, a terceira razão é explicada pela expressão das emoções, que ocorre, principalmente, nas sílabas tônicas.

Posteriormente, em 1991, Prins *et al* verificam que a gagueira, na fala espontânea, está mais relacionada com as sílabas acentuadas.

Diferentemente dos autores, Hubbard (1998), ao contrapor a posição da gagueira em início de palavra e em sílabas acentuadas, em meio de palavras, verifica que o fenômeno ocorre com mais frequência em sílabas iniciais de palavras. O autor, segundo Cruz e Frota (2009), também conclui que a gagueira está associada ao nível da frase entoacional. Para isso, ele vê a necessidade da investigação da gagueira nas palavras prosódicas e nos sintagmas fonológicos e entoacionais, a fim de observar se um nível métrico específico é vulnerável ao discurso fluente.

Posteriormente, em 2006, Arbisi-Kelm afirma que as disfluências gagas são resultado do enfraquecimento do planejamento e produção subsequente do material prosódico proeminente. De acordo com Sônia Cruz e Frota (2009) o autor constata, da comparação com informantes sem distúrbio na fala, que os gagos apresentam frases intermediárias menores e menos acentos tonais do que os não gagos.

Ao comparar a fala espontânea com a leitura, Cruz Sônia e Frota (2009) dizem que Arbisi-Kelm constatou, para a fala espontânea, uma preferência ao *locus* entoacional em início das frases entoacionais. Por outro lado, para a leitura, o autor observou que as sílabas tônicas salientes atraem mais as disfluências nos gagos do que nos não gagos. A hipótese do autor é a de que as disfluências gagas “ocorrem antes do alvo concreto das mesmas, pelo que são consideradas antecipatórias. Conclui que, de facto, o sistema de monitorização desempenha um papel determinante na localização das disfluências” (FROTA; CRUZ, 2009, p. 5).

Em Portugal, os estudos que trataram desse fenômeno no âmbito da Fonologia Prosódica e Entoacional apareceram apenas recentemente. O estudo de Cruz (2009) foi um dos primeiros a explorar a gagueira do ponto de vista da prosódia. A autora traçou o padrão da gagueira, a partir da análise comparativa entre dois indivíduos adultos gagos e dois considerados fluentes. Para isso, ela analisou parâmetros como: a organização prosódica do discurso controlado e espontâneo da palavra prosódica/clítico ao sintagma entoacional; a estruturação entoacional do discurso (tipos e distribuição de acentos tonais e tons de fronteira) e, o lugar prosódico. Cruz verificou, para o português europeu (PE) que, em relação aos falantes fluentes, os gagos não apresentaram um padrão tonal reduzido. De acordo com ela, também não foram observadas diferenças marcantes no que diz respeito à frequência de ocorrência de eventos disfluentes em conversações e em leituras.

Há também a dissertação de mestrado de Condelipes (2010), “Produção de fala por indivíduos gagos”. Nesse trabalho a autora analisou a gagueira levando em consideração o ponto de vista linguístico e psicolinguístico, a partir de três tarefas 1) leitura em voz alta; 2) do resumo de leitura e, 3) da leitura de palavras isoladas por oito sujeitos gagos e oito não gagos. Por meio desse *corpus* o objetivo de Condelipes foi observar o nível de fluência dos dois grupos de informantes, além de verificar como os gagos planejam e verbalizam as informações que dispõem. Os resultados revelaram a maior presença de disfluências na primeira tarefa para os dois informantes. Além disso, a autora também observou que a maior

diferença entre os indivíduos gagos e não gagos está na frequência com que os eventos ocorrem na leitura.

Em 2007, Fabrice Hirch, em sua tese de doutorado “Le bégaiement: perturbation de l’organisation temporelle conséquences spectrale”, observou a organização temporal da fala produzida por falantes adultos gagos. O autor considera que a gagueira se manifesta na fase de transição entre os elementos consonânticos e vocálicos, e por isso observou os momentos intravocálicos por meio da análise do V.O.T. O estudo de Hirch revelou que a perturbação da coordenação motora ocorre tanto nas produções fluentes quanto disfluentes de falantes gagos. Segundo ele, isso revela que os efeitos da gagueira são latentes. Hirch também verificou que o tempo de reação da laringe é maior nos falantes gagos. O autor diz que isso se deve às contrações involuntárias da laringe nos falantes gagos, causando aberturas e fechamentos descontrolados do órgão.

Mais recentemente, Ivanna Didirkova, em “Origines articulatoires de discours hors-norms” (2016), analisou o comportamento articulatorio das produções gagas e não gagas de falantes com disfluência gaga e comum. O trabalho foi realizado por meio de dois procedimentos. O primeiro é um método de nasofibrosopia, que através de uma câmera inserida no nariz permite visualizar o comportamento da laringe e das cordas vocais. O segundo método foi feito por meio de uma articulografia eletromagnética (electromagnetic articulography - EMA), que permite medir os movimentos dos lábios, línguas e mandíbulas durante o processo de fala. Nos dois procedimentos participaram três informantes, sendo que para o primeiro foram dois falantes gagos e um de controle (não gago) e, no segundo, dois falantes normais e um gago. Para o primeiro procedimento, os resultados da pesquisa mostraram que a laringe de pessoas gagas comporta-se de forma diferente de pessoas não gagas, bem como há uma variação nos gestos articulatorios conforme o tipo de disfluência. Foram verificados movimentos inadequados de adução e abdução das cordas vocais e movimentos irregulares da laringe apenas em pessoas gagas. No entanto, para os movimentos não linguísticos, como engolir e tossir, na simulação de esforço e na leitura silenciosa, a laringe comportou-se da mesma forma para falantes gagos e não gagos. O segundo método, por sua vez, identificou que os articuladores (língua, lábios e mandíbulas), em falas normais e patológicas, são postos em movimento mesmo antes do início da fala. O trabalho de Didirkova é bastante relevante para os estudos atuais, pois mostra a importância do estudo das questões articulatorias em pessoas que apresentam gagueira. No entanto, assim como ela

afirma nesse artigo, é preciso analisar um número maior de informantes, a fim de se ter uma visão mais ampla do comportamento da laringe.

1.5.4 Os estudos sobre a gagueira a partir da perspectiva linguística no Brasil

No Brasil, até onde podemos apurar, os trabalhos, no campo linguístico, são ainda mais escassos, restringindo-se às questões clínicas do fenômeno e/ou a questões mais específicas e, mesmo assim, relacionados mais com a disfluência comum do que com a gaga. Por isso, iremos apresentar os dois tipos de trabalhos.

Dentro dos estudos linguísticos, Ester Scarpa é a pioneira no estudo das disfluências no país. Na maior parte dos seus trabalhos, a autora parte do estudo da fala infantil. Os trabalhos subsequentes, seja no campo da linguística, ou da fonoaudiologia, que partem da análise linguística fazem referência, de um modo ou de outro aos estudos da autora. Foi a partir do seu trabalho que os estudos sobre fluência e disfluência ganharam maior ênfase. A autora, no entanto, não trata da gagueira especificamente, voltando-se mais aos estudos sobre a disfluência comum em crianças¹⁴.

Entre os trabalhos de Ester Scarpa podemos citar os seguintes: *Sobre o Sujeito Fluente* (1995) e *(Ainda) sobre o sujeito fluente* (2006). Mais recentemente, a autora participou da organização do *Caderno de Estudos Linguísticos* (2012), com volume temático referente à (Dis)fluência. Essa edição é muito interessante, pois está inteiramente voltada aos estudos de disfluência, fluência e gagueira, contando com a participação de linguistas e fonoaudiólogos. Como exemplo de um desses artigos, temos o estudo de Scarpa e Fernandes-Svartman “A estrutura prosódica das disfluências em português brasileiro” que mostrou que a ocorrência das disfluências não é aleatória no discurso, ocupando espaços específicos no interior do domínio prosódico e o trabalho de Chacon e Villega “Hesitações na fala infantil: indícios da complexidade da língua”. No artigo, os autores procuram identificar o funcionamento das hesitações na fala das crianças, a partir da leitura de parlendas.

Ainda no volume, mas tratando especificamente da gagueira, temos os artigos de Barbosa e Merlo (2012), “Análise acústica da fala suavizada: estudo de caso em gagueira”, que será destacado posteriormente e também o de Carneiro e Scarpa (2012), “Singularidade nas manifestações de fala gagas”. Neste trabalho, as autoras procuram mostrar a singularidade dos sinais linguísticos na fala dos gogos. Segundo elas, esses eventos são frequentemente

¹⁴ A intenção deste capítulo é apresentar os estudos que trataram principalmente da gagueira. No entanto, não poderíamos deixar de destacar os trabalhos de Ester Scarpa, já que eles foram base para os estudos linguísticos sobre gagueira no país.

deixados de lado em vista de um diagnóstico clínico. De acordo com Carneiro e Scarpa, as falas gagas são singulares no que se refere aos sons estranhos à língua e diferentes no sentido de que esses sons e esses sujeitos, não são os mesmos. As autoras procuram mostrar que a busca pela homogeneização dessas falas, em detrimento de um diagnóstico ou etiologia, deixam de lado o que torna essas falas singulares.

Também há os trabalhos orientados por Scarpa, que seguem o mesmo viés linguístico e teórico. Entre eles, podemos citar a dissertação de Ramos (2008), “Hesitações e rupturas na fala infantil: as franjas da teoria e o lugar do sujeito na aquisição”. Neste trabalho, a fim de analisar a ocorrência de pausas e de repetições, Ramos estudou as hesitações na fala de uma criança. O objetivo da autora consistiu em mostrar que as disfluências são intrínsecas à linguagem. Segundo ela, esses fenômenos ocorrem devido à alternância entre a posição do sujeito com relação à língua. “Os momentos de hesitação são aqueles em que as dominâncias se alternam: dominância da língua, dominância da fala do outro, e a dominância do sujeito” (RAMOS, 2008, p. 7).

Depois, temos o trabalho de Delfino (2009) que estudou a prosódia das disfluências de reparo. O autor buscou comparar no que os aspectos prosódicos presentes na fala espontânea, com presença de disfluências de reparo, diferenciavam-se da fala considerada fluente. O autor caracteriza as disfluências de reparo como trechos de frases em que o falante, ao perceber que passou uma informação errada, se autocorrige. Segundo ele, as disfluências não podem ser consideradas simplesmente como erros que acontecem durante o processamento linguístico, uma vez que “podem servir como evidência do funcionamento dos sistemas de produção e compreensão da linguagem” (DELFINO, 2009, p.xiii).

No campo da prosódia, a partir da tese de Celeste (2010), “A prosódia na expressão de atitudes na fala de indivíduos com e sem gagueira”, Celeste e Reis (2011) estudam os aspectos prosódicos na expressão de atitude em falantes com gagueira, a fim de observar como esses informantes expressam a atitude na variação melódica da entoação da fala. As expressões analisadas foram: a neutra, a de certeza e a de dúvida. Para isso, os autores coletaram amostras de 12 informantes com a presença de distúrbio na fala e de 12 sem o distúrbio. Para a análise acústica, analisaram a frequência fundamental (F0) ao longo dos enunciados, nas seguintes posições: dentro das vogais da sílaba tônica saliente, da pré-tônica e da pós-tônica. No artigo, Celeste e Reis (2001) concluíram que as variações melódicas foram mais significativas na expressão de certeza e de dúvida nos falantes sem gagueira. Já os informantes com gagueira quase não apresentaram diferenciações nas expressões de atitude.

Em 2012, Merlo trata em sua tese de doutorado, “Dinâmica temporal de pausas e hesitações na fala semiespontânea”, da não aleatoriedade das pausas demarcativas e das hesitações. Segundo a autora, esses dois fenômenos não devem ser tratados como algo supérfluo ou até mesmo como erro na fala. Para ela, esses eventos “sinalizam as decisões do falante sobre o planejamento conceitual e a formulação linguística” (MERLO, 2012, p. 20), para as pausas demarcativas e para as hesitações, respectivamente. Além disso, Merlo também considera esses eventos como “fenômenos dinâmicos da língua falada”, que emergem a partir de necessidades específicas e de estilos de fala do informante.

No artigo “Análise acústica da fala suavizada: estudo de caso em gagueira” (2012), Barbosa e Merlo estudam um caso de gagueira a partir da suavização, método utilizado no tratamento da gagueira. De acordo com os autores, essa estratégia reduz a frequência da gagueira, porém, também altera a prosódia da fala. O trabalho consistiu na leitura, por um informante gago, de 21 frases veículo, iniciadas pelas plosivas [p, t, k]. Como resultado, os autores observaram que a diminuição das hesitações gaguejadas com a técnica aplicada sugere que a suavização é um modo hipoarticulado da fala.

Bohen (2009) também se interessa pelas questões linguísticas para dar conta da gagueira. A autora, atrelando à linguística a fonoaudiologia verifica as características das palavras gaguejadas. Em sua tese, ao analisar 1326 palavras, Bohem concluiu que 1) os bloqueios e as repetições são mais frequentes; 2) 97% da gagueira ocorreu em sílabas iniciais; 3) a tonicidade, bem como a faixa etária e o gênero não interferiram na característica das palavras gaguejadas; 4) ao comparar as hesitações normais com a gagueira, ela concluiu que as hesitações ocorrem entre as palavras, e a gagueira dentro delas; e, por fim, 5) concluiu que a regularidade da gagueira é a regularidade da linguagem, já que para entender a primeira é preciso compreender a segunda.

Como podemos ver, no Brasil, os estudos linguísticos sobre gagueira ainda são poucos. Como foi destacado, há um número pequeno de trabalhos que discutem especificamente a gagueira. A maior parte deles está voltada ao estudo da disfluência comum, visando à análise, por exemplo, das hesitações e autocorreções. Quando há trabalhos com uma abordagem linguística direcionados à gagueira, grande parte deles vem de fonoaudiólogos. Como vimos, se formos levar em consideração o tempo dos estudos sobre a gagueira, veremos que os estudos linguísticos são os mais antigos, datando da primeira metade do século XX. Arbisi-Kelm (2006) deixa evidente em sua tese o seu desapontamento diante da incredulidade de muitos em aceitar a gagueira como uma desordem na linguagem. Essas

afirmações mostram a importância e a necessidade de desenvolver trabalhos desse tipo, a fim de contribuímos com os estudos clínicos.

1.6 Caracterização do fenômeno

Esta seção tem por intuito apresentar algumas definições sobre a gagueira. A maior parte dos conceitos vem da literatura especializada. Ao levantar essas questões, nosso objetivo é compreender melhor o fenômeno, a fim de analisar os nossos dados de um modo mais adequado.

Segundo Spinelli (1983, p. 99), a gagueira caracteriza-se por fenômenos linguísticos como “[...] repetições de sílabas, palavras ou conjunto de palavras, *prolongamentos* de sons, *hesitações* e *bloqueios*”. Além dessas características, o autor também considera a desorganização dos relatos e os desvios de assunto, ocasionados, ou por situações que levam a uma instabilidade emocional ou até mesmo por um problema fisiológico.

Bohen acrescenta, definindo-a como

[...] uma ruptura na expressão oral de um indivíduo, caracterizada por desvios no fluxo, suavidade, ritmo, velocidade e/ou esforço com as quais as unidades fonológicas, lexicais, morfológicas e/ou sintáticas são faladas. Há interrupções significativas na sequência estabelecida de sílabas e palavras emitidas em um determinado tempo. (BOHEN, 2003, p. 41).

A literatura especializada diz que a gagueira tem início entre os três e sete anos de idade, sendo mais comum em pessoas do sexo masculino. Por acreditar ser uma desordem no desenvolvimento da fala, Valerie LaPorte (1997) diz que em metade dos casos, a gagueira começa gradualmente, desenvolvendo-se ao longo dos meses. Já na outra metade deles, começa repentinamente, em cerca de duas semanas. A autora propõe uma lista, para o inglês, de oito comportamentos comuns associados à gagueira, conhecidos, segundo ela, como *The eight Danger Signs of Stuttering*¹⁵. Veja a seguir:

1. Repetição de parte de palavras (sílabas), que normalmente ocorre na sílaba inicial, “*da-da-da-daddy*”;
2. Substituição do *schwa* ou por vogal fraca na repetição. Por exemplo, ao invés da criança dizer “*bay-bay-baby*”, ela diz “*buh-buh-baby*”;
3. Prolongamento de um som, que consiste em segurá-lo por mais tempo do que o necessário, como “*mmmmommy*”;

¹⁵ Os oito sinais de perigo da gagueira. [Tradução nossa]

4. Tremores. Esse evento caracteriza-se, segundo ela, por movimentos incontrolláveis da mandíbula, lábios ou língua;
5. Aumento da frequência (variação melódica) e intensidade durante uma repetição ou prolongamento de um som;
6. Esforço e tensão, dificuldade para produzir um som, que pode vir associado a movimentos faciais, dos lábios, dos membros, de troncos e piscar de olhos;
7. Momentos de ansiedade, que se evidenciam por uma expressão de medo ao se aproximar de palavras de difícil pronúncia;
8. Omissão/anulação. Refere-se a um número incomum de pausas ou pausas mais longas do que o normal, à omissão de certas palavras ou pronúncias de palavras completas, ao atraso de uma pronúncia até que se possa dizê-la sem dificuldades.

A autora, a partir da compilação de uma variedade de teorias e representações de comportamento das pessoas que gaguejam, também apresenta uma descrição dos estágios evolutivos da gagueira, juntamente com a faixa etária. Ela salienta a importância de não generalizar estes comportamentos, uma vez que nem todas as pessoas que gaguejam irão apresentá-los. Observe o quadro 1, a seguir:

Quadro 2 – Os estágios de desenvolvimento da disfluência gaga, adaptado de Valerie Laporte (1997). Disponível em <http://www.d.umn.edu/~cspiller/stutteringpage/onset.htm>

Idade	Tipo de disfluência	Características
Entre 2 e 6 anos (Estágio 1)	Disfluência normal	<ul style="list-style-type: none"> • Ocorre devido à dificuldade de aprendizagem da linguagem oral pela maioria das crianças; • As disfluências tendem a ser repetições de sílabas individuais, palavras inteiras, interjeições e correções; • Pausas breves são normais, pois se refere ao planejamento da fala pelas crianças; • Não há presença de tensão, frustração ou ansiedade; • Ocorre no período em que a criança está aprendendo a andar ou aprimorando as habilidades motoras; • Há períodos de fluência e disfluência; • Mudanças comportamentais também podem ocasionar esse tipo de disfluência.
Entre 2 e 6 anos (Estágio 2)	Disfluência intermediária	<ul style="list-style-type: none"> • As disfluências tendem a ser repetições e sons prolongados; • Mais do que duas disfluências juntas, por exemplo, <i>myyyyyyy-my-pup-pup-py</i>; • Há períodos de fluência e disfluência que vem e vão em ciclos; • A criança demonstra uma pequena conscientização e preocupação com sua disfluência. Não há frustração por falha na mensagem.
Entre 2 e 6 anos	Início da	<ul style="list-style-type: none"> • As disfluências ocorrem mais comumente no início de

(Estágio 3)	gagueira	<p>palavras ou frases;</p> <ul style="list-style-type: none"> • As crianças tendem a ser mais disfluentes quando estão ansiosas, tristes, ou sob algum tipo de tensão; • As repetições são normalmente parte de palavras em oposição à palavra inteira. Envolvem a vogal <i>schwa</i>; • A gagueira vem e vai em ciclos, podendo ou não ser desencadeada por momentos de <i>stress</i>; • A criança demonstra consciência da dificuldade da fala, além de frustração por não conseguir passar a mensagem.
Entre 2 e 6 anos (Estágio 4)	Gagueira intermediária	<ul style="list-style-type: none"> • Os tipos de disfluência incluem repetições, prolongamentos e bloqueio; • A gagueira começa a ser crônica, sem períodos de fluência; • Começa a aparecer comportamentos físicos secundários junto com a gagueira, como piscar de olhos, movimentos de membros, lábios, entre outros; • A gagueira começa a aumentar em períodos de ansiedade, tristeza ou tensão.
A partir de 14 anos e adultos (Estágio 5)	Gagueira avançada	<ul style="list-style-type: none"> • A gagueira é caracterizada por interrupções frequentes e perceptíveis; • O contato visual diminui e a pessoa usa truques para disfarçar os períodos de disfluência gaga; • A pessoa antecipa a gagueira, tem medo e evita falar; • A pessoa identifica-se como gaga, há conscientização, timidez e vergonha.

Fonte: Adaptado de (VALERIE LAPORTE, 1997).

Para Wingate (1964), a gagueira é uma ruptura da fluência verbal, que se evidencia por meio de prolongamentos e repetições involuntárias. O autor diz que ela também pode vir acompanhada de movimentos corporais e emoções negativas, como medo ou irritação.

Segundo Oliveira e Gargantini (2003), também estão associados à gagueira questões como os distúrbios respiratórios, que incluem inspirações ou expirações prolongadas, interrupção da respiração, soltura repentina do ar. Há ainda, segundo as autoras, os distúrbios fonatórios, tais como, as alterações na qualidade de voz, mudanças bruscas ou falta de variação no *pitch*, alteração da velocidade, que pode ser rápida ou lenta. E, por fim, questões relacionadas à articulação, como a não sincronização entre os músculos articulatórios, a anormalidade entre os movimentos mandibulares e a fonação, tremores labiais. Além disso, as autoras também acrescentam que os falantes gagos evitam certos sons e, com o intuito de substituí-los, produzem sequências sem sentido.

Do lado linguístico, Carneiro e Scarpa (2012, p. 156), baseadas em De Lemos¹⁶ (1992, 2001 e 2006), definem a gagueira como fenômenos

[...] observados ou sentidos na fala do gago, como pausas, repetições, bloqueios, interjeições, tidos como disfluências que perturbam a fluência verbal. Acontece assim uma manifestação de linguagem que, como tal, implica a língua, ou seja, está submetida ao seu funcionamento.

Celeste (2010), por sua vez, conceitua a gagueira como uma alteração prosódica, um “distúrbio da comunicação humana que afeta diretamente a fluência da fala, causando interrupções na cadeia segmentar devido a erros na programação motora temporal, com sucessivas tentativas de retomada da fluência” (CELESTE, 2010, p.34).

Uma pessoa falante de uma língua de ritmo acentual também pode apresentar gagueira quando começa a aprender a falar no ritmo acentual. Segundo Abercrombie (1965, 1967) e Cagliari, (2002, 2007) isso ocorre porque há evidências de que as crianças começam a falar utilizando o ritmo silábico e depois passam ao acentual. Isso revela o esforço da criança na tentativa de reestruturar a sua fala sobre o ritmo acentual.

O erro faz com que ela [a criança] repita sílabas até que consiga reestruturar o ritmo, que pode então ser silábico, e portanto mais familiar à criança, ou representar uma tentativa nova de pôr o restante em ritmo acentual, forçando a produção do trecho seguinte no ritmo novo (CAGLIARI, 2007, p. 141).

A gagueira também pode vir associada a questões emocionais. Acredita-se que há um aumento nos desvios de fala quando uma pessoa leva um choque emocional, se vê em situação de risco ou muito desconfortável, como vimos em alguns estudos . Porém, outros trabalhos consideram que relacionar o aumento da disfluência gaga a questões sentimentais, como situações emotivas, momentos de insegurança ou ansiedade, apesar de ser uma afirmação muito comum e popularmente difundida, é um argumento que deve ficar no passado. De acordo com Carneiro e Scarpa (2012), essas asserções devem ser tratadas simplesmente como algo imaginário, que nasceu devido a uma tentativa de homogeneização dos sujeitos, que levam ao surgimento

[...] de mitos, como o mito da ansiedade, o mito da evitação de situações de fala, o mito da evitação de palavras temidas, o mito de que o gago possui pouco vocabulário, dentre outros. São mitos, presentes em uma certa

¹⁶ A discussão apresentada na citação de Carneiro e Scarpa (2012) sobre a questão da gagueira estar submetida ao funcionamento da linguagem é muito mais complexa e aprofundada em De Lemos (1992, 2001, 2006). No entanto, não é objetivo deste trabalho destacar essas questões, e sim mostrar a visão linguística das autoras em relação à gagueira.

literatura e também no senso comum, tomados ora como causas, ora como consequências da gagueira [...] (CARNEIRO; SCARPA 2012, p. 156).

De acordo com as autoras, esses mitos são cada vez mais tratados como contraevidências para a definição da gagueira. Trata-se de um sujeito que está sob o efeito de um funcionamento linguístico-discursivo que, incontrolavelmente, coloca sua fala à deriva, em meio a uma imprevisibilidade. Essas contraevidências “põem falante e ouvinte à escuta dos movimentos da língua; deslocam do sentido para o corpo. Geram tensão no corpo do falante. Geram tensão no ouvinte” (CARNEIRO; SCARPA 2012, p. 156). Com os não gogos, isso não ocorre, uma vez que os sinais linguísticos não colocam a língua à deriva, à margem na sociedade. Dessa maneira, a gagueira seria, então, “o efeito da tensão entre o reconhecimento e o estranhamento dos movimentos da língua em uma fala”. (CARNEIRO E SCARPA, 2012, p. 157).

Do nosso ponto de vista, essa tensão ocorre por uma questão de empatia fonética, neste caso, o que denominamos de empatia negativa. Esse termo ainda não havia sido usado na linguística para tratar de questões relacionadas à disfluência gaga.

A *Empatheia*, palavra de origem grega que significa “paixão” ou ser “muito afetado”, foi objeto de estudo de diferentes áreas. No entanto, recebeu maior destaque no campo de estudo da Psicologia, sendo introduzido primeiramente por Lipps em 1909.

Dentro da Psicologia, esse fenômeno foi conceituado de diferentes maneiras ao longo dos anos, mas a que mais nos interessa é a que caracteriza esse fenômeno como “uma capacidade através da qual as pessoas compreendem umas às outras, sentiam e percebiam o que acontece com os outros, como se elas mesmas estivessem vivenciando as experiências alheias” (SAMPAIO et al, 2009, p. 214). Na mesma linha, Cameron (2011), diz que a empatia diz respeito ao modo como as pessoas se conectam com as outras, no sentido de compreendê-las, bem como entender os seus sentimentos. Ainda segundo a autora, em um momento de interlocução, a empatia pode aumentar ou diminuir ao longo da conversação. O que também nos leva a considerar a empatia como uma questão de relativização, uma vez que o desconforto causado na fala do outro é sentido de modo diferente pelos ouvintes. Nesse sentido, é preciso salientar que esse sentimento não acontece apenas quando ouvimos a fala de pessoas gagas, mas também a de pessoas não gagas. Por exemplo, uma pessoa que fala com uma velocidade de fala menor ou que hesita muito, causa um estranhamento no ouvinte. Portanto, a empatia também pode ser considerada como uma questão de gradiência.

Abercrombie (1965), bem como Cagliari (2012a) utilizaram esse termo para explicar o ritmo das sílabas. Segundo eles, espera-se da percepção da fala uma igualdade entre aquilo que se ouve e aquilo que é interiorizado pelo falante.

All rhythm, it seems likely, is ultimately rhythm of bodily movement. Language rhythm is thus something which belongs primarily to the speaker, rather than the hearer; something which arises out of the speaker's movements, and especially out of the muscular movements which produce the air-stream¹⁷ (ABERCROMBIE, 1965, p. 19).

Como diz Abercrombie (1965), o ritmo é um movimento corporal, que surge dos movimentos do falante, especialmente dos musculares. Para que a percepção da fala aconteça é necessário, explica Abercrombie, que o ouvinte e falante se identifiquem, pois durante esse processo o ouvinte reproduz em sua mente os mesmos mecanismos linguísticos que o seu interlocutor produziu durante a sua fala. No entanto, em uma fala gaga, há uma falha tanto na produção do interlocutor quanto na reprodução dela pelo seu interlocutor. Os movimentos não são sentidos da mesma forma, uma vez que o ouvinte não consegue reproduzir os movimentos articulatórios da mesma forma. Conforme Cagliari (2012a) não ocorre nessa interação entre gago e não gago o que ele chama de “comunicação fonética específica, padronizada no corpo e na mente” (CAGLIARI, 2012a, p. 35). Isso leva, então, a uma empatia negativa, a uma sensação de estranhamento entre falante e ouvinte, porque este sente dificuldade em pronunciar a fala do outro, gerando, conseqüentemente, o preconceito.

Aos 82 anos de idade Van Riper escreve:

Eu gaguejei todos os dias da minha vida. Acho que tenho uma gagueira incurável. Todo mundo tem seu próprio demônio pessoal e o meu é a gagueira. Depois que entendi que a gagueira é um problema e aprendi a lidar com ela (sem evitá-la, escondê-la ou lutar contra ela), meu demônio deixou de me deter. Tive uma gagueira grave com longos bloqueios acompanhados por contorções faciais e espasmos, que não apenas provocavam a rejeição de meus ouvintes, como também tornou minha comunicação quase impossível. (Fonte: IBF – Instituto Brasileiro de Fluência)

O depoimento do fonoaudiólogo descreve situações que são comuns no discurso da maioria das pessoas gagas. A maior queixa das pessoas gagas é o preconceito de quem os ouve, o estranhamento, que está bem claro quando Van Riper fala que os episódios disfluentes

¹⁷ Parece provável, em última análise, que todo ritmo é um movimento corporal. Ritmo de língua é algo que pertence primeiramente ao falante, ao invés do ouvinte; algo que surge dos movimentos do falante, e especialmente dos movimentos musculares que produzem a corrente de ar. [Tradução nossa].

causam uma sensação de rejeição no ouvinte. Infelizmente, o gago, tornou-se um personagem caricato, em canções, piadas populares, em vídeos do *youtube* e entre outros meios de comunicação. O desconcerto do ouvinte em relação à fala gaga ocorre por uma questão de empatia fonética negativa.

Os episódios empáticos, de acordo com Sampaio et al (2009) podem gerar diferentes tipos de afetos, como por exemplo, sentimentos de perturbação ou desconforto, piedade ou compaixão. Dentro dos estudos fonéticos, a empatia, ligada à gagueira, seria a capacidade de sentir biologicamente os mecanismos que fazem com que a pessoa fale.

[...] existen razones psicológicas para pensar que la generación de sonidos vocales de dos interactuantes estaría directamente relacionada con lo que sucede entre ellos a nivel psicológico durante la interacción. Dicho de otro modo: Desde un enfoque de la empatía como fenómeno interaccional, corporeizado, resulta factible esperar que en el experimentar el mundo como el otro (Bohart y Greenberg, 1997) de la empatía se diesen fenómenos observables a nivel vocal, es decir fonético. Por lo tanto, sería esperable encontrar un vínculo entre convergencia fonética y empatía. Al ser un fenómeno preconsciente, corporal y espontáneo, la convergencia fonética constituiría una herramienta para acercarse experiencialmente a la vivencia del otro¹⁸ (QUEZADA et al. 2012, p. 150).

Nesse sentido, o descontrolo do gago sobre a sua fala cria um fenómeno de falta de empatia fonética no interlocutor, devido à reprodução da fala do seu interlocutor, seja ela “errada” (disfluência gaga) ou não. Isso, portanto, leva à rejeição ao gago, gerando diferentes formas de preconceito. O não conhecimento da gagueira e os espasmos físicos atrelados a ela também levam à exclusão e ao distanciamento do gago, seja na escola, no ambiente ocupacional, nas conversas do dia a dia. Em fim, no convívio social.

Tudo o que é diferente, causa estranhamento, leva à exclusão. Diante disso, um dos queremos mostrar que os eventos sentidos na fala do gago são os mesmos de uma fala normal, com disfluência comum. De uma maneira geral, a diferença parece estar na frequência em que esses fenómenos ocorrem, sendo mais intensos nos portadores de gagueira. É possível, que a aceitação da fala gaga pelo ouvinte, ajude o seu portador a sentir-se mais seguro e, com isso, lidar melhor com a gagueira.

¹⁸ [...] existem razões psicológicas para pensar que a geração de sons vocais de dois falantes estaria diretamente relacionada com o que acontece com eles em nível psicológico durante a interação. Em outras palavras: a partir de um enfoque na empatia como um fenómeno interaccional, corporizado, é possível esperar que a experiência de mundo como o outro (Bohart y Greenberg, 1997) gere fenómenos observáveis a nível vocal, ou melhor, fonético. Portanto, seria esperado encontrar um vínculo entre a convergência fonética e a empatia. Por ser um fenómeno pré-consciente, corporal e espontâneo, a convergência fonética se constituiria em uma ferramenta para experiência a vivência do outro. [tradução nossa].

1.7 Sobre a articulação

And of course phonetics is absolutely essential to the student of linguistics. It is virtually impossible to do serious work in linguistics without a thorough knowledge of phonetics.
(CATFORD, 2001, p. 2)

A articulação também influencia a produção de disfluências na fala. Sendo assim, será apresentada uma introdução geral às questões fonéticas que envolvem a produção da fala. Para isso, abordamos questões referentes às etapas de produção da fala, como a aerodinâmica, a fonatória, a articulatória e a acústica. Seguido disso, apresentamos os outros componentes fonéticos da fala, os suprasegmentos, como por exemplo, a tessitura, a variação melódica, a duração, entre outros¹⁹. Vale salientar que o que será apresentado é resultado de uma visão geral, que procura abarcar mais especificamente as questões de interesse desta tese²⁰.

O ato de falar envolve a maior parte do corpo. De acordo com Catford (2001), quando falamos com alguém ocorre uma sequência de eventos. Segundo o autor, em detrimento da necessidade comunicativa, o falante conceptualiza um evento de modo particular, e em seguida codifica-o, de acordo com o que foi estabelecido pela gramática da língua. O código linguístico é então externalizado e apreendido pelo ouvinte por meio de uma série de eventos. Esses eventos iniciam-se com o falante e culminam com a decodificação e conceptualização do enunciado pelo ouvinte. O processo de codificação e decodificação, bem como a conceptualização não fazem parte do domínio fonético.

Catford chama esses eventos de fases da fala, dividindo-as em sete etapas. A primeira fase diz respeito à programação neurolinguística, que está relacionada à seleção, sequenciamento e tempo das etapas subsequentes. Nessa etapa, as ideias se juntam aos sons correspondentes do que se quer dizer em uma determinada ordem, a partir das regras da língua (CAGLIARI, 2007; MASSINI-CAGLIARI; CAGLIARI, 2008).

Em seguida, o processo é neuromuscular. É o momento da transmissão dos impulsos neurais e contração dos músculos individuais. É a preparação do corpo para dizer o que foi planejado. (MASSINI-CAGLIARI; CAGLIARI, 2008).

¹⁹ É comum utilizar o termo suprasegmento como sinônimo de prosódia. No entanto, alguns autores discordam dessa utilização. Por exemplo, Crystal (1969) define a prosódia como um conjunto de características fonológicas que têm uma relação variável com as palavras selecionadas, em oposição aos elementos (como por exemplo, os fonemas, o significado lexical) que têm uma identificação e relação direta com essas palavras. Já os suprasegmentos, segundo o autor, seriam indícios de um efeito vocal que abrange mais de um segmento no enunciado, como por exemplo, a variação melódica.

²⁰ Para entender mais sobre o assunto, consultar: Catford (2010, 1977); Cagliari (2007); Ladefoged e Johnson (2011) e Ladefoged (1995, 1983, 1975).

Como resultado da contração muscular ocorrida na etapa neuromuscular, a terceira fase consiste na posição e movimento de todos os órgãos do trato vocal, como por exemplo, a contração da caixa torácica, a aproximação das cordas vocais²¹, entre outros. De acordo com Catford (2001), o movimento dos órgãos na fase anterior age sobre o ar que está contido dentro do trato vocal, comprimindo-o ou dilatando-o, por meio de diferentes movimentos, como jatos rápidos, explosões repentinas, correntes suaves, entre outros. Essa movimentação do ar caracteriza a fase aerodinâmica.

A quinta fase é a acústica. O movimento do ar durante a etapa aerodinâmica produz ondas sonoras que são percebidas pelo ouvinte. Durante esse processo uma onda sonora transportada pelo ar irradia da boca do falante atingindo o ouvido de qualquer um que esteja em uma distância audível, incluindo o próprio falante.

O ar que chega ao aparelho auditivo recebe as vibrações das ondas sonoras que são transmitidas ao ouvinte por meio dos pequenos ossos do ouvido médio e interno. Nesse caminho ocorre a estimulação dos terminais nervosos dos nervos auditivos. Os impulsos nervosos presentes nesses terminais viajam do nervo auditivo ao cérebro, gerando a sensação do som. Esse processo de estimulação e transmissão dos impulsos neurais constitui a sexta etapa de produção da fala.

Por fim, a sétima etapa é a de identificação neurolinguística. Nesse momento, os sinais neuroreceptivos são identificados como sons específicos da fala.

De acordo com Catford (1977), dentro de todo processo de fala, a fase acústica é a mais remota da fala como atividade humana. A sua função é simplesmente transmitir as informações da fala do falante ao ouvinte, “*it is not itself part of the human communicative event, which is the linguistic phonetician’s central interest*”²² (Catford, 1977, p. 47). Mesmo assim, não deixa de ser necessária aos estudos fonéticos, uma vez que o desenvolvimento de técnicas dos últimos 30 ou 40 anos tem feito da fase acústica uma das mais acessíveis. Além disso, ela é necessária em cada fase do processo de fala. Muitas vezes, os dados acústicos nos ajudam a entender o que ocorre com a fase aerodinâmica. Segundo Catford (1977), é uma das fontes principais de informação sobre a fase perceptual.

O sinal acústico forma-se quando os órgãos vocais estão em movimento. Essa movimentação agita as moléculas de ar que são propagadas para fora em todas as direções, até

²¹ As cordas vocais (termo mais popular) também são conhecidas como pregas vocais. A primeira nomenclatura está relacionada ao pensamento, diga-se de passagem, errôneo, de que essas estruturas vibravam de modo semelhante às cordas de um violino.

²² “ela não faz parte do evento comunicativo humano, que é o centro de interesse do linguísta foneticista. [Tradução nossa].

atingir o ouvinte (HARRINGTON, 2010). A fonética acústica descreve os diferentes tipos de sinal acústico.

Fisicamente, o som é o movimento oscilatório das moléculas de ar. Segundo Catford (1977), esses movimentos podem ser de dois tipos: irregulares ou regulares. Os do primeiro tipo caracterizam um fluxo de ar turbulento, resultado das mudanças na velocidade do ar. Já os do segundo tipo são resultado de vibrações regulares das cordas vocais.

Uma oscilação completa é chamada de *ciclo* (*cycle*). A frequência (F0) é o número de ciclos (vibrações) completos em um segundo, sendo representados por *cps* (ciclos por segundo) ou, de acordo com a notação internacional em *Hz* (Hertz).

A escala de frequência que uma pessoa jovem pode ouvir está entre 20 e 20000 Hz. A maioria das frequências que são importantes na fala situam-se entre 50 e 8000 Hz. A variação melódica é proporcional à F0, ou seja, quanto menor a frequência F0 menor a variação melódica e, quanto maior a F0, maior a variação melódica. Na fala masculina, a variação melódica média é de 100 Hz, enquanto a da mulher é de 200 Hz.

A amplitude também está presente na onda sonora. Ela está relacionada à medida da magnitude durante um ciclo de uma onda sonora. Em princípio, a amplitude determina a intensidade do som, desde que a frequência seja constante. A sua unidade de medida é o *decibel*, dB. Devido à natureza da onda sonora, a amplitude ou pressão do som varia constantemente de um momento para o outro.

A intensidade, por sua vez, diz respeito à quantidade de energia transmitida ao longo de uma onda sonora. A intensidade acústica dos diferentes sons varia de falante para falante. De acordo com Catford (1977), a intensidade equivale ao dobro da amplitude. A amplitude e a intensidade são correlatos físicos da altura.

Do ponto de vista acústico, o trato vocal funciona como um gerador de sons e como um sistema variável de ressonadores que respondem às frequências dos sons gerados.

Os sons da fala são produzidos no trato vocal, que é formado pelo aparelho respiratório, laringe, boca e nariz.

Os quatro processos principais do mecanismo de produção da fala são: o processo articulatório, formado pela cavidade oral (lábios, língua e dentes), que é responsável por gerar sons específicos. O segundo é o processo oro-nasal, formado pela cavidade nasal. O terceiro caracteriza o processo fonatório, que é responsável por modular o ar que passa através da laringe. Em outras palavras, é a ação das cordas vocais.

Entre as cordas vocais forma-se uma abertura chamada de glote. Ao passar pela laringe, a corrente de ar é acusticamente modulada (MASSINI-CAGLIARI; CAGLIARI, 2008). Durante o processo fonatório, a vibração das cordas vocais produz sons sonoros. Quando a corrente de ar passa pela laringe, sem alteração, ocorre a produção dos sons desvozeados.

O último processo é o do fluxo de ar (*airstream process*), que é formado pelo sistema respiratório. Nessa fase, a saída de ar pelos pulmões constitui a etapa principal. De acordo com Ladefoged e Johnson (2010), a corrente de ar que passa através da boca e do nariz é determinada por meio do processo oro-nasal. Já os movimentos da língua e dos lábios em interação com o palato e a faringe são parte do processo articulatorio.

As possibilidades articulatórias da estrutura da laringe e cordas vocais permitem a produção de diferentes tipos de fonação e segmentos, como os sons vozeados e desvozeados; a produção da oclusiva glotal, por exemplo, que ocorre devido ao fechamento da glote; produz sons do tipo *creaky voice*, que são os sons graves produzidos por meio da corrente de ar lenta que passa pelos pulmões. (MASSINI-CAGLIARI, CAGLIARI, 2008; CATFORD, 2001, LADEFOGED, 2010).

Outra importante função da articulação consiste na configuração final da corrente de ar, gerando um som de tipo e qualidade específicos. Nesse sentido, as consoantes podem ser classificadas conforme o modo e lugar de articulação, vibração ou não das cordas vocais e características do mecanismo aerodinâmico envolvido.

Levando em consideração o mecanismo aerodinâmico, a direção do ar determinará se ela será do tipo egressiva ou ingressiva, como mencionado. Os sons egressivos caracterizam a saída do ar do corpo e os ingressivos, a sua entrada. Segundo Massini-Cagliari; Cagliari (2001) os sons do primeiro tipo também podem ser denominados como plosivos e os do segundo tipo de implorivos.

Quanto ao modo de articulação, os sons podem ser do tipo, oclusivo, nasal, fricativo, africado, lateral, vibrante, retroflexo ou aproximante. Já em relação aos lugares de articulação eles podem ser, bilabial, labiodental, dental, alveolar, palatoalveolar, alveopalatal, palatal, velar, uvular, faringal ou glotal.

Nos sons vocálicos, o ar passa livremente pela cavidade oral, e por isso, são sempre vozeados. Eles são classificados de três maneiras, 1) de acordo com a posição da língua na área vocálica, podendo ser, alta, médio-alta, médio-baixa e baixa; 2) por meio das regiões

articulatórias, anterior, central e posterior e, 3) conforme a posição dos lábios durante a sua produção, labializado (arredondado) ou não-labializada (não-arredondados)²³.

Dentre as possibilidades de articulação das consoantes, vamos tratar especificamente 1) das oclusivas bilabiais, alveolares e velares; 2) das fricativas labiodental e alveolar; 3) das nasais bilabial e alveolar e, 4) da lateral dental. Essa opção deve-se ao fato de que grande parte das disfluências ocorrem para esses segmentos, em posição de sílaba inicial.

Nos sons com modo de articulação oclusivo há um bloqueio total à corrente de ar. Ladefoged e Johnson (2010) dividem as oclusivas em orais e nasais. Sobre as oclusivas orais, os autores explicam que se além do fechamento articulatorio da boca, o palato mole estiver elevado, bloqueando o trato nasal, a corrente de ar estará completamente obstruída. Isso faz com que haja uma pressão na boca, formando então a oclusiva oral. Quando há uma separação dos articuladores, uma corrente com pequenas explosões de ar é liberada, e os sons bilabiais [p, b]; alveolares [t, d] e velares [k, g] são formados.

No que diz respeito às oclusivas nasais, Ladefoged e Johnson explicam que se o ar está parado na cavidade oral, mas o palato mole está abaixado, o ar passa pelo nariz, produzindo uma nasal oclusiva. Os sons desse tipo são a nasal oclusiva bilabial [m], a nasal oclusiva alveolar [n], e a nasal velar oclusiva [ŋ].

Quanto aos lugares de articulação, de acordo com Massini-Cagliari e Cagliari (2008), os sons oclusivos são produzidos por meio de um fechamento dos lábios. Já os alveolares, quando a parte da frente da língua toca os dentes incisivos superiores. Nos sons velares o dorso da língua está aumentado e toca o palato mole.

As fricativas, por sua vez, caracterizam a saída do ar dos pulmões. Nos segmentos fricativos, o ar, ao sair dos pulmões, passa por um caminho turbulento, produzindo o som fricativo. Os sons [f], [v] e [s] são fricativos, mas cada um é articulado em um lugar diferente. Por exemplo, os sons [f] e [v] são produzidos com os dentes superiores tocando os lábios inferiores, por isso, são chamados de labiodentais. Já o som [s] é produzido quando o ápice da língua toca os alvéolos dos dentes incisivos superiores. Os lugares de articulação são os obstáculos encontrados pelo ar. A consequência acústica da passagem de ar turbulenta é a geração de uma energia aperiódica.

²³ Não é nosso interesse fazer uma descrição detalhada do modo e lugar de articulação das consoantes, bem como da posição da língua, região articulatória e posição dos lábios das vogais. A nossa intenção consiste apenas em fazer uma caracterização geral das suas possibilidades articulatórias, já que elas são a base do entendimento de algumas questões que serão apresentadas nesta tese. Além disso, o estudo fisiológico, acústico e aerodinâmico é a preocupação central da fonética. Para uma revisão mais detalhada consultar: Ladefoged e Johnson (2010); Massini-Cagliari, G e Cagliari, L. C (2008); Catford, J. C (2001).

As nasais que podem ser do tipo bilabial e alveolar são representadas pelos sons [m] e [n], respectivamente. Os sons nasais são produzidos na cavidade nasal. Segundo Catford (2001), a área nasal é uma câmara de revestimento da membrana mucosa que não é capaz de produzir movimentos voluntários. No entanto, segundo o autor, nas fossas nasais, no orifício faringo-nasal e nas áreas terminais da cavidade nasal pode ocorrer controles voluntários, e conseqüentemente, diferentes variedades de articulação.

Nos sons nasais, o véu palatino encontra-se abaixado. Como há um fechamento completo da boca para a produção do som bilabial [m] e entre a ponta da língua e a parte posterior dos dentes superior e inferior, para o som alveolar [n], todo o ar usado para sua produção é desviado através do nariz. De acordo com Catford, nos sons nasalizados tanto as fossas nasais quanto a boca, estão abertas, o que possibilita a saída do ar saia por dois caminhos.

Shriberg (2001) confirmou por meio dos seus dados uma maior tendência para a produção de disfluências na oclusiva bilabial. Segundo a autora, a causa dessa inclinação ainda não é clara. Mas pode, segundo ela, refletir a tendência em cortar o *gesture* durante o fechamento dos lábios.

Por fim, no modo de articulação lateral, conforme Ladefoged e Johnson (2010) ocorre uma obstrução da corrente de ar em um ponto ao longo do centro do aparelho fonador. Durante a produção desses segmentos, há um fechamento incompleto entre os dois lados da língua e o céu da boca. Os sons laterais também podem ser classificados como aproximantes, uma vez que não há uma interrupção do ar ou qualquer ruído fricativo. O segmento representativo da lateral alveolar é o [l].

Como vimos, o critério para distinguir um tipo de articulação de outra é aerodinâmico e acústico, conforme a presença ou ausência da turbulência do fluxo de ar através do canal articulatório.

Por meio da análise aerodinâmica é possível medir a pressão subglótica e o fluxo oral. (Oliveira *et al*, 2013). De acordo com Oliveira *et al* (2013) essas medidas são utilizadas para a avaliação clínica da função laríngea, que atuando de forma não invasiva, ajudam na detecção das patologias do trato vocal.

Os pulmões servem como uma fonte de pressão constante durante a fonação, e essa pressão pode ser utilizada para determinar a eficiência com que a laringe converte a energia aerodinâmica para a energia acústica. À medida que a laringe funciona como um transdutor de aerodinâmica, a eficiência vocal, definido como eficiência sonora (dB), dividido pela energia

aerodinâmica (fluxo oral), indica as condições da saúde bucal (OLIVEIRA *et al*, 2013, p. 120).

Ainda de acordo com os autores, essas medições ajudam a entender melhor as características físicas das vias aéreas e as propriedades mecânicas dos tecidos da laringe. Além disso, eles explicam que a resistência glótica depende do tamanho da via aérea, e que outros fatores contribuem para a sua resistência, assim como o grau de adução das cordas vocais e a velocidade das partículas que passam por meio da glote.

Tendo em vista essas questões, Oliveira *et al* fizeram um estudo das medidas aerodinâmicas em falantes do PB (Português do Brasil) que não apresentavam queixas vocais. Essas medições foram feitas por meio do programa EVA (Método multiparamétrico de avaliação vocal objetiva assistida), desenvolvido na França, no Laboratoire *de parole et lagage* da Universidade de *Aix-en-Provence*.

Segundo os autores, esse programa permite a análise da maioria dos parâmetros de produção de fala, como som, intensidade, pressão aerodinâmica, entre outros. Para a análise, foram selecionados 40 informantes, sendo 20 do sexo masculino e 20 do sexo feminino, entre 18 e 45 anos. As vozes dos falantes foram gravadas através dos sinais acústicos e aerodinâmicos, como valor de referência para a pressão subglótica e fluxo oral. Para a gravação, os autores utilizaram um equipamento acoplado ao EVA para a medição do fluxo de ar, de modo que houvesse a vedação necessária.

Os autores pediram aos informantes que gravassem a sequência “Papai papa o papá de Paulo”. De acordo com o Oliveira *et al* (2013), a opção por essa sequência deve-se ao fato de que o fonema /p/ é o mais favorável para as medições da pressão subglótica. Eles explicam que o fato dos lábios estarem fechados e a glote aberta durante a produção desse fonema, leva a um equilíbrio da pressão subglótica. Os autores também analisaram as medidas do fluxo oral dos informantes.

A partir dessas medições e testes estatísticos, os pesquisadores chegaram a um resultado do que seria o padrão de normalidade para a pressão subglótica e fluxo oral. Além disso, eles também verificaram uma diferença em relação ao gênero e as medidas de intensidade, eficiência da laringe e F0 para a pressão subglótica. Quanto ao fluxo oral, a diferença ocorreu apenas para a frequência fundamental²⁴.

²⁴ Os resultados estatísticos detalhados da análise feita por Oliveira *et al* (2013) podem ser vistos integralmente por meio do artigo “Análise das medidas aerodinâmicas no português brasileiro por meio do método multiparamétrico de avaliação vocal objetiva assistida (EVA)”. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-18462013000100014>.

Os valores apontados no trabalho dos autores não ajudam a afirmar que momentos de disfluência, seja ela comum ou gaga, estejam relacionados a variações na pressão subglótica e no fluxo oral, já que não temos o equipamento necessário para desenvolver esse tipo de análise. No entanto, os valores revelam o que seria o padrão de normalidade para a fala do PB, em detrimento desses dois parâmetros, além de possibilitarem comparações futuras.

Os resultados obtidos por meio dessa pesquisa também indicam que as disfluências podem estar relacionadas a questões aerodinâmicas. Ainda assim, seria preciso desenvolver uma análise mais detalhada, com fonemas mais adequados à medição, pois os autores trabalharam com frases pré-estabelecidas e, o nosso *corpus* diz respeito a gravações de fala semiespontânea. Talvez, esse fato leve a uma diferença nos resultados. A opção pela análise do fonema /p/ deve-se a possibilidade da influência aerodinâmica na produção das disfluências, uma vez que grande parte dos bloqueios concentram-se durante a produção das oclusivas. Resta-nos saber, portanto, se essa tendência também ocorreria para as fricativas e nasais.

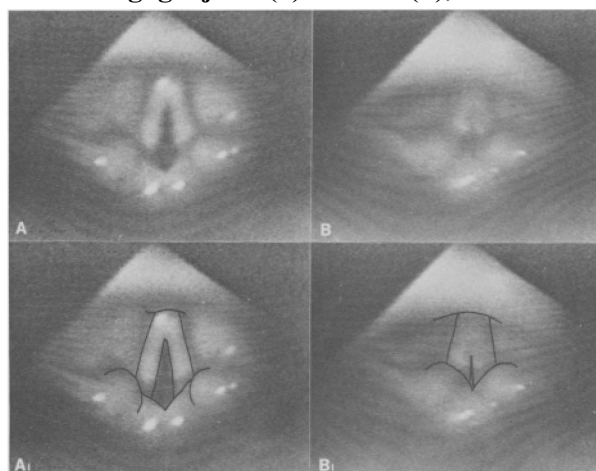
Dessa maneira, as disfluências ocorreriam em detrimento do bloqueio do ar, em determinados momentos, durante a produção dos sons. Em outras palavras, o ar ficaria retido, dificultando a produção dos sons, principalmente das consoantes oclusivas bilabiais e alveolares, das fricativas labiodentais e alveolares e das nasais bilabial e alveolar. Esse fato também nos permite sugerir que essa seja a maior diferença entre a disfluência gaga e comum. Pois, nos parece, que no caso das disfluências gaga, o ar ficaria recluso por mais tempo até a sua liberação, o que levaria a produção dos espasmos. Por outro lado, na disfluência comum, a reclusão do ar seria quase nula, o que a deixaria quase imperceptível para quem ouve²⁵.

Ao analisar a laringe durante a produção de eventos com disfluência gaga e fluentes, Conture *et al* (1977) perceberam que a laringe, em grande parte das vezes, não se comporta da mesma forma²⁶. Conforme os autores, o mau funcionamento da laringe faz com que os gagos apresentem tempos de reação de vozeamento mais lento e variável, além de VOTs maiores, que aumentam à proporção que o lugar de articulação se afasta da cavidade oral. Veja na imagem 1, abaixo, o comportamento da laringe em um indivíduo fluente e em outro com disfluência gaga.

²⁵ A análise dos dados permitirá que seja feita uma conclusão mais categórica sobre a hipótese em questão

²⁶ De acordo com Freeman e Ushijima (1978, p. 538), por quase um século e meio autores como, Arnott, 1828; Muller, 1833; Hunt, 1861; Wyke, 1971, entre outros, propuseram estudos que correlacionavam à disfluência gaga ao componente da laringal.

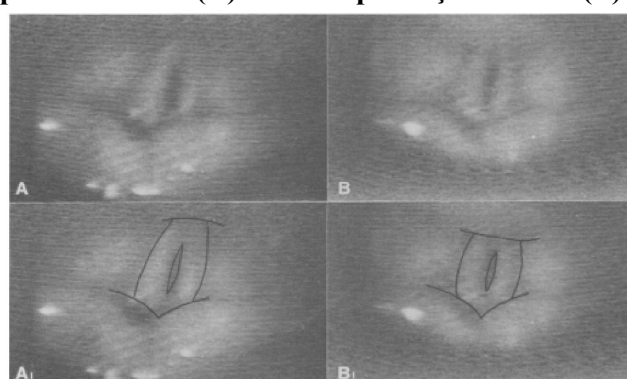
Figura 1 – Imagem do comportamento da laringe durante a produção de uma palavra gaguejada (a) e outra (b), fluente.



Fonte: (Conture *et al*, 1977, p. 261)

A figura 1a representa o informante com disfluência gaga durante um período de repetição da palavra *raibow*. Já a 1b caracteriza a produção fluente da mesma palavra por um informante que não apresenta disfluência gaga. Como podemos observar, em 1a as cordas vocais ficam separadas durante a repetição disfluenta. Já em 1b elas estão muito próximas. No entanto, por meio do mesmo procedimento, os autores verificaram que durante a produção de um som alongado não houve alteração da abertura da laringe. Conture *et al* observaram que tanto na produção gaga quanto na fluente ocorreu uma pequena aproximação entre as cordas vocais. Observe a seguir:

Figura 2 – Exemplificação do comportamento da laringe durante a produção da vogal alongada [i] na palavra *tea* em (A) e da sua produção fluente (B).



Fonte: Conture *et al* (1977, p. 666).

Os resultados mostram, então, que o comportamento da laringe varia conforme o tipo de disfluência gaga. De acordo com Conture *et al*, uma hipótese para isso seria o fato de que

cada tipo de disfluência esteja associada a um padrão diferente de esforço muscular da laringe. Para exemplificar, eles dizem que um evento disfluyente pode relacionar-se com a atividade inapropriada dos músculos adutores e abdutores, enquanto que outro tipo, pode estar relacionado apenas com a atividade inapropriada dos músculos adutores.

Conture *et al* também consideram que o comportamento da laringe pode estar relacionado a uma variedade de outros fatores, assim como a polarização de tonicidade da laringe, o seu nível de atividade muscular, o local inapropriado de ocorrência da atividade muscular da laringe no enunciado, entre outros. Conforme os autores, mesmo que as análises feitas indiquem que o comportamento inadequado da laringe esteja associado, na maior parte das vezes, às disfluências gagas, não é possível afirmar que ele seja a sua causa. Esse esclarecimento só será possível a partir de novas investigações.

Em 1978, Freeman e Ushijima, no artigo “*Laryngeal muscle activity during stuttering*”, também observaram a atividade da laringe durante produção gaga e fluente. Assim como Conture *et al* (1977), os autores verificaram, por meio de uma eletromiografia, que os eventos disfluentes vinham acompanhados por uma alta atividade dos músculos laringais e por perturbações entre a reciprocidade normal dos músculos adutores e abdutores da laringe. Por meio da análise dos dados, Freeman e Ushijima mostraram que há uma forte relação entre a atividade anormal dos músculos da laringe com os momentos de gagueira.

Mais tarde, em 1985, Conture *et al* confirmaram por meio de novos dados, que durante a produção de eventos gagos, a laringe não se comporta normalmente. Os autores acreditam que esse comportamento anormal deva-se também a falhas nos eventos respiratórios e articulatórios. Por exemplo, eles dizem que a abdução inapropriada da laringe observada em seus dados durante a repetição de uma sílaba ou som, pode estar relacionada com inspirações breves e/ou devido a não liberação do contato articulatório supraglotal²⁷.

As disfluências também afetam vários aspectos fonéticos da fala, como a duração, a entoação, a qualidade de voz, a taxa de elocução, a tessitura, o ritmo (fenômenos prosódicos), bem como a coarticulação e a intensidade²⁸.

Cagliari (1982) diz que durante a coarticulação ocorrem modificações fonéticas de vizinhança, devido à interferência das características articulatórias de um som sobre o outro. Esse fenômeno pode ocorrer tanto no início quanto no fim de um enunciado. Segundo Shriberg (1999, 2001), quando ocorre disfluência, há uma mudança na base do padrão

²⁷ Por meio de uma pesquisa sobre os trabalhos atuais de Edward Conture, verificamos que o último trabalho do autor relacionado ao comportamento anormal da laringe foi o de 1985. No entanto, ele não deixou de trabalhar com a disfluência gaga.

²⁸ Sobre a intensidade, ver a introdução deste capítulo.

coarticulatório. De acordo com ela, na fala fluente, os articuladores se movem normalmente em direção às posições do articulador para o início da próxima palavra. No entanto, no caso das disfluências essa relação não pode ser assumida. A coarticulação é regida pela próxima palavra do plano fonético do falante no momento em que ela é enunciada, e não pela sequência de palavras que é, em última análise, produzida. Segundo Shriberg, na fala fluente, o plano e o *output* final são consistentes, mas com as disfluências, ou o conteúdo lexical que segue pode estar indisponível temporariamente ou o plano pode mudar rapidamente.

Conforme Catford (2001, p. 163), “[...] *speech is a continuum; a continuous flux of initiatory, phonatory, and articulatory states and movements, constantly changing, often overlapping and interpenetrating and influencing each other*”²⁹. Assim, quando tratamos esses sons individualmente, diz o autor, estamos quebrando essa cadeia de eventos em uma série de segmentos. No entanto, esse tipo de estudo não deixa de ser importante, pois permite analisar os sons de modo mais detalhado.

Os fenômenos prosódicos não são somente uma característica dos segmentos individuais, mas também da fala em um contínuo. Segundo Massini-Cagliari e Cagliari (2008) são esses fenômenos que permitem o fluir da fala. Além disso, eles podem atuar como uma forma de sinalizar um erro ao ouvinte (LEVELT; CUTLER, 1983).

A tonicidade e a variação melódica estão relacionadas à iniciação e à fonação. A primeira diz respeito ao modo como a energia muscular liberada conduz a iniciação, ou seja, o órgão que inicia o fluxo de ar na fala. Conforme Massini-Cagliari e Cagliari (2008), uma sílaba será tônica ou átona a partir do seu grau de saliência.

De acordo com os autores, as sílabas tônicas podem ser de três tipos: 1) de acento primário, como por exemplo, a sílaba *zi*, da palavra *cafezinho*; 2) de acento secundário, como a sílaba *ca* ainda na palavra *cafezinho* e 3) com acento frasal. No caso de (3), os autores afirmam que a sílaba é acentuada no nível da palavra, recebendo um acento frasal, como no enunciado:

- *Vou tomar um cafezinho.*

A entoação determina o acento frasal. A variação da posição do acento frasal também muda o foco do enunciado.

Quanto ao acento secundário, Massini-Cagliari e Cagliari explicam que de acordo com as normas da língua, uma sequência muito grande de sílabas átonas não é permitida. Assim, a

²⁹ [...] a fala é um contínuo; um fluxo contínuo de iniciação, fonação, e movimentos e estados articulatórios, em constante mudança, muitas vezes se sobrepondo, se interpenetrando e influenciando uns aos outros [Tradução nossa].

fim de formar uma sequência rítmica mais regular, essas sílabas recebem o que os autores chamam de esforço extra. Nesse caso, o acento secundário seria considerado um efeito de regras de eurritmia da língua.

A questão lexical também define um acento secundário, como por exemplo, os derivados de *-íssimo*, e *-(z)inho*, nas palavras *bèlíssima* e *hotèlzinho*. (MASSINI-CAGLIARI; CAGLIARI, 2008, p. 115).

A entoação ou variação melódica está relacionada à fonação e inclui características da qualidade de voz. A variação na frequência de vibração das cordas vocais leva a produção de diferentes variações melódicas (CATFORD, 2001). Essa mudança na frequência pode ser produzida por dois modos: 1) pelo alongamento e tensão das cordas vocais. Nesse caso, quanto maior a tensão, maior a variação melódica, ou 2) pela mudança de pressão subglotal. A variação melódica será proporcional à variação da pressão subglotal.

Cada sílaba apresenta uma altura melódica. Nas línguas tonais, como o chinês, essa altura é fixa, já nas línguas entoacionais, como o português, os padrões melódicos variam de acordo com o enunciado. Por exemplo, as frases interrogativas apresentam um padrão entoacional ascendente³⁰. Para Abercrombie (1967), são essas as características que levam a entoação a ser, provavelmente, um dos elementos mais importantes da fala, já que o fluxo conversacional depende tanto das flutuações tonais quanto dos indicadores afetivos que, conseqüentemente, acarretam mudanças na tessitura.

A tessitura diz respeito à variação melódica na voz de uma pessoa quando esta fala normalmente (CRYSTAL, 2000). Conforme Cagliari (2002), a tessitura ocorre de acordo com os ajustes da laringe que podem aumentar ou abaixar a sua F0. É o elemento prosódico que permite reconhecer uma pessoa sem nunca tê-la visto (CAGLIARI, 1992a).

Normalmente, a tessitura é constante. Porém, as intenções discursivas do falante podem determinar variações. Por exemplo, segundo Cagliari (2002), quando queremos destacar um contexto, usamos uma tessitura mais alta. Por outro lado, contextos mais irrelevantes, encaixados, ou fora do lugar sintático característico, apresentam uma tessitura mais baixa. Expressões de raiva, fúria, alegria, desespero, entre outras, também estão relacionadas a essas variações. A tessitura também tem o papel de revelar a estrutura do discurso, por isso devem ser estudadas juntamente com o ritmo, entoação e qualidade de voz.

A qualidade de voz é uma das propriedades fonéticas que individualiza o falante (CAGLIARI, 2007). Isso ocorre conforme as questões anatômicas e articulatórias. A questão

³⁰ Para uma descrição mais detalhada da entoação veja: Halliday (1970) e Cagliari (2007).

anatômica, segundo Cagliari (2007), está relacionada às diferentes dimensões do aparelho fonador dos homens, mulheres e crianças. E, a articulatória, diz respeito ao modo como as pessoas articulam os sons. A qualidade de voz é produzida no nível da glote, como por exemplo, um tipo de fonação trêmula (ou *creaky voice*). De acordo com Ladefoged e Johnson (2010), na fonação do tipo *creaky voice* (ou laringalização), as cordas vocais não estão completamente fechadas. Essa fonação caracteriza-se por apresentar uma variação melódica baixa, com uma frequência que varia entre 20 Hz e 50 Hz.

Ao dividir o processo fonatório, fisiologicamente, em quatro fases, sendo elas 1) tensão da laringe e dos músculos respiratórios; 2) início da respiração; 3) adução final das cordas vocais e 4) iniciação da vibração das cordas vocais, Duez (2001) afirma que a fonação do tipo *creaky voice* ocorre devido a uma falta de ajustamento entre as fases dois e três. De acordo com a autora, essa fonação é frequente nas pausas preenchidas, mas não exclusiva, pois também pode ser observada em outros momentos de disfluência (SHRIBERG, 1999). Shriberg (1999), por exemplo, verificou que o *creaky voice* pode ocorrer após a interrupção do enunciado pelo falante. Em outras palavras, ao perceber que errou, o falante interrompe o enunciado para corrigi-lo usando uma qualidade de voz do tipo *creaky voice*.

A velocidade de fala, segundo Oliveira *et al* (2013, p. 124) é definida como “a taxa de variação da distância percorrida por uma partícula no tempo”. A velocidade do fluxo do ar é maior quando o trato vocal encontra-se desobstruído. Ao citar o princípio de Bernoulli, Oliveira *et al* explicam que o estreitamento das cordas vocais faz com que a velocidade do fluxo de ar seja maior. Em consequência disso, há uma diminuição da pressão.

Sobre esse fenômeno prosódico, Abercrombie (1967, p. 96) diz:

By tempo is meant speed of speaking, which is best measure by rate of syllable-succession. It is a feature which, like loudness, is varied from time to time by the individual speaker. Some people employ more variation in tempo than others, but everyone has a norm which is characteristic of his usual conversational style³¹.

Conforme Massini-Cagliari e Cagliari (2008), as variações de velocidade levam a modificações fonéticas. De acordo com os autores, quanto maior for a taxa de elocução, maior será a tendência “para a centralização vocálica, para a queda de segmentos, para a co-

³¹ Tempo significa velocidade de fala, cuja melhor medida é dada pela velocidade de sucessão das sílabas. É uma propriedade que assim como a intensidade, varia de tempo em tempo, e de acordo com cada falante. Algumas pessoas apresentam mais variação no ritmo do que outras, mas todas apresentam uma norma que é característica do seu estilo usual de fala [Tradução nossa].

articulação, para a perda de qualidades articulatórias e conseqüente perda de inteligibilidade da fala” (MASSINI-CAGLIARI; CAGLIARI, 2008, p. 117). A diminuição da velocidade também é um problema, pois de acordo com os autores, o falante começa a ter dificuldades de articulação e o ouvinte de percepção. Por outro lado, os autores explicam que a variação na velocidade também pode ter outras finalidades. Por exemplo, a desaceleração da velocidade pode evidenciar o final do turno discursivo ou enfatizar um argumento. Já o seu aumento evita a intromissão do interlocutor.

A duração está relacionada à articulação e, desse modo, à duração das sílabas (ABERCROMBIE, 1967). Daniel Jones (1960) refere-se a esse elemento prosódico como o período de tempo que ele se mantém em uma frase ou palavra. Independentemente dos segmentos, ela está sistematizada ao nível das sílabas, diz Cagliari (1992).

De acordo com Cagliari (2007), os sons só podem ser avaliados quanto à sua duração se comparados uns com os outros, por isso a duração é uma medida relativa. Além disso, de acordo com o autor, o fato de cada língua ter a sua duração prevista no léxico faz com que elas tenham a chamada duração intrínseca (CAGLIARI, 1994). Ainda segundo o autor, essa duração atribui valor e peso às sílabas, que são importantes para que os processos fonológicos sejam ou não aplicados. As durações referentes aos processos fonológicos dizem respeito ao alongamento ou encurtamento dos segmentos. Para Cagliari, o recurso do alongamento no Português é utilizado para significar qualidades atributivas. Veja, a seguir, os exemplos retirados de Cagliari (1992, p. 142):

- a) João tem uma ca:sa!
- b) O cavalo co:rre

Conforme o autor, no exemplo (a), o alongamento da sílaba *ca*, representado pelo sinal diacrítico de dois pontos (:), modifica o significado literal da palavra *casa*, que agora passa a significar que a casa é de excelente qualidade, grande, bonita, entre outras. Em (b), por sua vez, o alongamento da sílaba *co*: faz uma alusão ao fato do cavalo não apenas correr, mas correr muito rápido, mais do que seria o normal.

As vogais e as consoantes podem ser articuladas de modo variável em sua duração. De acordo com Catford (2010), há uma tendência para as vogais abertas terem uma duração maior do que as vogais fechadas. Segundo o autor, uma razão para isso é o fato das vogais abertas exigirem um movimento articulatório maior do que as vogais fechadas.

Para Cruttenden (1986), a relevância da duração como uma característica prosódica é difícil de ser avaliada porque há, frequentemente, influências diferentes na duração absoluta de um segmento ou sílaba.

A duração dos segmentos e das sílabas está diretamente relacionada ao estudo do ritmo da fala. De acordo com Abercrombie (1967), todas as línguas possuem um ritmo, embora as hesitações e outras pausas tendam, às vezes, a disfarçar esse fato. Conforme o autor, o ritmo emerge claramente durante os momentos em que a fala é fluente e ininterrupta.

Na fala, o ritmo caracteriza a recorrência periódica de algum tipo de movimento, produzindo a expectativa de que a regularidade da sucessão continuará. De acordo com Cagliari (2007, p. 132), o ritmo caracteriza-se por uma simetria, “uma harmonia resultante de certas combinações e proporções regulares”. A repetição e a expectativa são, conforme o autor, duas propriedades fundamentais para a percepção do ritmo. Por repetição, entende-se a segmentação do contínuo do movimento da fala em pedaços. Segundo Cagliari, esses pedaços ou unidades rítmicas (como veremos a seguir) apresentam uma duração “que pode ser medida e controlada pelo observador e, portanto, podem ser comparadas com a expectativa que se tem delas” (CAGLIARI, 2007, p. 132-133).

Os movimentos relacionados à produção do ritmo da fala são correspondentes ao processo de produção da sílaba e da tonicidade. Juntos, eles compõem o mecanismo pulmonar de fluxo do ar. O ritmo na fala é essencialmente um ritmo muscular, e os músculos envolvidos nesse processo são os respiratórios. Sendo assim, em relação ao processo aerodinâmico, Cagliari (2012) explica que o ritmo depende do volume de ar que é controlado na fala, produzindo, assim, as unidades com mais ou menos volume, ou intensidade.

Abercrombie (1967) explica que o ritmo é sentido empaticamente pelo ouvinte através dos movimentos do falante. Segundo ele, os sons que ouvimos são pistas. No entanto, o autor diz que para termos essa apreensão intuitiva e imediata é necessário que falante e ouvinte tenham a mesma língua materna, caso contrário, essa empatia fonética não será possível, uma vez que os sons não serão reconhecidos como pistas precisas dos movimentos que os produzem. Dessa maneira, não é só uma medida física, mas também cognitiva (CAGLIARI, 2012). Ainda segundo o autor,

O ritmo depende da velocidade com que cada unidade fonética é pronunciada e da identificação de saliências fonéticas, tudo interpretado fonologicamente; ou seja, o ritmo é um fenômeno idealizado na mente do falante, mas que tem apoio na percepção auditiva e até nos mecanismos articulatórios de produção e de percepção dos sons da fala. É um fenômeno envolvido com o mecanismo aerodinâmico da fala, com a duração

articulatória dos segmentos, com a categorização temporal das unidades de fala, com a produção e percepção desses elementos e, finalmente, é um fenômeno comprometido com o sistema sonoro da língua (CAGLIARI, 2012, p. 33).

Conforme Cagliari (2007), o ritmo pode ser fixo ou variado. O ritmo fixo diz respeito à repetição de um padrão constante. O autor ilustra o ritmo fixo através do Latim, pois seria formado por uma sequência de sílabas longas e breves. A duração das sílabas longas é o dobro da breve. Essa duração pode ser representada pela isocronia das batidas das sílabas tônicas e através do barulho constante das máquinas. Já o ritmo variado caracteriza uma sucessão regular não constante de marcas rítmicas no decorrer no tempo. Segundo Cagliari (2007, p. 133):

Numa língua cujas sílabas têm uma duração predeterminada, os intervalos entre as ocorrências de sílabas tônicas marcará um ritmo variado, exceto na poesia de forma fixa; e as sílabas marcarão um ritmo fixo. Nas línguas que se caracterizam por terem os intervalos entre as sílabas tônicas ocorrendo em períodos de tempo iguais, terão um ritmo fixo com relação a esse fato, e um ritmo variado com relação à duração individual das sílabas.

A partir da tradição dos estruturalistas americanos, como Pike (1947) e foneticistas ingleses, como Abercrombie (1967), as línguas são caracterizadas de acordo com o tipo ritmo que apresentam, dividindo-se em línguas de ritmo silábico e acentual. As línguas de ritmo silábico caracterizam-se por apresentar sílabas com duração aproximadamente igual. Como exemplo, podemos citar o Francês, o Yorubá, o Japonês e o Espanhol entre outras. Por outro lado, quando a classificação é feita de acordo com a ocorrência de sílabas acentuadas no enunciado, temos uma língua de ritmo acentual. Em línguas desse tipo, as sílabas acentuadas ocorrerão em intervalos de duração aproximadamente iguais ou isocrônicos. “Entre duas sílabas acentuadas poderá ocorrer, por exemplo, uma ou cinco sílabas não acentuadas. Nesse caso, a duração individual das cinco sílabas será proporcionalmente muito mais reduzida do que a duração da única sílaba não-acentuada entre duas sílabas tônicas” (CAGLIARI, 2007, p.134). Como exemplo desse caso podemos citar o Inglês, o Russo, o Português e o Árabe, entre outras (ABERCROMBIE, 1967; CAGLIARI, 2007).

Uma pessoa falante de uma língua de ritmo acentual costuma apresentar disfunção / gagueira quando começa a aprender a falar no ritmo acentual, deixando o silábico. (ABERCROMBIE, 1965, 1967; CAGLIARI, 2002, 2007).

O ritmo é gerado por vários elementos que estão presentes na dinâmica da fala. Cagliari (2007) atribui a esses fenômenos o nome de “unidades rítmicas da fala” e diz que o ritmo é melhor compreendido a partir do entrelaçamento dessas das unidades.

As unidades rítmicas são, de acordo com o autor, as sílabas, as moras, os pés e os intervalos, o grupo tonal, as pausas, o impulso e o repouso, o *icto* e a *rêmis* e, por fim, a *ársis* e a *tésis*.

As sílabas são o suporte para a construção da fala que, por sua vez, é uma modificação do sistema respiratório. De acordo com Cagliari, essa modificação sempre apresentará uma duração maior do que a necessária para a produção do enunciado, isto é, “nenhum enunciado se inicia com a primeira sílaba representando o primeiro pulso torácico, suporte da sílaba” (CAGLIARI, 2007, p.135). Entretanto, o foneticista afirma que deve haver uma programação da duração das sílabas antes da articulação dos sons. Caso haja um erro nesse momento, ressalta o autor, teremos um caso de gagueira.

As moras são unidades de percepção da duração das sílabas e dos segmentos. No caso das sílabas, as moras medem a extensão das pulsações torácicas e, no caso dos segmentos, o intervalo entre duas proeminências vocálicas. A mora também pode medir pausas breves correspondentes à realização de uma sílaba ou unidade.

Os pés e intervalos caracterizam, respectivamente, línguas de ritmo acentual e línguas de ritmo silábico. Os pés dizem respeito ao intervalo de duração entre duas sílabas tônicas e determinam o aparecimento de um ritmo do tipo fixo. Os intervalos, por sua vez, representam o segmento que compreende duas sílabas tônicas. Os intervalos caracterizam o aparecimento de um ritmo do tipo variado.

O grupo tonal (GT), por sua vez, é maior do que os pés. Cada GT pode ser formado por um ou mais de um pé. Cada pé apresenta uma sílaba tônica saliente, ou seja, que contém a informação mais importante do enunciado.

As pausas também são consideradas como uma unidade rítmica, podendo ocorrer em várias posições do enunciado. De acordo com Cagliari, as regiões em que ela ocorre podem diferenciar significados, sendo necessárias para que o ritmo não seja desfeito. Segundo o autor, elas preenchem o vazio sonoro “entre o que foi dito antes e o que se lhe segue” (CAGLIARI, 2007, p.135).

O impulso e o repouso também proporcionam um ritmo à fala. Conforme Cagliari, as sílabas apresentam um movimento ondular composto por um impulso, por um pico e por um repouso. “Os impulsos variam de intensidade de acordo com a tonicidade silábica provocada pelas pulsações torácicas. Impulso e repouso produzem um fluxo rítmico que pode ser analisado em termos da tonicidade das sílabas e das moras das unidades” (CAGLIARI, 2007, p.139).

O *icto* e a *rêmis* são, respectivamente, a saliência tônica e o repouso nos segmentos átonos. Nos enunciados, essas unidades rítmicas unem-se em um fluxo contínuo, incorporando também as pausas.

Por fim, a *ársis*, representada pelas saliências e a *tésis* pelos vales correspondem “ao efeito final da somatória de todos os parâmetros da fala, produzindo um movimento ondulatório ou fluxo rítmico” (CAGLIARI, 2007, p.140). Elas são responsáveis por produzir uma sensação própria, que é gerada por meio da superposição das outras unidades rítmicas (CAGLIARI, 2007).

É importante não confundirmos ritmo com velocidade. Conforme Massini-Cagliari e Cagliari (2008), o ritmo refere-se ao modo como as línguas organizam a substância fonética no tempo a partir da proeminência das sílabas e acentos. Porém, o mesmo padrão rítmico, segundo eles, pode ser pronunciado com uma velocidade de fala maior ou menor.

No processo de produção da linguagem o ritmo, bem como a velocidade são essenciais. O falante pode alterar a velocidade de fala como uma estratégia discursiva, como por exemplo, diminuir a velocidade a fim de permitir a maior compreensão do seu ouvinte ou ainda, aumentar a velocidade, de modo a ganhar o turno conversacional. Por outro lado, com o ritmo isso não é possível, uma vez que a organização e compreensão da fala ficam comprometidos. A partir disso, ao considerarmos que o discurso é também estratégico, apresentamos, de modo sucinto, o ponto de vista cognitivo a partir dos pressupostos de Willem Levelt (1989).

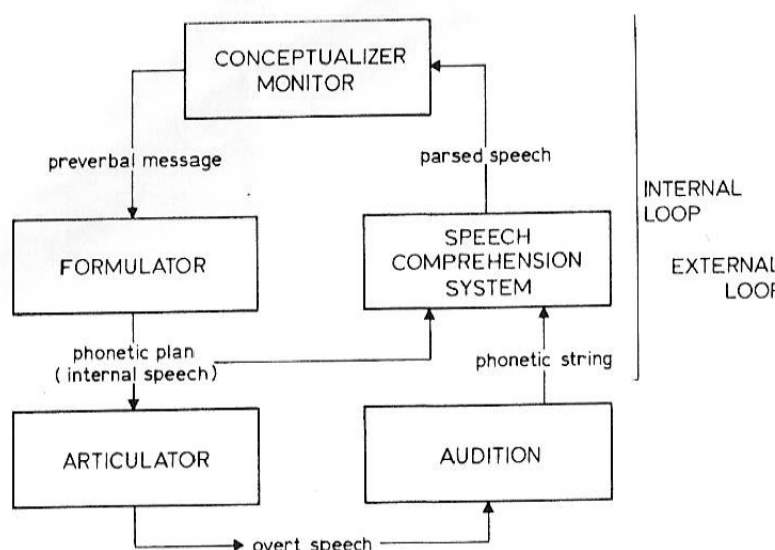
1.8 O modelo de levelt (1989)

A partir dos trabalhos de Garret (1975, 1980), Levelt (1989) desenvolve um modelo psicolinguístico de produção de fala. De acordo com Brum-de-Paula (2005, p. 7), o autor enriqueceu a proposta de Garret,

[...] criando um modelo de produção da linguagem capaz de dar conta das diferentes operações mentais realizadas pelo falante quando intenções, pensamentos e sentimentos são transformados em discurso articulado.

Observe, a seguir, o modelo de auto-monitoramento da fala proposto pelo autor.

Figura 3 – Levelt e o modelo de Auto-monitoramento da fala.



Fonte: Levelt (1989, p. 470).

Levelt (1989) divide o processo de produção de fala em 1) conceitualizador; 2) formulador e 3) articulador. A etapa referente ao conceitualizador caracteriza o primeiro mecanismo de produção de fala. É nesse momento que a mensagem é produzida e o falante, através da monitoração, consegue acessar o discurso antes que ele seja produzido, conseguindo, conseqüentemente, detectar o erro e corrigi-lo. Nessa etapa, através do *External Loop* o falante monitora a sua fala antes da produção e ainda a controla por meio do *Internal Loop*.

A construção da mensagem pré-verbal é o primeiro passo para a produção da fala. Ela tem início a partir de alguma intenção comunicativa. A partir dessa intenção, o falante selecionará a informação desejada. Com a continuidade do discurso, os objetivos intermediários podem ser acionados ou bloqueados. A maior tarefa do falante será, então, manter o controle do discurso enquanto constrói a mensagem.

De acordo com Levelt, a intenção comunicativa é a responsável pelos atos de fala. O caminho percorrido da intenção comunicativa até a mensagem envolve mais de um passo. Frequentemente, um objetivo inicial de fala é expandido em subobjetivos. Por exemplo, se a intenção for informar uma rota para determinado lugar, o falante deve criar todo um plano. Ele terá que planejar e ordenar os vários objetivos secundários que serão acionados, e, para

cada um deles, terá que decidir qual o ato de fala que deverá ser utilizado, ou seja, se fará uso de uma asserção, de uma interrogação, entre outras.

Para cada ato de fala utilizado há o planejamento da informação que será expressa pelo interlocutor, a fim de atingir o objetivo inicial. No entanto, algumas informações podem ser repassadas sem que tenham sido explicitamente formuladas. Isso só é possível porque o falante conta com a capacidade cooperativa do interlocutor em inferir no objetivo do enunciado, que traz apenas uma fração da afirmação que foi repassada.

Na segunda fase do processo de produção da fala ocorre a codificação gramatical. É a etapa em que se inicia a estrutura sintática e fonológica da mensagem, bem como o recebimento do ritmo e da prosódia. É o momento da construção do plano fonético. Levelt supõe que o processo formulativo é derivado do léxico. Para ele, a codificação gramatical e fonológica são medidas por entradas léxicas. O autor denomina essa suposição de Hipótese Lexical. De acordo com essa hipótese, a mensagem produzida pelo informante não pode ser desencadeada por uma forma sintática específica, como uma construção passiva, mas por outro lado, deve ser mediada por itens lexicais, que por meio de suas propriedades gramaticais levam o código gramatical a gerar uma estrutura sintática específica.

Conforme Levelt, o léxico mental do falante é um repositório de conhecimento sobre as palavras da sua língua. Do ponto de vista da produção da linguagem, cada item do léxico é uma lista de pelo menos quatro tipos de características. Há, primeiro, a especificação do termo, que é dada pelo significado. Nessa etapa há um conjunto de condições que devem ser preenchidas na mensagem para que um determinado item seja selecionado. Por exemplo, na entrada *eat* o significado é algo para comer por prazer ou para nutrição. O segundo item seria um conjunto de propriedades sintáticas que inclui a categoria de cada entrada, a de *eat*, por exemplo, seria uma entrada verbal. A terceira etapa diz respeito à especificação morfológica do item, como a raiz que no caso da palavra *eat* mostra o tempo verbal e a pessoa gramatical. A quarta etapa, por sua vez, seria a forma fonológica que inclui o acento e a sílaba: no caso de *eat* teríamos uma palavra monossilábica, com uma sequência de vogais e consoantes.

Levelt relaciona a velocidade de fala à teoria do acesso lexical. Conforme o autor, uma pessoa com velocidade de fala normal produz aproximadamente 150 palavras por minuto. Por outro lado, se estiver sobre pressão, essa taxa pode dobrar. Segundo Levelt, em média, um falante de inglês, adulto, possui um vocabulário ativo de aproximadamente 30.000 palavras, mas esse número pode variar de acordo com o informante. No dia a dia, o falante escolhe

entre essas palavras continuamente de duas a cinco vezes por segundo. Contudo, diz o autor, a questão está em saber como a velocidade e a escolha precisa entre esses lemas é feita.

Em uma fala fluente a taxa de erro é considerada baixa. Alguns estudos, segundo Levelt, mostram que a maioria dos erros são lexicais. Conforme o autor, eles estariam relacionados aos desvios para a obtenção dos lemas. Os erros são causados pela incorporação conceitual (*conceptual intrusion*) e incorporação associativa (*associative intrusion*). A primeira ocorre quando o lema selecionado é perturbado por atividades simultâneas de dois ou mais conceitos. A segunda, quando a seleção do lema é interferida por associações entre os lemas (ou palavras associadas).

Levelt propõe três classes de erros: 1) o *blend*, que seria a substituição e a troca de palavras. Nesse caso, teríamos duas palavras condensadas em uma, ou seja, dois lemas são recuperados e competem pelo mesmo espaço sintático; 2) a substituição, que pode ser de vários tipos, entretanto a mais frequente é a que envolve um tipo de relação semântica entre o alvo e o erro e 3) a troca de palavras, que é resultado de fragmentos de diferentes mensagens que foram ativadas ao mesmo tempo. Esse evento ocorre devido à simultaneidade de acesso de diferentes lemas por diferentes fragmentos da mensagem.

A última etapa é a articulatória. Nela, há o acionamento dos órgãos responsáveis pela produção de fala³². A articulação está relacionada à execução motora do plano fonético. *Before it becomes articulated, a phonetic plan can be temporarily buffered.* (LEVELT, 1989, p. 454).

Para que a fala seja iniciada uma quantidade mínima do plano fonético deve ser reunida e entregue ao articulador. No entanto, a articulação de uma palavra não pode ser iniciada antes que a sua primeira sílaba seja liberada. Antes disso, a segunda parte da sílaba deve estar disponível. Isso é uma condição necessária para a articulação fluente das sílabas das palavras. Quanto maior o número de itens no *buffer*, mais difícil fica a recuperação de cada um deles. O tempo extra, necessário para recuperar o próximo item é frequentemente garantido pelo alongamento do item da última sílaba. Na fala fluente, o papel do *buffer* articulatório pode ser bastante limitado.

Sendo assim, para que o falante inicie o processo de articulação da primeira palavra do enunciado é necessário que a mensagem que será comunicada seja gerada. Seguido disso, o falante gera a estrutura da superfície completa para a mensagem. E, depois, começa a construir o plano fonético para o enunciado. É durante esse processo que as disfluências

³² Para mais informações conferir o item 5.1.2 referente ao processo articulatório.

acontecem no discurso. Conforme Levelt, a codificação de mensagens, a formulação e a articulação podem ocorrer paralelamente, mas há uma diferença na frequência e nos períodos do enunciado que estão sendo construídos.

Levando em consideração essas três fases, o discurso representará o que o falante pretende comunicar. Conforme Levelt, durante a conversação o falante busca em sua memória cognitiva o que deseja enunciar. Retomando o exemplo em que o falante tem como intenção ilustrar determinada rota, o autor explica que durante o assunto, o falante busca em sua memória cognitiva pistas relevantes do seu mapa cognitivo da cidade, que fica armazenado em sua memória. A partir disso ele é capaz de recuperar pontos estratégicos da rota, como igrejas, comércios e bancos para informar o interlocutor.

O autor afirma que há sugestões de que em longos monólogos os falantes alternam lentamente entre fases que gastam muita atenção na informação recuperada e fases em que se concentram para finalizar a mensagem. Isso leva à produção de pausas³³ no discurso, que ocorrem, principalmente, em longos períodos de fala ou em monólogos. Alguns autores, como Henderson *et al* (1966) e Goldman-Eisler (1968) (cf. Levelt, 1989) sugerem que a distribuição das pausas na fala deve-se à alternância no ritmo e à preocupação do falante com a elaboração do objetivo e da informação que deve ser recuperada.

³³ De acordo com Levelt (1989), o período de silêncio não deve durar mais do que 200 ou 250 ms. Falaremos mais detalhadamente sobre isso no capítulo 7, que trata da metodologia.

2 QUADRO TEÓRICO

2.1 O modelo autossegmental métrico

O objetivo deste capítulo é fazer uma apresentação do modelo desenvolvido por Pierrehumbert (1980). Para a análise dos dados desta tese juntamos a versão de 1980, com a modificação feita por Beckham e Pierrehumbert de 1986, adaptando os modelos de acordo com os nossos objetivos³⁴. Feito isso, faremos uma breve introdução ao modelo de Halliday (1970), uma vez que a divisão dos enunciados em grupos tonais seguiu a teoria desenvolvida por ele.

Em 1980, com a tese “The phonology and phonetics of English Intonation” Pierrehumbert marcou o início de um novo período dentro da pesquisa entoacional ao apresentar um modelo que separava a representação fonológica da implementação fonética. O modelo proposto pela autora teve sua evolução a partir de contribuições da fonologia gerativa.

O gerativismo nasceu com Noam Chomsky, em 1957, com o lançamento do livro *Syntactic Structures*. Com o intuito de opor-se ao modelo behaviorista existente na época até então, o modelo procura investigar os aspectos fonológicos das línguas a partir do seu componente fonológico, como um conjunto de subsistemas em interação, cada um deles apresentando aspectos particulares e específicos.

Uma gramática que parte dos princípios gerativistas tem como objetivo geral definir e elaborar regras para a construção de enunciados, tal qual como é proposto no modelo de Pierrehumbert, já que ela parte de uma abstração fonológica e, para dar conta dessa abstração, ela precisa investigar as regras que mapeiam as representações fonológicas em representações fonéticas. É importante salientar que por ter a formulação de regras como objetivo principal, a fonologia gerativa clássica impediu o desenvolvimento de mecanismos apropriados à análise prosódica (MATEUS, 2001). Foram essas limitações que propiciaram o surgimento da teoria autossegmental, iniciada por Goldsmith.

Assim, quando se faz uma análise fonológica seguindo a abordagem de Pierrehumbert (1980), é essencial levar em conta os pressupostos gerativistas. Pois, de acordo com Cagliari (2012), a análise entoacional de um enunciado nesse modelo exige também uma análise da grade métrica. Pela grade métrica definem-se as sílabas tônicas, descrevendo a relação de força entre elas, e os pés métricos, em um nível superior.

³⁴ As adaptações utilizadas serão apresentadas no capítulo referente à metodologia.

O modelo de Pierrehumbert (1980) é conhecido como teoria métrica autossegmental (AM) da entoação, rotulado primeiramente por Robert Ladd. Dentro dessa abordagem, a interpretação dos tons vem de análises acústicas que podem também ter uma avaliação auditiva.

Sobre essa teoria, Ladd (1996) afirma que ela possui como objetivo fonológico ser capaz de caracterizar os contornos adequadamente em uma cadeia de elementos categoricamente distintos e, como objetivo fonético, proporcionar um mapeamento dos parâmetros acústicos, a partir dos elementos fonológicos.

Segundo o autor, a teoria métrica-autossegmental recebe esse nome porque, do lado autossegmental, trata das propriedades fonéticas e fonológicas dos tons e da relação dos *pitch accents*³⁵ com a cadeia segmental. Já do lado métrico, é uma teoria geral das relações sintagmáticas na fonologia.

De acordo com Ladd, a teoria AM entende a entoação a partir de quatro princípios gerais. O primeiro deles diz respeito à linearidade da estrutura tonal, que consiste na caracterização do contorno entoacional em uma sequência de eventos tonais, como os acentos melódicos associados às sílabas proeminentes da cadeia segmental e aos tons adicionais, associados a pontos específicos da cadeia segmental. O segundo princípio diz respeito à distinção entre o *pitch accent* e a tonicidade (*stress*). O *pitch accent* refere-se às pistas perceptuais concretas para o acento ou proeminência de determinadas sílabas ao longo do enunciado e está ligado à organização prosódica. A tonicidade está relacionada à proeminência melódica das palavras, que são determinadas por outros parâmetros, como duração e intensidade, além da frequência fundamental (F0). A tonicidade, segundo o autor, deve ser entendida como pistas acústicas. O terceiro princípio, por sua vez, consiste na análise dos eventos entoacionais em níveis tonais, de modo que os *pitch accents* e os tons de fronteira (*edge tones*) são descritos por meio dos tons alto (H *high*) e baixo (L *low*), ou a partir da combinação desses tons. O *pitch accent* é considerado pelo autor como o evento fonológico mais importante em que um contorno melódico pode ser analisado. Por fim, o último princípio está ligado à tendência da F0 em um enunciado, como a declinação. De acordo com Ladd, esses são os princípios básicos para uma boa análise entoacional.

A abordagem de Pierrehumbert foi construída a partir da teoria métrica de Liberman (LIBERMAN, 1975; LIBERMAN e PRINCE, 1977) e sobre a análise de Bruce (1977) para o acento de palavra do sueco.

³⁵ O *pitch accent* refere-se a um elemento entoacional saliente como, por exemplo, um pico na curva melódica. Manteremos o termo inglês, já que se trata de um termo técnico de uma teoria específica.

O pioneiro nesse modelo de análise fonológica foi John Goldsmith (1976), denominando-o de “autossegmental”. Para o autor, a fonologia autossegmental é uma tentativa de fornecer um entendimento mais adequado do lado fonético da representação linguística. O modelo foi utilizado primeiramente para a descrição dos tons nas línguas tonais e, posteriormente, deu origem à Geometria de Traços (cf. CLEMENTS, 1985).

Atualmente, o modelo de Goldsmith para as línguas entoacionais tem sido muito utilizado, principalmente depois dos trabalhos de Pierrehumbert (1980), t' Hart *et alii* (1990), Ladd (1996), Hirst e Di Cristo (1998), entre outros. Descrições da entoação do português nesse modelo têm sido feitas por Frota e Vigário (1999), Moraes e Stein (2006), Fernandes (2007), entre outros.

2.1.1 Características do modelo teórico

No que diz respeito propriamente às características teóricas do modelo de Pierrehumbert (1980), podemos dizer que a autora desenvolveu uma teoria simples e eficiente. A abordagem é uma representação abstrata que foi desenvolvida para estudar a entoação do inglês britânico (*standard*). Nesse modelo a entoação é vista como um fenômeno fonológico. Assim, por meio da descrição fonológica ela explica vários detalhes sobre a medição instrumental da frequência fundamental, estabelecendo novos padrões fonéticos. Com isso, Pierrehumbert evidencia os diferentes padrões entoacionais de um texto, a fim de mostrar como um mesmo padrão pode ser implementado em textos com padrões acentuais distintos. Pois, de acordo com ela, a mesma melodia alinha-se a uma mesma sentença de diferentes modos, o que corresponde aos diferentes modos de se atribuir os acentos frasais (*phrase accents*)

A diferença no modo de representação de um enunciado está na intenção do falante, no contexto em que o enunciado foi dito. Para outro falante a notação pode ser completamente diferente. Tudo dependerá do propósito comunicativo. Por isso, há um número infinito de pronúncias de um mesmo enunciado. Assim, Pierrehumbert propõe uma série de regras que transformam essas representações em realizações fonéticas, já que segundo ela, essas regras explicam as características quantitativas da F0 e também as características que podem ser consideradas qualitativas, como a configuração geral da curva entoacional.

Para descrever o padrão fonológico proposto a autora leva em consideração três fatores 1) o gramatical 2) a grade métrica e 3) a frequência fundamental (F0). O componente gramatical é responsável pelos tons da frase entoacional. As unidades fonológicas são geradas a partir de uma estrutura gramatical de estado finito, que define os acentos melódicos e que

está alinhada à sílaba tônica. Eles tomam como base o padrão métrico do texto e mais dois tons adicionais que caracterizam a entoação no final do enunciado (os *pitch accents* e os *edge tones*). A grade métrica determinará quais sílabas são acentuadas e quais não são, além de descrever a relação de força entre as sílabas tônicas. A frequência fundamental, por sua vez, alinha os tons com o texto, de acordo com a estrutura do tom e a representação métrica do texto, o que explica a relação do controle do padrão entoacional com a variação melódica.

A partir disso, Pierrehumbert propõe o seu sistema de notação, fazendo assim, uma associação fonológica entre os tons e o texto. Em seu modelo, as unidades fonológicas são descritas pelos tons H (*high/alto*) e L (*low/baixo*) e pelas possíveis combinações entre eles (*bitonals*). Esses tons são gerados por um sistema fonológico, como parte da gramática da língua, não há outros tons além desses. O não reconhecimento de um tom intermediário, nesse caso, o médio, traz alguns problemas para o modelo teórico da autora, pois, algumas vezes, ela acaba traduzindo um tom médio como um tom alto ou baixo.

Ao trabalhar com dois tons, Pierrehumbert procura fazer um apanhado entre os modelos anteriores de quatro níveis tonais, como os de Trager e Smith (1951), Pike (1945) e Liberman (1975) e teorias que procuraram explicar a entoação em termos de mudança de F0 ao invés de níveis de F0, como a de Bolinger (1951, 1958), Ladd (1980) e Clark (1978). De acordo com a autora, uma teoria que leva em consideração os alvos tonais é interessante porque facilita a descrição de como o mesmo padrão entoacional alinha-se com os diferentes textos além de ajudar a resolver muito dos problemas com as teorias de quatro tons, e de ter a vantagem de descrever a entoação em sequências de níveis tonais. Para Pierrehumbert, a anotação com quatro níveis tonais gera uma ambiguidade, pois confunde diferenças tonais com diferenças na escolha do *pitch range* (tessitura, escala melódica). Gussenhoven (2002) corrobora com a afirmação de Pierrehumbert, pois, para ele, a representação por dois tons torna mais clara a distinção entre representação fonológica e o contorno melódico, uma vez que além de descrever somente o que é linguisticamente possível, caracteriza os contrastes da língua.

Para a descrição entoacional, Pierrehumbert faz uma distinção entre o *stress* (tonicidade da sílaba ou ritmo), o *tune* (tom), o *phrasing* e o *pitch range* (tessitura). O *stress* refere-se à tonicidade da sílaba no enunciado ou ao padrão rítmico. A sílaba tônica saliente, que contém a informação principal é denominada por ela de *nuclear stress*. A tonicidade contém informações sobre a duração, amplitude e características espectrais dos segmentos da fala. O tom é a origem abstrata da frequência fundamental. Por exemplo, uma diferença entre

os tipos de frase (afirmativa, declarativa, interrogativa, entre outras) é uma diferença de tom. O padrão de tonicidade é independente do tom, já que um determinado tom pode ser aplicado a diferentes padrões entoacionais e um padrão tônico pode ser produzido com diferentes tipos de tons. O tom também ajuda a marcar as fronteiras dos enunciados, indicados por meio da duração de um padrão e por uma pausa.

O *phrasing*, por sua vez, refere-se ao modo como um enunciado complexo é dividido. Cada frase entoacional (*intonational phrase*) ajuda na escolha de um tom. De acordo com Beckman e Pierrehumbert (1986) há dois níveis frasais: a *intermediate phrase* (frase intermediária) e a frase entoacional. A frase entoacional pode ser formada por uma ou mais frases intermediárias.

O *pitch range* diz respeito à tessitura, que pode variar de acordo com a intenção do falante. Por exemplo, um falante pode aumentar o seu tom de voz com a intenção de enfatizar uma informação específica em um enunciado. A tessitura, no contexto geral do enunciado, interage com o significado dos tons, fornecendo a sua interpretação no contexto.

Segundo Hirschberg e Pierrehumbert (1990), o abaixamento final e a tessitura estabelecem uma relação com a interpretação entoacional. “*They are especially important in conveying the hierarchical segmentation of the discourse*”³⁶ (HIRSCHBERG; PIERREHUMBERT, 1990, p. 279).

Além disso, as autoras (*op. cit.*) observaram que o abaixamento final reflete o grau de finalidade de um enunciado, isto é, quanto maior for a diminuição da curva entoacional no final do enunciado, maior será a sensação de que o enunciado chegou ao fim, completou um tópico.

Dentro de sua abordagem, o contorno entoacional é formado por uma sequência de *pitch accents* e *edge tones*. Os *pitch accents* são caracterizados por um tom H ou L, ou pela combinação entre eles. O tom principal do *pitch accent* é simbolizado por um asterisco (*) – *starred tone* – e representa a sílaba tônica frasal, o foco entoacional do enunciado como, por exemplo, H* e L*. O *pitch accent* também pode ser formado por um tom secundário, que marca a sílaba átona. Ele pode vir antes ou depois do tom com asterisco. Esse tom é indicado por um hífen sobrescrito (H⁻ L⁻). Atualmente, a maioria dos pesquisadores não faz uso desse diacrítico, pois já pressupõe-se que a sílaba que segue ou precede a sílaba com asterisco (tônica) seja átona.

³⁶ Eles são especialmente importantes em transmitir a segmentação hierárquica do discurso. (Tradução nossa).

Os *edge tones*, por sua vez, estão divididos em acentos frasais e tons de fronteira. Os acentos frasais podem ser formados por um tom (H ou L) ou por um par de tons (H+L). As sílabas do acento frasal nunca serão marcadas com um asterisco, pois elas são menos salientes que os *pitch accents*. Esse tom é atribuído entre o último acento frasal e o tom de fronteira, onde ocorre a variação mais notável da F0. A importância fonética do *acento frasal* é o controle da F0 entre o último *pitch accent* da frase intermediária e o começo da próxima, ou o final da frase entoacional.

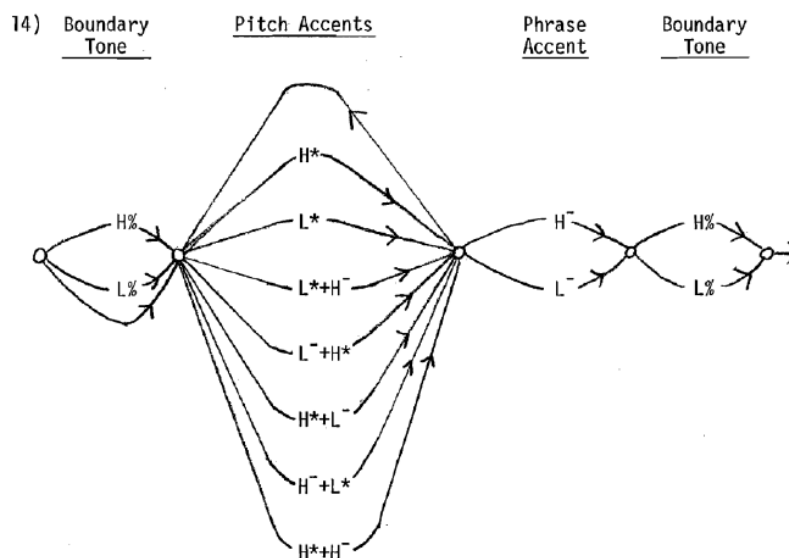
Os tons de fronteira estão associados ao final da frase entoacional e referem-se ao valor fonológico de um tom. Eles são indicados por um símbolo de porcentagem (%) ao lado de cada unidade tonal (H% ou L%) e sempre serão formados por um único tom. Esses tons estão associados às fronteiras dos domínios prosódicos e não estão relacionados à marcação de proeminência. O tom de fronteira ocorre no acento frasal, independentemente do padrão entoacional anterior. O acento frasal, por sua vez, localiza-se logo após o acento nuclear. Pode acontecer de um tom de fronteira ocorrer no início de um enunciado. Nesses casos, ele será atribuído quando o valor da F0 for alto, de tom H inicial e contorno descendente. Na maioria das vezes, o tom de fronteira inicial ocorre quando o enunciado é muito longo e o foco está no fim da frase.

O símbolo de porcentagem (%) também é utilizado para indicar uma marcação de pausa no enunciado, como no exemplo: “*This is the cat % that ate the rat % that stole the cheese*” (PIERREHUMBERT, 1980, p. 21).

Em seu modelo, os diacríticos (*, %, e -) propostos não estão relacionados ao valor tonal, ou seja, os tons do tipo H*, H- H% são igualmente tons altos, não há uma diferença de altura entre eles, mas, sim, na maneira como eles estão associados ao texto. Com essas informações, a autora propõe então uma gramática de estado finito da entoação, que gera, no enunciado, um conjunto de sequências tonais bem definidas. Observe a seguir o padrão de notação simplificado³⁷ do modelo teórico:

³⁷ O padrão mostrado na figura é simplificado, pois não há identificação do acento frasal extra utilizado com as *tag questions*.

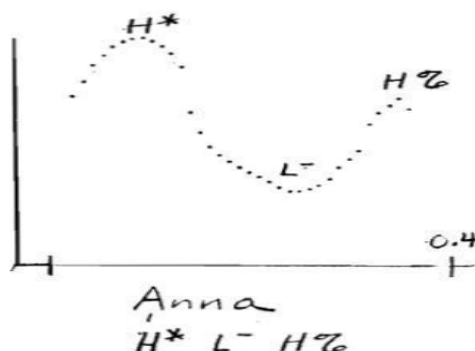
Figura 4- Representação simplificada do modelo autosssegmental de Pierrehumbert.



Fonte: Pierrehumbert, 1980, p. 29.

Nessa gramática, os tons são formados por um ou mais *pitch accents*, que vêm seguidos de um acento frasal e um tom de fronteira obrigatórios. Os tons de fronteira ocorrem independentemente do padrão entoacional que os precede e do acento frasal, que é localizado logo depois do acento nuclear. Os tons de fronteira iniciais também estão representados. Normalmente, eles serão ascendentes. Por outro lado, se o tom de fronteira final L% for antecedido por acento frasal do tipo L, teremos um contorno entoacional do tipo descendente. Já se o tom de fronteira final L% for precedido por um acento frasal do tipo H, teremos um contorno entoacional nivelado. Os *pitch accents*, os acentos frasais e os tons de fronteira são, de acordo com essa gramática, os tons bem formados para a frase entoacional. Veja como esses três elementos distribuem-se no enunciado:

Figura 5- Distribuição do *pitch accent*, acento frasal e tom de fronteira no enunciado “Anna”.



Fonte: Pierrehumbert, 1996, p. 255.

Pierrehumbert destaca sete possibilidades de *pitch accents*: H^* , L^* , L^*+H^- , L^-+H^* , H^*+L^- , H^-+L^* , H^*+H^- . Entre eles estão representados os tons simples (*single tones*) e os tons compostos (*bitones*). Os tons simples são equivalentes aos pés monossilábicos e os tons compostos assemelham-se aos pés bissilábicos, no sentido em que eles são classificados por uma relação de força. Eles representam um espraiamento dentro de uma sílaba. Em relação a esses eventos tonais, o tom H^* é caracterizado pelo pico local e vem alinhado à sílaba tônica. Já o tom L^* caracteriza o vale local. Paradigmaticamente, L^* é mais baixo que H^* , no mesmo contexto. O valor fonético de L^* diminui se a proeminência aumenta. Nas mesmas circunstâncias, ocorre o aumento de H^* . A diminuição de L^* e o aumento de H^* não são simétricos, uma vez que L e H estão em diferentes relações com o falante.

A autora não atribui padrões sintáticos aos tons, já que eles são extremamente dependentes do contexto. A sua intenção é descrever a entoação a partir da observação das características da frequência fundamental (F0). Para ela, não interessa se a frase é afirmativa ou interrogativa, mas, sim, o que o falante fez e não o que o falante quis dizer com o enunciado (CAGLIARI, 2011).

Esses eventos tonais geram regras de reajuste entoacional. Em seu modelo, essas regras não estão baseadas em um sistema de derivação (o que seria esperado, uma vez que seu modelo é de base gerativa), mas sim em regras locais contexto-sensitivas que mapeiam a sequência dos tons de acordo com os valores quantitativos que determinam o contorno da frequência fundamental (F0). Esse sistema representacional é chamado de *mapping* e caracteriza um padrão fonológico entoacional que [...] é atribuído a um enunciado (*text*) através de uma representação (*mapping*) que configura o enunciado entoacionalmente

(CAGLIARI, 2011, p.12). Essas regras que se aplicam da direita para a esquerda e incluem movimentos de *downstep* e *upstep*, regras de espreadimento (*spreading*), de formação de *plateau* e de implementação das linhas de base (*baseline units*). Elas são específicas da língua inglesa e estão mais ligadas aos tons do que aos padrões entoacionais.

As duas regras mais discutidas pela autora são as de *downstep* e de *upstep*. *Downstep* é um termo técnico que se refere à queda melódica que ocorre do começo ao fim do enunciado de modo contínuo. *Upstep*, por sua vez, é o contrário do que ocorre com o *downstep*, isto é, ocorre com movimento crescente da curva melódica de modo contínuo no enunciado. É uma regra local e está relacionada ao final do enunciado. Nessa regra há um aumento do L ou H que vier depois de um acento frasal do tipo H-.

Além do modelo apresentado, há mais duas versões para a análise entoacional de Pierrehumbert. A primeira foi a modificação feita por ela em colaboração com Mary Beckman (BECKMAN; PIERREHUMBERT, 1986; PIERREHUMBERT; BECKMAN, 1988), que teve como maior modificação a questão da análise das regras de *downstep*. O novo ponto de vista está no fato de que o fenômeno é desencadeado por meio do *pitch accent* e não pela alternância entre um tom alto e baixo (H L H). A segunda é uma versão modificada e simplificada da revisão que faz parte do sistema ToBI – *Tones and Break Indices* – (SILVERMAN *et al.* 1992), uma proposta padrão desenvolvida por um grupo de pesquisadores para a rotulagem das características prosódicas da fala digital, a partir da base de dados do inglês.

2.2 O modelo de Halliday (1970) e Cagliari (2007)

M.A.K. Halliday (1970/1973) descreve a entoação a partir de um modelo de gramática funcional³⁸. O autor desenvolveu seu modelo a partir da análise do inglês britânico, mais especificamente com a variedade do inglês RP (*received pronunciation*) do sudoeste da Inglaterra.

Para fazer uma análise a partir dessa metodologia é preciso levar em consideração quatro unidades, o grupo tonal, os pés, as sílabas e o fonema.

Os pés (P) são descritos de acordo com o ponto de vista de Abercrombie (1964), ou seja, como uma “[...] *unit of rhythm in English, and has a structure of two elements, ‘ictus’*

³⁸ Não iremos entrar em detalhes sobre o método desenvolvido pelo autor, uma vez que a análise entoacional deste trabalho seguiu a abordagem teórica de Pierrehumbert (1980) e Pierrehumbert e Beckham (1986). Sendo assim, para maiores detalhes sobre o modelo teórico de Halliday conferir os trabalhos: Halliday (1970, 1973); Cagliari (1982, 2007) e Carvalho (2013).

and ‘*remis*’³⁹ (HALLIDAY, 1963, p. 106). O fato de o pé ser a unidade de ritmo e não a sílaba refere-se, segundo Abercrombie, ao “*stress-timing*”. Um pé é marcado pelo espaço entre dois *ictus*. A unidade abaixo dos pés é formada pelas sílabas, que podem ser fracas ou fortes. A sílaba forte opera como um *ictus* e a sílaba fraca como uma *remiss*. A sílaba forte representa a tônica e a sílaba fraca a átona.

Todo pé apresenta, pelos menos, uma sílaba acentuada e um número relativo de sílabas fracas. Se um pé for seguido de uma pausa ou iniciar um grupo tonal, ele será considerado silencioso e caracterizado como incompleto. Por outro lado, os pés completos são aqueles que não apresentam uma sílaba átona no início de um grupo tonal.

A estrutura do grupo tonal (GT) é operada pelos pés. Assim como o pé, cada grupo tonal apresenta dois elementos em sua estrutura. O elemento tônico e pretônico. O elemento tônico é obrigatório e está presente em cada grupo tonal. Já o elemento pretônico é opcional.

A variação melódica do componente pretônico em um grupo tonal se realiza devido às características do tom da STS. Essa sílaba se individualiza pela maior variação do contorno melódico, ou seja, por apresentar a marca entoacional mais proeminente do enunciado, compreendendo a parte do enunciado que o falante julga como sendo a mais importante (o foco do GT). A STS, além de definir o GT, também contribui para a definição dos tons secundários, variantes dos tons básicos ou primários (CAGLIARI, 2007, p.174).

Os grupos tonais podem ser simples ou compostos. Um GT é considerado como simples quando apresenta somente uma sílaba tônica saliente. Nesse caso, ocorre apenas uma mudança de contorno melódico notável (CMN). Um CMN, por sua vez, é definido como um padrão do tipo que caracteriza um tom no sistema. Veja um exemplo de GT simples, a partir do enunciado abaixo (CAGLIARI, 2007, p. 165):

(a) //Eu não /acho que/ seja as/sim// (falando com certeza)⁴⁰.

Por outro lado, um GT composto acontece quando há duas sílabas tônicas salientes, apresentando, portanto, duas mudanças notáveis de contorno melódico, como na frase a seguir, também retirada de Cagliari (2007, p. 165).

(b) //Eu não/acho que/seja as/sim// (falando com hesitação, dúvida).

Os pés são marcados por uma isocronicidade fonológica, em outras palavras, por uma regularidade nas batidas tonais, já que “[...] *there is a tendency for strong syllables to occur*

³⁹ “[...] como uma unidade de ritmo em inglês, formada por uma estrutura de dois elementos ‘*ictus*’ e ‘*remis*’ [tradução nossa].

⁴⁰ As sílabas tônicas que representam a mudança no CMN aparecem sublinhadas nos exemplos (a) e (b)

at roughly regular intervals of time whatever the number of weak syllables, including zero, in between"⁴¹.

Juntamente com a atribuição de valores aos pés rítmicos, Halliday descreve os padrões de tonicidade, de tonalidade e dos tons. Esses três sistemas são independentes um do outro, mas não o são do sistema rítmico. A tonicidade está relacionada com a acentuação e é caracterizada por ser uma propriedade das sílabas e não das vogais. É responsável por definir o ritmo e marcar o ponto focal de cada unidade de atividade, relacionando-se, dessa forma, com o tema e o rema em um enunciado. A tonalidade, por sua vez, descreve os grupos tonais e suas relações sintáticas e semânticas, marcando um tipo de unidade em atividade na língua. Por fim, os tons marcam o tipo de atividade envolvida e descrevem os padrões entoacionais (padrões fonológicos e pragmáticos), caracterizando a variação melódica de um enunciado. É um padrão complexo, feito de uma posição simples entre polaridade e não polaridade, ou seja, pelo que é oposto, como, por exemplo, quando uma pergunta exige respostas do tipo sim ou não.

Diferentemente de Pierrehumbert (1980) o autor propõe o uso de cinco níveis de altura melódica: alto, meio-alto, médio, meio-baixo e baixo. Não há outros tons fora dessa pauta. A pauta entoacional pode ser usada tanto em análises auditivas quanto em análises acústicas. Neste último caso, os padrões entoacionais se adaptam à tessitura. Os padrões são chamados tons entoacionais ou simplesmente tons e são identificados através de números: 1, 2, 3, entre outros.

Halliday (1970) também não faz referência a tons mais baixos do que o tom baixo, nem a tons mais altos do que o tom alto. Às vezes, a variação da altura melódica nos tons alto e baixo pode apresentar um intervalo grande, mesmo sem uma influência específica da variação da tessitura. Nestes casos, o ouvido ajusta essas alturas a um valor linguístico determinado, igual ao dos tons alto ou baixo. Esse ajuste mostra a ação da interpretação fonológica sobre dados fonéticos de natureza física, detectados por meios técnicos de análise acústica. A escolha de um tom pode relacionar-se com os tipos de oração (afirmativas, interrogativas, entre outras), com as noções de modalidade (possibilidade, probabilidade, entre outras), com os atos de fala (como ordem, pedido, etc) ou ainda, de acordo com o comportamento linguístico do falante, bem como com suas atitudes (indiferença, polidez, surpresa, entre outras).

⁴¹ “[...] há uma tendência para que as sílabas fortes ocorram em intervalos de tempo mais ou menos regulares.

É importante salientar que Cagliari (1982/2007) adaptou o modelo de Halliday (1970) para os padrões entoacionais do português brasileiro (dialeto paulista). No entanto, essa modificação diverge muito pouco da análise feita pelo autor para o inglês britânico. Por exemplo, no que concerne aos tons primários, Halliday (1970; 1973) trabalhou com cinco tons primários simples e dois compostos, como mostrado. Cagliari (1982, 2007), por sua vez, propôs seis tons primários simples e três tons compostos.

Diante dessas considerações metodológicas Pierrehumbert foi responsável por inaugurar um novo período na pesquisa entoacional, trazendo um modelo que separa a representação fonológica da sua implementação fonética. Tanto assim, que nos dias atuais, a maior influência para os trabalhos em fonologia entoacional tem sido a sua tese de doutorado, mesmo que muitos estudiosos tomem como base apenas as ideias gerais do modelo, deixando de lado alguns pontos importantes para a análise. Nas palavras de Silverman *et al* (1992, p. 868), Janet Pierrehumbert “*have been particular influential over the last decade in such areas as speech synthesis, relationships between prosody and discourse, and laboratory phonology*”⁴².

Por sua vez, o trabalho com a transcrição dos tons a partir do modelo de análise proposto por Halliday não é uma tarefa fácil, uma vez que essa abordagem, além de exigir um treinamento auditivo muito específico, também requer uma habilidade fonética. Segundo Cagliari (2011), o treinamento exigido é semelhante ao que é feito para transcrição dos tons nas línguas tonais ou para a transcrição das qualidades vocálicas, que segue o método proposto por Abercrombie (1967) para as vogais cardeais. De acordo com Cagliari (2011, p.6),

Esse treinamento acostuma o ouvido com as variações de altura melódica, à moda das notas musicais, o que permite ao linguista transcrever os tons entoacionais com uma simples análise auditiva. Sem isto, obviamente, como acontece com as notas musicais, a pessoa se sentirá confusa e pode achar que tal empreitada é impossível, dada a variação que ouve.

São essas questões, portanto, que nos levaram a escolher esses dois modelos teóricos para a análise deste corpus. O modelo de Pierrehumbert (1980, 1986) é, atualmente, um dos mais utilizados em análises fonológicas, o que permite uma melhor interpretação dos dados pelo leitor. Além disso, é mais viável analisar as disfluências comuns e gags por meio dessa metodologia, devido às quebras constantes na curva entoacional. Uma interpretação feita

⁴² “têm sido especialmente influente ao longo das últimas décadas, em áreas como síntese de fala, nas relações entre prosódia e discurso e em fonologia de laboratório”[Tradução nossa].

apenas pelo modelo de Halliday (1970) e Cagliari (1982/2007) gera um desconforto no leitor, uma vez que exige um maior treinamento acústico. Mesmo assim, esse modelo teórico não poderia ser deixado de lado, pois a análise destes dados é primeiramente auditiva e depois acústica.

3 METODOLOGIA

3.1 Informantes

Participaram desta pesquisa seis informantes gagos e seis informantes não gagos. O tamanho da amostra é suficiente para o desenvolvimento desta pesquisa e justifica-se por dois motivos: o primeiro deles diz respeito à quantidade de enunciados segmentados para os informantes. Por exemplo, para os informantes gagos, a quantidade de enunciados variaram entre 38 e 72 grupos tonais (GTs). Já para os informantes não gagos essa variação ficou entre 34 e 54 GTs. O segundo motivo refere-se propriamente à análise acústica, que exige um cuidado maior, principalmente no que concerne à fala dos informantes gagos.

O parâmetro idade e sexo não foram levados em consideração nesta pesquisa. No grupo dos informantes gagos, foram gravados cinco falantes do sexo masculino e um do sexo feminino, sendo a mesma proporção para o grupo dos não gagos. A similaridade nessa variável não passou de uma coincidência, no entanto, ela confirma os dados da literatura especializada, no caso do primeiro grupo, que dizem que nas pessoas do sexo masculino a predominância aos distúrbios da fala é maior. A idade dos informantes gagos variou entre 18 e 65 anos e, a dos não gagos, entre 22 e 46 anos. A igualdade entre a idade dos informantes não foi objetivo desta tese.

No que diz respeito aos informantes gagos, é importante salientar a dificuldade de acesso a eles. Primeiro, porque eles configuram uma classe não tão comum na sociedade. E, segundo, porque muitos deles apresentam dificuldade de aceitação da patologia. Por isso, a participação do fonoaudiólogo foi muito importante nesta etapa. Foi ele quem fez o intermédio com alguns dos informantes antes da gravação, uma vez que já tratava ou tinha tratado de alguns dos informantes com o desvio de fala. Os outros informantes, que não foram identificados pelo fonoaudiólogo, relataram a patologia e mencionaram que já haviam sido diagnosticados por um clínico especializado como portadores de gagueira. Quanto aos informantes não gagos, foi perguntado a eles se possuíam alguma dificuldade na fala ou se já haviam sido diagnosticados com algum distúrbio de fala. Todos eles não se referiram a nenhuma dificuldade ou desvios de fala.

Vale acrescentar que os diferentes graus de gagueira, que vão desde a disfluência comum até graus mais avançados, relatados pela literatura especializada (GUITAR, 2006),

não foram levados em consideração para a gravação. Apenas levamos em conta o fato do informante ser ou não portador do desvio de fala.

Todos os informantes são falantes nativos do português brasileiro. Esse é um fato relevante, uma vez que são poucos os estudos, no campo linguístico, em língua portuguesa (PB), que fizeram uma comparação entre uma fala com disfluência gaga e outra com disfluência comum. Quanto aos dialetos, apenas um informante do grupo dos não gagos é falante do dialeto paulista. Um outro nasceu no estado de São Paulo, mas vive na cidade de Andradas há 21 anos. Já no grupo dos gagos, um informante é falante do dialeto carioca e o outro, do paulista. No entanto, o primeiro reside no interior de Minas Gerais (cidade da coleta de dados) há cinco anos e o segundo, há 38 anos. Todos os informantes não apresentam conhecimento linguístico sobre o assunto.

Quanto à escolarização, tanto no grupo dos informantes gagos quanto dos não gagos, o grau de instrução variou entre ensino médio incompleto e superior completo. Mais uma vez, essa questão não foi um parâmetro para a análise dos dados.

Todos os informantes receberam e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE).

Abaixo, seguem informações sobre os sujeitos que participaram desta pesquisa

3.1.1 Os informantes gagos

AP⁴³, 52 anos, nasceu no interior de São Paulo, em São José do Rio Pardo. Reside em Andradas (MG) há 38 anos e é de profissão vendedor. Relatou ser gago desde a infância, mas diz não precisar a idade que os pais aperceberam-se do desvio. O informante possui Ensino Médio incompleto.

PA, 18 anos, nasceu em Andradas e é estudante. Ele diz que sua mãe percebeu que ele era gago aos seis anos de idade, aproximadamente.

MAU, 65 anos, nasceu no Rio de Janeiro e vive em Andradas há cinco anos. Ele relata que já viveu em outras cidades do Brasil e também no Chile. Atualmente, ele é aposentado e possui Ensino Médio incompleto. O informante também conta que a disfluência gaga foi percebida quando ele tinha cinco anos, por sua mãe. Ele diz que a idade contribuiu, em seu caso, para uma maior aceitação da sua condição.

⁴³ A abreviação é a mesma utilizada para a identificação dos informantes durante a análise dos dados.

VA, 51 anos, nasceu em Andradadas. É de profissão escriturário e possui Ensino Médio completo. O informante não sabe precisar quando identificou a gagueira, mas também relata que foi na infância.

CAM, 22 anos, nasceu em Andradadas e é de profissão manicure. A informante possui Ensino Médio completo e relata que a sua gagueira foi percebida aos três anos. Ela diz que os períodos de ansiedade, nervosismo e depressão acentuam os momentos de disfluência.

JO, 53 anos, também nasceu em Andradadas. É engenheiro civil e empresário. O informante possui Ensino Superior completo e diz não precisar o período em que a gagueira teve início, mas também diz que foi na infância. Assim como MAU, relata que a experiência o ajudou a lidar melhor com os desvios da fala e contribuiu para a sua melhor aceitação.

Dentre os informantes gagos, apenas AP seguiu com o tratamento fonoaudiólogo. Os outros informantes tiveram acompanhamento clínico por um período, mas não deram continuidade. Todos relataram dificuldade em conviver com a disfluência. Além disso, todos relacionaram o aumento da gagueira em momentos de tensão.

3.1.2 Os informantes não gagos

MA, de 59 anos, nasceu em Espírito Santo do Pinhal (SP), mas vive em Andradadas há 35 anos. É técnico em prótese dentária e possui Ensino Médio completo.

CA, de 49 anos, nasceu em Andradadas. É de profissão gerente comercial, e possui Ensino Médio completo.

CL, de 45 anos, também nascida em Andradadas, possui Ensino Médio incompleto, e não desenvolve uma atividade profissional.

LE, de 44 anos, nasceu em Andradadas. O informante possui Ensino Médio incompleto e, atualmente, é auxiliar de classificação de café.

NE, de 23 anos, nasceu em Espírito Santo do Pinhal (SP), mas vive em Andradadas desde os três anos. Possui graduação em Biologia. No entanto, exerce a profissão de Gerente de Relacionamento.

Por fim, BR, de 28 anos, nasceu em Andradadas. O informante possui graduação em odontologia, exercendo a profissão de dentista.

Em relação aos informantes não gagos, percebemos que os que estão mais habituados com situações que exigem um pouco mais de interação conversacional, devido à profissão que exercem, sentiram-se mais confortáveis com a gravação e, conseqüentemente, produziram menos disfluências. O tema da gravação também contribuiu para que as disfluências

aumentassem ou diminuíssem, devido à familiaridade ou não com o assunto. Quanto ao grupo dos não gogos, esse fato também foi observado.

O início da gravação também exigiu atenção, pois alguns informantes, principalmente do grupo dos falantes com disfluência gaga, relataram que estavam nervosos com a gravação e, por isso, estavam gaguejando mais. No entanto, no decorrer da gravação a ansiedade dos informantes dos dois grupos diminuiu e a gravação continuou normalmente. É importante lembrar, mais uma vez, que questões relacionadas à emoção, ansiedade e nervosismo, não são o foco desta tese, mas eles indicam a maior produção de disfluências, seja ela gaga ou não.

3.2 MATERIAIS E PROCEDIMENTOS

3.2.1 A coleta de dados

Os dados foram registrados com um minigravador de voz digital, com microfone embutido, da marca Sony (modelo ICD-PX312). A gravação do *corpus*, para os informantes gogos, foi processada na clínica de uma fonoaudióloga. A intenção inicial era utilizar a cabine de audiometria da sala do consultório, uma vez que seria o ambiente de melhor isolamento acústico. No entanto, como os informantes sentiram-se desconfortáveis no local e, sendo a proposta deste trabalho formar um *corpus* de fala semiespontânea, a opção foi fazer uma gravação mais informal, como uma conversa informal, um diálogo, a fim de deixá-los mais confortáveis com a situação e, concomitantemente, nos aproximarmos o melhor possível de uma fala espontânea. Mesmo assim, a gravação foi feita na sala do consultório. É válido salientar que apenas dois, dos seis informantes não foram gravados no ambiente relatado. Isso aconteceu porque um deles tinha sofrido uma fratura na perna e o outro não tinha condições de ir ao local.

Quanto aos informantes não gogos, as gravações foram realizadas em ambientes com o mínimo de ruído possível, para não interferir na qualidade do áudio. O ambiente não foi o mesmo para os informantes gogos e não gogos, pois os primeiros já tratavam ou haviam sido tratados pela fonoaudióloga, o que facilitou e permitiu o contato com eles. No entanto, o objetivo foi o mesmo para os dois grupos, ou seja, uma entrevista informal.

Os sujeitos não receberam informações sobre o que deveriam fazer. Foi informado apenas que estávamos desenvolvendo uma pesquisa para a faculdade. Em seguida, procuramos incentivá-los a falarem sobre assuntos pessoais, como algo que tivesse marcado de modo positivo ou negativo suas vidas. Também tratamos de assuntos que estavam em pauta nos jornais, no período (janeiro a julho de 2015), como temas relacionados com a

seleção brasileira, times de futebol, o incêndio na boate Kiss, no Rio Grande do Sul, entre outros.

As gravações foram processadas acusticamente com o software PRAAT, desenvolvido por Paul Boersma e David Weenink (1992-2014, versão 5.3.70), do *Institute of Phonetic Sciences (University of Amsterdam)*, acesso disponível gratuitamente em <http://www.fon.hum.uva.nl/praat/>, e, em seguida, foram analisadas auditivamente pela pesquisadora, para que fosse possível verificar os momentos de disfluência.

Na sequência, o texto falado foi segmentado em enunciados correspondentes aos grupos tonais, e de acordo com os eventos de interesse (presença de pausa, inserções, repetições, alongamentos, entre outros). Também foi feita uma transcrição ortográfica dos enunciados.

3.2.2 Dados da análise acústica

Para a interpretação dos dados, não foi nosso intuito observar um fenômeno em específico. O que fizemos foi verificar os eventos disfluentes nos dois tipos de fala e, em seguida, quantificá-los e analisá-los.

Os GTs variaram entre os informantes, no entanto, esse fato não influenciou a análise. Isso ocorreu por dois motivos: o primeiro, porque alguns informantes apresentaram mais momentos de disfluência do que os outros, o que pode estar relacionado a maiores momentos de tensão. Por outro lado, a menor quantidade de eventos disfluentes também pode estar relacionada à familiaridade dos informantes com o assunto. Vale lembrar, que um dos objetivos da gravação foi abordar temas que fizessem parte do cotidiano dos entrevistados ou que fossem temas atuais nos meios de comunicação. O segundo motivo que levou a essa diferença deve-se ao fato da gravação ter sido mais longa para alguns informantes.

Veja a distribuição dos GTs para os informantes gagos e não gagos na tabela 1 abaixo:

Tabela 1 - Distribuição do número de Grupos tonais por informante.

Informantes com disfluência gaga	Número de GTs	Informantes com disfluência comum	Número de GTs
AP	35	MA	38
PA	38	CA	34
MAU	58	CL	47
VA	66	LE	37
CAM	72	NE	53
JO	44	BR	39

Fonte: própria.

Quanto aos fenômenos disfluentes, sejam eles gogos ou não, foram encontradas as ocorrências que serão apresentadas a seguir. Para alguns eventos foram feitas subdivisões, que foram chamados de disfluência complexa. Consequentemente, os eventos disfluentes que aparecem sem subdivisões foram nomeados de disfluência simples. A representação que será utilizada durante a análise dos fenômenos aparece entre parênteses. Vale lembrar, que a definição dada para cada um dos eventos é apenas um resumo do que pode ser visto no capítulo oito referente à análise dos dados.

As disfluências comum e gaga observadas durante a análise dos dados foram classificadas em simples e compostas. As disfluência simples foram:

- **Inserção** (INS) – Caracteriza-se por fones de preenchimento que variam conforme a língua. Neste trabalho, foram consideradas como inserção as vogais: [ɛ], [i] e os marcadores discursivos como, /né/, /aí/, /então/, /tão/, /sabi/. A opção em considerar essas palavras como inserção deve-se ao fato de elas serem consideradas como recursos fáticos, ou seja, são utilizadas como “retroalimentação do interlocutor”, com o objetivo de manter ou tomar o turno (SCLIAR-CABRAL; MARTIM; CHIARI, 1981). Normalmente, o falante usa essas palavras como uma pergunta retórica ou com a intenção de reafirmar o que foi dito. Elas podem ocorrer no início, meio ou fim de GT. É mais comum que elas ocorram em início e fim de GT do que em meio. Abaixo apresentamos como essa disfluência foi simbolizada juntamente com alguns exemplos:

- INS_p – inserção de palavra (né, aí, então);
- INS_v – inserção de vogal (é, i);
- INS_e – inserção de vogal e palavra (i aí), de modo que juntas caracterizam uma palavra.

- **Pausa** – (PS) As pausas caracterizam o momento de silêncio para a demarcação dos grupos entoacionais (CRUTTENDEN, 1986; CAGLIARI, 1992). Elas podem ocorrer entre os grupos entoacionais, chamadas de pausa demarcativa (CRUTTENDEN, 1986) ou no interior de uma unidade entoacional, denominadas de pausas hesitativas⁴⁴. As pausas foram identificadas como:

- PS_i – pausa silenciosa ou respiratória em início de enunciado;
- PS_e – pausa em meio de enunciado (silenciosa ou respiratória);
- PS_f – pausa silenciosa ou respiratória em fim de enunciado.

⁴⁴ A metodologia utilizada para a anotação das pausas será apresentada na sequência.

- **Alongamento** – (AL) Caracteriza-se por um aumento na duração de um fone. Esse aumento pode ocorrer em início, no meio (nos disfluente gags) ou no fim de palavra e, ainda, em vogais isoladas. Neste trabalho, o alongamento está representado na transcrição ortográfica pelo símbolo de dois pontos (:) depois da palavra alongada;

- **Correção**⁴⁵ – (COR) Ocorre quando o informante erra a produção anterior e a corrige na sequência do enunciado. Em nossos dados dividimos as correções em dois tipos:

- A correção em que o informante mantém parte da informação produzida no segmento com erro na correção;
- A correção em que o informante esquece o segmento com erro, realizando um novo que não apresenta partes do anterior.

- **Repetição** – (RP) É a repetição não intencional de um evento, que pode ocorrer com palavras isoladas ou em sequência, ou em parte de palavras como um fone ou sílaba. Quando as RPs ocorreram em sílaba, a análise acústica foi feita para a sílaba repetida, seja ela dupla ou tripla, e para a sílaba na produção completa, isto é, na sílaba em palavra. Durante a análise dos dados foram identificadas as seguintes repetições:

- 2RP – repetição dupla de palavra;
- 3RP – repetição tripla de palavra;
- 4RP – repetição quádrupla de palavra;
- 5RP – repetição quádrupla de palavra;
- RP_{p2} – repetição de palavra dissílaba;
- 2RP_{p2} – repetição dupla de palavra dissílaba;
- 2RP_{p3} – repetição dupla de palavra trissílaba;
- RP_s – repetição de sílaba;
- RP_{sf} – repetição de sílaba final;
- RP_{sp} – repetição de sílaba em palavra;
- RP_{ps2} – repetição de sílaba em dissílabos;
- 2RP_{ps} – repetição dupla de sílaba em palavra;
- 3RP_{ps} – repetição tripla de sílaba em palavra;
- 4RP_{ps} – repetição quádrupla de sílaba em palavra;
- 3RP_s – repetição tripla de sílaba inicial;
- 3RP_v – repetição dupla de vogal;
- 3RP_v – repetição tripla de vogal;

⁴⁵ Esse fenômeno também é conhecido como “erros de performance” (ZELLNER, 1994).

- 4RP_v – repetição quádrupla de vogal;
- 6RP_v – repetição sêxtupla de vogal;
- 2RP_{pv} – repetição dupla de vogal em palavra;
- RP_{sip} – repetição de sílaba incompleta em palavra;
- RP_{pi} – repetição de palavra incompleta;
- RPs_i – repetição de sílaba incompleta;
- RP_{si3} – repetição de sílaba incompleta em trissílabo;
- 2R_{sip} – repetição dupla de sílaba incompleta em palavra.

Além da repetição de palavra, de vogal e de sílaba em palavra isolada, ocorreu também a repetição de parte do enunciado, que foi assim identificada:

- RP_e – repetição de estrutura;
- RP_{epi} – repetição de estrutura em palavra incompleta
- 3RP_e – repetição tripla de estrutura;
- 4RP_e – repetição quádrupla de estrutura.

• **Bloqueio** – (BLO): evidencia-se pela obstrução temporária de um segmento.

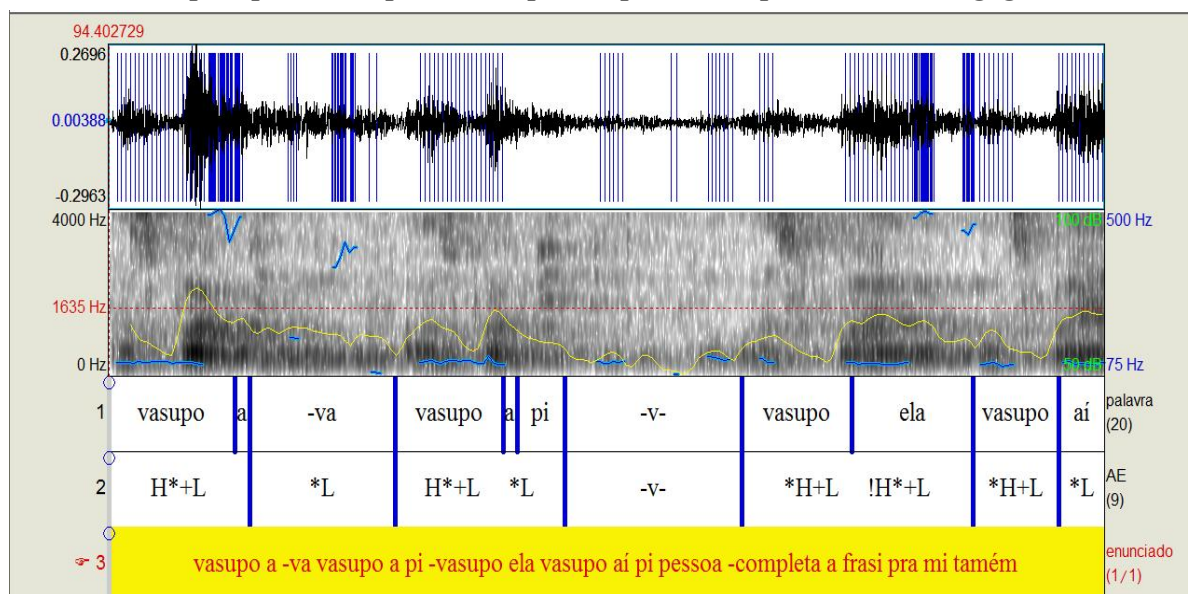
Ele pode ocorrer tanto em vogais quanto em consoantes oclusivas (bilabiais, alveolares e velares), fricativas (labiodentais ou alveolares), nasais (bilabial ou alveolar) ou consoantes laterais (dental) em início de palavra. O bloqueio leva a uma suspensão rápida do enunciado. Em nossos dados foram observados:

- BLO_e – bloqueio em consoante oclusiva;
- BLO_v – bloqueio em vogal.

Através da análise do bloqueio, verificamos, em grande parte das vezes, a ocorrência de um tipo de pausa antes do evento disfluyente, que ocorre junto com o período de soltura de algumas consoantes. As consoantes fricativas, nasais e laterais, bem como as vogais, não apresentam momentos de soltura, mas mesmo assim apresentaram pausas. Diante disso, a nossa intenção será a de observar se a pausa anterior pode ser indicativa ou não de bloqueio. Essa pausa pode ou não ser do tipo silencioso. As PSs silenciosas referem-se a momentos em que não ocorre, auditivamente, qualquer tipo de ruído. Veja na figura 6 que segue um exemplo de pausa em que não há silêncio⁴⁶.

⁴⁶ Os exemplos apresentados neste capítulo são frutos de trabalhos apresentados em Congressos e publicados em Anais de Congressos.

Figura 6 – Representação do bloqueio -v- da palavra *vasupo* no enunciado “vasupo a -va vasupo a pi -v- vasupo ela vasupo aí”, produzido pelo informante gago AP.



Fonte: própria.

A figura 6 é somente uma exemplificação da disfluência, por isso, é apresentado apenas parte do enunciado produzido pelo informante AP. Como vemos, ocorre um bloqueio, com pausa anterior na consoante fricativa labiodental [v], antes da terceira repetição da palavra “vasupo”. O PRAAT considera momentos como esses como períodos de silêncio. No entanto, os pulsos (linhas horizontais em azul) indicam que há participação das cordas vocais no momento que antecede ao BLO. Dessa maneira, não temos um caso de pausa silenciosa, mas sim, situações que indicam uma pausa articulatória, com a configuração de uma articulação fricativa. Em outras palavras, o informante interrompe o enunciado para produzir o bloqueio. Casos desse tipo foram simbolizados com a letra inicial da sílaba bloqueada, no momento da interrupção e entre hifens, como por exemplo, -v-. O bloqueio que não for antecedido por uma interrupção é simbolizado somente por um hífen anterior à palavra, como, por exemplo, -va (que, nessa figura, também indica um BLO). Sendo assim, as interrupções foram classificadas em:

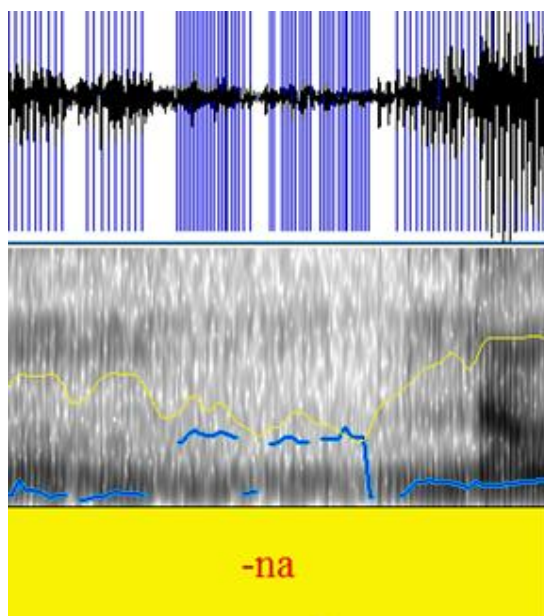
- a) **Bloqueio com pausa silenciosa** – quando não há presença de sonorização, o falante apenas interrompe a fluência por meio de um silêncio. A presença de articulação é indicada por meio do espectro, que pode ser da produção de uma consoante ou vogal que antecede o BLO;
- b) **Bloqueio com pausa respiratória** – quando há sonorização (percepção da respiração durante a análise auditiva);

c) **Bloqueio com pausa articulatória**⁴⁷ – quando ocorre uma PS com presença de pulsos;

d) **Bloqueio com ausência de pausa** – pode acontecer que o informante não apresente nenhum dos três tipos de pausa. Nesse caso, o BLO é identificado pela forma de onda da produção, que para os informantes gogos tende a ser interrompida e para os não gogos, tende a reduzir significativamente.

Observe as imagens 7 e 8, abaixo, que mostram essa diferença no caso do BLO com ausência de PS.

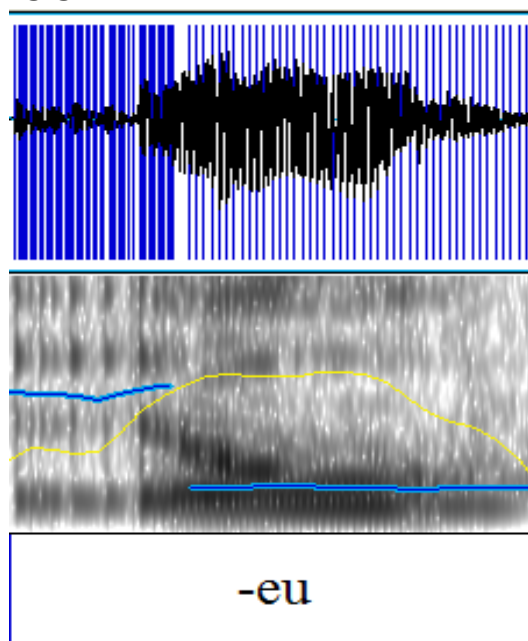
Figura 7 – Exemplo de segmento com redução de forma de onda em BLO sem PS para o informante gago PA.



Fonte: própria.

⁴⁷ É importante deixar claro que há articulação em todas as PSs. No entanto, chamamos essa PS de articulatória devido à presença dos ciclos de onda que não ocorrem nos outros momentos.

Figura 8 - Exemplo de segmento com interrupção de forma de onda em BLO sem PS para o informante não gago CR.



Fonte: própria.

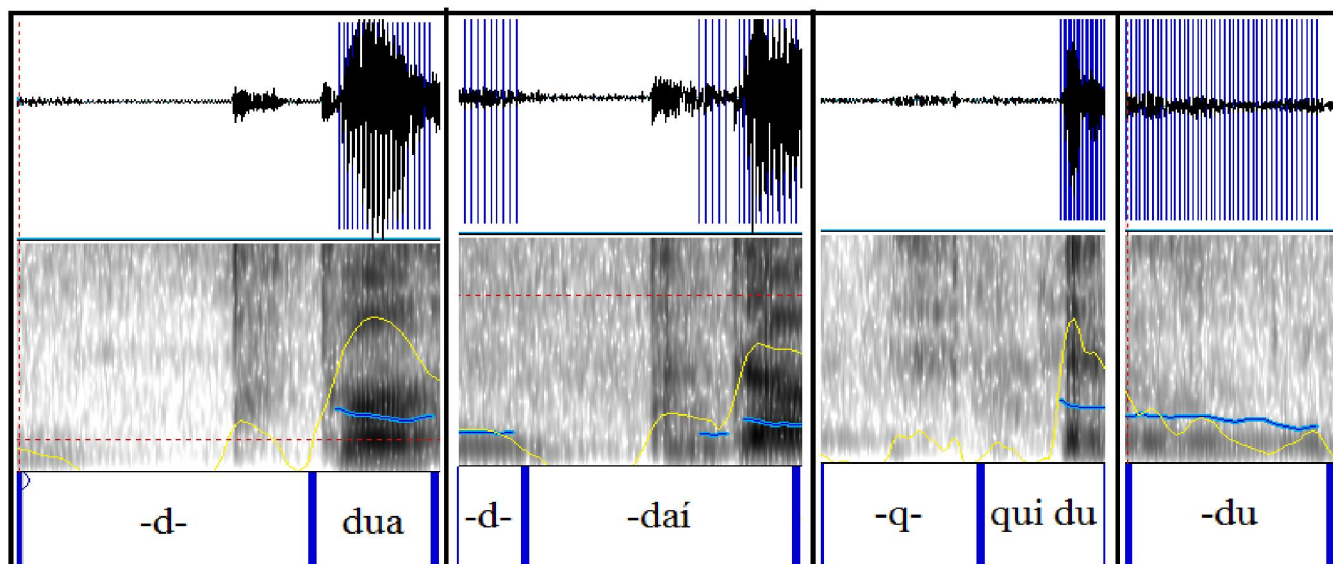
Nas figuras 7 e 8 os informantes gago (PA) e não gago (CR) apresentam diferentes configurações acústicas para o mesmo tipo de PS. Observe que para PA ocorre uma redução significativa na forma de onda enquanto que para CR temos momentos de interrupção devido à soltura das cordas vocais.

Diante disso, consideramos que os BLOs, em consoantes ou vogais seriam antecidos por algum desses três tipos de interrupção. Caso esse fato se justifique, os bloqueios poderão ser melhor identificados, tanto no meio linguístico quanto na literatura especializada. Já que esse tipo de disfluência tem sido verificada apenas auditivamente. Além disso, a confirmação dessa teoria evidenciará que tanto as disfluências comuns quanto as gagas, além de poderem ser utilizadas como estratégias discursivas para a construção do fluxo de fala, também estão relacionadas a questões articulatórias. Sendo assim, resumidamente, podemos ter as seguintes PSs de bloqueio:

- Bloqueio com pausa articulatória (ciclos de onda) – do tipo 1;
- Bloqueio com pausa, mas sem pulsos (nesse caso não ocorre marcação de ciclos de onda) – do tipo 2;
- Bloqueio sem pausa anterior (irregularidades na forma de onda) – do tipo 3;
- Bloqueio com pausa respiratória (presença auditiva de respiração) – do tipo 4.

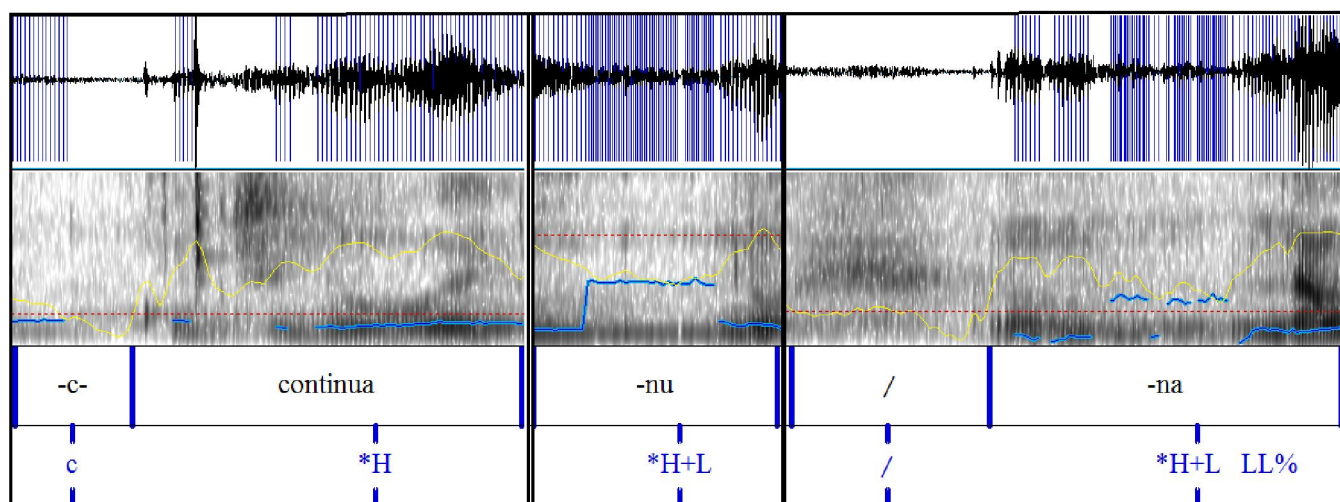
Veja na sequência (figs. 9, 10 e 11) as imagens que ilustram esses quatro tipos de PSs de bloqueio:

Figura 9 – Representação da possibilidade das três ocorrências de bloqueio observadas para as disfluências comuns.



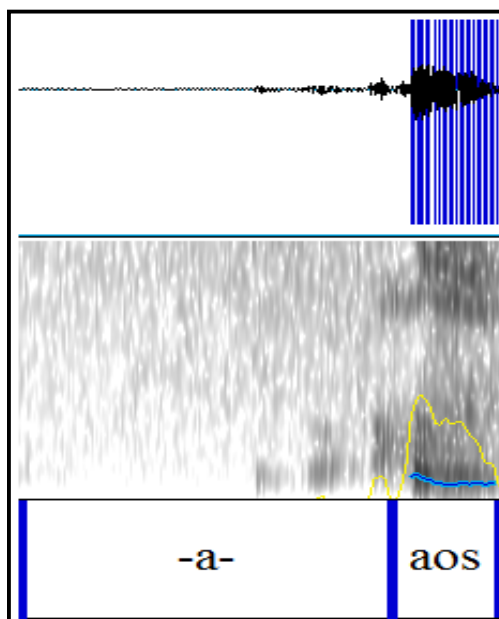
Fonte: própria.

Figura 10 – Representação da PS de bloqueio silenciosa para os informantes não gags.



Fonte: própria.

Figura 11 – As quatro possibilidades de pausas de BLO para as disfluências gagas.



Fonte: própria.

As figuras 9, 10 e 11 representam as possibilidades de pausas de bloqueio observadas para os informantes não gagos (fig. 9 e 10) e gagos (fig. 11), respectivamente.

Na figura 9, da esquerda para a direita estão ilustrados os três tipos de BLOs resultantes dos enunciados analisados. A primeira imagem refere-se ao enunciado EN1 “a perspectiva do café essi anu -continua divagar”, em que ocorre PS articulatória antes do BLO_c. A segunda, ao BLO_c sem PS respiratória ou articulatória no EN26 “/é um -nu sei si é Andradas porque eu nu/ eu convivu aqui”. Por fim, a última ilustra uma PS respiratória em EN31 “i elis num pensam -na união”.

Na figura 10, a PS de bloqueio é silenciosa. Ela está representada pelo segmento /-a-/, referente à vogal bloqueada. Como podemos ver, ela diferencia-se da PS articulatória em relação à presença de pulsos. Já em relação à PS articulatória, a diferença está justamente na articulação. Na PS respiratória (conf. fig. 9c), o espectro não mostra uma estrutura de formantes, diferentemente da figura 10, em que vemos a presença da articulação de /a/ por meio dos formantes.

Há algumas evidências que provam que a primeira PS é articulatória. Por exemplo, notamos que no segmento -c- há a presença dos pulsos, o que indica vozeamento, bem como uma F0 que segue a mesma linha da palavra seguinte “continua”. Comparativamente, na terceira imagem, não há a presença de pulsos e os espectros não são iguais. A análise auditiva

mostra que a respiração é bem clara. Por outro lado, em EN26 não há presença de nenhum dos dois tipos de pausa.

Na figura 11, também da esquerda para a direita, temos uma PS silenciosa, indicada pela primeira imagem, em que ocorre articulação da consoante /-d/. Em seguida, ocorre uma PS articulatória, com presença de pulsos e F0 também para o segmento /-d/. Depois, ocorre uma PS respiratória para o segmento /-q/. Nesse caso, temos uma articulação que não produz efeito articulatório, uma PS do ponto de vista aerodinâmico. Por fim, temos um BLO em que não ocorre nenhum tipo de PS. Como vemos, os momentos de pausa em BLO são do mesmo tipo para os dois grupos de informantes

- **Falso começo** – (FC) Caracteriza o segmento que é interrompido antes da sua conclusão. Quando o falante realiza um FC ele esquece esse elemento, deixando uma sequência incompleta, que não tem sentido com o restante do GT;

- **Palavra Incompleta** – (PI) Caracteriza-se por ser um recurso que o falante lança mão para não gaguejar. Em casos como esse, o informante não termina a produção da palavra, deixando-a incompleta. Diferentemente do FC, ela tem sentido com o restante do enunciado. Quando esse fenômeno ocorre, a duração da palavra é menor;

- **Hesitação** – (HES) Esse tipo de disfluência, caracterizada como hesitação, representa momentos em que o falante apresenta um pensamento suspenso, ou seja, como não conseguiu completar o raciocínio e precisa de algum tempo para colocá-lo em ordem, usa recursos para manter o turno conversacional, como /ah/ e /eh/. Em nossos dados foram verificadas HES apenas do tipo /ah/. Além disso, elas podem ter a função pragmática de surpresa ou ansiedade. Geralmente, ocorrem junto com uma qualidade de voz do tipo *creaky voice*.

Durante a análise dos eventos percebemos que a maioria das ocorrências de disfluência simples não ocorriam isoladamente, ou seja, não tínhamos apenas uma repetição de palavras ou somente uma inserção. Esses eventos vinham acompanhados por outros fenômenos. Por exemplo, a repetição de uma palavra juntamente com a inserção ou inserção de vogal com alongamento, entre outros. Dessa maneira, chamamos essas ocorrências de disfluência complexas, sendo elas⁴⁸:

- INS_v+BLO – inserção de vogal com bloqueio;
- INS_v+AL – inserção de vogal alongada;

⁴⁸ É importante salientar que foram apresentadas apenas as disfluências complexas mais comuns entre os informantes. As disfluências complexas mais específicas de cada falante serão vistas ao longo das análises.

- INS_p+AL – inserção de palavra alongada;
- INS_e+AL – inserção de vogal e palavra com vogal alongada;
- $2INS_v+COR$ – inserção dupla de vogal com correção;
- RP_s+BLO_c – repetição de sílaba com bloqueio em consoante;
- $2RP_p+BLO_c$ – repetição dupla de palavra com bloqueio em consoante;
- $3RP_p+INS_v$ – repetição em palavra trissílaba com inserção de vogal;
- $3RP_p+PS_d$ – repetição em palavra trissílaba com pausa silenciosa

entre as produções;

- RP_e+PI – repetição de estrutura com palavra incompleta;
- $3RP+PI$ – repetição em palavra trissílaba com palavra incompleta;
- $3RPe+PI$ – repetição tripla de estrutura com palavra incompleta;
- $3RPp+PS$ – repetição tripla de palavra com pausa;
- $3RP_v+PS_e$ – repetição tripla de vogal com pausa entre as produções;
- $2RP_v+PS_{2e}$ – repetição tripla de vogal com pausa entre as produções;
- $2RP_v+INS_p$ – repetição dupla de vogal com inserção de palavra;
- $RP_{ps2}+PS$ – repetição de sílaba em palavra dissílaba com pausa silenciosa;
- $2RP_{p2}+INS_v$ – repetição dupla de palavra dissílaba com inserção de vogal;
- $2RP_p+AL_{tp}$ – repetição dupla de palavra dissílaba com alongamento
- $2RP_p+AL+PS$ – repetição dupla de palavra dissílaba com alongamento e

pausa silenciosa;

- $RP_{si}+PS$ – repetição de sílaba incompleta com pausa silenciosa;
- $RP_{pi}+PS$ – repetição de palavra incompleta com pausa;
- $RP_{epi}+INS_p$ – repetição de estrutura com palavra incompleta com inserção

de palavra;

- RP_e+INS_p – repetição de estrutura com inserção de palavra;
- RP_e+BLO_v – repetição de estrutura com bloqueio em vogal;
- RP_e+INS – repetição de estrutura com inserção;
- RP_e+INS_v – repetição de estrutura com inserção de vogal;
- $RPe+BLO+PS$ – repetição de estrutura com bloqueio e pausa silenciosa;
- REP_e+2RP_v – repetição de estrutura com repetição dupla de vogal;
- $2RP+AL_{pm}$ – repetição dupla de palavra com alongamento em meio de

palavra;

- BLO+PS – bloqueio com pausa silenciosa;
- COR+AL+PSe – correção com alongamento e pausa entre as estruturas;
- 3RP+PS – repetição tripla de palavra com pausa entre as produções;
- COR+RP_{pi} – correção com repetição incompleta de palavra;
- COR+PS – correção com pausa silenciosa;
- COR+PI – correção com palavra incompleta
- COR+PS_e+INS_v – correção com pausa entre as palavras e inserção de vogal;
- COR+PS_e+BLO - correção com pausa entre as palavras e bloqueio;
- COR+INS_v – correção com inserção de vogal;
- COR+2INS_p – correção com inserção dupla de palavra;
- COR+INS_v - correção com inserção dupla de vogal;
- COR+2RP_v – correção com repetição dupla de vogal.

3.3 Metodologia para análise dos fenômenos prosódicos e acústicos

Como vimos, Cruttenden (1986) divide as pausas em: silenciosas e preenchidas. As pausas silenciosas estão relacionadas ao silêncio total. Já as pausas preenchidas são as utilizadas para preencher o silêncio como, por exemplo, *hum*. Nesta tese, a pausa preenchida foi tratada como inserção. Foram analisadas em nossos dados as pausas silenciosas e respiratórias em início, meio e fim de GT.

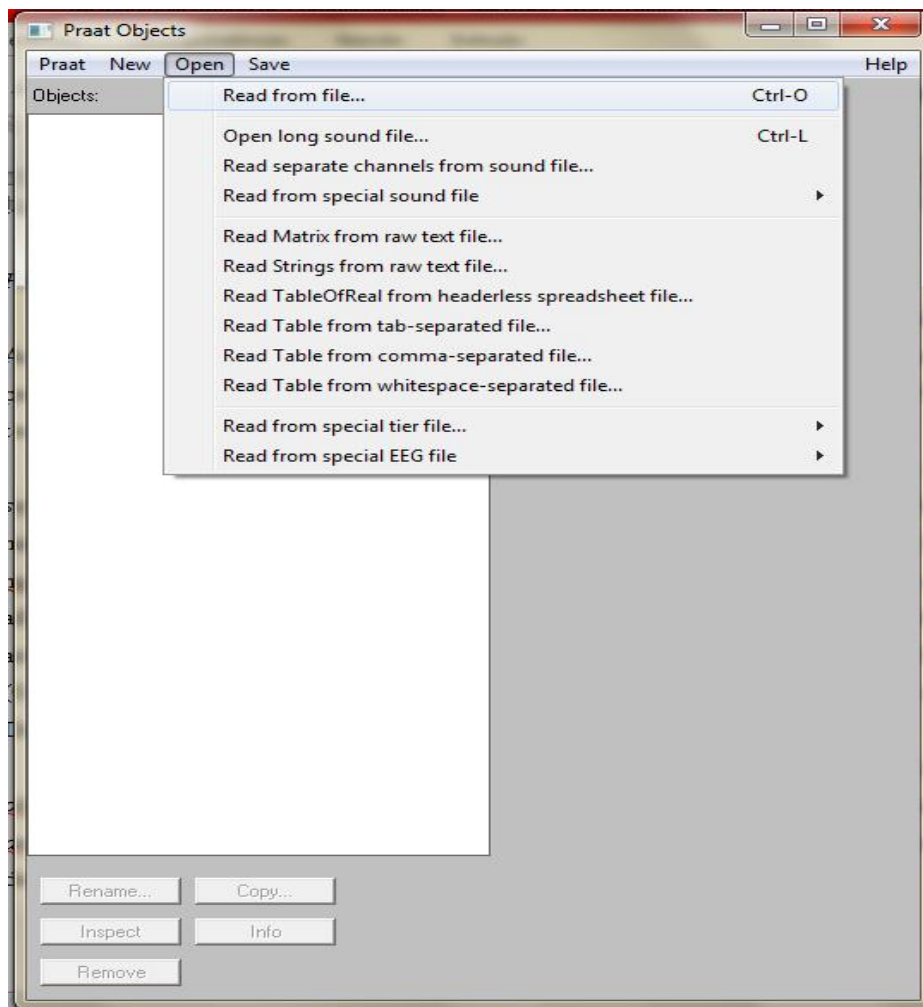
Para a notação das pausas utilizamos o recurso disponibilizado pelo *software*, que permite detectar os silêncios⁴⁹. Segundo Cruttenden (1986), o limiar mínimo para que a pausa seja percebida em diferentes níveis varia de um segundo (1000 ms) até um quarto de um segundo (0,25 segs ou 250 ms). Laver (1995), por sua vez, afirma que o período de fechamento de um segmento plosivo e desvozeado pode chegar até 0,18 segs (180 ms). Levando em conta essas afirmações, o período de silêncio considerado foi de 200ms (0,02 segundos). Essa metodologia já havia sido utilizada por outros autores, como Delfino (2009).

O procedimento⁵⁰ consistiu, então, em: 1) abrir o som desejado para análise no PRAAT *objects*; 2) selecionar, na tela do PRAAT *objects* o enunciado e acionar o comando *annotate > to text grid (silences)* e 3) delimitar a duração do silêncio nos padrões desejados. Veja os três passos a seguir:

⁴⁹ É preciso mencionar que a nomenclatura silêncio deve-se unicamente ao nome da entrada do programa computacional. Quando falamos em silêncio, não estamos nos referindo à ausência de som, mas sim a um período em que não ocorre nenhum tipo de segmento. O PRAAT não diferencia períodos de ausência de som de momentos respiratórios. Isso é feito auditivamente.

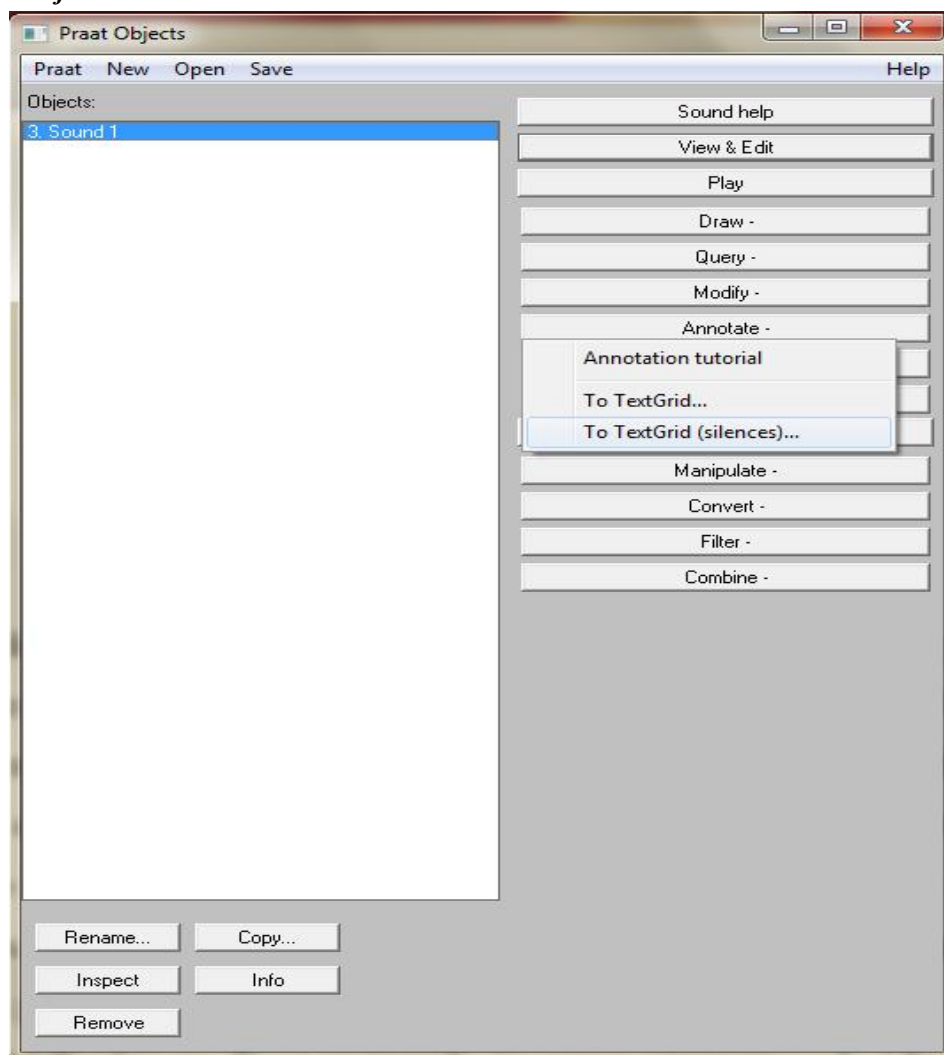
⁵⁰ As figuras 12, 13, 14 e 15 representam o modo de funcionamento do programa computacional PRAAT.

Figura 12 – Primeira etapa: seleção do som desejado para análise na tela do PRAAT *objects*.



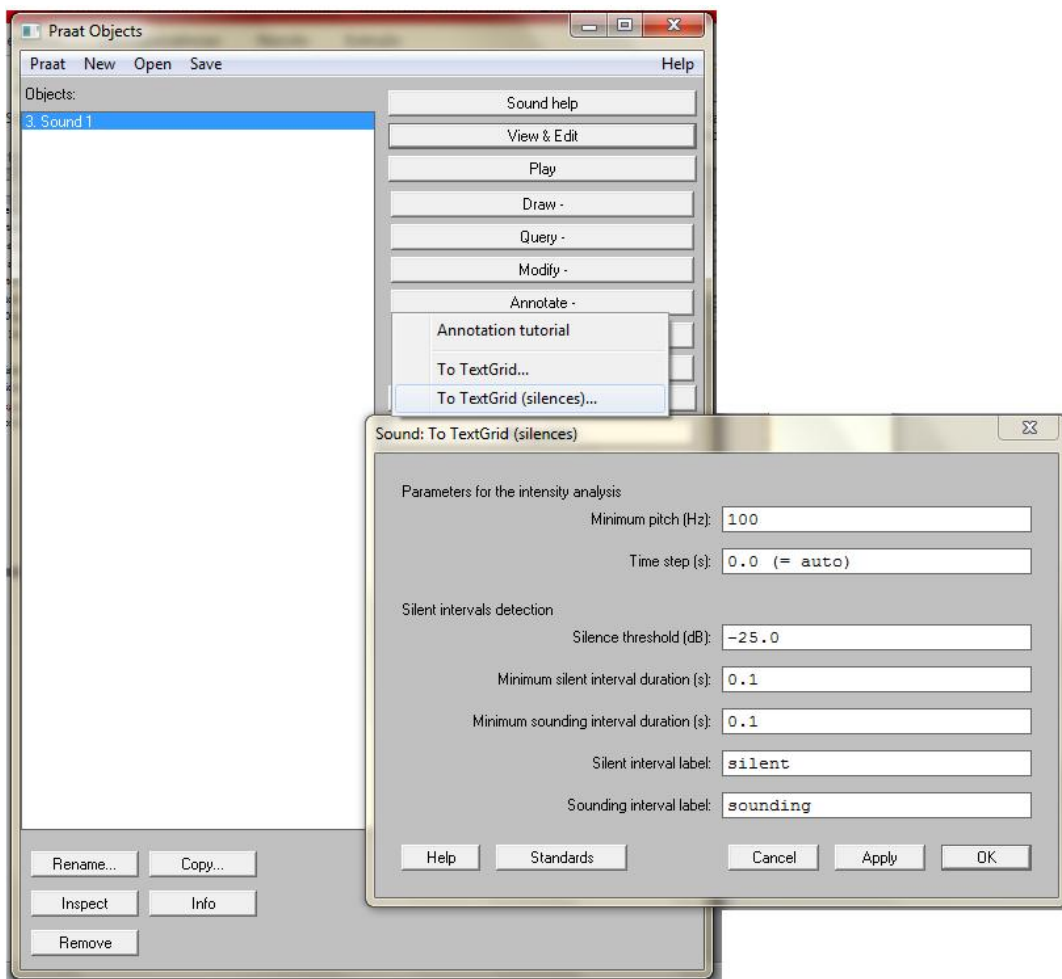
Fonte: própria.

Figura 13 – Segunda etapa: processo para a seleção do silêncio análise na tela do PRAAT *objects*.



Fonte: própria.

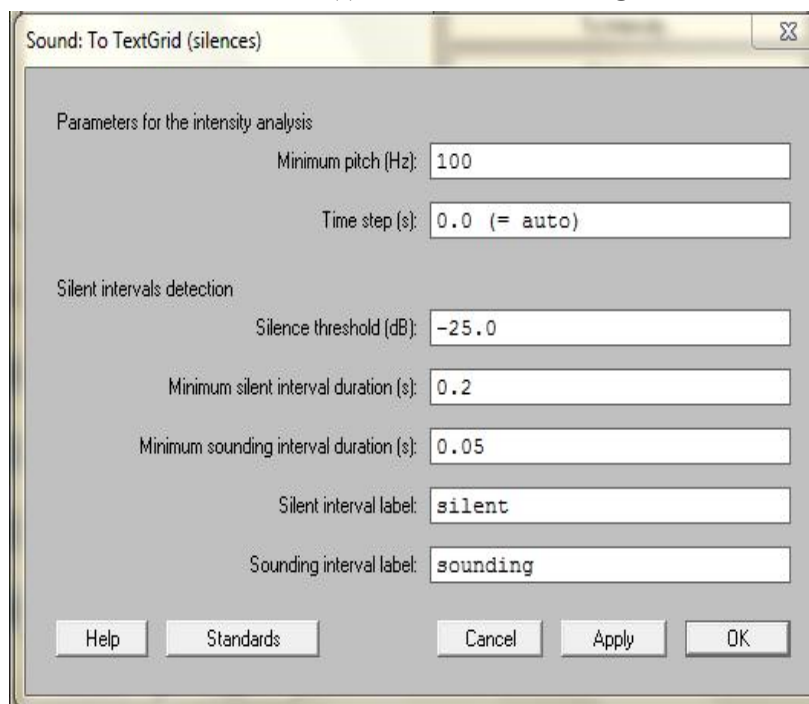
Figura 14 – Terceira etapa: janela do *software* para a notação dos valores referentes ao silêncio.



Fonte: própria

Feito isso, é preciso inserir os valores referentes à detecção do silêncio. O valor para o *minimum silent interval duration* (s), que define o intervalo de silêncio mínimo para que a duração ocorra é de 0,2s (200 ms). E, o *minimum sounding interval duration* (s), que corresponde ao intervalo mínimo para que o silêncio não ocorra foi anotado em 0.05s (50 ms). As outras medidas não precisam ser alteradas. Assim, temos:

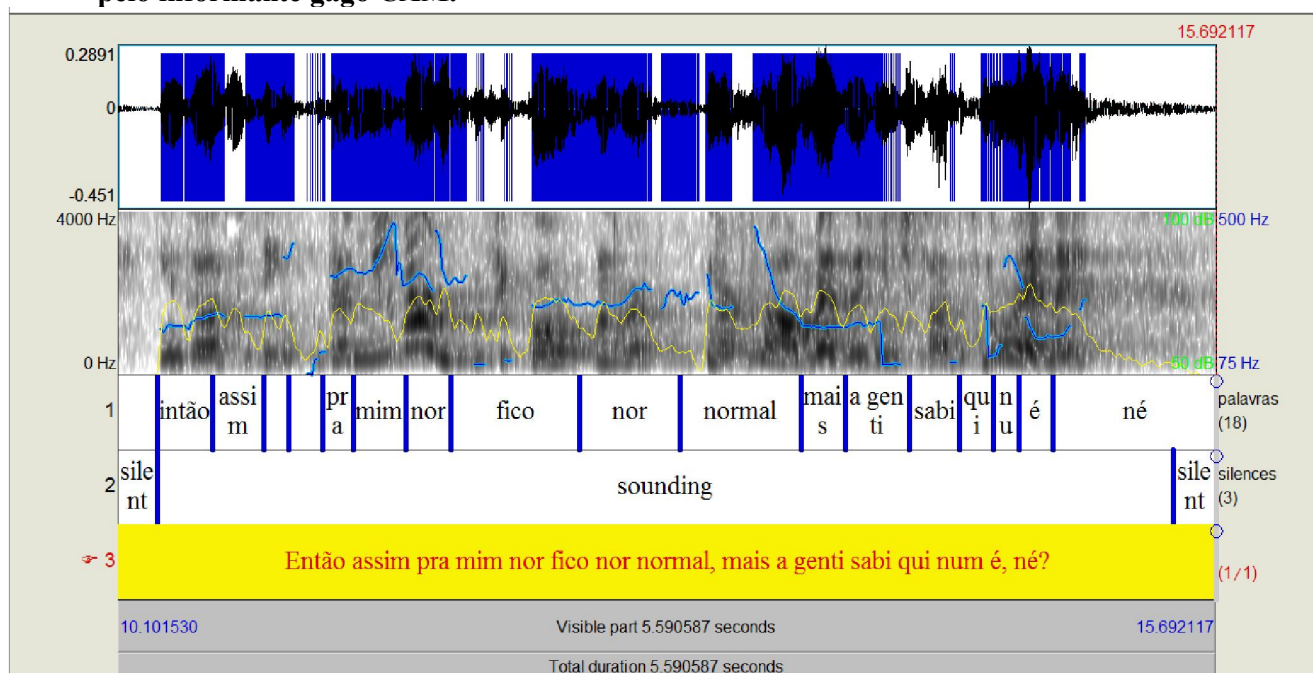
Figura 15 - Tela do PRAAT referente aos valores utilizados para anotação do silêncio em *minimum silent interval duration* (s) e *minimum sounding interval duration*.



Fonte: própria.

Estabelecidas as durações, o *software* cria uma nova camada de *text grid*. A partir disso, é possível selecionar os dois *text grid* e acionar o comando *merge*, no PRAAT *objects*. Esse comando insere a nova camada de silêncio junto com o espectrograma e a camada, já anotada, da segmentação das palavras e do enunciado, gerando a seguinte tela:

Figura 16 - Tela com a identificação dos *tiers* referentes à segmentação das palavras, dos silêncios (*silent*), dos períodos de sonorização (*sounding*) e enunciado completo, na frase “Então, assim pra mim nor fico nor normal, mais a gentis abi qui num é, né”, produzido pelo informante gago CAM.



Fonte: própria.

Na figura 16, as pausas silenciosas aparecem no início e no fim de GT. A primeira com duração de 0.201s e a segunda com duração de 0.219s.

É importante ressaltar que, mesmo utilizando o comando descrito para a detecção dos silêncios, foi preciso conferir auditivamente o resultado, pois houve momentos em que o programa considerou como silêncio o final de segmentos fricativos e períodos iniciais de oclusivas, principalmente, [p] e [t]. Esse critério de anotação das pausas foi importante, uma vez que permitiu verificar em quais momentos os silêncios são mais frequentes.

Depois disso, as PSs foram anotadas no GT de acordo com a sua posição (início, meio ou fim) e se eram silenciosas ou respiratórias. No último caso, essa diferenciação foi feita por meio de uma análise auditiva. Caso fossem silenciosas, não haveria qualquer possibilidade de som, mas se fossem respiratórias, poderia haver um ruído. Depois disso, somamos cada grupo de PSs e observamos a recorrência de cada uma delas no conjunto das gravações. As PSs que ocorreram junto com um evento disfluyente foram observadas separadamente das que antecederam eventos fluentes.

A intensidade⁵¹ também foi anotada para os eventos disfluentes a partir do recurso disponibilizado pelo PRAAT. No *software*, ela vem representada pela linha amarela. Optamos por analisar a intensidade de transição, inicial, final e máxima dos eventos. A primeira, diz respeito ao período que antecede a disfluência, que pode ou não ser disfluente. Com isso, nossa intenção foi verificar se havia um aumento ou diminuição da intensidade na passagem da fluência para a disfluência. Os eventos disfluentes que ocorreram em início de GT ou foram antecidos por uma PS não foram considerados. Vale salientar que, em grande parte das vezes, a intensidade final correspondia ao período de transição de um evento para o outro.

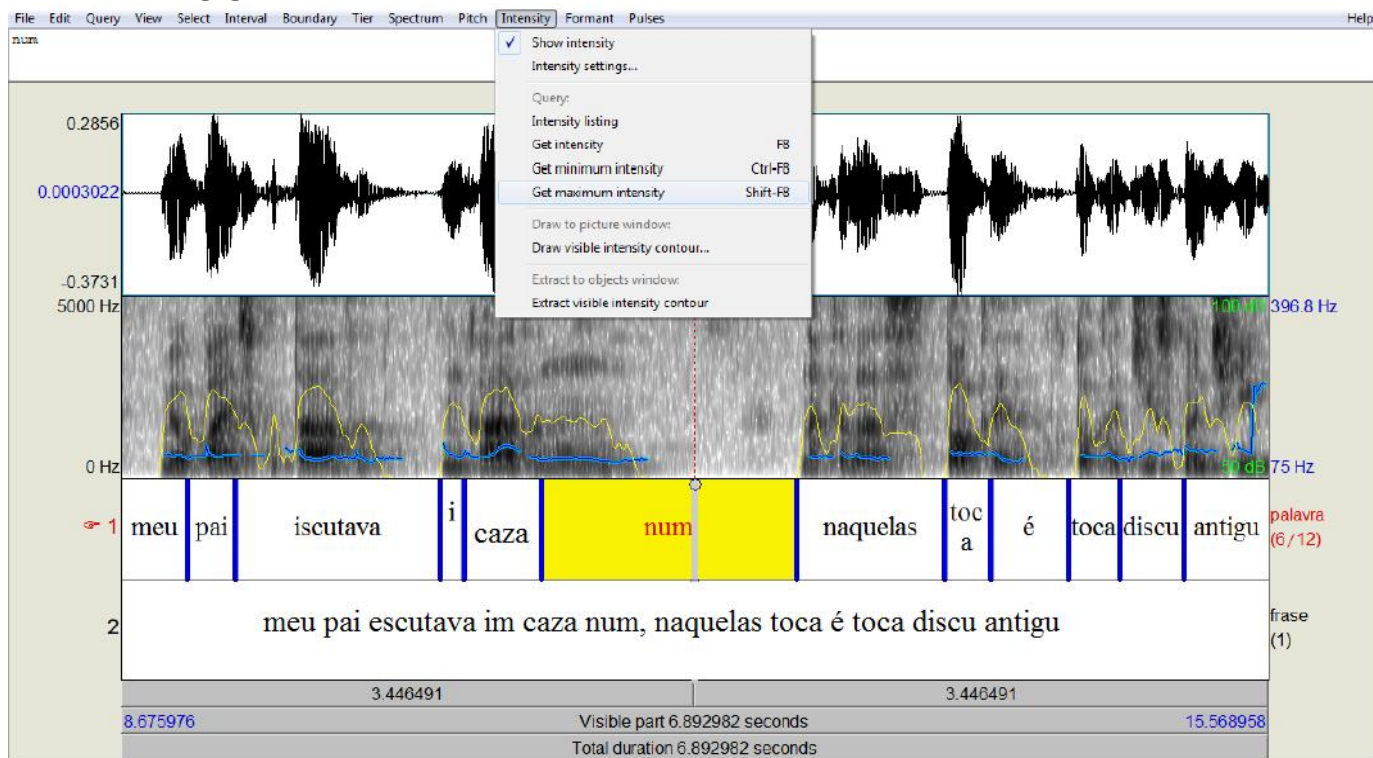
A intensidade inicial é anotada logo no início do segmento, bem como a intensidade final, no final da produção do segmento. A análise desses dois momentos também permite observar se a intensidade final diminui ou aumenta em relação à inicial. Além disso, será possível fazer uma análise comparativa desses momentos para os dois grupos de informantes.

A intensidade máxima é gerada por meio do comando do *software*, *intensity > get maximum intensity*. Para a obtenção da intensidade de transição, inicial e final, é possível optar por três comandos: 1) selecionar, na linha amarela, a posição desejada e verificar o número que aparece, em verde, do lado esquerdo da tela; 2) selecionar, na linha amarela, o período referente à posição desejada e selecionar na tela do PRAAT *intensity > get intensity* e, por fim, 3) selecionar, na linha amarela, a posição desejada e clicar sobre o comando F8 do computador.

Veja abaixo, nas figura 17 e 18, a exemplificação da metodologia utilizada para a análise da intensidade.

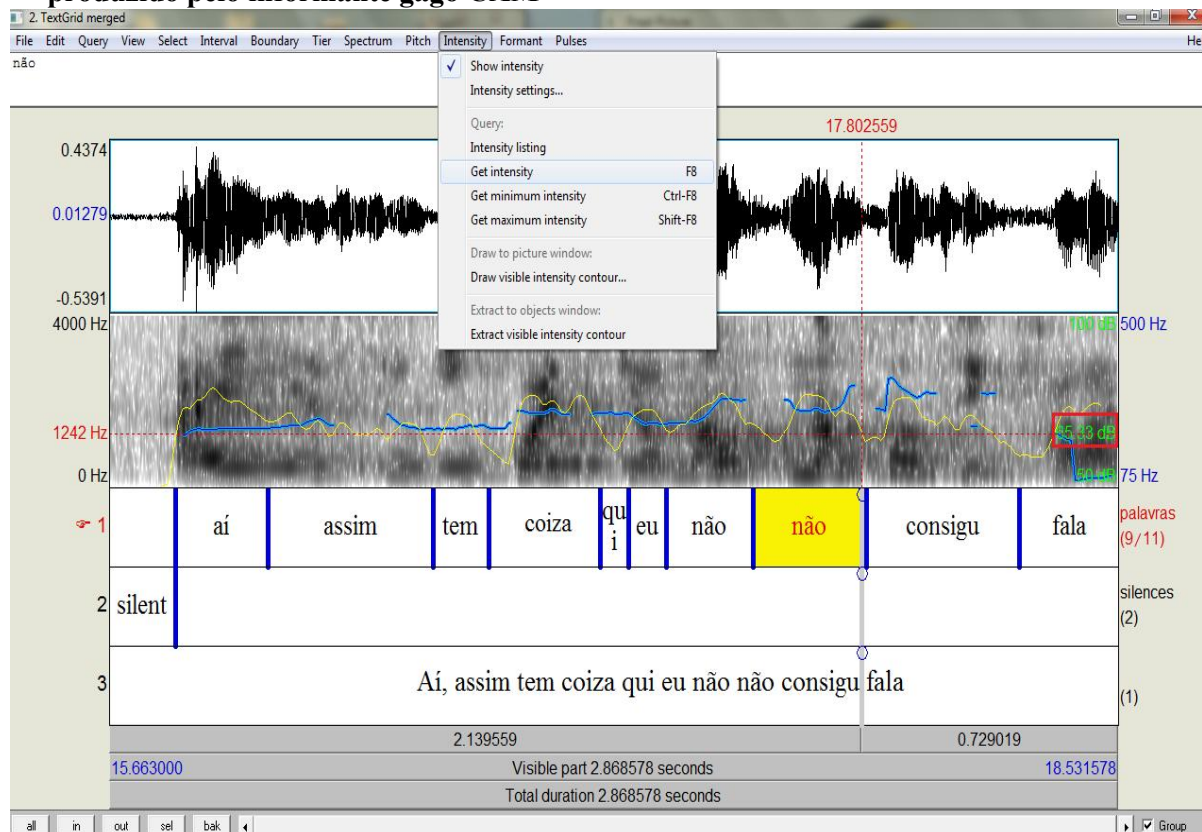
⁵¹ É importante mencionar a dificuldade de se mensurar a intensidade, que pode aumentar ou diminuir conforme a posição do gravador em relação ao falante. Dessa maneira, como não usamos um ambiente de gravação específico, optamos em manter o gravador na mesma posição, sobre a mesa, durante a gravação dos doze informantes, a fim de evitarmos maiores oscilações.

Figura 17 – Procedimento para análise da intensidade máxima no enunciado “meu pai escutava im caza num, naquelas toca é toca discu antigu” produzido pelo informante não gago BR.



Fonte: própria.

Figura 18– Procedimento para análise da intensidade no momento de transição, início ou fim de segmento no enunciado “Aí, assim tem coiza qui eu não não consigu fala” produzido pelo informante gago CAM”



Fonte: própria.

A figura 18 representa o procedimento para a análise da intensidade máxima. Já a figura 49 refere-se à análise da intensidade para transição, início ou fim de segmento. Nesse caso, a intensidade foi anotada no fim do segmento *não*. O seu valor é de 65,33 dB e aparece destacado pelo quadrado vermelho na figura.

Em relação à duração, Crystal (1997) a caracteriza como o tempo de articulação para a produção de um som, sílaba ou enunciado. Como vimos, a duração varia de acordo com a velocidade de elocução, ou seja, são fenômenos inversos, já que o aumento da velocidade leva a uma diminuição na duração dos segmentos. O ritmo, a entoação e a tonicidade, bem como o contexto fonológico, podem influenciar a duração.

Sendo assim, a duração foi analisada a partir: 1) da taxa de elocução; 2) do tempo de articulação e 3) da taxa de articulação.

1) **A taxa de elocução** resulta da divisão da quantidade de sílabas produzidas pela duração total do enunciado. As sílabas foram segmentadas através de uma análise auditiva e acústica, levando em consideração o espectrograma, a forma de onda, entre outros;

2) **O tempo de articulação** é resultado da subtração da duração total do enunciado pela duração total das pausas;

3) **A taxa de articulação**, por sua vez, resulta da divisão do número total de sílabas do enunciado pelo tempo de articulação.

Veja o quadro 3 abaixo que exemplifica essa relação:

Quadro 3 – Esquema para a análise das categorias temporais.

1) A taxa de elocução	$\frac{\text{número total de sílabas}}{\text{duração total do enunciado}}$
2) Tempo de articulação	$\text{duração total do enunciado} - \text{duração total das pausas}$
3) Taxa de articulação	$\frac{\text{número total de sílabas}}{\text{tempo de articulação}}$

Fonte: própria

Segundo Cagliari (2007), a duração relaciona-se ao ajustamento do ritmo, interagindo também com as regras de acento, entoação, qualidade de voz e velocidade de fala.

Por sua vez, a frequência fundamental⁵² refere-se ao efeito acústico produzido pela frequência de vibração das cordas vocais (CRISTÓFARO, 2011). A F0 dos eventos disfluente foi medida em dois momentos:

1) A partir da análise da frequência fundamental mínima F0 (min) e máxima F0 (max). Esses valores são obtidos a partir dos comandos *pitch > get minimum pitch* e *pitch > get maximum pitch* do *software*. A média aritmética desses dois valores F0 (max)/ F0 (min) gera a variação de tessitura dos informantes, que também foi medida neste trabalho;

2) A partir da análise da frequência fundamental média F0(m). Essa medida também é gerada pelo *software*, a partir do comando *pitch > get pitch*.

⁵² Como a análise da F0 é um processo mais frequente em análises acústicas por meio do PRAAT, optou-se por não apresentar uma captura de tela para esse evento, bem como foi feito para outros fenômenos. Mesmo assim, para maiores informações é possível conferir a dissertação de Carvalho (2013), disponível on-line por meio do endereço:
http://www.athena.biblioteca.unesp.br/F/9UK9GUHTDRT5T43ULNY9CY2QPXNCU6SMXU6DX9JYXJ5G3RG81L06608?func=service&doc_library=UEP01&doc_number=000710861&line_number=0001&func_code=WEB-BRIEF&service_type=MEDIA.

Feito isso, foi possível observar em qual tipo de disfluência a F0 é maior. Além disso, para as RPs, esse parâmetro permitiu observar se segmentos iguais são significativamente diferentes, bem como a análise da duração e da intensidade. Vale salientar ainda, que em casos de repetições de mais de dois segmentos, tivemos as seguintes possibilidades:

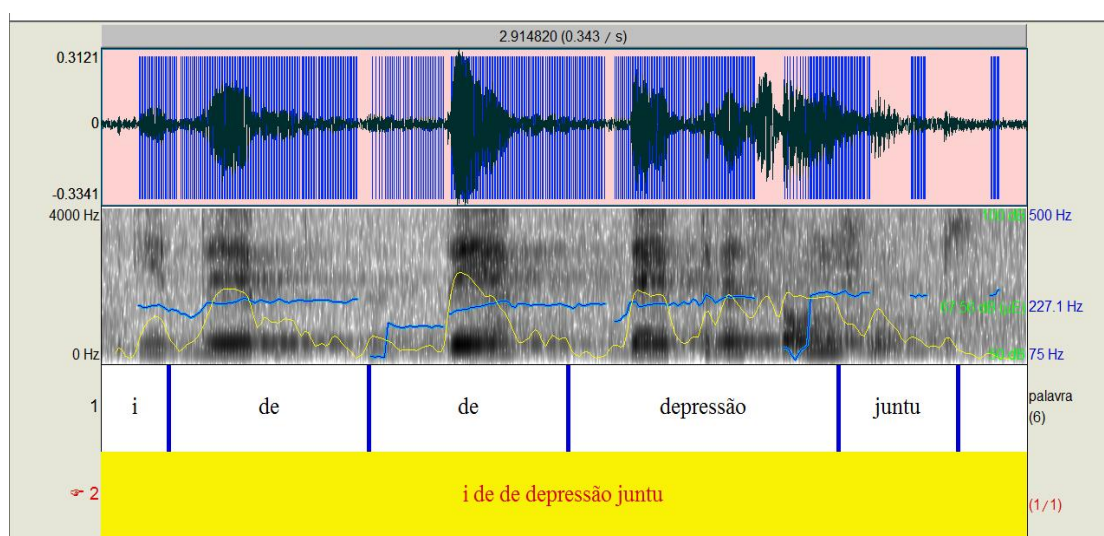
- Padrão da F0 decrescente – F0 que diminui sucessivamente ao longo das RPs;
- Padrão da F0 crescente – F0 que aumenta sucessivamente ao longo das RPs;
- Padrão em forma de onda crescente – F0 que começa com uma frequência mais alta, diminui e aumenta;
- Padrão em forma de onda decrescente – F0 que começa com uma frequência mais baixa, aumenta e diminui.

Os informantes do sexo feminino apresentam uma escala de frequência mais alta do que os do sexo masculino, como mencionado, além de uma qualidade de voz mais aguda. Isso fez com que a presença de erros na F0 fosse maior, gerando problemas para a análise acústica. As disfluências que apresentaram esse fenômeno não foram analisadas acusticamente para esse parâmetro. No entanto, para a análise entoacional, modificamos o padrão *default* do *software*.

Já os pulsos glotais estão relacionados à vibração das cordas vocais. No *software*, eles estão representados por linhas azuis verticais. Através da análise dos pulsos podemos observar as irregularidades na marcação dos ciclos das ondas, o que possibilita a comparação entre os eventos disfluentes nos dois grupos de informantes.

Veja na imagem 19, abaixo, o exemplo desse parâmetro.

Figura 19– Identificação dos pulsos por meio das linhas azuis verticais no enunciado “i de de depressão junto”, produzido pelo informante gago CAM.



Fonte: própria.

Para a análise acústica dos informantes gagos e não gagos, os pulsos foram analisados apenas em situações de correção (COR), repetições de sílabas, palavras, vogais e estrutura (RP_s; RP_p; RP_v; RP_e). Para as outras disfluências, foram feitas apenas algumas considerações, pois como são segmentos diferentes, fica difícil estabelecer comparações.

Finalmente, a análise entoacional dos dados foi feita de acordo com o modelo teórico de Pierrehumbert (1980) e da versão modificada do modelo (BECKHAM; PIERREHUMBERT, 1986). O modelo auditivo desenvolvido por Halliday (1970) e adaptado para o português do Brasil por Cagliari (1982, 2007) foi utilizado para a segmentação do enunciado em GTs. Intencionamos aqui que o padrão entoacional das repetições, no mesmo enunciado e informante, se mantém, uma vez que repetição conservaria a articulação. Por outro lado, os outros tipos de disfluências teriam padrões entoacionais diferentes, devido aos diferentes tipos de frases e as suas posições no enunciado. No entanto, é provável que as inserções, se apresentarem a mesma estrutura sintática, mantenham o mesmo contorno entoacional.

Resumidamente, no modelo da autora, a curva entoacional é formada por uma sequência de *pitch accents* e *edge tones*. Os tons nesse modelo são representados por H (*high* - alto) ou L (*low* - baixo), ou pela combinação entre eles. Nesse caso, temos os tons binários (H+L ou L+H). O *pitch accent* é simbolizado por um asterisco (*) como, por exemplo, H* ou L*. Ele representa o foco entoacional do enunciado. As sílabas átonas não precisam ser anotadas. Os *edge tones* dividem-se em acentos frasais e tons de fronteira. Os acentos frasais podem ser formados por um tom (H ou L) ou por um par de tons (H+L). As sílabas do acento frasal não são anotadas pelo símbolo (*), já que são menos salientes do que as do *pitch accent*. Os tons de fronteira são representados pelo símbolo de porcentagem (%), como por exemplo, H% ou L%. Eles ocorrem no final da frase entoacional. O símbolo de porcentagem também é utilizado para a marcação de pausa no enunciado.

Pierrehumbert (1980) também trabalha com regras de reajuste entoacional, como as regras de *downstep*, *upstep*, regras de espraiamento, entre outras. Os enunciados analisados exigiram apenas as regras de *downstep* (queda melódica contínua do início ao fim do enunciado) e *upstep* (movimento crescente contínuo da curva melódica no enunciado). Para a análise das regras de *downstep* utilizamos a versão adaptada de 1986⁵³. Esses eventos foram anotados como !H, para o *downstep* e ¡H para o *upstep*.

⁵³ Para mais informações sobre o modelo conferir o capítulo 5.

Para análise dos dados foi preciso fazer algumas adaptações no modelo entoacional. Os ajustes foram feitos apenas para os eventos disfluente do enunciado, a fim de diferenciá-los dos eventos fluentes. Veja a seguir os símbolos utilizados para a leitura da análise entoacional nos enunciados:

- (#) – Esse símbolo foi usado para representar as pausas em fim de enunciado e a pausa silenciosa;
- (*) – O símbolo do asterisco segue a mesma proposta de Pierrehumbert (1980). No entanto, para os eventos disfluente eles foram anotados antes do tom H ou L, como por exemplo, *H ou *L. Isso foi feito apenas com o objetivo de diferenciar as disfluências do resto do enunciado;
- (%) - O símbolo de porcentagem foi utilizado com a mesma intenção de Pierrehumbert (1980) de anotar as pausas entre os grupos tonais e para anotação dos tons de fronteira;
- (/) - a barra inclinada indica a presença de pausas respiratórias nos enunciados.

Por meio dos procedimentos apresentados será possível comparar com maior exatidão as diferenças entre as disfluências gaga e comum que são acusticamente significantes.

Na sequência, apresentamos a primeira parte das análises, que diz respeito à análise individual de cada um dos informantes de cada grupo. Iniciamos com a análise dos informantes com disfluência comum e depois, com os disfluente gagos. Para o primeiro informante de cada grupo foi feita uma análise mais detalhada dos eventos. A partir deles, seguimos com uma análise mais geral. No entanto, os eventos que divergiram do que foi observado para o primeiro informante receberam uma análise mais específica.

Na segunda parte da análise comparamos os resultados obtidos, a fim de que fosse possível observar as semelhanças e diferenças entre os dois tipos de fala.

4 RESULTADOS E DISCUSSÃO

Este capítulo divide-se em duas partes. Na primeira delas será feita uma análise mais detalhada dos parâmetros acústicos e entoacionais observados para as disfluências comum e gaga. Na sequência, faremos uma análise comparativa dos resultados obtidos entre os dois grupos a fim de observar semelhanças e/ou diferenças.

4.1 Informantes com disfluência comum

O informante CR

O informante CR, em um total de 34 enunciados realizou 93 disfluências. O número máximo de disfluências por enunciado foi de quatro. Dentre elas, duas foram pausas. Do total de disfluências, quatro foram complexas e 89 simples. A sua taxa de elocução foi de 4,358 sil/ms, a de articulação de 5,185 sil/ms e o tempo de articulação aproximado de 104ms. O gráfico 1 caracteriza as disfluências de uma maneira geral, ou seja, não estão indicados os subtipos das disfluências. Veja a seguir.

Gráfico 1 - Relação das disfluências simples para o informante não gago CR



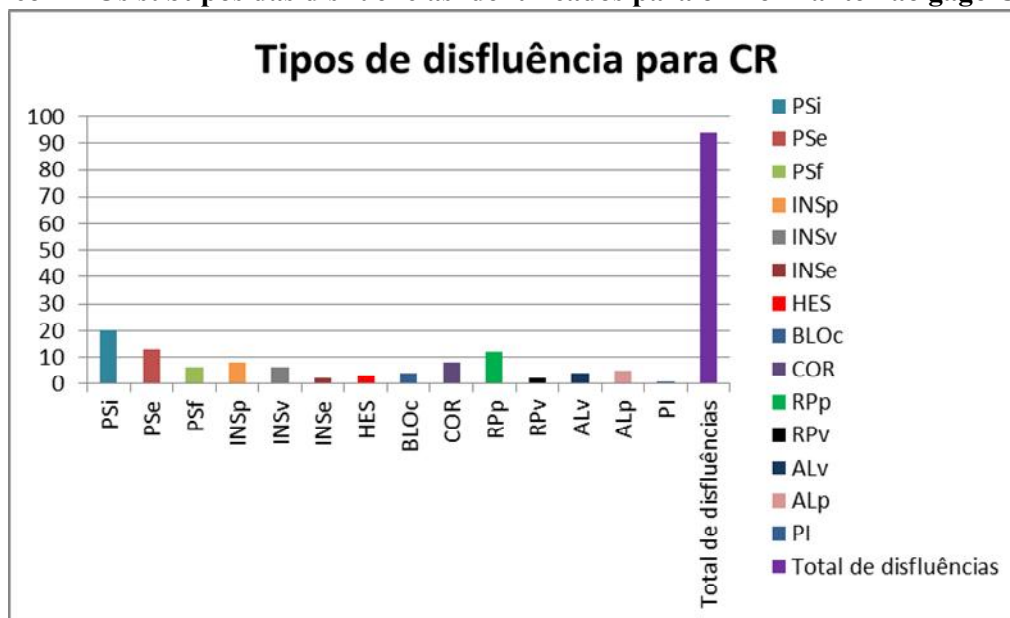
Fonte: própria.

De acordo com o gráfico, o evento mais frequente é a PS (38 produções), que aparece em aproximadamente 41% das produções. Já o menos frequente é a produção de PI, com 1% das ocorrências. Quanto à proporção de eventos no total de enunciados, ocorrem em segundo lugar as INSS, com aproximadamente 17% das ocorrências, depois as RPs com

aproximadamente 15%, os ALs com aproximadamente 9,5%, as CORs com 8,5% e, por fim, os BLOs e HESs, com respectivamente 4,2% e 3,2% das produções.

O segundo gráfico, por sua vez, traz as informações mais específicas. Observe:

Gráfico 2 – Os subtipos das disfluências identificados para o informante não gago CR.



Fonte: própria.

O gráfico 2 revela que entre os tipos de PS analisadas (PS_i , PS_e e PS_f), a mais frequente é a do tipo PS_i , com aproximadamente 19% das ocorrências. As outras pausas representam respectivamente 16% (PS_e) e 5,3% (PS_f). Como vimos no gráfico 1, foram produzidas um total de 39 pausas. Dentre elas, 68% foram de pausas respiratórias, 26,3% de pausas silenciosas e 5,3% de pausas articulatórias. Em segundo lugar aparecem as repetições de palavra, com 12,6% das produções. Nesse caso, há uma diferença do gráfico geral, em que as inserções (17%) eram mais comuns do que as repetições (16%). Em relação à subdivisão das repetições temos as seguintes ocorrências:

1. Para as RP_p s

- $3RP_p$ (sendo 1 complexa do tipo $3RP_p + PS_e$) – 2,1%;
- $2RP_p$ – 6,5%;
- $2RP_{p3}$ – 1%;
- $2RP_{p2}$ – 1%.

2. Para as RP_v s

- $2RP_v$ – 1%;
- $3RP_v$ – 1%.

3. Para as RP_s

- a) 2RP_s – 2,1%

Dentre as RPs, a mais frequente são as repetições de palavras, que aparecem em 13% dos eventos. Em seguida, temos as repetições de vogais e sílabas, com 2,2% das produções.

Os subtipos das inserções ficaram assim caracterizados:

1. Para as INSs

- a) INS_p em início de GT – 5,3% (4 /intão/; 1 /aí/);
 b) INS_f em fim de GT – 3,2% (2 /né/; 1 /aí/);
 c) INS_p em meio de GT – 3,2% (1 aí; 2 assim);
 d) INS_v em início de EN – 4,2% (2 simples /i/; 1 complexa /i/; 1 complexa /é/);
 e) INS_v em meio de EN – 1,1% /i/;
 f) INS_e – 2,1% (1 /i aí/ em início de EN e 1 /i/ em início de EN).

Quanto às INSs, a produção de inserção, seja de vogal, consoante ou inserção complexa (vogal e consoante), acontece com maior frequência em início de enunciado. O marcador discursivo, considerado neste trabalho como INS, mais usado pelo informante é “intão” e, a vogal mais frequente /i/.

As CORs, BLOs, HESs e PI, mantém a porcentagem apresentada no gráfico geral. Vale salientar que, para esse informante, ocorreram apenas bloqueios em consoantes, sendo que dentre as quatro ocorrências, duas foram em consoantes oclusivas velares [k] e duas de nasais alveolar [n]. Além disso, três deles foram analisados com uma pausa articulatória antes da produção.

Por fim, os alongamentos não se diferenciaram muito do que foi visto no gráfico geral, sendo 4,3% de AL em vogal e 5,4% em palavras. Dentre eles, tivemos as seguintes ocorrências:

1. AL_v

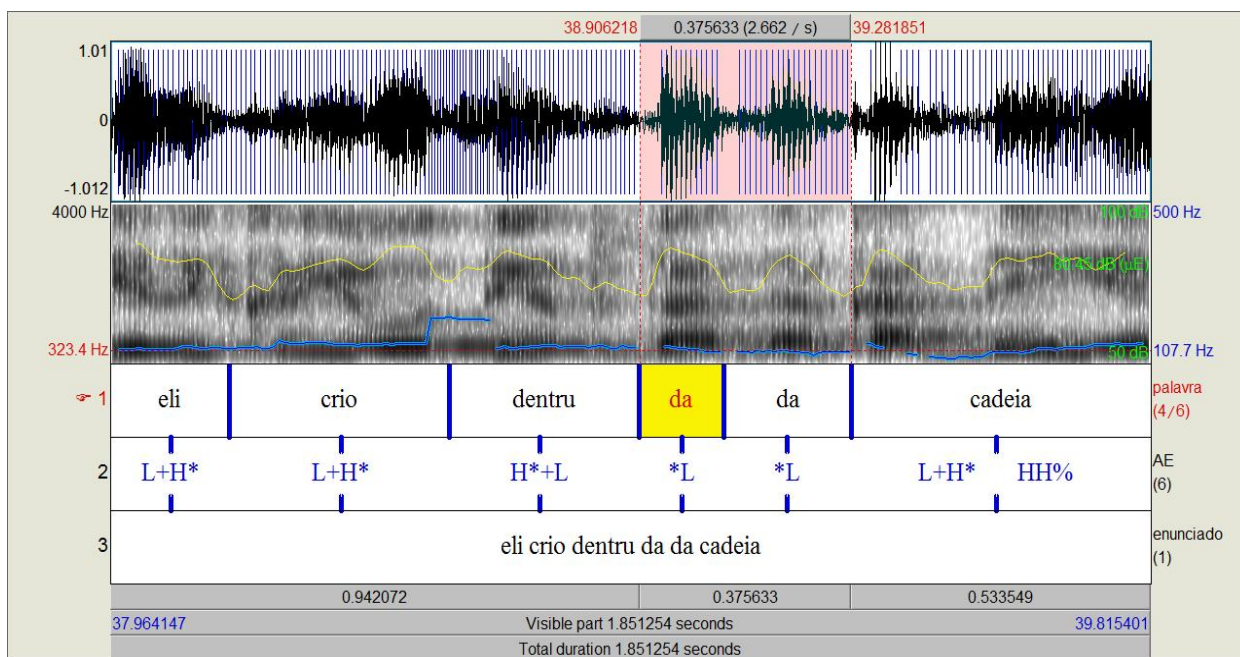
- a) 1 AL_v - /a/;
 b) 3 AL_v (complexo), sendo (INS_e + AL_v /é/; INS_e + AL_v /i/ e INS_e + AL_v /é i/).

2. AL_p

- a) 1 AL_p em início de palavra;
 b) 1 AL_p em fim de palavra
 c) 2 AL_p.

Quanto à análise entoacional das repetições, o padrão entoacional se manteve para a maioria delas, sejam elas de palavras, sílabas ou vogais. Nas 14 repetições encontradas, 13 delas apresentaram padrões entoacional simples, variando entre *H e *L e apenas uma apresentou um padrão composto. Veja o exemplo que segue:

Figura 20 – Exemplificação do padrão entoacional da repetição para o informante CR, no enunciado resumido EN6 “eli crio dentro da da cadeia”.



Fonte: própria.

A fim de facilitar a análise, a figura acima é apenas um recorte do enunciado completo EN6 “/i eli crio dentru *da da* cadeia eli crio um personagim#”, em que ocorre, além da 2RP_p, uma pausa respiratória em início de enunciado (PS_i), uma inserção de vogal /i/ (INS_v) e uma pausa silenciosa em fim de enunciado (PS_f).

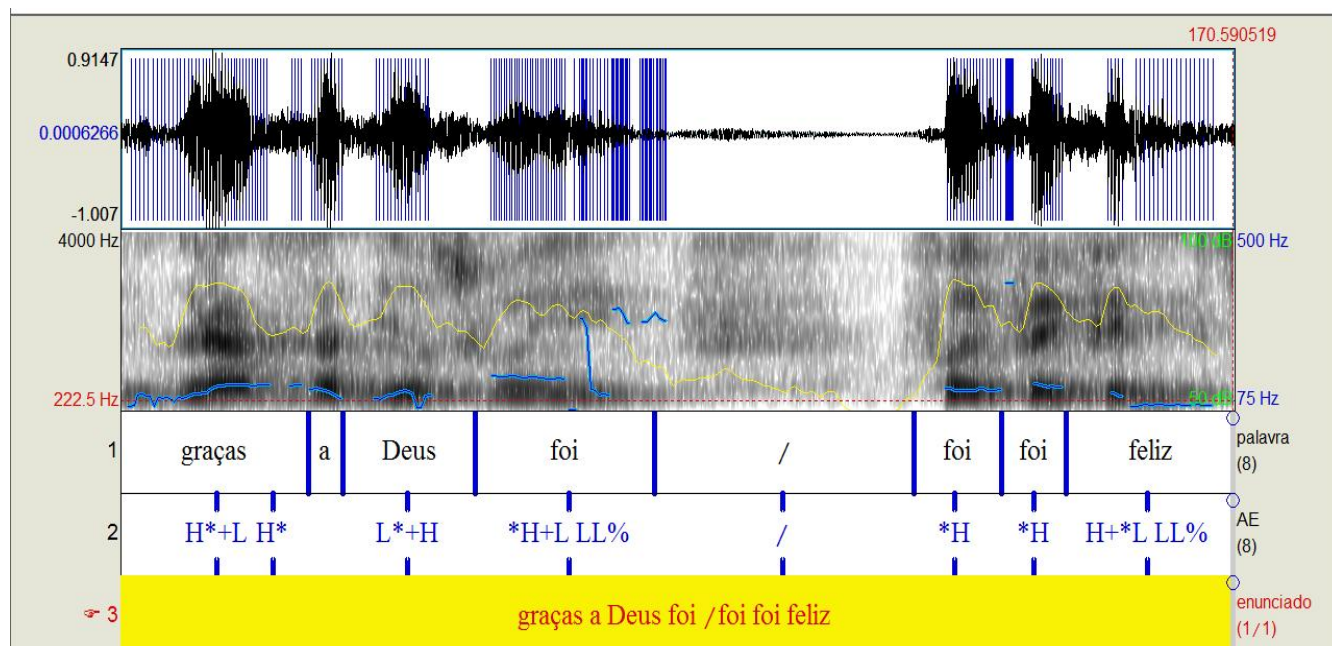
No exemplo, observamos que a configuração entoacional é praticamente a mesma para as duas repetições /da/, /da/, *L. Se observarmos mais atentamente, veremos que a intensidade (linha amarela da figura) também apresenta uma silimilaridade para as duas ocorrências. A diferença maior está na duração das duas repetições, o que será analisada posteriormente.

No enunciado completo, exemplificado mais acima, observamos ainda que após a repetição o informante retoma a estrutura “eli crio” do começo da frase. Nesse caso, após o erro, CR recupera parte da informação inicial, como se estivesse inserindo um novo GT. Isso mostra que, incoscientemente, o falante sabe que cometeu um ato falho e com o intuito de

retomar a fluência, recupera parte do enunciado. No entanto, em um total de 14 repetições, esse fenômeno ocorreu apenas para o enunciado em questão e para o EN19.

Como dissemos, a configuração entoacional não apresentou o mesmo padrão apenas para um dos enunciados, que está caracterizado a seguir.

Figura 21 – Padrão entoacional da repetição no enunciado EN19 “graças a Deus foi/ foi foi feliz” produzido por CR.



Fonte: própria.

Como vemos, a primeira RP da palavra /foi/ apresenta o padrão *H+L e as outras duas repetições *H e *H. A frequência máxima da vogal /i/ na primeira RP é muito alta: de 315 Hz⁵⁴. Isso ocorreu devido a um erro de estação, o que não leva a variação no padrão entoacional. Após essa produção, ocorre uma PS_e silenciosa e, posteriormente, mais duas repetições da palavra /foi/, com padrão *H e *H.

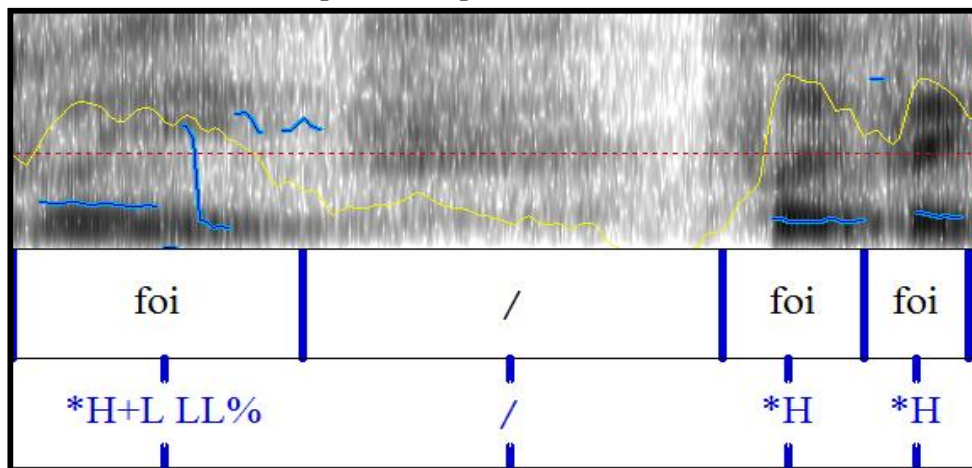
Ao olharmos para intensidade das três RPs, veremos que elas apresentam praticamente o mesmo padrão decrescente. Além disso, se observarmos o início das três produções, veremos que elas são muito próximas. Tanto isso é verdade, que o pico entoacional é o mesmo nos três eventos, *H. É possível que essa diferenciação deva-se à PS, que ocorre após a primeira RP, gerando um caso de disfluência complexa do tipo 3RP_p+ PS_e. A PS_e entre as produções ocasiona uma quebra no GT⁵⁵. É como se tivéssemos um caso de pensamento

⁵⁴ Os valores de F0 e intensidade foram aproximados para todos os informantes.

⁵⁵ Vale salientar que nesse momento da entrevista o informante estava um pouco emocionado, o que também pode ter contribuído para que essa RP_e acontecesse.

suspenso⁵⁶, que precisa ser completado para que o enunciado não perca o sentido e o fluxo de fala continue. Contudo, essa diferenciação não descaracteriza o fato de que para as RPs a configuração articulatória permanece a mesma, o que pode ser identificado pelo espectro do enunciado. Observe essa afirmação mais detalhadamente na figura que segue:

Figura 22 – Representação detalhada do espectrograma das repetições no enunciado “graças a Deus foi/ foi foi feliz”, produzido por CR.



Fonte: própria.

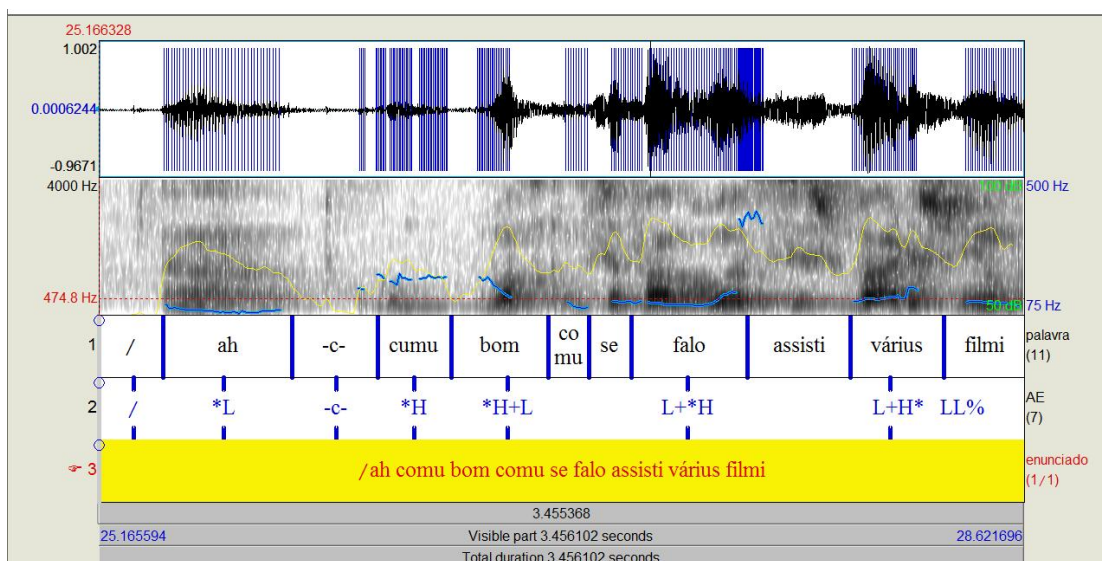
A imagem recortada mostra que o espectrograma das três repetições são muito parecidos. A diferença está no modo como o informante pronuncia o primeiro /foi/, ou seja, CR a pronuncia como uma palavra dissílaba, diferentemente das outras produções, que seguem o padrão de monossílabos.

Quanto aos outros tipos de disfluências analisadas para o informante CR, foi verificado que:

1) as três HES, presentes em EN4, EN20 e EN28, mantiveram o mesmo padrão entoacional *L. As três HES foram /ah/. Dentre elas, EN4 e EN20 ocorreram em início de palavra e EN28 em meio de palavra. Observe um exemplo que segue:

⁵⁶ No modelo de Halliday (1970) e Cagliari (1992, 2007) o padrão entoacional do tipo suspenso é representado pelo tom 3 e suas variantes.

Figura 23– Exemplo de HES no enunciado EN4“ah como bom como se falo assisti vários filmi”.



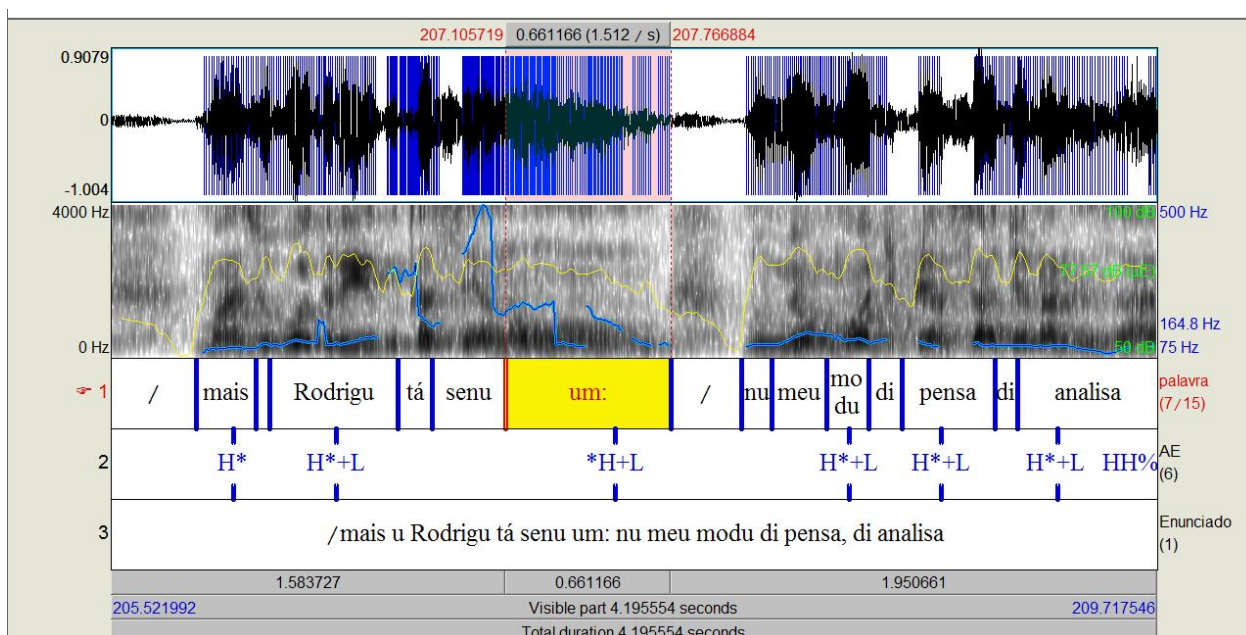
Fonte: própria.

A figura acima traz informações da forma de onda, espectro, pulsos intensidade e F0 das disfluências. Nesse caso, assim como para EN28 a HES ocorre em início de enunciado, logo após a PS silenciosa. As HESs apresentaram uma qualidade de voz do tipo *creaky voice*. Essa característica é comprovada por uma F0 mais baixa, com mínima de 79 Hz e máxima de 106 Hz, bem como através dos pulsos (linhas horizontais em azul) mais próximos. As HES são utilizadas para dar continuidade ao enunciado que segue, como forma de concatenar o pensamento. Mas também podem ter uma função pragmática de ironia, decepção, contentamento ou surpresa. Em CR, a HES (cf. fig.54) liga as ideias. Nesse caso, o falante faz uma união lógica de ideias por meio da HES quando lembra que o seu interlocutor já tinha mencionado que ele havia assistido a vários filmes. Em EN20 “ah u u u Rodrigu ta senu uma surpresa” traz um tom de contentamento, uma retomada positiva, mas também não deixa de relacionar as ideias. Por outro lado, em EN28 “/é um povu qui vivi ah fazenu fofo”, o tom é de decepção diante de um fato. É interessante notar que a F0 é ainda mais baixa, sendo a máxima de 91 Hz e a mínima de 75 Hz. Casos em que as hesitações mantivessem o mesmo padrão entoacional não eram esperados. No entanto, todas apresentaram o mesmo padrão entoacional do tipo *L.

Quanto ao alongamento, os AL_p que ocorreram em palavras monossílabas, mantiveram o mesmo padrão entoacional (*H+L) nas palavras /um/ e /qui/, respectivamente em EN24 e EN28. Esse também não era um fenômeno esperado, mas nesse caso foi apenas uma coincidência.

Para os AL_{ps} em final de palavra os padrões entoacionais foram diferentes ($*L+H$ e $*H+L$). Para o AL_p em início de palavra e AL_v não é possível fazer comparações, uma vez que foram encontradas apenas um evento de cada. Os AL_s têm a função de estender o pensamento. Em outras palavras, ajudam o informante a lembrar ou a programar o que será articulado na sequência. É como se o informante deixasse o pensamento suspenso muito rapidamente, recuperando-o logo em seguida. Dessa maneira, ao estender a sílaba inicial/final ou palavra, o falante consegue recuperar e/ou formular a próxima sequência. Veja:

Figura 24 – Exemplo de alongamento em EN24 “/mais u Rodrigu tá senu um:/ nu meu modu di pensa, di analisa”, produzido por CR.



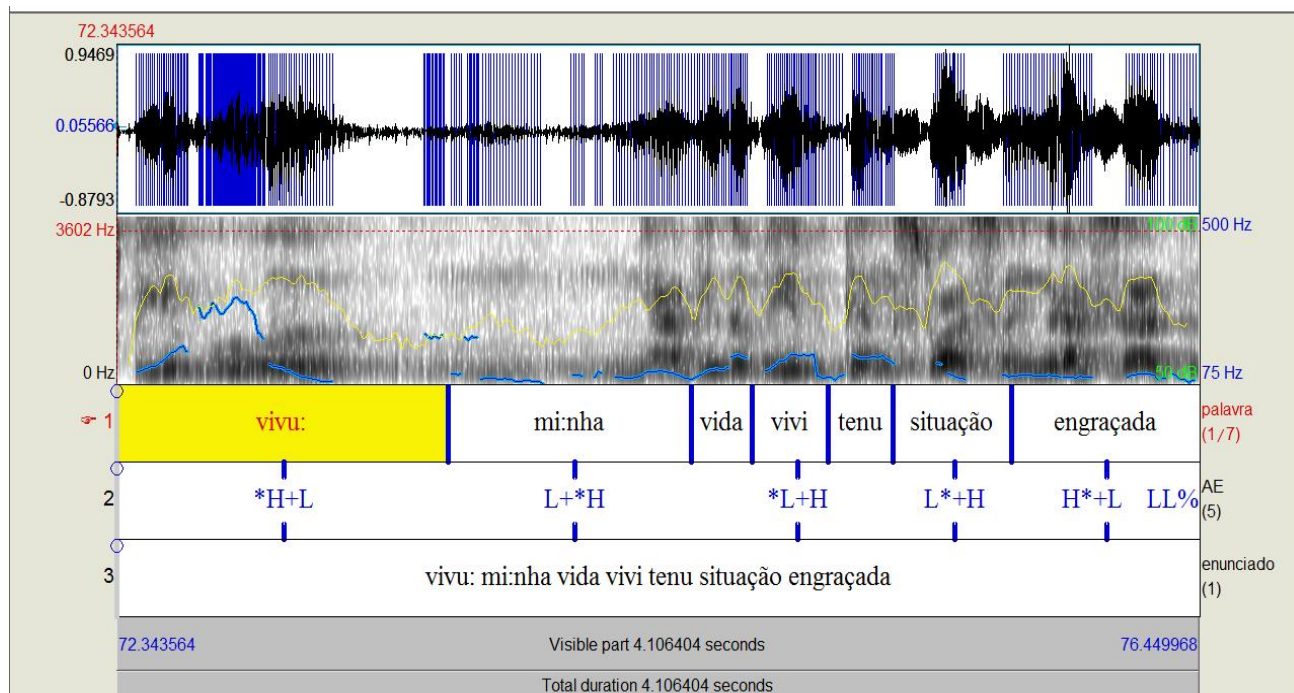
Fonte: própria

Em EN24 “/mais u Rodrigu tá senu um:/ nu meu modo di pensa, di analisa”, exemplificado pela figura 24, a palavra alongada /um:/ aparece selecionada em amarelo. A duração da palavra alongada foi de 0.661s. Dentre todas as palavras do enunciado, a alongada foi a que teve a maior duração. Os outros segmentos com durações maiores são /Rodrigo/, com uma duração de 0.513s, e /analisa/, com uma duração de 0.562s. Mesmo assim, esses valores não caracterizam alongamentos, uma vez que são segmentos maiores do que a palavra /um:/.

Como vemos, o falante alonga a palavra com o intuito de buscar um adjetivo para a pessoa referida. Mesmo assim, ele não o encontra e cria uma nova estratégia para manter o turno conversacional até que formule a palavra desejada. Durante esse processo, CR ainda realiza uma pausa respiratória de 0.283s. Percebemos que esse recurso é mais utilizado

quando o alongamento ocorre na palavra inteira e na vogal, já que quando o alongamento ocorre em início ou fim de sílaba, além dessa estratégia, o informante utiliza mais um recurso. Veja:

Figura 25– Exemplo de AL_p em fim e início de palavra em EN11 “vivu: :minha vida vivi tenu situação engraçada”, produzida por CR.



Fonte: própria

A figura traz dois casos de alongamento. O primeiro em final de sílaba, /vivu:/ e o segundo em início, /mi:nha/. No primeiro alongamento, o informante não consegue atingir o objetivo da sua argumentação e corrige o enunciado. É como se tivéssemos dois GTs “vivu:” e “minha vida vivi tenu situação engraçada”. A segunda sílaba alongada /vu:/ tem duração de 1.021s. Por outro lado, a primeira sílaba da palavra /vi/, tem duração de 0.224s. O total da duração da palavra é de 1.256s, ou seja, o AL final refere-se a 81% da duração da palavra total. Por outro lado, o alongamento inicial da palavra que segue é um pouco menor, mas não deixa de ser significativo. A duração de /mi:/ é de 0.733s e a da sílaba posterior é de 0.180s. A palavra tem uma duração total de 0.921s. Nesse caso, o AL inicial ocupa 92% da duração total palavra. Assim, em relação à palavra, o segundo AL seria proporcionalmente maior do que o primeiro. Para fins ilustrativos, nesse enunciado, as sílabas têm durações médias de 0.135s, o que prova que as duas sílabas apresentadas são alongadas. Além disso, as duas sílabas alongadas vêm acompanhadas de uma qualidade de voz do tipo *creaky voice*, que pode ser

verificada por meio dos pulsos mais espaçados e F0 mais baixa, sendo a mínima de 81 Hz para /vu:/ e 77 Hz para /mi:/.

Como foi mencionado, o AL em palavra total tem como finalidade manter o interlocutor no discurso para recuperar o pensamento e, conseqüentemente, o fluxo de fala. Por outro lado, o AL em início ou fim de palavra, além de ter a estratégia de manter o turno conversacional, no primeiro caso, /vu:/ passa a ideia de uma incerteza, tanto assim, que o informante age como se encerrasse o GT. O segundo AL, /mi:/ por sua vez, também usa o AL de maneira enfática, chamando atenção para ele, para a sua vida.

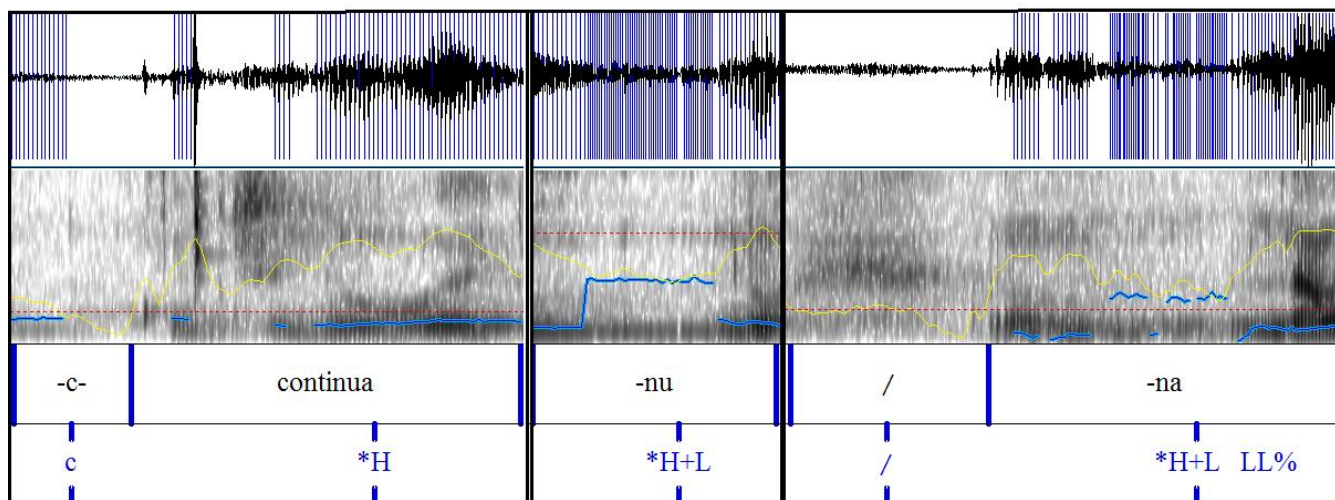
Assim como ocorreu para o AL_p em início de palavra e para o AL_v, não é possível dizer categoricamente que o padrão entoacional das PIs são diferentes, pois em CR elas representam apenas 1% das ocorrências. No entanto, como veremos, os padrões são diferentes, uma vez que os segmentos também são diferentes.

Quanto ao BLO, foram encontradas quatro ocorrências, respectivamente em: EN1 /-continua/, EN4 /-cumu/, EN26 /-nu/ e EN31 /-na/. Dentre elas, EN26 não apresentou PS articulatória nem respiratória; EN31 apresentou uma PS respiratória e EN1 e EN4 apresentaram pausas com padrão articulatório.

Como foi descrito no capítulo referente à metodologia, a PS articulatória é um tipo de pausa que mantém a configuração articulatória da consoante bloqueada, podendo ou não estar presente a vibração das cordas vocais. Isso pressupõe que os BLOs seriam identificados por meio de uma PS anterior que mantivesse a configuração articulatória da consoante ou vogal bloqueada. Para o informante em questão a teoria foi confirmada para 50% dos casos.

Em relação a CR, os padrões entoacionais obtidos para foram: *H (EN1), *H (EN4), *H+L (EN26) e *H+L (EN31). Como vemos, os dois enunciados com PS articulatória mantiveram o mesmo padrão entoacional *H e os BLOs que seguiram sem a PS articulatória também, mas agora com o padrão *H+L. Para fins ilustrativos, retomamos a figura 9, apresentada na metodologia que representa os três tipos de pausas.

Figura 26 – Representação da possibilidade dos três tipos de bloqueio observados para as disfluências comuns.



Fonte: própria.

Da esquerda para a direita estão ilustrados os três tipos de BLOs resultantes dos enunciados analisados. A primeira imagem refere-se ao enunciado EN1 “a perspectiva do café essi anu continua divagar”, em que ocorre PS articulatória antes do BLO_c. A segunda imagem refere-se ao BLO_c sem PS respiratória ou articulatória no EN26 “/é um -nu sei si é Andradas porque eu nu/ eu convivu aqui”. Por fim, a última ilustra uma PS respiratória em EN31 “i elis num pensam na união”.

Há algumas evidências que provam que a primeira PS é articulatória. Por exemplo, notamos que no segmento -c- há a presença dos pulsos, o que indica a presença das cordas vocais, além disso, ocorre uma F0 que segue a mesma linha da palavra seguinte “continua”. Além disso, o espectro de -n- segue o mesmo padrão da palavra seguinte. Comparativamente, na terceira imagem, não há a presença de pulsos e os espectros não são iguais. A análise auditiva mostra que a respiração é bem clara. Por outro lado, em EN26 não há presença de nenhum dos dois tipos de pausa. A sua F0, bem como em EN31, não segue um padrão constante. Outro fato que poderia justificar o BLO seria o tipo de segmento anterior. No entanto, nos três exemplos apresentados eles são diferentes, uma vez que antes de EN1 ocorre uma vogal posterior alta [u], em EN26 uma nasal bilabial [m] e em EN31 uma fricativa alveolar [s]. Outra característica dos BLOs diz respeito à duração maior da consoante bloqueada⁵⁷. Mesmo diante dessas características, parece que a identificação dos BLOs é mais clara auditivamente. É possível, que não exista um padrão para a identificação dos BLOs para

⁵⁷ Essas questões serão vistas mais especificamente no item referente à análise acústica.

as disfluências comuns, ou seja, eles poderão ser seguidos de uma PS articulatória ou respiratório, ou ainda, não apresentar nenhuma delas.

No que diz respeito às correções, de 93 casos de disfluências, apenas 8,6% delas foram de CORs, o que no conjunto geral dos dados não é muito significativo. O padrão entoacional analisado da passagem do erro para a correção foram os seguintes:

- | | |
|--|------------------------------------|
| a) EN4 - cumu [bom comu] | >*H [*H+L]; |
| b) EN7 - quandu eli [qui eli] | >*H+L *H [*H+L *H]; |
| c) EN11 - vivu [minha vida] | >*L+H [*L+H *L+H]; |
| d) EN21 - eu inda fui uns [uns não] | >*H L+*H L+*H L+*H HH % [L+*H *L]; |
| e) EN25 - trazê [fazê] | >*H [*L+H]; |
| f) EN26 - porque eu nu [eu convivu aqui] | >*H+L *L LL% [L+*H]; |
| g) EN27 - tem um [tem pessoas] | >*H *H+L [*H *H+L]; |
| h) EN29 - eli [e são] | >*L [*H *H]. |

Optamos por anotar o padrão entoacional de todas as palavras que se encaixavam na estrutura de erro e correção, isso foi feito para fins comparativos. Os exemplos acima apresentam apenas a parte inicial da correção e não o GT completo.

Dos enunciados selecionados, observamos que durante a correção o informante apresenta duas estratégias de correção: 1) manter a palavra do erro na correção (EN4, EN7, EN21, EN26, EN25 e EN27). Nesse caso, pode ou não aparecer uma pausa (EN21 e EN26) ou 2) mudar significativamente os GTs (EN11 e EN29). Vale salientar que em EN11 ocorreu um AL_p em final de palavra no erro /vivu:/ e na correção um AL_p em início de palavra /mi:nha/. No decorrer dessas estratégias, o informante mantém a mesma frequência entoacional tanto no erro quanto na COR em apenas dois enunciados (EN7 e EN27) que dizem respeito à estratégia dois. Nos outros enunciados que mantêm parte do erro na COR, apenas em EN21 o informante segue o mesmo padrão entoacional da palavra recuperada no erro “eu inda fui uns [uns não]”.

De uma maneira geral, esses resultados mostram que o informante, ao perceber que errou, na maioria das vezes, muda a F0 e, conseqüentemente, o padrão entoacional. EN7 e EN27 foram os únicos que fugiram à regra, para esse informante, como já exemplificado. No entanto, dentre esses enunciados, apenas EN7 manteve a F0 constante, ou seja, sem quebra na estrutura durante a passagem do erro para a correção, como veremos na análise acústica.

Em um contexto geral das disfluências, as INSs foram, depois da pausa, as mais frequentes. Dentre elas, a mais comum foi a INS de palavra e a menos comum, a INS de estrutura. Retomando os exemplos, os tipos de INSs utilizados pelo informante foram /né/, /intão/, /assim/ e /aí/ para INS de palavra, /i/ e /é/ para as vogais e, por fim, as INS de estrutura foram /i aí/ e /é i/.

As INS de palavras também são chamadas de marcadores discursivos. Mas, neste trabalho, elas foram caracterizadas como uma inserção, uma vez que além de serem um recurso discursivo, caracterizam-se como uma disfluência, uma vez que o informante a utiliza com o intuito de ganhar tempo para o processamento final da mensagem.

No caso de CR, a INS em início de GT são as mais frequentes. Depois aparecem as INS em final de enunciado e, por último, as em meio de GT. Vale salientar que em meio de GT, entre todas as realizadas, três foram de INS_p e uma de INS_v . Em início de GT a INS_p mais comum é /intão/. Por sua vez, a INS_v mais frequente é /i/. Já em fim de enunciado só ocorreram INS_p , sendo a mais utilizada /né/. Quanto às INS_e s, elas foram realizadas sempre em início de enunciado e estão divididas em: inserção de vogal com palavra e INS de vogal dupla. Elas foram caracterizadas dessa forma porque ocorrem simultaneamente. No caso do primeiro tipo, essa INS_e funciona como se fosse uma estrutura completa, isto é, uma palavra. Já o segundo tipo (INS de vogal dupla) não pode ser caracterizada da mesma forma, pois no conjunto do enunciado não possui um sentido completo, mas, como ocorrem em sequência, foram considerados como uma estrutura. A suposição feita foi a de que o padrão das INSs seria o mesmo quando se tratasse da mesma estrutura sintática. No entanto, esse não foi o resultado esperado para 100% dos casos. Veja os padrões obtidos:

a) Em início de GT para as palavras:

- EN3 /intão/ – $H+*L >$ “intão u mercadu devi permace nus níveis qui istão”;
- EN7 /aí/ – $*L+H >$ “aí quandu eli qui eli conseguiu”;
- EN10 /intão/ – $L+*H >$ “#intão eu eu recomendaria essi filmi”;
- EN13 /intão/ – $*L >$ “intão eu achu ingraçadu aquela aquela dia lá”;
- EN33/intão/ – $L+*H >$ “intão /hoji assim falanu hoji eu to /muitu decepionadu”.

b) Em início de GT para as vogais:

- EN6 /i/ – $*L >$ /i eli crio dentru da da cadeia eli crio um personagim#/;
- EN8 /é/ – $*L >$ /é num num jeitu mirabulanti fugi da cadeia/;
- EN23 /i/ – $*L >$ /#i todú mundu tem erru/;

- EN31 /i/ – *L > /i elis num pensam -na união/.

c) Em início de GT para inserção de estrutura:

• EN9 /i aí/ – L+*H > “i aí eli conseguiu pega tudu essi dinheru i viveu a vida deli /im aspas feliz né”;

• EN34 /é i/ – *L *L > “é i / assim eu honestamenti eu to sem perspectiva nenhuma#”.

d) Em final de GT para palavras:

• EN9 /né/ - *L > /i aí eli conseguiu pega tudu essi dinheru i viveu a vida deli /im aspas feliz né/;

• EN12 /né/ – *L > /a genti /qui faiz a vida se ingraçada o tristi né/;

• EN14 /aí/ – *L > /quandu eu assisti essi essi episódiu du padri aí#/.

e) Em meio de GT para palavras:

• EN32 /aí/ > L+*H > “cria mintiras aí issu vai disgastanu a genti”;

• EN33 /assim/ > L+*H /intão /hoji assim falanu hoji eu to /muito decepcionadu/;

• EN34 /assim/ > *L+H “é i / assim eu honestamenti eu to sem perspectiva nenhuma#”.

f) Em meio de GT para as vogais:

• EN9 /i/ > *H “i aí eli conseguiu pega tudu essi dinheru i viveu a vida deli /im aspas feliz né”.

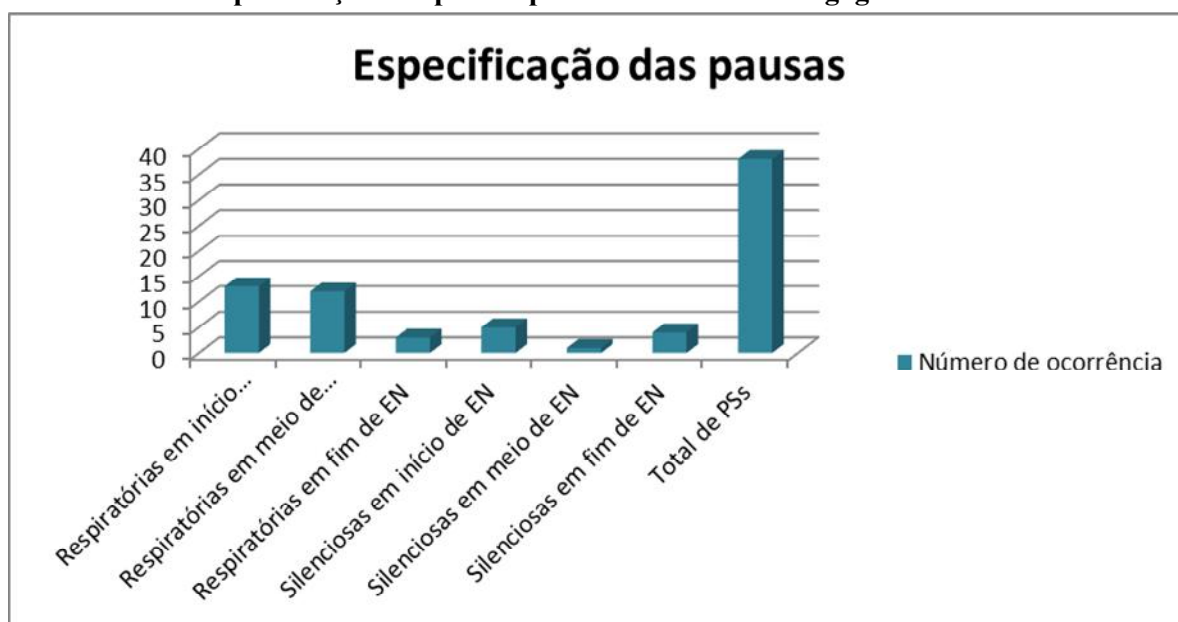
As análises entoacionais evidenciam que o padrão se manteve para INS_v em início de GT (*L) e INS_p (*L) em final de GT. Esse padrão indica que o informante, em início e fim de GT, utiliza uma F0 mais baixa para as INSs. Para os outros casos foi obtido apenas dois padrões iguais para a INS em meio de GT (EN32 e EN33). No entanto, a tendência da Hz para esse subtipo de disfluência apresentou um padrão crescente L>H no sentido da Hz. Para as vogais, em meio de GT, não foi possível fazer comparações, uma vez que tivemos apenas uma realização. Em início de GT ocorre uma variação na tendência da Hz, mas a mais comum é no sentido crescente, sendo apenas EN3 no sentido da queda de Hz. Dentre as INS_p s em início de GT, apenas EN13 foi classificado com uma estrutura simples (*L), isso ocorreu porque o informante manteve uma F0 constante e uma qualidade de voz mais grave.

Para as INS_p s em final de GT, o marcador discursivo em EN9 e EN12 tem uma função interrogativa retórica, ou seja, o informante interroga afirmativamente o seu interlocutor, como se já estivesse implícita a concordância entre interlocutor e falante. EN14, por sua vez,

não tem a intenção de fazer um questionamento, mas sim de retroceder à informação anterior, mostrando a importância dela para o que está sendo dito. Em início de GT, as INS_{ps} analisadas têm como função dar sequência ao GT, sem que ele perca a conexão com o GT anterior. As INS_{vs} , em início de GT, parecem funcionar como preenchedoras, ou seja, o informante, com dificuldades de dar sequência ao discurso, usa uma vogal para suprir essa ausência até que ele encontre a palavra desejada. Já a INS_e divide-se entre dar sequência ao enunciado (EN9) e preencher o vazio (EN34). Para a INS_v em meio de enunciado a finalidade é coesiva e conclusiva. Por fim, a INS_p em meio de GT do tipo /assim/, bem como a do tipo /ai/ tem finalidade sequenciadora.

Como mencionado, em um total de 38 pausas, 18 ocorreram em início de enunciado, 15 em meio de enunciado e cinco em fim de enunciado. Dentre elas, aproximadamente 66,8% foram de pausas respiratórias, 26% de pausas silenciosas e, 5%, articulatórias. O gráfico abaixo apresenta a tendência de cada tipo de PS para uma posição específica do enunciado.

Gráfico 3 – Especificação das pausas para o informante não gago CR.



Fonte: própria.

O gráfico 3 identifica as pausas 1) respiratórias e 2) silenciosas, bem como traz informações sobre as suas posições no enunciado: 1) início; 2) meio ou 3) fim. Ao analisarmos o gráfico, verificamos que para CR as pausas ocorrem com mais frequência em início e meio de enunciado, com respectivamente 34% e 31% aproximadamente das ocorrências. Em final de enunciado, as PSs respiratórias aparecem em aproximadamente 8% das ocorrências. Por outro lado, as PSs silenciosas em fim de enunciado somam 10% do total

das pausas verificadas para a entrevista completa, 13% em início de enunciado e 1% para meio de enunciado. Já as PSs silenciosas em início de GT ocorreram para 13% dos dados e, em meio de enunciado, para 1%.

As PSs silenciosas, de uma maneira geral, são as menos frequentes. Porém, quando elas ocorrem, são mais frequentes em início de enunciado (13%). Portanto, independentemente do tipo de pausa, a preferência é que elas ocorram em início de enunciado para CR. A maior frequência de PSs entre todas as disfluências indica que para CR elas são mais importantes para a construção do enunciado.

A duração média das pausas mostrou que, em início de enunciado, as respiratórias duram aproximadamente 0.518s, em meio de enunciado, 0.706s e, em fim de enunciado 0,513. Já as silenciosas, em início de enunciado, têm duração de 0.357s, em meio de enunciado, de 0.265 e, em fim de enunciado a duração média é de 0.514s. Isso mostra que as PSs respiratórias tendem a ter uma duração maior em meio de enunciado e, as silenciosas, em fim de enunciado.

Quanto à intensidade, todas as PSs, independentemente do tipo, foram identificadas com intensidade. Essa intensidade diz respeito a uma questão aerodinâmica e acústica, ela aparece nas PSs devido à turbulência do ar. Devido a esse fato, esses resultados não foram levados em consideração.

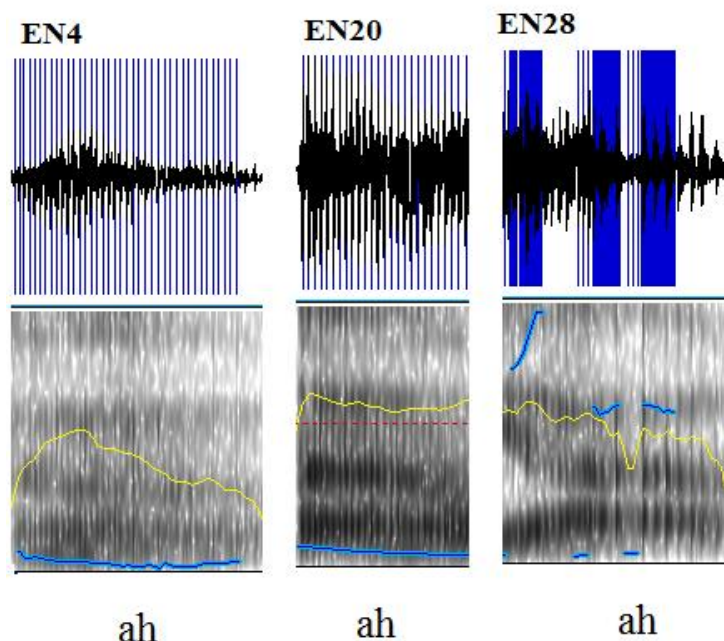
A hesitação (HES) foi um dos tipos de disfluência menos frequente para CR, contando apenas 3,2% (EN4, EN20 e EN28) das produções. Veja as ocorrências:

- EN4 “/ah -comu bom comu se falo assisti várius filmi”;
- EN20 “#ah u u u Rodrigu ta senu uma surpresa muito agradável pra cidadi”;
- EN28 “/é um povu qui vivi ah fazenu fofo”.

Em EN4 e EN20 elas foram identificadas em início de enunciado e logo após uma PS. Por sua vez, em EN28, em meio de enunciado. Nas duas primeiras produções, elas antecedem um BLO (EN4) e, uma 3RP_v (EN20). Já em EN28 não foram identificadas disfluências que seguem ou antecedem a HES. Elas foram identificadas nos seguintes enunciados:

Como mencionamos, os três tipos de HES apresentam uma qualidade de voz do tipo *creaky voice*. Esse tipo de qualidade de voz evidencia-se por pulsos mais próximos e F0 baixa. Veja a imagem que caracteriza essa afirmação:

Figura 27 - Representação das hesitações nos enunciados EN4, EN20 e EN28 produzidos por CR.



Fonte: própria.

A imagem identifica as hesitações em EN4, EN20 e EN28. Como vemos, a F0 é baixa para as três ocorrências. Em EN28 ocorre um erro de gravação, no entanto, essa interferência não descaracteriza o padrão grave realizado pelo informante, pois dentre esses três enunciados, em EN28, CR apresentou uma qualidade de voz mais grave do que a de EN4 e EN20.

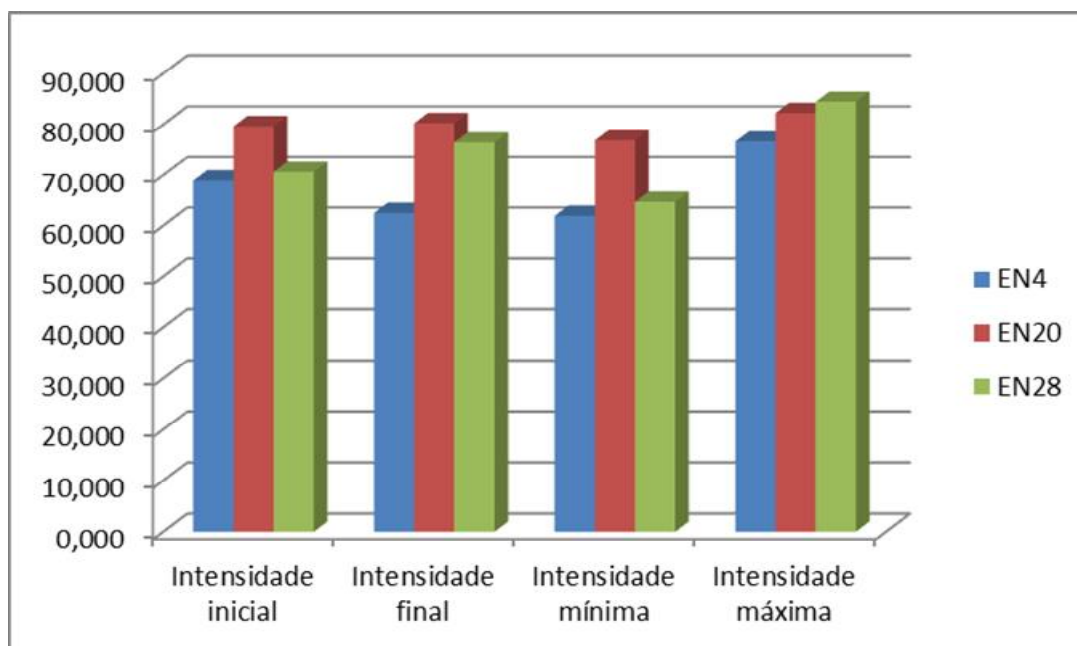
Quanto à F0, os valores referentes à frequência mínima e máxima foram:

- EN4 - F0 mín (79 Hz) F0 max (106 Hz)
- EN20 - F0 mín (92 Hz) F0 máx (107 Hz)
- EN28- F0 mín (86 Hz) F0 máx (0 Hz)

Por meio dos valores da F0 máxima e mínima, vemos que a frequência manteve-se praticamente constante. De acordo com os tons propostos por Halliday (1970) e Cagliari (1992/2007) e, de acordo com a variação de tessitura proposta por Carvalho (2014), vemos que as variações de F0 para as hesitações em CR são típicas de um tom baixo. Fato que é corroborado por meio da análise entoacional feita no modelo de Pierrehumbert (1980/1986) que apresentou um tom do tipo *L, como vimos. Em EN28, devido a problemas de leitura do *software*, não foi possível identificar a F0 máxima.

Quanto à intensidade, veja o gráfico 4 abaixo:

Gráfico 4 – Análise das intensidades inicial, final, mínima e máxima para a hesitação hesitação para o informante não gago CR.



Fonte: própria.

O gráfico 4 apresenta os valores referentes à intensidade inicial, final e intensidades mínimas⁵⁸ e máximas para cada uma das três produções. Ao analisarmos o gráfico juntamente com a figura 58, observamos que EN4 apresentou a menor intensidade entre todas as ocorrências quando comparado com EN20 e EN28. EN20. Por sua vez, EN4 apresentou uma intensidade menor do que EN28 apenas em relação à intensidade máxima. Já EN28 apresentou intensidades mais próximas a EN20, sendo a maior diferença em relação à intensidade mínima. Dentre as intensidades anotadas, a maior variação esteve entre as intensidades mínima e final. Não foi possível verificar a intensidade de transição da fluência para a fluência em EN4 e EN21, uma vez que elas ocorrem em início de enunciado, vindo de uma PS. Por outro lado, em EN28 a passagem da fluência para a disfluência mostrou uma diminuição na intensidade 80 dB > 78 dB. Comparando os resultados do gráfico ainda com a figura 58, vemos que EN20 apresentou as intensidades mais altas. Isso mostra que a posição da hesitação no enunciado não determina um padrão para as intensidades.

Nas três produções os pulsos foram: 39 (EN4), 25 (EN20) e 54(EN28). Se formos compará-los com a duração veremos que eles são proporcionais, ou seja, quanto maior a duração, maior o número de pulsos. A duração de EN4 foi de 0.481s, a de EN20 de 0.268s e a

⁵⁸ A intensidade mínima foi analisada para este informante apenas para fins ilustrativos.

de EN28, de 0.483s. Além disso, as durações mostram que o informante não mantém o padrão duracional nas hesitações, seja na mesma posição ou em posições diferentes.

A disfluência do tipo PI foi a que menos ocorreu para CR: apenas 1% das produções. Em casos como esse, a palavra é recuperada pelo contexto. Esse tipo de disfluência foi identificada em final de enunciado (EN28), com o segmento /fofo/. A sua duração é de 0.599s. Normalmente, na PI, o informante apresenta uma velocidade de fala maior do que a normal. Mas esse não foi o caso dessa produção, uma vez que o informante simplesmente parou a produção nas duas sílabas iniciais. Vale salientar que essa PI não ocorreu por um erro de segmentação do pesquisador. O informante simplesmente termina esse GT por meio dessa produção. É possível que o informante tenha terminado o GT nesse momento, pois pressupõe que o seu interlocutor, no conjunto do enunciado, tenha compreendido a palavra em seu sentido completo.

No caso dessa PI, a F0 é mais baixa e termina com um padrão ascendente. Assim, a F0 mínima foi de 97 Hz e a máxima de 114 Hz. Na passagem da estrutura fluente para a disfluente ocorre uma diminuição na intensidade que vai de 76 dB para 74 dB. A intensidade máxima é de 84 dB A intensidade final é de 79 dB, aumentando em relação à inicial.

No conjunto das disfluências, o BLO também não foi muito significativo, somando 4,3% das ocorrências. Como foi mencionado, em EN1 (cf. fig.56) e EN4 (cf. fig.57) ocorreu uma PS articulatória. Já em EN26 (cf. fig.57) não ocorreu nenhuma PS antes do BLO e, em EN31 (cf. fig.57), uma PS respiratória.

Em relação à análise acústica do BLO com PS articulatória, em EN1 a palavra com bloqueio apresentou uma duração de 0.679s, enquanto o período articulatório apresentou uma duração de 0.204s. Quanto à intensidade, no momento de transição ela não foi analisada. Já a intensidade final aumentou em relação à inicial. A média da F0 mostra uma frequência mais baixa na 126 Hz.

Em EN4, também com PS articulatória, a duração é de 0,318s, que é aproximadamente a metade da duração de EN1. O período articulatório apresenta uma duração de 0.276s, com uma diferença de 0.072s em relação a EN1. Quanto à intensidade, também ocorre um aumento da final em relação à inicial. A média da F0 é de 189 Hz

Em EN26 não ocorreu PS, o bloqueio é percebido auditivamente e também por meio de algumas tendências acústicas. Por exemplo, a F0 começa mais baixa com 115 Hz e aumenta rapidamente para 227 Hz. Esse valor representa um erro, um ruído muito agudo. Essa tendência não é verificada em EN1 e EN4, mas nesses enunciados a F0 também não

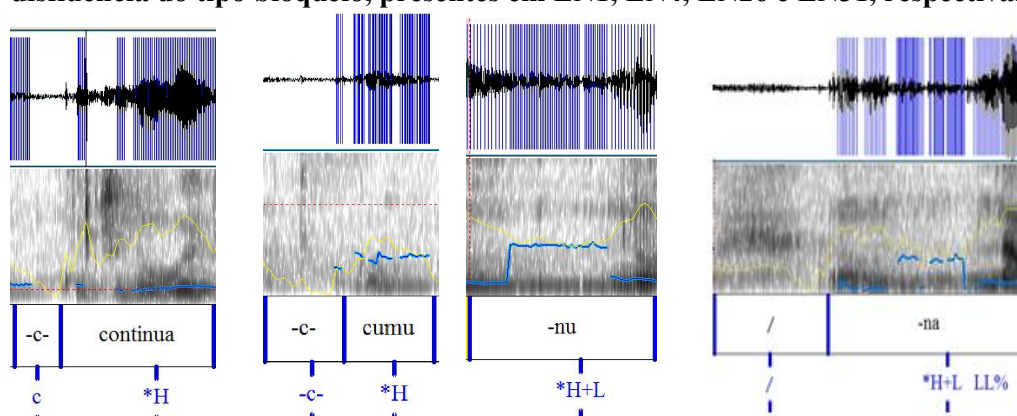
segue um padrão constante, apresentando variações. EN31, por sua vez, apresenta o mesmo padrão da F0 de EN26, com uma intensidade que começa com uma frequência de 101 Hz, passando rapidamente para 198 Hz. EN26 foi anotada com uma duração de 0.451s que, de acordo com o tamanho do segmento representa uma produção mais longa. Quanto à intensidade, há uma diminuição na transição do período fluente para o BLO_c, com conseqüente aumento na intensidade final.

Por fim, EN31 tem uma duração de 0.880s. Assim como acontece para EN26, EN31 apresenta uma duração maior do que a normal para o segmento.

A duração da PS, respiratória, nesse caso, foi de 0.495s, que é maior do que as articulatórias. Assim como aconteceu em outras produções houve um aumento da intensidade em final de enunciado.

Em relação aos pulsos, o esforço vocal foi maior para as consoantes nasais dentais [n] do que para as oclusivas velares [k]. Observe a imagem a seguir:

Figura 28 – Caracterização dos pulsos, intensidade e frequência fundamental na disfluência do tipo bloqueio, presentes em EN1, EN4, EN26 e EN31, respectivamente.



Fonte: própria.

Nas figuras estão identificados os pulsos (linha vertical azul), a F0 (linha horizontal em azul), a intensidade (linha amarela) e a forma de onda obtida nas palavras. As imagens identificam 1) bloqueios com pausas articulatória (primeira figura); 2) bloqueio articulatório com pausa silenciosa (segunda figura); 3) bloqueio com ausência de PS (terceira figura) e 4) bloqueio com PS respiratória (última figura). Vale salientar, que mesmo sendo respiratória observamos a presença de articulação.

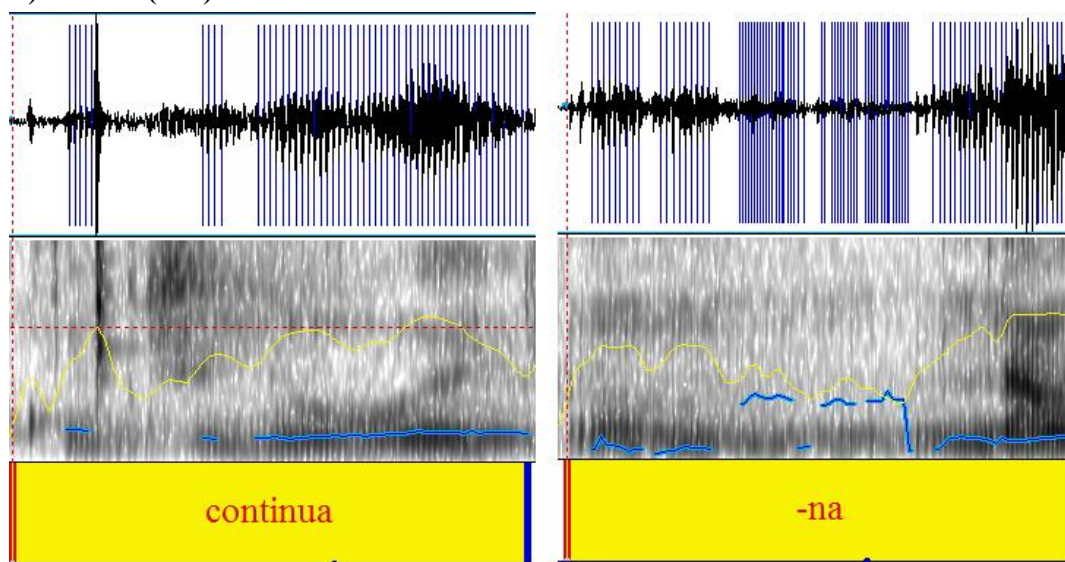
Dessa maneira, é possível dizer que o bloqueio tende a ocorrer com uma PS anterior, bem como com um aumento da intensidade final. A análise das PSs articulatórias mostrou um esforço vocal menor durante a sua articulação. A duração das PSs articulatórias também

evidenciou que elas mantêm um padrão, com uma média de 0,240s. Por outro lado, a duração da PS em EN31 prova que ela é respiratória, uma vez que esteve dentro da média proposta para as PSs respiratórias, que foi de 0.518s.

Sendo assim, resumidamente, observamos uma duração média maior para as nasais alveolares (0,665s), enquanto que as oclusivas velares apresentaram uma duração média de 0,498s, bem como a intensidade máxima (76 dB e 82 dB). Já a F0 foi maior para as oclusivas do que para as nasais (157 Hz e 155 Hz).

Outra questão interessante observada durante a análise dos BLOs refere-se a uma interrupção na forma de onda. Observamos que tanto para as oclusivas quanto para as nasais, CR interrompe a sua produção. Veja o exemplo a seguir para EN1 e EN31.

Figura 29 – Representação da interrupção da forma de onda para os enunciados EN1 (29a) e EN31 (29b).



Fonte: própria.

Para fins didáticos, a figura 29 identifica apenas o momento em que o fenômeno ocorre. A imagem da esquerda (29a) mostra a interrupção na forma de onda para a oclusiva e a da direita (29b), para a nasal. O que ocorre nessas produções é uma repetição da fonação, devido à soltura das cordas vocais. Em 29a observamos dois momentos de soltura. O informante solta a obstrução da oclusiva e não dá sequência a vogal seguinte, bloqueando mais uma vez. Somente após as duas solturas, que duram, respectivamente, 0,127s e 0,036s, ele retoma a fluência. Veja que o segundo momento é mais breve do que o primeiro. Em 29b verifica-se o mesmo processo de 29a. Mas, nesse caso, as interrupções continuam ao longo do segmento /-na/. No que se refere à consoante nasal, as suspensões duram, no primeiro momento, 0,027s e, no segundo, 0,047s. Essa característica mostra que não há continuidade

no segmento durante a produção do bloqueio, sendo construído por etapas. O que temos, então, são pedaços muito curtos de articulação separados por pausas muito breves. A tentativa de fluência é perdida, portanto, por meio dessas breves sequências de pausas.

A análise da INS foi dividida entre inserção de palavra e inserção de vogal. Dentre elas, subdividimos a análise entre as que ocorrem em início, meio e fim de enunciado.

Em início de enunciado CR produziu quatro INS_ps com a palavra /intão/ e uma com /ái/. Os enunciados em que elas ocorreram foram: EN3, EN7, EN10, EN13 e EN33. A produção de /intão/ foi verificada para EN3, EN10, EN13 e EN33. Dentre essas realizações, apenas EN10 ocorreu após uma pausa.

Quanto à INS_p de /intão/, verificamos que a sua duração apresentou uma média de 0.304s. Os pulsos, por sua vez, apresentaram uma variação maior, sendo para EN3 31 pulsos; EN10 20 pulsos; EN13 9 pulsos e EN33 43 pulsos. Esse resultado mostra que em EN13 a qualidade de voz é mais grave. A análise da média da F0 além de provar a qualidade de voz *creaky voice* do informante (83 Hz), também evidenciou uma tessitura mais baixa. Já para a análise da intensidade não foi possível verificar a transição da fluência para disfluência, uma vez que elas ocorrem em posição inicial. A análise da intensidade final mostrou uma tendência para o aumento da intensidade em EN3, EN10 e EN13. No entanto, essa tendência não foi verificada para EN33 porque depois dessa disfluência ocorre uma PS respiratória, que caracteristicamente leva a uma diminuição da intensidade. A média das intensidades máxima (73 dB) mostra que para essa produção a intensidade é mais baixa, assim como veremos durante a análise.

A análise de /ái/ (EN7) revelou uma duração menor (0.230s) do que a média observada para /intão/ (0.304s). Os pulsos (29) também caracterizaram uma qualidade de voz mais grave do que algumas produções de /intão/. A média da F0 (143 Hz) identifica que para /ái/ o informante utiliza uma tessitura mais alta do que para /intão/. De acordo com a análise proposta em Carvalho (2014) esse tom seria identificado como médio. Por fim, a análise da intensidade mostrou que, para /ái/, a intensidade média em início de palavra é mais alta (80 dB) do que a média identificada para /intão/ (75 dB). Quanto à intensidade final, há uma diminuição. Essa diferença pode estar relacionada com a produção do segmento que segue essa disfluência, uma vez que em EN3, EN10 e EN13 são produzidas vogais e em EN7, temos a produção de uma oclusiva velar [k].

Por sua vez, em meio de enunciado foram produzidas INS_ps com palavra /assim/ e /ái/. Quanto ao primeiro tipo, a duração foi de 0.400s (EN33) e 1.033s (EN34). A média da F0

mostrou que EN34 (170 Hz) apresenta uma tessitura mais alta do que EN33 (106 Hz). A análise dos pulsos foi de 42 para EN33 e de 84 para EN34. Por fim, houve uma diminuição da intensidade em EN33, na passagem da fluência para a disfluência. Já em EN34, ocorreu um aumento da intensidade. No caso de EN34, essa diferença já é esperada, uma vez que a INS_p ocorre depois de uma PS respiratória. Em EN33 há uma diminuição da intensidade e, em EN34, um aumento. A média das intensidades máxima mostra que nos dois exemplos elas são próximas, com uma intensidade de 75 dB para EN33 e de 73 dB para EN34.

Para /ái/, em meio de enunciado, a duração é de 0.366s, maior do que a mesma produção em início de enunciado. Os pulsos também são maiores nessa posição, somando 39. A média da F0 também distancia as duas produções, já que EN32 apresentou uma tessitura mais baixa do que /ái/ em início de GT. Na transição da fluência para a disfluência ocorreu um aumento da intensidade. Por outro lado, as duas produções mostram uma tendência para a diminuição da intensidade final quando comparada à inicial. A pequena variação entre as médias das intensidades (80 dB em início de enunciado/ 77 dB em meio de enunciado) revela uma proximidade entre as intensidades.

Finalmente, em fim de enunciado as INS_p s foram /ái/ (EN14) e /né/ (EN9 e EN12). Em relação ao primeiro tipo, foi verificada uma duração menor do que para as posições em início e meio de enunciado. Os pulsos para a INS_p em final de GT foram menores (16) do que os ocorridos para as INS_p s em meio de GT (39), o que evidencia um padrão mais grave para esse tipo de INS_p em final de enunciado. A média da F0 mostrou uma variação maior do que em EN7 e EN32. No entanto, essa diferença mais alta deve-se a um erro no *software*. Se analisarmos a F0 mínima (82 Hz) e máxima (87 Hz) veremos que elas são baixas, comprovando, assim, a sua qualidade mais grave. A análise da intensidade caracterizou uma diminuição na transição do período fluente para o disfluente, diferente de EN32. A intensidade final diminuiu em relação à inicial, bem como anotado para as outras posições. A média das intensidades máxima (75.dB) também evidencia um padrão para os três momentos.

Para o segundo tipo de INS_p /né/, em final de enunciado, a duração das duas produções foi de 0.462s (EN9) e de 0.267 (EN12), com uma diferença de 0.195s. A diferença entre os pulsos foi de 10, sendo 34 para EN9 e de 24 para EN12, o que identifica um esforço vocal menor para EN12. A média da F0 foi de 112 Hz para EN9 e 101 Hz para EN12. Como podemos ver, a diferença entre as médias da F0 é pequena. A intensidade, por sua vez, caracteriza um aumento na transição da fluência para a INS_p . Em final de enunciado a

intensidade aumenta para EN12 e diminui para EN9. A diferença (4 dB) entre as médias das intensidades (70 dB/ 74 dB) mostrou a proximidade entre elas.

Foram realizadas quatro INSs de vogal em início de enunciado, sendo duas simples, com vogal /i/ (EN6/ EN31) e duas complexas (com alongamento - INS_v+AL). Dentre elas, uma foi com a vogal /é/ (EN8) e a outra com vogal /i/ (EN23). Quanto à INS_v simples, a primeira (EN6) ocorreu depois de PS respiratória e teve uma duração de 0.129s. Para EN31 não foi identificada PS e a sua duração foi de 0.071s. A análise da F0 identificou uma tessitura mais baixa para essas produções, sendo a média de ambas 126 Hz e 105 Hz, respectivamente. Os pulsos somaram 14 para EN6 e de 6 para EN31. A média das intensidades máximas, por sua vez, mostrou que elas são próximas, sendo a de EN6 de 75 Hz e a de EN31 de 74 Hz. A intensidade final apresenta uma tendência ascendente.

As inserções com alongamento apresentaram uma duração de 0.688s para EN8 e 0.588s para EN23. Esse resultado prova que elas diferenciam-se das durações simples. Os pulsos, por sua vez, foram de 59 e 24. Vale salientar que ocorreu um erro de leitura do *software* em EN23. Para as INSs complexas, a tessitura foi baixa (92 Hz e 90 Hz), mais baixa do que a anotada para as INSs simples. A intensidade também seguiu uma tendência para o aumento final. A intensidade máxima foi maior para EN8 (81 dB) do que para EN23 (74 dB).

Em meio de enunciado, as INSs também ocorreram para a vogal /i/ nos enunciados EN9 e EN22. Em relação às durações de EN9 e EN22 (0.236s e 0.404s), elas foram maiores do que as observadas para as INSs em início de enunciado (mas menores do que as INSs com alongamento). Os pulsos variam entre 53 (EN53) e 43 (EN22). Quanto à F0 média, ocorreu um erro em EN9. Mesmo assim, por meio da frequência mínima (136 Hz) e máxima (189 Hz), foi possível observar um tom mais alto do que o encontrado para as INSs em início de enunciado. EN22 manteve a tessitura baixa, com uma média de 103 Hz.

A análise da intensidade mostrou um aumento na passagem da fluência para a disfluência em EN22 e para EN9, um abaixamento. Por outro lado, a intensidade final aumentou, diferentemente do que havia sido caracterizado para as INS_v em início de enunciado. Por fim, a intensidade máxima média foi próxima entre as duas ocorrências, sendo de 74 dB para EN9 e de 78 dB para EN22.

O alongamento, por sua vez, foi verificado para palavras e vogais. O AL do primeiro tipo foi dividido em alongamento de sílaba em palavra e alongamento em palavra (AL_p). Dentre os alongamentos de sílaba em palavra foram encontrados dois tipos: 1) alongamento em fim de palavra e 2) alongamento em início de palavra. O AL do tipo 1) ocorreu para dois

enunciados (EN8 e EN11) para as palavras /mirabulanti:/ e /vivu:/, sendo o primeiro em meio de enunciado e, o segundo, em início. A duração da palavra total foi de 1.144s para /mirabulanti:/ e 1.256s para /vivu:/. A sílaba final alongada, por sua vez, teve uma duração de 0.780s e de 1.121s. Isso mostra que as sílabas restantes apresentaram uma duração de 0.664s e 0.235s, respectivamente. Em relação à palavra /mirabulanti:/ as outras quatro sílabas restantes teriam uma duração de 0.166s. O tamanho dos segmentos mostra que sílaba /vu:/ da palavra /vivu:/ tem uma duração maior do que a /ti:/ da palavra /mirabulanti:/. A análise da F0 mostrou uma tessitura mais baixa para /mirabulanti:/. Os pulsos não destoaram muito, sendo de 138 para EN8 e de 132 para EN11.

Na transição da fluência para a disfluência houve um aumento da intensidade para EN8, que não foi verificado para EN11, por ter ocorrido em início de enunciado. Já a intensidade final diminuiu para EN8 e aumentou em EN11. A média das intensidades máximas foram próximas para as duas produções: 78 dB (EN8) e 76 dB (EN11).

O AL de sílaba inicial ocorreu apenas em EN11, ele foi realizado na sílaba /mi:/ da palavra /mi:nha/. A duração da palavra total é de 0.921s e a do segmento alongado de 0.706s, restando uma duração de 0.215s para a sílaba /nha/. A F0 também caracterizou uma tessitura baixa (101 Hz). Os pulsos somam 74. A análise da intensidade identifica um aumento na transição da fluência para a disfluência, bem como um aumento da intensidade final. A média da intensidade, por sua vez, caracteriza uma intensidade mais baixa (71 dB) do que a analisada em início de enunciado.

O AL em palavra total ocorreu para o artigo /um:/ (EN24) e para o pronome /qui:/, ambas em meio de enunciado. As durações foram de 0.661s e 0.749s, que são próximas da duração das sílabas em EN8 (fim de palavra) e EN11 (início de palavra). A F0, assim como ocorreu nas outras posições apresentou uma frequência também baixa. Em EN24 também ocorre um erro de leitura.

Os pulsos somam 99 para EN24 e para EN28. A análise da intensidade para EN24, bem como está identificado pela figura acima (cf. fig. 59), mostra uma diminuição na transição da fluência para a disfluência (78 dB > 78 dB), assim como para a intensidade final (65 dB). EN28 mantém o padrão da transição, mas a intensidade final aumenta. Em EN24 e EN28 a intensidade final mantém o padrão observado para o AL em final de palavra.

O último tipo de AL foi o identificado para as vogais⁵⁹. Para CR, encontramos apenas uma ocorrência: em EN18 para a vogal /a:/. O AL_v ocorreu em meio de enunciado, a sua

⁵⁹ Vale salientar que os outros tipos de ALs foram analisados nas INSS complexas.

duração foi de 0.555s, próximo ao observado para o AL_p em final e início de sílaba, bem como para o AL em palavra completa. A F0 também caracteriza um tom baixo, que é próprio desse informante. Os pulsos somam 54, e a análise da intensidade mostra um aumento na transição e uma diminuição da intensidade final, bem como para EN8, EN24 e EN28. A média da intensidade (79 dB) mostrou a sua proximidade com EN8, EN11, EN24 e EN28.

A repetição para CR ocorreu em palavras e sílabas em palavras. Quanto à RP de palavras foram observadas RPs duplas e triplas. A RP tripla de palavra ($3RP_p$) ocorreu em EN5 e EN19, para as palavras monossílabas /dum/ e /foi/. EN5 caracteriza uma disfluência simples. As durações foram de 0.286s (RP_1), 0.307s (RP_2) e, 0.291s (RP_3), com uma média de 0.294s. Como vemos, a maior duração foi de 0.307s, que ficou entre as outras duas produções. Mesmo assim, a diferença de RP_2 para as outras foi de 0.021s (RP_1) e 0.016s (RP_3). Já entre RP_1 e RP_3 a diferença foi de 0.005s. A maior diferença ocorreu entre RP_1 e RP_2 . Mesmo sendo palavras iguais, os pulsos variaram entre 22 (RP_1), 32 (RP_2) e 28 (RP_3). Mais uma vez, maior para a segunda RP. A análise da F0 média mostrou que as $3RPs$ seguem um padrão muito próximo, sendo de 99 Hz (RP_1), 99 Hz (RP_2) e 99 Hz (RP_3). A intensidade média também mostrou essa proximidade, sendo de 75 dB para RP_1 , de 75 dB para RP_2 e de 75 dB para RP_3 . Nas $3RPs$ a transição tanto na passagem da fluência para a disfluência, quanto entre as duas disfluências ocorreu uma diminuição na intensidade. No entanto, apenas para RP_2 a intensidade final aumentou.

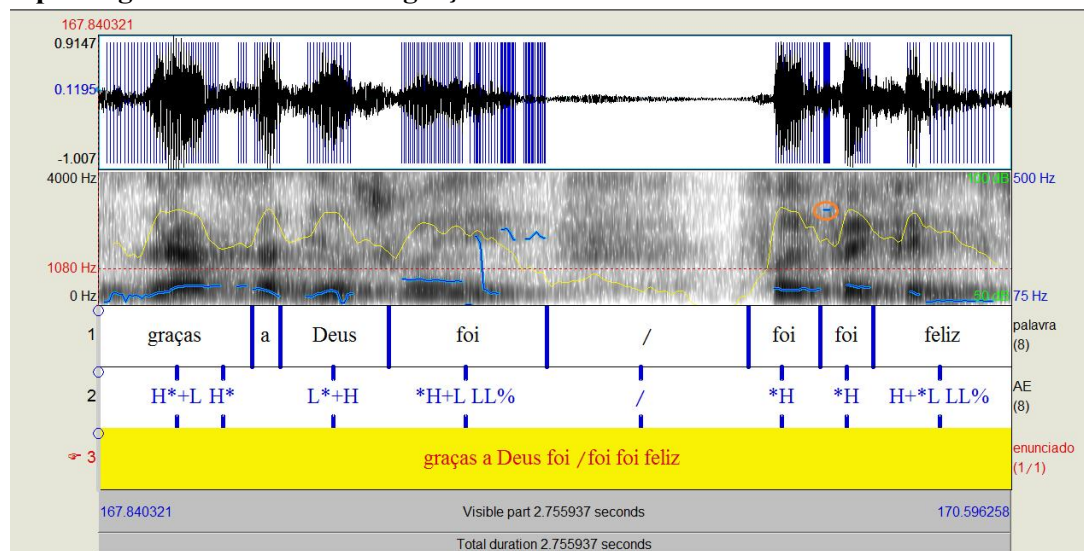
Em EN19, por sua vez, a disfluência foi complexa, uma vez que entre as produções ocorreu uma PS respiratória. Por isso, esse tipo de disfluência foi anotada como $3RP_p + PS_e$. Essa PS ocorreu depois da primeira RP. Essa PS funciona como uma estratégia articulatória de recuperação da fluência, que durou 0.608s. Entretanto, ela não ajuda CR a voltar a fluência, quebrando mais uma vez o discurso com mais duas RPs. Além disso, o informante CR alonga um pouco mais a produção da primeira repetição /foi/, que dura 0.476s, sendo maior do que a duração das duas produções juntas (0.027s e 0.193s).

A média da F0 mostrou que as frequências são diferentes para as três RPs (146 Hz, 124 Hz e 132 Hz, respectivamente). No entanto, a diferença entre elas é pequena. Os pulsos variaram entre 66, 17 e 19, respectivamente. Por fim, a intensidade foi muito próxima entre as duas RPs finais (81.410 dB e 81.269 dB), com uma diferença de 0.141 dB. Já para a primeira RP a intensidade foi mais baixa com 76 dB aproximadamente. Quanto à transição da fluência para a disfluência, não é possível propor um padrão, uma vez que a PS respiratória que ocorre

entre as produções leva a um abaixamento e posterior aumento da intensidade. Por outro lado, a intensidade final diminuiu para todas as produções.

Veja a figura 30 que exemplifica algumas das características apresentadas:

Figura 30 – Identificação da frequência fundamental, intensidade, forma de onda e espectrograma no enunciado “graças a Deus foi/foi foi feliz”.



Fonte: própria.

Observe, para a primeira RP, a quebra na estrutura de F0, bem como a PS que segue a repetição. Veja, ainda, a proximidade entre as duas últimas RPs quanto à duração, F0 e intensidade, que por sua vez, parecem ter o mesmo pico. A diferença entre elas é mais visível para os pulsos. Por outro lado, vemos que o momento em que os pulsos estão mais unidos deve-se a um ruído. Esse fenômeno aparece identificado pelo círculo em vermelho na frequência (F0).

Ainda para a RP_p foram identificadas seis ocorrências com repetição dupla em monossílabos (2RP_p): EN6; EN8; EN10; EN15 /tem/, /tem/; EN22 /pra/, /pra/ e EN30 /pra/, /pra/. Quanto às durações, todos os enunciados tiveram pequenas variações. Veja a duração para cada enunciado, bem como a diferença na variação anotada entre parênteses:

- EN6 – 0,180s – 0,179s (0,001s);
- EN8 – 0,316s – 0,312s (0,004s);
- EN10 – 0,173s – 0,170s (0,003s);
- EN15 – 0,236s – 0,219s (0,017s);
- EN22 – 0,241s – 0,228s (0,013s);
- EN30 – 0,241s – 0,140s (0,101).

Como podemos ver, a maior diferença ocorre em EN30. As outras diferenças são muito pequenas. Vemos, ainda, que a primeira repetição de EN22 e EN30 apresentam a mesma duração. Mesmo que as repetições tenham ocorrido para a mesma palavra nos dois enunciados, esse foi resultado apenas uma coincidência, uma vez que para os outros informantes esse similaridade não ocorreu. A média da F0 mostrou um resultado próximo para cada dupla de RPs. Apenas em EN10 houve uma quebra de estrutura, devido à interferências na gravação. Observe:

- EN6 – 111 Hz – 105 Hz (6 Hz),
- EN8 – 96 Hz – 97 Hz (1 Hz);
- EN10 – 115 Hz – 139 s (24 Hz);
- EN15 – 113,629 Hz – 113,086 Hz (0,543 Hz);
- EN22 – 94 Hz – 95 Hz (1 Hz);
- EN30 – 113 Hz – 124 Hz (11 Hz).

Os pulsos, por sua vez, não mantiveram o mesmo padrão em todos os eventos. A análise da média da intensidade também apresentou uma pequena variação, uma vez que entre os grupos elas mantiveram uma mesma tendência. Por exemplo, em EN10 os resultados obtidos foram 75 dB para a primeira RP /eu/ e 74 dB para a segunda RP/eu/. Em EN22 e EN30, que apresentaram a mesma palavra repetida, a intensidade foi próxima tanto na individualidade dos enunciados quanto na comparação entre eles. Veja:

- EN22 – 80,946 dB – 80,916 dB (0,03 dB),
- EN30 – 80,631 dB – 81,236 dB (0,605 dB).

Entre as repetições, a maior diferença verificada diz respeito à variação da intensidade no momento de transição da fluência para a disfluência e para a intensidade final. Em EN6 e EN15 a intensidade aumentou na transição. Já para EN10 houve uma variação, ou seja, na primeira RP a intensidade aumentou e, na segunda, diminuiu. Já em EN8 a intensidade diminuiu para as duas ocorrências. Por sua vez, para EN22 e EN30 as RPs foram iniciais e seguidas de PS na primeira RP, mas na segunda RP foram identificados uma diminuição e aumento da intensidade. A intensidade final teve uma tendência ascendente. A intensidade máxima, por sua vez, diminuiu da primeira para a segunda RP em 57% dos casos.

Para CR também foi verificada repetição dupla em palavra dissílaba $2RP_{p2}$ /essi/ (EN14) e trissílaba $2RP_{p3}$ /aquele/ (EN15). De maneira contrária do que vimos nas 2RPs para monossílabos, nesses dois casos a diferença na duração foi significativa. Em EN13 as durações foram: 0.425s e 0.729s e em EN14, 0.181s e 0.326s. Isso ocorreu porque, no

primeiro caso, o informante ri durante a primeira RP, o que contribui para o erro na F0 e, na segunda RP, ele respira. Já em EN14, a qualidade de voz fica tão grave que interfere na gravação. A tessitura do informante é baixa.

A análise da intensidade durante a transição da fluência para a disfluência, mostrou um aumento para EN13 e uma diminuição para EN14. No entanto, entre as RPS ocorreu o contrário, ou seja, em EN13 houve uma diminuição e em EN14 um aumento. Já a média das intensidades máximas comprovou a semelhança entre as RPs.

O informante também realizou uma sequência de repetições iguais no mesmo enunciado (EN17), do tipo /ca/. É importante salientar ainda que as duas RP_s não ocorreram simultaneamente. Elas são separadas por uma pausa, que indica um novo GT. No que se refere às durações, a primeira dupla de RPs duraram, respectivamente, 0.205s e 0.284s, como uma diferença de 0,079s entre elas. O segundo grupo, por outro lado, apresentou as durações de 0.234s e 0.243s, o que leva a uma diferença menor entre as RPs (0.009s). Isso mostra que, no segundo momento, o informante chegou à fluência mais rapidamente. Mesmo assim, vemos que CR procura manter a articulação para as quatro RPs. Os pulsos também mostram a proximidade entre as RPs, uma vez que no primeiro GT os pulsos foram 10 e 18 e, no segundo, 10 e 19. A média das F0 também mostra o tom mais baixo dos eventos e a qualidade de voz grave. Em relação às intensidades, houve uma diminuição tanto na transição da fluência para a disfluência quanto entre as RPs nos dois eventos. De modo semelhante, a intensidade aumentou nos dois grupos de RPs. Já a média da intensidade mostrou que a primeira RP do primeiro GT, bem como a segunda RP do segundo GT, é mais alta. No entanto, a diferença entre a média da intensidade nos dois grupos foi de 5,115 dB e 5,035 dB.

Por fim, o último tipo de RP foi o verificado para as vogais. Em CR essa disfluência foi tripla (3RP_v) e ocorreu com a vogal /u/ em EN20. A duração das três ocorrências foram: 0.236s, 0.440s e 0.129s. Como vemos, a segunda RP é a que dura mais, com uma diferença de 0.204s da primeira e 0.311s da terceira. Quanto à F0, apenas a primeira RP_v apresentou um padrão constante. Nas outras duas ocorreu interferência na gravação. Contudo, nas três RP_vs a média da F0 evidenciou um padrão baixo (100 Hz, 94 Hz e 93 Hz), sendo as duas últimas com uma qualidade de voz mais grave. Os pulsos variaram entre 24, 61 e 18. No período de transição da fluência para a disfluência (primeira RP_v) a intensidade aumentou, bem como na transição das RP_vs. A média das intensidades máxima mostrou que a primeira RP_v foi anotada com uma intensidade maior do que e a terceira. Mesmo assim, a diferença da maior intensidade em relação às outras duas foi de 2.051 dB e 4.046 dB, o que as aproxima.

Diante das observações feitas para as repetições, verificamos que há uma tendência para que a intensidade máxima e a frequência fundamental sigam um mesmo padrão. A F0, por sua vez mantém uma tessitura baixa. Normalmente, ela é mais constante para a primeira repetição, quebrando na segunda ou terceira RP, principalmente em momentos em que há uma qualidade de voz do tipo *creaky* voice. Quanto à duração, o resultado também é parecido, mas a diferença, às vezes, é um pouco maior. Algumas vezes, essa diferença deve-se a momentos de risadas, respirações ou pausas, tanto na própria repetição ou entre elas. A transição da fluência para a disfluência, no conjunto das produções mostra uma tendência para a diminuição.

O último tipo de disfluência realizada por CR foi a correção. Essa disfluência refere-se ao fato do falante se autocorrigir quando percebe que produziu uma sentença inadequada. Como vimos, isso pode ser feito por meio duas estratégias: 1) o informante corrige o erro, mas mantém parte da informação do GT errado e 2) o informante constrói um novo GT.

Delfino (2009) fez uma análise das correções, chamadas por ele de disfluências de reparo. Em relação à F0, a sua análise mostrou que, quando comparado ao trecho com problema, não há alteração significativa do contorno melódico durante a correção. No entanto, o autor observou para 50% dos dados que a sílaba tônica saliente se desloca para o trecho da correção. Para CR a análise feita não corroborou com os dados do autor, já que em 62,5% das correções o trecho de maior intensidade manteve-se no erro. Diferentemente de Delfino, esses resultados mostram que o falante tende a não contrastar a informação durante a correção. Por outro lado, a análise da intensidade para CR mostrou, assim como nos resultados obtidos pelo autor, que esse parâmetro tende a ser maior na correção. Como mencionamos, a intensidade foi analisada em momento de transição, fim de enunciado e por meio da intensidade máxima. Em relação à intensidade de transição, houve um aumento na passagem da fluência para a correção em 62,5% dos casos. Já na transição do erro para a correção, houve uma diminuição da intensidade em 70% dos casos. A intensidade máxima tende a ser maior na correção em 62,5%. A intensidade final, por sua vez, também aumentou em 62,5% dos casos.

A duração também seguiu o mesmo padrão, ou seja, o informante tende a aumentar a duração no momento na correção (62,5%). O que se justifica, na maior parte das vezes, pelo fato de que o informante tende a trazer informações novas no reparo, aumentando assim o GT. Em outras palavras, o falante utiliza um tempo maior para formular a correção. Por fim, a análise dos pulsos não mostrou uma diferença significativa entre o erro e a correção. Isso mostra que o informante tende a manter a mesma qualidade vocal nos dois momentos (erro e

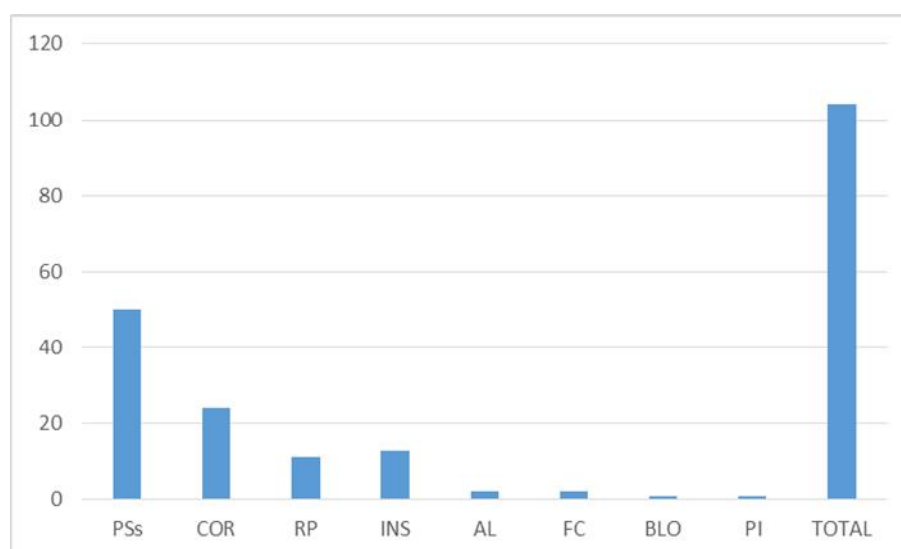
correção). É importante lembrar que em EN21 e EN26 ocorreu uma PS antes da correção. Isso mostra que nesse momento o informante planejou mais a sua fala do que nas outras correções.

O informante MA

Em 38 enunciados, o informante não gago MA realizou 104 disfluências. O máximo de disfluência observada por enunciado foi de seis, sendo três do tipo PS. Dentre elas, 18 foram complexas. A sua taxa de elocução foi de 4,264 sil/ms, a de articulação de 5,623 sil/ms e o tempo de articulação aproximado foi de 101ms.

De um ponto de vista geral, o gráfico abaixo identifica as disfluências para MA:

Gráfico 5 – Os tipos de disfluências realizadas pelo informante não gago MA.



Fonte: própria.

O gráfico 5 traz apenas a indicação geral das disfluências. A especificação de cada uma será vista na medida em que as analisamos. Como vemos, MA realiza oito tipos de disfluências no decorrer da sua fala. A que ocorre com mais frequência são as PSs, depois as CORs, INSs e RP. As menos expressivas são, respectivamente, os ALs e FCs, seguidos do BLO e PI.

Quanto à PS, aparecem identificadas apenas as PSs simples, ou seja, que ocorreram em início, meio ou fim de enunciado, mas sem virem acompanhadas de um outro tipo de disfluência comum. MA realizou ao longo da gravação 64 pausas, o que representa 62% do total de disfluências. Dentre elas, o maior número de pausas ocorreu para final de enunciado, com 31% das ocorrências. Depois, tivemos as:

- PSs silenciosas em início de enunciado – 28%;

- PSs silenciosas em meio de enunciado – 21%;
- PSs respiratórias em início de enunciado – 7,8%;
- PSs respiratórias em fim de enunciado – 6,3%;
- PSs respiratórias em meio de enunciado – 4,7%.

Como podemos ver, as PS silenciosas são as mais frequentes para MA, representando 81,3%, aproximadamente de todas as PSs. Em contrapartida, as respiratórias, representaram 18,7%.

As PSs mais longas foram as silenciosas em meio de enunciado, com uma duração total de 10,247s. Já as de duração menor foram as respiratórias em fim de enunciado, 1,244s. Depois, aparecem as PS silenciosas em início de enunciado (9,069s); em seguida, as silenciosas em final de enunciado (8,092s) e as PS respiratórias em meio de enunciado (1,973s) e em início de enunciado (1,929s).

Depois das PSs, é mais comum que ocorra na fala de MA CORs representando 23,3% do total das disfluências. As CORs dividiram-se entre simples e complexas. As CORs complexas foram as mais expressivas (70,8%). Elas ocorreram com:

- COR com BLO;
- COR com PI;
- COR com PS;
- COR com RP;
- COR com INS;
- COR com AL.

As CORs variaram entre os tipos 1 e 2. Para MA a diferença entre os dois tipos foi pequena, sendo o total de CORs do tipo 1) de 54% aproximadamente. Veja abaixo dois casos que exemplificam as duas possibilidades de correções:

- EN28 /Só di sábadu > só di sexta, sábadu/;
- EN24 / primera cumida queu fi queu fu queu cu fu fui nua festa > foi bacalhoada/.

Em EN28 é perguntado para MA quais os dias que ele costuma cozinhar para os amigos. Primeiramente ele responde que é apenas aos sábados, mas ele se lembra de que às sextas-feiras também é comum, acrescentando essa informação no GT que diz respeito à COR. No segundo exemplo foi perguntado a ele qual a primeira comida que ele havia cozinhado. Nesse caso ocorre um erro de processamento da mensagem, de modo que o falante não entende a pergunta de início, cometendo vários erros até chegar à resposta correta. Como

vemos, há a introdução de palavras incompletas /fi/, /fu/, /cu/, que nos dão a ideia de que ele, num primeiro momento, vai responder adequadamente, por meio de /fi/. Porém, como a mensagem chega de modo errado para ele, MA se atrapalha, variando entre /fu/, que pode estar relacionado à outra informação /fui nua festa/ e /cu/, que remeteria a resposta do seu interlocutor /cuzinhei/. Sendo assim, como ele não consegue colocar o pensamento em ordem, mantendo uma linha de raciocínio, opta por apagar a informação anterior e vai direto à resposta /foi bacalhoda/.

Para a análise acústica das CORs, nos atemos apenas ao que ocorre no erro e na construção do novo enunciado. A análise dos outros tipos de disfluências será feita com os seus respectivos tipos. Dentre as 24 CORs, 50% delas ocorreram em início de enunciado e as outras em meio de enunciado.

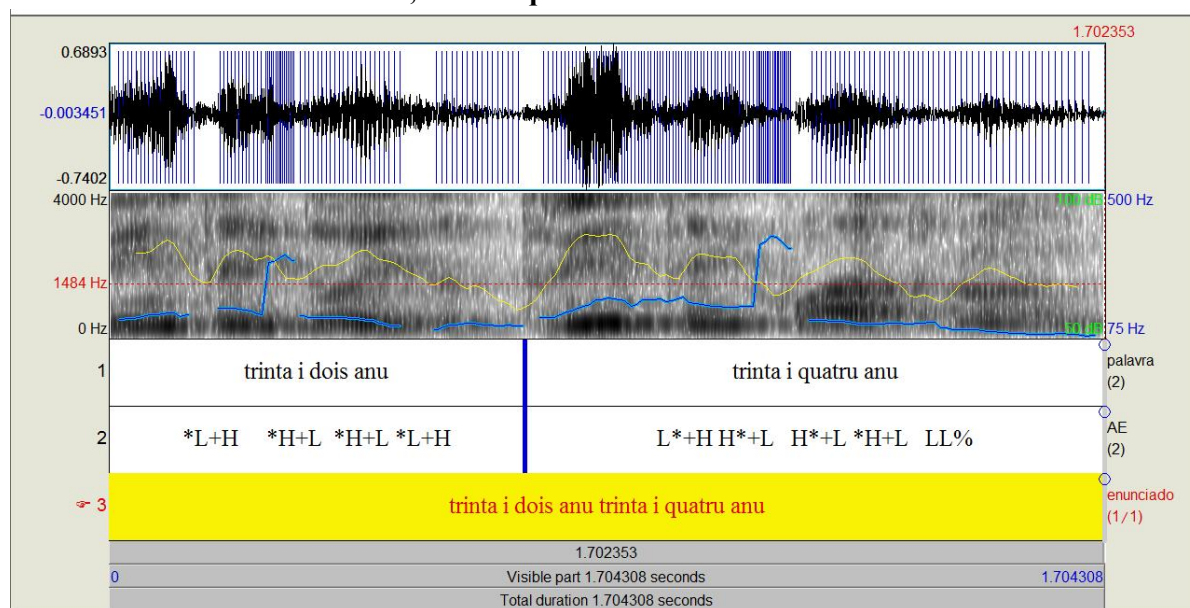
Quanto à duração, a análise mostrou que para 79% dos enunciados com COR, aproximadamente, houve um aumento da duração no momento da correção. Isso mostra que na volta à fluência o falante aumenta a duração em mais de 50% das produções. Já para a análise da média da F0, tivemos os seguintes resultados: 45,8% dos dados apresentaram uma F0 maior no momento da correção e 54,2% dos dados, uma tendência para a diminuição da média da F0 na COR. Mesmo a diferença não sendo muito grande, ela mostra que no reparo MA diminui a F0. Quanto aos pulsos, 70% deles foram maiores no momento da correção. Isso ocorreu, pois na maior parte das vezes os problemas de gravação estão no erro, como, por exemplo, os momentos que o informante respira ou produz algum segmento mais agudo, que leva a um aumento da F0 e, conseqüentemente, aglomeração dos pulsos.

Em relação à intensidade, foram observados os momentos de transição e intensidade máxima. Quanto aos momentos de transição foram analisadas a transição da fluência para a disfluência (no erro) e de disfluência para fluência (ou correção). Assim, foram analisados 48 momentos de transição de intensidade. Os resultados mostraram que dos 48 momentos, 12 ocorreram em início de enunciado e 10 foram antecidos por uma PS antes da correção. Nesses casos, o informante erra no primeiro GT faz uma PS e, em seguida, corrige o GT anterior. Por outro lado, os momentos que não iniciaram o GT ou que não foram antecidos por PS mostraram a maior tendência para que o informante reduza a intensidade tanto na transição da fluência para o erro (18,7%), quanto do erro para a correção (16,6%). Por fim, a análise da intensidade máxima revelou para 62,5% dos dados o aumento da intensidade na correção.

Finalmente, quanto à análise entoacional, não há uma padronização entre o erro e a correção, uma vez que temos uma mudança nos GTs. O máximo que pode ocorrer é que o pico entoacional seja o mesmo, fato que foi verificado para 50% dos dados.

A seguir apresentamos um exemplo de enunciado que exemplifica os resultados obtidos:

Figura 31 – Exemplificação dos parâmetros acústicos e do padrão entoacional para MA no enunciado “trinta i dois anu, trinta i patru anu”.



Fonte: própria.

O enunciado (EN14), apresentado na figura 31, identifica uma COR simples, uma vez que o informante não acrescenta mais nenhum outro tipo de disfluência. Ele apenas corrige a idade para a pergunta feita pelo interlocutor sobre o período de tempo que ele reside na cidade que mora atualmente (Andradas). Como vemos, a duração é maior na COR, sendo de 0,993s, enquanto que no erro é de 0,711s. As duas produções são parecidas. Veja que ocorre um erro, que se deve a uma produção muito aguda do segmento /i/, tanto na COR quanto no erro. Esse foi um dos casos em que a F0 média não seguiu a tendência para o abaixamento na correção. No entanto, a diferença entre elas foi pequena: 1,757 Hz. Os pulsos e a intensidade máxima, por sua vez, confirmam os resultados obtidos. A intensidade na transição exemplifica um caso em que a COR inicia o GT, mas não mostra a tendência para o seu abaixamento na correção, mostrando um aumento de 1,33 dB em relação a anterior. Por fim, a análise do padrão fonológico mostra um caso em que o pico entoacional é o mesmo nos dois períodos.

A INS foi o segundo tipo de disfluência mais frequente, com 12,5% das ocorrências. Elas variaram entre INS de palavra e INS de vogal e entre simples e complexas. Em relação

às INS_{ps} , apenas uma ocorreu dentro de COR. Já quanto às vogais, três foram alongadas. Foram observadas três tipos de INS_{ps} /né/, aí/ e /intão/. A primeira foi a mais frequente (38,4%) e ocorreu em final de GT. A INS_p que ocorreu junto com a COR, mas especificamente no erro, acontece em meio de enunciado. No entanto, como na sequência ocorre uma PS para início da COR, essa INS_p tem a mesma função sintática de interrogação afirmativa retórica das que foram identificadas em fim de enunciado. Veja abaixo:

- /queu comecei desdi d udus dozi, né/

Em seguida, aparece a INS_p do tipo /aí/ (23%) e, por último, /intão/ (7,7%), em início de GT. Apenas uma ocorrência do tipo /aí/ ocorreu em meio de GT. No entanto, como ela acontece após uma PS, apresenta a mesma função sintática das que ocorrem em início de GT. Sendo assim, /aí/ e /intão/ apresentam uma função sintática de sequenciadora, dando continuidade ao pensamento anterior. Observe o exemplo:

- “aí interrompi uns dia, um anu di profissão pa fazê cursinho”.

Para as $INSs$ de vogais, MA realizou apenas inserções do tipo /é/. Dentre elas, duas ocorreram em início de GT e três em meio de GT. A funcionalidade expressiva dessas inserções podem ser de dois tipos ou 1) a de ganhar tempo para dar continuidade ao discurso ou 2) afirmativa. Nesse caso, seria uma resposta positiva a uma interrogação. As inserções alongadas caracterizam-se por serem do primeiro tipo. Observe:

- /é depoi fiz cur várius cursos/
- /depois é- montei um labor vim resolvi vim pra Andradas/

O primeiro exemplo identifica uma INS_v do tipo 2. MA, ao ser interrogado sobre sua atuação, ou seja, se faz cursos, utiliza essa INS_v de forma afirmativa. Já no segundo exemplo temos uma INS_v alongada, com a funcionalidade de ter tempo para construir o resto do pensamento. Nesse caso, mesmo utilizando essa estratégia, o informante não consegue chegar a uma conclusão. MA muda de ideia em um momento em que a mensagem já havia sido produzida e, isso faz com que ele erre e tenha que fazer uma COR.

Em relação à análise acústica, observamos que a duração das INS_{ps} do tipo /né/ não variaram significativamente, sendo a diferença entre a maior e a menor de 0,067s (0,358s e 0,291s). No entanto, apenas uma produção não manteve esse padrão, tendo uma duração de 0,109s.

A F_0 , por sua vez, caracteriza-se por apresentar uma tessitura mais alta, sendo que a F_0 mínima apresentou uma média de aproximadamente 150 Hz e a máxima de 165 Hz. Os pulsos, por sua vez, não podem ser considerados como parâmetros para análise já que não

mantiveram o mesmo padrão. Finalmente, a transição da intensidade revelou para 100% dos casos um aumento do período da fluência para a INS_p . Por sua vez, a intensidade máxima manteve uma média, sendo a maior de 78 dB e a menor de 70 dB.

A INS_p do tipo /aí/ também mostrou uma tendência para a padronização da duração (0,351s e 0,321s). A tessitura, por sua vez, manteve um padrão mais baixo, sendo a mínima de 103 Hz e a máxima de 126 Hz. A intensidade máxima também não variou significativamente, com uma diferença de 3 dB entre elas. Em relação à /intão/ tivemos apenas um caso. Essa produção apresenta uma duração maior, de 0,451s, bem como uma tessitura mais alta do que as outras produções.

Quanto às INS s de vogal, as alongadas foram as que apresentaram as maiores durações (0,421s; 0,426s e 0,561s). Dentre as produções não alongadas a variação foi maior, sendo uma de 0,069s e a outra de 0,224s. Isso pode ter ocorrido porque a primeira duração identifica uma produção em início de enunciado e a segunda, em meio de enunciado. A F_0 caracteriza uma tessitura mais baixa, assim como em /aí/. Os pulsos em INS_v alongada são maiores do que em INS_v não alongada (7 e 72, por exemplo). Mas isso é esperado, uma vez que o esforço vocal é maior. Por fim, quando não foi antecedida por PS ou ocorreu em início de GT, diferentemente das INS_{ps} , as INS_v s apresentaram uma diminuição na transição da intensidade. Finalmente, no que diz respeito à intensidade máxima, a diferença foi maior entre elas, de modo que a variação entre a maior e a menor foi de 15 dB.

Finalmente, para a análise fonológica, o padrão fonológico variou apenas para a INS_p /né/ em *H+L *L e *H. De todos os enunciados com esse tipo de inserção, somente um deles apresentou o padrão *L. O padrão mais recorrente foi o do primeiro tipo, que apareceu em três casos. Em seguida, temos o padrão *H em duas situações. O padrão que mais destoa dos outros é *L. Porém, nesse caso tivemos uma INS muito grave, por isso, esse resultado. Quanto aos outros casos, o padrão não é o mesmo, mas vemos, por outro lado, que eles apresentam o mesmo pico entoacional *H. A diferença entre eles está no fato de que MA enfatizou mais as INS s quando o padrão foi *H+L. Para /aí/, o padrão obtido foi L+*H e, para /intão/ *H+L. Já para as vogais, em início de enunciado, o padrão foi *H+L e, em meio de enunciado, *L. Como vemos, em início de GT o informante começa com uma produção mais alta do que quando faz uma INS_v em final de GT.

Depois, o terceiro tipo de disfluência mais recorrente para MA foram as repetições. Isoladas, foram produzidas 11 RPs, o que representa, aproximadamente, 10,6% do total. Em COR, tivemos três casos de RPs. MA realizou sete tipos de repetições:

- 1 Repetição de vogal simples (RP_v simples) - /u u/ ;
- 1 Repetição de vogal complexa (RP_v complexa) - /i- i/;
- 1 Repetição de palavra simples (RP_p simples) - /du du/;
- 1 Repetição de palavra incompleta (RP_{pi}) - /du dus/;
- 5 Repetições de palavra complexa (RP_p complexa), sendo 3 em COR;
- /da/da/
- /um co > um a um amigu/
- 2 Repetições de estrutura (RPe) - /i não i não/;
- 1 Repetição de estrutura com palavra incompleta (RP_{epi}) - /u ganhu u ga u ganhu/;

Entre esses tipos de RPs, tivemos apenas dois casos de RPP tripla e um de RPe tripla. Quanto à análise acústica das vogais, como só tivemos dois casos (EN8 complexa e EN17 simples), sendo um diferente do outro, é mais difícil fazer comparações. Mesmo assim, para a RP_v complexa há uma diminuição da duração na segunda RP. Isso porque, como se trata de uma RP com alongamento na primeira vogal /i-/, esse é um fato esperado. Tanto assim que tivemos uma duração de 0,632s para a primeira vogal e de 0,072s para a segunda. A análise da F0 revelou uma frequência maior para a repetição complexa. No segundo caso, a produção foi surda. Vale ressaltar que, para os dois casos, tivemos uma produção mais grave, de tessitura mais baixa. Os pulsos, por sua vez, só puderam ser analisados para EN8, mostrando um esforço vocal menor para a segunda RP. Finalmente, a análise da intensidade máxima e transição da intensidade da primeira repetição para a segunda foi maior em EN8. Como nesse enunciado a RP ocorreu no início da frase, a transição da primeira RP não foi analisada. Em EN17, por sua vez, a intensidade foi menor tanto na transição da fluência para a repetição, quanto da primeira repetição para a segunda. A intensidade máxima também foi menor. Essa diferença não foi auditivamente significativa: 0,352 dB.

Para as repetições de palavras, tanto em COR quanto simples, consideramos primeiramente, a análise das RPs duplas. As ocorrências simples serão analisadas individualmente. Dessa maneira, foram observadas nove produções, consequentemente, 18 segmentos. Dentre os nove enunciados, três ocorreram em início de frase e, dos 18 segmentos, apenas dois deles foram antecidos por pausa. A análise da duração mostrou para esses eventos, assim como na COR, uma tendência para o aumento da duração na segunda repetição (55% das produções).

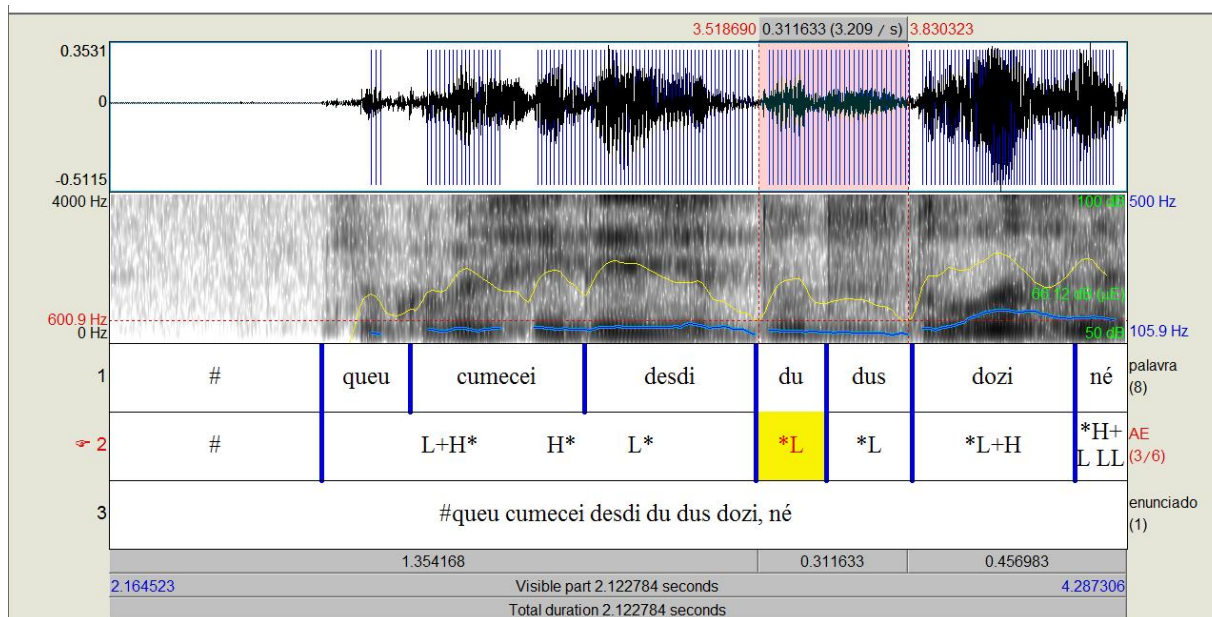
Em relação à média da F0, tivemos duas repetições sem realização de F0 na segunda RP. Nesse caso, foi considerado como um caso de diminuição da frequência fundamental. Além disso, deixamos de fora uma produção, pois no segundo evento ocorreu um erro que impossibilitou a análise, devido à realização de um evento muito agudo. Assim, dentre os oito casos, 62,5% deles tiveram uma diminuição na frequência na segunda RP. Quanto aos pulsos, também seguindo as mesmas condições propostas, a análise mostrou que em relação à primeira RP, eles tendem a diminuir (57%, aproximadamente). Por fim, a análise da intensidade mostrou que no momento da transição da fluência para a primeira RP, excluindo os momentos em que a RP ocorreu em início de enunciado, a tendência é para a diminuição da intensidade (60%), bem como na passagem da primeira RP para a segunda (85,7%). A intensidade máxima, por sua vez, variou em apenas um enunciado. Sendo assim, cinco reduziram a intensidade (55,5%) na segunda RP e quatro aumentaram de intensidade (44,5%). Isso mostra a disposição das RPs em manterem a mesma estrutura.

Finalmente, no que diz respeito às produções com mais de duas RPs, ocorre o que chamamos de forma de onda na duração, isto é, na transição da primeira para a segunda produção a duração diminui e, da segunda para a terceira, aumenta. Para a F0, a maior tendência é para a diminuição. Dos dois casos observados, apenas em um deles aconteceu o aumento da F0 na terceira RP, bem como foi observado para os pulsos. Isso revela um menor esforço vocal. Para a intensidade, no momento da transição a tendência é para o aumento nos três segmentos. Em relação à intensidade máxima, da primeira RP para a segunda, observa-se uma diminuição na intensidade. Já da segunda RP para a terceira, um aumento.

Portanto, o conjunto das RPs mostrou que para MA a duração tende a aumentar, já a F0 e a intensidade máxima, os pulsos e momentos de transição, têm a maior disposição para a diminuição.

É importante salientar que as diferenças observadas não foram grandes, mas ajudam a distinguir as RPs e a propor alguns parâmetros. Se deixarmos os dados numéricos de lado, veremos que as repetições não se distanciam significativamente. O fato da F0 e da intensidade serem mais constantes, mostra que esse informante varia pouco a sua fala, o que pode ser considerado como uma estratégia discursiva. Veja a figura que segue:

Figura 32 – Representação da duração, intensidade, F0 e pulsos para a RP_{pi} no enunciado “# queu comecei desdi d udus dozi, né”, realizado pelo informante não gago MA.



Fonte: própria.

A figura 32 identifica, na seleção em rosa, uma RP_p, do tipo /du dus/. Nesse caso, o informante acrescenta mais um segmento à palavra /s/, deixando a primeira RP incompleta. Como vemos, apesar de haver uma diferença entre as RPs, elas são relativamente pequenas. A primeira apresentou uma duração de 0,146s e a segunda, de 0,178s. Essa pequena diferença ocorre também para a F0 (109 Hz e 103 Hz) e para a intensidade máxima (71 dB e 64 dB). Se observarmos atentamente a figura veremos que as duas produções são muito parecidas. Essa constatação foi percebida para a maior parte das repetições, não sendo verificada apenas para os casos de alongamento.

A análise fonológica também justifica a afirmação, uma vez que tivemos apenas um enunciado EN38 (com repetição tripla do tipo RP_e que não seguiu o mesmo padrão para as estruturas repetidas. Mesmo assim, isso ocorreu apenas na primeira RP. As outras duas mantiveram o mesmo padrão.

Em seguida, com uma produção bem menor do que as outras disfluências, temos o AL e o FC, ambos em 1,92% das produções.

O AL foi verificado para vogais e consoantes. Como vimos, o primeiro tipo ocorreu apenas com as INs de vogais /é/. Em relação ao AL de palavra, MA realizou duas produções, uma simples /morei-/ e outra complexa /aos-/. A produção complexa aconteceu com BLO. Os

dois alongamentos foram verificados em final de palavra, sendo o primeiro em início de GT e o segundo, em meio de GT:

- “morei: dexa eu lembra aqui”,
- “trabalhei dus dozi -aos: dizenovi anus”.

Por meio dos dois exemplos, bem como já tinha sido visto para o AL_v em INS, essa disfluência ocorre quando o informante precisa de ajuda para lembrar de algo ou de alguma situação, o que é muito evidente no primeiro exemplo, no qual ele faz o alongamento e continua o GT, mencionando algo que ele precisa lembrar, nesse caso, do lugar onde morou.

A análise acústica dos AL_{ps} mostra uma duração maior para os dois casos: /morei:/ 0,941s e /-aos:/ 0,824s. Nas duas produções a tessitura também é mais baixa, com uma média de 117 Hz e 117 Hz. Os pulsos não foram analisados comparativamente devido à sua variação. Por fim, em relação à intensidade, em período de transição, um dos AL_s ocorreu em início de GT /morei:/ e o outro, /-aos:/ foi antecedido por uma PS de BLO. Devido a sua posição, a intensidade inicial nas duas produções foi mais alta e a final mais baixa como, por exemplo, 69 dB e 44 dB para o AL em /-aos/.

A análise fonológica foi diferente nas duas produções, já que temos alongamentos em palavras diferentes. Para o AL do tipo /-aos:/, ela é do tipo *H, com padrão mais alto e sem variação significativa. Já para em /morei:/, temos um tom *bitonal* do tipo L+*H, que começa mais baixo e depois aumenta.

A disfluência do tipo FC identifica situações em que o falante vai iniciar um assunto, mas durante a produção da mensagem compreende que esse não é o melhor caminho e não continua o raciocínio, deixando-a fora de contexto. Para esse informante foram analisados três casos de FC, sendo que um deles ocorreu com COR:

1. “como queu si é cheguei na profissão”;
2. “fui trabalhar em/aí disiti da/da faculdade;”
3. “pes eu achu qui as meus amigu gosta di mim”.

Nos três exemplos os FCs são /si/, /em/ e /pes/. Observe que em todas as produções a informação não se encaixa no contexto, de modo que o informante não completa o GT de acordo com a mensagem que essas palavras produziriam. Caso o informante optasse por usá-las, teríamos um novo GT. É importante salientar que /pes/, considerado como um FC, nos dá a ideia de uma disfluência do tipo PI. No entanto, como ela se enquadra nas informações acima, foi considerada como um FC.

Nos três casos, a análise acústica mostrou uma F0 média variada para as três produções. Para /si/ tivemos uma produção surda, em /em/, uma tessitura mais alta 135 Hz. Já para /pes/, tivemos uma tessitura mais baixa 106 Hz. A duração também variou 0,194s, 0,382s e 0,232s. A intensidade, em momento de transição, seguiu o padrão de abaixamento para os dois primeiros exemplos (1 e 2). Para o último (3), ela não foi analisada, pois ocorreu em início de GT. A intensidade máxima foi maior para /em/ e /pes/ e menor para /si/. Finalmente, a análise fonológica revelou um tom simples alto para /em/ (*H) e baixo /pes/ (*L). Já o FC do tipo /si/ não apresentou F0.

A disfluência do tipo BLO aparece entre os casos menos expressivos de disfluência para MA, bem como a do tipo PI. Levando em consideração o conjunto de disfluências, foram identificados três casos de BLO, sendo dois em vogal e um em consoante. Entre eles, apenas um caso foi de BLO simples /-marcelu/, por isso, consideramos essa disfluência como uma das menos recorrentes. Os outros casos foram de bloqueios complexos, pois ocorreram junto com COR, sendo eles: /-aos/, /-um/. Como vemos, temos uma situação de BLO em consoante nasal bilabial [m], uma em vogal alta [a] e uma em vogal baixa [u]. Os três bloqueios seguiram um dos padrões observados para esse tipo de disfluência, sendo dois com pausa articulatória /-marcelu/ e /-um/, e, um /-aos/, com PS silenciosa, mas com articulação da vogal anterior.

A análise acústica para essa disfluência revela que a PS que caracteriza o BLO é maior para a produção /-aos/, com uma duração de 1,048s. Já o segmento teve uma duração de 0,350s. Por sua vez, a duração da palavra /-um/ foi de 0,176s e a PS de 0,402s. O segmento /-aos/ durou 0,743s e a PS 0,243s. Nos dois primeiros casos, a PS de bloqueio é maior do que o segmento e, no terceiro, menor. A F0 média revelou uma tessitura baixa para as três produções. A intensidade máxima, por sua vez, foi mais alta para o BLO em consoante 79 dB. Quanto à intensidade inicial e final, para todos os casos, tivemos uma intensidade que começa mais alta, já que a produção do segmento inicia-se após uma PS e acaba mais baixa. No que diz respeito à análise fonológica, todas elas iniciaram-se com um pico mais alto, mas os segmentos /-marcelu/ e /-aos/ foram bitonais, com o padrão *H+L e o segmento /-um/ manteve um tom alto simples *H.

Junto com o BLO aparece a disfluência do tipo PI, com aproximadamente 1% das ocorrências. Essa porcentagem é referente apenas à produção simples dessa disfluência, uma vez que em COR elas são mais recorrentes, sendo verificadas em cinco casos. Observe as ocorrências:

1. “ondi tivi meus#impre primeru impregu di protéticu#”;
2. “#depois é montei um labor vim resolvi vim pra Andradas#”;
3. “é depois fiz cur várius cursus”;
4. “# na profis du du du que se fala”;
5. “i tinha um funcio eu cuzinhava a noiti”;
6. “#pur ca du aumentu du material qui são tu quasi todus importadu”.

As PIs representadas nos exemplos são, respectivamente, /impre/, /labor/, /cur/, /profis/, /funcio/, /ca/, todas ocorreram em meio de enunciado. Apenas no primeiro exemplo ela foi identificada após uma PS. Os cinco primeiros casos ocorrem junto com COR e o último em enunciado isolado. Como vemos, elas são mais recorrentes em CORs, o que mostra que uma das estratégias desse informante, ao perceber o erro, é reduzir algumas palavras, ou seja, ao perceber que vai errar ele para a produção da mensagem no meio do caminho para que a volta à fluência seja mais rápida. A PI simples representada pelo exemplo EN6, por sua vez, mostra situações de queda de palavras. Esse fenômeno é comum na fala das pessoas e diz respeito a questões rítmicas.

No entanto, a PI simples apresenta uma velocidade de fala maior do que quando ocorre com COR. Por exemplo, nos casos 1 e 3 temos as palavras que foram reduzidas e a sua produção completa. Em 1, a duração da PI foi de 0,406s e a da palavra completa de 0,402s. Já em 3, a PI durou 0,306s e a completa 0,596s. A PI no exemplo 1 foi ligeiramente menor do que a sua palavra completa, isso pode ter ocorrido porque essa PI apareceu logo após uma PS, de modo que MA demora um pouco mais para formular a mensagem novamente e retomar a fluência. Por outro lado, no caso 3 vemos que a tendência é para que o segmento completo dure mais do que o incompleto.

A análise da F0 mostrou tessituras mais baixas para as PIs. Apenas o caso 3 apresentou uma tessitura mais alta (186 Hz). A tessitura mais baixa foi identificada para a PI simples (caso 5), 96 Hz. As outras variaram entre 119 Hz e 112 Hz.

Quanto à transição de intensidade da fluência para disfluência, apenas para o caso 1 a PI não foi analisada, pois ocorreu após uma PS. Dentre os outros casos, elas mantiveram a tendência de abaixamento, sendo verificadas para 60% das ocorrências. Quanto ao aumento ou diminuição da intensidade inicial para a final, em 50% dos casos a intensidade final aumentou em relação à inicial e, em 50% deles, diminuiu. A intensidade máxima, por sua vez, variou, sendo maior para a PI do tipo /cur/, de 81 dB e menor para /funcio/, 65 dB.

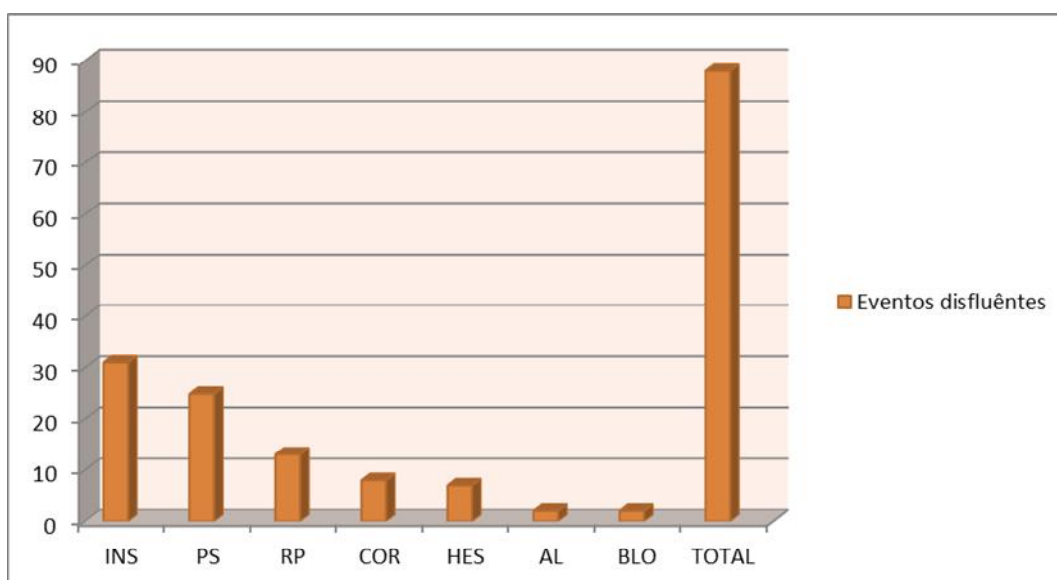
A análise fonológica para a PI complexa apresentou uma variação entre tons simples e compostos. No entanto, todos os casos apresentaram o pico entoacional no tom alto *H e *H+L. Já a PI simples caracterizou-se por apresentar um tom baixo simples do tipo *L.

O informante LE

O informante não gago LE realizou 88 disfluências em um total de 37 enunciados. Entre elas ocorreram disfluências simples (62,2%) e complexas (31,8%). O número máximo de disfluências por GT foi de sete, sendo duas do tipo PS. A sua taxa de elocução foi de 5,754 sil/ms, a de articulação foi de 6,351 sil/ms e o tempo de articulação de 84ms.

Veja a distribuição dos eventos disfluentes na fala de LE.

Gráfico 6 – Distribuição dos eventos disfluentes realizados pelo informante não gago LE.



Fonte: própria.

O gráfico 6 mostra que LE realizou oito tipos de disfluências, organizadas em ordem decrescente de produção: INSS; PSs, RPs; CORs; HESs, ALs e BLOs. Diferentemente de MA, o informante realizou mais INSSs do que PSs. Dentre as outras disfluências, a proporção foi parecida. No entanto, LE realizou uma disfluência do tipo HES, que não foi observada em MA.

As disfluências do tipo INSSs aparecem em aproximadamente 35% dos eventos. Elas dividiram-se entre INSS_vs (6,5%); INSS_ps (as mais frequentes, com 90% das realizações) e INSSs

complexas (3,5%). As INSs complexas ocorreram junto com AL e PS. As INSs realizadas por LE foram:

- INS_p em início de GT - /tão/, /intão/ e /ái/;
- INS_p em fim de GT - /né/ e /sabi/;
- INS_v em meio de GT - /é/;
- INS complexa em meio de GT - /é/.

Assim como tínhamos observado para os outros informantes, as INSs do tipo /intão/ e /ái/ ajudam a sequenciar o enunciado, bem como /é/ em meio de GT. Vale salientar que em nenhuma das produções /é/ teve um sentido afirmativo. Além disso, como podemos ver, diferentemente dos outros informantes, LE realizou uma variação de /intão/ > /tão/. No contexto dos enunciados esse segmento tem a mesma função semântica de /intão/. Veja:

- EN8 /tão fais cincú anu quela é camada/.

Nesse caso, LE usa a INS_p /tão/ a fim de dar sequência a sua fala. Nesse enunciado ele explica a situação da sua mãe que vive acamada há cinco anos.

A INS_p /né/ manteve a função de interrogativa retórica, bem como a INS_p do tipo /sabi/ em fim de GT.

No que diz respeito à análise acústica, a média da duração mostrou que a INS complexa é a que apresentam a maior duração (0,649s). Isso ocorreu pois ela foi identificada com a disfluência do tipo AL. Já a INS que teve a menor duração foi a de palavra em início de GT (0,246s). As outras durações foram de 0,429s para a INS de palavra em fim de GT e 0,395s para a INS_v .

A média da F0 mostrou que a entoação é maior para a INS_p em fim de enunciado (167 Hz), seguida da INS_p de início de GT (159 Hz); INS complexa (153 Hz) e INS_v (144 Hz). Quanto aos pulsos, o esforço vocal também é maior para a INS_p em fim de GT (70). No entanto, aparecem na sequência as INS_v s (50); a INS complexa (49) e, por fim, as INS_p s em início de GT.

Quanto à intensidade no momento de transição e, deixando de lado as INSs que iniciaram o GT ou foram antecedidas por PSs (dois casos), em todos os tipos de INSs houve uma diminuição da intensidade na transição da fluência para a disfluência. As INS_p s em fim de enunciado apresentou uma diminuição em 60% dos casos, já a INS_c e as INS_v s totalizaram 100% dos eventos. Em relação à intensidade final, os resultados variaram entre os tipos de INSs. No entanto, a tendência foi para a diminuição da intensidade quando comparada com a intensidade inicial. Sendo assim, a INS complexa e a INS_p em final de GT apresentaram uma

diminuição na intensidade em 100% dos casos para o primeiro tipo e 95% para o segundo tipo. Para a INS de palavra em início de GT foi observado um aumento na intensidade final em 85% dos casos. Por fim, para a INS_v verificou-se em 50% dos casos uma aumento na intensidade e em 50% uma diminuição. Assim, como vimos, as INSs do tipo complexa e de vogal representaram o menor número de ocorrências, por isso, fica mais difícil fazer comparações. A intensidade máxima variou pouco entre as INSs, mesmo assim, ao fazermos comparações por meio da média da intensidade máxima, vemos que a INS_p em início de GT apresentou a maior intensidade 82 dB. Na sequência, aparece a INS complexa, com uma intensidade de 82 dB; a INS_p em fim de GT com 80 dB e, por último, a INS_v, com uma intensidade de 76 dB, aproximadamente.

A análise dos padrões fonológicos confirmou a tendência das INSs do mesmo tipo e mesma posição no enunciado apresentarem o mesmo padrão entoacional. Para /tão/ e /intão/, os padrões variaram entre *H+L e *H. A INS_p /intão/ (um caso), apresentou o tom composto *H+L, já /tão/ variou entre o tom composto *H+L e simples *H, sendo que a maior tendência foi para o primeiro tipo. Como vemos, os padrões não são exatamente iguais, mas o pico entoacional é o mesmo. Para a INS_p do tipo /né/, os padrões também variaram. No entanto, dentre as 17 ocorrências de /né/, apenas 11% aproximadamente (ou dois casos) foi para o padrão *H+L, e o restante, 88% para o padrão *H.

A INS_p do tipo /sabi/ em fim de GT, apresentou um tom do tipo *L+H, os outros mantiveram o padrão *H+L ou *H. Isso pode ter ocorrido por, nesse enunciado, o informante apresentar uma qualidade de voz mais grave. Por sua vez, /aí/, em início de GT manteve o padrão alto (*H+L). Já a INS de vogal variou entre (*L+H) e (*H+L). Esse fato pode ter ocorrido porque o primeiro padrão é seguido de uma PS e o segundo não. Finalmente, o único caso de INS complexa apresentou o padrão *H+L.

Depois das INSs, as PSs correspondem a 28,4% dos eventos disfluentes. O tipo mais frequente são as PSs de fim de enunciado (52%), seguidas respectivamente das PSs de início de GT (36%) e meio de GT (12%). As PSs que antecederam eventos disfluentes contabilizaram 32%. Em relação à duração, as PSs em fim de enunciado também foram as que apresentaram a maior duração (4,610s). Depois aparecem as PSs em início de GT (3,052s) e as PSs em meio de GT (0,983s). As PSs silenciosas foram as mais frequentes, com uma porcentagem de 76%, enquanto as respiratórias representam 24% do total.

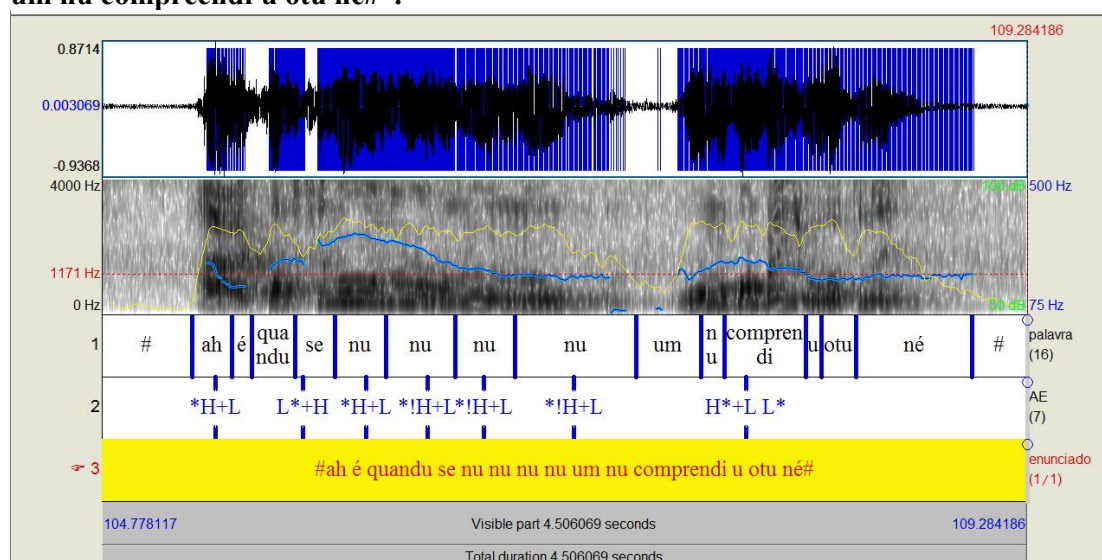
As RPs, por sua vez, representam 14,8% aproximadamente do total de disfluências realizadas por LE. Dentre elas, apenas um tipo foi de RP complexa, que ocorreu junto com

AL. Esse tipo de disfluência ocorreu em palavra (RP_p , 61,5%), sendo um caso de RP quádrupla. O restante das repetições foi de RP s duplas de palavra, que variaram entre palavras monossílabas (83%) e dissílabas (17%). Também foram verificadas RP s de sílaba em palavra (RP_{ps} , 23%). Para esse tipo de RP um caso foi de repetição dupla de sílaba. Por fim, LE também realizou RP de estrutura (RP_e , 15,5% aproximadamente). Não foram identificadas RP s em vogal. Veja, na sequência, os exemplos para cada um dos tipos de RP s:

1. RP_{ps}
 - a. EN7 /minha mãe tem deu *au* austrózia/;
 - b. EN35 /aí conseguiu *lu lu* luga pa dona da farmácia/.
2. RP_p
 - a. EN17 “a turma *nu nu* joga futebol não”;
 - b. EN29 “#agora graças a Deus já vorto vorto a chuva né#”;
 - c. EN26 “/ah é quando se *nu nu nu nu* um nu compreendi u otu né#”.
3. RP_e
 - a. “ah eu gostu asisti *um jogu um jogu* di futebol, né”.

Por meio da análise acústica da duração não foi possível estabelecer um padrão geral. Isso porque, para as RP_{ps} e RP_p s em 50% dos casos a duração diminuiu e, em 50% aumentou. Para a RP_p complexa a duração diminuiu. Fato que é esperado, uma vez que o AL ocorreu na primeira RP . Já para a RP_e , nos dois casos, a duração diminuiu. Por sua vez, a RP quádrupla seguiu um padrão de forma de onda crescente, ou seja, a duração começou menor, em seguida aumentou e assim sucessivamente. Observe o exemplo:

Figura 33– Exemplificação de uma forma de onda crescente para uma repetição do tipo 4 RP_p realizada pelo informante não gago LE no enunciado “ah é quando se *nu nu nu nu* um nu compreendi u otu né#”.

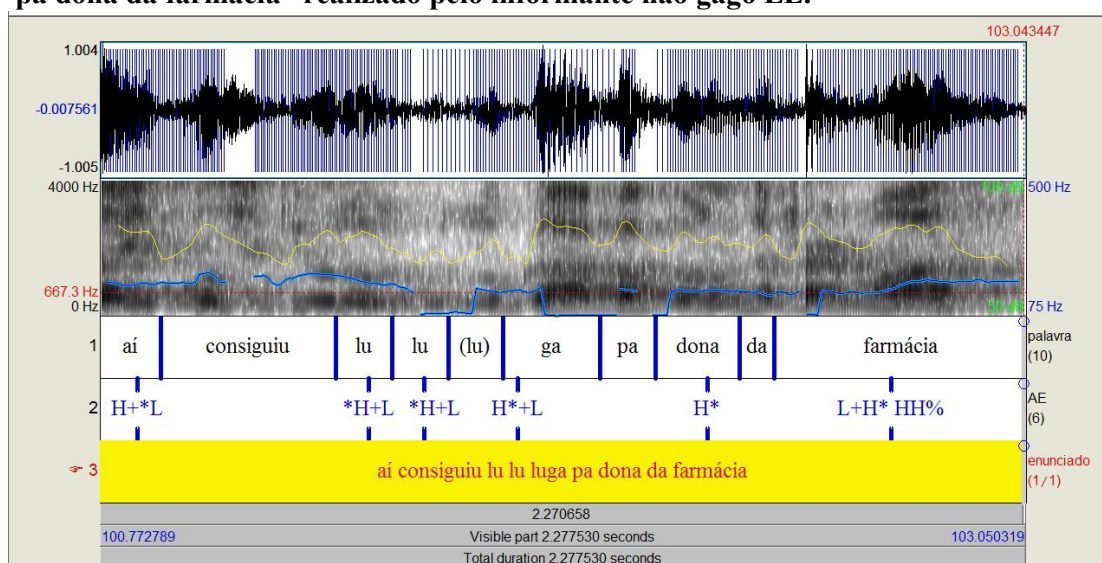


Fonte: própria.

Observe, nesse exemplo (cf. fig. 33), que o informante realiza quatro RPs do segmento /nu/ que duram, respectivamente, 0,249s; 0,337s; 0,288s e 0,591s. Além desse tipo de disfluência, LE realiza no mesmo enunciado uma HES e uma INS_p final. As curvas da F0 (linha em azul) e intensidade (linha em amarelo) mostram padrões mais altos. Esse fato deve-se a dois motivos: o primeiro deles é que, diferentemente dos outros informantes não gago, LE apresenta uma qualidade de voz mais aguda. A outra razão deve-se ao fato de ter ocorrido algumas interferências na gravação, o que determinou uma qualidade de gravação mais baixa.

Por fim, a RP_{ps} dupla manteve a mesma duração para as duas sílabas. Veja na figura abaixo:

Figura 34 – Exemplo de repetição de sílaba dupla no enunciado “aí conseguiu lu lu luga pa dona da farmácia” realizado pelo informante não gago LE.



Fonte: própria.

Como podemos ver na figura 34, LE mantém a duração de 0,138s para as duas sílabas /lu/ e /lu/. A sílaba realizada na palavra completa que representa a volta à fluência tem uma duração de 0,133s, ou seja, ela diminui em comparação às RPs. Na primeira RP a entoação é mais alta. Na segunda, ela começa mais alta e depois diminui, levando a uma qualidade de voz mais grave, como podemos ver por meio dos pulsos, bem como na sílaba em palavra.

Quanto à F0, dentre os casos em que não foram observados erros, a entoação teve uma tendência à diminuição. Como apresentado nas figuras 33 e 34, para a $4RP_p$ a F0 começou mais alta e veio diminuindo sucessivamente, o que determinou um padrão do tipo *downstepp*. Já para a $2RP_s$ a F0 também começa mais alta, diminui bruscamente e depois aumenta. Para a RP_{ps} de uma sílaba, a F0 diminuiu nos dois casos. Já para a RP_e os resultados variaram entre aumento (50%) e diminuição (50%) da F0, bem como para a RP_p . Por fim, a

RP_p complexa aumentou a entoação na volta à fluência. Em relação aos pulsos, para as RP_{ps} de um segmento o esforço vocal é menor na volta à fluência, bem como para a RP_p complexa. Se formos comparar o esforço vocal inicial da 2RP_{ps} e 4RP_p com o final, veremos que ele é menor no primeiro caso (cf. figura 64) e maior no segundo (cf. figura 65).

Finalmente, para a intensidade, obtivemos os seguintes resultados: no que se refere à intensidade no momento de transição, para as RP_{ps} de uma ou duas sílabas a intensidade diminuiu tanto na transição da fluência para a disfluência (66,6% dos eventos), quanto da disfluência para a fluência (100% dos eventos). Para a RP complexa, como temos apenas um caso, a intensidade aumentou na passagem da fluência para a disfluência e diminuiu na transição disfluência para a disfluência. Já para a RP_e, dos dois casos realizados, um deles foi antecedido por PS e o outro apresentou um aumento na intensidade na transição da fluência para a disfluência. Por outro lado, na passagem da disfluência para a disfluência a intensidade diminuiu. Para a 4RP_p a intensidade aumentou na transição da fluência para a disfluência e depois diminuiu sucessivamente nas outras RPs. Por fim, para a RP_p dupla a tendência é para o aumento da intensidade na passagem da fluência para a disfluência (66,6% dos casos) e para a transição da disfluência para a disfluência a tendência é para o abaixamento (66,6% dos casos).

No que diz respeito à intensidade final, para as RP_{ps}, incluindo a 2RP_{ps} a intensidade final tem a tendência ao abaixamento, bem como para a 4RP_p e para a RP_{pe}. No caso da RP_p complexa a intensidade final aumenta na transição da fluência para a disfluência e diminui na transição da disfluência para a disfluência. Por fim, para a RP_p, a tendência é para o aumento da intensidade (58%, aproximadamente, dos casos). Finalmente, a média da intensidade máxima revelou uma tendência ao aumento para a segunda RP_p (66,6%). Já para a RP_{ps}; 2RP_{ps}; RP_p complexa; RP_e e 4RP_p a tendência foi para a diminuição da intensidade máxima para segunda RP ou última RP, no caso das 2RP_{ps} e 4RP_p.

A análise entoacional confirmou que os padrões entoacionais se mantêm para as RPs. No caso do informante LE o padrão se manteve para 100% dos casos. A única diferença observada foi o *downstep* observado para EN26 (cf. figura 33). Esse fenômeno ocorre quando o tom alto subsequente é ligeiramente mais baixo do que o tom alto anterior.

Seguido da RP, temos a COR, que caracteriza, aproximadamente, 9% do total de disfluências. A COR dividiu-se entre eventos simples e complexos. Dentre eles, o primeiro tipo foi o mais frequente, com 87,5% das ocorrências, enquanto a complexa (um caso) representou 12,5% dos eventos e ocorreu com uma PS_e. Dentre as CORs, em apenas 12,5%

dos eventos, que equivale a um caso, LE recupera parte da informação do erro na COR. Já nos outros 87,5% dos casos o informante muda completamente a informação na COR.

Quanto à análise acústica, LE diminui a duração na COR complexa no momento da correção. Por outro lado, a COR simples mostra uma tendência para o aumento da duração na COR em, aproximadamente, 57% dos casos. Em relação à F0 média, a entoação é maior na na COR simples (57% aproximadamente) do que na COR complexa. Os pulsos seguem o mesmo padrão e porcentagem para a COR simples, enquanto para a COR complexa ocorre uma diminuição do esforço vocal no momento da COR.

A análise da intensidade no momento de transição mostrou para a COR complexa um aumento na passagem da fluência para o erro. A transição do erro para a correção não foi observada para esse caso por ser antecedida por uma PS. Quanto à COR simples, na passagem da fluência para o erro a tendência é para a diminuição da intensidade (57%, aproximadamente, dos casos). Por outro lado, do erro para a COR o fenômeno inverte-se, ou seja, ocorre um aumento da intensidade em aproximadamente 57% dos casos. No que diz respeito à intensidade final, a tendência é para o abaixamento (64,3%) na COR simples. Por sua vez, para a COR complexa foi identificado um abaixamento da intensidade final no erro e um aumento na correção. Finalmente, a intensidade máxima foi maior na correção tanto para a COR simples (71,4% dos casos) quanto para a COR complexa.

Em relação à análise fonológica, os padrões foram diferentes. Eles variaram entre simples e compostos na passagem do erro para a correção. Nessa transição do erro para a COR, a tendência foi para a diminuição da entoação em aproximadamente 71,4% dos casos. A COR complexa também apresentou uma diminuição na intensidade do erro para a correção.

Depois das CORs, ocorrem as HESs, representando 7,95% dos eventos disfluentes. As HESs ocorreram em 100% dos casos em início de GT, assim como observado para os outros informantes, sejam eles gagos ou não. Pragmaticamente, para LE, elas não indicam surpresa ou lembrança de um fato. O informante as insere no discurso com a intenção de responder ao seu interlocutor, mas com uma maior variação. É como se, para ele, as questões feitas pelo interlocutor despertassem um significado maior, que faz com que ele use /ah/ ao invés de, por exemplo, /olha/, com um significado semântico afirmativo. Esse fato que é confirmado por meio da análise entoacional, como veremos. Observe três exemplos em que esse fenômeno ocorre:

- EN12 “ah gostu gostu tudu serviço queu façú eu gostu lá sabi”;
- EN17 “ah eu gostu asisti um um jogu di futibol né”;

- EN20 “/ ah tem: bastanti coisa né?”.

Quanto à análise acústica, a duração das HESs não manteve uma proximidade, variando entre 0,149s (menor duração) e 0,645s (maior duração). A média da duração desse tipo de disfluência foi de 0,317s. A F0 média manteve, para a maior parte dos eventos, uma frequência mais alta, que variou entre 141 Hz e 198 Hz, com uma média de 174 Hz. Os pulsos também variaram, mostrando um esforço vocal variado para cada HES.

Em relação à intensidade, só foram analisados os momentos de intensidade final e intensidade máxima. Não analisamos o período de transição, pois, como dissemos, as HESs ocorreram em início de GT. Sendo assim, a intensidade final apresentou um aumento de 100% dos casos. Por sua vez, a intensidade máxima apresentou um padrão mais alto, sendo a menor de 77 dB e a maior de 84 dB. Como vemos, mesmo sendo mais altas, elas não são muito diferentes. A média da intensidade máxima foi de 76 dB.

A análise entoacional das HES mostrou que elas seguem o padrão das RPs e INSSs, ou seja, mantém o padrão fonológico. Dos casos observados, todos mantiveram um pico entoacional alto *H. Em apenas um caso o padrão foi de tom alto simples (*H). Nos outros eventos o padrão foi composto *H+L, o que indica que o informante variou mais a entoação.

Os alongamentos, por sua vez, foram os menos frequentes na fala de LE. Essa disfluência representa 2,3% do total de eventos. Dentre as ocorrências, uma foi em palavra (AL_p) monossílabo /ca:/, em meio de GT e a outra em vogal (AL_v) em início de GT /i:/. Não foram verificados ALs complexos.

Quanto ao AL_p, a sua duração foi de 0,631s. A média da F0 foi alta, de 183 Hz, bem como os pulsos. A intensidade no momento de transição e final aumentou e a máxima foi de 83 dB.

Para o AL_v a duração foi um pouco maior do que o AL_p 0,642s. No entanto, a F0 média foi um pouco menor, mas ainda mantém um padrão alto com uma frequência de 178 Hz, assim como observado para os pulsos. A intensidade não foi analisada no momento de transição, pois ocorreu em início de GT. Por outro lado, a intensidade final também aumentou em relação à inicial. Já a intensidade máxima manteve um padrão alto, próximo ao que foi observado para a AL_p, de 83 dB.

A análise entoacional mostrou um padrão variado para os ALs. O padrão entoacional para o AL_p foi de *L+H, enquanto, para a AL_v, o pico entoacional foi em um tom alto *H+L. Nesse sentido, LE inicia o AL_p com um tom mais baixo e aumenta a entoação e, no caso da vogal, ocorre o contrário.

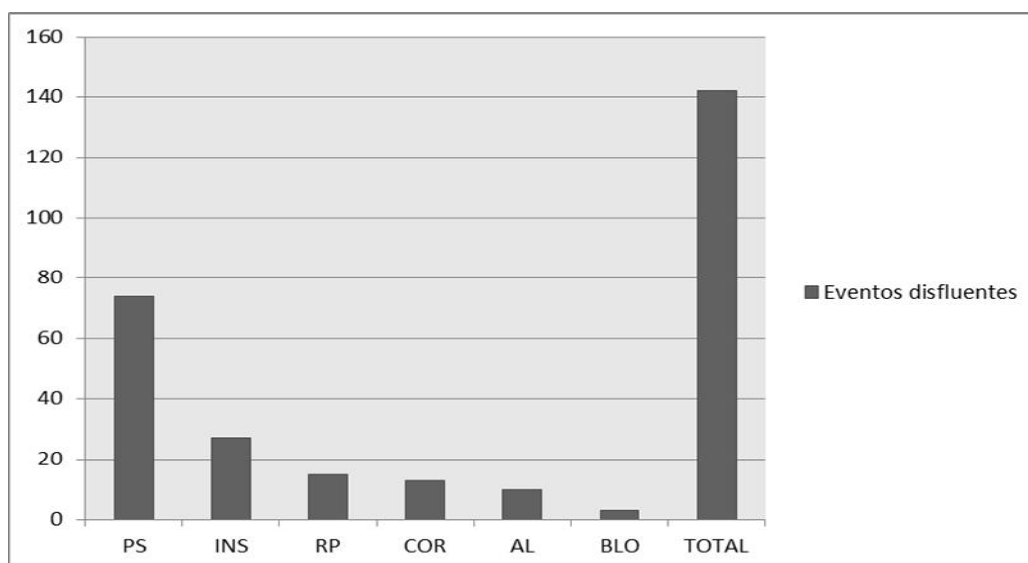
O BLO equiparou-se à produção dos ALs também com 2,3% do total de eventos. Todos eles ocorreram em consoantes oclusivas (BLO_c), sendo uma oclusiva palatal [t], no segmento /-trabalho/, e a outra em oclusiva velar [k], para o segmento /-qui/. Ambas as produções foram analisadas com uma PS de BLO articulatória. Apenas a oclusiva velar [k] apresentou interrupção na forma de onda.

Em relação à análise acústica, a duração das PSs de BLO foi maior apenas para o BLO_c em oclusiva velar. A média das PSs foi de 0,213s, enquanto a dos segmentos foi de 0,311s. Quanto à F0 média, a maior frequência ocorre para o segmento [k], de 172 Hz. A oclusiva palatal apresentou uma F0 média de 165 Hz. Já o esforço vocal é maior para o segmento [t]. No que diz respeito à intensidade, em nenhum dos casos ela foi observada no período de transição, pois vem de uma PS. A intensidade final, por sua vez, aumentou em relação à inicial para os dois segmentos. Por fim, a intensidade máxima é maior para a oclusiva palatal (85 dB) do que para a oclusiva velar (77 dB).

Quanto ao padrão entoacional, os segmentos apresentaram apenas tons simples altos *H.

O informante BR

O informante não gago BR realizou 142 disfluências em 39 enunciados, divididas entre simples e complexas. A porcentagem de disfluências complexas foi de aproximadamente 8,5%. O restante (91,5%) foi de disfluências simples. O máximo de disfluências por GT foi sete. Dentre elas, três foram de PSs. A taxa de elocução de BR foi de 3,799 sil/ms, já a de articulação de 5,030 sil/ms, e o tempo de articulação aproximado de 143ms. Veja, no gráfico 7, a distribuição dos eventos disfluentes na fala de BR:

Gráfico 7- Distribuição dos eventos disfluentes realizados pelo informante não gago BR.

Fonte: própria.

BR oscila sua fala entre seis tipos de disfluências, dispostas, aqui, em ordem decrescente de realização: PSs; INSS; RPs; CORs; ALs e BLOs.

As PSs correspondem à maior frequência de disfluências aproximadamente, 52%, dos eventos. Dentre elas, a mais comum ocorreu em meio de enunciado (40%), depois aparecem as de início de enunciado (25%), seguidas das PSs em final de GT (23%) e, por fim, das PSs em meio de enunciado (12%). No total, 35% aproximadamente das PSs antecederam eventos disfluentes. As PSs em meio de enunciado também foram as que apresentaram a maior duração (26,532s). Em seguida aparecem as PSs em final de enunciado (9,244s) e, por fim, as de início de enunciado (7,951s). Entre PSs silenciosas e respiratórias, estas apresentaram a maior duração.

Depois das PSs, as INSS foram as mais frequentes, representando 19% dos eventos. As INSS dividiram-se entre INS_v (mais frequente, com 62% do total); INS_p e INS_e . O último evento caracteriza-se por apresentar uma INS de dois tipos. Entre as INSS, apenas duas foram complexas, ocorrendo junto com AL.

As INSS realizadas por BR foram:

- INS_p em início de GT - /intão/, /aí/;
- INS_p em fim de GT - /né/, /assim/;
- INS_v em início de GT - /é/, /i/;
- INS_v em meio de GT - /é/, /i/;

- INS complexa em início de GT - /é/;
- INS_e em início de GT - /intão é/, /é i/;
- INS_e em fim de GT - /i tal/.

Assim como tínhamos observado para os outros informantes, as INSs do tipo /intão/ e /ái/ ajudam a sequenciar o enunciado, bem como /i/ e /é/ em início e meio de GT; /intão é/ e /é i/. Diferentemente de MA e LE, a INS_v do tipo /é/ não apresentou um significado afirmativo. A INS_p /né/ também ocorreu no final do primeiro GT, e caracterizou uma interrogação afirmativa retórica. Diferentemente do que vinha sendo observado, BR realizou uma INS_p do tipo /assim/ em final de GT, sendo que o mais comum é que ela ocorra em início ou meio de GT, com função sequenciadora. Em final de GT, /assim/ apresentou a função semântica de comparação e exemplificação. Observe:

- EN27 “uma facultadi achu qui im Pasadina na California i elis são bem típicus nerds americanus assim”.

Nesse enunciado BR fala sobre uma série que se passa em um canal por assinatura *The big ban theory*. Durante a caracterização das personagens usa a INS_p /assim/ com o intuito de comparar e exemplificar as personagens com os jovens americanos.

Por fim, outro tipo de INS que ainda não havíamos identificado foi a INS de estrutura /i tal/. Veja:

- EN39 “alguns ie ero ieróglifus i tal”.

Em EN39 BR discorre sobre algumas séries de ficção científica e as teorias que envolvem a construção de alguns monumentos históricos. Ao tratar dessas questões, BR primeiro erra a pronúncia e se autocorrige. Na sequência, utiliza a INS /i tal/. Nesse caso, a sua função é de preenchimento. O informante supõe que o seu interlocutor tenha entendido o seu pensamento e, ou por não lembrar mais de outros fatos ou por querer terminar o assunto, usa o preenchedor /i tal/, que formalmente equivale à /entre outras coisas/, por exemplo.

A análise acústica revelou que a duração das INS_vs é maior em início de GT (0,359s) do que em fim de GT (0,293s). O mesmo padrão foi verificado para a F0 média e pulsos. Já para as INS_ps a duração foi maior para os segmentos em fim de GT (0,459s), seguidos das INSs em início de enunciado (0,357s) e, por fim em meio de GT (0,316s). Quanto à F0 média e pulsos, elas foram maior em início de GT (124 Hz e 35), depois aparecem em meio (119 Hz e 29) e fim de enunciado (112 Hz e 25). As INSs complexas e as INSs de estrutura, por sua vez, apresentaram uma duração maior, já que temos segmentos maiores, sendo respectivamente 1,018s e 0,700s. Por outro lado, a INS complexa apresentou uma F0 média

menor do que os outros eventos (103 Hz), e a F0 média da INS de estrutura equiparou-se à INS_p em meio de GT (118 Hz). Já o esforço vocal foi maior para a INS complexa.

No que diz respeito à intensidade no momento de transição e, deixando de lado as INSs que iniciaram o GT ou foram antecidas por PSs, em 100% dos eventos as INS_vs aumentaram a intensidade na transição da fluência para a disfluência. Já as INS_ps diminuíram a intensidade em 66% das produções. As INS complexas e de estrutura não foram observadas porque todas ocorreram ou em início de GT ou depois de uma PS. Em relação à intensidade final, aproximadamente 65% das INS_vs aumentaram a intensidade em relação à inicial. Quanto às INS_ps, 80% delas também seguiram o mesmo padrão para a intensidade final. Já 100% dos casos das INSs complexas e de estrutura diminuíram a intensidade. Por fim, a intensidade máxima foi maior para as INSs de estrutura (77 dB). Em seguida, aparecem, respectivamente, as INS_ps em início de GT (69 dB); as INS_ps em meio e fim de GT (68 dB); as INSs complexas (65 dB) e, por fim, as INS_vs em meio de GT (61 dB).

A análise dos padrões fonológicos confirmou a tendência das INSs manterem o padrão para segmentos do mesmo tipo. Em relação às INS_vs, apenas a de início de GT não manteve o mesmo padrão, variando entre um tom alto *H e baixo simples *L. Essa diferença também foi verificada para as INSs complexas. Mesmo que ambas tenham sido do mesmo tipo e alongadas /:é/, elas variaram entre um tom simples (*L) e composto (*H+L). As INSs complexas também não apresentaram o mesmo padrão, variando entre L+*H; *L+H e L+*H. Como vemos, todas começam com um tom baixo. A diferença está nos picos entoacionais. As INS_ps também mantiveram o mesmo padrão para os eventos do mesmo tipo. Por fim, as INSs que apresentaram apenas uma ocorrência de cada evento foram anotadas com os seguintes tons:

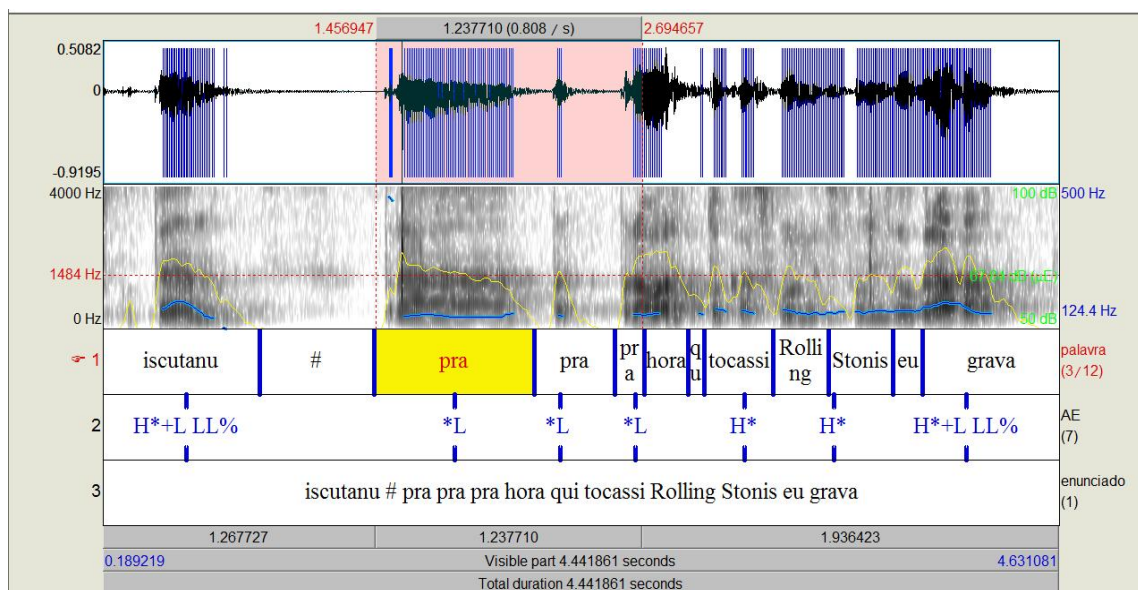
- /né/ - *L+H;
- /assim/ - *H+L;
- /aí/ - L+*H.

As RPs representam 10,5% do total de disfluências realizadas por BR. Esse tipo de disfluência ocorreu em vogal (RP_v, 40%), palavra (RP_p, 53%) e sílaba em palavra (RP_s, 6,7%). Elas também se dividiram entre simples (66,5% aproximadamente) e complexas (33,5% aproximadamente). As RPs complexas ocorreram com AL (dois casos) e com PS_e (100% dos eventos, incluindo as RPs que também apresentaram AL). Dentre todas as RPs, apenas duas delas foram triplas e ocorreram em RP_p, sendo uma em RP_p simples (cf. fig. 66 abaixo) e a outra em RP complexa. Além disso, as RP_ps ocorreram em palavras monossílabas (40%)

dissílabas (40%) e trissílabas (20%). Vale salientar que os casos de segmentos desvozeados ou, *creaky voice*, ocorreram sempre na última RP, independentemente de serem duplas ou triplas.

No que diz respeito à análise acústica, a duração mostrou para as RPs triplas uma diminuição progressiva nos dois casos. Veja um exemplo a seguir:

Figura 35 – Exemplo de diminuição da duração para uma RP tripla no enunciado “iscutanu# pra pra pra hora qui tocassi Rolling Stonis eu grava” produzido pelo informante não gago BR.



Fonte: própria.

Como vemos na figura 35, BR diminui a duração na passagem de uma RP para a outra. A primeira tem uma duração de 0,743s, a segunda, de 0,372s e a última, de 0,140s. As RPs de palavra, vogal, sílaba e RP complexa também apresentaram um tendência para a diminuição da duração em respectivamente 66,6%, 60%, 100% e 75% dos casos.

Quanto à F0, a tendência também é para que a frequência diminua na segunda repetição. Fato que também foi observado para a RP de sílaba. No entanto, para a RP tripla essa tendência não se mantém. Como está exemplificado na figura 66, nesse caso especialmente, a F0 média aumenta progressivamente (112 Hz > 114 Hz > 116 Hz). Por outro lado, na outra ocorrência, há uma diminuição progressiva, que pode destacar-se já que a terceira RP é desvozeada (124 Hz, 118 Hz e 0 Hz). Os pulsos seguiram a mesma tendência observada para a F0 média, isto é, o esforço vocal foi menor na segunda RP para as palavras, vogais e sílabas. Quanto às RPs triplas, os pulsos seguiram um padrão de onda para o exemplo identificado na figura acima e mantiveram o padrão da F0 no segundo caso.

Por fim, observamos para a intensidade no momento de transição da fluência para a primeira RP uma diminuição para a RP_p (100%) e RP complexa (66,6%). Para as RP_vs esse parâmetro não foi observado, uma vez que os eventos ou foram antecidos por PS ou ocorreram em início de GT, bem como para a RP de sílaba. Já para as RPs triplas, o caso que não foi antecedido por PS (exemplificado na figura 66 acima) apresentou uma diminuição na intensidade. No que diz respeito à transição da intensidade da primeira RP para a segunda, verificamos que para as RPs de palavra o padrão inverte-se, ou seja, em 100% dos casos a intensidade aumenta. Esse parâmetro não foi observado para a RP complexa, uma vez que todas foram antecidas por uma PS. Quanto às RPs de vogal, em 60% dos casos ocorre uma diminuição da intensidade. Já para os outros eventos, a intensidade nesse momento aumenta para a RP de sílaba e, para as RPs triplas, em um evento ela diminui (cf. figura 66) e no outro aumenta na última RP, uma vez que a intermediária é antecida por uma PS.

A intensidade final apresenta uma tendência para o abaixamento para a RP_p (100%) e para a RP_v (60%). Já para a RP complexa a tendência é para o aumento (62,5%) da intensidade final. As RPs triplas dividiram-se entre o aumento no evento simples e abaixamento no evento complexo. Finalmente, a RP_s, bem como as RPs triplas não apresentaram uma tendência, uma vez que na repetição a intensidade final diminuiu e na sílaba em palavra completa ela aumentou.

Quanto à intensidade máxima, há um aumento em 100% dos casos na segunda repetição para as RP_ps. Para as RPs complexas o padrão se iguala em 50% dos eventos diminui e em 50% deles. No que diz respeito às RPs triplas, nos dois eventos o padrão mantém uma forma de onda, de modo que a intensidade começa mais alta, diminui e depois aumenta. Esse fenômeno está identificado na figura 66 por meio dos picos na curva de intensidade (em amarelo). Para as RP_vs a tendência é para a diminuição da intensidade na segunda RP (60%). Por sua vez, para a RP_s a intensidade final diminui na sílaba repetida e aumenta na sílaba em palavra completa.

A análise entoacional confirmou que os padrões entoacionais se mantêm para as RPs. Em relação às RP_ps, em apenas um caso o padrão variou de *L para *H+L. Em relação aos outros casos de RPs, tivemos um caso para RP complexa em que o padrão não se manteve. Nos segmentos desvozeados o padrão também não poderia ser o mesmo, já que não há frequência. As RP_s também mantiveram o mesmo padrão (*H) para a sílaba repetida e para a sílaba em palavra.

Seguido da RP, ocorre a disfluência do tipo COR, que caracteriza, aproximadamente, 9% do total de disfluências. A COR dividiu-se entre eventos simples e complexos. Dentre as CORs, o primeiro tipo foi o mais frequente, com 61,5% das ocorrências. As CORs complexas ocorreram junto com PI; PS_e, AL e INS_v. Observe os exemplos:

1. “eu *proc* > *pensava* em alguma coisa relacinada a minha habilidadi manual”;
2. “u >/eu aconselhu sim”;
3. “ti > *tevi uma fasi*: muito boa du Chicagu Bulls”;
4. “qui chama *chi é* > *Chariots of the Gods*”.

Os eventos com erro e a sua respectiva correção aparecem em itálico. O símbolo maior do que (>) corresponde à correção feita pelo informante.

No primeiro exemplo temos uma COR de palavra incompleta. No segundo, o informante introduz um segmento, faz uma PS respiratória que o ajuda a recuperar a fluência e corrige o GT, mudando totalmente o que havia sido programado. No terceiro, BR erra na conjugação verbal e na sequência corrige o GT, alongando a sílaba final do segmento /fasi:/. Por fim, no quarto exemplo o informante erra a sílaba inicial da banda e, com o intuito de recuperar a fluência insere a vogal /é/. Dentre todos os tipos de CORs, em 61,5% dos eventos o informante recupera parte da informação do erro na COR.

A análise acústica revelou que, em 60% dos casos que na COR complexa a duração da correção é menor do que no erro. Já na COR simples ocorreu o contrário, já que em 75% dos casos a duração aumentou na correção. Quanto à F0 média, em 80% dos eventos a F0 aumenta no momento da correção. Por outro lado, na COR simples não foi identificado uma tendência, uma vez que em 50% dos eventos a F0 diminuiu e em 50% aumentou. Em relação aos pulsos, tanto na COR simples quanto na COR complexa a tendência é para que o esforço vocal seja maior na correção e não no erro. Vale salientar que os momentos desvozeados ocorreram apenas no erro.

A análise da intensidade revelou, deixando de lado os momentos de PS (duas ocorrências) e COR em início de GT (cinco ocorrências), para a COR complexa, que em 66% dos eventos a intensidade aumentou na passagem do erro para a correção. Já no momento de transição do erro para a correção o padrão equiparou-se, ou seja, em 50% dos casos a intensidade aumentou e, em 50%, diminuiu. Já para a COR simples em 83% dos casos a intensidade diminuiu na passagem da fluência para o erro e, em 87,5% aumentou na transição do erro para a COR. Quanto à intensidade final, em 56% dos casos ela aumentou na COR simples. Por sua vez, a COR complexa não apresentou um padrão, uma vez que em 50% dos

casos a intensidade final diminuiu e em 50% deles aumentou. Por fim, para a intensidade máxima houve um aumento em 100% dos casos no momento de correção para a COR complexa. Para a COR simples a tendência também foi para o aumento. No entanto, esse padrão foi observado em 75% dos casos.

Em relação à análise fonológica, os padrões são diferentes, assim como vem acontecendo para os outros informantes. Sendo assim, eles variaram entre simples e complexos na passagem do erro para a correção. Nessa transição, os padrões mantiveram-se altos na passagem do erro para a correção ($H > H$) em aproximadamente 46% dos casos e, em 23% deles, a frequência diminuiu ($H > L$). Os demais eventos representaram períodos sem sonorização. Nesses casos, o padrão entoacional aumentou (mais frequente) e diminuiu (um evento).

Os alongamentos representam 7% das produções. Os eventos foram simples e complexos. O primeiro tipo ocorreu apenas para palavras, sendo elas monossílabas (quatro eventos); dissílabas (quatro eventos) e trissílabas (um evento). Vale lembrar que os ALs complexos ocorreram juntamente com a INS_v , como vimos. Dentre os eventos, foram verificados um AL_p em início de sílaba e o restante em sílaba final de palavra. A média de duração das sílabas alongadas foi de 0,662s para os ALs em palavras dissílabas em sílaba final; 0,627s para monossílabos; 0,588s para palavras trissílabas e 0,407s para palavras dissílabas em sílaba inicial. A mesma proporção foi verificada para a F_0 média, ou seja, a F_0 é maior para o AL em fim de sílaba em palavras dissílabas. Mesmo assim, a maior diferença entre elas é de 6 Hz. Por sua vez, a análise dos pulsos mostrou um esforço vocal maior para o AL em palavras trissílabas. Na sequência aparecem os ALs em monossílabos, seguidos do AL em início de sílaba e final de sílaba em palavras dissílabas.

Quanto à intensidade no momento de transição, três produções foram antecidas por PSs. O restante, por sua vez, apresentou para 85% dos dados uma diminuição da intensidade na passagem da fluência para a disfluência. Em relação à intensidade final, 60% das produções diminuíram a intensidade em relação à inicial. Por fim, a intensidade máxima foi maior para os ALs em sílaba final (75 dB), seguidos do AL em palavra trissílabas (70 dB) dos ALs em início de sílaba (69 dB) e, por fim, dos monossílabos (66 dB).

A análise entoacional mostrou um padrão variado para os ALs. Os padrões variaram entre tons simples (*L e *H) e compostos (*H+L; *L+H; H+*L e L+*H). Mesmo com essas variações, a tendência foi para que os ALs iniciassem com um tom alto (H, 70%) e pico entoacional também alto (60%).

Finalmente, o BLO foi a disfluência menos comum para BR, representando 2% do total de eventos. Todas elas ocorreram em consoantes oclusivas bilabiais, sendo duas em [p], nos segmentos /-principalmente/ e /personalidade/ e uma em [b], para o segmento /-bem/. Dentre essas produções, duas apresentaram PS de BLO silenciosa, ambas para [p], e uma articulatória, em [b]. Além disso, duas delas apresentaram interrupção na forma de onda, sendo uma em [p], no segmento /-personalidade/ e a outra em /b/, /bem/

Em relação à análise acústica, a duração das PSs de BLO não foi maior do que nenhum segmento. A sua média foi de 0,249s e a média da duração dos segmentos de 0,631s.

Quanto à F0 média, a maior frequência ocorre para a consoante [b] (147 Hz), enquanto as do tipo [p] apresentaram uma F0 de (116 Hz e 128 Hz). Em relação aos pulsos, o esforço vocal também foi maior para [b]. Por fim, a intensidade em momento de transição não foi observada para nenhum dos eventos. A intensidade final revelou para 66% dos casos uma tendência para o aumento. Finalmente, a intensidade máxima também seguiu a tendência da F0 e pulsos, sendo maior para o segmento /b/.

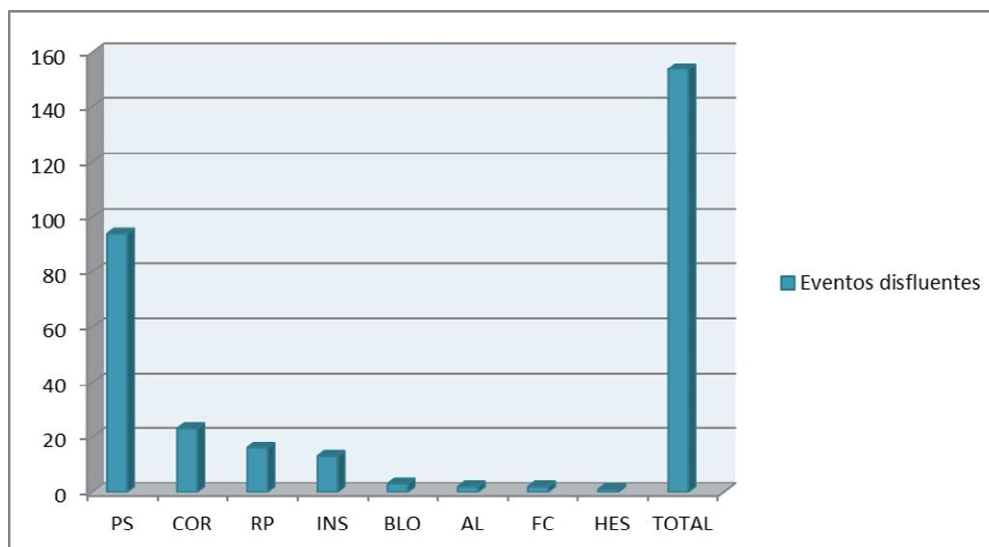
Quanto ao padrão entoacional, os segmentos apresentaram apenas tons compostos, dos tipos H+*L em /-personalidade/; *L+H em /-principalmente/ e, H+*L em /-bem/.

O informante NE

O informante não gago NE realizou 154 disfluências em 53 enunciados, divididas entre simples e complexas. A porcentagem de disfluência complexa foi de aproximadamente 8,5%, sendo 91,5% de disfluências simples. O máximo de disfluências por GT foi de cinco. Dentre os eventos, quatro foram do tipo PS. NE apresentou uma taxa de elocução de 5,119 sil/ms, uma taxa de articulação de 6,748 sil/ms e um tempo de articulação aproximado de 142ms.

Veja abaixo (gráfico 8) a distribuição dos eventos disfluentes na fala de NE:

Gráfico 8 – Distribuição dos eventos disfluentes realizados pelo informante não gago NE.



Fonte: própria.

O informante varia sua fala entre oito tipos de disfluências, dispostas em ordem decrescente de realização: PSs; CORs; RP; INSs; BLOs; AL; FC e HES.

As PSs correspondem à maior frequência de disfluências: aproximadamente 61% dos eventos. Dentre as disfluências, a mais comum é a de início de GT (38,3%), seguidas das PSs em final de GT (34%) e, por fim, das PSs em meio de GT (27,7%). No total 23,5% aproximadamente, das PSs antecederam eventos disfluentes.

Quanto à duração, as PSs em início de enunciado também foram as que apresentaram a maior duração (19,157s). Em seguida aparecem as PSs em meio de GT (13,165s), e, por fim, as de fim de GT (13,032s). Como podemos ver, a diferença entre os dois últimos tipos é pouco significativa. Entre PSs silenciosas e respiratórias, a última apresentou a maior duração (29,925s).

Depois das PSs, as CORs caracterizaram, aproximadamente, 15% do total de disfluências. Elas dividiram-se entre eventos simples (61%) e complexos (39%). As CORs complexas ocorreram junto com PS_e (seis casos), INS_v (dois casos) e RP_v (um caso). Dentre as CORs, o informante manteve parte da informação contida no erro na correção em apenas 26% dos casos.

A análise acústica revelou, para 78,6% dos casos que na COR a duração aumentou na correção, bem como observado para a COR simples (55,5%). Já para a F0 média, tanto para a COR simples quanto para a COR complexa, a tendência é para que a entoação diminua (78%)

no momento da correção para os dois tipos de COR. É válido salientar que quando houve uma qualidade de voz do tipo *creaky voice*, ela sempre ocorreu na correção. No que diz respeito aos pulsos, tanto na COR simples quanto na COR complexa a tendência é para que o esforço vocal seja maior na correção e não no erro, sendo a porcentagem de 55,5% para a COR complexa, e de 71% para a COR simples.

A análise da intensidade revelou, deixando de lado os momentos de PS em início de GT, uma tendência para a diminuição da intensidade na transição da fluência para o erro (55,5%) e do erro para a correção (71%). Já para a COR simples os resultados equipararam-se para os dois momentos observados. Quanto à intensidade final, seja para a COR simples ou para a COR complexa os resultados foram iguais, ou seja, em 50% dos casos, nos dois tipos de correção a intensidade final aumentou em relação à inicial e, em 50%, diminuiu. Por fim, a análise da intensidade máxima foi diferente para a COR simples e COR complexa. No primeiro caso, a intensidade máxima aumentou em 57,2% dos casos. Por outro lado, no segundo caso, ela diminuiu para 66,6% dos casos.

Em relação à análise fonológica, os padrões foram diferentes na passagem do erro para a correção. Eles variaram entre tons simples e compostos na passagem de um momento para o outro. Dessa maneira, nessa transição, os padrões mantiveram-se o tom baixo (L>L) em 26% dos casos e alto (H>H) em apenas 4% dos dados. Os demais eventos variaram, por exemplo, em casos em que o erro começou com um tom alto, a correção apresentou um tom baixo, e assim sucessivamente. No entanto, vale salientar que para BR a tendência foi para que a correção iniciasse com um tom baixo.

Seguido da COR aparecem as RPs, totalizando 10,4% aproximadamente das disfluências realizadas por NE. As RPs foram verificadas em palavra (RP_p, 68,8%) em sílaba em palavra (RP_s, 18,8%) e em vogal (RP_v, 12,4%). Também foram verificadas RPs complexas (18,8%), sendo todas elas realizadas com PS_c. O informante realizou apenas RPs duplas, sendo 12,5% em palavras dissílabas e 87,5% em monossílabas. Os segmentos desvozeados ou com uma qualidade de voz do tipo *creaky voice* foram identificados na última RP.

No que diz respeito à duração, em todos os casos de RPs ela diminuiu na segunda repetição. Por sua vez, a média da F0 apresentou um padrão variado. A repetição do tipo 2RP_p apresentou uma diminuição da entoação em 87,5% dos casos. Já para a RP_v, os dois casos verificados apresentaram uma qualidade de voz do tipo *creaky voice*. Na RP complexa, por sua vez, dentre as três ocorrências, ocorreu um caso de diminuição da entoação e, nos outros dois, uma qualidade de voz do tipo *creaky voice*. Por fim, para a RP_{ps} a tendência foi para o

aumento da entoação em 66,6% dos eventos. Os pulsos apresentaram o mesmo padrão observado para a análise da F0 média.

Em relação à intensidade no momento de transição da fluência para a primeira RP e da disfluência para a segunda RP, ocorreu uma tendência ao aumento nos dois momentos para a RP_{ps} (66% nos dois momentos) e uma diminuição para a repetição do tipo 2RP_p (66% e 77%, respectivamente). Já para a RP_v, a transição da fluência para a primeira RP não foi observado, pois elas ou iniciaram o GT ou foram antecidas por PS. Já na passagem da disfluência para a segunda RP, a intensidade diminuiu em 100% dos casos. Por fim, a intensidade para a RP complexa só pode ser observada na transição da fluência para a disfluência, uma vez que a segunda RP foi antecida por uma PS_e. Sendo assim, em 66% dos casos a intensidade diminuiu. Quanto à intensidade final, não foi possível chegar a uma padronização, pois para todos os tipos de repetição os dados igualaram-se, isto é, em 50% dos casos a intensidade aumentou em relação à final e, em 50%, diminuiu. Finalmente, a análise da intensidade máxima revelou para a RP_v e RP complexa uma diminuição da intensidade em 100% dos casos na segunda repetição. Já a RP_{ps} apresentou uma tendência para o aumento da intensidade (66%). Por fim, a repetição do tipo 2RP_p apresentou resultados iguais, o que dificulta o estabelecimento de parâmetros.

A análise entoacional, mais uma vez, manteve o padrão para segmentos iguais. A diferença ocorreu apenas para as repetições com a qualidade de voz *creaky voice*, em que não ocorre variação melódica. No entanto, a primeira RP apresentou um tom baixo *L, mantendo uma entoação baixa. No que diz respeito às RPps de sílaba, apenas um caso apresentou um tom diferente, variando de um tom baixo para um tom alto.

As INSS representaram 8,5% aproximadamente dos eventos disfluentes. Elas dividiram-se entre INS_v (46%); INS_p (23%) e INS_e (15,5%) e INS complexas (15,5%). Dentre as INSS mais frequentes, a INS_v em meio de GT foi a mais comum para NE. O último tipo de inserção ocorreu com alongamento em vogal. As INSS realizadas por NE estão assim distribuídas:

- INS_p em início de GT - /intão/, /tão/;
- INS_p em fim de GT - /né/;
- INS_v em início de GT - /é/, /i/;
- INS_v em meio de GT - /é/;
- INS complexa em início de GT - /i/;
- INS_e em início de GT - /intão é/;

Assim como tínhamos observado para os outros informantes, as INSs do tipo /intão/ e /tão/ ajudam a sequenciar o enunciado, bem como /i/ e /é/ em início e meio de GT e também /intão é/. A INSs complexa /i/ também é sequenciadora, entretanto ela difere das outras INSs no sentido em que o informante demora mais tempo para processar a mensagem e, por isso, tem que alongar a vogal ou palavra. A INS_v do tipo /é/ não apresentou um significado afirmativo. As demais INSs mantêm as características observadas para os outros informantes.

A análise acústica revelou a maior duração média para a INS complexa (0,730s), fato esperado, pois ocorre junto com AL. Na sequência temos: as INS_{v,s} em início de GT (0,476s); a INS_e (0,384s); as INS_v em meio de GT (0,340s); a INS_p final (0,223s) e, por fim, as INS_{p,s} em início de GT (0,185s).

Em relação à F0 média, o padrão não se manteve. Nesse caso, a maior frequência foi observada para a INS_e (124 Hz). Depois, aparecem: as INS_v em meio de GT (120 Hz); a INS_p final (119 Hz); as INS_{v,s} em início de GT (114 Hz); a INS complexa (110 Hz) e as INS_{p,s} em início de GT (102 Hz). Como podemos ver, a diferença entre as frequências são pequenas, o que mostra que o informante variou pouco a entoação. Os pulsos também seguiram o mesmo padrão apenas para a INS complexa, que apresentou o maior esforço vocal, e foi seguida das INS_{v,s} em início de GT; INS_e; INS_p final; das INS_{v,s} em meio de GT, além das INS_{p,s} em início de GT.

Em relação à intensidade no momento de transição e, deixando de lado as INSs que iniciaram o GT ou foram antecidas por PSs (INS_v em início de GT; INS complexa; INS_p inicial e INS_e), as INSs em início de GT (80% dos casos) e a INS_p em final de GT (um caso) apresentaram um aumento na intensidade. Por sua vez, a intensidade final aumentou para a INS_e (um caso); para as INS_{p,s} em início de GT (100%, dois casos) e para a INS_e (um caso). Para as INS_{p,s} inicial (dois casos) e INS_{v,s} (seis casos, 83% aproximadamente), a intensidade final aumentou. Finalmente, para as INS_{v,s} em início de GT a intensidade aumentou para a vogal /é/ e diminuiu para a vogal /i/. Por fim, a média da intensidade máxima foi maior para a INS_e de estrutura (85dB). Depois dela, temos: as INS_{v,s} em meio de GT (76 dB,); a INS_p em final de GT (74 dB); as INS_{p,s} (72,866 dB) e INS_{v,s} (72,751 dB) em início de GT, com uma diferença de 0,115 dB, e, por fim, a INS complexa, com uma intensidade máxima de 71dB.

A análise dos padrões fonológicos confirmou a tendência das INSs manterem o mesmo padrão entoacional. Mesmo variando entre /tão/ e /intão/, as duas INSs em início de GT mantiveram o tom baixo *L, bem como a INS_p em final de GT, a INS_v /é/ e a INS complexa, ambas em início de GT. As INS_{v,s} em meio de GT também apresentaram a maior

tendência para o tom baixo *L. Apenas um enunciado apresentou o tom composto *H+L. A INSe, por sua vez, foi anotada com o padrão L+*H.

Seguido das INs ocorre a disfluência do tipo BLO. O BLO foi verificado para 1,95% dos eventos disfluentes. Eles ocorreram apenas para consoantes, sendo dois em oclusiva velar [k], nas palavras /-consigu/ e /-queli/ e um em oclusiva bilabial [p], na palavra /-podi/. As pausas de BLO foram do tipo 2 (silenciosa e sem articulação e pulsos), para a oclusiva bilabial [p] e velar [k] e, do tipo 1 (articulatória) para a oclusiva velar na palavra /-queli/.

Em relação à análise acústica, a duração das PSs de BLO não foi maior do que nenhum segmento. A sua média foi de 0,287s e a média da duração dos segmentos foi de 0,392s. Quanto à F0 média, a maior frequência ocorre para a consoante [p] (182 Hz) e depois as oclusivas velares (152 Hz). Por outro lado, o esforço vocal é maior para as oclusivas velares.

A intensidade em momento de transição não foi observada para nenhum dos eventos, já a intensidade final aumentou para 100% dos casos. Finalmente, a intensidade máxima foi maior para as oclusivas bilabiais (83 dB), sendo a média das oclusivas velares de (74 dB).

Em relação ao padrão entoacional, os segmentos apresentaram tons compostos, para os segmentos /-consigo/ e /-queli/, respectivamente com os tons *L+H e *H+L e, um tom simples alto *H para o segmento /-podi/.

Os ALs e FCs foram verificados para 1,3% das produções. O primeiro tipo ocorreu em palavra (AL_p) /:tem/ e vogal /:i/ (AL_i). A maior duração foi verificada para o AL_p (0,954s), já o AL_v apresentou uma duração de 0,797s. A frequência média do alongamento foi de 122 Hz para o AL_p e 117 Hz para o AL_v. A análise da intensidade no momento de transição foi observada apenas para o AL_p, uma vez que o alongamento em vogal ocorreu em início de GT e após uma PS. No caso do AL_p, a intensidade no momento de transição diminuiu. Já a intensidade final aumentou em relação à inicial para o AL em vogal e diminuiu para o AL em palavra. Por fim, a intensidade máxima também é maior para o AL em palavra (78 dB), sendo a intensidade máxima para o AL em vogal de 76 dB.

O padrão entoacional, por sua vez, foi o mesmo para os dois tipos de AL. Ele foi anotado com o tom composto *H+L. Esse resultado foi apenas uma coincidência, pois, como vimos para os outros informantes, a tendência é que os alongamentos mantenham padrões entoacionais distintos.

O segundo tipo de disfluência, FC foi identificado nos seguintes enunciados:

- EN39 “um dia seu tivé moranu nu *fu*, num senu funcionáriu públicu lá na/nu Riu#”;
- EN52 “/entri otus *u* eli tem Libertinagi”.

Nos exemplos os FCs aparecem em itálico e são, respectivamente, os segmentos /fu/ e /u/. Veja que o informante apenas insere os segmentos no enunciado. Como vimos, esses segmentos não tem ligação com o que vem antes ou depois deles. No entanto, em EN39, o FC /fu/ está relacionado ao segmento /funcionáriu/. Mesmo assim, ele não caracteriza uma PI, uma vez que o informante não conclui o seu pensamento no primeiro GT /um dia seu tivé moranu nu *fu*/, iniciando o próximo GT com uma nova informação.

A análise da duração variou entre os segmentos diferentes, sendo de 0,337s para /fu/ e 0,087s para /u/. O segmento /fu/ é surdo, por isso, não foi analisado em relação à F0 média e pulsos. Já o segmento /u/ apresentou uma F0 média de 128 Hz. Por sua vez, a intensidade no momento de transição diminuiu para /fu/ e aumentou para /u/. A intensidade final diminuiu para os dois segmentos. A intensidade máxima foi menor para /fu/, de 44 dB, enquanto para /u/ foi de 76 dB.

A análise entoacional não foi anotada para /fu/, já que não apresentou frequência. Já para /u/ foi anotada com o tom simples baixo *L.

Por fim, com o menor número de ocorrência temos a HES, com 0,65% dos casos. Essa disfluência ocorreu para o segmento /ah/, bem como verificado para os outros informantes. No entanto, ela foi identificada não em início de GT, mas sim em meio de GT e após uma PS, assim como ocorreu para o informante CL (como veremos abaixo). Veja abaixo:

- EN4 “a árvori / *ah* corta ávori, planta outra”.

Observe que, assim como analisado para as INs do tipo /né/, que ocorrem normalmente em final de GT mas que podem acontecer em meio de GT, essa HES mantém a função pragmática das hesitações que ocorrem em início de GT. Isso ocorre porque temos dois GTs: no primeiro o informante introduz uma informação /a árvori/. Na sequência, ele para, fazendo uma PS respiratória e inicia um novo GT, por meio da disfluência hesitativa /*ah* corta ávori planta outra/. A função pragmática dessa hesitação não indica surpresa, mas uma resposta com maior variação entoacional. Tal fato também fica comprovado por meio da análise entoacional que foi anotada com o tom composto *H+L.

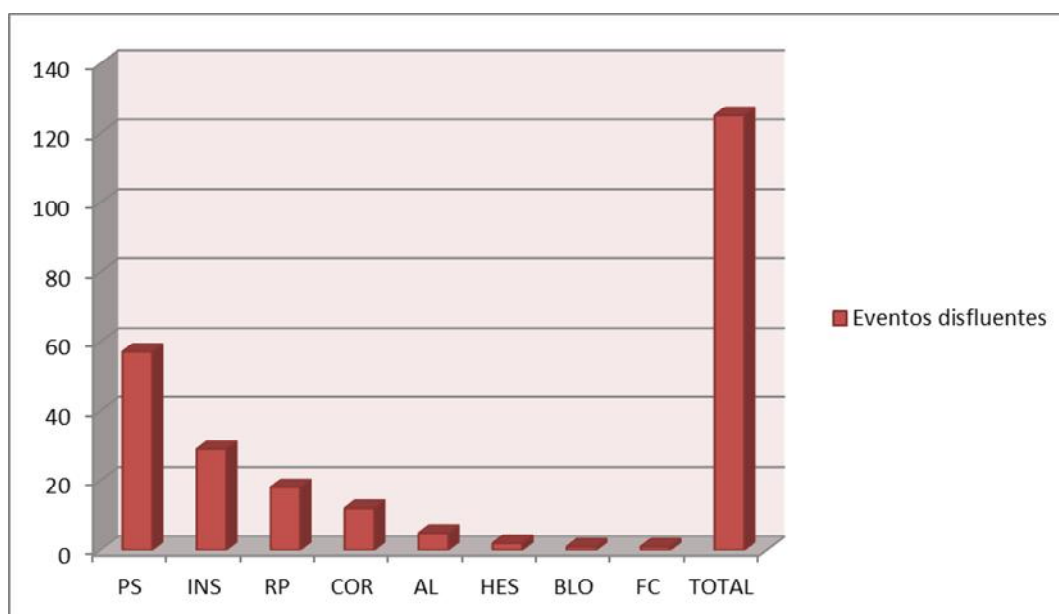
Essa HES apresentou uma duração de 0,369s. A média da F0, se comparada a das outras disfluências é alta, já que apresentou uma frequência de 178 Hz. A intensidade de

transição não foi analisada e a intensidade final aumentou quando comparada à inicial. A intensidade máxima também foi alta, de 83dB.

O informante CL

O informante não gago CL é o único informante do sexo feminino do grupo dos não gagos. O seu tipo de fala caracteriza-se por ser mais agudo, por isso a presença de frequências mais altas e maior interferência na gravação. O informante realizou 125 disfluências em 47 enunciados, divididas entre simples (97,6%) e complexas (2,4%). O máximo de disfluências por GT foi de seis. Dentre essas disfluências duas foram do tipo PS. CL apresentou uma taxa de elocução de 4,306 sil/ms, de articulação de 5,060 sil/ms e, um tempo de articulação aproximado de 136 sil/ms. A seguir apresentamos no gráfico 9 a distribuição dos eventos disfluentes na fala de CL:

Gráfico 9 – Distribuição dos eventos disfluentes realizados pelo informante não gago CL.



Fonte: própria.

CL varia entre oito tipos de disfluências, dispostas, aqui, em ordem decrescente de realização: PSs; INSs; RPs; CORs; ALs; HESs; BLO e FC.

As PSs, assim como ocorreu para CR, MA e BR, correspondem a maior frequência de disfluências, como mostra o gráfico. Essa disfluência foi verificada em 45,6% dos eventos. Dentre elas, a mais comum é a do tipo PS_e (em meio de enunciado), que foi anotada para

36,8% das pausas. Com uma diferença muito pequena temos as PSs de início de enunciado, representando 35% dos eventos. Por fim, aparecem as PSs de final de GT, somando 28,2% do total. As PSs que antecederam ou sucederam os eventos disfluente representam 52,2% das pausas.

Em relação à duração, as PSs em meio de enunciado também foram as que apresentaram a maior duração (10,183s). Em seguida aparecem as PSs em início de GT (8,080s) e, por fim, as de fim de enunciado (5,686s). Entre PSs silenciosas e respiratórias, as últimas foram as mais frequentes e de maior duração (13,808s). As PSs silenciosas tiveram uma duração total de 10,797s.

Depois das PSs, as INSS foram as mais frequentes, representando 23,2% dos eventos. As INSS dividiram-se entre INS_v e INS_p (mais frequente, com 75,8% do total) As INSS realizadas por CL foram:

- INS_p em início de GT - /intão/, /tão/;
- INS_p em fim de GT - /né/; /sabi/;
- INS_p em meio de GT - /intão/;
- INS_v em início de GT - /é/; /i/;
- INS_v em meio de GT - /é/;

Assim como tínhamos analisado para os outros informantes, as INSS do tipo /intão/ e /tão/ ajudam a sequenciar o enunciado, bem como /i/ e /é/ em início e meio de GT e /intão/ em meio de GT. A INS_v do tipo /é/ não apresentou um significado afirmativo. A INS_p /né/ ocorreu no final do primeiro GT. No entanto, assim como para os outros informantes, ela também indicou uma interrogação afirmativa retórica, bem como /sabi/.

A análise acústica revelou a maior duração para as INS_{ps} (0,537s). Em seguida aparecem as INS_{vs} em início de GT (0,420s); as INS_{ps} em meio de GT (0,362s); as INS_{ps} em fim de GT (0,341s) e, finalmente, a INS_v em meio de GT (0,171s). Quanto à F0 média, a maior entoação é verificada para a INS_v em meio de GT. Na sequência aparecem, em ordem decrescente, as INS_{ps} em fim de GT; as INS_{ps} em início de GT; a INS_p em meio de GT e, por fim, as INS_{vs} em início de GT. Quanto aos pulsos, o esforço vocal é maior para as INS_{vs} em início de enunciado. Depois, temos as INS_{ps} iniciais; as INS_{ps} finais; a INS_v em meio de GT e, por último, a INS_p em meio de GT.

No que diz respeito à intensidade no momento de transição e, deixando de lado as INSS que iniciaram o GT ou foram antecidas por PSs (INS_{ps} em início de GT e INS_{vs} em início e meio de GT), a INS_p em meio de GT caracterizou-se por um abaixamento na

intensidade e, as INS_{ps} em final de GT apresentaram uma tendência para o aumento da intensidade (83%). Em relação à intensidade final, as INS s em início (60%) e meio de GT (caso único) apresentaram um aumento na intensidade se comprada à inicial, assim como as INS_{vs} em início de GT (83%). Já as INS_{ps} em final de GT (80%) e a INS_v (caso único) em meio de GT caracterizaram uma intensidade final que diminuiu em relação à inicial. Finalmente, a análise da intensidade máxima revelou para a INS_p em meio de GT a maior intensidade (83 dB). Em seguida, com diferenças pouco significativas temos as INS_{ps} em início de GT (82 dB); a INS_v em meio de GT (79 Hz); as INS_{ps} em final de GT (75 dB) e as INS_{vs} em início de GT.

A análise dos padrões fonológicos confirmou a tendência das INS s do mesmo tipo manterem para a maior parte dos casos o mesmo padrão entoacional. Em relação às INS_{ps} em início de GT, 60% dos casos apresentaram o padrão, *H+L. Os outros casos apresentaram o tom composto L+*H. Como vemos, mesmo sendo diferentes, o pico entoacional continua. Já a INS_p em meio de GT apresentou o padrão L+*H. Por sua vez, as INS_{ps} em final de GT do tipo /né/ variaram os padrões entre os tons: L+*H (29%) ; *H+L (7%) ; *L+H (56%) e *L (8%). Se levarmos em conta a porcentagem de ocorrência de cada padrão entoacional, bem como os picos entoacionais, temos a maior tendência para o tom *L+H, com pico entoacional no tom baixo.

As RPs representam 14,4% do total de disfluências realizadas por CL. Esse tipo de disfluência ocorreu em palavra (RP_p , 38,8%) em sílaba de palavra (RP_s , 11%, aproximadamente) em estrutura (RP_e , 22,2) e por fim, em vogal (RP_v , 28%). Dentre todas as RPs, apenas uma delas foi tripla, do tipo $3RP_p$. As RP_{ps} ocorreram em palavras monossílabas (85,7%) e em palavra trissílaba (14,3%). Quanto às repetições de sílaba, em um evento a RP foi de uma sílaba (RP_{ps}) e no outro de sílaba dupla (RP_{ps2}). Observe o exemplo:

- EN12 “a genti tem a pastoral *voca* vocacional”;
- EN13 “*pa* pastoral da criança”.

Veja que, em EN12, o informante não repete apenas a sílaba inicial, que seria /vo/ de /vocacional/, mas sim, as duas sílabas iniciais /voca/. Já em EN13 ocorre uma RP em início de GT.

No que diz respeito à análise acústica, a duração mostrou para a RPs tripla uma duração em forma de onda, que começa mais baixa, aumenta e, em seguida, diminui (0,188s; 0,282s e 0,177s). Já para a RP_{ps} , em 100% dos casos a duração aumentou na sílaba em palavra, bem como para a RP_v (80%). Por outro lado, para as $2RP_{ps}$, a duração diminuiu na

segunda repetição (66,6%). Por fim, para a RP_e a duração equiparou-se (50% diminuiu e 50% aumentou). A análise da F0 média revelou para a disfluência do tipo $3RP_p$ o mesmo padrão de onda observado para a duração. Para a RP_{ps} , na repetição de uma sílaba, a entoação diminuiu e, na de duas sílabas, aumentou. Por sua vez, a RP_e manteve o padrão da duração. Fato que também foi verificado para as $2RP_{ps}$ (50% a entoação diminuiu e, em 50% aumentou). Finalmente, para a RP_v a tendência foi para a diminuição da F0 (60%). Os pulsos seguiram o mesmo padrão da F0 média para as disfluências do tipo $3RP_p$; RP_{ps} e RP_e . Já para a RP_v , o esforço vocal foi maior para a segunda repetição (80%) e, para a disfluência do tipo $2RP_p$ em 83% dos casos, aproximadamente, o esforço vocal foi menor na segunda repetição.

No que diz respeito à transição da intensidade da primeira RP para a segunda verificamos que para a $3RP_p$ a intensidade aumentou nos três momentos. Já para RP_{ps} a intensidade diminuiu, tanto na transição da fluência para a disfluência - ou da disfluência para a disfluência – quanto para a RP_v e $2RP_p$. Para a RP_e a tendência foi para a diminuição (66,6%) na passagem da fluência para a primeira RP_e e aumento (75%) na transição da primeira RP_e para a segunda.

A intensidade final, por sua vez, apresenta uma tendência para o aumento nos seguintes eventos: RP_s (60%); RP_e (62%) e $2RP_p$ (66,6%). Já para a disfluência do tipo $3RP_p$ a tendência é para o abaixamento da intensidade (66%). Por fim, para a RP_v os resultados equipararam-se para a repetição de sílaba dupla e para a repetição de sílaba simples, a intensidade aumentou em 100% dos casos.

Quanto à intensidade máxima, ocorre um aumento crescente, pouco significativo, para a disfluência do tipo $3RP_p$ (69,470 dB; 70,632 dB e 70,782 dB). Já para a RP_e e RP_v a tendência é de 75% e 80%, respectivamente, para o aumento na segunda repetição. A $2RP_p$, por sua vez, apresenta uma tendência para a diminuição (83%). Por fim, para a RP_s ocorre uma diminuição na intensidade para a repetição das duas sílabas iniciais da palavra e, para a repetição de sílaba inicial, um aumento. A diferença entre as intensidades da primeira para a segunda repetição foram pequenas em todos os eventos, sendo a diferença máxima de 6 dB.

A análise entoacional confirmou que os padrões entoacionais se mantêm para as RPs. Em relação às $2RP_{ps}$, em apenas um caso o padrão variou entre um tom simples e composto (*H+L> *H), mesmo assim o pico entoacional continuou no tom alto. Para a $3RP_p$ a primeira repetição apresentou um tom baixo (*L) simples, enquanto as outras foram anotadas com um tom simples alto (*H). Já nas repetições de sílaba, estrutura e vogal, não houve variação no padrão.

Na sequência aparece a COR, com 9,6% das produções. Dentre todas as CORs, CL manteve parte da informação errada em 41,6% dos casos. Além disso, esse evento foi o único que apresentou disfluências complexas, representando 25% do total de ocorrências. As CORs complexas ocorreram juntamente com PS_e, AL e RP. Veja abaixo:

5. “di te qui i lá *na na/ >* nu presídiu”;
6. “a- a > us ministu da eucarístia são movimentus#”;
7. “a genti tem qui traz e elis *pa # >* pra juntú lá né”.

Os eventos com erro e a sua respectiva correção aparecem em itálico. O símbolo maior do que (>) corresponde à correção feita pelo informante. As disfluências que ocorrem concomitantemente com a COR ajudam o informante a ganhar tempo para fazer a correção necessária, que diferentemente da COR simples, demora mais para ser processada. Observe que, no primeiro exemplo, o informante precisa fazer uma 2RP_p e uma PS respiratória para atingir à mensagem final. Já no segundo exemplo, CL precisa alongar e repetir a vogal [a] e, no último exemplo, é necessária apenas uma PS silenciosa.

A análise acústica revelou para 66,6% dos casos que, na COR complexa, a duração da correção é menor do que a do erro. Já na COR simples ocorreu o contrário, já que em 77,7% dos casos a duração aumentou na correção. Quanto à F0 média, em 66,6% dos eventos ocorre o aumento no momento da correção para a COR complexa. Por outro lado, na COR simples a tendência é para a diminuição da F0 média durante a correção. Quanto aos pulsos, na COR complexa a tendência é para que o esforço vocal seja menor na correção (também em 66,6%). Já para a COR simples ocorre o contrário, sendo o esforço vocal maior na correção (também em 77,7%).

A análise da intensidade mostrou, deixando de lado os momentos de PS e COR em início de GT (quatro ocorrências, no total), tanto para a COR simples e como para a complexa, uma tendência para a diminuição, seja na transição da fluência para o erro ou do erro para a COR. Essa tendência se mantém para a COR complexa em relação à intensidade final. Já para a COR simples os resultados equipararam-se, uma vez que em 50% dos dados a intensidade aumentou e, em 50%, diminuiu. Quanto à intensidade máxima, para a COR simples ela aumentou na correção (55,5%). Por outro lado, para a COR complexa ela diminuiu (66,6%).

Em relação à análise fonológica, os padrões são diferentes. Eles variaram entre simples e complexos na passagem do erro para a correção. Nessa transição, apenas três casos

mantiveram o mesmo tom. Os demais eventos apresentaram variações entre tons altos e baixos.

Os ALs, por sua vez, representam 4% das produções de CL, distribuídos apenas em disfluências simples. Esses eventos foram verificados para vogais em início -/é/ e /i/- e meio de GT -/é/-, e para palavras em início de sílaba /:muitas/. O AL_v foi o mais frequente (quatro eventos) e o AL_p, o menos (um evento).

A duração média dos AL_vs em início de GT foi de 0,951s e, em meio de GT de 0,992s. Já a F0 média foi maior para os ALs em início de GT. Quanto aos pulsos, a diferença foi de um entre as duas ocorrências (164 para o AL em início de GT e 163 em meio de GT). No que diz respeito à intensidade de transição, somente o AL em meio de GT pode ser analisado. Nesse caso, ocorreu um abaixamento na intensidade. Quanto à intensidade final, tanto em início quanto em meio de GT houve um aumento. Já em relação à intensidade máxima, a diferença foi pequena entre os AL_vs. Em início de GT a média foi de aproximadamente 77 dB e, em meio de GT, de 76 dB, aproximadamente.

Por sua vez, o AL_p apresentou uma duração total de 0,716s, enquanto a sílaba alongada durou 0,483s (mais de 50% do total do segmento). Quanto à F0 e pulsos, nesse segmento a frequência foi muito alta, que foi gerada, principalmente, devido a uma interferência na gravação. Mesmo assim, por meio de alguns ajustes, verificamos que ela é maior do que a F0 e os pulsos anotados para as vogais. Em relação à intensidade de transição foi verificado um abaixamento, bem como para a intensidade final. Por fim, a intensidade máxima foi de aproximadamente 83 dB, maior do que as dos ALs em vogal.

Diferentemente do que observamos para os outros informantes, a análise entoacional mostrou um padrão frequente para os AL_vs. Ambos foram anotados com um tom composto do tipo *H+L. O AL_p também apresentou um tom composto, no entanto, o pico entoacional foi no tom baixo, *L+H.

Após o AL, a disfluência menos frequente é a HES, com 1,6% do total de eventos. As duas ocorrências foram do tipo /ah/, sendo uma em início de GT e a outra em meio de GT, ambas depois de uma PS e no mesmo enunciado. Como a HES em meio de GT ocorre após uma PS, ela indica um novo GT, por isso pode ser considerada também como uma HES de início de enunciado. Observe as duas ocorrências:

- EN36 “/ ah sabi assim tipu uma # ah -b- batizo”.

Nos dois eventos, a função não é a de surpresa ou lembrança, assim como verificado para LE. CL as usa com a intenção de responder ao interlocutor por meio de uma variação entoacional maior (como em BR).

As duas ocorrências duram, respectivamente, 0,693s e 0,515s, com uma média de 0,604s. Em início de GT a HES é mais grave, com uma média de frequência de 99 Hz e, em meio de GT, ela é mais aguda, com uma média de 192 Hz. Os pulsos também indicam o esforço vocal maior para o primeiro tipo. Quanto à intensidade final, nos dois momentos ocorre uma diminuição. Já a intensidade máxima é maior para a HES em meio de GT (79 dB) do que em início (66 dB).

No que concerne à análise entoacional, a HES em início de GT apresenta um tom baixo simples *L, já que, como vimos, a informante apresentou uma qualidade de voz mais grave. Em meio de GT a variação entoacional é maior, sendo anotada por um tom composto com pico entoacional alto *H+L.

Finalmente, o BLO e o FC foram as disfluências com o menor número de ocorrências, aparecendo apenas 0,8% do total de eventos. A disfluência do primeiro tipo ocorreu em consoante oclusiva bilabial [b]. Junto com ela ocorreu uma PS de BLO do tipo 1 (articulatória). A duração da PS foi de 0,865s e não ultrapassou a duração do segmento, que foi de 1,108s. A F0 média apresentou uma entoação alta de 193 Hz. A intensidade final abaixou, quando comparada à inicial e a intensidade máxima foi de 83 dB. O padrão entoacional foi anotado com um tom composto do tipo *H+L.

No segundo tipo de disfluência (FC) há uma mudança completa na intenção discursiva. Veja no exemplo:

- “é foi um crescimentu *is* né, crescimentu na fê mesmu#”.

O FC é representado pela palavra em itálico */is/*. Observe que a intenção inicial de CL era dizer que o seu crescimento foi espiritual, mas ele muda o turno, deixando o primeiro GT apenas como uma afirmação “é foi um crescimento né”. A duração dessa disfluência é de 0,465s e não apresentou sonorização. A intensidade de transição aumentou na passagem da fluência para a disfluência, já a intensidade final diminuiu. A intensidade máxima, por sua vez, foi de 68dB. Finalmente, não foi possível analisar o padrão entoacional, já que o segmento não apresentou frequência.

4.2 Informantes com disfluência gaga

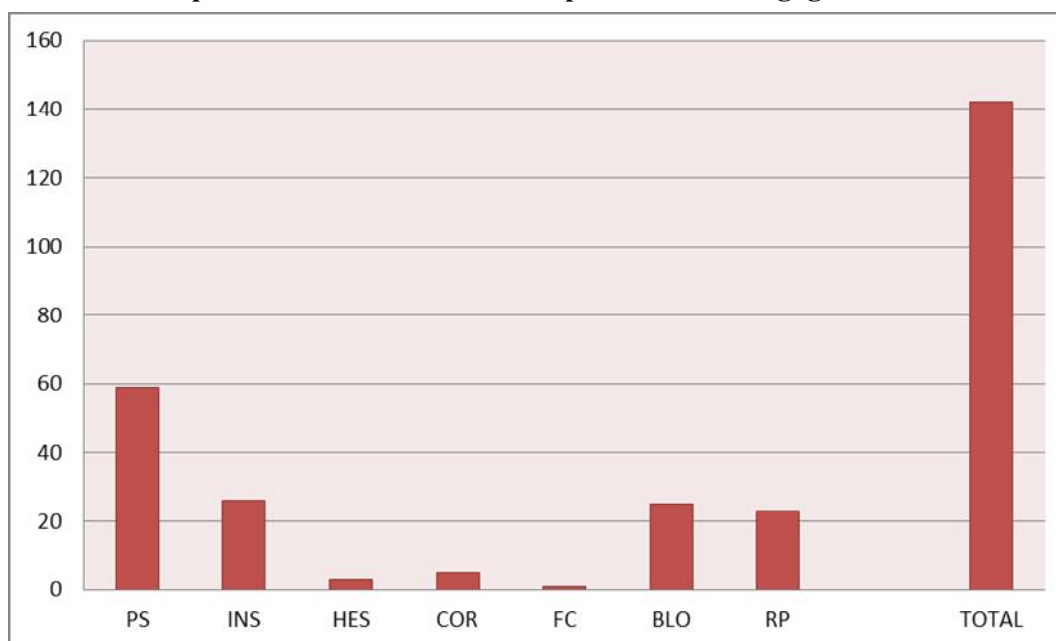
O informante PA

Em 38 enunciados, o informante gago PA realizou 142 disfluências. O máximo de disfluência observada por enunciado foi nove. Para PA as PSs foram as mais representativas, com 41,5% do total. Dentre elas, o informante tende a usar pausas simples (76,2%) e silenciosas (69,5%). A maior frequência de PSs para PA acontece em meio de enunciado em: 47,5% dos casos, aproximadamente. Para esse informante elas são indicativas de disfluência no enunciado.

Na maior parte das vezes elas são silenciosas (30,5%), e indicam a interrupção total do fluxo de fala. No geral, as PSs respiratórias (30,5%) atuam ajudando o informante a retomar o fluxo, mas também indicam ansiedade. As PSs em início de enunciado ocorrem em aproximadamente 39% dos casos e são, frequentemente, silenciosas (20%). Nesse caso, elas ajudam o informante a preparar a fala. As PSs em fim de enunciado representaram 22% dos casos, e a maioria é silenciosa (aproximadamente 17% dos casos), indicando que o enunciado chegou ao fim. A sua taxa de elocução foi de 5,084 sil/ms, a de articulação de 6,727sil/ms, e um tempo de articulação aproximado foi de 93ms.

A duração das pausas em início, meio e fim de enunciado foram, respectivamente: 5,152s, 10,761s e 7,498s. As pausas em meio de enunciado foram as maiores. Entre pausas silenciosas e respiratórias, o primeiro tipo foi o maior com uma duração de 20,352s.

O gráfico 10 a seguir mostra os tipos de disfluências que ocorreram ao longo da gravação. Na sequência, abordaremos de modo mais específico algumas das suas singularidades.

Gráfico 10 – Os tipos de disfluência realizados pelo informante gago PA.

Fonte: própria.

Depois das PSs é mais comum que PA realize INSs (18,3%). Na sequência, aparecem o BLO (17,6%), a RP (16,2%), a COR (3,5%), a HES (2,1%) e, por fim, o FC (0,7%).

Quanto à INS, a maioria ocorreu com palavras. Em um total de 26 ocorrências, elas representam 53,8% das produções. Elas variaram entre /sabi/ (mais frequente), /né/, /ué/ e /daí/. As INS_{ps} do tipo /sabi/ ocorreram sempre em fim de enunciado, bem como as do tipo /ué/. Para as do tipo /né/, verificamos duas em final de enunciado e uma em meio de enunciado. Por sua vez, as do tipo /daí/ foram encontradas em início (uma ocorrência) e meio de enunciado (uma ocorrência). A INS_p do tipo /sabi/ tem uma função interrogativa, que tem como justificar ao seu interlocutor o que foi dito. Veja dois exemplos:

- EN1 “qué a primera conversa nossa, daí eu to meu timidu, sabi”,
- EN5 “ah queu ficu conversanu muito é na sala, sabi”.

Já as do tipo /né/ (em meio e fim de enunciado), bem como /ué/, têm a função interrogativa retórica, assemelhando-se a “não é?” e “você não acha?”. Observe os exemplos:

- EN16 “i academia é bom tamém, né”;
- EN30 “purque fai quantu anu quela nu quela nu ganha –dua seleção / boa ué#

O marcador discursivo /daí/, por sua vez, nos passa a ideia de continuidade, tanto assim que ele ocorreu em início de enunciado, dando sequência ao argumento anterior, e em meio de enunciado. É um marcador discursivo que se assemelha ao /intão/ Veja a seguir:

- EN24 “daí a gente ia cantanu di lá até nu até nu campu”,

- EN25 “eu vo po uma música i né –daí eu cantei”.

Ainda em relação à inserção de palavras, PA produziu duas INSp,s complexas, uma com bloqueio, que está representada pelo enunciado EN25 (acima), e outra com uma vogal, do tipo /ishi i/. Neste caso (EN12), a INSp transmite a ideia de decepção diante de um acontecimento. Observe:

- EN12 “vai co vai vai i ishi i vai acontece muitas tragédia”.

Por fim, as INSp,s de vogais representaram 43,3% do total, sendo nove do tipo /i/ e duas do tipo /é/. As INSp,s de vogal complexas ocorreram apenas com /i/ (cf. EN12 e EN25). As INSp,s ocorreram em início de enunciado (duas produções) e em meio de enunciado (sete ocorrências). Já as do tipo /é/ ocorreram apenas em meio de enunciado. No contexto das produções, a INSp /i/ não tem uma finalidade coesiva ou sequenciadora. O informante a introduz no discurso com a intenção de evitar uma possível disfluência. Veja nos exemplos:

- EN18 “eu nu vo fala i qui não porque vai qui i qui aparece uma oportunidade boa”,
- EN28 “#porque eu queria# i fala muito #i# i dipressa#”.

Nos exemplos, as INSp,s /i/ aparecem em contextos propícios para que o informante produza um bloqueio, ou seja, antes de oclusivas /q/, /d/ e fricativas /f/, como veremos. Além disso, em EN28 ocorrem pausas antes e depois da INSp, o que representa outra tentativa de volta a fluência e controle da fala. Já a INSp do tipo /é/ foi verificada apenas em um enunciado. Apenas quando essa INS ocorre em início de enunciado observamos uma função de continuidade discursiva. Assim como para /i/, ela possui uma função estratégica no discurso, mas nesse caso, além de evitar momentos de BLO, ela ajuda o informante a sequenciar sua fala, assim como pode ser visto a seguir:

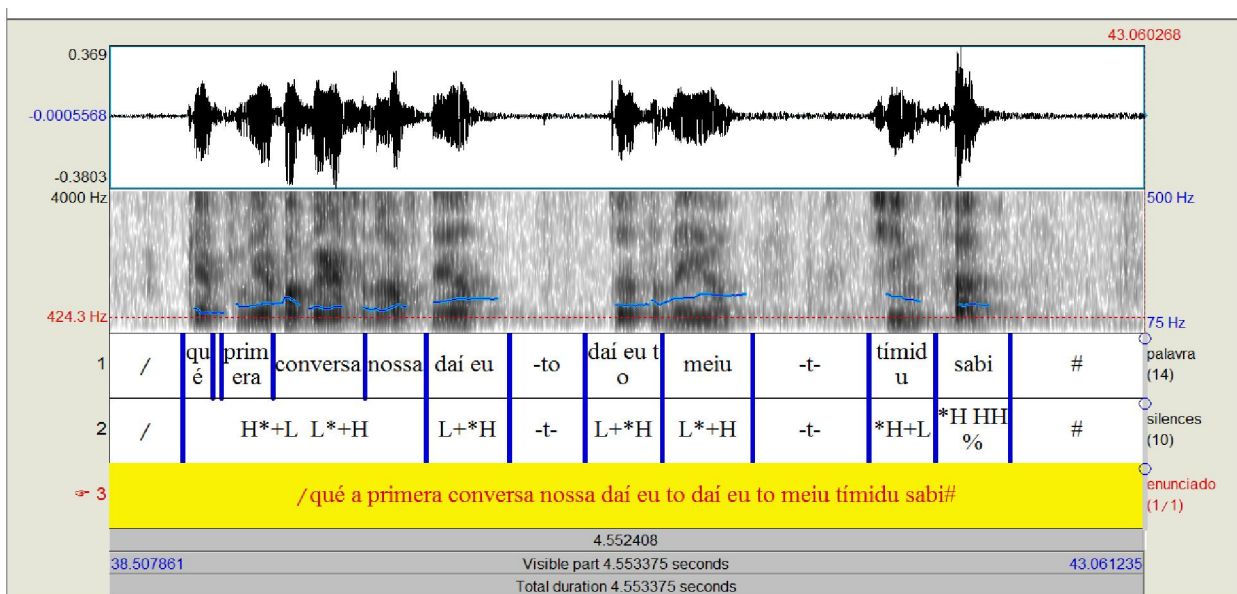
- EN5 “ah queu ficu é conversanu muito é na sala sabi”.

Quanto à análise fonológica, observamos que ela corroborou com o fato de que inserções do mesmo tipo manteriam o mesmo padrão entoacional. Sem exceção, para /sabi/, o padrão entoacional foi de *H. Para /né/, em final de enunciado, o padrão obtido foi de *H. Já em meio de enunciado foi de *H+L. Para /ué/ também não houve variação, mantendo o padrão *H+L. Para a INSp o padrão também se manteve. Veja a seguir os padrões obtidos em cada posição.

- /i/ em início e meio de enunciado => *H+L
- /é/ em meio de enunciado => *H

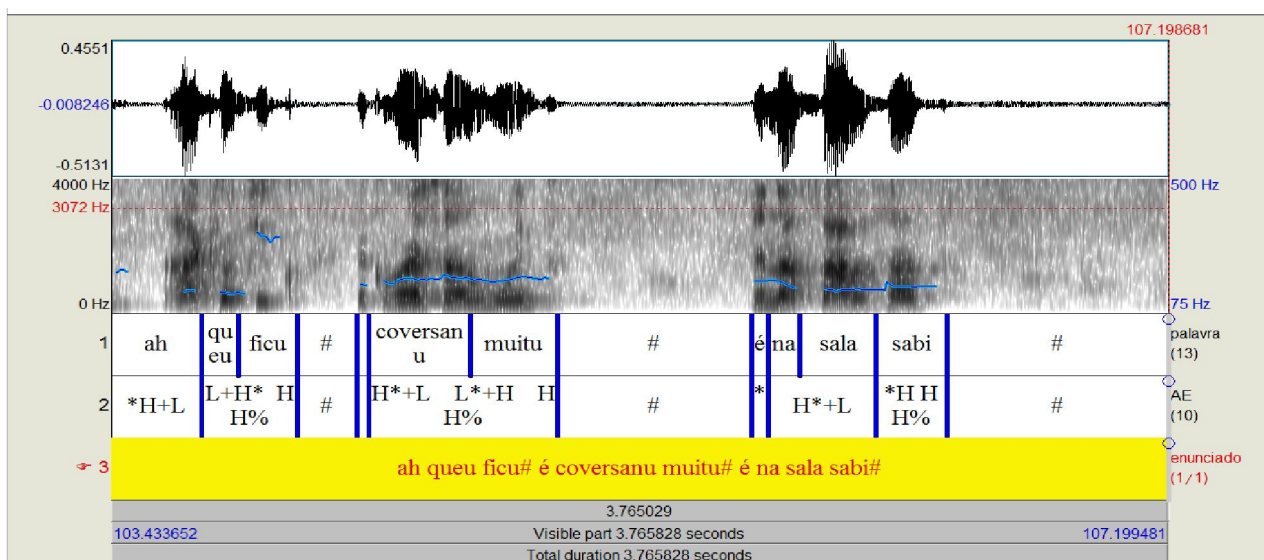
Abaixo apresentamos dois enunciados por meio das figuras 67 e 68 que exemplificam os resultados obtidos:

Figura 36 – Representação da curva entoacional para a INSp /sabi/ em final de frase no enunciado “/qué a primera conversa nossa daí eu to meu tímidu sabi”.



Fonte: própria.

Figura 37 – Representação da curva entoacional para a INSp /sabi/ em final de frase, para o enunciado EN5 “/ah queu ficu # é conversanu muito # é na sala sabi#”.



Fonte: própria.

As figuras 36 e 37 representam os enunciados EN1 e EN5. A curva da F0, comparativamente às outras palavras, mostra que a INSp do tipo /sabi/ é alta e não apresenta variação, por isso foi analisada com *H.

A análise acústica também mostrou uma tendência à padronização. Quanto à duração, a INSp do tipo /sabi/ variou entre 0,217s e 0,326s, com uma diferença de 0,109s. A média da F0 também não foi diferente. A diferença entre a maior e a menor foi de 22 HZ, bem como para a intensidade máxima. Em relação à transição da fluência para a disfluência, em todas as produções houve uma diminuição da intensidade.

A INS do tipo /né/ seguiu o mesmo padrão da INS anterior. No entanto, a duração foi menor, variando entre 0,101s e 0,161s. Quanto à intensidade, apenas em um enunciado houve um aumento durante a transição da fluência para a disfluência.

Para /ué/ a duração ficou na margem de /né/, variando entre 0,165s e 0,154s. Vale salientar que para EN7 e EN33 a duração foi a mesma para as duas produções (0,154s). A intensidade máxima também manteve uma mesma margem, com uma diferença pouco significativa. No entanto, para o momento de transição da intensidade, apenas um enunciado manteve a característica de abaixamento. A INSp do tipo /daí/ manteve os padrões duracionais de /sabi/. A intensidade máxima, por outro lado, foi próxima a das outras INSs, variando entre 76 dB e 79 dB. Diferentemente das outras INSps, a transição da intensidade não pode ser comparada, já que uma INSp ocorreu em início de enunciado e a outra após uma PS.

Em relação à INSp complexa, não é possível fazer comparações devido à diferença entre elas. Mas, quanto à transição da fluência para a disfluência, foi observado um aumento na intensidade, diferentemente das anteriores.

Para a INS de vogal, a do tipo /i/, em meio de enunciado, apresentou a maior duração (0,106s). A média das F0 também foi próxima. Dentre as ocorrências, duas foram surdas e uma apresentou uma qualidade de voz do tipo *creaky voice*. Quanto à transição da intensidade, não foi possível fazer comparações, pois esse tipo de disfluência ou ocorreu em início de enunciado ou depois de PS. Assim, a tendência é de que a intensidade comece mais alta. A intensidade também se manteve próxima, variando entre 62 dB e 74 dB. A INS do tipo /é/ também ocorreu depois de PS. A sua duração, média da F0 e intensidade máxima também se enquadram entre os padrões obtidos para /i/.

Quanto à intensidade final, para as INSps, em 100% dos casos ocorre uma diminuição. Já para as INSvs, em 77% dos casos há um aumento.

É importante salientar que há uma tendência para que a INS_v ocorra após uma pausa, pois, de 11 ocorrências, apenas em duas esse padrão não foi verificado, uma vez que ocorreram em início de enunciado.

O bloqueio, por sua vez, foi a segunda maior incidência de disfluência para PA. Os BLOs em consoante representaram 78,6% aproximadamente do total e o BLO em vogal, aproximadamente, 21,4%. Os bloqueios complexos representam, aproximadamente, 43% dos eventos. Sendo assim, tivemos:

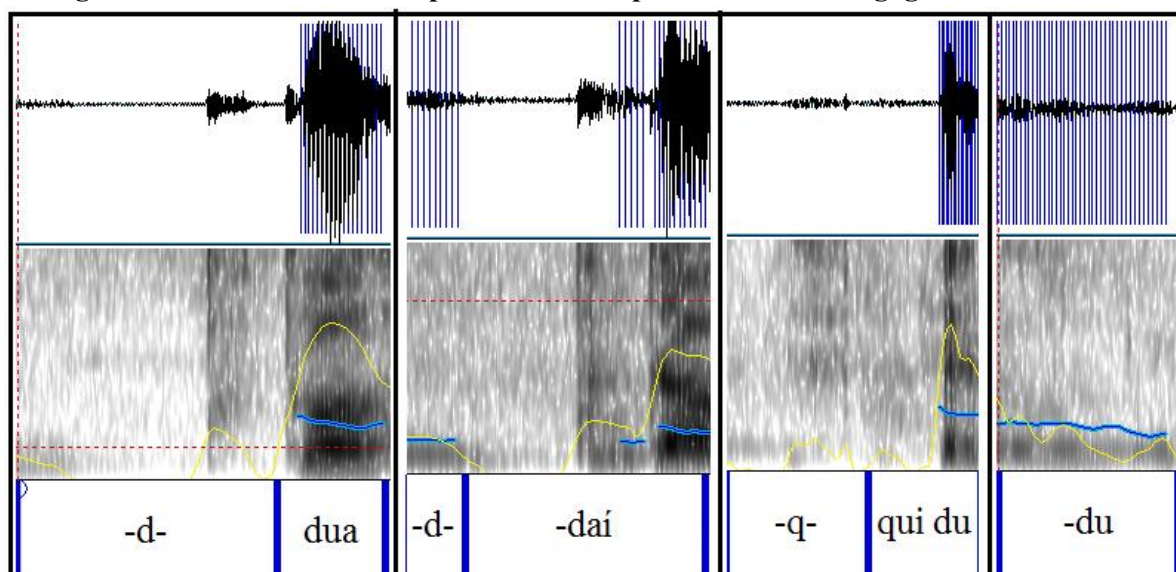
- 17 ocorrências de BLO em consoantes oclusivas [p], [t], [k], [d] como, por exemplo: /-palavra/, /-timidu/, /-quebradu/ e /-daí/;
- 4 ocorrências para fricativas [s], [v], [x] como, por exemplo: /-se/, /-viciadu/, /-reparadu/;
- 1 ocorrência para consoante nasal [n] como, por exemplo: /-nu/.

Para as vogais, PA realizou:

- 3 ocorrências em vogal anterior meio fechada [e] como por exemplo: /-eu/;
- 2 ocorrências em vogal anterior fechada [i] como, por exemplo: /-inroscada/;
- 1 ocorrência em vogal posterior fechada [u] como por exemplo: /-um/;
- 2 ocorrências em vogal anterior aberta [a] como, por exemplo: /a/.

Como havíamos proposto, os BLOs puderam ser identificados a partir de quatro pontos: 1) bloqueio com pausa articulatória (ou seja, na pausa anterior ao BLO há presença de pulsos e articulação); 2) bloqueio sem pulsos (nesse caso ocorre uma pausa silenciosa que pode ou não apresentar articulação); 3) bloqueio sem pausa anterior e 4) bloqueio com pausa respiratória. Esses exemplos podem ser verificados a partir da figura 42, retomada do capítulo referente à metodologia.

Figura 38 – Possibilidades de pausas de BLO para o informante gago PA.

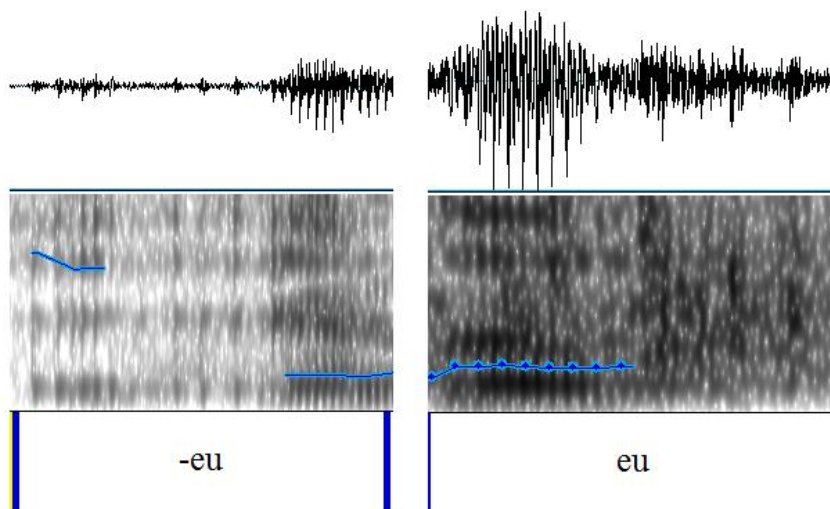


Fonte: própria.

Como vimos, na primeira imagem temos uma PS silenciosa, em que ocorre articulação da consoante /-d/. Na sequência, está representada uma PS articulatória, com presença de pulsos e F0 também para o segmento /-d/. Depois, temos uma PS respiratória para o segmento /-q/, sem presença de articulação. Por fim, aparece identificado um BLO em que não ocorre nenhum tipo de PS.

O bloqueio é um tipo de disfluência mais complexa de ser observada visualmente, devido às variações nas suas ocorrências. No entanto, há uma tendência geral para que, quando ele ocorra, mantenha alguma dessas características. Para o informante não gago, CR, foi verificado, ainda, interrupções na forma de onda em momentos de BLO. No entanto, esse padrão não foi identificado para o informante gago PA. Para esse falante, tanto nos BLOs em consoante quanto em vogal, a forma de onda no momento do BLO tende a redução, como se tivesse ocorrido um erro de gravação, e não há interrupções. Veja a diferença na forma de onda entre uma palavra sem bloqueio e outra com bloqueio:

Figura 39 – Comparação da forma de onda e do espectrograma em um segmento bloqueado (imagem esquerda) /-eu/ e em um sem bloqueio em momento fluente do informante gago PA.



Fonte: própria.

Na figura 39 temos duas produções do mesmo segmento para o mesmo informante. No entanto, a primeira produção ocorre em um momento de disfluência gaga e a segunda em um momento de fluência. Como podemos ver, a forma de onda é bem diferente nas duas produções. Na imagem da esquerda ela é muito reduzida, diferentemente da figura da direita. A curva da F0 já mostra que houve um erro na produção disfluente, que pode ter sido gerado ou por um ruído de gravação ou pelo próprio BLO. Contudo, descartamos essa última opção, uma vez que isso não ocorreu para mais de 50% dos dados. Além disso, o espectrograma nos mostra uma produção recorrente do mesmo segmento, nesse caso, da vogal [e], o que não acontece na segunda produção. Esse fenômeno de redução foi identificado para vogais e consoantes.

Em relação aos três momentos encontrados para a análise visual do BLO verificamos para esse informante que:

- Todos os casos de BLO em consoante apresentaram pausas articulatórias silenciosas sem pulsos;
- O BLO em vogal variou entre pausa articulatória sem e com pulsos e entre pausas articulatórias silenciosas;
- Os BLOs complexos variaram entre os quatro pontos mencionados. Mesmo assim, a tendência maior foi para pausas articulatórias sem pulsos.

A análise acústica do BLO_c revelou uma média de duração dos segmentos de 0,365s e das PSs de 0,344s. Já as médias das durações obtidas para o BLO_v foram de 0,248s para os segmentos; 0,210 para as PSs articulatórias e 0,434s para as PSs silenciosas. Por fim, para os BLO_cs, os resultados foram: 0,277s para os segmentos; 0,341 para as PSs articulatórias; 0,265s para as PSs silenciosas e 0,368s para a PS respiratória. Diante desses resultados, observamos que a duração dos segmentos em BLO simples são maiores do que a dos segmentos em BLO complexos e dos BLO em vogais. A média da PS articulatória também é maior.

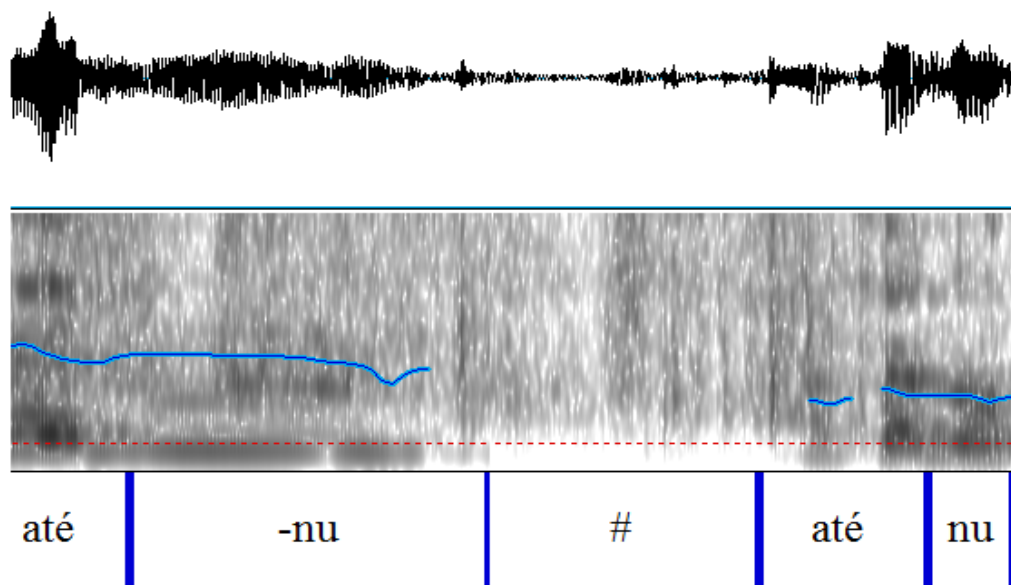
É importante lembrar que de 12 BLOs complexos tivemos apenas um BLO em vogal, que ocorreu para o mesmo enunciado e, em vogais repetidas. A média da duração dessas duas vogais bloqueadas foi de 0,346s. Também vale ressaltar que dos 12 BLOs complexos, dez ocorreram com palavras repetidas, sendo uma delas repetição de vogais, um com INS_p e um com COR. Quando o BLO ocorre em palavras, vogais ou estruturas repetidas, a tendência geral é que o informante faça o BLO na primeira palavra.

A média da F0 para os três tipos de BLO foram: 169 Hz para o BLO em consoante, 149 Hz para o BLO em vogal e 113 Hz para o BLO complexo. Esses resultados mostram que o informante apresenta uma entoação mais alta quando o BLO ocorre em consoante, e mais baixa quando ocorre em estruturas complexas. Se formos fazer uma aproximação entre esses dois momentos, o BLO em vogal está mais próximo de uma tessitura alta do que baixa. A análise dos pulsos, por sua vez, mostrou um padrão mais grave e esforço vocal menor para os BLOs em vogal e estrutura complexa, com uma média de 23 e 25 pulsos, respectivamente. Já o BLO em consoante é mais agudo e, conseqüentemente, apresenta um esforço vocal maior, com uma média de 36 pulsos.

Finalmente, em relação aos parâmetros de intensidade, não levamos em consideração a intensidade no momento de transição, uma vez que a maior parte dos casos foi antecedida por uma PS de BLO. Sendo assim, a média da intensidade máxima foi maior para o BLO em consoante (78 dB) e menor para as vogais e BLO complexos (73 dB e 72 dB, respectivamente). Ainda assim, a diferença entre a intensidade maior e menor não é significativa, variando entre 4 dB e 5 dB. Auditivamente, esses valores não são distintivos. Ao analisarmos os dados das intensidades individualmente, observamos que, no que se refere às consoantes, as oclusivas alveolares [d] e as fricativas alveolares são as que apresentam a maior intensidade, com 82,960 dB e 82,859 dB, respectivamente. As outras consoantes também apresentam uma intensidade mais alta, mas um pouco menor do que as anteriormente

mencionadas, variando entre 72 dB e 77 dB. Houve casos em que a intensidade nessas consoantes estava mais baixa, mas, isso ocorreu devido a uma produção surda de PA. As vogais mantiveram o mesmo padrão das consoantes, mas com valores um pouco menores, variando entre 70 dB e 81 dB. Isso mostra que a intensidade não é o melhor parâmetro para diferenciar bloqueios em oclusivas e vogais, já que as intensidades tendem a manter o mesmo padrão. Mesmo assim, as oclusivas e fricativas tendem a ter durações maiores do que o das vogais. No que se refere à intensidade no momento de transição da fluência para a disfluência, não é possível fazer comparações, uma vez que em apenas um enunciado o informante produziu um BLO com ausência de PS. Isso ocorreu, para esse informante, em um BLO complexo, em que há repetição de uma estrutura maior /até -nu/ /até nu/. Nesse exemplo, o informante faz uma pausa posterior, que ocorre antes da segunda RP_e. Veja na imagem 71 que segue:

Figura 40 – Análise do bloqueio em repetição de estrutura para PA no enunciado “daí a genti ia cantanu di lá até -nu até nu campu”.



Fonte: própria.

A fim de facilitar a visualização, apresentamos apenas o momento do BLO na repetição de estrutura em EN24. Como vemos, a PS silenciosa não aparece antes do BLO, mas, sim, depois. Veja que a duração é maior no bloqueio /-nu/. Além disso, também vemos uma redução na forma de onda, como mencionado e um aumento na F0 no momento que o

informante apresenta a dificuldade de soltura. A duração maior no bloqueio, bem como o aumento da F0, é comum para todos os casos em que ocorre repetição. O espectrograma também mostra essa dificuldade de soltura. Os formantes estão mais claros na fluência do que na disfluência.

Quanto à análise da intensidade final, comparativamente à inicial, tanto em vogais (66%) quanto em consoantes (68%), ocorre um aumento final.

A análise fonológica para esse tipo de disfluência mostrou um padrão para o BLO. Para o BLO simples em consoante, de oito resultados, seis foram analisados com o padrão *H+L, nos segmentos [t, d] e [k] e dois com o padrão *H, nos segmentos [s] e [p]. No BLO de estrutura, a inserção com BLO /-daí/, as RPs de segmentos [k], [v] e [d] também foram analisados com o padrão *H+L. Ainda no BLO de estrutura, tivemos um caso com o padrão entoacional *H, no segmento bloqueado [d]. Dois enunciados não foram mencionados porque foram produzidos sem sonoridade. As vogais também seguiram uma tendência à padronização. Os fonemas [e] foram classificados como *H e, os fonemas [i] com o padrão *H+L. Por fim, o fonema [u] foi classificado com o padrão *H. Veja que há uma tendência para que pontos de articulação iguais tenham o mesmo padrão entoacional. Em nossos dados, tivemos uma variação entre [t] e [d]. No entanto, o primeiro ocorreu em BLO simples e o segundo em BLO de estrutura. Ainda assim, o pico entoacional foi o mesmo *H.

Quanto às RPs, o informante produziu repetições simples e repetições de estrutura, de sílabas e palavras. Em um total de 23 RPs, 21,7% foram de repetições simples e 73,2% de repetições complexas. A presença maior de RPs complexas do que simples está relacionada ao fenômeno da empatia negativa que propusemos como hipótese neste trabalho. Também foram observados quatro casos de repetição de vogal complexa e de RP de sílaba simples.

As repetições complexas em palavras são as mais frequentes, seguidas da RP complexa em vogal e depois da RP de sílaba complexa. Com a mesma quantidade de produção seguem as RPs de estrutura e as RPs simples. Por fim, com menor produção temos a RP de sílaba simples. Assim, o informante realizou as seguintes repetições:

- 1 - 2RP_e simples > / é muito/ /é muito/;
- 4 - RP_v complexa, sendo:
 - 5RP_v + PS > / i/i i i i/;
 - 2RP_v + INS_p > / i ishi i/;
 - 2RP_v + PS > /i/i/;
 - 2RP_v + 2BLO_v > /-a -a/.
- 2 - RP_p simples, sendo:

3RP_p > /si si si/;
 4RP_{p2} > /daí daí daí daí/

- 11 - RP_p complexa, sendo:
 - 3 do tipo > 2RP_p+BLOc > /-qui qui/;
 - 1 do tipo > 2RP_{p4}+BLOc > /-viciadu viciadu/;
 - 1 do tipo > 5RP_p+BLOc > /da da da da -da/;
 - 3 do tipo > 2RP_p+PS > / eu /eu/;
 - 1 do tipo > 2RP_{p2}+PS > /uma /uma/;
 - 1 do tipo > 3RP_p+PS > /eu eu/ eu/;
 - 1 do tipo > 2RP_p+PS+COR > /fu/fu/queu/.

- 1 RP_{ps} simples > /ma mais/

- 3 RP_{ps} complexa, sendo:
 - RP_{ps2}+PS > /mi minha/;
 - RP_{psi3}+PS > /ci ciumenta/;
 - RP_{psi4}+BLOc > /-re -reparadu/.

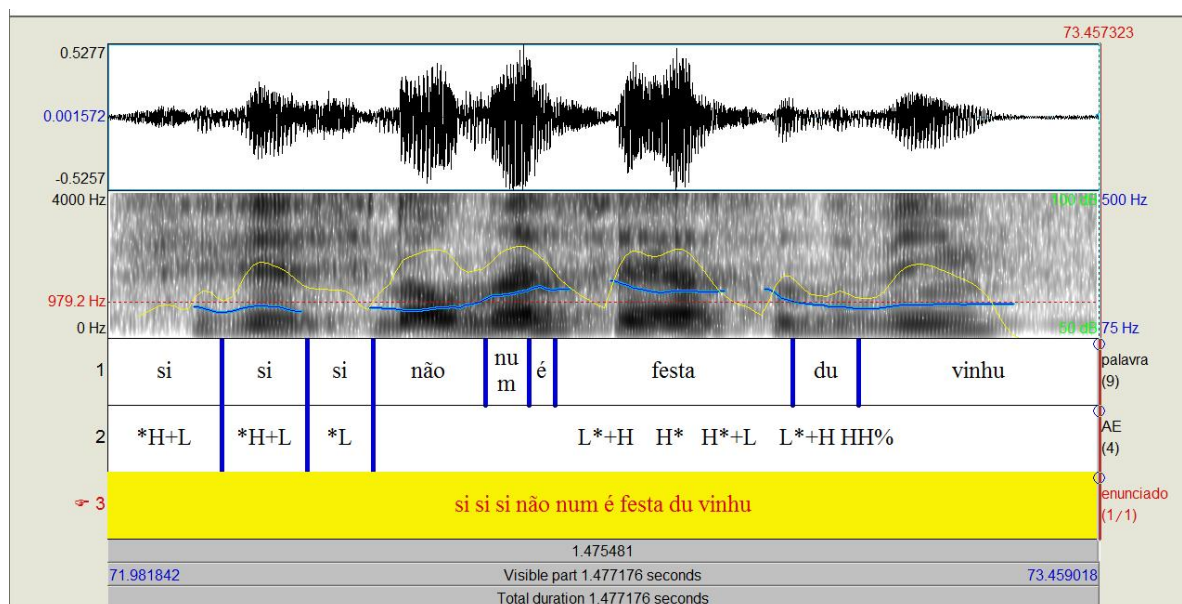
A análise acústica mostrou, para a RP_e simples, uma diferença duracional de 0,161s para EN10 e 0,086s para EN30. Para os dois exemplos, a duração menor está na busca pela volta à fluência. A média da F0 apresenta uma diferença maior entre as repetições em EN10, sendo de 61 Hz. Em relação à EN30 a diferença é de 15 Hz valor que, auditivamente, em relação às diferenças observadas é significativo. Mais uma vez, a maior diferença está na primeira palavra repetida. A análise dos pulsos revelou um padrão mais grave na volta da fluência. Por fim, a análise da intensidade máxima mostrou ser maior para a primeira RP, bem como a intensidade mínima. Já a análise da transição da intensidade nos dois momentos variou entre os enunciados. Por exemplo, em EN10, na primeira RP tivemos um pequeno aumento de 2 dB e, na segunda RP, um abaixamento. Por outro lado, em EN30, os dois momentos foram de abaixamento da intensidade.

Quanto à análise da RP_v complexa, observamos que o informante varia as durações na tentativa de voltar à fluência. Por exemplo, em EN6, enunciado em que o informante repete cinco vezes o mesmo segmento [i], a duração é de respectivamente 0,119s; 0,163s; 0,155s; 0,097s e 0,137s. Como vemos, PA começa com uma duração que já é mais alta para esse segmento. Depois, na tentativa de recuperar a fluência, faz uma pausa e aumenta a duração, mas ainda assim não consegue recuperá-la. Em seguida, diminui a duração mais uma vez. No entanto, PA não volta à fluência e, mais uma vez, aumenta intensidade, que fica maior do que a primeira. Supõe-se, assim, que para o informante, a duração ideal seria a última, nesse momento. EN12 mantém a diminuição de intensidade observada para RP_e. Contudo, ela diminui aproximadamente 50% da duração inicial: 0,130s > 0,065s. Para EN12, EN32 e

EN35 também ocorre o mesmo processo de diminuição da intensidade. A F0 varia em EN6, sendo maior para a última RP. É importante salientar que na quarta RP, o segmento é desvozeado. Nos outros exemplos não é possível estabelecer comparações, pois os segmentos repetidos não apresentam sonorização. Em EN12 a segunda RP é surda, em EN32 a primeira e, em EN35, ambas. Mesmo assim, a contagem dos pulsos mostra que a recuperação da fluência apresenta sempre um padrão mais grave do que nos períodos de tentativa de volta à fluência. Por fim, a análise da intensidade corroborou com o que foi observado para as RP_s, mostrando que a intensidade diminui na recuperação da fluência, apenas para EN35 houve um aumento de 3 dB. Mesmo em EN6, que apresentou um maior número de RP_s, a oscilação entre tentativa e perda de fluência mostrou um respectivo aumento e diminuição da intensidade. Dentre as quatro ocorrências obtidas para a RP_v, apenas EN12 não apresentou uma pausa na transição da fluência para a disfluência. Isso mostra que uma das estratégias desse informante durante a tentativa de volta a fluência é fazer um momento de pausa.

A RP_p, por sua vez, também seguiu os mesmos padrões dos tipos de RPs anteriores. No entanto, como essas foram RPs triplas (EN14) e quádruplas (EN23), observamos que a duração ocorre como uma forma de onda ou seja, a duração começa mais alta, diminui e depois aumenta, e assim sucessivamente, como verificado em EN6. Observe no exemplo 41 que segue:

Figura 41– Observação das durações para as 3RP_ps da palavra /si/ no enunciado EN14 “si si si não num é festa du vinhu”.



Fonte: própria

As durações para as RP_{ps} /si/, /si/ /si/ são, respectivamente, 0,164s; 0,127s e 0,097s. Elas mostram essa oscilação em forma de onda e a tendência para a diminuição na última RP_{ps}. A análise da F0 para EN14 também confirmou os resultados anteriores. No entanto, em EN23 observamos um aumento na F0, entretanto, foi pequeno (4 Hz), o que auditivamente não é significativo. Os pulsos, por sua vez, também confirmaram os dados anteriores. Finalmente, a análise da intensidade máxima não confirmou os resultados anteriores apenas para EN23. Ou seja, na última RP, que representa a recuperação da fluência, houve um aumento na intensidade em relação à RP anterior. Já na transição da fluência para a disfluência houve uma diminuição na intensidade.

A RP de sílaba simples é um caso diferente dos anteriores. Como foi mencionado, PA produziu apenas um exemplo desse tipo. Em relação à duração, mais uma vez os resultados foram os mesmos, a primeira RP_{ps} é maior do que a segunda. Contudo, esse era um caso esperado, pois na primeira produção, o informante aumenta a duração na tentativa de pronunciar a palavra completa. Já na segunda, que é a volta à fluência, ele diminui a duração do segmento para que ele se adeque à palavra. Quanto à média da F0, no que diz respeito apenas à comparação da sílaba incompleta /ma/ com a sílaba na palavra /mais/, a F0 é maior na sílaba da palavra, bem como a intensidade. Quanto aos pulsos, eles foram exatamente os mesmos: 13 e 13. Por fim, a transição da fluência para a disfluência mostrou um aumento da intensidade nos dois casos.

A análise da RP_{ps} complexa seguiu o padrão de diminuição da duração da RP_v simples. Por outro lado, a média das F0 corroborou os resultados obtidos para os outros tipos de RPs, isto é, uma diminuição da F0 na recuperação da fluência. Os pulsos também mostram uma tendência ao padrão mais grave na recuperação da fluência. Diferentemente do que ocorreu para a RP_s simples, os pulsos não foram iguais. É importante salientar que, na análise da palavra completa tanto para a RP_s simples quanto para a complexa, o segmento é mais agudo do que na repetição da sílaba. Por fim, a análise da intensidade média para as três ocorrências mostrou ser mais alta na sílaba da palavra completa em EN9 e EN15, assim como foi verificado para a RP_s simples, e mais baixa em EN36. Já a transição da intensidade pode ser observada apenas para EN9, uma vez que nos outros exemplos elas foram antecedidas por uma PS. Assim, em EN9 houve uma diminuição na sílaba incompleta e um aumento na transição da intensidade na sílaba da palavra completa.

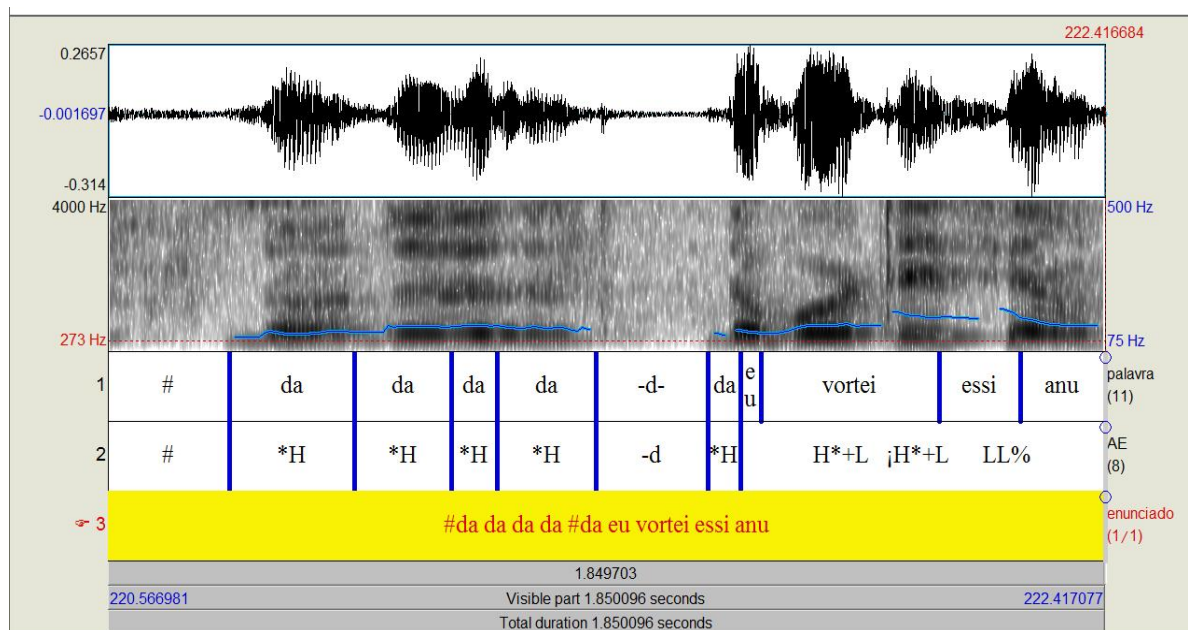
Finalmente, para as RPs complexas, observamos que, em relação à duração houve uma tendência unânime para a sua diminuição. No caso de EN11 e EN37, em que tivemos uma RP

quádrupla, o padrão de onda se manteve. A análise da média da F0 manteve o resultado de uma frequência mais alta na primeira palavra repetida. Em EN37 a média oscilou mais entre as repetições, e a volta a fluência apresentou uma diferença a mais de 3 Hz em relação à última RP, que representa a volta à fluência. A análise dos pulsos, de uma maneira geral, manteve os resultados anteriores, apenas EN11 destoou do padrão anterior. Finalmente, a análise da intensidade máxima mostrou que, para 63,6% dos dados, a tendência é para o abaixamento da intensidade na volta à fluência. Quanto à transição da intensidade, da fluência para a disfluência ou da disfluência para a disfluência, no caso de repetições de mais de duas palavras, das 26 RPs produzidas (somando todas as repetições obtidas nos enunciados), 34,6% mostraram um abaixamento, 19,2% resultaram no aumento da intensidade e 45,8% não foram analisadas, uma vez que foram antecedidas por pausas.

Quanto à análise fonológica, como tem sido observado para as RPs, há uma tendência geral para que elas mantenham o mesmo padrão⁶⁰. Quando há diferenciação, ela deve-se a casos em que o primeiro segmento é sonoro e o segundo é surdo, do contrário, elas se mantêm. Vale ressaltar que a maior diferença ocorre para as repetições de sílabas. Dentre as quatro RPs de sílaba produzidas por PA, nenhuma delas foi analisada com o mesmo padrão fonológico. Isso pode estar relacionado ao fato de, no primeiro segmento, termos apenas uma sílaba e, no outro, uma sílaba em segmento completo. Veja um exemplo em que o padrão entoacional é mantido:

⁶⁰ Como sabemos, a análise de Pierrehumbert (1980) e Beckham e Pierrehumbert (1986) é uma análise de base acústica. No entanto, como analisado em Carvalho (2013) ela não deixa de ter algumas bases auditivas. Sendo assim, é importante ressaltar que quando dizemos que há uma tendência para que a F0 seja mais alta nas repetições e, depois, afirmamos que a tendência geral é para que o padrão fonológico seja o mesmo, não estamos sendo contraditórios. Essa diferença que ocorre entre os dois eventos não é auditivamente (cf. Halliday, 1970; Cagliari, 2007) e nem acusticamente significativa para análise. Dessa maneira, não interfere nos resultados. O que ocorre, é que há diferenças acústicas entre um momento e outro, mas mesmo que essa diferença não seja significativa, é preciso que ela seja mencionada. Veja essa afirmação na figura 73, que identifica RPs.

Figura 42– Comparação da F0 ao longo das cinco repetições produzidas por PA no enunciado “#da da da da #da eu vortei essi anu”.



Fonte: própria.

Veja por meio da figura 42, que a F0 tem uma tendência a manter o mesmo padrão nas cinco repetições. Além disso, também aparece a PS de BLO, sem pulsos, identificada por -d-.

A COR não foi um dos eventos de disfluência mais frequentes para PA, bem como a HES e o FC. Em relação ao primeiro evento, bem como aconteceu para os informantes não gagos, PA optou por duas estratégias para desfazer o erro: 1) manter parte da informação do GT anterior na correção. Nesse caso tivemos quatro ocorrências em um total de cinco produções e 2) esquecer o GT anterior, criando um novo. Para essa estratégia foram verificadas apenas uma ocorrência. Veja um exemplo para cada situação:

- Situação do tipo 1): vai co > vai
- Situação do tipo 2): da > di

Em 80% das produções, as CORs ocorreram em início de GT e, entre elas, 20% foram antecedidas por uma PS inicial. Em relação aos outros enunciados, 60% apresentaram uma pausa antes da COR. No que diz respeito à análise acústica, verificamos para 60% dos dados uma tendência para que o informante aumente a duração no momento da correção. Já em relação à média da F0 durante o erro e a correção, em 100% dos casos PA aumentou a F0 no GT que traz a informação correta, ou seja, que mostra a recuperação da fluência, bem como analisado para a maior parte das RPs. Os pulsos, por sua vez, foram maiores durante a correção (80% dos casos). Esse aumento que há um esforço vocal maior e uma tessitura mais aguda, se comparada ao momento anterior. Quanto à intensidade, no momento de transição,

fica difícil fazer comparações, uma vez que, como vimos, PA inicia o GT com uma COR, que pode apresentar PS e, ao passar do erro para o evento corrigido, também opta por uma PS (60%). Como vimos, essa é uma estratégia utilizada por ele para a recuperação da fluência. Já no que se refere à intensidade máxima, em 80% dos casos o informante aumentou a intensidade no momento da COR, como ocorreu para a maior parte das RPs. O único enunciado que não seguiu esse padrão foi o que caracterizou a situação do tipo dois (àquela em que o falante muda os GTs). Quanto à intensidade final, em 70% dos casos a intensidade diminuiu.

Finalmente, no que diz respeito à análise entoacional, em 80% dos dados o informante muda o padrão entoacional. Esse fato já era esperado, pois há uma diferença na construção dos GTs. No entanto, o informante mantém o tom alto tanto no erro, quanto na correção em 40% dos casos, e o baixo, em 10% deles. No restante, o padrão se diferencia.

Foram realizadas três produções de HES. Todas elas ocorreram com a palavra /ah/ e em início de enunciado. No nível pragmático elas nos passam uma ideia de indecisão sobre o que vai ser dito, algumas vezes de embaraço e, ainda, funcionam como uma lembrança, ou seja, o informante lembrou-se da situação que gostaria de falar e usa a HES para enfatizar /ah.../. Veja os enunciados que apresentaram HES:

- EN4 - /ah tevi di -di professor/;
- EN5 - /ah queu ficu# é coversanu muito# é na sala sabi#/;
- EN16 - /ah i ac / i academia é bom também, né?/.

EN4 e EN5 identificam uma situação de embaraço. Já EN16 a lembrança positiva de uma informação.

Quanto à análise acústica da HES, as durações foram 0,184s (EN4); 0,239s (EN5) e 0,129s (EN16). A média da F0 também variou. Para EN4 ela foi de 103 Hz e para EN16 de 113 Hz. Já em EN16 o informante apresentou um estilo *creaky voice* que, por ser muito grave, não apresentou F0. A diferença entre as frequências de EN4 e EN16 é de aproximadamente 10 Hz, o que auditivamente e acusticamente não é significativo para o contexto. Os pulsos mostraram um padrão mais grave, com uma diferença de 1 Hz entre EN4 e EN16. EN5 não apresentou F0 e, conseqüentemente, não apresentou pulsos. Finalmente, a intensidade inicial não foi analisada, pois as HES ocorreram em início de GT. A intensidade máxima foi próxima nas três situações: EN4 69 dB; EN5 74 dB e EN16 74 dB. A diferença máxima entre elas foi de 5 dB, aproximadamente. Por fim, quanto à intensidade final, ocorreu um aumento em 66% dos casos.

Finalmente, a análise fonológica foi a mesma para EN4 e EN16 (*H+L). Em EN5 ela não foi anotada, pois não apresentou variação na F0.

O FC representou apenas um evento. Essa disfluência se caracteriza pela introdução uma palavra fora de ordem no enunciado. É como se a palavra não se encaixasse no enunciado, ficando vazia na produção. Observe:

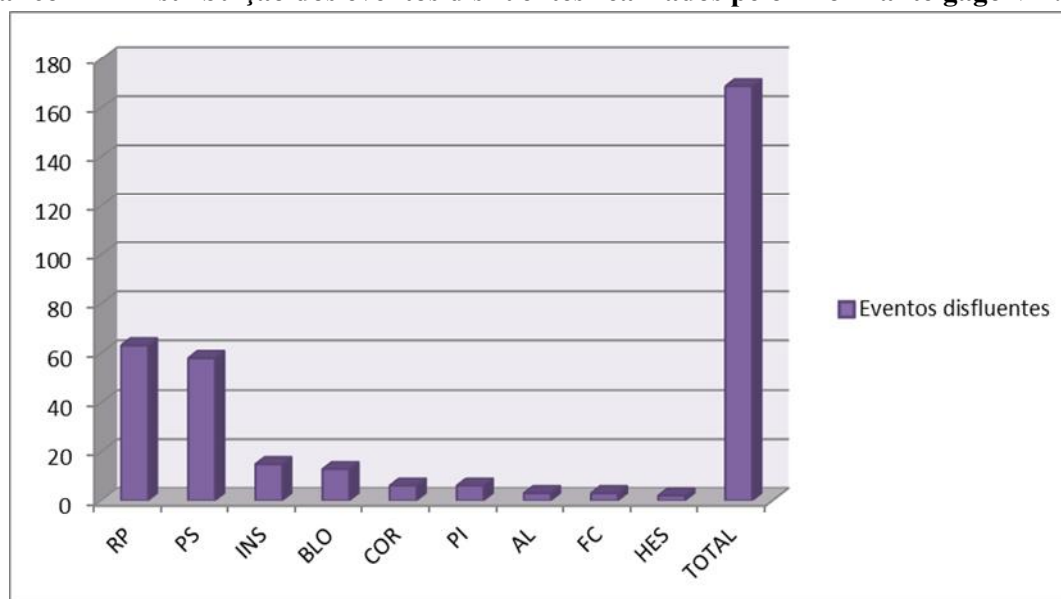
EN22 /# i daí eu consigo eu/ i daí eu fui tenta# anda né queu nunca tinha -quebradu nada#/.

Veja que PA começa o enunciado /i daí eu consigo/. Na sequência, ele deveria utilizar “i daí eu fui tenta [...]”, mas perde a sequência e introduz uma palavra fora do contexto /eu/. Esse fato pode ter ocorrido na tentativa do informante evitar a INS. No entanto, como ocorreu em apenas um evento, descartamos essa opção. Essa disfluência é diferente da COR, porque nela há uma sequência discursiva e no FC não. Quanto à análise acústica, esse evento ocorreu em meio de enunciado. A sua duração foi de 0,154s. Uma duração que, em relação aos outros segmentos sem disfluência, pode ser considerada grande. Por exemplo, no mesmo enunciado, a duração da palavra /eu/ sem disfluência que ocorre antes do FC é de 0,079s e a do segundo /eu/, que ocorre seguido de /i daí/, a duração é de 0,041s. A média da F0 e a intensidade máxima, comparativamente às outras produções da mesma palavra, são relativamente menores. O padrão fonológico para esse evento foi *H.

O informante VA

O informante gago VA realizou 169 disfluências em 66 enunciados, que também variaram entre simples e complexas. O segundo tipo de disfluências complexas representou 6,5% dos dados e o restante (93,5%) foram de disfluências simples. O máximo de disfluências por GT foi de seis, não sendo nenhuma do tipo PS. VA apresentou uma taxa de elocução de 5,211sil/ms, de articulação de 6,308sil/ms, e um tempo de articulação de 146ms. Veja no gráfico 11 a distribuição dos eventos disfluentes na fala de VA.

Gráfico 11 – Distribuição dos eventos disfluentes realizados pelo informante gago VA.



Fonte: própria.

Como vemos, VA varia a sua fala entre nove tipos de disfluências, dispostas da seguinte forma: RPs, PSs; INSs; BLOs; CORs; PIs; ALs; FCs e HESs.

As RPs foram as mais frequentes para esse informante, representando 38%, aproximadamente dos eventos. Os tipos de RPs verificadas foram:

- RP_{ps} (44,4%);
- 2RP_p (30,2%);
- RP_c (9,5%);
- RP_e (7,9%);
- RP_v (3,2%);
- 3RP_v (1,6%);
- 3RP_p (1,6%);
- 4RP_p (1,6%).

Por meio das porcentagens, que aparecem ao lado de cada ocorrência de cada RP, as repetições de palavras foram as mais frequentes em VA, com 96,8% dos casos, enquanto que as RPs em vogal representaram apenas 3,2%. É importante destacar que, entre as repetições duplas de palavras (2RP_p), apenas um caso ocorreu em palavras dissílabas. Já entre as repetições de sílaba, três foram do tipo RP_{ps2}, ou seja, repetição de sílaba dupla como, por exemplo, / atra atravessa/, e três foram do tipo 2RP_{ps}, ou seja, repetição dupla de sílaba como, por exemplo, /le le levo/. Os segmentos desvozeados, ou com uma qualidade de voz do tipo *creaky voice*, foram identificados tanto na primeira quanto na última RP.

No que diz respeito à análise acústica da duração, observamos as seguintes tendências:

- Para as repetições do tipo $2RP_p$ (62,2%), RP_{ps} (82,2%) e RP_c (66,6%) há uma diminuição na duração na passagem para a segunda RP. Vale salientar que as RP_{ps2} mantiveram a mesma tendência. No entanto, as $2RP_{ps}$ apresentaram uma tendência para o padrão do tipo onda em 66,6% dos casos;
- As RP_{es} , por sua vez, apresentaram uma tendência para o aumento da duração na segunda estrutura em 60% dos casos;
- As $3RP_{vs}$ e $3RP_{ps}$ também apresentaram um padrão do estilo onda, que iniciou-se com uma duração mais alta;
- A repetição do tipo $4RP_p$, por outro lado, manteve o padrão das repetições triplas até a terceira repetição. Já na quarta, a duração seguiu a tendência da última repetição, ou seja, aumentou;
- Por fim, para a RP_v não foi possível estabelecer uma tendência, uma vez que, em 50% dos dados, a duração aumentou e, em 50% diminuiu.
- No que se refere à F0 média e aos pulsos os resultados foram:
 - Para a repetição do tipo $3RP_p$ não foi possível propor um padrão, já que as três palavras repetidas foram surdas;
 - Para a $3RP_v$ ocorreu um aumento crescente da F0 média e pulsos;
 - Para a $4RP_p$, a segunda e a terceira RP foi influenciada por um erro de gravação, uma vez que VA produziu uma qualidade de voz muito aguda. Já em relação à primeira e à última RP, a F0 média foi menor na última repetição. Os pulsos mantiveram a mesma característica da F0;
 - As repetições do tipo $2RP_p$ (73,7%); RP_{ps} (77,7%); RP_v (100%); RP_e (60%) e RP_c (100%) apresentaram uma tendência para a diminuição da F0 média. A RP_{ps2} manteve essa tendência em 66,6% dos casos. Já as $2RP_{ps}$ apresentaram o padrão de onda, começando com uma F0 mais alta e acabando com uma F0 mais baixa. Os pulsos não mantiveram a mesma tendência apenas para as repetições do tipo $2RP_p$ e RP_v , que variou entre um maior esforço vocal (50%) e menor esforço vocal (50%).

Por fim, quanto à intensidade de transição, final e inicial foram verificados os seguintes padrões:

- Para as repetições do tipo $3RP_p$; $3RP_v$ e $4RP_p$ a intensidade no momento de transição tanto da fluência para a disfluência quanto da disfluência para a fluência, aumentou, respectivamente, em 66,6%; 66,6% e 75% dos casos. Quanto à intensidade final,

comparativamente à inicial, a intensidade diminuiu em 66,6% dos casos para as repetições do tipo $3RP_p$ e $3RP_v$. Já para a repetição do tipo $4RP_p$, ocorreu uma diminuição da intensidade em 75% dos casos. Por fim, a intensidade máxima revelou uma tendência de aumento crescente a $3RP_p$ e uma tendência para a queda para a $3RP_v$. Para a $4RP_p$ há uma tendência à queda até a terceira RP, mas na última RP (transição da terceira para a quarta repetição) a duração aumenta;

- Para a RP do tipo $2RP_p$, tanto da passagem da fluência para a disfluência (64,3%) quanto da transição da disfluência para a disfluência (68,4%), a intensidade diminuiu. Esse padrão manteve-se para a intensidade final (65,8%) e para a intensidade máxima (73,8%). A diminuição nos três momentos de intensidade também foi verificado em 100% dos casos para a RP_v ;

- No que diz respeito à RP_e , a intensidade também diminuiu na transição da fluência para a disfluência (66,6%) e aumentou na passagem para a disfluência (66,6%). Quanto à intensidade final, os resultados equipararam-se (em 50% diminuiu e em 50% aumentou). Por sua vez, a intensidade máxima mostrou uma tendência para a diminuição da intensidade na segunda estrutura em 60% dos casos;

- Quanto à RP_c , os resultados equipararam-se com a intensidade no momento de transição da fluência para a disfluência. Já para a transição da disfluência para a disfluência, a intensidade aumentou em 71,4% dos casos. Para a intensidade máxima os resultados mantiveram-se iguais;

- Por fim, a RP_{ps} apresentou um abaixamento da intensidade nos dois momentos de transição, sendo de 84,2% na passagem da fluência para a disfluência e de 79%, aproximadamente, na transição entre as repetições. As intensidades final (63,3%) e máxima (61%), por outro lado, apresentaram uma tendência para o aumento da intensidade. Por fim, no que diz respeito às variações da RP_{ps} , observamos, para a RP_{ps2} , uma diminuição em 66,6% dos casos na passagem da fluência para a disfluência e um aumento de 66,6% entre as transições. A intensidade final também aumentou em 66,6% dos casos. Já a intensidade máxima diminuiu em 66,6% dos casos. Por outro lado, a $2RP_{ps}$ caracterizou-se por uma diminuição da intensidade nos dois momentos de transição, sendo de 66,6% na passagem da fluência para a disfluência e de 100% entre as repetições. Finalmente, em relação à intensidade final e máxima, a intensidade aumentou em 77,7% dos casos para o primeiro tipo e apresentou um padrão de aumento crescente em 66,6% dos casos para o segundo tipo.

Por fim, a análise entoacional manteve o mesmo padrão para a maior parte das repetições, como pode ser observado na distribuição a seguir:

- Para as repetições do tipo $3RP_p$, $3RP_v$ e $4RP_p$, o padrão se manteve apenas para o segundo tipo. Para a $3RP_p$, a primeira repetição manteve um padrão alto (*H+L) e as duas últimas apresentaram uma diminuição (*L+H). Já a $4RP$ variou nas quatro ocorrências. A primeira repetição iniciou-se com um tom composto (*L+H). Na segunda ocorreu um erro, gerado por uma frequência mais aguda. Na terceira o padrão manteve-se baixo, mas com tom simples (*L). No entanto, na última repetição a entoação aumentou e continuou com um tom simples (*H);

- Para as repetições do tipo $2RP_p$ o padrão se manteve em 86,6% dos casos. Os casos em que a repetição não se manteve foram resultado de repetições sem sonorização na segunda repetição;

- As RP_{vs} mantiveram o mesmo padrão em 100% dos casos;

- As repetições complexas mantiveram o mesmo padrão em 66,6% dos casos;

- As RP_{ps} e suas variações também apresentaram uma tendência para que as repetições mantivessem o mesmo padrão entoacional. Sendo assim, para a RP_{ps} a tendência foi de 68,2% dos casos, aproximadamente. Para a RP_{ps2} o padrão manteve-se em 66,6% dos casos. E, por fim, para a $2RP_{ps}$ o padrão manteve-se em 100% dos casos;

- As repetições de estrutura foram o único tipo de repetição que não manteve o mesmo padrão entoacional para a maior parte dos dados, uma vez que o padrão continuou em apenas 20% dos casos.

Seguidas das repetições, as PSs representam o segundo tipo de disfluência mais frequente para VA, com 35%, aproximadamente, dos eventos. Dentre elas, a mais comum é a de final de GT (57%, aproximadamente), seguidas das pausas de início de GT (38%, aproximadamente) e, como as menos frequentes aparecem as pausas em meio de enunciado (5%, aproximadamente). No total, 31%, aproximadamente, das PSs, antecederam eventos disfluentes.

Quanto à duração, as PSs em fim de enunciado também foram as que apresentaram a maior duração (14,965s). Na sequência, temos as PSs em início de GT (10,997s) e, por fim, as que ocorrem em meio de GT (1,294s). Entre PSs silenciosas e respiratórias, aquelas apresentaram a maior duração (15,797s), enquanto estas apresentaram uma duração total de 11,459s.

Depois das PSs, as INSs aparecem com 9% dos eventos disfluentes. Elas dividem-se em INS_{ps} (60%) e INS_{vs} (40%). As INSs realizadas por VA estão assim distribuídas:

- INS_p em início de GT - /assim/, /tão/, /ái/;
- INS_p em fim de GT - /né/, /ta/;
- INS_p em meio de GT - /ái/;
- INS_v em início de GT - /é/, /i/;
- INS_v em meio de GT - /é/, /i/;

A INSs do tipo /tão/, assim como ocorreu tanto para os informantes gogos quanto não gogos, é uma variação da INS_p do tipo /intão/. Em início de GT, a sua função é a de sequenciar o enunciado, bem como /ái/. Já a do tipo /assim/ funciona como uma exemplificação. Observe no exemplo abaixo:

- EN4 “*assim*, i a genti era muito amigu”.

Observe, por meio do exemplo, que VA usa o segmento /assim/ como um recurso de exemplificação para o interlocutor ao tratar da sua relação com uma determinada pessoa. Também podemos dizer que essa palavra funciona como um elemento de coesão para a sequência discursiva.

Os segmentos /i/ e /é/ em início e meio de GT mantêm a ideia de sequência discursiva, assim como a INS_p /ái/ em meio de GT. A INS_p em final de enunciado, do tipo /né/, mantêm a função semântica de afirmação retórica. Por outro lado, a INS_p /tá/, que também ocorre em fim de enunciado, foi utilizada apenas com função afirmativa. Veja abaixo:

- EN35 “#nu fundu eu nu gostava # ta/.

No exemplo, VA discorre sobre o olhar do outro em relação à gagueira. O informante inicia o GT com uma PS silenciosa e, antes da INS, repete a pausa para depois afirmar o que havia sido dito, sem grandes variações entoacionais.

No que concerne à análise acústica das INSs, foi observado uma duração média maior para as INS_p em final de GT (0,262s). Em seguida, temos a INS_p em meio de GT (0,228s); as INS_{vs} em meio de GT (0,140s); as INS_{ps} em início de GT (0,135s) e, por fim, as INS_{vs} em início de GT, com uma duração média de 0,083s. Em relação à F0 média, as INS_{vs} e a INS_p em meio de GT não foram observadas, pois elas foram realizadas com uma frequência muito aguda, que gerou um erro de gravação. Dentre as inserções restantes, as INS_{vs} em início de GT apresentaram a maior frequência (155 Hz). Depois, aparecem as INS_{ps} em fim de GT (150 Hz) e, por fim, as INS_{ps} em início de GT (140 Hz). Quanto aos pulsos, o esforço vocal é

observado para as INS_{ps} em fim de GT e, na sequência, aparecem as INS_{ps} e as INS_{vs} em início de GT.

A análise da transição da intensidade não foi observada para as INS_{ps} e INS_{vs} em início de GT, uma vez que iniciaram o GT e, algumas, foram antecedidas por PSs. Sendo assim, no momento de transição, a intensidade aumentou na passagem da fluência para a disfluência para a INS_p em meio de enunciado e para as INS_{vs} em meio de GT. Por sua vez, para as INS_{ps} em fim de GT a intensidade diminuiu em 66,6% dos casos. A intensidade final, quando comparada à inicial, revelou que a intensidade diminuiu para as INS_{vs} em início de GT em 66,6% dos casos e para as INS_{ps} em início e fim de GT em 75% dos casos. Por outro lado, a intensidade aumentou para as INS_{vs} em meio de GT em 66,6% dos casos, e para a INS_p em meio de GT. Finalmente, quanto à intensidade máxima, a média dos resultados mostrou que a maior intensidade ocorre para a INS_p em meio de GT (86 dB). Em seguida, temos as INS_{ps} em início de GT (81 dB) e as INS_{ps} em fim de GT (79 dB). Por último, aparecem as INS_{vs} em início e meio de GT com uma intensidade (média) máxima praticamente igual, sendo de 77,105 dB para a primeira e de 77,115 dB para a segunda.

A análise dos padrões fonológicos confirmou a tendência das $INSs$ preservarem o mesmo padrão entoacional. Dessa maneira, os padrões obtidos para as inserções foram:

- /tão/ e /ái/ em início de GT - *L+H para 100% dos casos;
- /né/ e /ta/ em final de GT - *L+H para /ta/ e *L+H para né em 66,6%. O segmento que não apresentou o mesmo padrão foi analisado com o tom baixo *L;
- /ái/ em meio de GT - *H (caso único);
- /é/ e /i/ em meio de GT - *H para 100% dos casos;
- /é/ e /i/ em início de GT – foram os que mais variaram, uma vez que, entre as duas realizações do primeiro segmento, uma delas foi analisada com tom *H+L e a outra não apresentou sonorização. Já o segmento /i/ apresentou o tom simples *H.

Na sequência, a disfluência do tipo BLO aparece em 9% dos eventos disfluentes. Os BLOs foram verificados em consoantes e vogais, sendo o mais frequente em consoante (84,6%). Em vogais, eles ocorreram apenas para a vogal anterior meio aberta [ɛ]. Já para as consoantes eles foram verificados para:

- a nasal alveolar [n] – duas realizações no segmento /-nu/;
- as oclusivas velares [k] e [g], respectivamente nos segmentos /-gagueja/, /-queli/ e /conversanu/;
- as oclusivas bilabiais [p] e [b], respectivamente nos segmentos /-pra/ e /-bobu/;

- a oclusiva alveolar [t], no segmento /-tranquilu/;
- as fricativas labiodentais [f] e [v], respectivamente nos segmentos /-falanu/ e /-vai/;
- a fricativa alveolar [s] no segmento /-sempri/.

A relação acima mostra que é mais comum para VA realizar bloqueios em consoantes oclusivas. Quanto ao tipo de PS de bloqueio, o BLO_v apresentou, para 100% dos dados, uma PS articulatória (tipo 1). Já o BLO_c dividiu-se em PSs articulatórias (45,5%), em PSs silenciosas sem articulação (27,3%) e com articulação (9%) e BLOs sem pausa (18,2%).

Em relação à análise acústica, a duração das PSs de BLO foi maior do que os BLO_vs em 100% dos casos. Já para as consoantes, em relação aos segmentos que apresentaram algum tipo de PS, a tendência para 55,5% dos casos foi para que o segmento apresentasse uma duração maior do que as pausas de bloqueio. A média das pausas de bloqueio para as vogais foi de 0,869s, enquanto a do segmento foi de 0,183s. Por outro lado, a média das pausas em consoante foi de 0,609s. Já a média dos segmentos foi de 0,449s. Separando os bloqueios em relação ao modo de articulação, verificamos que as consoantes nasais foram as que apresentaram a maior duração (0,579s). Em seguida temos as consoantes oclusivas (0,425s), as consoantes fricativas e, por fim, as vogais (183s).

Quanto à F0 média, verificamos que a diferença entre as frequências é praticamente mínima, principalmente entre as consoantes oclusivas (170,365 Hz) e fricativas (170,340 Hz). As consoantes nasais variaram um pouco mais do que as anteriores, com uma frequência média de (174,600 Hz). As vogais, por sua vez, apresentaram uma F0 média de 153 Hz, aproximadamente. Em relação aos pulsos, o maior esforço vocal ocorre para as oclusivas. Em seguida, ocorrem as nasais e as fricativas e, por fim, as vogais, com o menor esforço vocal.

Em relação à intensidade de transição, somente dois segmentos /-nu/ e /-conversanu/ não foram antecidos por PS de bloqueio. Nesse caso, ambos apresentaram um abaixamento na intensidade. Em relação à intensidade final, não houve diferença significativa entre os modos de articulação, sendo que em 91% dos casos a intensidade reduziu. Os BLO_vs também apresentaram a mesma tendência, mas, nesse caso, em 100% dos dados. Por fim, a média da intensidade máxima também revelou uma diferença pequena entre os segmentos, sejam eles em vogais ou consonantes. Sendo assim, as oclusivas apresentaram uma intensidade aproximada de 85 dB e as fricativas de 82 dB. Já as consoantes nasais e as vogais foram os segmentos que mais se aproximaram com uma intensidade de 80,556 dB para o primeiro caso e 80,971 dB para o segundo caso.

Quanto aos padrões entoacionais, os segmentos do tipo nasal alveolar apresentaram um padrão que variou entre *H+L e *L. Já as consoantes oclusivas variaram entre *H+L e *H, assim como foi verificado para as fricativas. Em relação às vogais, um dos segmentos não apresentou sonorização e outro foi anotado com o padrão *H+L. Como vimos para os outros informantes, o padrão entoacional dos bloqueios tende a ser variado. No entanto, para VA, mesmo que os tons variem entre simples e compostos, a tendência é para que as oclusivas e fricativas mantenham um pico entoacional alto.

As CORs caracterizaram 3,5% do total de disfluências. Elas dividiram-se entre eventos simples (83,3%) e complexos (16,7%). As CORs complexas ocorreram junto com PS_e (dois casos), PI (três casos)⁶¹ e RP_e (um caso). Dentre as CORs, o informante manteve parte da informação contida no erro durante a correção em apenas 20% dos casos.

A análise acústica revelou, para a COR simples (um caso), uma diminuição da duração no momento da correção. Já para a COR complexa, a tendência foi para um aumento da duração em 60% dos casos. Quanto à F0 média, a tendência foi para o aumento da frequência no momento da correção para as CORs simples e complexas, sendo a tendência de 60% para o último tipo. No que diz respeito aos pulsos, na COR simples o esforço vocal é menor na correção. Já na COR complexa o esforço vocal foi maior na correção para 100% dos casos.

A análise da intensidade revelou, deixando de lado os momentos de PS e COR em início de GT, um aumento para a COR simples. Já a COR complexa apresentou um aumento da intensidade na passagem da fluência para a disfluência em 100% dos casos. Por outro lado, na transição do erro para a correção, a tendência foi para a diminuição da intensidade em 66,6% dos casos. Quanto à intensidade final, seja para a COR simples ou para a COR complexa, a tendência foi para o aumento da intensidade, com uma porcentagem de 80% dos casos para a COR complexa. Por fim, a análise da intensidade máxima foi diferente para a COR simples e COR complexa. No primeiro caso, a intensidade máxima diminuiu. No segundo caso, ela aumentou na correção em 100% dos casos.

Em relação à análise fonológica, os padrões foram diferentes na passagem do erro para a correção. Eles variaram entre tons simples e compostos na passagem de um momento para o outro. Durante essa transição, os padrões mantiveram-se o tom alto no erro e na correção em 60% dos casos. Os demais eventos (40%) variaram entre tons altos e baixos na passagem do erro para a correção.

⁶¹ Vale salientar que uma dessas PIs ocorreu juntamente com PS. Em outras palavras, uma mesma disfluência apresentou, concomitantemente, uma pausa e uma palavra incompleta.

A disfluência do tipo PI também foi anotada para 3,5% dos eventos disfluentes. Dentre os eventos, apenas dois deles apresentaram um erro de gravação, devido a uma qualidade de voz mais aguda. A duração média desses eventos foi de 0,314s. A média da F0 média, por sua vez, mostrou uma frequência entoacional de 156 Hz. Os pulsos variaram entre um segmento e outro, por isso, não foi possível estabelecer padrões. Quanto à intensidade, a tendência foi para a redução tanto em relação ao momento de transição (83,3%) quanto em relação à intensidade final (83,3%). Por fim, a média da intensidade máxima entre os segmentos foi de aproximadamente 72 dB, sendo a maior diferença entre os segmentos de 17 dB.

Finalmente, a análise entoacional mostrou um padrão que variou entre simples e complexo, com uma porcentagem de 33,3% para o tom *H+L; de 33,3% para o tom *H; 16,7% para tom L+*H e 16,7% para o som simples *L. Como vemos, a tendência foi para que as PIs apresentem um tom com pico entoacional alto.

Os ALs e FCs foram verificados para 1,8% das produções. A disfluência do tipo AL ocorreu apenas em palavra (AL_p), sendo todos os alongamentos em sílaba em final de palavra. Os segmentos alongados foram /seis:/, /di:/ e /trabalhanu:/. A duração dos segmentos completos foi de 0,793s, 0,327s e 1,255s. Já as sílabas alongadas apresentaram uma duração de 0,516s, 0,327s e 0,665s. Deixando de lado o segmento alongado em palavra monossílaba /di:/, os ALs representaram mais de 50% do total dos segmentos. A F0 média e os pulsos variaram entre os segmentos, sendo maiores para /trabalhanu:/,

A intensidade no momento de transição apresentou uma tendência de redução em 66,6% dos casos, padrão e porcentagem que também foram verificados para a intensidade final. Por fim, a intensidade máxima também foi maior para o segmento /trabalhanu:/, sendo de 85 dB. No entanto, a diferença da intensidade desse segmento em relação ao de menor intensidade máxima foi de aproximadamente 8 dB. A média das intensidades máximas foi de 79 dB.

O padrão entoacional foi o mesmo para os três casos de AL. Ele foi anotado com o tom composto *H+L.

O segundo tipo de disfluência, FC, mantém a característica de não ter relação com os outros elementos do GT. A análise acústica mostrou uma duração média entre os elementos de 0,179s. Em relação à F0 e pulsos, apenas um segmento pode ser analisado, já que, dentre os outros dois eventos, um deles apresentou um erro de gravação e o outro foi desvozeado. Dessa maneira, não foi possível estabelecer um padrão. A análise da intensidade, por sua vez, no momento de transição, apresentou uma redução da intensidade em 100% dos casos. Por

outro lado, a intensidade final apresentou uma tendência para o aumento em 66,6% dos casos. Já a intensidade máxima variou pouco com uma diferença de aproximadamente 5 Hz entre a menor e a maior. A média entre os segmentos foi de 83 dB, aproximadamente.

A análise entoacional não foi anotada para o segmento desvozeado. Já para os outros dois segmentos a tendência é para que entoação se mantenha mais baixa. Sendo assim, os padrões obtidos foram *L e *L+H. A maior variação, representada pelo último padrão, deve-se ao erro de gravação.

Por fim, a disfluência menos significativa para VA foi a do tipo HES, que representou 0,01% dos casos. Como de costume, essa disfluência ocorreu para o segmento /ah/ em início de GT. A função pragmática desse segmento mantém os padrões verificados para os outros informantes, passando-nos a ideia propriamente de hesitação quando questionado pelo interlocutor. Observe os enunciados.

- EN1 “ah, eu nu lembu seis seti anus seis anus é seti anus/;
- EN44 “ah, trabalhanu só nu te telefoni”.

Veja que, nos dois exemplos, o informante hesita ao responder o questionamento. Isso pode estar relacionado ao fato de VA estar inseguro quanto à resposta. No primeiro exemplo essa característica fica clara, pois o informante se autocorrige, fazendo uma INS de vogal /é/ até chegar a uma resposta.

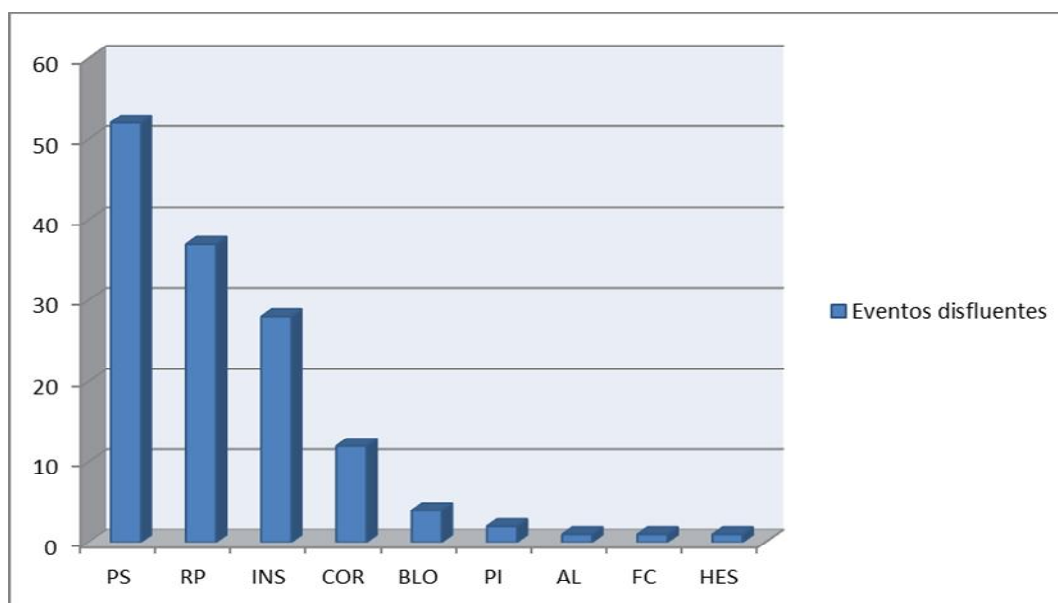
A média da duração para as duas ocorrências dessa HES foi de 0,332s. A média da F0 e os dos pulsos variaram, já que EN1 apresentou uma frequência mais alta, com uma média de 183 Hz e EN44, com uma qualidade de voz do tipo *creaky voice*, e com uma F0 média de 77 Hz. Em relação à intensidade, no momento de transição ela não foi analisada, pois se refere a segmentos em início de GT. A análise da intensidade final também não permitiu comparação, pois em um dos segmentos ela aumentou (EN1) e no outro diminuiu (EN44). Quanto à intensidade máxima, a diferença entre elas foi a mais significativa entre os outros eventos disfluentes, sendo de 83 dB para EN1 e 68 dB para EN44.

A análise entoacional apresentou para os dois casos, um tom simples, sendo alto (*H) para EN1 e baixo (*L) para EN44, o que se deve à qualidade de voz, no último caso.

O informante JO

O informante gago JO realizou 138 disfluências em 44 enunciados. Foram verificadas disfluências simples e complexas, sendo que as últimas contabilizaram um total de aproximadamente 19% dos enunciados. O máximo de disfluências realizadas por GT foram de seis, sendo três PSs (máximas) por enunciado. JO apresentou uma taxa de elocução de 4,881sil/ms, de articulação de 5,657sil/ms, e um tempo de articulação aproximado de 140ms. Veja a seguir a distribuição dos eventos disfluentes na fala de JO.

Gráfico 12 – Distribuição dos eventos disfluentes realizados pelo informante gago JO.



Fonte: própria.

De acordo com o gráfico 12 JO, assim como VA varia a sua fala entre nove tipos de disfluências, dispostas da seguinte forma: PSs; RPs; INSS; CORs; BLOs; PIs; ALs; FCs e HESs.

As PSs foram as mais frequentes para esse informante, representando 38%, aproximadamente dos eventos. Dentre elas, com uma diferença de apenas uma PS da final (38%), a mais comum foi a PS de início de GT (40%). A menos frequente foi a pausa de meio de enunciado, que representou 22% delas. No total, 46% aproximadamente das PSs foram antecedidas por eventos disfluentes.

Quanto à duração, as PSs em início de enunciado também foram as que apresentaram a maior duração (7,962s). Na sequência, temos as PSs em fim de GT (7,316s) e, por fim, as de meio de GT (4,229s). Entre PSs silenciosas e respiratórias, o primeiro tipo apresentou a maior duração (12,221s), enquanto o último apresentou uma duração total de 7,286s.

Após as PSs o segundo maior número de disfluências foi verificado para as RPs (26,8%), que estão distribuídas da seguinte forma:

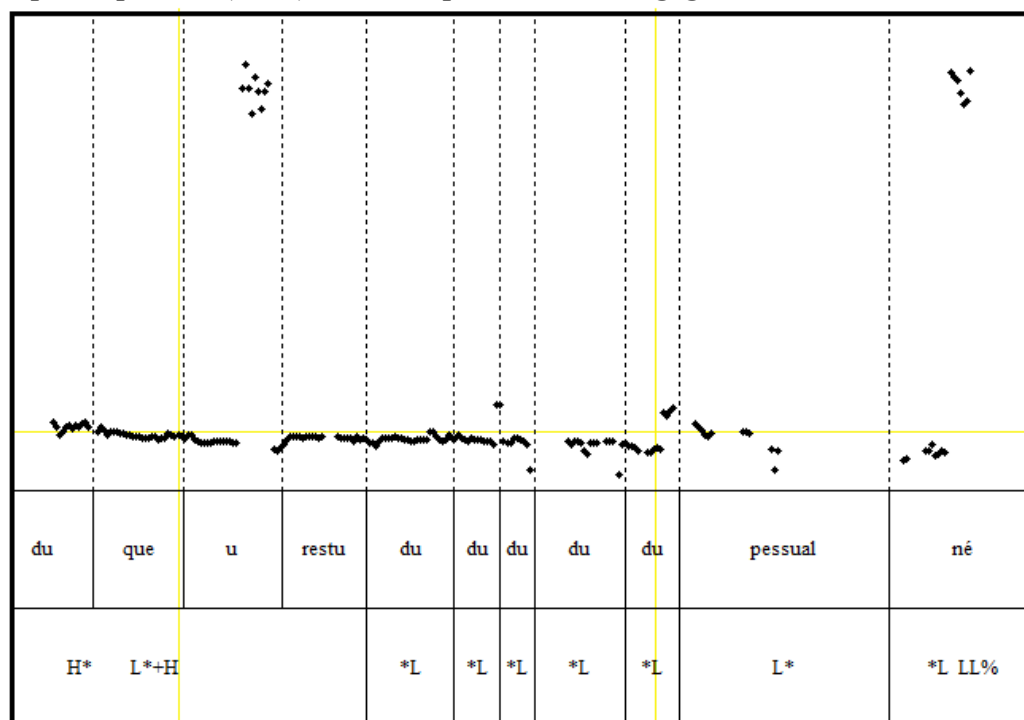
- 2RP_p (27%);
- RP_c (29,8%);
- 2RP_e (29,7%);
- 2RP_v (8,1%);
- 3RP_p (2,7%);
- 5RP_p (2,7%).

A distribuição mostra que as RPs de palavra foram as mais frequentes, aparecendo em 81% do total de repetições. Entre as repetições duplas de palavras (2RP_p) um caso ocorreu em palavras dissílabas e um em palavra trissílaba. JO não realizou repetições de sílaba. Os erros de gravação que foram identificados nas repetições (que foram dois) ocorreram na segunda palavra repetida. As repetições complexas (RP_c) foram realizadas concomitantemente com pausas (quatro casos); PIs (quatro casos); ALs (dois casos) e, BLO (um caso). É importante salientar que os ALs e o BLO ocorreram na segunda repetição.

No que diz respeito à análise acústica da duração observamos as seguintes tendências:

- Para as repetições do tipo 2RP_p (66,6%); RP_c (55,5%) e RP_e (63,6%) há um aumento da duração na segunda repetição;
- Já para a 2RP_v ocorre o contrário de modo que há uma tendência para a diminuição da duração na segunda repetição em 66,6% dos casos;
- A 3RP_p (um caso) apresenta um padrão em forma de onda, que começa com uma duração menor (0,485s) que depois aumenta (0,791s) e, por fim, aumenta mais uma vez (0,604s);
- A repetição quádrupla, por sua vez, apresenta uma duração que não segue nenhuma tendência. Veja a figura abaixo:

Figura 43 – Representação de um trecho do enunciado para a visualização da repetição quintupla de palavra (5RP_p) realizada pelo informante gago JO.



Fonte: própria.

A fim de facilitar a visualização, a imagem 43 refere-se a um fragmento do enunciado completo de EN12 “eu tenho um racio raciocíniu ma um poco mais lógico du que u restu *du du du du* pessual né”.

Na figura aparecem identificados a F0, a segmentação do enunciado na primeira camada e, na segunda camada, a análise fonológica. Como podemos ver, o informante repete cinco vezes o segmento /du/. A segmentação do enunciado mostra a duração variada de cada uma das palavras, que medem, respectivamente, 0,272s; 0,146s; 0,107s; 0,283s e, 0,169s. Nas três primeiras repetições a duração segue um padrão decrescente, depois, na quarta repetição há um aumento na duração com posterior redução.

No que se refere à F0 média e aos pulsos os resultados foram:

- Para as 2RP_p (77,7%); 2RP_v (66,6%) e RP_c (55,5%) ocorre uma tendência para a diminuição da F0 média na segunda repetição, bem como para os pulsos. No entanto, para o último as porcentagens foram respectivamente: 88,8%, 100% e 55,5%;
- Para a RP_c, por sua vez, ocorre uma aumento da F0 média (72,7%) e pulsos (63,6%) na segunda estrutura repetida;
- Para a 5RP_p, assim como foi observado para a duração há uma variação na F0 média e pulsos, uma vez que começa com uma frequência mais baixa (119 Hz), depois

aumenta (121 Hz), volta à F0 inicial (119 Hz), diminui novamente (113 Hz) e, por fim, volta à frequência inicial. É importante salientar que os valores das frequências foram aproximados. No entanto, a diferença entre a primeira e a terceira RP foi de 0,294 Hz e, entre a primeira e a última foi de 0,469 Hz;

- Por fim, a 3RP_p apresenta uma F0 que diminui progressivamente ao longo das repetições e um esforço vocal que varia em forma de onda, começando mais baixo, aumentando e, diminuindo novamente.

Quanto à intensidade de transição, final e inicial foram verificadas as seguintes características:

- Em relação a 2RP_v, não foi possível analisar a intensidade na passagem da fluência para a primeira RP, uma vez que dentre as repetições, duas delas iniciaram o GT e uma foi antecedida por uma PS. Por outro lado, a intensidade na transição da primeira RP para a segunda apresentou uma diminuição em 100% dos casos. A intensidade final também apresentou uma tendência para a diminuição em 66,6% dos casos, bem como foi observado para a intensidade máxima (100%), quando comparamos as duas repetições;

- Quanto a 2RP_p, a intensidade diminuiu tanto na passagem da fluência para a primeira RP quanto da primeira RP para a segunda em 77,7% das repetições. As intensidades final e máxima seguiram o mesmo padrão, diminuindo em 55% dos casos em relação ao primeiro parâmetro e, em 60% para o segundo;

- As repetições complexas, tanto no momento de transição da fluência para a disfluência quanto da disfluência para a disfluência também seguiram a mesma tendência da 2RP_p, diminuindo em 66,6% nos dois momentos observados. Por outro lado, houve um aumento da intensidade final em 58,3% das repetições, bem como da intensidade máxima (55,5%);

- A repetição de estrutura, por sua vez, manteve a tendência para a diminuição da intensidade apenas na transição da fluência para a disfluência (85,7%). Já na passagem da primeira repetição para a segunda, a intensidade aumentou em 60% dos dados. A intensidade final aumentou em 54,5% dos casos e, a intensidade máxima diminuiu em 54,5% das repetições;

- A 3RP_p revelou uma tendência para o aumento da intensidade nas transições. Já a intensidade final revelou uma diminuição da intensidade em duas das três RPs. Por fim, a intensidade máxima diminuiu progressivamente durante as três repetições (76 dB, 74 dB e 58 dB);

▪ Finalmente, a 5RP_p mostrou que dentre as cinco repetições, três delas (60%) aumentaram a intensidade na passagem de uma para a outra. Por outro lado, a intensidade diminuiu também em 60% das repetições. Já a intensidade máxima também abaxou progressivamente. Mesmo assim, as diferenças não foram significativas (72 dB; 68 dB; 63,918 dB; 63,153 dB; 60 dB).

Por fim, a análise entoacional manteve a tendência de realização do mesmo padrão em mais de 50% das repetições. Observe:

▪ A 2RP_p manteve o mesmo padrão em 80% dos casos, variando entre *H e *L. Os dois casos (20%) que destoaram apresentaram o seguinte padrão *H *H+L. Como vemos, a variação está entre um tom simples na primeira RP e composto na segunda. Mesmo assim, o pico entoacional continuou no tom alto;

▪ A 2RP_v, em 66,6% dos casos mantém o padrão com o tom *L. A repetição que varia também apresenta o tom baixo na primeira RP. Contudo, a segunda RP é surda;

▪ As RPs de estrutura mantêm o padrão em aproximadamente 82% dos casos, variando entre *L, *H e *L+H;

▪ A RP complexa foi o único caso de repetição que manteve o mesmo padrão entoacional em apenas em 45,5% dos casos, variando entre tons simples e complexos. Os casos em que o padrão não se manteve foram identificados para todas as PIs e, para duas das repetições que ocorreram junto com PS;

▪ A repetição do tipo 5RP_p como podemos ver na figura 43 conservou o mesmo tom baixo (*L) para todas as repetições;

▪ Finalmente, a 3RP_p manteve o mesmo tom composto *H+L nas duas primeiras repetições, modificando-o significativamente na terceira RP por meio de um tom simples baixo *L.

Depois das RPs, as INSS aparecem em 20% dos eventos disfluente. Elas dividiram-se INSS_ps (50%) e INSS_vs (50%). Dentre elas, também foram verificadas INSSs complexas, que ocorreram apenas em vogal com alongamento e INSSs de estrutura. As inserções realizadas por JO foram assim distribuídas:

- INSS_p em início de GT - /intão/, /ái/;
- INSS_p em fim de GT - /né/, /assim/, /ta/;
- INSS_v em início de GT - /é/, /i/;
- INSS_v em meio de GT - /é/;
- INSS_e em início de GT - /tão é/, /i é/;

- INS_c em início de GT - /:é/;
- INS_c em meio de GT - /:é/.

As INS s simples, as INS s de palavra em início de GT e de vogal em início e meio de GT mantiveram a função semântica sequenciadora, bem como as INS s de estrutura. As INS s complexas também dão ideia de sequência no enunciado. No entanto, o alongamento de um segmento mostra que o informante demora mais tempo para voltar à fluência, ou seja, a programação da mensagem dura mais do que quando ele faz apenas uma inserção. Finalmente, as INS_{ps} do tipo /né/ conservam a característica de uma pergunta retórica afirmativa, assim como /ta/. Por sua vez, o segmento /assim/ nos passa a ideia de exemplificação. Observe o enunciado:

- EN2 “i chegava nu nu na na sala deli *assim#*”.

No enunciado, JO conta sobre uma passagem que viveu com um professor do curso de Engenharia. Ao discorrer sobre o assunto, realiza uma INS com função de exemplificar a sua fala. Contudo, o informante a desloca para o final do GT, o que ainda não havia sido observado para nenhum dos dois grupos de informantes.

No que se refere à análise acústica das INS s, a maior duração foi verificada para as $INS_{e,s}$ (1,197s). Na sequência temos as $INS_{c,s}$ (0,834s); as INS_{ps} em início de GT (0,694s); a INS_v em meio de GT (0,500s); as INS_{ps} em fim de GT (0,353s) e, por fim, com a menor duração, as $INS_{v,s}$ em início de GT (0,289s).

Em relação à F_0 média, a maior frequência também ocorreu para as $INS_{e,s}$ (142 Hz). Depois, aparecem as INS_{ps} em início e fim de GT, com uma diferença de 1 Hz, sendo de 114 Hz para a primeira e 115 Hz para a segunda. Diferença essa que também foi verificada para as $INS_{v,s}$ em início de GT (103 Hz) e para as INS_s complexas (102 Hz). Finalmente, a menor frequência média foi identificada para a inserção de vogal em meio de GT (90 Hz). Quanto aos pulsos, o maior esforço vocal ocorre para as INS s de estrutura, que são seguidas das INS_{ps} em início de GT, das INS_s complexas, da INS_v em meio de GT, das INS_{ps} em fim de GT e, por fim, com o menor esforço vocal temos as $INS_{v,s}$ em início de GT.

A análise da transição da intensidade não foi observada para as INS_{ps} , $INS_{v,s}$, INS s complexas e INS s de estrutura, uma vez que iniciaram o GT e, algumas, foram antecedidas por PS s (INS_v em meio de GT). Sendo assim, apenas as INS_{ps} em final de GT foram anotadas. Não foi possível estabelecer um padrão, uma vez que os resultados equipararam-se. No que concerne à intensidade final, tanto as $INS_{v,s}$ em início de GT quanto às de meio de GT apresentaram um aumento da intensidade em 100% dos casos. Já as INS s complexas e INS s

de palavra em início de fim de GT apresentaram uma diminuição da intensidade em, respectivamente, 60%, 75% e 70% dos casos. Por fim, a média da intensidade máxima revelou a maior intensidade, para as INSS de estrutura (75 dB). Em seguida, ocorrem as INSS em início de GT (71 dB), as INSS_c (70 dB), INSS_v em início de GT (68 dB), as INSS_p em fim de GT (67 dB) e, por fim, a INSS_v em meio de GT (64 dB). Como podemos ver, a diferença entre elas também é muito pequena (11 dB).

A análise dos padrões fonológicos confirmou a tendência das INSSs manterem o mesmo padrão entoacional. Dessa maneira, os padrões obtidos para as inserções foram:

- /intão/ e /ái/ em início de GT;
- /intão/ *L+H para 75% dos casos;
- /ái/ *L (caso único);
- /né/, /ta/ e /assim/ em final de GT
- /né/ *L em 66,6% dos casos e *L+H em 33,3% dos casos;
- /ta/ *L em 66,6% dos casos;
- /assim/ *L+H (caso único);
- /é/ e /i/ em início de GT – *L para 100% dos segmentos do tipo /é/. Já o segmento /i/ variou entre *L, *L+H e *H. Mesmo assim, a tendência é para o tom com pico entoacional baixo;
- /é/ em meio de GT - *L (caso único);
- /tão é/ *H e /i é/ *L, para os casos de INSS_e;
- /:é/ *L em 60% dos casos.

As CORs representam 8,7% do total de disfluências. Elas dividiram-se entre eventos simples (66,6%) e complexos (33,3%).

As CORs complexas ocorreram junto com PS_e (dois casos), repetição de palavra (um caso) e inserção de vogal (um caso). Dentre as CORs, o informante manteve parte da informação contida no erro na correção em apenas 8% dos casos.

A análise acústica revelou para a COR complexa uma diminuição da duração na correção em 75% dos casos. Já para a COR simples os resultados equipararam-se, em 50% houve diminuição e em 50%, aumento. Quanto à F0 média, a tendência é para a diminuição da entoação na correção para a COR complexa (75%) e aumento para a COR simples (62%). Por sua vez, os pulsos revelaram um maior esforço na correção para a COR simples. Já para a COR complexa os resultados equipararam-se.

A análise da intensidade revelou, deixando de lado os momentos de PS e COR em início de GT, para a COR simples, uma diminuição da intensidade tanto na passagem da fluência para o erro (66,6%) quanto do erro para a correção (75%). Quanto à COR complexa, os resultados equipararam-se nos dois momentos observados. Por sua vez, a intensidade final equiparou-se para a COR simples. Já para a COR complexa, a tendência foi para o aumento em 62,5% dos casos. Por fim, a análise da intensidade máxima também foi diferente para a COR simples e COR complexa. No primeiro caso, a intensidade máxima não revelou uma tendência, uma vez que em 50% dos casos ela diminuiu na correção e, em 50% aumentou. No segundo caso, houve uma tendência para a diminuição na correção em 75% dos casos.

Em relação à análise fonológica, os padrões foram diferentes na passagem do erro para a correção. Eles variaram entre tons simples e compostos na passagem de um momento para o outro. Durante essa transição, os padrões mantiveram-se o tom baixo tanto no erro quanto na correção em 50% dos casos. Nos demais eventos (50%) os padrões variaram entre tons altos e baixos na passagem do erro para a correção.

Após as CORs ocorre a disfluência do tipo BLO em aproximadamente 3% dos eventos disfluentes. Os BLOs foram verificados apenas em consoantes, sendo elas:

- 2 ocorrências em consoante oclusiva bilabial surda [p], nos segmentos /-passa/ e /publicassi/;
- 1 ocorrência em consoante oclusiva bilabial sonora [b], no segmento /-bom/;
- 1 ocorrência em consoante oclusiva velar surda [k];
- 1 ocorrência em consoante oclusiva alveolar sonora [d], no segmento /-dizenovi/.

Em relação ao tipo de PS de bloqueio, todos os eventos caracterizaram-se por não apresentar PS antes do segmento. A duração média deles foi de 0,711s. A análise individual dos segmentos revela que as oclusivas bilabiais apresentaram a maior duração (0,757s), seguidas da oclusiva alveolar (0,745s) e, por fim, da oclusiva velar (0,588s). A F0 média dos segmentos foi de aproximadamente 112 Hz. Mais uma vez, a análise individual revelou uma F0 média maior para as oclusivas bilabiais (123 Hz), seguidas da oclusiva alveolar (100 Hz). Não foi possível analisar a F0 média da oclusiva velar, devido a um erro de gravação. Quanto aos pulsos, tanto as oclusivas bilabiais quanto a alveolar tiveram o mesmo resultado. Portanto, não se diferenciaram quanto ao esforço vocal.

Por sua vez, a análise da intensidade de transição mostrou uma tendência de abaixamento para 75% dos casos, assim como a intensidade final. Já a intensidade máxima

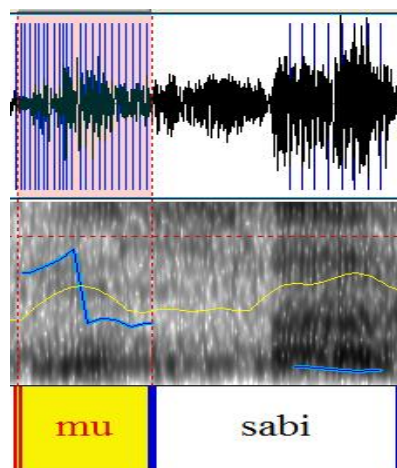
apresentou uma média 70 Hz. Individualmente, a diferença entre a maior (oclusiva alveolar) e a menor (oclusiva velar) intensidade máxima foi muito pequena, sendo de 0,987 dB.

Quanto aos padrões entoacionais, 100% dos segmentos apresentaram um tom simples. As consoantes oclusivas bilabiais surdas foram analisadas com o tom baixo *L, bem como a oclusiva alveolar sonora [k]. Por outro lado, a oclusiva velar surda foi analisada com o tom alto *H.

A disfluência do tipo PI é um dos tipos menos frequentes realizados por JO, contabilizando aproximadamente 1,4% dos eventos. Dentre as duas ocorrências, uma apresentou um erro de gravação, devido a uma qualidade de voz mais aguda e a outra foi desvozeada. Por isso, a F0 e os pulsos não foram anotados. A duração média desses eventos foi de 0,130s. Quanto à intensidade, durante a transição não foi possível estabelecer padrões, uma vez que em um segmento ela aumentou e no outro diminuiu. Já a intensidade final aumentou em 100% dos casos. Por fim, a média da intensidade máxima entre os segmentos foi de 66 dB, sendo diferença entre os segmentos de 18 dB.

Finalmente, quanto à análise entoacional, não foi atribuído um padrão fonológico para o segmento surdo, bem como para o segmento com o erro de gravação. Observe abaixo que mostra um erro de extração.

Figura 44 – Representação de um erro de gravação no enunciado EN35 “é qui todú mu sabi queu so, todú mundu fica falanu cumigu, vo isconde porque?”, realizado por JO.



Fonte: própria.

A figura 44 identifica apenas um trecho do enunciado completo EN35 “é qui todú mu sabi queu so, todú mundu fica falanu cumigu, vo isconde porque?”. A disfluência aparece selecionada em amarelo no segmento /mu/. Como podemos ver, a linha referente à frequência

entoacional apresenta um padrão que foge da normalidade. Fato que pode ser comparado ao segmento /sabi/. A F0 média fornecida pelo o programa para essa PI é de 279 Hz, enquanto que para /sabi/ é de 111 Hz.

O AL foi verificado em 0,7% das produções. Essa disfluência ocorreu em final de sílaba de palavra. É importante lembrar que também ocorreram alongamentos em vogal. No entanto, assim como vimos, eles foram realizados apenas em estruturas complexas, com repetições e inserções.

O segmento alongado ocorreu em /va:/, no enunciado EN20 “/queli va: /é é intende”. A sua duração foi de 0,886s. Já a F0 média de foi de 146 Hz. As intensidades no momento de transição, bem como a final, diminuíram. A intensidade média, por sua vez, foi de 75 dB. O padrão entoacional desse segmento foi anotado com o tom composto *H+L.

Por sua vez, a disfluência do tipo FC também não apresentou relação com os outros elementos do GT. No conjunto das disfluências, a sua ocorrência foi insignificante, bem como para AL, com 0,7% das ocorrências.

O segmento que caracterizou essa disfluência aparece no enunciado abaixo:

- EN28 “/ nossa com dificultadi pa fala qualque *pa* Araraquara”.

O FC ocorre no segmento em *itálico* /pa/. É possível que se JO o continuasse, poderia realizar o segmento /palavra/. A justificativa para essa disfluência seria a de que o informante a deixou de lado, vazia no enunciado, porque, provavelmente, iria realizar um outro tipo de disfluência, isto é, um BLO_c, já que há uma tendência para que o informante bloqueie oclusivas bilabiais.

A análise acústica mostrou para esse segmento uma duração de 0,314s. A F0 média foi de 146 Hz. A análise da intensidade, por sua vez, diminui no momento de transição, assim como a final, comparativamente à inicial. Já a intensidade máxima apresentou uma duração de 62 dB aproximadamente. Por fim, a análise entoacional foi anotada pelo tom simples alto *H.

Finalmente, junto com o AL e FC, ocorre a disfluência do tipo HES, também com 0,7% dos eventos. Como de costume, essa disfluência ocorreu para o segmento /ah/ em início de GT. A função pragmática desse segmento mantém o significado, nesse caso de surpresa, verificado para os outros informantes gagos e não gagos.

A duração dessa HES foi de 0,280s. A média da F0 foi de 164 Hz. Em relação à intensidade, no momento de transição ela não foi analisada, pois é um segmento em início de GT. A análise da intensidade final diminuiu em relação à inicial e a intensidade máxima foi de

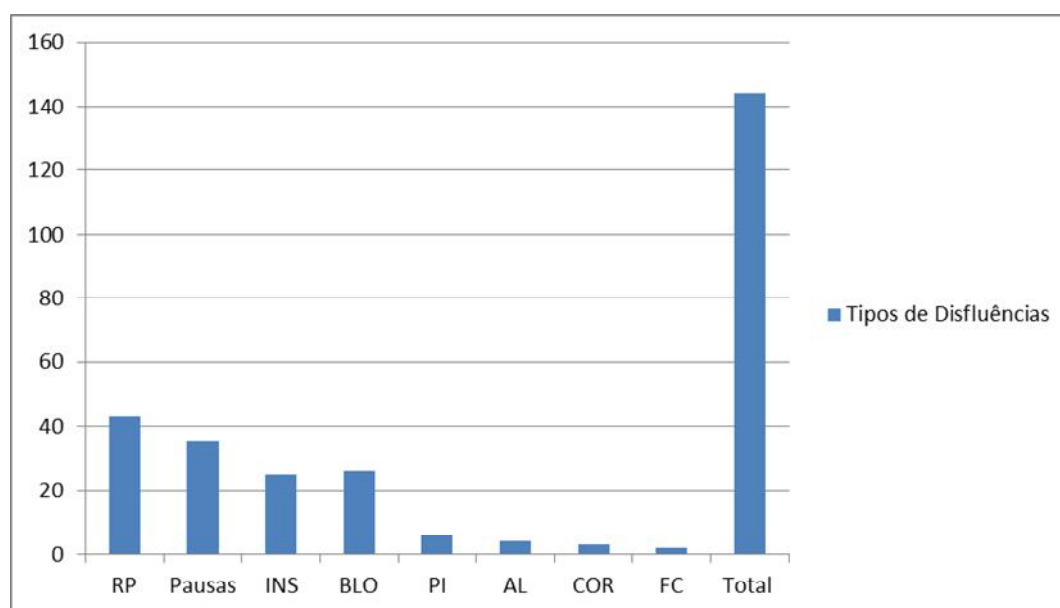
74 dB. A análise entoacional apresentou um tom composto, com pico entoacional alto do tipo *H+L.

O informante AP

O informante gago AP realizou 144 disfluências em 35 enunciados. Dentre elas, 24 foram complexas. O número máximo de disfluências por enunciado foi de oito. Dentre esses eventos ocorreu apenas uma PS. A sua taxa de elocução foi de 4,933sil/ms, de articulação de 5,555sil/ms, e um tempo de articulação de 140ms.

Observe a seguir o gráfico 13 que caracteriza as disfluências realizadas por AP ao longo da gravação.

Gráfico 13 – As variações de disfluência gaga na gravação de AP.



Fonte: própria.

Antes de iniciarmos a análise acústica, é importante salientar que AP utiliza com frequência as palavras /vamu supo/ e sua variação /va supo/. Essa expressão funciona como um apoio para de fala, ajudando-o a formular as ideias que ocorrem após a sua produção. Na maior parte das vezes, após essa expressão temos um período fluente. No entanto, essas expressões também ocorrem com repetições e bloqueios. Observe o exemplo:

▪ EN5 /é -o /é -va supo ansioso né -vamu supo -vamu supo -vamu supo quandu a pessoa ti pressiona u ce/.

AP inicia o GT com um BLOv /-o/, com o intuito de produzir a palavra /ansioso/. Como não consegue, faz uma INSv /é/ seguida da expressão /-va supo/, que também é

bloqueada. Por fim, encerra o primeiro GT com a INS_p /né/ e, a fim de dar mais um exemplo ao entrevistador, faz uso, mais uma vez, do apoio de fala, repetindo-a mais três vezes até conseguir produzir a mensagem final e retomar a fluência.

A partir dessas considerações, o gráfico mostra que AP produz mais RPs, depois PSs, seguidos, respectivamente, dos BLOs e INSs. Por fim, aparecem as PIs, os ALs, as CORs e os FCs.

A RP foi classificada entre simples e composta. No total, ela representou aproximadamente 30% dos casos. Elas foram distribuídas no contexto geral da gravação como:

- RP complexa (40%);
- RP_e (20%);
- RP_{ps} (16%);
- RP_p (12%);
- RP_v (9%).

Como podemos ver, as RPs complexas foram as mais frequentes para AP. As RPs complexas envolveram palavras e vogais, mas as primeiras foram as mais comuns. Na sequência, aparecem as RPs de 2) estrutura; 3) de sílaba, que foram verificadas em palavras de cinco, três e duas sílabas; 4) de palavras, sendo uma RP_p dupla, três triplas e uma quádrupla e, por fim, as repetições 5) de vogais, variando entre repetições duplas (um caso), triplas (um caso), quádruplas (um caso) e, quádruplas (um caso).

Em relação à RP_{ps} , AP realizou apenas duas RP_{ps} dupla de sílaba: /na na naquelas/ e /va va vasupo/. As RP_{ps} de vogal foram encontradas apenas para repetições do tipo /é/. Vale ressaltar, ainda, que para a RPe , AP costuma anunciar primeiro o foco, para em seguida repetir o padrão. Veja nos dois exemplos abaixo:

- EN18 “/eli eli tava eli tava/”;
- EN22 “/eu pegu pe pegu uma coisa daquela li i ergu pra cima#”.

Em EN18, AP anuncia o foco e o retoma anaforicamente por meio de uma RP_e , bem como em EN22. Essa característica da fala do informante funciona como um recurso discursivo que o ajuda a recuperar a mensagem, mas não evita a disfluência.

A análise dos padrões acústicos revelou os seguintes resultados:

1) RP_p

- **Duração** – Em 40% dos dados ocorre uma diminuição progressiva da duração; em 40% há uma duração em forma de onda, em que ocorre o abaixamento e o aumento da duração e, em 20% deles houve apenas um aumento na duração;

- **F0 média** – Em 60% dos dados, a entoação segue o padrão de diminuição e aumento da F0 e, em 40% a diminuição da F0 na segunda RP. É importante salientar que em 60% das palavras repetidas a F0 média foi maior para a segunda ou terceira palavra. Mesmo assim, no contexto geral, temos uma F0 que varia pouco, sendo a mínima de 90 Hz e a máxima de 108 Hz;

- **Pulsos** – Em 80% dos dados o esforço vocal é menor na segunda ou terceira RP_p;

- **Intensidade** – Em relação à quantidade de dados, apenas uma RP de um GT foi descartada, pois ocorreu em início de enunciado. Sendo assim, a passagem da primeira RP para a segunda ou terceira palavra, revelou para 82% dos casos uma diminuição da intensidade, enquanto que para 18% houve um aumento. Já em relação à transição da fluência para a disfluência (primeira palavra repetida), houve um empate, de modo que em 50% dos dados a intensidade diminuiu e em 50% deles aumentou. Quanto à intensidade final, em 62,5% dos casos a intensidade final aumentou, comparativamente à intensidade inicial. Por fim, a análise da intensidade máxima mostrou que para as RPs duplas, a intensidade diminuiu para a segunda palavra (um caso). Por sua vez, para os casos de duas ou mais RPs, a intensidade varia, de modo que para 50% dos casos ela mantém uma intensidade inicial maior e, 50%, menor.

2) RP_v

- **Duração** – Para as RPs em vogais os resultados mostram uma variação na duração, sendo que em 25% dos dados houve uma duração que variou em forma de onda (4RP_v) e, em 25% houve um aumento progressivo (3RP_v). Nos outros 50% dos casos houve uma diminuição na duração (2RP_v e 5RP_v);

- **F0 média** – Em 75% dos dados, a entoação segue o padrão de diminuição e aumento da F0 em forma de onda, sendo a F0 maior em 100% dos casos para a primeira RP_v e, em 15% ocorre a diminuição da F0 na segunda RP. Vale salientar que esse padrão ocorreu apenas para a RP_v dupla. Assim como para a RP_v, no contexto geral, temos uma F0 que varia pouco, sendo a mínima de 101 Hz e a máxima de 120 Hz;

- **Pulsos** – Em 75% dos dados o esforço vocal é maior na segunda ou terceira RP_v;

- **Intensidade** – Em relação à quantidade de dados, apenas uma RP_v de um GT foi descartada, pois ocorreu após uma PS. Dessa maneira, a passagem da primeira RP para a segunda ou terceira palavra, revelou para 60% dos casos uma diminuição da intensidade, enquanto que 20% aumentou. Já em relação à transição da fluência para a disfluência (primeira vogal repetida), também houve um empate, de modo que em 50% dos dados a intensidade diminuiu e em 50% deles aumentou. Quanto à intensidade final, os valores equipararam-se percentualmente, ou seja, em 50% dos casos a intensidade variou em forma de onda e, em 50% aumentou progressivamente. No entanto, a tendência (75%) é para que a intensidade seja maior na primeira RP_v.

3) RP_s

- **Duração** – Para as RP_s, em 100% dos casos a duração é menor na sílaba em palavra, e não na sílaba individual. Nos casos em que a repetição de sílaba é dupla, a sílaba intermediária apresenta uma duração maior do que a primeira sílaba;

- **F0 média** – Dos casos analisados, três deles apresentaram uma sílaba em palavra desvozeada, ambos em /*pessoa*/. Dentre as outras RP_ss, em 50% dos casos a F0 é maior na sílaba em palavra e, em 50% na sílaba individual. Assim como foi observado para as demais RPs, a F0 média apresentou pouca variação, de modo que a F0 mínima foi de 80 Hz e a máxima de 92 Hz;

- **Pulsos** – Deixando de lado o caso em que não houve sonorização, em 90% dos casos o esforço vocal é menor na última RP_s;

- **Intensidade** – Em relação à quantidade de dados, apenas uma RP_s foi descartada, pois ocorreu em início de GT. Dessa maneira, a passagem da primeira RP_s para a segunda sílaba isolada ou para a sílaba em palavra, revelou para 66% dos casos uma diminuição da intensidade. Já em relação à transição da fluência para a disfluência (primeira sílaba), também houve um empate, de modo que em 50% dos dados a intensidade diminuiu e em 50% deles aumentou, bem como verificados para os outros tipos de RPs. Quanto à intensidade final, os valores também foram iguais, sendo que em 50% dos casos aumentou e, em 50% diminuiu. Em relação à intensidade máxima, para as repetições de sílabas simples, em 60% dos casos a intensidade aumenta. Já para as RPs em que ocorre a repetição da sílaba inicial de palavra duas vezes, há uma variação, sendo verificados um caso de diminuição progressiva, e um de variação em forma de onda. Em 57% dos casos a entoação se mantém mais alta na primeira sílaba.

4) RP_e

- **Duração** – Para as repetições de estrutura dupla, em 57% dos casos a duração diminuiu. Já para as RPs com mais de duas estruturas repetidas, em um caso a duração variou em forma de onda e, em um caso aumentou nas três primeiras RPs e diminuiu na última;

- **F0 média** – Quanto à F0 média a tendência é para a diminuição da entoação para as RPs duplas (57%). Para as RPs com mais de duas estruturas, a tendência também é para que a última tenha uma entoação menor. A variação da entoação para esse tipo de RP é um pouco maior do que as anteriores, com uma diferença de 31 Hz, sendo a F0 mínima de 102 Hz e a máxima de 133 Hz. Mesmo assim, podemos dizer que a entoação varia pouco;

- **Pulsos** – No que diz respeito à RP_e dupla, o esforço vocal é menor na segunda RP (57%). Já para as repetições de mais de duas estruturas não há uma tendência, uma vez que dos dois casos, em um o esforço vocal foi maior na última e RP_e e no outro, foi menor;

- **Intensidade** – Dentre as RP_s_e analisadas, foram descartadas apenas dois segmentos para a análise da transição da fluência para a disfluência, uma vez que um ocorreu em início de GT e outro após uma PS. Sendo assim, foi verificado para 85% dos casos um aumento na intensidade. Já na passagem da primeira RP_e para a segunda ou terceira, em 53% dos casos a tendência foi para a diminuição da intensidade. Quanto à intensidade final, em 59% dos casos ela aumentou. Em relação à intensidade máxima para as repetições duplas, a tendência é de 71% para que ela seja menor na última RP_e e, conseqüentemente, maior na primeira. Já para as repetições de mais de duas estruturas, a variação continua. No entanto, comparando a intensidade da primeira e da última RP_e, a tendência para que a intensidade máxima seja maior na primeira RP_e continua.

5) RP_c⁶²

- **Duração** – Para as RPs complexas duplas, seja de RP_s, RP_p ou RP_e, a tendência é de 66% para a diminuição da duração na segunda repetição. Para os outros casos há variação entre diminuição progressiva (um caso), variação em forma de onda (um caso) e aumento (um caso), sendo que para 66% deles a duração final é menor do que a inicial;

- **F0 média** – Quanto à F0 média, para as repetições duplas o maior número de casos é de 53% para a diminuição da entoação na segunda RP. Já para os outros casos, a tendência é para o aumento (66%). A tendência a entoação constante continua para esse tipo de RP_c, de modo que a F0 mínima é de 88 Hz e a máxima de 108 Hz;

⁶² Por se tratar de um caso mais complexo, como o próprio nome já diz, as RP_cs serão observadas de uma maneira geral. As observações específicas serão feitas quando necessário.

- **Pulsos** – No que diz respeito à RP_c dupla, o esforço vocal é menor na segunda RP (60%). Já para as repetições de mais de duas estruturas a tendência (66%) é para que o esforço vocal seja maior na última repetição;

- **Intensidade** – Dentre as RP_s analisadas, em momento de transição da fluência para a primeira RP foram descartados oito segmentos, por os mesmos motivos das RP_s anteriores e, ainda, por serem antecidos por uma PS de BLO. Dessa maneira, em 80% dos casos a intensidade aumentou. O mesmo procedimento foi feito para a transição da primeira para a segunda e terceira RPs. Nesse caso, a tendência também foi para o aumento da intensidade (60%). Quanto à intensidade final, em 62,5% dos casos houve uma diminuição. Por fim, a intensidade máxima revelou um aumento para 53% dos casos. Já para os três casos de RP tripla (dois casos) e dupla (um caso), a tendência é para que a intensidade final seja menor do que a inicial (66%).

A partir das observações feitas, o conjunto geral das RPs revela que a maior tendência é para a diminuição da duração (60%), bem como para a F0 média. A média das frequências mínima e máxima mostrou uma tendência do informante em variar pouco a entoação. Os pulsos também seguiram a mesma proporção, sendo o esforço vocal menor na segunda RP. Quanto à intensidade, no momento de transição da fluência para a primeira RP, a tendência foi para a diminuição (80%). Quanto à transição da disfluência para a disfluência, ou seja, para a segunda ou terceira RP, não foi possível estabelecer um padrão, uma vez que os resultados se assemelharam, variando entre diminuição e aumento da intensidade. Quanto à intensidade final, o padrão foi o mesmo, de modo que para 40% dos casos a intensidade diminuiu; em 40% há uma equivalência entre o aumento e diminuição da intensidade, e em 20% a intensidade final aumentou em relação à inicial. Finalmente, quanto à intensidade máxima, também não há um padrão, uma vez que em 40% dos dados a intensidade máxima aumentou na segunda e terceira RP; em 40% os valores equipararam-se e, em 20% dos casos, diminuiu.

No que diz respeito à análise fonológica, assim como foi observado para os outros informantes, as repetições mantêm o mesmo padrão entoacional para a maior parte dos eventos. É importante acrescentar que para as RPs de sílaba pode acontecer de haver uma variação entre um tom simples na sílaba isolada e um tom composto na sílaba em palavra. Fato que é esperado, uma vez que o informante muda a tonicidade na palavra. Mesmo assim, a análise entoacional isolada da sílaba na palavra revela o mesmo padrão. Além disso, quando

o padrão variou entre simples e composto para os outros tipos de RPs, mas o pico entoacional se manteve .

Depois das RPs, as PSs foram as mais comuns para AP. Em 144 disfluências, aproximadamente 25% delas foram de PSs. Dentre elas, as PSs em início de enunciado foram as mais frequentes com aproximadamente 39% das ocorrências. Depois aparecem as PSs em meio de enunciado 36% e, por fim, as PSs em fim de enunciado 25% das ocorrências. De todas as ocorrências, as PSs silenciosas foram as mais frequentes (65% das ocorrências). Em relação à duração, as PSs respiratórias são mais longas do que as silenciosas, com uma média de duração de 0,722s, enquanto que as PSs silenciosas duram 0,676s. As pausas em meio de enunciado podem aparecer antes ou depois de uma disfluência ou, como é o caso das RPs, entre uma repetição e outra, ajudando o informante a recuperar a fluência. Observe:

- EN27 /-va supo eu já eu já eu já eu já ap /apartei muitas briga né/

Em EN27 o informante inicia o GT com a expressão de apoio bloqueada em fricativa [v], /-va supo/, que é seguida por uma repetição quadrúpla de estrutura (4RPe) “eu já eu já eu já eu já”. Na sequência, ocorre uma repetição de sílaba incompleta (RP_{psi}), que é interrompida por uma PS respiratória para retomar a fluência.

A disfluência do tipo BLO foi verificada em 18% dos casos, variando entre vogais e consoantes, e entre BLOs simples e complexos. Os BLOs em consoante foram maiores (59%) do que em vogais (41%). Os bloqueios complexos ocorreram com INS_{ps} (dois casos), COR (um caso) e RPs (10 casos). Sendo assim, levando em consideração os momentos em que os BLOs aparecem com outros casos de disfluência, eles representam 27% do total de eventos. Portanto, seria o segundo tipo de disfluência mais recorrente para AP.

Os segmentos bloqueados foram:

1) Bloqueios em vogal:

- 3 casos para a vogal meio fechada anterior [e];
- 3 casos para a vogal aberta anterior [a];
- 3 casos para a vogal meio aberta anterior [ɛ];
- 2 casos para a vogal fechada posterior [u];
- 1 caso para a vogal aberta posterior [ɔ];
- 1 caso para a vogal anterior fechada [i];
- 1 caso para vogal meio fechada posterior [o].

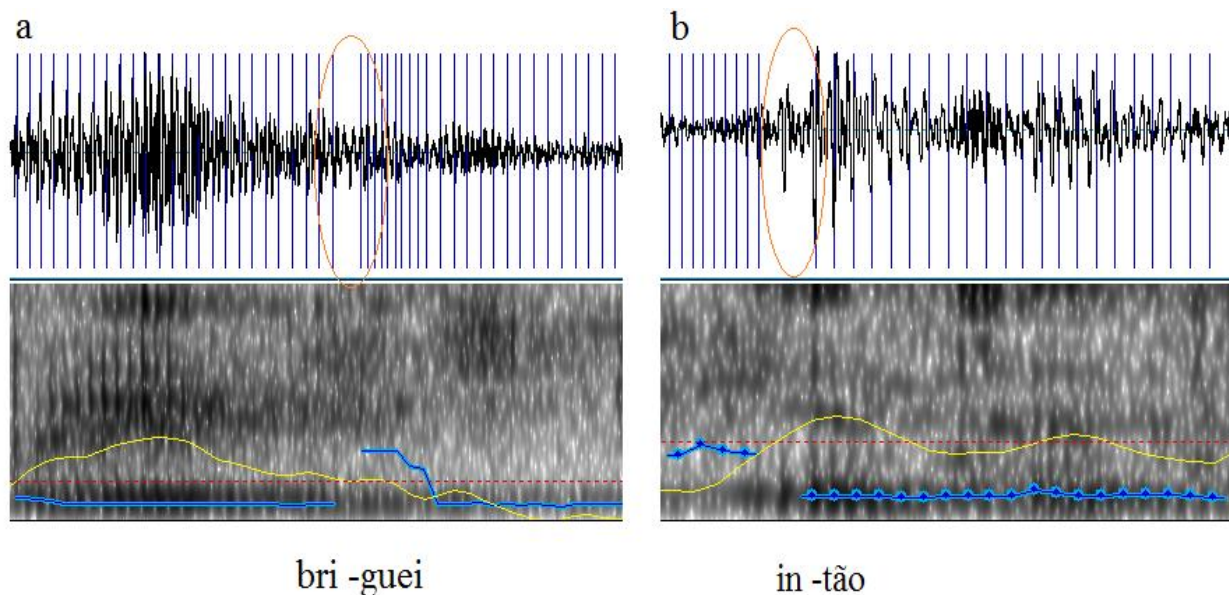
2) Bloqueios em consoante:

- 12 casos para a consoante fricativa labiodental [f];

- 4 casos para a consoante oclusiva velar [k];
- 2 casos para a consoante oclusiva bilabial [p];
- 2 casos para a consoante oclusiva bilabial [b];
- 2 casos para a consoante oclusiva alveolar [t];
- 1 casos para a consoante oclusiva velar [g].

Entre as vogais e consoantes, AP realiza mais momentos de BLO na vogal meio fechada anterior [e] e na consoante fricativa labiodental [f]. Normalmente, os BLOs costumam ocorrer em início de palavra. No entanto, para AP foi verificado dois casos em meio de palavra, um para a consoante oclusiva alveolar [t], em /in-tão/ e o outro para a consoante oclusiva velar [g], em /bri-guei/. Observe:

Figura 45 – Comparação entre os momentos de em meio de palavra para o informante gago AP.

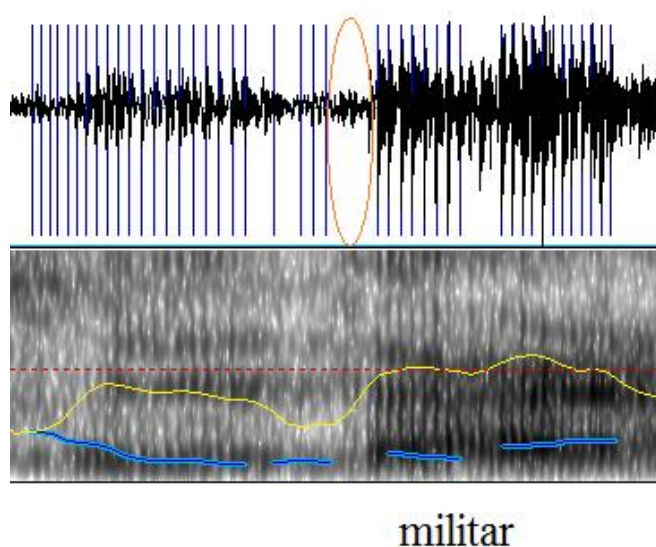


Fonte: própria.

A figura 45a e 45b exemplificam dois casos em que ocorre bloqueio no início da segunda sílaba da palavra. A fim de facilitar a visualização optamos por apresentar apenas os momentos de BLO de cada GT. A figura 45(a) representa o enunciado EN23 “Eu nunca bri brigue bri assi nunca bri-guei cum ninguém” e a figura 45(b) o enunciado EN26 “ in-tão eu fiz cursu militar”. O círculo em vermelho caracteriza o momento em que a disfluência ocorre. Observe que tanto para a figura 45a quanto para 45b há um corte na curva da F0 no momento do BLO. Em 45a, os pulsos também mostram essa quebra na estrutura, evidenciando a

diminuição do esforço vocal. A consoante [t], por sua vez, é naturalmente desvozeada, conseqüentemente, ela não teria presença de pulsos. No entanto, o BLO pode ser visualizado por meio da forma de onda mais irregular no momento da sua produção, diferentemente do que ocorre na mesma consoante sem BLO. Observe:

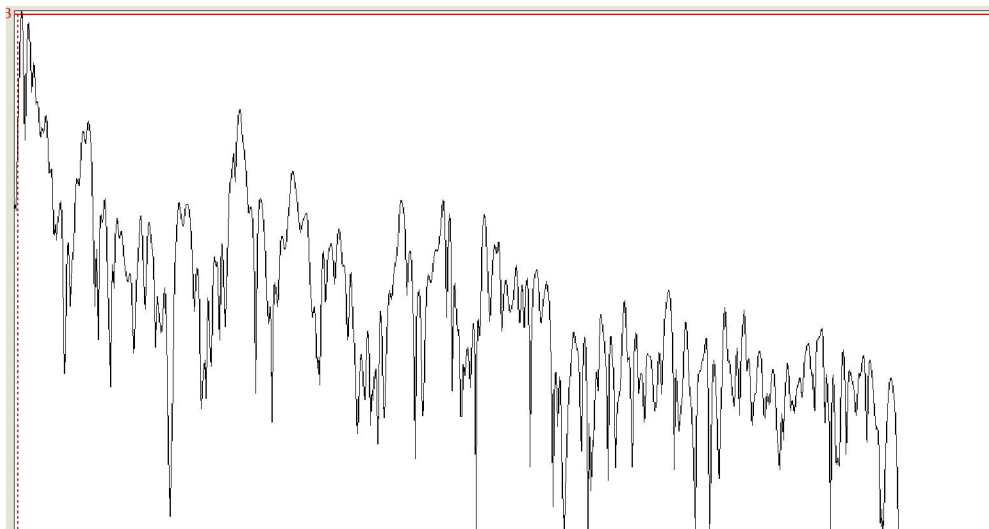
Figura 46 – Representação do segmento [t] sem BLO no enunciado EN26 “in-tão eu fiz cursu militar”.



Fonte: própria.

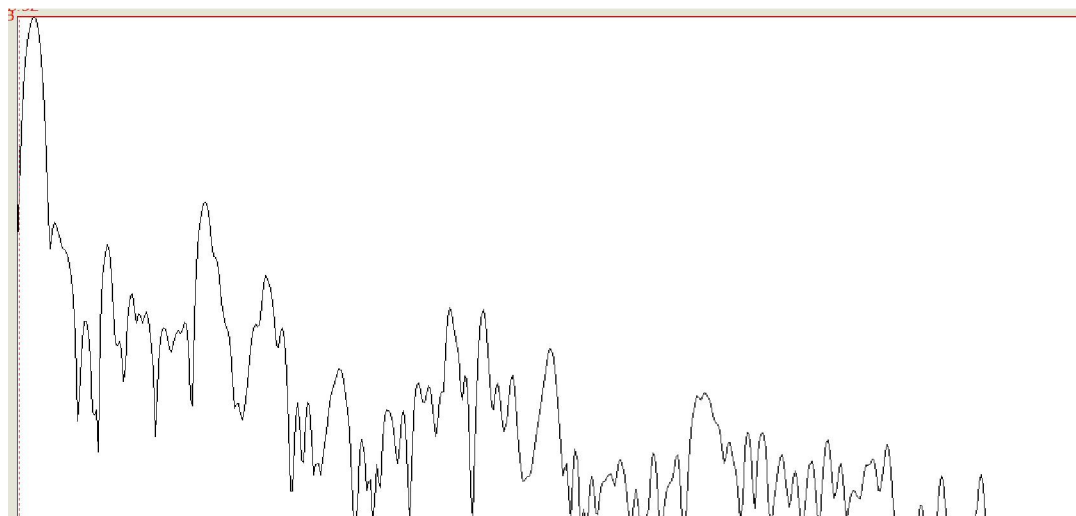
Ao compararmos as figuras 45b e 46 vemos que a forma de onda da primeira é mais irregular do que a da segunda. Vemos, ainda, que ocorre uma ruptura na curva da F0. No entanto, é um corte normal, que ocorre em momentos de desvozeamento. O espectro dos dois momentos caracteriza ainda melhor os dois segmentos.

Figura 47 – Espectro da forma de onda da consoante bloqueada [t] da palavra com disfluência/in-tão/.



Fonte: própria.

Figura 48 – Espectro da forma de onda da consoante sem bloqueio [t] da palavra sem disfluência /militar/.



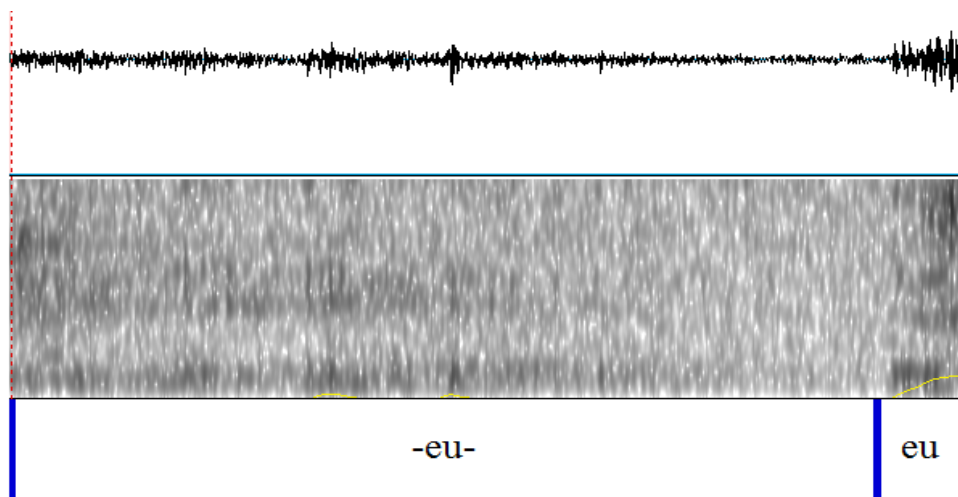
Fonte: própria.

A figura 47 representa o segmento bloqueado e a figura 48 o segmento sem a disfluência. Como podemos ver, na primeira figura o espectro varia mais do que na segunda, além de apresentar períodos maiores. É provável que essa seja uma maneira de evidenciar os BLOs quando não ocorrem em início de sílaba e, ainda, para segmentos vozeados e desvozeados. No entanto, fica difícil fazermos comparações, uma vez que esse fenômeno ocorreu apenas para AP.

Os BLOs em início de palavra variaram entre os tipos 1, 2 e 3. Não foram verificados nenhum BLO com PS respiratória (tipo 4). Os BLOs do tipo 1 foram os mais frequentes, ocorrendo em aproximadamente 49% dos casos. Em seguida, aparecem os BLOs do tipo 2 em 30% dos casos, aproximadamente e, por fim, os do tipo 3 em 21% dos casos. Em relação às vogais, a PS de bloqueio do tipo 2 foi a mais comum, sendo verificada para 60% dos casos, aproximadamente. Já para as consoantes, a PS de BLO mais frequente foi a do tipo 1, ocorrendo em 42% do total. Os BLOs apresentaram uma redução na forma de onda.

A análise acústica do BLOc simples mostrou uma média de duração dos segmentos com PS do tipo 1 de 0,379s e, das PS com bloqueio de 0,301s. Os segmentos bloqueados em consoante com PS do tipo 2 tiveram uma duração de 0,157s, enquanto que os segmentos de 0,473s. Já as BLOs complexas com PS do tipo 1 tiveram uma duração média de 0,370s e os segmentos de 0,445s. No mesmo contexto, as PS de tipo 2, duraram em média de 0,350s, enquanto que os segmentos sem bloqueio duraram em média 0,401s. Finalmente, as vogais com PS do tipo 1 (um caso) apresentaram uma PS de bloqueio média de 1,542s e, o segmento de 0,809s. Já as PSs de bloqueio do tipo 2 duraram em média 0,641s e, os segmentos, 0,189s. A média geral mostra que as PSs articulatórias são maiores (0,763s) do que as PS silenciosas (0,382s). É importante salientar que dentre os casos de PSs do tipo 2, incluindo vogais e consoantes, 54% delas apresentaram articulação. Observe na figura que segue:

Figura 49 – Representação de uma PS de bloqueio com presença de articulação, produzida pelo informante AP no enunciado EN32 “mais eu achu qui eu achu qui -eu achu eu achu qui cada coisa du #i vasupo eu separu cada coisa nu seu lugar”.



Fonte: própria.

No exemplo da figura 49, a PS de bloqueio está identificada pelo segmento /-eu-/. Nesse caso, ocorre uma PS do tipo 2, pois como podemos ver não há pulsos, bem como na produção do segmento /eu/, que foi bloqueado, uma vez que é desvozeado. Auditivamente, a

percepção de que ocorre articulação nesse momento de PS é mais fácil do que visualmente, uma vez que há a percepção da atuação das cordas vocais. Já visualmente, a articulação pode ser observada por meio da forma de onda que é mais reduzida, mas não deixa de apresentar variações. Em momentos de silêncio total ela quase não apresenta picos. E, ainda, identificamos essa articulação através do espectro, que mostra com menor intensidade a presença dos formantes um e dois na mesma linha do segmento /eu/.

A média da F0 para os três tipos de BLO foram: 107 Hz para o BLO em consoante; 83 Hz para o BLO em vogal e 84 Hz para o BLO complexo. Esses resultados mostram que o informante apresenta uma entoação mais alta quando o BLO ocorre em consoante e mais baixa quando ocorre em vogal. O informante gago PA também apresentou uma entoação mais alta nos bloqueios em consoante, mas mais baixa nos BLOs complexos. A análise dos pulsos mostra, no geral, que o esforço vocal para os três tipos de BLO é aproximado, sendo de 21 para o BLO em consoante; 22 para o BLO em vogal e, 24 para o BLO complexo. Mas, comparativamente, ele é um pouco maior na estrutura complexa.

Finalmente, em relação à intensidade, não levamos em consideração o momento de transição da fluência para disfluência, uma vez que a maior parte dos casos foi antecedida por uma PS de BLO. Em uma visão geral, a média da intensidade máxima não variou muito entre os três tipos de BLOs, sendo de 68 dB para o BLO complexo; 67 dB para o BLO em vogal e, de 66 dB para o BLO em consoante. Uma visão mais específica dos casos mostra, ao compararmos consoantes oclusivas e fricativas, que as últimas apresentam uma intensidade máxima um pouco maior (67,861 dB) do que as oclusivas (67,313 dB). Como vemos, auditivamente, essa diferença é imperceptível. Por outro lado, comparando as consoantes oclusivas bilabiais, alveolares e velares e, a fricativa labiodental, verificamos que a oclusiva velar [k] tem a maior intensidade (70,728 dB). Quanto às vogais, a maior intensidade máxima ocorre para a vogal aberta anterior [a]. Por fim, a análise da intensidade final revelou tanto para as vogais quanto para as consoantes um aumento, sendo de 70% para a BLO_v e de 65% para a BLO_c.

A análise do padrão entoacional revelou para as vogais (em BLO simples e BLO complexo) um padrão que varia entre o tom alto (50%) e baixo (50%) simples (*H e *L). Para o BLO em consoante (simples e complexo), foi verificado para a fricativa labiodental [v] um padrão entoacional que varia entre simples e composto e, entre alto e baixo. No entanto, a maior tendência é para o tom baixo (75%), tanto no padrão simples quanto no composto, uma vez que o último tem o pico entoacional no tom baixo (*L, *L+H). Os segmentos oclusivos

bilabiais [p, b] foram anotados em 100% dos casos com o tom simples baixo *L, bem como o oclusivo alveolar [t]. Por sua vez, o segmento velar [k] variou entre tons simples e compostos, com pico entoacional nos tons alto e baixo (*H+L, *L+H e *L). No entanto, a maior tendência foi para os tons compostos do tipo *L+H (50%) e pico entoacional no tom baixo (75%). Sendo assim, de uma maneira geral, os BLOs tendem a ter para AP um padrão entoacional do tipo *L.

Seguido do BLO, a INS representa 17,4% das produções. Dentre elas, ocorreram INSs em palavra (INS_p) e em vogal (INS_v). A INS_p foi a mais constante, com 80% das produções, enquanto que as vogais representaram 20% do total. Essa disfluência também foi classificada entre simples e complexas. No entanto, foram identificados apenas dois casos de INS_p. Sendo assim, a sua disposição no enunciado foi:

1. INS_p simples:

- 8 INS_p em fim de GT para a palavra /né/;
- 2 INS_p em meio de GT para a palavra /né/;
- 5 INS_p em meio de GT para a palavra /assim/;
- 2 INS_p em meio de GT para a palavra /ái/;
- 1 INS_p dupla em meio de GT para a expressão /intão aí/.

2. INS_p complexa:

- 1 INS_p complexa em início de GT para a palavra /intão/,
- 1 INS_p complexa em meio de GT para a palavra /intão/.

3. INS_v simples:

- 2 INS_v em início de GT para a vogal /é/;
- 2 INS_v em meio de GT para a vogal /é/;
- 1 INS_v em meio de GT para a vogal /i/.

Assim como para os outros informantes (gagos e não gagos), a INS_p do tipo /né/ em final de GT tem uma função de afirmação interrogativa retórica. Essa característica também foi observada para as INS_p /né/ em meio de GT. Veja:

▪ EN5 /é -o /é -v supo ansioso né -vamu supo -vamu supo -vamu supo quando a pessoa ti pressiona u ce/.

Como podemos ver, AP conclui o primeiro GT com essa INS_p, e continua a sequência discursiva, iniciando um novo GT. Já as INS_ps /assim/, /ái/ e /intão/ (simples e complexas), em início e fim de GT, indicam uma sequência de pensamento, bem como a INS_p dupla /intão aí/.

Em relação às vogais, a INS_v /é/ em início e fim de GT caracteriza uma afirmação positiva diante de um comentário do interlocutor. Por exemplo, quando perguntamos a AP se era verdade que a disfluência gaga ocorria com mais frequência em determinadas situações, ele responde afirmativamente por meio da INS_v /é/, como se fosse /sim/, como no exemplo:

- EN5 “é, va va vamu vasupo dependi a ocasião tamém né”.

Por outro lado, as $INS_{v,s}$ do tipo /é/ em meio de GT tem a mesma função de sequência discursiva das $INS_{p,s}$ /assim/ e /intão/ e /ái/. Já a INS_v /i/ tem a função de adicionar ideias. Observe:

- EN34 “mais eu achu qui eu achu qui -eu achu eu achu qui cada coisa du #i vasupo eu separu cada coisa nu seu lugar”.

Em EN34 AP começa a falar sobre determinado assunto, mas não consegue terminá-lo, devido aos vários bloqueios e repetições. Sendo assim, ele deixa a informação incompleta e adiciona por meio da INS_v outra ideia ao GT. Essa INS_v , que ocorre após uma PS silenciosa, ajuda o informante a não cometer mais disfluências, diferentemente do que foi observado para os outros tipos de INS s, que não têm essa mesma função.

Acusticamente, os parâmetros variaram de acordo com o tipo de INS . Sendo assim, para as vogais a duração foi maior para as $INS_{v,s}$ em início de GT /é/, com uma média de 0,362s. Na sequência, em meio de GT /i/ tem uma duração maior (0,276s), que as $INS_{v,s}$ do tipo /é/ em meio de GT, com uma média de 0,170s. A análise da média da F0 seguiu a mesma proporção da duração. No entanto, ela foi maior para as INS_v em meio de GT do tipo /é/, com uma média de 116 Hz do que a INS_v em meio de GT do tipo /i/ com uma frequência média de 111 Hz. Os pulsos, por sua vez, mantiveram a mesma proporção da F0 média. Quanto à intensidade, para as $INS_{v,s}$ em início de GT ela não foi analisada, bem como em meio de GT, uma vez que apenas uma INS_v não ocorreu após uma PS. Já em relação à intensidade inicial e final, em 60% dos casos houve um aumento na intensidade final. A intensidade máxima, por sua vez, foi maior em início de GT (78 dB), e menor para a INS_v em meio de GT do tipo /é/ (64 dB).

No que diz respeito às $INS_{p,s}$, chegamos aos seguintes resultados:

- INS_p do tipo /intão/ em início de GT – 0,627s;
- INS_p do tipo /né/ em fim de GT – 0,341s;
- INS_p do tipo /intão aí/ em início de GT – 0,301s;
- INS_p do tipo /assim/ em meio de GT – 0,266s;
- INS_p do tipo /intão/ em meio de GT – 0,266s;

- INSp do tipo /né/ em meio de GT – 0,220s;
- INSp do tipo /ái/ em meio de GT – 0,170s.

Os valores mostram que em início de GT a INSp /intão/ é maior do que as outras inserções. Depois, temos a INS do tipo /né/ em final de GT, seguida, respectivamente, das INSp /intão aí/, /assim/, /intão/ em meio de GT, /né/ em meio de GT e /ái/.

Quanto à F0 média, os resultados mostraram uma entoação maior para a INSp de estrutura /intão aí/ (130 Hz). Depois, aparecem em ordem decrescente as seguintes inserções:

- INSp do tipo /intão aí/ em início de GT – 130 Hz,
- INSp do tipo /intão/ em meio de GT – 119 Hz ;
- INSp do tipo /intão/ início de GT – 118 Hz;
- INSp do tipo /né/ em meio de GT – 116 Hz;
- INSp do tipo /né/ em fim de GT – 114 Hz;
- INSp do tipo /assim/ em meio de GT – 112 Hz;
- INSp do tipo /ái/ em meio de GT – 108 Hz.

Como vemos, a não ser pela INSp /intão aí/, em que ocorreu um erro de gravação, a análise da F0 média mostrou que a entoação tende a variar pouco. Em relação aos pulsos, os resultados seguiram as mesmas proporções obtidas para a F0 média.

Por fim, em relação à intensidade em início de enunciado, em 17 casos (59%) (deixando de lado os momentos iniciais e após PSSs), houve uma diminuição da fluência para a INSp, enquanto que em 41% deles houve um aumento. É importante destacar que para a INSp do tipo /ái/ houve um aumento na transição para as duas ocorrências. Já quanto à intensidade final e inicial, em 60% dos casos ela diminuiu, enquanto que em 40% deles aumenta. Vale salientar que para /intão/ em início e meio de GT e para a INSp /intão aí/ a intensidade final aumentou em relação à inicial em todos os casos. No entanto, como tivemos apenas uma ocorrência para todos os casos, os resultados serão melhor observados no conjunto dos dados. Finalmente, observamos uma intensidade máxima maior para a INSp /né/ em fim de GT (97 dB), seguida, respectivamente da INSp /intão aí/ (79 dB); /intão/ em meio de GT (73 dB); /intão/ início de GT (72 dB); /né/ em meio de GT (71 dB); /assim/ (68 dB) e, por fim, com a menor intensidade a INSp /ái/ (67 dB).

Quanto à análise fonológica, apresentaram um padrão igual na mesma posição no GT. Assim, as INSp do tipo /é/ em início de GT foram anotadas com um tom alto *H. Já em meio de GT com um tom baixo *L, bem como a INSp /i/ em meio de GT. Em relação às INSp do

tipo /né/ em fim e meio de GT, o padrão variou entre *L, *H e *L+H. Em 50% desses casos tivemos um tom alto *H e em 50% um tom baixo. No último caso, consideramos o tom composto com pico entoacional no tom baixo *L+H. Já para a INS_p do tipo /assim/, os tons variaram entre simples e compostos e, entre tons altos e baixos (*L, *H e *H+L). No entanto, em 60% dos casos a tendência foi para o pico entoacional no tom alto, seja ele simples ou composto *H e *H+L. Para /ái/, por sua vez, tivemos apenas tons simples, variando entre alto e baixo (*H e *L), bem como para /intão/. Finalmente, a INS_p dupla /intão aí/ foi anotada com o tom composto L+*H.

A PI, por sua vez, representa 4% das disfluências. Essa disfluência também ocorreu concomitantemente com dois outros eventos: COR (duas produções) e RPe (três produções). Sendo assim, junto com esses eventos elas representam aproximadamente 7% do total de disfluências. As PIs estão relacionadas à queda das palavras durante a fala, que diz respeito a processos rítmicos. Observe a seguir os dois exemplos:

- EN3 /é cas pessoa vem mi cobra, as própria pessoa vem mi cobra/
- EN25 /#a faiz tempu eu tinha eu tinha que é eu so é tinha dizo -é só tinha dizoitu eu tinha dizoitu anu/

As PIs são, respectivamente /cas/ (EN3) e /dizo/ (EN25). Em EN3 o informante reduz a produção /que as/ para /cas/. Já em EN25, ocorre uma queda que se deve a uma sobreposição de ideias. Observe que AP inicia o GT tentando lembrar a idade que tinha quando um incidente marcou sua vida. Quando ele lembra, ele apenas pronuncia o início da palavra /dizo/, sobrepondo a mesma ideia, mas agora de modo completo /-é só tinha dizoito/, reafirmando-a mais uma vez no final do GT /eu tinha dizoitu anu/.

Acusticamente foram analisadas 12 PIs, uma vez que em EN9 tivemos um caso de PI dupla em RPe, /depo/, /depo/. Desse modo, as PIs apresentaram uma duração que variou de acordo com o tamanho do segmento. Por exemplo, em EN8, a PI /pi/ teve uma duração de 0,100s, enquanto que em EN7, a PI /conhe/ teve uma duração de 0,330s. Os pulsos também mostraram um esforço vocal variado. Em relação à F0, tivemos duas produções sem sonorização (EN8 /pi/ e EN19 /fa/), e uma (EN3 /cas/) em que ocorreu um erro de gravação devido a uma produção mais aguda. Dentre os outros casos, a F0 média evidenciou uma tessitura nivelada para entoação, uma vez que a F0 mínima teve uma média de 74 Hz e a máxima de 102 Hz.

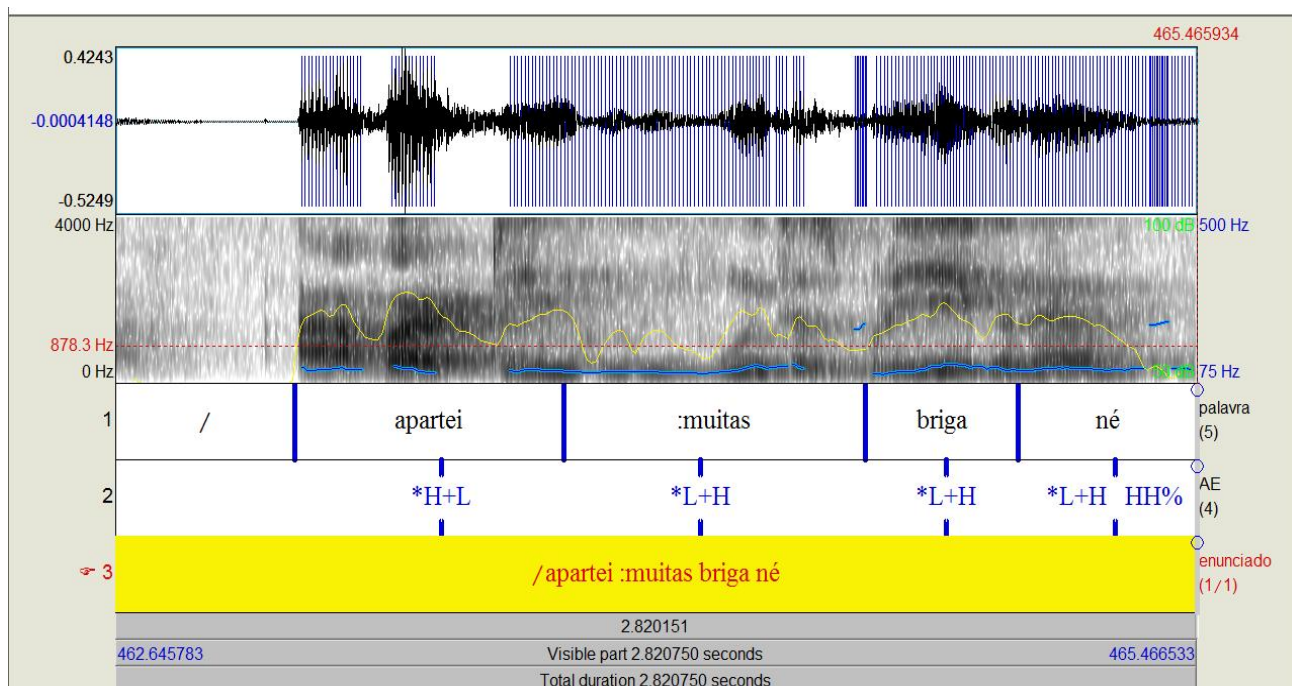
Quanto à intensidade, apenas EN13 ocorreu após uma PS. Todas as PIs foram identificadas em meio de enunciado. Sendo assim, a intensidade na transição revelou para

100% dos casos uma diminuição. Quanto à intensidade inicial e final, em 75% dos casos houve uma diminuição da intensidade inicial para a final. Em relação à intensidade máxima, a maior intensidade foi de 75 dB (PI simples) e a menor de 58 dB (PI complexa). Vale salientar que esses resultados não indicam que PIs complexas apresentam intensidade máxima menor do que PIs simples. Eles apenas evidenciam as diferenças verificadas. Dessa maneira, há uma tendência para a variação da intensidade máxima nas PIs.

Finalmente, a análise fonológica mostrou uma variação entre tons simples e tons compostos (*L, *H e *H+L). Dentre os exemplos, deixando de lado os casos sem sonorização, tivemos dois casos de tom baixo *L em PI (EN23 /brigue/ e *L /dizo/) e seis casos com o tom alto, sendo eles: EN7 /pala/; EN7 /conhe/ e EN24 /so/ com o padrão *H e, para EN13 /ormal/; EN9 /depo/ e EN29 /fa/, o padrão *H+L.

O alongamento também variou entre simples e complexos. Juntos eles representaram, aproximadamente 3,5% dos dados. Os alongamentos complexos ocorreram apenas com RP_e. Assim como foi verificado para os outros informantes (gagos e não gagos) o AL em final de palavra deve-se a situações em que o falante aumenta a duração do segmento, a fim de ter tempo para formular e/ou lembrar-se do final da mensagem. Já em início de segmento, ajuda o informante, nesse caso o AP, a evitar o BLOc. Esse caso ocorreu para o enunciado EN27 “/-vasupo eu já eu já eu já eu já eu já ap /apartei :muitas briga, né/”. Veja um fragmento da figura que identifica esse AL_p:

Figura 50 – Representação do AL_p inicial realizado pelo informante AP no enunciado “/apartei :muitas briga, né”.



Fonte: própria.

A figura 50 identifica o AL_p inicial para a palavra /:muitas/. Como vemos, AP alonga mais a primeira sílaba /:mu/ (0,551s) do que a segunda /tas/ (0,240s). A consoante nasal bilabial [m] é um segmento mais propício a ocorrer BLO, assim como os segmentos fricativos e oclusivos. Isso pode ter levado AP a alongar a sílaba inicial, a fim de não bloqueá-la. No entanto, essa estratégia não é intencional na fala do informante, uma vez que ele não alonga os outros segmentos que também são alvo de BLO. A figura também mostra, como veremos abaixo, uma intensidade que tem uma tendência a redução tanto na passagem do momento fluente para o disfluente quanto da intensidade inicial para a final. Além disso, por meio da curva da F0 também vemos uma tessitura com pouca variação entoacional.

Os ALs simples ocorreram em palavra (AL_p) e vogal (AL_v). O primeiro tipo ocorreu em início de sílaba /:muitas/ (EN27) e final de sílaba /dependi:/ (EN11). Já o segundo tipo ocorreu apenas para a vogal /é:/. Os ALs complexos foram identificados em palavra, mais especificamente em final de sílaba /tava:/ (EN16). Vale salientar que após os ALs de palavra o informante sempre volta à fluência. Por outro lado, nos ALs de vogais o informante apresenta momentos de disfluência. No caso de AP, após as duas AL_vs temos uma pausa

(EN21) e um bloqueio (EN14). Isso indica, que o AL de vogal não ajuda o informante a retomar a fluência.

Em relação à duração, quanto ao alongamento de palavra, o segmento /dependi:/ (EN11) durou 1,023s e a sílaba alongada 0,666s. Já o ALp em início de sílaba /:muitas/ (EN27) apresentou uma duração total de 0,785s e, a sílaba alongada durou 0,551s. Por fim, o ALp complexo /tava:/ (EN16) apresentou uma duração total de 0,751s e, a apenas a sílaba alongada de 0,431s. Como vemos, os momentos de alongamento duram mais de 50% do total da palavra, sendo, respectivamente, 65% para EN11; 52% para EN27 e, 57% para EN16.

No ALv as durações também são expressivas, sendo de 0,889s para EN14 e, 0,775s para EN21. Quanto à média da F0, AP mantém a tendência ao nivelamento da entoação. Fato que é justificado pela média da F0 mínima, de 100 Hz e máxima, de 123 Hz. Por fim, a análise da intensidade mostrou para 100% dos casos de AL_p uma diminuição na intensidade, na transição da fluência para a disfluência. Quanto à intensidade inicial e final, apenas para o ALp complexo houve um aumento da intensidade final, em relação à inicial. Para os AL_ps simples a intensidade final diminuiu. Para os ALs em vogais, dos dois casos, um deles ocorreu após uma PS (EN21) e, no outro, houve um aumento na transição da fluência para a disfluência (EN14). Já em relação à intensidade inicial e final, para os dois casos a intensidade final diminuiu. Por fim, houve uma variação pequena para intensidade máxima, sendo a diferença entre a maior (76 dB) e a menor (72 dB) intensidade de 4 dB.

No que diz respeito à análise fonológica para os AL_ps, todos os casos apresentaram um tom composto. No entanto, em fim de palavra o pico ocorreu em tom alto L+*H (EN11) e *H+L (EN16). Já para o AL observamos que em fim de palavra o pico entoacional ocorreu para tom baixo *L+H. Por fim, para as vogais, o padrão variou entre simples *L (EN21) e composto *H+L.

A disfluência do tipo COR representou 2% do total das disfluências. Em todas elas o informante optou por manter parte da informação do GT com erro na informação corrigida. Observe os enunciados:

- EN3 */é cas pessoa vem mi cobra, as própria pessoa vem mi cobra/*,
- EN24 */agora uma vei qui mi deu um so -um tapa já tento mi da um tapa eu catei u braço deli quebro quasi qui quebrei u braço deli, né/*.

Em EN3 temos uma COR “é cas pessoa vem mi cobra > as própria pessoa vem mi cobra”. Já em EN24 temos uma COR dupla “um so > um tapa” e, “eu catei u braço deli quebro > quase qui quebrei u braço deli”. Nos três exemplos AP traz uma informação do

segmento que deve ser corrigido na correção. Em EN3 a COR é enfática, o informante, ao falar sobre a sua disfunção, enfatiza que não é apenas ele que se preocupa em fazer um tratamento para melhor a gagueira, mas as /próprias pessoas/. É importante salientar que tivemos duas CORs complexas: uma em EN3 que ocorreu junto com PI e, a primeira COR de EN24, que ocorreu com PI e BLO_v.

Dentre os três casos, apenas EN3 iniciou o erro no início do GT. Os outros ocorreram em meio de GT. A análise da duração mostrou 100% dos eventos uma diminuição da duração no momento da correção. Por exemplo, em EN3 o erro teve uma duração de 1,343s enquanto que a COR teve uma duração de 1,074s. A análise da F0 média também revelou que na correção o informante tende a diminuir a entoação nos três casos analisados. Fato que também foi observado para os pulsos, ou seja, o esforço vocal é maior no erro. Quanto à intensidade, no momento de transição, foram analisados dois casos para o erro, uma vez que EN3 iniciou o GT e três para a correção. Dos momentos analisados, não houve um padrão, pois para 50% dos dados foi observado um aumento e diminuição na intensidade. Já para a COR, em 66% dos casos houve uma diminuição da intensidade e em 33% um aumento. Isso mostra uma maior tendência para a diminuição da intensidade na passagem do erro para a COR. A intensidade final, em relação à inicial tem uma tendência maior para o abaixamento (66%) dos casos. Por fim, comparando a intensidade máxima nos momentos do erro e COR, observamos que a intensidade é sempre maior no erro. No entanto, essa diferença é pequena, sendo a maior de 4 dB (EN3) e a menor de 0,515 dB (EN24, primeira COR). Isso mostra que o falante tende a manter uma intensidade com pouca variação nos dois momentos de fala.

Quanto à análise fonológica, como é de costume, os padrões são variados. Apenas para EN3 tivemos o mesmo pico entoacional. Mesmo assim, eles variaram entre simples (*L) e compostos (*L+H).

Por fim, depois da COR, o FC foi a disfluência com menor número de ocorrências (1,4%). Ela foi identificada para os enunciados EN28 e EN34. Veja:

- EN28 “#vasupo /vasupo seu *fu* intão purqu -eu so eu -eu so /eu so eu so uma pessoa du bem”,
- EN34 “mais eu achu qui eu achu qui -eu achu eu achu qui cada coisa *du* #i vasupo eu separu cada coisa nu seu lugar”.

Nos dois exemplos, o FC aparece em itálico /fu/ e /du/, respectivamente. Para os informantes não gagos observamos que o FC ocorre simplesmente porque o informante optou por mudar de assunto, deixando o segmento incompleto ou, devido a uma sobreposição de

ideias. Para os informantes gagos, a nossa suposição é a de que eles também deixem o segmento incompleto não só por uma questão de esquecimento ou sobreposição de ideias, mas também para evitar a disfluência gaga. Como vemos, para a AP, o FC ocorre em /fu/ e /du/, segmentos fricativo [f] e oclusivo [d] que tem uma maior tendência para o BLO. Observe que nos dois enunciados há uma modificação no GT após o FC.

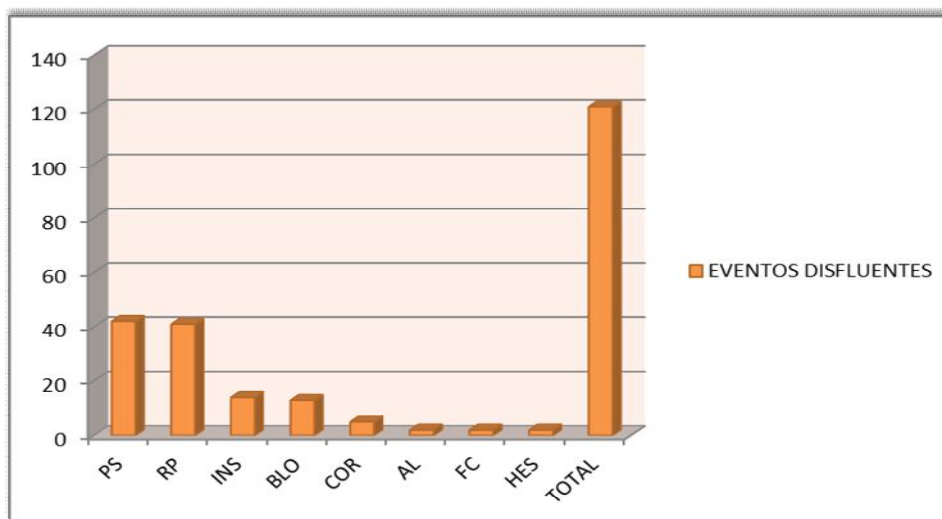
A análise acústica do FC mostrou uma duração variada para os dois casos, sendo EN34 maior (0,368s) do que EN28 (0,243). A média da F0 mostrou que EN34 tem uma frequência mais alta do que EN28. No entanto, a análise da tessitura indica, mais uma vez, uma frequência que é nivelada pela entoação (F0 mínima 127 Hz e, F0 máxima 137 Hz). Os pulsos também variam conforme o esforço vocal, sendo maior para EN34.

Por fim, a intensidade tanto no momento de transição da fluência para a disfluência quanto em relação à intensidade inicial e final, revelou para 100% dos casos uma tendência ao abaixamento. A análise da intensidade máxima também varia, sendo maior para EN34. No que diz respeito à análise entoacional, os dois casos apresentaram um tom alto, mas EN28 foi analisado com um tom composto *H+L e EN34 com um tom simples *H.

O informante MAU

MAU realizou 121 disfluências em 58 enunciados. As disfluências complexas representaram aproximadamente 13% dos eventos, enquanto as simples, 87%. Para MAU, o máximo de disfluências por GT foi de seis, sendo duas do tipo PS. A sua taxa de elocução foi de 3,349sil/ms, a de articulação de 4,148sil/ms, e apresentou um tempo de articulação aproximado de 139ms. A seguir, o gráfico 14 representa a distribuição dos eventos disfluentes na fala desse informante.

Gráfico 14 – Distribuição dos eventos disfluentes realizados pelo informante gago MAU.



Fonte: própria.

A fala de MAU caracterizou-se por apresentar oito tipos de disfluências, dispostas, em ordem, da seguinte forma: PSs; RPs, INs; BLOs; CORs; ALs; FCs e HESs. É importante salientar que MAU não realizou nenhum tipo de PI isolada. No entanto, ela foi verificada em disfluências complexas junto com CORs e RP_es.

As PSs foram as mais frequentes para MAU, representando 35% dos eventos. Dentre elas, a mais comum é a de final de GT (57%), seguidas das pausas de início de GT (31% aproximadamente) e, por fim, das pausas em meio de enunciado (8% aproximadamente). Dentre as PSs, 64%, aproximadamente, antecederam eventos disfluentes.

Quanto à duração, as PSs de fim de enunciado também foram as que apresentaram a maior duração (13,020s). Na sequência, aparecem as PSs em início de GT (6,049s) e, por fim, as de meio de GT (4,573s). Entre PSs silenciosas e respiratórias, o segundo tipo apresentou a maior duração (12,242s), enquanto que as silenciosas, uma duração de 6,827s.

Seguido das PSs, a disfluência com o maior número de ocorrência foi a RP em 34%, aproximadamente, dos eventos. Os tipos de RPs realizadas foram:

- RP_{ps} (17%);
- 2RP_p (30%);
- RP_c (34%);
- RP_e (17%);
- 3RP_p (2%).

As repetições de palavras foram as mais frequentes, sendo verificadas em aproximadamente 95% dos casos. As RPs de vogal não apareceram isoladamente. Elas foram verificadas apenas em 5% dos casos de repetições complexas (RP_{c,s}). Quanto às repetições duplas de palavras (2RP_p), um caso ocorreu em palavra dissílaba e outro em palavra trissílaba. Já entre as repetições de sílaba, tivemos uma repetição do tipo RP_{ps2}, ou seja, repetição de sílaba dupla, em /via viajando/ e um caso do tipo 2RP_{ps} em /fa fa fala/. No que se refere às repetições complexas, tivemos os seguintes casos:

- 2RP_e + PI – três casos;
- 2RP_e+PS – um caso;
- RP_{ps}+PS – um caso;
- RP_{ps2}+PS – um caso;
- 2RP_p+PS – um caso;
- 2RP_e+PI+BLO_c – um caso;
- 2RP_e+BLO_c – um caso;
- 2RP_p+BLO_c – um caso;
- RP_{ps}+BLO_c+PS – um caso;
- 2RP_v + AL_v+ PS – um caso;
- 3RP_e+PI+PS – um caso;
- 4RP_v+PS – um caso.

É importante salientar que os BLOs nas repetições apresentaram uma PS articulatória e o AL_v ocorreu na segunda vogal repetida. Além disso, vale lembrar que, de um modo geral, os erros de gravação gerados por uma qualidade de voz mais aguda ou devido a algum tipo de interferência, bem como os segmentos desvozeados ou com uma qualidade de voz do tipo *creaky voice*, foram identificados tanto na primeira quanto na última RP.

Sendo assim, no que diz respeito à análise acústica da duração observamos as seguintes tendências:

- Para as repetições do tipo 2RP_p (58%), RP_{ps} (66,6%) e 2RP₂ (71%) ocorre uma diminuição na duração na passagem para a segunda RP. Vale salientar que as RP_{ps2} mantiveram a mesma tendência. No entanto, a 2RP_{ps} (um caso) apresentou uma tendência para o padrão do tipo onda;
- As RP_{c,s}, por sua vez, apresentaram uma tendência para o aumento da duração na segunda estrutura em 66,6% dos casos. A repetição quádrupla de vogal seguiu a mesma tendência, uma vez que da primeira para a segunda RP a intensidade aumentou e, da terceira

para a quarta também. Já a repetição tripla de estrutura apresentou um padrão onda, iniciando com uma duração maior, mas que diminui na segunda RPe e, na terceira aumentou mais uma vez;

- A $3RP_p$, por sua vez, apresentou um padrão de duração decrescente.

No que se refere à F0 média e aos pulsos os resultados foram:

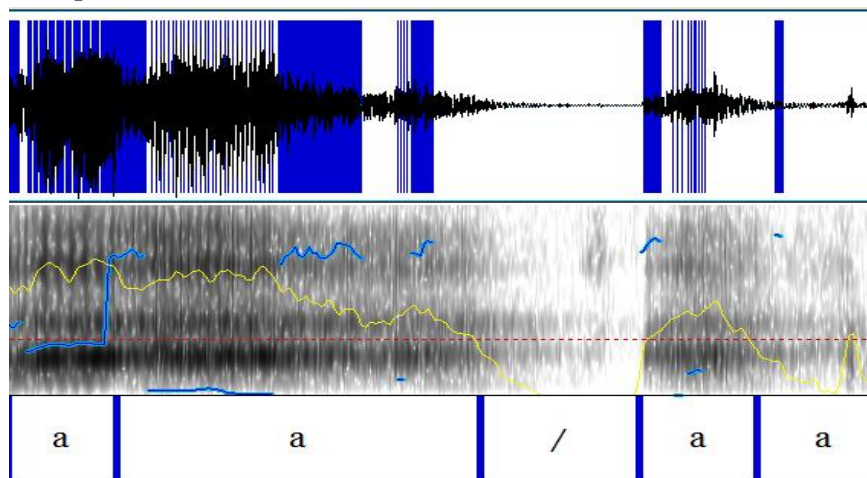
- Para a repetição do tipo $3RP_p$ ocorre um padrão de onda, uma vez que a F0 inicia-se mais alta, diminui e aumenta na terceira repetição. Já os pulsos diminuíram ao longo das repetições;

- As repetições do tipo $2RP_p$ (60%) e RP_{ps} (66,6%) apresentaram uma diminuição na F0 média na segunda RP. Quanto à $2RP_{ps}$, a F0 média aumentou ao longo das repetições. Os pulsos seguiram a mesma proporção. Desse modo, o esforço vocal foi menor na segunda RP. Já para a $2RP_{ps}$ o esforço vocal foi menor na primeira repetição, aumentou na segunda, diminuindo novamente na terceira RP;

- A $2RP_e$ apresentou um aumento da frequência na segunda estrutura em 71% dos casos. Já os pulsos mostraram um esforço vocal menor para a segunda repetição;

- Finalmente, as RP_{cs} revelaram uma tendência para a diminuição da frequência em 62,5% dos casos. Os pulsos, por sua vez, revelaram que o esforço vocal foi maior na segunda repetição em 62,5% dos casos. Essa tendência diz respeito às repetições complexas duplas. Para a repetição tripla, a F0 média revelou um padrão decrescente e os pulsos, um padrão de onda, sendo o esforço vocal maior na primeira repetição. Já a repetição do tipo $4RP_v$ apresentou um padrão melódico apenas na primeira RP, mas mesmo assim com indicação de erro. Observe o trecho do enunciado referente a essas repetições:

Figura 51 – Repetição do tipo 4RP_v realizada por MAU no enunciado EN26 “#aí você começa a a a a pensa”.



Fonte: própria.

A figura 51 representa parte do enunciado completo EN26 “#aí você começa a a a a pensa”. Observe que entre as RPs ocorre uma PS_e respiratória. Por isso, essa disfluência enquadra-se em um evento complexo. Como podemos ver por meio da F0, em todos os segmentos repetidos há uma interferência na gravação. Como dissemos, a primeira RP permitiu a análise da F0 média. No entanto, isso só foi possível ao deixarmos de lado, durante a seleção, o erro. A partir disso, a F0 média anotada para esse segmento foi de 184 Hz. Nas outras RPs essa seleção não foi possível. A segunda e a terceira repetição apresentaram uma qualidade de voz do tipo *creaky voice* e a última, um segmento desvozeado. A análise da duração mostra, por sua vez, uma proximidade entre a primeira (0,313s), terceira (0,339s) e quarta (0,383s). Por outro lado, a segunda RP varia significativamente entre elas, apresentando uma duração de 1,048s. Fato que sugere um segmento alongado. A PS_e entre as RPs durou 0,459s, revelando uma diferença de 0,076s entre a última RP. Isso sugere que o informante faz uma pausa que tem praticamente o mesmo tamanho das três RPs.

Por fim, quanto à intensidade de transição, final e inicial foram verificados os seguintes padrões:

- A repetição do tipo 3RP_p apresentou uma intensidade de transição que aumentou nas duas últimas RPs. Já a intensidade final, diminuiu nas duas primeiras e aumentou na última RP. Por fim, a intensidade máxima aumentou no decorrer das repetições, medindo, respectivamente, 72 dB, 76 dB e 84 dB;
- A repetição de estrutura (2RP_e) apresentou uma intensidade de transição que diminuiu em 100% dos casos na passagem da fluência para a primeira RP e que aumentou em

71% dos casos na transição entre as duas disfluências. As intensidades final e máxima, por sua vez, diminuíram respectivamente em 57% e 85% dos eventos;

- A $2RP_p$, por sua vez, apresentou uma diminuição na transição da fluência para a disfluência em 100% das repetições. Por outro lado, na transição entre as disfluências, a intensidade aumentou em 58% dos eventos repetidos. A intensidade final, por sua vez, também aumentou em 62,5% dos eventos. Quanto à intensidade máxima, não foi possível estabelecer padrões, pois as intensidades mantiveram-se em 50% dos casos altas e em 50% baixas;

- Em relação às repetições complexas, no que concerne aos eventos duplos, a intensidade de transição diminuiu tanto na transição da fluência para a disfluência (100%) quanto entre as disfluências. A intensidade máxima, por sua vez, aumentou em 58% dos casos e, a intensidade final apresentou a mesma impossibilidade de padronização das $2RP_{ps}$. Em relação à repetição quádrupla, a intensidade de transição diminuiu nos dois primeiros e no último segmento. O terceiro antecedeu uma PS, (cf. fig. 82). A intensidade final, por outro lado, aumentou em relação à inicial, nas quatro RPs. Por fim, a intensidade máxima diminuiu ao longo das repetições: 83 dB; 65 dB; 63 dB e 54 dB. Quanto à repetição tripla, não foi possível analisar a intensidade de transição, uma vez que ela iniciou o GT e as outras duas RPs foram antecedidas por PSs. A intensidade final, por sua vez, aumentou nos três eventos. Já a intensidade máxima diminuiu da primeira para a segunda RP e aumentou da segunda para a terceira RP;

- Finalmente, as RPs de sílaba apresentaram uma diminuição da intensidade nos dois momentos de transição, em 80% dos casos na passagem da fluência para a disfluência e em 83% na transição entre os eventos disfluente. A $2RP_{ps}$, por sua vez, iniciou o GT na primeira RP, na segunda caracterizou-se por uma diminuição da intensidade e, na terceira, que indica a sílaba em palavra, por um aumento. Quanto à intensidade final, a intensidade aumentou em 58% das RPs de sílaba e, na $2RP_{ps}$, aumentou nos três momentos. Finalmente, a intensidade máxima equiparou-se para as RP_{ps} . Já para a $2RP_{ps}$, ela aumentou progressivamente durante as produções.

Por fim, a análise entoacional manteve a tendência do mesmo padrão na maior parte das repetições. Observe a análise feita a seguir:

- As RP_{ps} mantiveram o mesmo padrão em 71% das repetições, variando entre *L+H e *H+L. Os dois padrões que variaram mantiveram o mesmo pico entoacional, de modo que a variação centrou-se em tons simples e compostos. Por exemplo, uma das sílabas

repetidas foi anotada com o tom *H e, a sílaba em palavra apresentou o tom *H+L. A repetição do tipo 2RP_{ps} e RP_{ps2} manteve o mesmo padrão entoacional;

- A 3RP_p foi analisada com o tom alto simples *H nos três momentos;
- Em relação às 2RP_{ps} as repetições mantiveram o mesmo padrão fonológico em 83% dos eventos. Dentre as duas RPs que não apresentaram o mesmo padrão, uma delas manteve o pico entoacional (*H *H+L) e a outra não (*L+H *H+L);

- A 2RP_e confirmou a tendência para 85% dos dados. A única repetição de estrutura que não conservou o mesmo tom também não manteve o pico entoacional, sendo analisada com o padrão *H+L *L+H;

- Por fim, a repetição complexa, mesmo apresentando mais de uma disfluência no segmento, confirmou a tendência para 64% dos eventos. Os padrões diferentes foram verificados para a 4RP_p, (cf. fig. 82), em que apenas a primeira RP apresentou uma F0, e mesmo assim, com interferência. Além dessa RP, a repetição tripla de estrutura diferenciou-se das outras produções apenas na primeira repetição (*H *L+H *L+H). Os outros três eventos que não mantiveram o padrão ocorreram para a RP_{ps}+PS, RP_{ps}+BLOC+PSe e 2RP_v + AL_v+PS.

Depois das RPs, as INSS aparecem em 11,6% dos eventos disfluentes. Elas dividiram-se INSS_{ps} (64%) e INSS_{vs} (36%). As INSS realizadas por MAU estão assim distribuídas:

- INSS_p em início de GT - /intão/, /ái/;
- INSS_p em fim de GT - /né/;
- INSS_p em meio de GT - /ái/, /intão/;
- INSS_v em início de GT - /é/, /i/;
- INSS_v em meio de GT - /é/, /i/.

As inserções realizadas por MAU não se diferenciaram semanticamente das analisadas para os outros informantes. Sendo assim, os segmentos /i/ e /é/ em início e meio de GT estão relacionados com a ideia de sequência discursiva, assim como as INSS_{ps} em início e meio de GT /ái/ e /intão/. A INSS_p em final de enunciado do tipo /né/ mantém a função semântica de afirmação retórica.

Por meio da análise acústica, a maior duração foi anotada para as INSS_{ps} em final de GT (0,559s). Em seguida, temos as INSS_{ps} em meio de GT (0,485s); as INSS_{ps} em início de GT (0,388s); as INSS_{vs} em início de GT (0,371s) e, as INSS_{vs} em meio de GT (0,321s). Quanto à F0 média, as INSS_{vs} em meio de GT foram as que apresentaram a maior frequência (185 Hz). Na sequência, aparecem as INSS_{ps} em meio de GT (175 Hz); as INSS_{ps} em fim de GT (168 Hz); as

INS_{p,s} em início de GT (167 Hz) e, por fim, as INS_{v,s} em início de GT (146 Hz). Como podemos ver, a diferença entre as frequências médias não é significativa entre as INS_{p,s} de início e fim de GT. A análise dos pulsos, por sua vez, identificou um maior esforço vocal para as INS_{p,s} em fim de GT. Depois, temos as INS_{v,s} em meio de GT; as INS_{p,s} em início de GT, as INS_{v,s} em início de GT e, as INS_{p,s} em meio de GT.

A análise da transição da intensidade não foi observada para as INS_{p,s} e INS_{v,s} em início de GT e, algumas, foram antecedidas por PSs. Sendo assim, no momento de transição, a intensidade equiparou-se para as INS_{p,s} em fim de GT. Em relação às INS_{p,s} de meio de GT, o segmento /intão/ foi antecedido por PS e o segmento /ái/ apresentou uma diminuição da intensidade. Já as INS_{v,s} em meio de GT apresentaram uma diminuição da intensidade em 100% dos casos. A intensidade final, quando comparada à inicial, revelou uma diminuição da intensidade para as INS_{p,s} em início (80%), meio (100%) e fim (100%) de GT. Já as INS_{v,s} em meio de GT apresentaram uma tendência para o aumento da intensidade em 66,6% dos casos e, as INS_{v,s} em início de GT apresentaram um padrão que se equiparou entre o aumento e diminuição da intensidade.

Finalmente, quanto à intensidade máxima, a média dos resultados mostrou que a maior intensidade ocorre para a INS_p em meio de GT (85 dB). Em seguida, temos as INS_{p,s} em início de GT (84 dB); as INS_{v,s} em meio de GT (82 dB); as INS_{v,s} em início de GT (81 dB) e, com a menor intensidade, as INS_{p,s} em fim de GT (79 dB). A diferença entre as inserções, seja de palavra ou vogal, não é significativa entre os segmentos.

A análise dos padrões fonológicos também confirmou a tendência das INSSs manterem o mesmo padrão entoacional para a maior parte dos segmentos. Os padrões obtidos para as inserções foram:

- /intão/ e /ái/ em início de GT – *L+H para 100% dos casos;
- /né/ em final de GT – foram os únicos segmentos do mesmo tipo que variaram, apresentando os tons *L+H e *H;
- /ái/ em meio de GT – *H+L (caso único);
- /intão/ em meio de GT – *L+H (caso único);
- /é/ e /i/ em meio de GT – O primeiro segmento foi analisado com o tom *L e o segundo com o tom *H;
- /é/ e /i/ em início de GT – *H para 100% dos casos.

Após as INSSs ocorre a disfluência do tipo BLO em 11% aproximadamente das disfluências. Os BLOs foram verificados em consoantes e vogais, sendo o mais frequente em

consoante (77% dos eventos aproximadamente). Para as vogais, eles ocorreram em vogal anterior meio fechada [e] e vogal anterior aberta [a].

Já para as consoantes eles foram verificados em:

- nasal bilabial [m] – três realizações nos segmentos /-muitus /, /muitu/ e /-meu/;
- oclusiva velar [k] – três realizações nos segmentos /-curta/, /-casa/ e /caçadu/;
- oclusiva bilabial [p] – duas realizações nos segmentos /-perde / e /-pressona/;
- oclusiva alveolar [d] – uma realização no segmento /-dessis/;
- fricativa alveolar [s] – uma realização no segmento /-sabi/.

A relação acima mostra que é mais comum para MAU, assim como observado para VA, realizar bloqueios em consoantes oclusivas.

Quanto ao tipo de PS de bloqueio, para o BLO_v foi verificado um evento sem pausa para a vogal [e] e duas PSs articulatórias nos dois casos que tiveram a vogal [a] bloqueada. O BLO_c, por sua vez, apresentou para 60% dos casos bloqueio sem PS, para 30% bloqueio com PS articulatória e, para 10% um bloqueio do tipo dois (sem pulsos, mas com articulação). Quanto à duração dessas pausas, em BLO_v, elas foram maior do que os segmentos. Por outro lado, para o BLO_c a PS foi maior apenas em oclusiva alveolar.

A análise mais específica dos BLOs em vogal mostra uma maior duração para o segmento /-e/ (0,533s), enquanto que a média dos segmentos do tipo /a/ foi de 0,138s. A F0 média segue a mesma tendência, sendo maior para o bloqueio em /-e/ e menor para o BLO em /-a/. Fato que também foi verificado para os pulsos, revelando o maior esforço vocal para /-e/. Quanto à intensidade de transição, apenas o segmento /-e/ foi analisado, uma vez que os bloqueios em /-a/ antecederam uma pausa de bloqueio. Nesse caso, a intensidade aumentou na transição da fluência para o BLO. Já a intensidade final diminuiu em /-e/ e variou em /-a/, uma vez que em um evento aumentou e no outro diminuiu. Finalmente, a intensidade máxima não seguiu a mesma tendência observada para os outros parâmetros, de modo que foi maior para /-a/. Mesmo assim, a diferença entre os dois tipos de segmentos (/a/ 86 dB e /-e/ 81 dB aproximadamente) foi pequena, de 5 dB.

No que diz respeito à análise acústica dos BLOs em consoante, a maior duração foi identificada para a consoante /-d/ com uma duração de 0,998s. Na sequência, aparecem as consoantes do tipo /-p/ (0,900s); /-m/ (0,857s); /-s/ (0,703s) e /c/ (586s). Quanto à F0 média, a maior frequência ocorre para o segmento /-m/ (196 Hz). Depois, temos os segmentos do tipo /-s/ (198 Hz); /-d/ (183 Hz); /-p/ (180 Hz) e /c/ (138 Hz). Em relação aos pulsos, o maior

esforço vocal foi identificado para as consoantes desvozeadas bilabiais [p]. Em seguida, aparecem, em ordem decrescente, os segmentos do tipo /-m/; /-d/; /-s/ e, por fim /-c/.

Em relação à intensidade de transição, 1) os segmentos /-s/ e /-d/ foram antecidos por uma PS de bloqueio; 2) o bloqueio em consoante /-p/ apresentou um caso com PS de bloqueio e, o outro caso apresentou uma diminuição na intensidade; 3) o bloqueio em consoante /-m/ variou, de modo que em um caso a intensidade aumentou e no outro diminuiu e, por fim 4) o único segmento do tipo /-c/ apresentou uma diminuição da intensidade. A intensidade final, por sua vez, aumentou em 100% dos casos nos bloqueios em /-c/, /-d/ e /-s/. Já em /-m/ a tendência foi para o abaixamento em 66,6% dos eventos. Por sua vez, em /-p/, em um segmento a intensidade final aumentou e no outro diminuiu.

Quanto aos padrões entoacionais, os segmentos iniciados com uma oclusiva bilabial surda [p] foram anotados com o padrão *H+L. Já com os iniciados com a nasal bilabial [m] o padrão entoacional variou entre os tons compostos *H+L e *L+H (66,6%), assim como foi anotado para os segmentos iniciados com as oclusivas velares sonoras [k]. Por sua vez, os iniciados em oclusiva alveolar sonora [d] e fricativa alveolar surda [s] foram anotados com o tom composto *L+H. Os padrões entoacionais verificados para os BLOs em vogal foram *H+L, para a vogal anterior meio fechada [e] e, *H+L e *L+H para a vogal anterior aberta [a].

As CORs caracterizaram 4% do total de disfluências. Elas dividiram-se entre CORs simples (60%) e complexas (40%). As CORs complexas ocorreram junto com PI e AL_v (um caso) e BLO_v (um caso). A COR com BLO_v ocorreu para o segmento /-u/, sem PS de bloqueio. Dentre as CORs, o informante manteve parte da informação contida no erro na correção em 60% dos eventos.

A análise acústica revelou tanto para a COR simples quanto para a complexa um aumento da duração na correção em 100% dos casos. Quanto à F0 média, para a COR simples, a frequência diminuiu no momento da correção em 66,6% dos eventos. Já para a COR complexa, ocorreu um aumento da F0 em 100% dos casos. Em relação aos pulsos, o esforço vocal aumentou durante a correção, tanto para a COR simples quanto para a complexa em 100% dos eventos.

Por sua vez, a análise da intensidade mostrou, deixando de lado os momentos de COR em início de GT (um caso em COR simples), uma diminuição da intensidade na passagem da fluência para o erro em 100% dos casos para a COR simples. Ainda em relação à COR simples, na transição do erro para a correção, a intensidade também diminuiu, mas nesse caso,

em 66,6% dos eventos. Já em relação à COR complexa, não foi possível estabelecer padrões no momento de transição da fluência para o erro, nem do erro para a correção, uma vez que nos dois momentos a intensidade aumentou e diminuiu em 50% dos dados. No que concerne à intensidade final, observamos um abaixamento da intensidade em 100% dos eventos para a COR complexa. Por sua vez, a COR simples apresentou resultados que se equipararam. Finalmente, a análise da intensidade máxima revelou para a COR simples uma diminuição em 66,6% dos casos durante a correção. Por outro lado, a COR complexa não permitiu uma padronização, pois em 50% dos eventos a intensidade diminuiu na correção e em 50% aumentou.

No que se refere à análise fonológica, os padrões foram diferentes na passagem do erro para a correção, variando entre tons simples e compostos. Durante essa transição, os padrões mantiveram o tom alto tanto no erro quanto na correção em 40% dos casos. Os demais eventos (60%) variaram entre tons altos e baixos na passagem do erro para a correção.

Os ALs, FCs e HESs representaram 1,7% aproximadamente do total das disfluências. O primeiro tipo ocorreu apenas para as vogais, sendo elas /i:/ em início de GT e /é:/ em meio de GT. A duração dos segmentos foi respectivamente de 0,841s e 0,917s. A média dos eventos foi de 0,879s. Em relação à F0 média, a maior entoação ocorre para o segmento /i/. Por sua vez, os pulsos equipararam-se, não havendo distinção entre o esforço vocal. Quanto à intensidade, houve uma diminuição no momento de transição da fluência para a disfluência para o segmento /é/. Já para /i/ esse parâmetro não foi analisado, pois o AL_v ocorreu em início de GT. Em relação à intensidade final, em /i/ ocorre um aumento, enquanto que em /é/ um abaixamento. Por fim, a intensidade máxima foi maior para /é/. Contudo, a diferença entre os segmentos foi pequena, de 5dB.

Por fim, o padrão entoacional dos dois segmentos, apesar de apresentarem tons compostos, também foi diferente. O segmento /i/ apresentou um tom composto com pico entoacional baixo *L+H. Já o /e/ foi anotado com um tom com pico entoacional alto *H+L.

A disfluência do tipo FC foi realizada nos segmentos /as/ e /si/. As durações variaram entre um e outro, sendo de 0,408s para o primeiro e 0,737s para o segundo. No que diz respeito à F0 média e pulsos, não foi possível realizar análises, uma vez que no primeiro FC houve um erro de gravação e o segundo foi desvozeado, bem como ocorreu para o informante VA. A análise da intensidade, por sua vez, tanto no momento de transição quanto à final apresentou uma redução da intensidade para os dois casos. Já a intensidade máxima variou

entre um evento e outro em aproximadamente 4 dB. A média da intensidade máxima dos dois eventos foi de 85 dB, aproximadamente.

A análise entoacional não foi anotada para o segmento desvozeado. Já o segmento /as/ foi anotado com o tom simples *H.

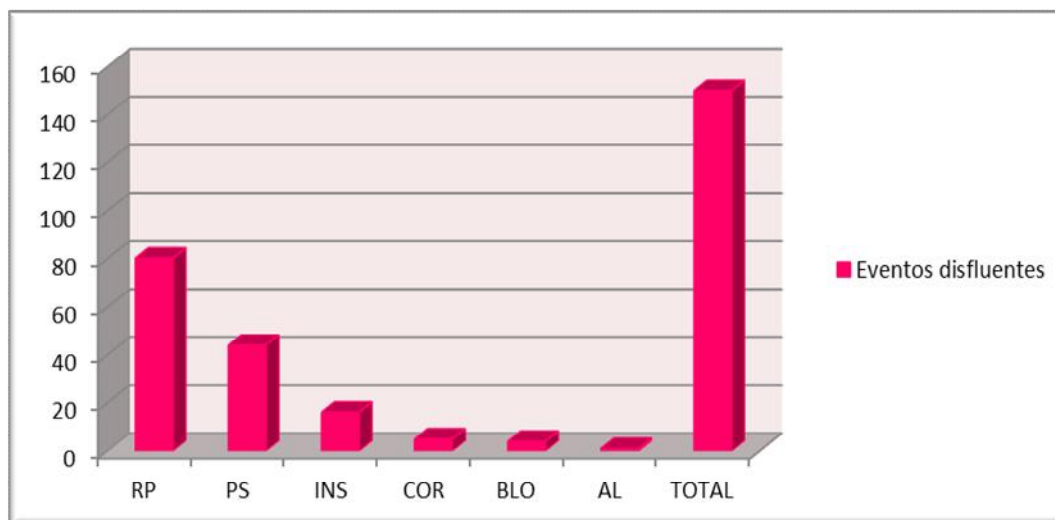
Finalmente, a disfluência do tipo HES, conforme foi verificado para os outros informantes, ocorreu para o segmento /ah/ em início de GT. Os dois eventos, mantêm a função pragmática de resposta mais enfático (hesitar) à pergunta do interlocutor.

A média da duração para esses eventos foi de 0,413s. A análise da F0 média e pulsos também variou entre os dois segmentos. Em relação à F0 média, uma das HESs foi anotada com uma frequência foi de 174 Hz, e a outra com 148 Hz. Em relação à intensidade, no momento de transição ela não foi analisada, pois são disfluências de início de GT. A análise da intensidade final aumentou nos dois eventos, quando comparada à inicial. A intensidade máxima, por sua vez, apresentou uma variação de 6dB, aproximadamente entre as HESs. Finalmente, os eventos foram anotados com o mesmo tom composto *H+L.

O informante CAM

O informante gago CAM foi o único do sexo feminino no grupo dos falantes com disfluência gaga. Assim como ocorreu para CL (do grupo dos não gagos). A fala de CA também se caracterizou por ser mais aguda, gerando frequências mais altas e maior presença de interferência na gravação. CAM realizou 150 disfluências em 72 enunciados, também divididas entre simples e complexas. As disfluências complexas contabilizaram 8,7% dos casos aproximadamente, conseqüentemente, as disfluências simples representam 91,3% dos casos. O máximo de disfluências por GT foi cinco, sendo uma do tipo PS. A taxa de elocução de CAM foi de 3,958sil/ms, a de articulação de 4,351sil/ms, e um tempo de articulação aproximado de 230ms. A seguir apresentamos no gráfico 15 a distribuição dos eventos disfluentes na fala de CAM.

Gráfico 15 – Distribuição dos eventos disfluentes realizados pelo informante gago CAM.



Fonte: própria.

CAM realizou seis tipos de disfluências, RPs, PSs; INSs; CORs BLOs e AL, respectivamente. Dentre todos os informantes, CAM foi a que menos apresentou variações no conjunto de possibilidades das disfluências.

Como podemos ver, as RPs foram as mais frequentes, aparecendo em mais da metade dos enunciados do informante (53,3%). A distribuição das RPs estão identificadas abaixo. Vale salientar que CAM foi o informante que apresentou mais variações entre as possibilidades de repetições.

- RP_{ps} (15%);
- $2RP_{ps}$ (8,7%);
- $2RP_p$ (30,2%);
- $3RP_{ps}$ (1,25%);
- $4RP_{ps}$ (2,5%);
- $5RP_{ps}$ (1,25%);
- RP_{ps2} (10%);
- $2RP_{ps2}$ (1,25%);
- $3RP_{ps2}$ (1,25%);
- RP_c (16%);
- $2RP_e$ (8,7%);
- $2RP_v$ (2,5%);
- $3RP_p$ (2,5%);

- 4RP_p (1,25%).

As repetições de palavras foram as mais frequentes, com uma ocorrência de 97,5% dos casos. Dentre as repetições duplas de palavras (2RP_p), quatro ocorreram em palavras dissílabas e uma em trissílaba. Os segmentos desvozeados ou com uma qualidade de voz do tipo *creaky voice* foram identificados tanto na primeira quanto na última RP. Em relação às RPs complexas, elas ocorreram com PIs; RPs; RP_{ps}; 3RP_p; 2RP_e PSs e BLOs. No último caso, os BLOs ocorreram em consoantes oclusivas alveolares desvozeadas [t] (dois casos) e oclusivas bilabiais desvozeadas [p] (dois casos). De todos os casos de BLO apenas um [p] foi realizado com pausa articulatória do tipo 2.

No que diz respeito à análise acústica da duração foram verificados os seguintes resultados:

- Para a 2RP_v a duração diminuiu em 100% das repetições;
- A 4RP_p apresentou uma duração que diminuiu ao longo das repetições;
- A 3RP_p revelou para um caso uma duração que diminuiu na segunda RP e aumentou na terceira (onda) e, para o outro, uma duração decrescente, bem como observado para a 4RP_p;
- A 2RP_{ps2} apresentou uma duração com padrão de onda, mas que começou mais baixo, aumentou na segunda RP e depois diminuiu;
- A 3RP_{ps2} diminuiu progressivamente;
- A 4RP_{ps}, dos os dois casos, um deles apresentou uma duração que aumentou da primeira para a segunda e RP e que depois diminuiu progressivamente. Já o outro apenas revelou uma diminuição da duração ao longo das RPs;
- A duração da 3RP_{ps} aumentou nas três primeiras repetições e diminuiu na última, que diz respeito à duração da sílaba em palavra;
- A 5RP_{ps}, por sua vez, revelou uma duração que diminuiu na passagem da primeira para a segunda RP, que depois aumentou na terceira RP, diminuindo mais uma vez na quarta e, por fim, aumentando na sílaba em palavra;
- Para a 2RP_e, RP_{ps2}, 2RP_p, RP_c e RP_{ps} a duração diminuiu em respectivamente em 71%, 87%, 63%, 60% e 91% das repetições;
- Finalmente, para a 2RP_{ps} os resultados equipararam-se, ou seja, em 50% das repetições a duração aumentou da primeira para a segunda RP e diminuiu na sílaba em palavra e, nos outros 50% diminuiu progressivamente em relação à RP inicial.

No que se refere à F0 média e aos pulsos os resultados foram:

- Para a 3RP_p tivemos dois padrões de onda, sendo que em um caso houve uma diminuição da duração primeira para a segunda RP e um aumento da segunda para a terceira e, um caso em que a duração aumentou na segunda RP e diminuiu na terceira. Já os pulsos mantiveram um padrão crescente para os dois casos;
- Para a 4RP_p a F0 média diminuiu na passagem da primeira para a segunda RP e aumentou ao longo das RPs restantes. Já os pulsos apresentaram uma diminuição da primeira para a segunda RP, um aumento da segunda para terceira e terceira para a quarta;
- A 2RP_v mostrou um aumento da F0 média e pulsos em 100% dos casos;
- A 2RP_{ps2} revelou uma F0 e pulsos que aumentaram na segunda RP e diminuíram na sílaba de palavra;
- 3RP_{ps2} tanto para a F0 quanto para os pulsos aumentou da primeira para a segunda RP e, na sequência diminuiu;
- Para a 4RP_{ps} observamos para um caso, uma F0 média com um padrão onda, que começa com uma diminuição, depois aumenta, diminui novamente e termina com um aumento da F0 média. Os pulsos, por sua vez também mantiveram um padrão onda, mais diferentemente da F0, houve um aumento da primeira para a segunda RP. No outro caso de 4RP_{ps} tivemos para as duas últimas RPs uma RP surda e uma com erro. Já as duas RPs que seguiram a primeira repetição revelaram um aumento progressivo na F0 média. Para os pulsos houve uma diminuição e posterior aumento;
- Quanto a 3RP_{ps}, tanto para a F0 média quanto para os pulsos, ocorreu uma diminuição da primeira para a segunda RP, com posterior aumento e, mais uma vez, uma diminuição da última RP para a sílaba em palavra;
- Para a 5RP_{ps} a F0 diminuiu da primeira para segunda RP, aumentou na terceira, diminuiu na quarta e aumentou na última. Em relação aos pulsos, também houve uma diminuição da primeira para a segunda RP, mas por outro lado, uma diminuição na terceira e quarta RPs, com posterior aumento;
- Em relação à RP_e os resultados equipararam-se entre as estruturas. Já para os pulsos houve um aumento do esforço vocal em 71% dos casos para a segunda estrutura repetida;
- Para a RP_{ps2} a F0 média aumentou em 87% dos eventos e diminuiu para os pulsos em 62% deles. A 2RP_p revelou a mesma tendência, mas para respectivamente 72% e

54% dos casos. A mesma situação ocorreu para a RP_c , sendo para a F0 média um aumento em 60% dos eventos e para os pulsos uma diminuição também em 60% dos casos;

- A RP_{ps} , por sua vez, apresentou uma diminuição tanto para a F0 média (52%) quanto para os pulsos (76%);

- Por fim, a $2RP_{ps}$ variou para a F0 média entre 1) um aumento da primeira para a segunda RP e diminuição da segunda RP para a terceira (sílabas em palavra), 2) diminuição da primeira para a segunda RP e aumento para a sílaba em palavra e 3) aumento progressivo. Já para os pulsos, 80% dos casos revelaram uma tendência para o aumento da primeira para a segunda RP e uma diminuição do esforço vocal na sílaba repetida em palavra.

Quanto à intensidade de transição, final e inicial foram verificados os seguintes padrões:

- Para a $3RP_p$ entre as duas ocorrências a intensidade diminuiu na passagem da fluência para a disfluência em apenas um caso. Depois, na passagem entre as disfluências, ocorre uma diminuição em 66,6% dos casos. Em relação à intensidade final, não houve uma padronização. Já quanto à intensidade máxima, o padrão foi do tipo onda, diminuindo a intensidade da primeira para a segunda RP e depois aumentando;

- Quanto a $4RP_p$ a intensidade de transição aumentou nos dois momentos observados. A intensidade final, por sua vez, aumentou nas quatro RPs e, a intensidade máxima apresentou um padrão crescente;

- Em relação à $2RP_v$, dentre as duas ocorrências, uma delas iniciou o GT. A outra, por sua vez, revelou uma diminuição da intensidade no momento de transição. Já a intensidade na passagem entre as transições não possibilitou padronização, pois os resultados equivaleram-se. Fato que se manteve para a intensidade final, bem como para a intensidade máxima;

- A $2RP_{ps2}$ diminuiu nos dois momentos considerados. A intensidade final aumentou em 66,6% dos casos e a intensidade máxima apresentou um padrão onda que aumentou da primeira para a segunda RP e diminuiu na terceira;

- A $3RP_{ps2}$ revelou uma diminuição na transição da fluência para a disfluência e, na passagem entre as disfluências uma diminuição em 75% das RPs. Quanto à intensidade final houve um aumento em 100% das repetições. Já a intensidade máxima revelou um aumento da primeira para a segunda RP, seguida de uma abaixamento na segunda e terceira RPs;

- A $3RP_{ps}$, por sua vez, apresentou uma intensidade de transição que diminuiu na passagem da fluência para a disfluência, bem como entre as repetições (66,6%). Já a intensidade final apresentou valores iguais e, a intensidade máxima, um padrão crescente;
- A $4RP_{ps}$ revelou uma abaixamento na transição da fluência para a RP nos dois casos analisados e, uma tendência para o aumento da intensidade na passagem entre as repetições em 80% dos casos. Já a intensidade máxima variou, de modo que em um caso diminuiu da primeira para a segunda RP e depois seguiu com um padrão de onda, aumentando na terceira RP, diminuindo na quarta e aumentando mais uma vez na sílaba repetida em palavra;
- Para a $5RP_{ps}$ a intensidade diminuiu na transição da fluência para a primeira RP, diminuindo em 80% nas outras RPs. A intensidade final, por sua vez, apresentou uma tendência para o aumento em 66,6% dos eventos. Finalmente, para a intensidade máxima, foi observado u padrão decrescente nas três primeiras RPs, seguido de um aumento da intensidade e, depois, de um abaixamento na sílaba em palavra;
- Quanto à $2RP_e$ ocorreu uma diminuição da intensidade em 100% dos casos tanto durante a transição da fluência para a primeira RP_e quanto da primeira para a segunda RP_e . A intensidade final também diminuiu em 71% das RPs e, a intensidade máxima aumentou em 85% das repetições;
- A $2RP_{ps}$ aumentou em 55% dos casos na transição da fluência para a primeira RP e uma diminuição na transição entre as RPs em 52% dos eventos. A intensidade final abaixou em 62% dos casos e, a intensidade máxima apresentou um padrão de onda, que aumentou da primeira para a segunda RP e diminui da segunda para a RP em sílaba de palavra em 58% dos dados;
- Para a RP_{ps2} ocorreu uma diminuição da intensidade de transição nos dois momentos, sendo a tendência para o abaixamento de 75% na passagem da fluência para a primeira RP e de 85% entre as disfluências. A intensidade final, por sua vez, aumentou em 75% dos casos, bem como a intensidade máxima, mas nesse caso, em 62,5% dos eventos;
- A $2RP_p$ apresentou a mesma tendência de abaixamento observada para o tipo de repetição anterior, sedo de 72% no primeiro momento e 83% no segundo. Já a intensidade final aumentou em 79% dos casos e a máxima diminuiu em 66% deles;
- A RP_c também seguiu a mesma tendência verificada para a RP_{ps2} e $2RP_p$, sendo o abaixamento de 60% na transição da fluência para a primeira repetição e de 100% entre as

disfluências. Já a intensidade final e máxima diminuiu em 90% e 80% dos eventos, respectivamente;

- Finalmente, a 2RP_{ps} seguiu os mesmos parâmetros dos quatro tipos de RPs anteriores, diminuindo em 81% dos casos no primeiro momento e em 68% no segundo. A intensidade final e máxima aumentou, respectivamente, em 72% e 54,5% das RPs.

Por fim, a análise entoacional manteve a tendência para o mesmo padrão na maior parte das repetições. Veja abaixo:

- A 3RP_p apresentou para 100% dos eventos o padrão *H;
- A 4RP_p também manteve o mesmo tom *H para as três RPs e para a RP de sílaba em palavra;
- A 2RP_v apresentou para um caso, uma RP de padrão *creaky voice* e a outra com tom simples alto *H. E, para o segundo caso, uma RP de tom simples alto *H e a outra com tom simples baixo *L;
- A 2RP_{ps2} também não manteve os mesmos padrões apresentando os tons *L *L+H *H+L;
- A 3RP_{ps2}, por sua vez, apresentou o mesmo tom *H para as três primeiras RPs e, para a repetição da sílaba em palavra o tom composto *H+L;
- 3RP_{ps} manteve o padrão *H para a primeira e terceira RP, bem como para a sílaba em palavra. Já a segunda RP apresentou o tom baixo *L;
- A 4RP_{ps} apresentou para um caso o padrão *H *L *H *H *H+L. Como vemos, para esse casos ocorreram dois tons diferentes o segundo, que foi anotado como baixo e o terceiro que manteve o pico entoacional alto, mas agora em tom composto;
- A 5RP_{ps} manteve para 100% das repetições, inclusive para a repetição em sílaba de palavra o tom composto *H+L;
- A 2RP_e manteve o mesmo padrão entoacional para 86% dos casos. O único caso que variou apresentou uma qualidade de voz do tipo *creaky voice*;
- A 2RP_{ps} manteve o mesmo padrão para 57% dos casos. Dos eventos que não mantiveram o mesmo padrão, um deles variou em apenas uma RP, de modo que a segunda sílaba repetida e a sílaba em palavra apresentou o mesmo tom, sendo o padrão *L *H *H e, os outros dois casos apresentaram a primeira RP com uma qualidade de voz *creaky voice* e a segunda RP, bem como a sílaba repetida em palavra o mesmo tom alto *H;
- A RP_{ps2}, por sua vez, manteve o mesmo padrão para 87,5% dos casos;

- A $2RP_p$ apresentou o mesmo padrão para 75% das repetições. Dos três casos que variaram, dois deles foram anotados com o mesmo pico entoacional, variando em relação a um tom simples e o outro composto;

- Quanto à RP_c os padrões foram os mesmos em 46% dos casos. Dentre os casos que variaram, dois deles apresentaram uma qualidade de voz *creaky voice* para uma RP, mantendo o tom no outro evento e quatro deles apresentaram o mesmo pico entoacional nas RPs;

- Finalmente, para a RP_{ps} o padrão entoacional manteve-se para 63,6% dos casos. Dos resultados que variaram, dois mantiveram o mesmo pico entoacional e quatro apresentaram um tom desvozeado.

Após as repetições, as PSs representam o segundo tipo de disfluência mais frequente para CA, ocorrendo em aproximadamente 29% das disfluências. O tipo de PS mais comum foi a de final de GT (47,7%). Em seguida, aparecem as PSs em início de GT (36,5% aproximadamente) e, por fim, as PSs em meio de GT (15,8% aproximadamente). As pausas que antecederam eventos disfluentes contabilizam 47%.

Em relação à duração, as PSs em fim de enunciado também foram as maiores (11,580s). Na sequência, temos as PSs em início de GT (6,829s) e, por fim, as de meio de GT (2,910s). Entre PSs silenciosas e respiratórias, a maior duração foi verificada para as PSs do primeiro tipo, com uma duração de 13,985s. Já as PSs respiratórias tiveram uma duração total de 7,334s.

As INSS, por sua vez, aparecem em aproximadamente 11,7% dos eventos disfluentes. Elas dividiram-se INS_{ps} (68,8%), INS_v s (18,8%) e INS_e s (12,4%) As INSS realizadas por CA foram:

- INS_p em início de GT - /assim/, /intão/, /aí/;
- INS_p em fim de GT - /né/;
- INS_p em meio de GT - /assim/;
- INS_v em início de GT - /i/;
- INS_v em meio de GT - /é/;
- INS_e em início de GT - /intão assim/; /i assim/.

A não ser pela INS_p em final de GT, que tem a função semântica de uma afirmação retórica, com a intenção de reafirmar o que foi enunciado, a maior parte das outras INSS mantiveram a ideia principal de dar sequência a enunciação, bem como foi analisado para os outros informantes. Essa tendência também foi verificada para as INSS de estrutura. A única

INS que não manteve essa característica foi a do tipo /*assim*/, que apresentou uma função semântica de exemplificação. Observe o exemplo:

- EN69 “*assim* eu ve eu vo fala eu num consigo”.

Veja que CAM inicia o GT com a INS_p /*assim*/. Na sequência discursiva, ela conta como a gagueira influência a sua rotina, fazendo uma exemplificação por meio do segmento /*assim*/, o que acontece com ela, nesse caso, que não a deixa se comunicar bem. Na sequência da INS_p , a informante também realiza uma COR simples, por meio da troca de duas consoantes: /eu ve > eu vo/. Dentre os enunciados que apresentaram esse tipo de INS_p (três em início de GT e um em meio de GT), apenas mais um deles, em início de GT, manteve a função de exemplificação. Os outros nos passam a ideia de sequência discursiva.

No que concerne à análise acústica das INS_s , a média da duração revelou para as $INS_{e,s}$ (0,762s) a maior duração. Em seguida, temos as $INS_{v,s}$ em início de GT (0,384s); as $INS_{p,s}$ em início de GT (0,367s); a INS_p em meio de GT (0,356s); a INS_v em meio de GT (0,331s) e, por fim, as $INS_{p,s}$ em final de GT (0,326s). Quanto à F0 média, ela não foi observada para a maior a INS_e , pois a frequência muito aguda do informante gerou um erro de gravação. Dentre as outras INS_s , a INS_p em meio de GT foi a que apresentou a maior F0 (199 Hz), seguida das $INS_{p,s}$ em início de GT (195 Hz); da INS_v em meio de GT (193 Hz); das $INS_{v,s}$ em início de GT e, por fim, das $INS_{p,s}$ em final de GT. Os pulsos, deixando de lado as $INS_{e,s}$, seguiram a mesma tendência da F0 média.

A análise da transição da intensidade não foi observada para as $INS_{p,s}$, $INS_{v,s}$ e $INS_{e,s}$ em início de GT, uma vez que iniciaram o GT. Dessa maneira, no momento de transição, a intensidade aumentou na passagem da fluência para a disfluência para a INS_p em meio de enunciado e diminuiu em 66,6% dos casos para a INS_p de final de GT. A intensidade final, quando comparada à inicial, revelou um aumento apenas para as $INS_{p,s}$ em início de GT (57%). Já para a INS_v em meio de GT; $INS_{e,s}$ e $INS_{p,s}$ em meio e fim de GT ela diminuiu em 100% dos dados. Por fim, as $INS_{v,s}$ mantiveram o mesmo padrão: 50% dos casos aumentou e 50% diminuiu. Quanto à intensidade máxima, a média dos resultados mostrou a maior intensidade para a INS_p em meio de GT (77 dB). Em seguida, temos as $INS_{e,s}$ (74 dB) e as $INS_{p,s}$ em início de GT (73 dB). As $INS_{p,s}$ em fim de GT e a INS_v em meio de GT aparecem juntas com uma diferença de 0,150 dB, sendo a intensidade máxima da primeira INS de 72,763 dB e da última de 72,914 dB. Finalmente, a menor intensidade máxima foi verificada para as $INS_{v,s}$ em início de GT.

A análise dos padrões fonológicos confirmou a tendência das INSS manterem o mesmo padrão entoacional. Dessa maneira, os padrões obtidos para as inserções foram:

- /intão/, /ái/ e /assim/ em início de GT - *H para 100% dos casos com /intão/ e /ái/ e, *H+L para 100% dos casos de INS do tipo /assim/;
- /né/ em final de GT - *L+H para 66,6% dos eventos (dois casos) e *L para a outra INS_p;
- /assim/ em meio de GT - *H (caso único);
- /intão assim/ e /i assim/ - *H+L para os dois eventos;
- /i/ em início de GT – INS_v que mais variou. Dos dois casos, um apresentou o tom *L e o outro *H;
- /é/ em meio de GT - *H (único caso).

Na sequência, as CORs representam 3,3% do total de disfluências. O informante não realizou nenhuma COR complexa, apenas simples. Dentre elas, manteve parte da informação contida no erro na correção em 40% dos casos.

A análise acústica mostrou que a duração aumentou durante a correção em 60% dos casos. A média da F0 também aumentou no momento da correção, mas agora em 80% dos casos. Os pulsos mantiveram a mesma tendência em 60% dos eventos corrigidos. Em relação à intensidade, seja na transição da fluência para o erro ou do erro para a correção a intensidade diminuiu em respectivamente 60% e 80% dos eventos. A intensidade final e a intensidade máxima também mantiveram o mesmo padrão. Nesses casos, a intensidade final diminuiu em comparação à inicial em 60% dos dados e a intensidade máxima foi maior na correção do que no erro em 80% deles.

Por fim, a análise fonológica, mostrou o mesmo padrão entoacional tanto no erro quanto na correção em 40% das disfluências. Os outros eventos variaram entre tons simples e compostos, altos e baixos, na passagem do erro para a correção.

Depois das CORs, o BLO aparece em aproximadamente 2,7% dos eventos disfluentes. Eles foram verificados em consoantes (75%) e vogal (25%). Em vogal, o bloqueio ocorreu em vogal anterior aberta [a]. Já para as consoantes eles foram identificados em:

- oclusiva bilabial [p], duas ocorrências nos segmentos /-palavra/ e /-paralizo/;
- oclusiva alveolar [t], uma ocorrência no segmento /-trata/;

Quanto ao tipo de PS de bloqueio, o BLO_v apresentou uma PS articulatória do tipo 1. Já o BLO_c dividiu-se em PSs articulatória para o bloqueio no segmento /-palavra/, PSs silenciosas com articulação (tipo 2) no segmento /-trata/ e, por fim, um caso sem pausa, mais

especificamente para o segmento /-paralizo/. Dentre todas as PSs de BLO, apenas a de BLO_v foi maior do que o segmento bloqueado.

Em relação à análise acústica, o BLO_v apresentou uma duração de 0,194s. Quanto à F0 média, ocorreu um erro de gravação, gerando uma frequência muito alta. No que diz respeito à intensidade, a final aumentou em comparação à inicial e a máxima apresentou uma intensidade de 74dB, aproximadamente.

Quanto ao BLO em consoante, a maior duração ocorreu para a consoante [p] (1,021s) e a menor para [t] (0,537s). Em relação aos pulsos e F0 média não foi possível fazer comparações, uma vez que para a consoante [t] ocorreu um erro de gravação. A intensidade de transição foi medida apenas para o segmento /-paralizo/, sem PS de BLO, que apresentou uma diminuição na intensidade. Quanto à intensidade final, em 100% dos casos ela aumentou em relação à inicial. Por fim, a intensidade máxima foi maior para a consoante oclusiva desvozeada [p], com 77 dB, enquanto que a oclusiva alveolar desvozeada [t] apresentou uma intensidade de 70 dB aproximadamente.

No que diz respeito aos padrões entoacionais, o BLO_v foi analisado com o tom alto simples *H. Os BLOs em consoante, por sua vez, variaram entre tons simples e compostos. O segmento /-trata/ foi anotado com o tom composto com pico entoacional baixo *L+H. Já os segmentos /-paralizo/ e /-palavra/ variaram entre um padrão entoacional simples, para o primeiro e composto para o segundo segmento. No entanto, ambos apresentaram um pico entoacional alto. Sendo assim, eles foram anotados com os tons *H e *H+L, respectivamente.

A disfluência do tipo AL foi a menos frequente em CAM, aparecendo em 0,6% (um evento) do total de disfluências. O segmento alongado ocorreu em vogal /i/ em início de GT. A sua função semântica manteve a ideia de continuidade discursiva. O segmento durou 1,093s, com uma frequência média de 193 Hz. Quanto à intensidade, como foi uma disfluência em início de GT, ela não foi analisada. A intensidade final revelou um abaixamento em relação à inicial e, a intensidade média foi de 73 dB. Por fim, esse segmento foi anotado com o tom simples alto *H, o que revela uma entoação mais alta que se mantém constante.

4.3 Comparação geral dos resultados

4.3.1 GTs, PSs, tipos de disfluências, tessitura e aspectos temporais

No que diz respeito ao número de GTs foi observado que os informantes não gogos realizaram uma média de 41 GTs por informante. Já os gogos, uma média de 52. A diferença

de 11 GTs entre os informantes não pode ser considerada como significativa, uma vez que o número de GTs está relacionado principalmente à duração da gravação de cada informante, como vimos no capítulo referente à metodologia. De uma maneira geral, se formos separar os informantes de acordo com o número de GTs veremos que essa diferença continua pouco significativa. Observe a tabela a seguir.

Tabela 2 – Comparação entre o número de GTs dos informantes gagos e não gagos.

INFORMANTES GAGOS	INFORMANTES NÃO GAGOS	NÚMERO DE GTs	
		GAGOS	NÃO GAGOS
PA	MA	38	38
AP	CR	35	34
JO	LE	44	37
MAU	BR	58	39
VA	CL	66	47
CAM	NE	72	53

Fonte: própria.

A tabela 2 mostra uma tentativa de comparação dos GTs entre os informantes gagos e não gagos. Como podemos ver, há informantes que se aproximam mais do que outros. Provavelmente, um enunciado mais longo apresentará um número maior de GTs. Na tabela, vemos que em alguns casos (PA/MA e AP/CR) é irrelevante contar os GTs para inferirmos algo sobre as disfluências. No entanto, ao levarmos em conta os outros informantes, é possível depreender que os falantes gagos segmentam mais, criam mais GTs do que os não gagos. Para PA e MA não houve diferença. Para AP e CR, mesmo sendo pequena, AP tem 2,94% a mais de GTs do que CR. JO, por sua vez, tem 15,9% a mais de GTs do que o informante não gago LE. Por fim, os informantes gagos MAU, VA e CAM realizaram 32,7% a mais de GTs do que os informantes não gagos BR, CL e NE. Nesses casos, a diferença típica entre gagos e não gagos é que os primeiros realizam 30% a mais de GTs do que os últimos. Nesse sentido, em uma fala comum, um falante gago segmentaria 30% a mais do que um não gago.

A maior diferença entre os informantes está no tipo de disfluência que eles realizam. Os informantes gagos realizam mais disfluências complexas do que os não gagos. Observe a tabela abaixo:

Tabela 3 –Número de ocorrência das disfluências para os informantes gagos e não gagos.

	DISFLUÊNCIA SIMPLES	DISFLUÊNCIA COMPLEXA	TOTAL
INFORMANTES GAGOS	737	127	864
INFORMANTES NÃO GAGOS	648	58	706

Fonte: própria.

A tabela 3 mostra que os informantes gagos (14,7%) apresentaram mais disfluências complexas do que os não gagos (8,2%). Consequentemente, as disfluências simples são maiores para os não gagos (91,8%) do que para os falantes gagos (85,3%). A diferença de 6,5% entre as disfluências gagas pode parecer irrelevante. No entanto, acreditamos que ela contribui efetivamente para a diferenciação dos indivíduos, bem como para a questão da empatia fonética. Tal fato está intimamente relacionado com a distribuição do número de disfluências por enunciado realizada por cada falante. Veja na tabela que segue:

Tabela 4 – Distribuição do número de disfluências e pausas por enunciado para os informantes gagos e não gagos.

INFORMANTES GAGOS	Número de disfluências por GT	Número de pausas	INFORMANTES NÃO GAGOS	Número de disfluências por GT	Número de pausas
AP	8	1	BR	7	3
CAM	5	1	CL	6	2
JO	6	3	CR	4	1
MAU	6	2	LE	7	2
PA	9	1	MA	6	2
VA	6	0	NE	5	4

Fonte: própria.

Na tabela 4 estão identificados os informantes gagos e não gagos, bem como o número máximo de disfluências realizada por cada um deles em um mesmo GT. Ao lado da quantidade de disfluências, também está apresentado o número máximo de pausas (respiratória ou não), que ocorreu em cada um desses GTs. A média aritmética do número de disfluências revela a diferença de aproximadamente uma disfluência para os informantes gagos, sendo de seis disfluências por enunciado para os gagos e de cinco para os não gagos. Já para as PSs, a média foi de uma aproximadamente para os gagos e de duas para os não gagos. Esses resultados mostram uma diferença insignificante. No entanto, a análise mais aprofundada dos tipos de disfluência ocorridos e consequentemente o número de PSs revelam uma situação diferente.

De acordo com a tabela, todos os informantes não gogos realizaram algum tipo de PS no decorrer do GT. Já em relação aos gogos VA não realizou nenhuma PS. Individualmente, os informantes gogos realizaram no máximo nove disfluências por enunciado (PA), sendo que entre elas ocorreu apenas uma PS. Por sua vez, os não gogos realizaram no máximo sete disfluências por GT (BR e LE), sendo três PSs para BR e duas para LE. O informante NE, por sua vez, em um total de cinco disfluências, quatro foram de PSs. Por outro lado, o máximo de PSs realizadas por um informante não gago foi de três, em um conjunto de seis disfluências.

No que diz respeito aos tipos de disfluências em um mesmo enunciado os gogos realizam mais disfluências complexas e repetições do que os não gogos. AP, por exemplo, realizou as seguintes repetições em um mesmo GT: $4RP_v$, $2RP_v$ e $2RP_p$. E, em outro, $4RP_v$, $2RP_e+PI$, $2RP_p+BLO_v$. Para os não gogos não foram observadas mais de um tipo de repetição em um mesmo GT. Mesmo que eles realizem mais de uma disfluência por GT, não há duas repetições em um mesmo enunciado. Além disso, uma palavra ou vogal pode repetir para os não gogos o máximo de três vezes. Já para os gogos, essa repetição pode chegar a cinco vezes. Tal resultado é justificado pela média de RPs com mais de duas repetições realizadas por gogos e não gogos, que chegam a 5,5% do total de RPs, que somam 287 para o primeiro grupo e 0,6% para o segundo em um total de 87 RPs. A quantidade de RPs realizadas pelos informantes também confirma que os gogos realizam mais repetições do que os não gogos.

Quanto às outras disfluências, a mais comum para os não gogos são as PSs (83%), depois aparecem, em ordem, as INSSs, RPs, CORs, ALs, BLOs, HESSs, PIs e FCs. Já para os gogos elas dividiram-se entre RPs (50%) e PSs (50%). Depois, ocorrem INSSs, BLOs, CORs, ALs, FCs, HESSs e PIs. Como vemos, tanto para os informantes gogos quanto para os não gogos as HESSs, as PIs e os FCs são os menos frequentes. Já as INSSs costumam ser as mais frequentes depois das PSs para os não gogos e para os gogos, depois de PSs e RPs. É importante salientar que o número máximo e mínimo de disfluências refere-se à produção mais frequente dos informantes e não à quantidade de realizações. Por exemplo, em relação à quantidade, os falantes gogos realizaram 290 PSs e 287 RPs, como vimos. Mesmo assim, as PSs são maiores para os falantes não gogos, somando 338.

Finalmente, no que diz respeito aos tipos de disfluência para um mesmo informante, não foi verificada uma diferença significativa entre os grupos de falantes, sendo a média de 6,8 disfluências em cada gravação para os informantes com disfluência gaga e de 6,5 para a disfluência comum.

Quanto às disfluências complexas, elas também são verificadas para os informantes não gago, mas normalmente são mais frequentes com PSs. Por exemplo, pode ocorrer uma INS_p com PS, ou uma RP_p com PS, dentre outras variações de disfluências também com PS. Isso não inviabiliza que ocorra outros tipos de disfluências complexas, como por exemplo, CORs com INSs ou PIs. No entanto, é mais difícil que RPs, em disfluência comum, venham acompanhadas de outros tipos de disfluências do que as PSs. Para BR, por exemplo, dentre as 12 disfluências complexas realizadas, apenas duas delas foram do tipo $2RP_p+AL_p+PS_e$. Já para o informante gago VA, em um total de 11 disfluências complexas foram realizadas:

- 1 caso do tipo $3RP_e+PI$;
- 1 caso do tipo $2RP_e+PI$;
- 1 caso do tipo $2RP_p+PI+PS_e$;
- 1 caso do tipo $2RP_e+PI+PS_e$;
- 1 caso do tipo $2RPe+PSe$ ⁶³.

Como vemos, aproximadamente 16,6% da produção de RP complexa para BR ocorre com outro tipo de disfluência, enquanto que para VA elas representam aproximadamente 45,5%. É interessante ressaltar que essa diferença é mais saliente com as RPs, já que os outros casos de disfluência complexa são semelhantes para os dois grupos de informantes. Veja um exemplo na tabela que segue.

Tabela 5 – Comparação entre os tipos de disfluências complexas para os informantes gago JO e não gago MA.

JO	MA
COR+INS_v	COR+PSe+BLO _v
COR+PSe	COR+PSe
INS_v + AL	INS _v +AL
	COR+PI

Fonte: própria.

A tabela 5 acima identifica as disfluências complexas para os informantes JO e MA. Nela, foram excluídos os exemplos que se repetiam e os casos com RPs. Ao serem comparadas, as disfluências mostram casos que se assemelham para os dois grupos de informantes, como uma INS de vogal com AL ou uma COR com pausa entre o erro e a correção. Os casos de disfluências que ocorreram para um informante, mas não para o outro, como COR+PI para MA, não inviabiliza a afirmação. Esse tipo de disfluência também foi

⁶³ Vale destacar que os exemplos de RPs complexas apresentados (BR e VA) foram escolhidos por terem uma diferença de apenas uma disfluência complexa entre eles, a fim de facilitar as comparações.

observado para outros informantes gagos, como por exemplo, para o informante gago VA no enunciado que segue:

- EN26 “/tava **co** > **fazenu** si sigundu aí saí#⁶⁴”

EN26 representa um GT em que ocorre uma COR com PI. No exemplo, o erro e a correção aparecem grafados em negrito. O símbolo maior que (>) indica, do lado esquerdo o erro, que aparece por meio de uma palavra incompleta. Do lado esquerdo do símbolo temos a correção. Nesse caso, VA confunde o início do curso escolar com a sua realização. No primeiro momento a PI sugere que ele estaria começando o segundo ano, mas na sequência, ao se corrigir, abandonando a realização completa da palavra, ele diz que já estava cursando “fazendo” o segundo ano quando deixou a escola. Em EN26 também ocorre uma PS inicial e final representadas, respectivamente, por (/) e (#) e uma repetição de sílaba.

Retomando às PSs, vimos que elas são mais frequentes para os informantes não gagos, representando aproximadamente 48% do total de disfluências, enquanto que para os informantes com disfluência gaga, elas representam 33,6% aproximadamente. Dentre os seis falantes não gagos analisados, apenas um deles (LE) não apresentou a PS como disfluência mais realizada. Já para os gagos, apenas em três deles (JO, MAU e PA) a PS foi a disfluência mais frequente. Apesar de representar apenas 50% dos informantes gagos, em detrimento aos não gagos que chegam a 90%, esses resultados revelam a importância das PSs para o discurso.

Como mencionado, elas podem ser respiratórias ou silenciosas e ocorrer em início, meio ou fim de enunciado. A tabela que segue identifica a ocorrência geral das pausas para esses dois grupos de informantes.

Tabela 6 – Distribuição das pausas de acordo com o seu contexto no GT.

DISTRIBUIÇÃO DAS PSs	DISFLUÊNCIA GAGA	DISFLUÊNCIA COMUM
PS em início de GT	33,3%	33,3%
PS em meio de GT	0%	33,3%
PS em fim de GT	66,7%	33,3%

Fonte: própria.

A tabela 6 identifica a média de maior ocorrência das pausas, por informante, de acordo com a sua posição no GT. Diante disso, observamos que os informantes com disfluência gaga apresentaram um maior número de PSs (66,7%) em fim de GT (CAM, MAU, PA e VA) e uma menor ocorrência (33,3%) em início de GT (AP e JO). Por sua vez, os informantes com disfluência comum tiveram a mesma distribuição (33,3%) nos três

⁶⁴ A justificativa em apresentar esse exemplo é salientar os eventos comuns entre os dois grupos de informantes, uma vez que a tabela 6 apresentou casos dois casos que ocorreram para MA, mas não para JO.

contextos, sendo que CR e NE realizaram um maior número de PSs em início de GT, BR e CL em meio e LE e MA em final de GT.

Quanto à duração das PSs em início, meio e fim de enunciado, os resultados foram os seguintes:

Tabela 7 – Duração das PSs de acordo com a sua posição no enunciado.

DURAÇÃO DAS PSs em ms	EM INÍCIO DE GT	EM MEIO DE GT	EM FIM DE GT	TOTAL
DISFLUÊNCIA GAGA	44,809	24,969	61,860	131,638
DISFLUÊNCIA COMUM	57,760	72,506	44,679	174,945

Fonte: própria.

Como podemos ver na tabela 7, as PSs apresentam a maior duração para os informantes com disfluência comum em início e meio de GT. Para os disfluentes gogos, elas são maiores apenas em fim de GT. Pode-se dizer que a maior diferença entre as durações ocorre para as PSs em meio de GT (47,537ms). Em início de GT a diferença é de apenas 12,951ms, o que mostra que o começo do enunciado não influencia a produção de gagueira. Gagos e não gagos tendem a iniciar a fala bem, assim como a sua finalização, pois como podemos observar, a diferença das PSs em relação à produção gaga e não gaga, também é pequena (17,181ms). O problema maior ocorre no meio do enunciado, como vimos. Nesse momento, os informantes com disfluência comum conseguem articular melhor a sua fala por meio de PSs, diferentemente dos não gagos que acabam realizando mais disfluência. No conjunto, a duração total das PSs é maior para os informantes com disfluência comum, com uma diferença de 43,307ms dos informantes com disfluência gaga.

No que diz respeito às PSs respiratórias e silenciosas obtivemos os seguintes resultados:

Tabela 8 – Ocorrência das PSs silenciosas e respiratórias.

TIPO DE PS/INFORMANTES	PSs RESPIRATÓRIAS	PSs SILENCIOSAS
DISFLUÊNCIA GAGA	33,3%	66,7%
DISFLUÊNCIA COMUM	66,7%	33,3%

Fonte: própria.

A tabela 8 mostra que em relação à maior ocorrência por informante entre PSs respiratórias e silenciosas, houve uma inversão nos resultados quanto à produção dos informantes gagos e não gagos. Para o primeiro grupo, os falantes realizaram um maior

número de PSs silenciosas (AP, CAM, PA e VA). Por sua vez, o segundo grupo realizou mais PSs respiratórias (BR, CL, CR e NE). A análise da duração comprova os resultados obtidos. Observe na sequência a tabela de duração das PSs.

Tabela 9 – Duração das PSs respiratórias e silenciosas.

DURAÇÃO	PSs RESPIRATÓRIAS	PSs SILENCIOSAS
DISFLUÊNCIA GAGA	63,637	79,817
DISFLUÊNCIA COMUM	93,558	80,684

Fonte: própria.

A análise da duração das PSs para os informantes gagos e não gagos (cf. tab. 9) mostra que as PSs silenciosas foram as que apresentaram a maior duração para os informantes com disfluência gaga (79,817ms), enquanto que para os falantes com disfluência comum foram as respiratórias (93,558ms). Por sua vez, a comparação entre os grupos revela, mais uma vez, que os informantes com disfluência comum apresentam uma duração maior tanto para as PSs respiratórias quanto para as silenciosas. No entanto, essa diferença é pouco significativa para as PSs silenciosas, de apenas 0,867ms. Já para as respiratórias ela é de 29,921ms.

A disfluência do tipo INS, como vimos, podem ocorrer com palavras e vogais. As do primeiro tipo variaram para os informantes não gagos entre:

Quadro 4 – Os tipos de INSs para os informantes com disfluência gaga e comum.

TIPOS DE INS	GAGOS	NÃO GAGOS
/né/	√	√
/tá/	√	
/ué	√	
/daí/	√	
/intão/	√	√
/intão é/		√
/intão aí/	√	
/tão/	√	√
/aí/	√	√
/assim/	√	√

/sabi/	√	√
/tão é/	√	
/intão assim/	√	
/i/	√	√
/é/	√	√
/é i/		√
/i é/	√	
/i tal/		√
/i né/	√	
/ishi i/	√	

Fonte: própria.

O quadro 4 identifica as INSs de vogais e palavras para os informantes gagos e não gagos. O símbolo (√) evidencia a ocorrência da INS conforme o falante. A INS do tipo /né/ pode ocorrer em meio ou fim de GT. Quando ela é realizada em meio de GT indica que o enunciado apresenta dois grupos tonais, sendo que a INS_p /né/ representa o fim do primeiro GT em um mesmo enunciado. Observe abaixo:

- EN3 “é foi um crescimentu *is né* crescimentu na fé mesmu#”.

Em EN3, CL inicia o GT com uma INS_v /é/. Na sequência, ele realiza uma disfluência do tipo FC /is/, completando o GT com uma INS_p do tipo /né/, que fecha o GT com uma afirmação retórica. É como se CL subentendesse que o seu interlocutor tivesse entendido, de acordo com o contexto, que o que ele passou foi um crescimento. Dessa maneira, para confirmar, faz uso da INS_p /né/, fechando, assim esse GT. Em seguida, o falante explica por meio do segundo GT o tipo de crescimento a fim de encerrar a sequência discursiva.

Por sua vez, as INSs do tipo /tá/ e /ué/ são realizadas apenas em fim de GT e /daí/ em início ou meio, bem como /intão/ e /aí/. Já /intão é/ e /intão aí/ foram realizados em início de enunciado, assim como /tão/. Contudo, nesse contexto, o primeiro tipo ocorreu apenas para os não gagos. Vale salientar que /intão é/ também ocorreu em meio GT. A INS /assim/ foi encontrada em início, meio e fim de GT. No entanto, no primeiro contexto ela ocorreu apenas para os informantes com disfluência gaga. A INS /sabi/ ocorreu apenas em final de GT. As INSs /tão é/, /i é/, /intão assim/, /i assim/, /ishi i/ e /é i/ são realizadas em início de GT. No

entanto, os cinco primeiros casos ocorrem apenas para os informantes gagos e o último para os não gagos. Por outro lado, /i tal/ ocorre apenas para não gagos em fim de GT. Por fim, as INSs em vogais /i/ e /é/ podem ser realizadas tanto em início quanto em meio de GT para os gagos e não gagos. No mesmo contexto, elas variaram entre simples e complexas, sendo que a INS_v do tipo /i/ complexa (alongada) foi realizada apenas para os não gagos.

Como podemos ver, são poucas as INSs que ocorrem para um grupo de informante, mas não para o outro. Mesmo assim, vemos que os falantes gagos realizam mais variedades de INSs do que os não gagos, sendo 85% do total de INSs para os disfluente gagos e 55% para os informantes com disfluência comum.

Em relação ao tipo de INS mais recorrente para os informantes, foi observado tanto para os informantes gagos quanto para os não gagos que as INSs de palavras são as mais comuns. Mesmo assim, ocorrem algumas diferenças. Observe os dois quadros que descrevem essa afirmação.

Quadro 5 – Descrição das INSs para os informantes com disfluência gaga.

INFORMANTES GAGOS	PALAVRA E/OU VOGAL	CONTEXTO	TIPO
AP	PALAVRA	MEIO	/ASSIM/
CAM	PALAVRA	INÍCIO/FIM	/ASSIM/ e /NÉ/
JO	EMPATE	FIM (PALAVRA) INÍCIO (VOGAL)	/NÉ/ e /I/
MAU	PALAVRA	INÍCIO	/AÍ/
PA	PALAVRA	FIM	/SABI/
VA	PALAVRA	INÍCIO/FIM	/AÍ/ e /NÉ/

Fonte: própria.

Quadro 6 - Descrição das INSS para os informantes com disfluência comum.

INFORMANTES GAGOS	NAÕ	PALAVRA E/OU VOGAL	CONTEXTO	TIPO
BR		VOGAL	MEIO	/É/
CL		PALAVRA	FIM	/NÉ/
CR		PALAVRA	INÍCIO	/INTÃO/
LE		PALAVRA	FIM	/NÉ/
MA		PALAVRA	FIM	/NÉ/
NE		VOGAL	MEIO	/É/

Fonte: própria.

Nos dois quadros (5 e 6) a primeira coluna identifica o informante, a segunda, se a INSS é de palavra ou de vogal, a terceira o contexto em que ocorre e a última o tipo de INSS.

O quadro 5 mostra que para os informantes com disfluência complexa a INSS_p ocorre para 83% dos falantes. O resultado só não chega aos 100% dos casos porque para JO há um empate entre as INSSs de vogal e palavra. Já para os informantes com disfluência comum as INSSs de palavra ocorrem para 66,7% dos falantes. A maior variação entre os informantes ocorre para o contexto e tipo de INSS_p. Os informantes gagos variam mais. Observe que mesmo a INSS_p /né/ ocorrendo em três falantes (CAM, JO e VA), ela não é única, empatando com outro tipo de INSS (/assim/, /i/ e /ái/). Já para o contexto, não é possível estabelecer um padrão, uma vez que mesmo a INSS varia em início e fim de GT. Veja que há um empate para CAM, JO e VA. Por outro lado, para AP as INSSs são mais comuns em meio de GT, para MAU em início e para PA em fim de GT, Por sua vez, para os informantes com disfluência comum essas diferenças são menores. Observe por meio do quadro 6 que a INSS_p /né/ ocorre para 50% dos informantes (CL, LE e MA), sendo que os outros 50% variam entre /é/ (30%) e /intão/ (20%). Proporção que também é identificada para o contexto, uma vez que para fim de GT a proporção é de 50%, para meio de GT é de 30% e para início de 20%.

No que diz respeito à duração dessas INSSs, apresentamos, por meio da tabela 10, os seguintes resultados:

Tabela 10 – Duração total das INSs para os informantes gagos e não gagos.

INS/CONTEXTO	DURAÇÃO/ GAGOS	DURAÇÃO/ NÃO GAGOS
/NÉ/ (FIM)	7,814	14,408
/NÉ/ (MEIO)	1,028	0,531
/INTÃO/ (IN)	3,781	4,727
/INTÃO (MEIO)	1,047	0,782
/TÃO/ (IN)	0,096	1,626
/Á/ (IN)	2,863	1,173
/Á/ (MEIO)	0,758	0,687
/ASSIM/ (MEIO)	1,688	1,433
/ASSIM/ (FIM)	0,356	0,459
/SABI/ (FIM)	1,688	1,433
/I/ (IN SIMPLES)	2,009	0,503
/É/ (IN SIMPLES)	2,316	4,365
/I/ (MEIO SIMPLES)	1,377	1,744
/É/ (MEIO SIMPLES)	1,997	7,238
/É/ (IN COMPLEXA)	2,141	2,046
/É/(MEIO COMPLEXA)	1,761	0,657

Fonte: própria.

A tabela 10 identifica, na primeira coluna, as INSs comuns de palavras e vogais (simples e complexas) que ocorreram tanto para os informantes gagos quanto para os não gagos. Nela também está identificado o contexto de cada INS. A segunda e a terceira coluna, por sua vez, identificam a soma da duração (em ms) para os informantes gagos e não gagos, respectivamente. A partir disso, vemos que para as INSs de palavra a duração empatou para os informantes com disfluência gaga e comum, ou seja, em 50% das INSs foi maior para o primeiro grupo (/né/ em meio de GT; /intão/ em meio de GT e /ái/ em início e meio de GT) e em 50% para o segundo (INSs de palavras restantes). Já para as INSs de vogal simples, a duração foi maior para os falantes com disfluência comum em 75% das INSs (/é/ em início de GT, /i/ e /é/ em meio de GT). Por fim, para as INSs complexas em 100% dos casos (/é/ em início e /i/ em meio) a duração foi maior para os informantes com disfluência gaga.

No que diz respeito à F0 média, pulsos, intensidade de transição, final em relação à inicial e intensidade máxima, verificamos que a F0 média é maior para os não gogos em palavras, sendo maior para 80% dos falantes. Já para os informantes com disfluência gaga elas variaram para as INSs de vogais (50%) e INSs de palavras (50%). A mesma relação foi observada para os pulsos nos dois grupos de informantes. Em relação à intensidade de transição, como vimos no quadro 4 há uma variação no contexto em que as INSs ocorrem, o que impossibilita que seja proposta uma tendência para o padrão. Quanto aos informantes não gogos essa tendência também não foi possível, uma vez que para 10% dos informantes as INSs ocorrem em início de GT ou após uma PS; para os outros 10% a tendência é para que a intensidade aumente o padrão; para 40% ocorre um empate entre aumento e diminuição do padrão e, por fim, para os outros 40% a tendência é para que ocorra a diminuição no padrão. Sendo assim, para os dois grupos de informantes não é possível propor um padrão para a intensidade. Por sua vez, a análise da intensidade final, quando comparada à inicial, mostrou uma tendência para o abaixamento em 90% dos casos para os informantes com disfluência gaga e uma tendência para o aumento da intensidade para os informantes com disfluência comum em 80% das INSs. Finalmente, a intensidade máxima foi a única que apresentou a mesma tendência para os dois grupos de informantes, já que ao compararmos a média da intensidade máxima das INSs de palavras e de vogais, foi observado que ela é maior para as INS_{ps}, ocorrendo para 100% dos informantes gogos e para 90% dos não gogos.

Por fim, tomando como base as INSs que foram comuns para os dois grupos de informantes, a análise mostrou uma tendência para que INSs do mesmo tipo mantenham o mesmo padrão entoacional.

Em relação à disfluência do tipo HES, não foram todos os informantes que a realizam, tanto no grupo da disfluência gaga quanto na comum. Dos falantes analisados, 70% deles realizaram uma HES, seja para os gogos ou não gogos. As HES podem ocorrer em início ou meio de GT. Os falantes com disfluência gaga realizaram apenas HES em início de GT. Já dentre os informantes com disfluência comum, dois deles (CL e CR) realizaram HES em início e meio de GT, um em início (LE) e, por fim, apenas (NE) em meio. Todas as HESs foram do tipo /ah/.

Em relação à duração e aos outros fenômenos acústicos, as HESs comparadas foram as de início de GT, já que são comuns aos dois grupos. Nesse caso, ela foi maior para os informantes com disfluência comum, apresentando uma duração total de 3,618ms. Por outro

lado, a duração para os informantes com disfluência gaga foi de 2,324ms, sendo a diferença entre eles de 1,294ms.

Quanto à intensidade, a de transição não foi levada em consideração para as HESs em início de GT. No que diz respeito à intensidade máxima, podemos considerar que elas foram relativamente iguais, uma vez que a diferença entre elas foi de apenas 1 dB, sendo a média entre os informantes gagos de 78 dB e a dos não gagos de 77 dB. A intensidade final, por sua vez, não apresentou um padrão para os informantes com disfluência comum, uma vez que para um informante ela diminuiu em relação à intensidade inicial (CL), para o outro equiparou-se (CR), de modo que para uma HES aumentou e para a outra diminuiu. E, por fim, para o outro falante a intensidade final aumentou em relação à inicial (LE). Já para os informantes com disfluência gaga, para 60% deles ocorreu uma queda da intensidade final em relação à inicial. Finalmente, também com uma diferença pequena, a F0 média foi maior para os informantes com disfluência gaga, sendo de 132 Hz aproximadamente, enquanto que para os informantes com disfluência comum foi de 117 Hz aproximadamente.

Finalmente, a análise entoacional revelou para 90% dos informantes gagos o padrão *H+L, sendo que apenas um (VA) apresentou uma HES com uma qualidade de voz do tipo *creaky voice*. Por outro lado, os informantes com disfluência comum, dois deles apresentaram para início de GT uma qualidade de voz do tipo *creaky voice*. O outro informante (LE), por sua vez, variou entre *H e *H+L. No entanto, a porcentagem maior foi para o segundo padrão (86% aproximadamente). Isso mostra que as HESs para os dois grupos analisados iniciam com um tom mais alto (*H) e depois caem (L).

Assim como a HES, a disfluência do tipo FC também não foi comum para todos os falantes, ocorrendo para 50% dos informantes não gagos e para 90% dos informantes gagos. No que concerne à duração, ela também foi maior para o último grupo (2,076 ms), enquanto que para o primeiro foi de 1,698ms. Quanto à intensidade de transição, os informantes gagos apresentaram uma tendência para a diminuição do padrão para 80% dos informantes. Já para os não gagos não foi possível estabelecer um padrão, já que para CL a intensidade aumentou, para MA diminuiu e para NE variou entre as duas possibilidades. Por sua vez, a intensidade final apresentou uma queda em relação à intensidade inicial para 100% dos informantes não gagos. Para os informantes com disfluência gaga o padrão também foi esse, mas para 80% dos falantes. A intensidade máxima, foi maior para a disfluência gaga com uma média aproximada de 71 dB. Os falantes com disfluência comum apresentaram uma intensidade máxima média de 65 dB aproximadamente, que em comparação a dos não gagos, é pequena.

Finalmente, a F0 média foi menor para os informantes com disfluência comum, de 114 Hz, enquanto a dos informantes gogos foi de 124 Hz. Essa diferença de 10 Hz também não é auditivamente significativa.

Quanto ao padrão entoacional, fica difícil estabelecer comparações, uma vez que os FCs não são os mesmos para os informantes.

O AL, por sua vez, foi realizado para 100% dos informantes com disfluência comum e para 90% dos informantes com disfluência gaga. O falante do último grupo que não apresentou essa disfluência foi PA que, por sua vez, apresenta uma fala caracterizada por ser mais entrecortada. Para o primeiro grupo essa disfluência foi a quinta mais frequente e para o segundo, a sexta.

Para os informantes gogos a duração total do AL foi de 5,998ms, enquanto que a dos não gogos foi de 21,740ms, o que revela uma duração aproximadamente 78% maior do que a do primeiro grupo.

No que diz à análise da intensidade, foi verificado para 90% dos informantes com disfluência gaga uma intensidade de transição com um padrão de queda em relação ao contexto anterior. Já para os informantes com disfluência comum, 50% dos Als foram antecidos por uma PS, os outros 50% variaram entre aumento e diminuição no padrão. A intensidade final também foi contrária para os dois grupos. Para os informantes com disfluência gaga o padrão também foi de queda para 90% dos falantes. Já para os não gogos 70% dos falantes apresentaram um padrão com tendência ascendente. Por sua vez, a intensidade máxima equiparou-se para os dois grupos com uma média aproximada de 76 dB. Esse resultado equivalente pode ser apenas uma coincidência. No entanto, a análise das outras disfluências mostrou, até o momento, uma diferença pouco significativa entre os grupos. Finalmente, a F0 média foi maior para os informantes com disfluência gaga (154 Hz aproximadamente) do que para os não gogos (138 Hz).

A análise entoacional também se distingue entre os grupos uma vez que os contextos são diferentes.

As PIs também foram realizadas pelos dois grupos de informantes. Dentre os informantes não gogos, apenas 30% deles realizaram esse tipo de disfluência, já para os informantes com disfluência gaga a sua frequência foi de 50%. Para o primeiro grupo, esse tipo de disfluência ocupou a oitava posição dentre os eventos, perdendo apenas para o FC que ficou em último lugar. Por outro lado, para o segundo grupo a PI ficou em último lugar.

A duração total das PIs para os informantes gagos foi de 5,786ms e para os não gagos foi de 2,181ms. Portanto, maior para o primeiro grupo. No que concerne à intensidade, a análise mostrou para a de transição um padrão de queda para os dois grupos, sendo de 100% para os informantes com disfluência comum e de 70% para os disfluentes gagos. A intensidade final seguiu o mesmo padrão e porcentagem para os informantes gagos. No entanto, para os não gagos, não foi possível chegar a um resultado, uma vez que dentre os dois informantes, CR apresentou uma tendência crescente para a intensidade final e para MA os resultados equipararam-se, de modo que 50% da intensidade foi realizada com um padrão crescente e 50% com um padrão decrescente. Finalmente, a F0 média revelou uma frequência significativamente irrelevante para os dois grupos, sendo de 119 Hz para a disfluência comum e 118 Hz para a disfluência gaga. É importante salientar que dentre os informantes gagos, a F0 não foi medida apenas para JO porque das três realizações, uma foi surda e nas outras duas ocorreu um erro de gravação.

Para dar conta da análise as disfluências do tipo CORs, como vimos, foram divididas entre simples e complexas. No último caso, elas podem ser realizadas com uma PS, RP, AL, entre outros tipos de disfluências. No que diz respeito à sua posição em relação às outras disfluências, elas foram mais frequentes para a disfluência gaga, ocupando o quarto lugar. Por sua vez, para a disfluência comum ocupou o quinto lugar. No entanto, quanto à quantidade, os resultados foram opostos. Os informantes não gagos realizaram 88 CORs, enquanto os não gagos 35, ou seja, aproximadamente 71,5% a mais do que os informantes gagos. Todos os falantes analisados realizaram CORs.

As CORs podem ocorrer em início ou meio de GT, mas nunca em final. Nos dois contextos, o número de ocorrência foi maior para os informantes com disfluência comum, mas os dois grupos realizaram um número maior de disfluências em meio de GT. Observe a tabela 11 abaixo:

Tabela 11 – Descrição da quantidade e contexto das CORs de acordo com o grupo de informantes.

NÃO GAGOS	COR EM INÍCIO DE GT	COR EM MEIO DE GT	GAGOS	COR EM INÍCIO DE GT	COR EM MEIO DE GT
BR	3	10	AP	1	2
CL	5	7	CAM	1	4
CR	4	4	JO	1	10
LE	0	8	MAU	1	4
MA	12	12	PA	4	1

NE	5	18	VA	2	4
TOTAL	29	59	TOTAL	10	25

Fonte: própria.

A tabela 11 apresenta os informantes não gogos e gogos e o número de disfluências de acordo com a sua posição no enunciado. Como podemos ver, os não gogos realizam o maior número de CORs em início e meio de GT. No entanto, dentre as duas possibilidades, os dois grupos realizaram mais disfluências em meio de GT. Dentre os informantes não gogos, apenas LE não realizou uma COR em início de GT. Já em relação aos informantes gogos, todos realizaram os dois tipos de CORs. É importante observar ainda que os informantes com disfluência gaga realizaram, no mínimo uma COR em início e meio de GT e no máximo 10 em meio de GT. Por sua vez, os falantes com disfluência comum, a não ser por LE, realizaram no mínimo três CORs em início de GT e quatro em meio. Já o máximo de CORs em início de GT foi de 12 e em meio de 18.

Para esse tipo de disfluência, a duração foi analisada para as CORs simples e complexas. Nos dois casos, ela foi maior para a disfluência comum. A COR simples apresentou uma duração total de 16,001ms para os falantes não gogos e 3,623ms para os gogos. Já a COR complexa apresentou uma duração de 12,213ms para o primeiro grupo e 4,701ms para o segundo. Os resultados mostram, ainda, que as CORs simples são maiores do que as complexas para os falantes sem disfluência gaga. Por sua vez, as CORs complexas são maiores para os informantes com disfluência gaga. Ainda para a duração, analisamos se a sua tendência é para o aumento, no momento da correção, ou para a sua diminuição. A análise revelou uma tendência, para os dois grupos, para o aumento da duração, como caracterizado na tabela 12 abaixo:

Tabela 12 – Tendência da duração das CORs simples e complexas conforme o grupo de informantes.

DURAÇÃO	DIMINUIÇÃO DA DURAÇÃO		AUMENTO DA DURAÇÃO	
	COR SIMPLES	COR COMPLEXA	COR SIMPLES	COR COMPLEXA
CORs SIMPLES E COMPLEXAS				
INFORMANTES GAGOS	1,795	1,970	1,828	2,731
INFORMANTES NÃO GAGOS	3,052	4,512	12,949	7,701

Fonte: própria.

A tabela 12 descreve para os informantes gogos e não gogos a duração das CORs simples e complexas de acordo com a sua tendência para o aumento ou diminuição. Esses

valores foram medidos por meio da soma das durações de cada informante. Para isso, primeiro, observamos qual era a tendência da duração durante a passagem do erro para a correção para cada informante. Depois, subtraímos os valores e os separamos conforme a sua tendência e tipo de COR. Na sequência, os valores foram somados e chegamos ao resultado apresentado na tabela 12 acima. Como podemos ver, a tendência é para que ocorra um aumento da duração durante a passagem do erro para a correção para os falantes gogos e não gogos tanto para a COR simples quanto para a COR complexa.

No que se refere à F0 para os informantes com disfluência comum e gaga, a COR simples apresentou uma tendência ao aumento, no momento do reparo⁶⁵, para 70% dos informantes nos dois grupos. Já para a COR complexa⁶⁶ a tendência é para que 50% dos falantes seja com disfluência simples ou gaga apresentem uma diminuição da F0 na correção. No que diz respeito aos outros 50%, para 40% dos falantes com disfluência comum a tendência foi para o aumento durante a correção. Os outros 10% apresentaram resultados que se equipararam. Por outro lado, os outros 50% dos falantes com disfluência gaga apresentaram uma tendência ao aumento da F0 durante a correção.

Quanto à intensidade de transição há uma tendência entre os informantes com disfluência comum para a diminuição do padrão na correção em 50% dos eventos. O restante divide-se entre resultados comuns (40%) e tendência ao aumento do padrão (10%). Já para os informantes com disfluência gaga, a tendência também é para a diminuição do padrão no reparo, mas para 70% dos falantes. A COR complexa mantém a mesma característica e porcentagem para a disfluência comum. No entanto, os outros 50% dividem-se em 40% para o aumento do padrão e em 10% para resultados iguais. Por sua vez, na disfluência gaga, para 70% dos falantes não foi possível estabelecer um padrão, uma vez que os resultados mantiveram-se iguais para o aumento e diminuição da intensidade. Os outros 30% revelaram uma tendência à queda.

No que concerne ao padrão da intensidade final, foi verificada uma tendência ao aumento da intensidade para 50% dos falantes com disfluência comum. Por outro lado, 50% dos falantes com disfluência gaga revelaram uma tendência ao abaixamento do padrão da intensidade. Já para a COR complexa, 70% dos informantes com disfluência comum analisados apresentaram resultados que se equipararam entre aumento e diminuição do

⁶⁵ Vale ressaltar que os eventos acústicos estão sendo observados no momento em que o falante se corrige e não durante o erro.

⁶⁶ Também é importante salientar que para a análise da COR complexa, a informante gaga CAM não participa dos resultados obtidos, uma vez que realizou apenas CORs simples.

padrão. Os informantes com disfluência gaga também não apresentaram uma tendência, uma vez que 40% deles apresentaram resultados que se mantiveram iguais, 40% uma tendência ao aumento do padrão e 20% à diminuição.

Finalmente, a intensidade máxima foi maior no momento de reparo, em COR simples para 50% dos informantes com disfluência comum. No mesmo contexto, 70% dos falantes com disfluência gaga apresentaram um abaixamento durante a correção. O padrão obtido para a COR simples com os informantes não gogos manteve-se para a COR complexa, mas, nesse caso, para 100% dos informantes. Fato que não foi observado para a disfluência gaga, uma vez que 40% dos falantes apresentaram uma tendência ao abaixamento da intensidade, 40% para o aumento e 20% mantiveram os mesmos resultados.

As CORs também variaram conforme o padrão entoacional do erro e do reparo. Para os informantes com disfluência comum, 43,7% deles mantiveram o mesmo tom (L ou H) do erro no momento da correção. Para a disfluência gaga a porcentagem dos informantes que mantiveram o mesmo tom tanto no erro quanto no reparo foi de aproximadamente 47%. Como podemos ver, temos uma diferença aproximada de 4% entre os grupos, o que não é estatisticamente relevante.

A disfluência do tipo BLO também foi verificada para os informantes gogos e não gogos. Para os primeiros falantes essa disfluência ocupou o quarto lugar em ordem de realização e o último grupo, o sexto lugar, dentre as nove possibilidades de disfluências que os falantes gogos e não gogos realizaram. Para os informantes com disfluência comum foram verificados 16 casos de BLOs, sendo 14 em consoantes e dois em vogal. Já para a disfluência gaga foram identificados 74 casos, sendo 54 em consoantes e 20 em vogais. Como vemos, mesmo a proporção sendo maior para os informantes com disfluência gaga, ambos realizam um maior número de BLOs em consoantes. A tabela 13 abaixo identifica os tipos de consoantes e a porcentagem de ocorrência para cada informante.

Tabela 13 – Descrição da porcentagem de ocorrência dos BLOs em consoante para os informantes com disfluência comum e gaga.

CONSOANTES	DISFLUÊNCIA COMUM	DISFLUÊNCIA GAGA
[p]	21%	17%
[b]	15%	8%
[k]	36%	20%
[g]	0%	4%
[t]	7%	9%
[d]	0%	7%
[v]	0%	20%

[s]	0%	5%
[m]	7%	5%
[n]	14%	4%

Fonte: própria.

A tabela 13 descreve a porcentagem aproximada de ocorrência das consoantes para cada grupo de informantes de acordo com o total realizado, como especificado. Se separarmos os BLOs conforme o modo de articulação, veremos que as consoantes oclusivas são as mais frequentes para os dois grupos, ocorrendo em 79% dos BLOs para os falantes não gogos e em 65% dos casos para os falantes gogos. Dentre elas, a mais comum para os dois grupos é a oclusiva velar desvozeada [k]. Os falantes com disfluência comum diferentemente dos disfluentes gogos não realizaram oclusivas alveolares vozeadas [d] e oclusivas velares vozeadas [g]. Na sequência, os informantes gogos realizaram 25% de BLOs em fricativas labiodental [v] e alveolar [s], o que não foi identificado para os informantes com disfluência comum. Estes, por sua vez, produziram os BLOs em consoante nasal [m] e [n] como o segundo tipo de bloqueio mais frequente (21%). Já para os informantes com disfluência gaga o modo de articulação nasal ocupou o terceiro lugar (9%).

Por outro lado, para os BLOs em vogais tivemos a seguinte distribuição:

Tabela 14 – Descrição da porcentagem de ocorrência dos BLOs em vogais para os informantes com disfluência comum e gaga.

VOGAIS	DISFLUÊNCIA COMUM	DISFLUÊNCIA GAGA
[a]	50%	20%
[u]	50%	5%%
[e]	0%	30%
[ɛ]	0%	30%
[o]	0%	5%
[i]	0%	10%

Fonte: própria.

A tabela 14 identifica as vogais bloqueadas para a disfluência comum e gaga. Os informantes do primeiro grupo realizaram apenas dois tipos de vogal bloqueada [a] e [u]. Já os informantes com disfluência gaga realizaram, no total, 20 BLOs em vogal (identificadas na tabela acima), assim como foi mencionado. Dentre as possibilidades de regiões articulatórias, 90% foram realizadas em região anterior e 10% em posterior.

Como foi descrito, as pausas de bloqueios podem ocorrer da seguinte maneira:

- tipo 1 – pausa com presença de ciclos de onda;
- tipo 2 – pausa sem marcação de ciclos de onda;
- tipo 3 – ausência de pausa de bloqueio;

- tipo 4 – pausa respiratória (auditiva).

Diante dessas possibilidades, a mais frequente para os não gogos foi a do tipo 1, com 44% do total, assim como para os informantes com disfluência gaga, que representou aproximadamente 48% das ocorrências.

Quanto à duração dos segmentos bloqueados e das PSs que antecederam os BLOs, os informantes com disfluência gaga, no total, apresentaram uma duração maior tanto para a o segmento quanto para a pausa. Para os informantes não gogos a duração do segmento bloqueado foi de 8,380ms enquanto a dos gogos foi de 33,492ms. Já a duração das PSs foi de 5,613ms. para o primeiro grupo e 28,487ms para o segundo. A diferença entre eles é grande. No entanto, temos que levar em consideração a quantidade de ocorrência da disfluência para um e outro grupo. Mesmo assim, é possível afirmar que os informantes gogos demoram mais tempo para sair da PS de bloqueio e do bloqueio no segmento, uma vez que a comparação individual dos informantes revelou durações maiores nos dois casos, para 100% dos informantes com disfluência gaga.

No que diz respeito às intensidades, a de transição não foi observada, uma vez que a maior parte dos BLOs foi antecedida por PSs. Por sua vez, a média da intensidade máxima entre os informantes mostrou uma diferença de apenas 4 dB entre os grupos, sendo de 75 dB aproximadamente para os informantes com disfluência gaga e de 79 dB aproximadamente para os informantes com disfluência comum. A intensidade final também apresentou o mesmo padrão crescente para os dois grupos, sendo de 70% para a disfluência gaga e 90% para os não gogos. Finalmente, a média da F0 para os falantes também foi maior para a disfluência comum, de 153 Hz aproximadamente, enquanto que os informantes com disfluência gaga apresentaram uma média para a F0 de aproximadamente 145 Hz, o que também é uma diferença (8 Hz) auditiva insignificativa.

Por fim, a análise entoacional variou conforme o segmento bloqueado para os dois grupos em tons simples e compostos.

As RPs junto com as PSs, como mencionado, foi a disfluência mais realizada para os informantes com disfluência gaga, enquanto para os informantes com disfluência comum, foi a terceira mais frequente. Mesmo estando em uma posição de diferença, a quantidade total realizada por cada grupo foi consideravelmente relevante, sendo 287 RPs para os informantes gogos e 87 para os não gogos. Isso mostra que os informantes com disfluência gaga realizaram aproximadamente 77% a mais de repetições do que os não gogos. Outra diferença considerável entre eles, como vimos, diz respeito às RPs complexas, que representa, para a

disfluência gaga, 28,6% do total e para a disfluência comum, 18,3%. O máximo de RPs complexas realizadas por um informante não gago, por exemplo, foi de cinco (BR), enquanto que para um informante gago (PA) foi de 18 de um total de 23 repetições. Todos os informantes gagos realizaram uma RP complexa. Por outro lado, um dos informantes não gago (CL) não realizou nenhum tipo de RP complexa. O total de palavras repetidas pode chegar até cinco para a disfluência gaga, enquanto que para a disfluência comum, até três.

Dentre as possibilidades de RPs tivemos os seguintes resultados:

Tabela 15 – Porcentagem de ocorrência dos tipos de RPs⁶⁷.

POSSIBILIDADES DE RPs	INFORMANTES GAGOS	INFORMANTES NÃO GAGOS
RP _v	7%	18%
RP _p	52%	59%
RP _s	27%	13%
RP _e	14%	10%

Fonte: própria.

A tabela 15 descreve a porcentagem de RPs em vogal, palavra, sílaba e estrutura para os informantes gagos e não gagos. Para os dois tipos de disfluência, a RP em palavra é a mais realizada, 52% para o primeiro e 59% para o segundo. Em segundo lugar, temos a RP de sílaba para a disfluência gaga (27%) e a de vogal (18%) para a disfluência comum. Depois, aparecem a RP de estrutura para os informantes gagos (14%) e a RP de sílaba (13%) para os não gagos. Finalmente, temos a RP_v (7%) para o primeiro grupo e a RP_e (10%) para o segundo.

A partir dessa distribuição observamos o que ocorreu com a duração da segunda RP na passagem de uma repetição para outra, que poderia aumentar ou diminuir. Os resultados foram alcançados a partir da soma da diferença das durações dos informantes de cada grupo. Primeiro, foram anotadas as durações das repetições que aumentaram. Sendo assim, tivemos os seguintes resultados:

Tabela 16 – Duração total das repetições que aumentaram para os informantes com disfluência gaga e comum.

POSSIBILIDADES DE RPs	DURAÇÃO DISFLUÊNCIA GAGA	DURAÇÃO DISFLUÊNCIA COMUM
RP _v	2,428	0,806
RP _p	5,047	1,028
RP _s	15,340	0,174

⁶⁷ É importante deixar claro que há os subtipos de RPs, como, por exemplo, 3RP_e, 4RP_p, 5RP_{ps}, entre outras. No entanto, preferimos anotá-las de acordo com a sua maior frequência, a fim de facilitar a análise. A especificação de cada variação das RPs pode ser vista nos itens 8.1 e 8.2.

RP_e	0,963	0,983
-----------------------	-------	-------

Fonte: própria.

Dentre as possibilidades de RPs, os informantes gogos apresentaram uma maior duração para 90% dos casos (cf. tab. 16). O único tipo de RP para os informantes com disfluência comum que apresentou uma duração total maior do que a dos não gogos foi a repetição de estrutura (RP_e). Mesmo assim, a diferença entre elas foi de 0,020ms. Como vimos na tabela 16, a porcentagem de ocorrência das RPs é, em 80% dos casos, maior para os informantes com disfluência gaga. No entanto, mesmo nos eventos em que os informantes não gogos apresentaram uma maior porcentagem (RP_v e RP_p), os falantes com disfluência gaga apresentaram uma duração maior.

Quanto às durações que diminuíram na passagem de uma RP para a outra, os resultados foram:

Tabela 17 – Duração total das repetições que diminuíram para os informantes com disfluência gaga e comum.

POSSIBILIDADES DE RPs	DURAÇÃO DISFLUÊNCIA GAGA	DURAÇÃO DISFLUÊNCIA COMUM
RP_v	4,870	3,075
RP_p	9,116	5,918
RP_s	12,403	0,577
RP_e	3,066	0,693

Fonte: própria.

A tabela 17 identifica a soma das durações que diminuíram para as quatro possibilidades de repetições mostra para 100% dos casos a maior duração foi para os informantes com disfluência gaga.

Ao compararmos as tabelas 16 e 17 vemos que para 90% das RPs, tanto para os informantes com disfluência gaga quanto comum, a duração foi maior para a soma das durações que diminuíram. Esse resultado também revela que é mais comum que os falantes diminuam a duração durante a repetição dos segmentos. Os únicos casos em que esse resultado não se manteve foi para a RP de estrutura para os informantes com disfluência comum e para a RP de sílaba para os informantes com disfluência gaga.

No que diz respeito à F0 e variações de intensidade (transição, final e máxima), a comparação foi feita por meio da observação se esses fenômenos acústicos apresentavam uma tendência à queda ou aumento para a segunda RP. Dessa maneira, a RP_v, apresentou para a F0 uma tendência para a diminuição da frequência para os informantes com disfluência comum (60%) e gaga (70%). A intensidade de transição seguiu a mesma tendência, sendo a

porcentagem também de 60% para o primeiro grupo e 80% para o segundo. Já a intensidade final e máxima apresentou uma tendência ao aumento no padrão para 60% dos informantes com disfluência comum. Os informantes com disfluência gaga, por sua vez, não apresentaram um padrão, variando entre aumento e diminuição para a intensidade final e, para a intensidade máxima, 50% dos informantes apresentaram uma tendência para a queda na segunda RP e 50% variaram entre aumento e diminuição do padrão.

A RP_p , por sua vez, também apresentou uma diminuição na F0 na segunda RP para a disfluência gaga (90%) e comum (60%), bem como a intensidade de transição, sendo de 90% para as duas disfluências. Já a intensidade final aumentou para os dois grupos, 90% para os gogos e 60% para os não gogos. Finalmente, a intensidade máxima foi diferente para as disfluências, diminuindo para 70% dos informantes com disfluência gaga e aumentando para 60% dos falantes com disfluência comum.

Quanto à RP_s , não foi possível estabelecer um padrão da F0 para os informantes com disfluência comum, já que ela variou entre aumento, diminuição e, ainda, apresentou situações em que os informantes mantiveram 50% dos resultados para as duas possibilidades. Por outro lado, para a disfluência gaga houve a diminuição para 70% dos falantes. No que diz respeito à intensidade, houve um aumento no padrão da intensidade de transição, final e máxima para os informantes com disfluência gaga em respectivamente 70%, 90% e 80%. Já para a disfluência comum, a intensidade de transição diminuiu para 70% dos falantes, a final manteve as características da F0 e a máxima aumentou para 70% dos informantes.

Finalmente, a RP_e apresentou uma redução na F0 para 70% dos informantes com disfluência gaga na segunda estrutura repetida. Já para os falantes com disfluência comum, o padrão variou em redução e aumento para 50% dos informantes. Quanto à intensidade de transição, ocorreu o aumento do padrão para 90% dos falantes gogos. Os não gogos, por sua vez, não apresentaram um padrão definido, variando entre aumento, diminuição e em situações em que os informantes apresentaram as duas tendências. A intensidade final seguiu os mesmos critérios da de transição para a disfluência comum, enquanto que para a disfluência gaga houve uma diminuição no padrão para 60% dos falantes. A intensidade máxima diminuiu para 70% dos informantes com disfluência gaga e aumentou para 70% dos informantes com disfluência comum.

Portanto, a F0 apresentou para 100% das RPs a tendência para a redução da frequência na disfluência gaga. Já para a comum, 50%. A intensidade de transição, por sua vez, para 90% das RPs em disfluência comum diminuiu. Em contrapartida, para a disfluência gaga ela variou

em 50% para o aumento e 50% para a redução. Quanto à intensidade final, os informantes com disfluência gaga e comum apresentaram uma tendência ao aumento para 50% das RPs. Por fim, para intensidade máxima apresentou para 100% das RPs em disfluência comum um aumento no padrão. Já para a disfluência gaga, 50% das RPs foram analisadas com uma tendência para diminuição.

A análise dos fenômenos acústicos revalou que a F0 e variações de intensidade não distanciam significativamente os informantes gagos dos não gagos, pois mais de 50% dos padrões acústicos que acontecem para um grupo de informantes também é verificado no outro. A maior diferença entre a disfluência comum e gaga está, até o momento, na duração que é significativamente maior para a disfluência gaga.

Por sua vez, a tessitura foi analisada na fala dos informantes com disfluência gaga e comum por meio do programa PRAAT. Primeiro, foram anotadas a F0 mais alta e mais baixa⁶⁸ de cada um dos informantes durante a realização das disfluências. Depois, os valores de cada tipo (RPs, INSS, BLOs) de disfluência foram somados para cada falante. Na sequência, foi feita a média aritmética para cada informante e, por fim, para os grupos (gagos e não gagos). Veja na tabela abaixo a tessitura total de cada falante e a média aritmética para os grupos.

Tabela 18 – Descrição das F0 mínima e máxima e média aritmética para os informantes com disfluência comum e gaga.

INFORMANTES DISFLUÊNCIA COMUM	F0 MÍNIMA	F0 MÁXIMA	INFORMANTES DISFLUÊNCIA GAGA	F0 MÍNIMA	F0 MÁXIMA
BR	97,326	121,401	AP	98,274	120,761
CL	152,858	176,728	CAM	128,219	160,246
CR	106,013	129,046	JO	104,827	131,410
LE	138,575	175,328	MAU	166,428	134,041
MA	104,781	124,7923	PA	115,783	137,910
NE	99,968	117,058	VA	163,922	131,736
MÉDIA	116,586	140,725	MÉDIA	118,813	146,780

Fonte: própria.

⁶⁸ Esses valores são obtidos no programa por meio da seleção do segmento que se quer analisar. Na sequência, acionamos os comandos *pitch > get minimum pitch* (menor F0) e *pitch > get maximum pitch* (maior F0).

A média aritmética da soma das frequências mínima e máxima para as disfluências comum e gaga (cf. tab. 18) mostra que os informantes apresentaram uma variação de tessitura baixa. Observe que a diferença entre a variação de tessitura para os informantes com disfluência comum foi de aproximadamente de 24 Hz e para os informantes com disfluência gaga foi de aproximadamente 28 Hz. Veja ainda, por meio da comparação entre os grupos, que a diferença entre eles também é baixa. A diferença entre a F0 mínima foi de 2 Hz aproximadamente e entre a máxima foi de 6 Hz aproximadamente. Esse resultado não faz diferença auditivamente.

A análise individual dos falantes revela que os informantes gagos tendem a variar mais a tessitura do que os não gagos. Dentre eles, 66,7% dos falantes gagos apresentaram uma tessitura maior do que os não gagos (33,3%).

Finalmente, a tabela 19, abaixo, representa os resultados obtidos para os aspectos prosódicos temporais da entrevista dos falantes com disfluência gaga e comum. São apresentadas as médias obtidas para cada variável pesquisada.

Tabela 19 – Média dos aspectos prosódicos temporais dos informantes com disfluência gaga e comum.

MÉDIAS		
VARIÁVEL	DISFLUÊNCIA GAGA	DISFLUÊNCIA COMUM
Tempo total de elocução	174,652	148,355
Tempo de articulação	148	118
Taxa de elocução	4,569	4,600
Taxa de articulação	5,457	5,666
Número de sílabas	784	670
Duração de pausas	26,195	29,462

Fonte: própria.

Ao compararmos o tempo total de elocução dos dois grupos vemos que os informantes com disfluência gaga apresentaram um tempo médio mais longo (174 aprox.) do que os informantes com disfluência comum (148 aprox.). Esse resultado pode estar relacionado com o maior número de disfluências verificadas para a disfluência gaga.

O tempo de articulação, obtido pela subtração da duração total do texto pela duração das pausas, também foi maior para a disfluência gaga, com o valor médio de 148 (aprox.) e de 118 (aprox.) para a disfluência comum. Isso mostra que os falantes com disfluência gaga demandam um tempo bem maior do que os informantes não gagos para a articulação dos segmentos.

A taxa de elocução, que é mensurada a partir da divisão do número de sílabas emitidas pelo tempo de elocução, apresentou uma média de 4,569sil/ms para a disfluência gaga e 4,600sil/ms para a disfluência comum, o que não caracteriza uma diferença significativa entre os grupos. Isso mostra que os informantes apresentam uma taxa de elocução muito próxima e que não contribui para a sua separação.

A taxa de articulação, que é obtida através da divisão do número de sílabas emitidas pelo tempo de articulação, também indicou uma diferença pequena entre os grupos, sendo a média de 5,457sil/ms para a disfluência gaga e 5,666sil/ms para a disfluência comum. Isso mostra que há uma semelhança entre a produção do gesto articulatório para os informantes com disfluência gaga e comum.

Quanto ao número de sílabas, a média revelou um maior número de sílabas produzidas pelo grupo dos falantes gagos (784) do que para os não gagos (670). Essa diferença pode estar relacionada a processos fonológicos, como por exemplo, fenômenos de sândi.

Por fim, a média da duração das pausas, como vimos, foi maior para os informantes com disfluência comum (29,462ms), enquanto para a disfluência gaga foi de (26,195ms). Essa diferença pode significar uma menor organização (planejamento) discursiva na fala gaga, provocando, conseqüentemente, um número maior de disfluências.

CONCLUSÃO

O propósito deste trabalho foi entender, por meio dos fenômenos acústicos e prosódicos, o porquê da fala gaga causar uma empatia negativa no interlocutor. E, em consequência disso, tornar-se motivo de distanciamento, preconceito e negatividade. Fato esse que não é verificado em momentos de disfluência comum. A pressuposição era a de que, mesmo sendo fenômenos considerados diferentes, fosse possível, por meio de uma comparação, encontrar semelhanças dentro da singularidade de cada um dos dois tipos de fala.

Sendo assim, a comparação dos fenômenos acústicos para a disfluência gaga e comum revelou diferenças pouco expressivas para a F0 e os momentos de intensidade (transição, final e máxima). A análise da tessitura e dos padrões entoacionais também não distanciou os dois tipos de disfluências, uma vez que a análise do primeiro mostrou que tanto a disfluência comum quanto a gaga tendem a manter uma variação entoacional mais baixa. Quanto ao padrão entoacional, a variação melódica ocorreu de acordo com o tipo de informante e de disfluência. Apenas as disfluência do tipo RPs, HESs e INSs mantiveram, conforme o falante e contexto, para a maior parte dos dados, o mesmo padrão entoacional.

A duração foi o evento que mostrou maiores diferenças entre um e outro tipo de fala, sendo maior, em grande parte das vezes para as disfluências gagas. Esse resultado é um indício da empatia negativa em relação à fala de pessoas com disfluência gaga.

O ritmo está intimamente relacionado à duração. O fato do falante demorar mais para voltar à fluência, ou melhor, formular o conteúdo, faz com que ele tenha problemas com o ritmo da fala e, conseqüentemente, quebre a expectativa do ouvinte, levando, assim, à empatia (fonética) negativa em quem ouve.

Para Cagliari (2007, 2015), a expectativa é uma propriedade fundamental no processo de percepção do ritmo. Segundo ele, a ideia de ritmo está mais relacionada à expectativa do interlocutor do que propriamente a sua realização, no que diz respeito às quantidades absolutas de duração das unidades. O autor explica que a linguagem é o espelho da mente, de modo que a linguagem, em uso, permite observar como a mente funciona. Segundo ele, a expectativa está mais relacionada à espera do falante e do ouvinte do que ao sentido literal do enunciado. Nesse sentido, o falante deseja que o seu interlocutor compreenda não só o sentido literal do discurso, mas também a informação que está por trás do sentido literal.

Diante do processo comunicativo, falante e ouvinte criariam, então, sua própria expectativa que acaba sendo frustrada, na fala gaga, tanto para quem fala quanto para quem ouve. Na fala “normal”, a disfluência comum funciona mais como uma estratégia pragmática. É um processo cultural. Cada língua apresenta disfluências que são características da sua região. As disfluências também são fundamentais para pessoas que querem falar outra língua, pois se na língua materna ela mostra uma quebra na estrutura, um vazio cognitivo que espera ser preenchido, na língua estrangeira, a disfluência mostra segurança, confiabilidade, e mais do que isso, mostra fluência.

Outro fenômeno que pode contribuir para o sentimento de distanciamento entre a fala gaga e não gaga diz respeito à quantidade de disfluências por GT. Como vimos, as disfluências funcionam como estratégias que o informante usa na tentativa de recuperar a fluência. Todos os falantes analisados (gagos e não gagos) realizaram RPs, ALs, BLOs, PSs, entre outras.

De uma maneira geral, as INSs, os ALs e as RPs ajudam o informante a recuperar a mensagem, a fim de que ele ganhe tempo para formular o pensamento, que devido ao processo de produção falha durante a sua execução. Com as CORs e FCs a falha ocorre justamente durante a produção da mensagem. No entanto, o falante, seja ele gago ou não, tem consciência de que as realizou. Isso é tanto verdade que ocorre a correção da informação, que pode ou não apresentar partes da informação errada. Já com os FCs, não há recuperação do erro. A informação errada é deixada de lado e o processo comunicativo continua. As PIs, por sua vez, caracterizam um fenômeno rítmico.

As PSs devem-se à alternância no ritmo, demarcação dos GTs e à preocupação do falante com a elaboração do objetivo e da informação que deve ser recuperada. Para os informantes não gagos elas são mais distribuídas (início, meio e fim de GT) no enunciado, diferentemente da dos não gagos que tendem a ocorrer mais em início e fim de GT (cf. tab.7). Além disso, a sua duração é maior para a disfluência comum. Os informantes gagos tendem a apresentar uma duração maior para os outros eventos disfluentes. Esse resultado mostra que as PSs funcionam como um recurso discursivo importante para a construção do discurso, principalmente em meio de GT. O maior número delas em um GT diminui a percepção de falha e revela uma maior organização do pensamento.

Os falantes gagos, por sua vez, constroem a sua fala por meio de RPs e disfluências que duram mais do que a dos não gagos, aumentando o tempo de volta à fluência. Além disso, esses eventos são mais complexos e a sua quantidade em um único GT também é maior. O

falante gago realiza com mais frequência números maiores de RPs, que podem chegar até cinco, enquanto para os não gagos elas podem chegar até três, mas com uma frequência menor. Sendo assim, essa quebra mais frequente na estrutura rítmica na disfluência comum e o tipo mais frequente de preenchimento do erro, ou seja, para os não gagos as PSs e gagos as RPs leva ao aumento da empatia negativa e da expectativa na fala. Fato que confirma que a repetição, ao atrasar o fluxo de fala, seja na disfluência comum ou gaga, ajuda o informante a recuperar a próxima palavra e conseqüentemente, voltar à fluência. Para os informantes gagos, a disfluência caracteriza, então, um processo múltiplo, que vem carregado de mais fatores, ou seja, da maior complexidade e número de eventos em um mesmo GT.

Os BLOs, assim como as RPs são outro tipo de disfluência que contribuem mais para o distanciamento entre as disfluências, devido à sua complexidade. As outras disfluências, por sua vez, mesmo que sejam mais frequentes em um grupo do que em outro não contribuem diretamente para essa separação. Isso ocorre porque a maior parte dos BLOs vem acompanhados de PSs, e essas, por sua vez, também duram mais na disfluência gaga. O momento de soltura, antes da produção do segmento, demora mais para ocorrer, o que aumenta a percepção da falha.

Essas questões estão mais relacionadas ao processo articulatório, principalmente às questões aerodinâmicas. O informante gago demora mais tempo para atingir o padrão ideal de duração, por isso, ele gagueja mais vezes. O maior tempo de articulação faz com que ocorra uma mudança maior no processo respiratório, levando o informante gago a realizar um número maior de disfluências complexas por enunciado, dificultando, assim, a produção da mensagem. Além disso, o fato do tempo de articulação (cf. tab. 19) ser maior para os gagos revela que a gagueira está mais relacionada à questão articulatória, enquanto que a disfluência comum, à questão cognitiva.

Outra questão a favor da hipótese articulatória são os bloqueios. Como vimos, os gagos são os que mais realizam esse tipo de disfluência. Os nossos resultados mostraram uma maior frequência de BLO para as consoantes oclusivas (79%) que, por sua vez, caracterizam o bloqueio completo da corrente de ar, assim como foi observado por Shiriberg (2001) que observou a maior tendência de disfluências para a oclusiva bilabial.

A fim de facilitar a visualização desses resultados, observe o quadro descritivo 7 que segue:

Quadro 7 – Descrição geral dos fenômenos observados para a disfluência gaga e comum.

FENÔMENOS OBSERVADOS	DISFLUÊNCIA GAGA	DISFLUÊNCIA COMUM
Disfluência simples	737 ocorrências	648 ocorrências
Disfluência complexa	127 ocorrências	58 ocorrências
GT's	52 ocorrências	41 ocorrências
Nº de disfluências por GT	6 ocorrências	5 ocorrências
Nº de pausas por GT	1 ocorrências	2 ocorrências
Tipo mais frequente de disfluência	RPs	PSs
Tipo menos frequente de disfluência	PIs	FCs
Porcentagem de PSs	33,6%	48%
Porcentagem PS em início de GT	33,3%	33,3%
Porcentagem PS em meio de GT	0%	33,3%
Porcentagem PS em fim de GT	66,7%	33,3%
PSs respiratórias	33,3%	66,7%
PSs silenciosas	66,7%	33,7%
Duração total das PSs	131,638ms	174,945ms
Tipo de INS mais frequente	Palavras	Palavras
Duração total das INSS	32,720ms	43,812ms
Média da F0 (INS)		maior
Padrão da intensidade final (INS)	90% tendência à queda	80% tendência ao aumento
Padrão da intensidade máxima (INS)	maior em INS _{ps}	maior em INS _{ps}
HES/ tipo	/ah/	/ah/
Contexto de ocorrência das HES	início de GT	início, meio e fim de GT
Duração da HES	2,324ms	3,618ms
Padrão da intensidade máxima (HES)	78 dB	77 dB
Padrão da intensidade final (HES)	tendência para a diminuição da intensidade	variação no padrão
Média da F0 (HES)	Maior	
FC/ duração	2,076 ms	1,698ms
Padrão da intensidade de transição (FC)	tendência para a queda	padrão variado
Padrão da intensidade final	tendência para a queda	tendência para a queda

(FC)		
Padrão da intensidade máxima (FC)	71 dB	65 dB
Média da F0 (FC)	Maior	
AL/ duração	5,998ms	21,740ms
Padrão da intensidade de transição (AL)	tendência para a queda	sem padrão definido
Padrão da intensidade final (AL)	tendência para a queda	tendência para o aumento no padrão
Padrão da intensidade máxima (AL)	76 dB	76 dB
F0 média (AL)	Maior	
PIs/ duração	5,786ms	2,181ms
Padrão da intensidade de transição (PI)	tendência para a queda	tendência para a queda
Intensidade final (PI)	tendência para a queda	sem padrão definido
F0 média (PI)		maior
COR/ contexto de frequência mais comum	meio de GT	meio de GT
Duração total das COR's simples e complexas	8,324ms	28,214ms
Tendência da duração no reparo para as COR's simples e complexas	aumento	aumento
Tendência da F0 média durante o reparo para a COR simples	aumento	aumento
Tendência da F0 média durante o reparo para a COR simples	diminuição	diminuição
Tendência da F0 média durante o reparo para a COR complexa	variação no padrão	variação no padrão
Padrão da intensidade de transição durante o reparo (COR simples)	diminuição	variação no padrão
Padrão da intensidade de transição durante o reparo (COR complexa)	variação no padrão	variação no padrão
Padrão da intensidade final durante o reparo (COR simples)	diminuição	aumento
Padrão da intensidade final durante o reparo (COR complexa)	variação no padrão	variação no padrão
Padrão da intensidade máxima na passagem para o reparo (COR simples)	abaixamento	aumento
Padrão da intensidade	variação no padrão	aumento

máxima na passagem para o reparo (COR complexa)		
BLO/ segmento mais comum	consoantes oclusivas [k]	consoantes oclusivas[k]
Pausa de BLO mais frequente	pausa com presença de ciclos de onda (TIPO 1)	pausa com presença de ciclos de onda (TIPO 1)
Duração total dos segmentos bloqueados	33,492ms	8,380ms
Duração das PSs de BLO	28,487ms	5,613ms
Padrão da intensidade máxima (BLO)	75 dB	79 dB
Padrão da intensidade final	crescente	crescente
Média da F0		maior
RP _s	287	87
Total de RP _s complexas	28,6%	18,3%
Número máximo de palavras repetidas em um GT	5	3
Tipo de RP mais realizada	RP _p	RP _p
Duração da segunda RP	Maior	
Tendência da duração na segunda RP	diminuição	diminuição
Tendência da F0 média para as RP _s	100% diminui	50% diminui
Tendência da intensidade de transição para as RP _s	Varia	90% diminui
Tendência da intensidade máxima para as RP _s	50% diminuição	100% aumenta
Tendência da intensidade final RP _s	50% aumenta	50% aumenta
Média aproximada da F0 mínima para a análise da tessitura	1118 Hz	116 Hz
Média aproximada da F0 máxima para a análise da tessitura	146 Hz	140 Hz
Tempo de elocução (média aprox.)	174ms	148ms
Tempo de articulação (média aprox.)	148ms	118ms
Taxa. de elocução	4,569ms	4,600ms
Taxa de articulação	5,457ms	5,666ms
Número de sílabas	784	670
Duração de pausas	26,195ms	29,462ms

Fonte: própria.

O quadro 7 apresenta uma descrição geral dos eventos observados para a disfluência comum e gaga. No quadro, estão destacados os GTs, os tipos de disfluência, a duração dos eventos, a intensidade (transição, final e máxima) e a F0. Os resultados para a análise da F0 média e momentos de intensidade (máxima, final e de transição) não foram suficientes para

distanciar os dois tipos de disfluências, uma vez que as diferenças no padrão foram pequenas. A maior diferença está relacionada ao tipo de disfluência, à duração e aos aspectos temporais (cf. tab. 20).

Os informantes com disfluência gaga tendem a apresentar um maior número de RPs e BLOs⁶⁹ por GT do que os não gagos, assim como a complexidade dos eventos. A duração desses eventos também são maiores. Eles realizam um menor número de pausas do que os não gagos. A taxa de elocução e de articulação também foi menor para os informantes com gagueira. Por outro lado, o tempo de elocução, articulação e o número de sílabas foram maiores. Essas tendências revelam a desorganização do ritmo na fala gaga, tanto pelo maior duração nos eventos repetidos quanto ao menor número de pausas realizadas e maior tempo de articulação. O informante gago quebra a estrutura rítmica com mais frequência, o que, concomitantemente, ocasiona a quebra na expectativa do ouvinte. A percepção das falhas tornam-se, então, mais frequentes e, conseqüentemente, mais perceptíveis. Esses fatos também mostram que a gagueira está mais relacionada a questões articulatórias, devido a um erro durante a programação da duração das sílabas antes da articulação dos sons (Cagliari, 2007). Já com a disfluência comum, as falhas são cognitivas, ou seja, o falante cria estratégias para preencher os momentos vazios (disfluentes) durante a produção da fala. Mesmo assim, não podemos dizer que a disfluência comum deva-se apenas um processo cognitivo e que a disfluência gaga deve-se somente a questões articulatórias. Certamente, podem ocorrer falhas articulatórias em uma fala com disfluência comum, bem como a fala gaga não é construída apenas por elas. O falante gago, bem como o não gago, também se vale de estratégias para contruir o seu discurso. E, portanto, também é um processo cognitivo.

Mesmo diante dessas diferenças, acreditamos que o mais adequado não é separar o que se considera uma fala fluente de uma não fluente, diagnosticada por apresentar gagueira. A construção da fluência na fala só é possível por meio das disfluências. Ela pode ser comparada à elaboração de um texto, que para ser terminado, necessita de correções, apagamentos e até mesmo de uma pausa, para que possa ser retomado e, em seguida, finalizado. E, a fala, seja ela comum ou gaga, segue a mesma construção.

Esse ponto de vista, no entanto, não tem como objetivo enquadrar as duas falas em um mesmo nível. Mesmo porque não existem falas iguais. Os falantes não são os mesmos.

⁶⁹Destacamos as RPs e BLOs porque, ao nosso ver, são elas que causam maior estranhamento ao interlocutor. Diferentemente das INSSs, dos FCs, COR's, entre outros, que mesmo sendo chamadas aqui de disfluência, funcionam mais como estratégias discursivas para a produção da fala.

Contudo, os resultados mostraram, parafraseando Scarpa e Fernandes-Svartman (2012), que os limites que separam a (dis)fluência comum da gaga são fluidos.

As ideias, assim como um texto em construção, precisam de tempo para serem programadas. E isso, na fala, ocorre por meio das disfluências. É mais fácil recuperar palavras mais frequentes, mais comuns ao vocabulário do falante. A fluência pura, livre de erros, é artística, ensaiada, obtida por meio do texto pronto, decorado.

As pessoas gagas e não gagas realizam o mesmo tipo de disfluência. Cada um deles, na sua individualidade, tem um estilo de fala que lhe é próprio, mas cada uma dessas falas apresentam sinais linguísticos que lhes são específicos (Carneiro; Scarpa, 2012).

Vimos que a análise linguística das disfluências não pode ser descartada e tratada apenas como um caso acidental, já que a diferença entre esses dois tipos de fala também está relacionada a questões acústicas. A empatia negativa, que gera o preconceito, decorre do modo como esses eventos se apresentam.

O distanciamento também é gerado porque a gagueira é mais um reflexo sobre a produção, e a disfluência comum, mais do discurso. A última diz respeito a questões mais estilísticas. Os momentos disfluentes, são mais organizados, acontecendo em momentos de fala que não prejudicam significativamente o discurso. Na disfluência gaga há um colapso na organização das sílabas, já que o contexto fonético não permite que elas sejam reconhecidas como processos estilísticos.

Ao juntar questões qualitativas e quantitativas em uma análise bastante detalhada, o que também justifica a quantidade de informantes, este trabalho forneceu dados importantes para o melhor conhecimento do assunto, que até então foi muito pouco explorado. Além disso, a pesquisa também é relevante para a área da linguística, formando um *corpus* específico para a área da fonoaudiologia, esclarecendo tipos de disfluências e suas ocorrências, e para a área da educação, ajudando professores e pais a conhecer melhor a fala de alunos e filhos com disfluências e gagueiras.

O estudo apresentado está longe de esgotar o assunto sobre as disfluências gaga e comum. É preciso que mais trabalhos sejam desenvolvidos para aumentar os dados referentes à comparação entre esses dois fenômenos da fala. Seria interessante, para pesquisas futuras, a observação dos fenômenos por meio de outro tipo de *corpus* como, por exemplo, de leitura, seja de sentenças, de pares mínimos, de um pequeno trecho de uma história ou até mesmo um trabalho comparativo entre estes tipos de leitura. Há ainda, a possibilidade de aumentar o número de informantes e verificar se os resultados obtidos se mantêm.

REFERÊNCIAS

ABERCROMBIE, David. *Studies in phonetics and linguistics*. London: Oxford University Press, 1965.

_____. *Elements of General Phonetics*. Edinburgh, Edinburgh University Press, 1967.

ARNOLD, J. E.; FAGNANO, M.; TANENHAUS, M. K. Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research*, v. 32, n. 1, p.25–36, 2003.

AZEVEDO, Nadia Pereira da Silva Gonçalves. Uma análise discursiva da gagueira; trajetórias de silenciamento e alienação na língua. Mestrado em fonoaudiologia. 132f. Pontifícia Universidade Católica de São Paulo, 2002.

BÍBLIA. A. T. Êxodo. In: BÍBLIA. Português. *Bíblia sagrada: contendo o antigo e novo testamento*. Tradução de Frei João José Pedreira de Castro. São Paulo: Ave-Maria, 2001. p. 103-105.

BECKHAM, Mary. Elisabeth.; PIERREHUMBERT, Janet. Intonation structure of English and Japanese. *Phonology Yearbook*, 3, p. 255-310, 1986.

BERGMANN, Günther. Studies in stuttering as a prosodic disturbance. *Journal of Speech and Hearing Research* 29, 290-300, 1986.

BLOODSTEIN, OLIVER – *Stuttering- The search for a cause and cure*. Needham Heights: Allyn and Bacon, 1993.

BOERSMA, P.; WEENINK, D. Praat: doing phonetics by computer (Version 5.3) [Computer Program]. Disponível em: <[http://: www.praat.org/](http://www.praat.org/)>, 2014.

BOHNEN, Anelise. Junqueira. *Estudo das palavras gaguejadas por crianças e adultos: caracterizando a gagueira como um distúrbio de linguagem*. Tese de Doutorado. Estudos da Linguagem. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.

BOLINGER, D. A theory of pitch accent in English. *Word*, n. 14, p. 100-149, 1958.

_____. Intonation: level vs configurations. *Word*, n. 7, p. 199-210, 1951.

BORTFELD, H.; LEON, S. D.; BLOOM, J. E.; SCHOBBER, M. F.; BRENNAN, S. E. Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. *Language and Speech*, v. 44(2), p. 123-147, 2001. Disponível em: <<http://www.psychology.sunysb.edu/sbrennan-/papers/bortetal.pdf>>. Acesso em: 8 abr. 2015.

BRUCE, G. *Swedish word accents in sentence perspective*. Travaux de L'Institut de Linguistique de Lund, CWK Gleerup: Mamö, 1977.

BRUM-DE-PAULA, Mirian Rose. Da intenção à articulação: modelizações e análise proposicional. In: GUIMARÃES, E.; BRUM-DE-PAULA, M.R. (Orgs.). *Sentido e Memória*. Campinas: Pontes, 2005. Disponível em: <<http://wp.ufpel.edu.br/laboratorioelo/files/2014/05/sentido.pdf>>. Acesso em 17 de jun. 2015.

BÜCHEL, Christian; SOMMER, Martin. Mistério não solucionado: *o que causa a gagueira*. Tradução de Silva, Hugo e Revisão de Merlo, Sandra. *PLoS Biology*. v. 2, p. 160-164, fev. 2004. Disponível em: <<http://www.gagueira.org.br>>. Acesso em: 16 mar. 2015.

CAMERON, Lynne. Empathy: a review. *Living with Uncertainty*. Working Paper 1. 2011. Disponível em: <www.open.ac.uk/researchprojects/livingwithuncertainty/>. Acesso em: 20 set. 2016.

CAGLIARI, Luiz. Carlos. *An outline of a theory of expectation*. 2016, p.1-14 (ms).

_____. Línguas de ritmo silábico. *Revista Estudos Linguísticos*, Belo Horizonte, v. 20, n. 2, p. 23-58, julho/dezembro 2012a.

_____. *A entoação no componente fonológico da gramática*: Halliday (1970) e Pierrehumbert. 2012. p. 1-18 (ms).

_____. *A entoação no componente fonológico da gramática*. 2011, p.1-3 (ms).

_____. *Elementos de fonética do Português Brasileiro*. São Paulo, Paulistana, 2007.

_____. *Dossiê Prosódia*, 2002 (ms).

_____. Prosódia: algumas funções dos supra-segmentos. In: *Cadernos de Estudos Linguísticos*, Unicamp, n. 23, p. 137-151, jul/dez, 1992.

_____. Da importância da prosódia na descrição de fatos gramaticais. In: *Gramática do Português falado: níveis de análise lingüística*, ILARI, R. (Org.). Campinas: Ed. UNICAMP, v. 2, 1992a, p. 39-64.

_____. Elementos de fonética do Português Brasileiro. 1982. 192f Tese (Livre docência em Linguística) - Universidade Estadual de Campinas, Campinas, 1982.

CAGLIARI, L. C; MASSINI-CAGLIARI, G. O papel da tessitura dentro da prosódia portuguesa. In: CASTRO, I.; DUARTE, I. *Razões e emoção: miscelânea de estudos em homenagem a Maria Helena Mateus*, 2003, v.1, p. 67-85. Disponível em: <<http://www.fl.ul.pt/dlgr/mateus/mateus.htm>>: Acesso em 18 jun. 2010.

CARVALHO, MARIANE. A variação melódica da fala: um estudo da correlação entre uma análise acústica e uma análise auditiva da entoação. 2013. 228f. Dissertação (Mestrado em Linguística e Língua Portuguesa). Faculdade de Ciências e Letras, Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara, 2013.

CATFORD, J. C. *A practical introduction to phonetics*. Oxford: Oxford University Press, 2010.

_____. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press, 1977.

CLARK, M. M. *A dynamic treatment of tone with special attention to the tonal system of Igbo*. Bloomington: Indiana University Linguistics Club, 1978.

CLEMENTS, G. N. The Geometry of Phonological Features. *Phonology Yearbook 2*, Great Britain. p. 225-252. 1985.

CLUTER, A. Speakers conceptions of the functions of prosody. In: CUTLER, A.; LADD, D, R. (ed.). *Prosody: models and measurements*. Springer, Heidelberg. 1983.

CONTURE, E., GERALD, G., MCCALL, N.; BREWER, D. W. Laryngeal behaviour during stuttering. *Journal of Speech and Hearing Research*, v. 20, 1977 p. 661-668. Disponível em: <http://www.researchgate.net/profile/Edward_Conture/publication/22536987_Laryngeal_behavior_during_stuttering/links/00463530b7ff6a8457000000.pdf>. Acesso em: 16 jun. 2015.

CONTURE, Edward, G.; SCHWARTZ, Howard D.; BREWER, David W. Laryngeal behavior during stuttering: a further study. *Journal of speech and hearing research*, v. 28, 1985, p. 233-240. Disponível em:

<http://www.researchgate.net/profile/Edward_Conture/publication/19152196_Laryngeal_behavior_during_stuttering_a_further_study/links/00463530b7fb457e0e000000.pdf>. Acesso em: 16 jun. 2015.

CARNEIRO, Célia R.; SCARPA, Ester. Singularidades nas manifestações de fala gagas. In: *Cadernos de estudos linguísticos* (UNICAMP), Campinas, v. 54.1, n. 1 p. 155-156, 2012. Disponível em: <<http://www.iel.unicamp.br/revista/index.php/cel/article/view/2571/2004>>. Acesso em: 16 mar 2012.

CHACON, Lourenço.; VILLEGA, Cristyane Camargo Sampaio. *Hesitações na fala infantil: indícios da complexidade da língua*. In: *Cadernos de estudos linguísticos* (UNICAMP), Campinas, v. 54.1, n.1 p. 81-95. Disponível em: <<http://www.iel.unicamp.br/revista/index.php/cel/article/view/2571/2004>>. Acesso em: 24 mar 2015.

CHOMSKY, Noam. *Aspects of the theory of syntax*. Cambridge: MIT Press, 1965.

CLARK, Hebert, H.; FOX TREE, Jean, E. Using uh and um in spontaneous speech. *Cognition*, v. 84, p. 73-111, 2002. Disponível em: <http://www.columbia.edu/~rmk7/HC/HC_Readings/Clark_Fox.pdf> . Acesso em 13 abr. 2015.

CONDELIPES, Teresa Isabel. *Produção de fala por indivíduos gagos*. 2010. 152f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2010. Disponível em: <http://repositorio.ul.pt/bitstream/10451/3652/1/ulfl080866_tm.pdf>. Acesso em 23 de mar 2015.

CRISTÓFARO, Thais Silva. *Dicionário de fonética e fonologia*. São Paulo: Contexto, 2011.

CRUTTENDEN, Alan. *Intonation*. Cambridge: Cambridge University Press, 1986.

CRYSTAL, David. *Dicionário de linguística e fonética*. Rio de Janeiro: Jorge Zahar, p. 275, 2000.

CRUZ, Marisa. *Gaguez: em busca de um padrão prosódico e entoacional*. Dissertação de mestrado. Universidade de Lisboa, 2009. Disponível em: <<http://repositorio.ul.pt/handle/10451/518>>. Acesso em: 02 jul 2012.

CRUZ, Marisa.; SÓNIA, Frota. O sintagma entoacional na gaguez – evidências do PE. *Talk given at XXV National Meeting of Associação Portuguesa de Linguística*, Faculdade de Letras, Universidade de Lisboa, Lisboa, 2009. Disponível em: <http://labfon.letras.ulisboa.pt/texts/Cruz_Frota2010.pdf> . Acesso em: 20 de mar. de 2015.

_____. *Condelipes*. A gaguez e a investigação científica: contributos da linguística para a intervenção. Invited talk given at V Jornadas sobre a Gaguez, Associação portuguesa de gagos & Centro de Estudos Sociais da Universidade de Coimbra, Coimbra, 2011. Disponível em: <http://labfon.letras.ulisboa.pt/texts/Cruz_Frota_Condelipes_2011.pdf>. Acesso em: 23 de mar. de 2015.

DELFINO, Alexandre. *Estudo prosódico das disfluências de reparo*. Dissertação de Mestrado. Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

Didirkova Ivanna. Origines articulatoires des discours hors-normes : le cas du bégaiement. *Revue Signes, discours et société*. Disponível em: <<http://www.revuesignes.info/document.php?id=4817&format=print>>. Acesso em 21 jun. 2016.

DUEZ, Danielle. *Acoustico-phonetic characteristics of filled pauses in spontaneous French speech: preliminary results*, In: *DISS'01*, 2001, p. 41-44. Disponível em: <http://www.iscaspeech.org/archive_open/archive_papers/diss_01/dis1_041.pdf>. Acesso em: 20 mai. 2015.

_____. *La signification des pauses dans la production et perception de la parole*, 3-4, p. 275-300, 1997.

FINN, Patrick; INGHAM, Roger. The selection of "fluent" samples in research on stuttering: conceptual and methodological considerations. In Healey, Ch (org.). *Readings on research in stuttering*. Nova Iorque: Longman Publishing Group, 1991, p. 91-109.

FÁVERO, Leonor Lopes.; ANDRADE, Maria Lúcia da Cunha Victório de Oliveira.; AQUINO, Zilda Gaspar Oliveira. Correção. In: JUBRAN, Clélia Cândida Abreu Spinardi; Koch, Ingdore Grunfeld Villaça (Orgs.). *Gramática do Português falado*. Campinas: Editora da Unicamp, 2006. p. 255-273.

FREEMAN, Frances.; USHIJIMA, Tatsujiro. Laryngeal muscle activity during stuttering. *Journal of Speech and Hearing Research*, v. 21, 1978, p. 538-562. Disponível em: <<http://web.haskins.yale.edu/Reprints/HL0248.pdf>>. Acesso em: 16 jun. 2015.

FERNANDES, F. R. *Ordem, focalização e preenchimento em português: sintaxe e prosódia*. 2007. 445f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2007.

FERREIRA, Fernanda; BAILEY, Karl G. D. Disfluencies and human language comprehension. *Cognitive Sciences*. Elsevier, 2004.

FREIRE, Regina Maria; PASSOS, Maria Cristina Pascalicchio. *Gagueira: uma questão discursiva*. *Trab. linguist. apl.* [online]. 2012, vol.51, n.1, pp. 153-173. ISSN 0103-1813. Disponível em: <<http://www.scielo.br/pdf/tla/v51n1/v51n1a08.pdf>>.

FROTA, S.; VIGÁRIO, M. Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. In: APL, v. 1, 1999, Faro. *Actas do XV encontro nacional da APL*, Faro: APL, 1999. p. 533-555.

GOMES-KELLY, Roberta Ecleide de Oliveira. Fluir ou disfluir: eis a questão! Uma discussão sobre a gagueira e a psicanálise. In: COLOQUIO DO LEPSI IP/FE-USP, 4., 2002, São Paulo. *Proceedings online*. Disponível em: <http://www.proceedings.scielo.br/scielo.php?script=sci_arttext&pid=MSC000000032002000400021&lng=en&nrm=abn>. Access on: 19 mar. 2015.

GUIMARÃES, Rejane Vieira. *Breve revisão sobre a disfemia*. Disponível em: <<http://www.profala.com/arttf40.htm>>. Acesso em: 16 mar. 2015.

GUITAR, Barry. *Stuttering: an integrated approach to its nature and treatment*. 3rd ed. Philadelphia, Lippincott Williams & Wilkins, 2006.

GUSSENHOVEN, Carlos. Phonology of intonation. *Glott International*. 2002, v. 6, n.9/10. nov/dez. pp. 271-284.

_____. On the intonation of tonal varieties of English. In Markku Filppula, Juhani Klemola & Devyani Sharma (ed.). *The Oxford Handbook of World Englishes online*. 2014

Online Publication Date: Dec 2014. DOI: 10.1093/oxfordhb/9780199777716.013.29. Acesso em 09 abr. 2015.

Disponível em: <http://gep.ruhosting.nl/carlos/Gussenhoven_Handbook_VarEng.pdf>

HALLIDAY, M.A.K. The Tones of English. In: JONES, W. E.; LAVER, J. (Ed). *Phonetics in Linguistics*. London: Longman, 1973. p. 103-126.

_____. *A course in spoken english: Intonation*. London: Oxford University Press, 1970.

HARRINGTON, Jonathan. Acoustic Phonetics. In: HARDCASTLE, Willian J.; LAVER, John.; GIBBON, Fiona E. (Orgs.). *The handbokk of phonetic sciences*. Cambridge: Wiley-Blackwell, 2010. p. 80-129.

HEIKE, A. E. A content-processing view of hesitation phenomena. *Language and Speech*, 24(2), 1981, p. 147-160.

HIRSCH, Fabrice. *Le bégaiement : perturbation de l'organisation temporelle de la parole et conséquences spectrales*. 2007. 282p. Tese de doutorado. Université Marc Bloch. França, 2007. Disponível em: <http://scd-theses.u-strasbg.fr/572/>. Acesso em 18 jun. 2016.

HIRSCHBERG, Júlia; PIERREHUMBERT, Janet. The meaning of intonation contours in the interpretation of discourse. In: P. Cohen, J. Morgan, and M. Pollack, (ed.). *Intentions in Communication*. MIT Press, Cambridge, p. 271-311, 1990.

HIRST, D.; DI CRISTO, A. *Intonation system: survey of twenty languages*. Cambridge: Cambridge University Press. 1998.

JAKOBSON, Roman. *Lingüística e comunicação*. São Paulo, Cultrix, 1995.

JAKUBOVICZ, R. *A Gagueira: teoria e tratamento de adultos e crianças*. Rio de Janeiro: Revinter, 1997.

JONES, D. *An Outline of English phonetics*. Cambridge: Cambridge University Press, 1960.

KOCH, Ingedore Grunfeld Villaça. *Desvendando os segredos do texto*. 6ª ed. São Paulo: Cortez, 2009.

_____. *O texto e a construção dos sentidos*. São Paulo: Contexto, 2007.

_____. Especificidade do texto falado. In: JUBRAN, Clélia Cândida Abreu Spinardi; Koch, Ingdore Grunfeld Villaça (Orgs.). *Gramática do Português falado*. Campinas: Editora da Unicamp, p. 39-46, 2006.

KORMOS, Judit; DÉNES, Mariann. Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32 (2), p. 145-164, 2004. Disponível em: <<http://eprints.lancs.ac.uk/10747/1/fluency-preprint.pdf>>. Acesso em: 15 abr. 2015.

LADD, D. R. *Intonational phonology*. Cambridge: Cambridge University press, 1996.

_____. Levels versus configurations, revisited. In: F. B. Agard; G. B. Kelley, A. Makkai; V. B. Makkai (eds.). *Essays in honor of Charles F Hockett*. Leiden: E. J. Brill, p. 93-131.

_____. *The structure of intonational meaning: evidence from English*. Bloomington: Indiana University Press, 1980.

LADEFOGED, Peter.; JOHNSON, Keith. *A course in phonetics*. Boston: Cengage Learning, 2011.

LADEFOGED, Peter. *Elements of acoustic phonetics*. Chicago: University of Chicago Press, 1995.

_____. The linguistic use of different phonation types. In: BLESS, D.; ABBS, J. (Orgs.). *Vocal fold physiology*. San Diego: College Hill Press, 1983, p. 351-360.

_____. *A course in phonetics*. New York: Holt Rinehart and Winston, 1975.

LAPORTE, Valerie. *Onset and development*. 1997, Disponível em: <<http://www.d.umn.edu/~cspiller/stutteringpage/onset.htm>> Acesso em: 18 abr. 2014.

LENNON, P. Investigating fluency in EFL: a quantitative approach. *Language Learning*, 40, 387-412, 1990.

LEVELT, Willem.J.M.; CUTLER, A.: Prosodic marking in speech repair. IN: *Journal of Semantics* n.2, p. 205–217, 1983.

LEVELT, Willem, J. M. *Speaking: from intention to articulation*. Cambridge: London, 1989.

LAVER, John. *Principles of phonetics*. Cambridge: Cambridge University Press, 1995.

LIBERMAN, M; PIERREHUMBERT, J. *Intonational invariance under changes in pitch range and length*. In: ARONOFF, M.; OEHRLE, R.T. (ed.). *Language sound structure*. Cambridge, Mass.: MIT Press, p. 157-233.

LIBERMAN, M.; PRINCE, A. On stress and linguistic rhythm. *Linguistic Inquiry*, n. 8, p. 249-336, 1977.

LINDBLOM, B. Phonetic universals in vowel systems. In: OHALA, J. & J. J. JAEGER (Orgs.). *Experimental phonology*. Academic Press, 13 – 44, 1985.

LUCENTE, Luciana. *DaTo: um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea*. 2008. 72f. Dissertação (mestrado) – Universidade Estadual de Campinas, Campinas, 2008.

MASSINI-CAGLIARI, Gladis.; CAGLIARI, Luiz Carlos. Fonética. In: BENTES, A. C.; MUSSALIM, F. (Org.). *Introdução à linguística: domínios e fronteiras*. São Paulo: Cortez, 2001. p. 105-146.

MATEUS, Maria Helena. Mira. A investigação em fonologia do português. *DELTA* [online]. 2001, vol.17, n.spe, pp. 57-79. Disponível em:
<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502001000300005>

MARCUSCHI, Luiz Antônio. Fenômenos intrínsecos da oralidade. In: JUBRAN, Clélia Cândida Abreu Spinardi; Koch, Ingdore Grunfeld Villaça (Orgs.). *Gramática do Português falado*. Campinas: Editora da Unicamp, 2006. p. 47-70.

MARTINS, Roberto de Andrade. *Aristóteles e o estudo dos seres vivos*. Disponível em:
<<http://www.ghtc.usp.br/server/PDF/ram-Aristoteles-livro.PDF>>. Acesso em: 16 Mar. 2015.

MEIRA, Maria Isis Marinho. Causas da gagueira. Fórum online. Dia Internacional de atenção à gagueira. Disponível em: <http://www.abragagueira.org.br/causas_forum.asp?id=9> . Acesso em: 25 mar. 2015.

MERLO, Sandra. *Dinâmica temporal de pausas e hesitações na fala semi-espontânea*. Tese de Doutorado. Universidade Estadual de Campinas, Campinas, 2012.

_____. Hesitações na Fala Semi-Espontânea: análise por séries temporais. Dissertação de Mestrado. Campinas: IEL-UNICAMP, 2012.
Disponível em: <<http://libdigi.unicamp.br/document/?code=vtls000381660>> . Acesso em: 3 mar. 2015.

MERLO, Sandra.; BARBOSA, Plínio, Almeida. Análise acústica da fala suavizada: estudo de caso em gagueira. *Cadernos de Estudos Lingüísticos* . (UNICAMP), v. 54, p. 167-181, 2012.

_____. *Séries temporais de pausas e hesitações na fala espontânea*. Caderno de estudos linguísticos (54.1), Campinas, Jan./Jun. 2012, p. 11-24.

MATA, Ana. Isabel. *Para o Estudo da Entoação em Fala Espontânea e Preparada no Português Europeu*: metodologia, resultados e implicações didáticas. Dissertação de Doutorado, F.L. Universidade de Lisboa, 1999.

MONIZ, Helena Gorete Silva. *Contributo para a Caracterização dos Mecanismos de (Dis)fluência no Português Europeu*. Dissertação de Mestrado em Linguística – Linguística Portuguesa (Programa orientado para a Linguística Educacional). Faculdade de Letras, Universidade de Lisboa, 2006. Disponível em: <http://www.inesc-id.pt/pt/indicadores/Ficheiros/4006.pdf>. Acesso em 10 jun. 2015.

MORAES, J. A.; STEIN, C. C. Attitudinal patterns in brazilian portuguese intonation: analysis and synthesis. *Speech Prosody. Dresden*, maio, p. 2-5, 2006. Disponível em: <http://www.isca-speech.org/archive/sp2006/papers/sp06_223.pdf>. Acesso em: 22 set. 2009.

NAKATANI, C. H; HIRSCHBERG, J. A corpus-based study of repair cues in spontaneous speech. In: *Journal of the Acoustical Society of America*, 95 (3), p. 1603-1616, 1994.

O'CONNELL, D. C.; KOWAL, S. *Communicating with one another: toward a psychology of spontaneous spoken discourse*. New York: Springer. 2010.
Disponível em:
<http://www.academia.edu/3794060/Communicating_with_One_Another_Toward_a_Psychol

ogy_of_Spontaneous_Spoken_Discourse_Cognition_and_Language_A_Series_in_Psycholinguistics>. Acesso em: 28 abr. 2015.

OLIVEIRA, Ana Maria do Carmo Carvalho de; RIBEIRO, Ignês Maia; MERLO, Sandra; CHIAPPETTA, Ana Lúcia de Magalhães Leal. *O que fonoaudiólogos e estudantes de fonoaudiologia entendem por fluência e disfluência*. *CEFAC* [online]. 2007, vol.9, n.1, pp. 40-46. ISSN 1982-0216.

Disponível em: <<http://www.scielo.br/pdf/rcefac/v9n1/v9n1a04.pdf>>. Acesso em: 14 abr. 2015.

OLIVEIRA, Karina Vitor; Faria, Bárbara Silveira; Silva, Juliana Preisser Godoy; Reis, César; Ghio, Alain; Gama, Ana Cristina Côrtes. Análise das medidas aerodinâmicas no português brasileiro por meio do método multiparamétrico de avaliação vocal objetiva assistida (EVA). *Revista CEFAC*, São Paulo, v. 15, n. 1, p. 119-127, 2013.

Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-18462013000100014&lng=en&nrm=iso>. Acesso em: 12 mai. 2015.

OLIVEIRA, Maria Helena Mourão Alves; GARGANTINI, Marisa Bueno Mendes. Comunicação e gagueira. In: *Estudos de psicologia*. Puc Campinas, v. 20, n. 1, p. 51-60, 2003.

ORLANDI, Eni. *A linguagem e seu funcionamento: as formas do discurso*. Campinas, SP: Pontes, 1987.

PIERREHUMBERT, J. *The Phonology and phonetics of english intonation*. 1980. 402f. Tese (Doutorado em Linguística) – Department of linguistics and philosophy, Massachusetts Institute of technology, Indiana University Linguistics Club, 1980.

PIERREHUMBERT, J.; BECKMAN, M. *Japanese tone structure*. Cambridge: The MIT Press, 1988.

PIERREHUMBERT, J; HIRSCHBERG, J. The Meaning of Intonational contours in the Interpretation of Discourse. IN: Cohen, P; Morgan, J; Pollack, M. (ed.). *Intentions in Communication*, MIT Press, Cambridge. p. 271-311, 1990.

PIKE, Kenneth. Lee. *The intonation of American English*. Ann Arbor: The University of Michigan Press. 1945.

PRINS, David; HUBBARD, Carol P. KRAUSE, Michelle. Syllabic stress and the occurrence of stuttering. *Journal of Speech and Hearing Research* 34, 1011-1016, 1991.

QUEZADA, Camilo; ROBLEDO, Juan Pablo; ROMAN, Domingo; CORNEJO, Carlos. Empatía y convergencia del tono fundamental. *RLA* [online]. 2012, vol.50, n.2 [citado 2015-03-26], p. 145-165 .
Disponível em: <http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-48832012000200007&lng=es&nrm=iso>. ISSN 0718-4883. Acesso em: 26 mar. 2015.

RAMOS, Susy. *Hesitações e rupturas na fala infantil: as franjas da teoria e o lugar do sujeito na aquisição*. Dissertação de Mestrado. Universidade Estadual de Campinas, Campinas, 2008.

SAMPAIO, Leonardo Rodrigues; CAMINO, Cleonice Pereira dos Santos.; ROAZZI, Antonio. *Revisão de aspectos conceituais, teóricos e metodológicos da empatia*. *Psicol. cienc. prof.* [online]. 2009, vol.29, n.2, p. 212-227. ISSN 1414-9893.
Disponível em: <<http://www.scielo.br/pdf/pcp/v29n2/v29n2a02.pdf>>. Acesso em: 26 mar. 2015.

SCARPA, Ester, Mírian . Sobre O Sujeito Fluente. *Cadernos de Estudos Linguísticos* (UNICAMP), Campinas, v. 29, p. 163-184, 1995.

_____. (Ainda) sobre o sujeito fluente. In: Lier-de Vitto, M. F. (Org.). *Sobre a Aquisição, Patologias e Clínica de Linguagem*. 1ed. São Paulo: Editora da PUC-SP, p. 161-180, 2006.

_____. Disfluências e prosódia na fala infantil: primeiros resultados. In Barros, I; Rfken,K.; Acioli, M.; Azevedo, N.; Fonte, R.; Caiado, R & W. Cavalcanti (orgs). *Aquisição, desvios e práticas de linguagem*. Curitiba, Editora CRV, 2014.

SCARPA, Ester, Mírian.; FERNANDES-SVARTMAN, Flaviane. A estrutura prosódica das disfluências em Português Brasileiro. **Cadernos de estudos linguísticos** (UNICAMP), Campinas, v. 54.1, p. 25-40, 2012.
Disponível em: <<http://www.iel.unicamp.br/revista/index.php/cel/article/view/2571/2004>>. Acesso em: 16 jul 2013.

SCARPA, Ester, Mírian (Org.); NOVAES PINTO, Rosana. (Org.). *Caderno de Estudos Linguísticos - Volume temático sobre (dis)fluência*. 54. ed. Campinas: IEL - Universidade Estadual de Campinas, v. 1. 182p, 2012.

SCLIAR-CABRAL, L.; MARTIM, E; CHIARI, B. (1981). Fenômenos de pausa e hesitação em língua portuguesa. *Anais do IV Encontro Nacional de Linguística*, PUC- RJ, setembro de 1981.

SHRIBERG, Elizabeth. To "Errrr" is Human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, v. 31, n. 1, 2001, p. 153-169.
Disponível em: <http://www.sri.com/sites/default/files/publications/to_errr_is_human.pdf>.
Acesso em: 13 mai. 2015.

_____. Phonetic consequences of speech disfluency. *Proceedings of the International Congress of Phonetic Sciences*, vol. 1. San Francisco, 1999, p. 619-622.
Disponível em:
<<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.52.2699&rep=rep1&type=pdf>> .
Acesso em: 13 mai. 2015.

SILVERMAN, K.; *et al.* ToBi: a standard for labeling English prosody. *Proceedings of the 1992 International Conference on spoken Language Processing*. Banff: Canada, 1992.

SOUZA E SILVA, Maria Cecília Pérez.; CRESCITELLI, Mercedes Fátima de Cunha. Interrupção. In: JUBRAN, Clélia Cândida Abreu Spinardi; Koch, Ingdore Grunfeld Villaça (Orgs.). *Gramática do Português falado*. Campinas: Editora da Unicamp, 2006. p.71-86.

SPINELLI, Mauro. *Foniatría*. São Paulo, Moraes, 1983.

STARKWEATHER, C. W.; GIVENS-ACKERMAN, J. *Stuttering*. Austin, Texas: Pro-ed, 1997.

't HART, J.; COLLIER, R.; COHEN, A. *A perceptual study of intonation: an experimental phonetic approach to speech melody*. Cambridge: Cambridge University Press. 1990.

TRAGER, Geord. Leonard.; SMITH, Henry Lee. L. *An outline of English structure 41*. Oklahoma: Battenburg Press.

VASCONCELOS, José Leite. *A evolução da linguagem: ensaio anthropologico*. Porto: Typhographia occidental, 1886.

VAN, Charles Riper. *The nature of stuttering*. 2 ed., Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

ZELLNER, Brigitte. Pauses and the temporal structure of speech. In: *Fundamentals of speech synthesis and speech recognition*. Chichester: John Wiley and Sons Ltd, 1994, p. 41-62.

WINGATE, M. – A standard definition of stuttering. *Journal of Speech and Hearing Disorder.*, v. 29, p. 484-489. 1964.